Forensic identification from human fingermarks using direct PCR to improve DNA recovery

Jennifer E.L Templeton

MSc, BSc (Honours)

A thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy

School of Biological Sciences
Faculty of Science and Engineering
Flinders University

July 2017

TABLE OF CONTENTS

THESIS ABSTRACT

My original contribution to knowledge is the application of direct PCR for sampling 'touch' DNA obtained from fingermarks for the purpose of human identification. DNA from fingermarks can be crucial evidence in forensic cases where partial or smudged prints are obtained and hence cannot be used for classical fingerprinting. Advances in technology have facilitated the typing and interpretation of trace or low-level DNA from fingermarks. It is well-known in forensic science that fingermark traces may possess limited DNA. The flow-on effect of this is that the DNA profiling of fingermarks often yields little or no information that can be used to assist forensic investigations. For samples such as fingermarks, every effort needs to be made to reduce processes that are wasteful of DNA so that the success rate for DNA profiling is maximised. Standard processing in most forensic laboratories involves the sample going through a DNA extraction step, which is known to lose high percentages of DNA. One possible workflow that removes the DNA extraction step involves placing the sample directly into the Polymerase Chain Reaction (PCR). This process is called 'direct PCR' and has shown to be successful in other forensic applications where traditional DNA profiling failed.

This thesis examines the effectiveness of direct PCR to generate DNA profiles from fingermarks. Informative DNA profiles were obtained from swab fibres used to recover DNA from plastic, wood, glass and metal substrates. Further results highlight the potential for dusted fingerprints to be successfully profiled. A case report is included (see Chapter IV) demonstrating the application of the technique in real case work where DNA was recovered from a smudged fingermark on the surface of a drug seizure. Additionally, a mock case demonstration highlights that direct PCR can be used on samples subjected to environmental exposure.

In this thesis, the quality of a DNA profile is assayed by the relative peak height of the alleles, associated artefacts, allele 'drop-in' and 'drop-out'. In the first data chapter, a comparison to standard non-direct PCR was carried out (i.e. extraction followed by PCR). The data I obtained infer better quality STR-based DNA profiles recovered by direct PCR over conventional extraction. It was found that informative profiles can be generated from fingermarks left by an individual only 15 minutes after washing

hands. The use of a detergent-based nylon flocked swab was trialled and resulted in improved quality of the DNA profiles obtained.

The donors of the prints were able to be identified in a quicker time-frame than is currently possible with traditional methods that involve DNA extraction. Direct PCR reduces the opportunity for contamination by eliminating the multiple tube changes and additional steps required during an extraction. Consequently, there is a reduction in the cost of labour and reagents needed to process samples and a high through-put potential for case work exhibits. According to Forensic Science South Australia (FSSA), direct PCR is at times the only way to obtain a DNA profile from a crime scene exhibit. Ultimately, the work showed that direct PCR has a role to play in case work and proved to be reliable, robust and reproducible.

DECLARATION

I certify that this thesis does not incorporate without acknowledgement any material previously submitted for a degree or diploma at any university; and that to the best of my knowledge and belief it does not contain any material previously published or written by another person except where due reference is made in the text.

Signed. .....

Jennifer E.L Templeton

Date.....

ACKNOWLEDGEMENTS

Foremost, I would like to thank my principal supervisor Professor Adrian Linacre for his continuous enthusiasm, support, patience, motivation, and immense knowledge throughout my PhD study. Thank you for taking me on as your PhD student and for helping me to think outside the box and for sharing your wisdom. Your guidance has helped me in all aspects of my research and in writing this thesis. I have been priviledged to have known Adrian for over 13 years when I was first introduced to forensic science at Strathclyde University in Scotland where I obtained my Masters degree. I could not have asked for a better supervisor and mentor for my PhD study. I am extremely grateful for everything Adrian has helped me with.

Besides my principle supervisor, I would like to thank my other co-supervisors Dr Oliva Handt and Dr Dunan Taylor from Forensic Science South Australia, for their encouragement, insightful comments, and advice on my research and manuscipts. Thank you for taking time out of your busy day to read long drafts of my thesis chapters and manuscripts. I truly appreciate everything you have helped me with and the time you have invested in me.

I would also like to thank my fellow work colleagues from Flinders University for providing a supportive and collaborative work environment. In particular I would like to thank Alicia Haines for her support, inspirational work, camaraderie, and continual encouragement during my PhD years. I also thank both Renée Blackie and Sherryn Ciavaglia sincerely for making me feel so welcome when I first arrived and for their continual friendship, support and expert advice.

Financial assistance support for my PhD came from the Science and Industry Endownment Fund and the Attorney General's Department South Australia, and ultimately the tax payers of Australia.

On a personal note, I would also like to thank my family, especially my mum and brother Paul who are on the other side of the world but have still managed to be very supportive throughout my PhD years and listened to all aspects of my research, I love you very much. I am also grateful to the fond memories of my dad. Angela

Slater, a very good friend who is extremely generous with her time by looking after my son Patrick in order for me to focus on my studies. A huge thank you to my beautiful son Patrick who came into the world in the 2nd year of my PhD and although this presented a huge challenge to continue writing at home while looking after you, you have fully inspired me to keep going. Also thank you to my dog Angus who goes walking with me most nights to clear my head. Finally, a huge thank you to my husband Gary for helping to support me throughout the last 3 years and for everything you have helped me with. In particular listening to my oral presentations and project plans, giving me advice, and donating fingermark samples. Also, thank you for being an amazing husband and dad to our boys. I love you very much. The writing herein is dedicated to you and Patrick.

Finally thanks to the reader – if you have got this far it means that you have read at least one page of my thesis. Thank you!

LIST OF ABBREVIATIONS

µL	Microlitre
µM	Micromolar
° C	Degree Celcius
ABI	Applied Biosystems
AMEL	Amelogenin
ANZFSS	Australian and New Zealand Forensic Science Society
bp	Base Pair
BSA	Bovine Serum Albumin
BSAG	Biology Specialist Advisory Group
C	Cytosine
Ca^{+}	Calcium Ion
CCD	Charge-Coupled Device
CE	Capillary Electrophoresis
cf.	Compare
Cl^{-}	Chloride ion
cm	Centimetre
CODIS	Combined DNA Index System
Cu	Copper
DAB	DNA Advisory Board
dATP	deoxyadenosine Triphosphate
dCTP	deoxycytosine Triphosphate
dGTP	deoxyguanine Triphosphate
dNTP	deoxynucleotide Triphosphate
dTTP	deoxythymine Triphosphate
DMSO	Dimethyl Sulfoxide
DNA	Deoxyribonucleic Acid
DTT	DiThioThreitol
EDNAP	European DNA Profiling Group
EDTA	Ethylene-Diamine-Tetra-Acetic acid
ENFSI	European Network of Forensic Science Institute
EO	Ethylene Oxide
ESR	Environmental Science and Research Limited

FBI	Federal Bureau of Investigation
FSS	Forensic Science Service (UK)
FSSA	Forensic Science South Australia
G	Guanine
H^+	Hydrogen Ion
HVSI	Hypervariable Segment I of the Mitochondrial DNA Control Region
HVSII	Hypervariable Segment II of the Mitochondrial DNA Control Region
ID	Identification
INDEL	Insertion or Deletion
ISFG	International Society for Forensic Genetics
ISO	International Organization for Standardization
Ka	Kiloannus, a unit of time equal to one thousand years
kb	Kilobase
KCl	Potassium Chloride
kg	Kilogram
kV	Kilovolts
LCN	Low Copy Number
Mg^+	Magnesium Ion
$MgCl^2$	Magnesium Chloride
mL	Millilitre
mm	Millimetre
mM	Millimolar
MPS	Massively Parallel Sequencing
mtDNA	Mitochondrial DNA
NaCl	Sodium Chloride
NaOH	Sodium Hydroxide
NCIDD	National Criminal Investigation DNA Database
NDNAD	National DNA Database
ng	Nanogram
NGS	Next Generation Sequencing
Ni	Nickel
NIST	National Institute of Standards and Technology
nm	Nanometre

nu	Nuclear
OL	Off-Ladder
PBS	Phosphate Buffered Saline
PCR	Polymerase Chain Reaction
pg	Picogram
ProK	Proteinase K
QA	Quality Assurance
qPCR	Quantitative PCR
RFLP	Restriction Fragment Length Polymorphism
RFU	Relative Fluorescent Unit
RNA	Ribonucleic Acid
SAPOL	South Australia Police
SD	Standard Deviation
SDS	Sodium Dodecyl Sulphate
SNP	Single Nucleotide Polymorphism
Std Error	Standard Error
SOP	Standard Operating Procedure
SPSS	Statistical Package for Social Sciences
STR	Short Tandem Repeat
SWGDAM	Scientific Working Group on DNA Analysis Methods
T	Thymine
TE	10 mM Tris, 0.1 mM EDTA, pH 8.0
TH01	Human Tyrosine Hydroxylase Gene
Tm	Melting Temperature
Tris-HCl	Tris Hydrochloride
TWGDAM	Technical Working Group on DNA Analysis Methods
UV	Ultra Violet
Vmax	Maximum Velocity
VNTR	Variable Number Tandem Repeat
Zn	Zinc

LIST OF FIGURES

Chapter I:

Chapter II:

Chapter III:

Appendix - Supplementary Information (Chapter III):

Chapter IV:

Appendix - Supplementary Information (Chapter V):

LIST OF TABLES

ACHIEVEMENTS

Patent application

Provisional patent - Invention Title: Nucleic Acid collection device, 2016.

Poster and oral presentations:

1) "The International Society for Forensic Genetics" conference, Melbourne, Australia, 2013. ***Poster presentation.***

2) "The Australian and New Zealand Forensic Science Society" conference, Adelaide, Australia, 2014. ***Oral presentation and poster presentation.***

3) "The American Academy of Forensic Sciences" conference, Orlando, Florida, America, 2015. ***Oral presentation.***

Recognition for thesis work:

- Awarded the Inaugural Ross Vining Memorial Award presented by the Australian and New Zealand Forensic Science Society (ANZFSS) conference, in 2014, for the best South Australian student abstract.
- Cash prize, runner up at school level, in the Flinders University 3 minute thesis (3MT) competition in 2014.
- News feature article, in the journal Biotechniques "Fingerprinting the Fingerprints: The Past, Present, and Future of Molecular Forensics", May 2015.

Scholarships and grants awarded:

1) John Stocker postgraduate PhD top-up scholarship provided by the Science and Industry Endownment Fund (SIEF);

2) The Australian and New Zealand Forensic Science Society (ANZFSS) travel award and;

3) The Flinders University research higher degree travel scholarship.

CASEWORK

Case 1 - Drug seizure Methamphetamine case, South Australia. *The application of direct PCR to generate a DNA profile.*

PEER REVIEWED PUBLICATIONS

Publications are listed below that have been awarded during the writing of this thesis that I have contributed to as first or co-author, and show a depth of knowledge and understanding in areas of low-template DNA typing:

Reference	Impact Factor	Citations	Year of publication	Chapter in thesis
1. A Linacre, **JEL Templeton** (2014) Forensic DNA profiling: state of the art. **Research and Reports in Forensic Medical Science** 4:25-36.	Not listed	2	2014	Chapter I
2. **JEL Templeton,** Genetic profiling from challenging samples: Direct PCR of touch DNA. **Forensic Science International: Genetics Supplement Series** 4.1 (2013): e224-e225.	Not listed	9	2013	Chapter III
3. **JELTempleton**, D Taylor, O Handt, P Skuza, A Linacre. Direct PCR improves the recovery of DNA from various substrates. **Journal of the Forensic Sciences** (accepted April 2015).	1.705	3	2015	Chapter III
4. R Ottens, **JEL Templeton**, V Paradiso, D Taylor, D Abarno, and A Linacre. Application of direct PCR in forensic casework. **Forensic Science International: Genetics Supplement Series** 4, no. 1 (2013): e47-e48.	Not listed	9	2013	Chapter III

5.	**JEL Templeton**, A Linacre (2014). DNA profiles from fingermarks. **BioTechniques** 57(5):259-266.	2.754	11	2014	Chapter IV
6.	**JELTempleton**, O Handt, D Taylor, and A Linacre. DNA profiles from fingermarks: a mock case study. **Forensic Science International: Genetics Supplement Series** 5 (2015): e154-e155.	Not listed	1	2015	Chapter V
7.	L Bastien, P Brotherton, K Mitchell, **JEL Templeton**, VA Thomson, *et al*. "Late Pleistocene Australian marsupial DNA clarifies the affinities of extinct megafaunal kangaroos and wallabies." **Molecular biology and evolution** 32.3 (2015): 574-584.	9.105	7	2014	Discussion Appendix
8.	C Sarkissian,P Brotherton, Balanovsky, **JEL Templeton**, B Llamas, J Soubrier, V Moiseyev, V Khartanovich, A Cooper, W Haak, The Genographic Consortium. Mitochondrial Genome Sequencing in Mesolithic North East Europe Unearths a New Sub-Clade within the Broadly Distributed Human Haplogroup C1." **PloS One** 9.2 (2014): e87612.	3.730	12	2014	Discussion Appendix
9.	**JEL Templeton,** Brotherton, P. M., Llamas, B., Soubrier, J., Haak, W., Cooper, A., & Austin, J. J. DNA capture and next-generation sequencing can recover whole mitochondrial genomes from highly degraded samples for human identification. **Investigative Genetics** 4.1 (2013): 1-13.	Not listed	26	2013	Discussion Appendix

10. P Brotherton, H Wolfgang, **JEL Templeton**. Neolithic mitochondrial haplogroup H genomes and the genetic origins of Europeans. **Nature Communications** 4 (2013): 1764.	11.470	92	2013	Discussion Appendix
11. C Santos, M Fondevila, D Ballard, R Banemann, AM Bento, C Børsting, W Branicki, F Brisighelli, M Burrington, T Capal, L Chaitanya, R Daniel, V Decroyer, R England, KB Gettings, TE Gross, C Haas, J Harteveld, P Hoff-Olsen, A Hoffmann, M Kayser, P Kohler, A Linacre, M Mayr-Eduardoff, C McGovern, N Morling, G O'Donnell, W Parson, VL Pascali, MJ Porto, A Roseth, PM Schneider, T Sijen, V Stenzl, D Syndercombe Court, **JE Templeton**, M Turanska, PM Vallone, RAH van Oorschot, L Zatkalikova, Á Carracedo, C Phillips, EUROFORGEN-NoE Consortium.(2015) Forensic ancestry analysis with two simple capillary electrophoresis ancestry informative marker (AIM) panels: Results of a collaborative EDNAP exercise **Forensic Science International Genetics** 19, 56-67.	4.604	5	2015	Discussion Appendix
12. JEL Templeton, Duncan Taylor, Oliva Handt, Adrian Linacre. Typing DNA profiles from previously enhanced fingerprints using direct PCR. Under review in **Forensic Science International: Genetics (Accepted).**	4.604	-	May 2017	Chapter V

i. R Blackie, **JEL Templeton**, D Taylor, and A Linacre. PCR buffer enhancement of STR kits used for human identification. **International Journal of Legal Medicine**. ***(In preparation).***	N/A	N/A	In preparation	Chapter IV

Chapter I:

Introduction

Manuscript enclosed:

Linacre, A., & Templeton, J. E. L. (2014). Forensic DNA profiling: state of the art. Res Rep Forensic Med Sci, 4, 25-36.

1.1 Preface

This introduction describes the history of DNA typing for the purpose of human identification and details current methodology that is used to generate DNA profiles from fingermarks - a common form of trace evidence typical in forensic investigations. Chapter 1 concludes with a published review on the use of DNA profiling for human identification (see manuscript enclosed: ***Linacre, A., & Templeton, J. E. L. (2014). Forensic DNA profiling: state of the art. Res Rep Forensic Med Sci, 4, 25-36)***.

Recent advances in DNA technology, and sensitivity, have enabled biological samples that are invisible to the naked eye to be successfully analysed [1-5]. In spite of this, there are often low success rates for low-template and degraded DNA samples [2]. '*Every contact leaves a trace*' postulated by Dr Edmond Lochard is a mantra familiar to forensic science [6], but all too often this trace is too small to analyse, is invisible to the naked eye, or cannot be recovered from the surface on which it was deposited.

Individuals transfer their DNA to objects, simply by touch – so called 'touch' DNA [5]. The terms 'touch' DNA, contact DNA and trace DNA are used in the same context in the literature. When an object has been touched, epithelial cells or cell-free DNA may be transferred when complexed with oil or sweat residue [7-9]. Human skin is abundant in sweat ducts; approximately 2 – 4 million are distributed within the skin [10] (see Figures 1.1 and 1.2) and they contribute to the amount of sweat and potential cell-free DNA deposited by touch [7]. These features can be exploited when attempting to obtain a DNA profile from a touched object at a crime scene and used to associate or exclude witnesses and suspects.

1.2 'Touch' DNA

Since it was first recognised in 1997 that DNA could be retrieved from fingermarks [5], handled or touched items have become a common source of evidentiary material submitted to a forensic laboratory for analysis [2, 4, 11, 12]. However, it has been found that a significant number of crime-related surfaces, even though they should have DNA present, do not necessarily yield analytically useful material [2-4, 13].

More sensitive commercial kits have become available in the last few years which make it more possible to analyse trace material and to further appreciate the complex nature of ‘touch’ DNA samples [14-17].

Despite the likelihood of a useful deposit being present on touched items, forensic scientists encounter difficulties when attempting to recover and type DNA from a substantial number of these items [12, 13, 18]. Although the human genome is massive, forensic science currently only utilizes a small fraction of it for what is now termed ‘DNA profiling’.

Forensic science laboratories around the world make use of a very small part of the human genome – so-called short tandem repeats or STRs – as a highly discriminating barcode to identify individuals [19-22]. It is common knowledge in the forensic science field that many ‘touch’ DNA exhibits fail to generate profiles [2-4, 13]. Identifying and genetically characterising ‘touch’ DNA samples is challenging because DNA is often a mixture from more than one source, degraded [9, 23], limited in quantity (i.e. low-template) [2-4, 24], and may contain elements that co-extract with the DNA [25], hindering subsequent amplification. All of these aspects are discussed in this introductory chapter.

The standard process of isolating ‘touch’ DNA from a touched surface is swabbing or tape lifting (which may not be effective at recovering all of the DNA present), extraction/purification (where DNA is lost) and capillary electrophoresis (CE) to separate amplicons (amplified product) (see Figure 1.3). Loss due to extraction and purification is of particular importance when samples possess limited DNA to begin with, as is the case with trace DNA, and any additional loss due to extraction may prevent amplification of STRs [13, 26-33]. As a result, standard protocols often recover sub-optimal levels of nuclear DNA that can result in a poor quality STR-based DNA profile or fail to produce a profile at all [11, 32, 34].

The area of ‘touch’ DNA profiling gets much more complex depending on: the nature of the surface that was touched [18, 35-41]; pressure and friction used during transfer [4]; the time since DNA deposition [40]; the shedder status/characteristics of

the donor [41, 42]; quality and quantity of DNA deposited (i.e. low-template or degraded) [23, 43]; and presence of PCR inhibitors/environmental factors [25, 44]. Some of these aspects will be investigated in this thesis, more specifically, nature of substrate, time of deposition, and environmental exposure.

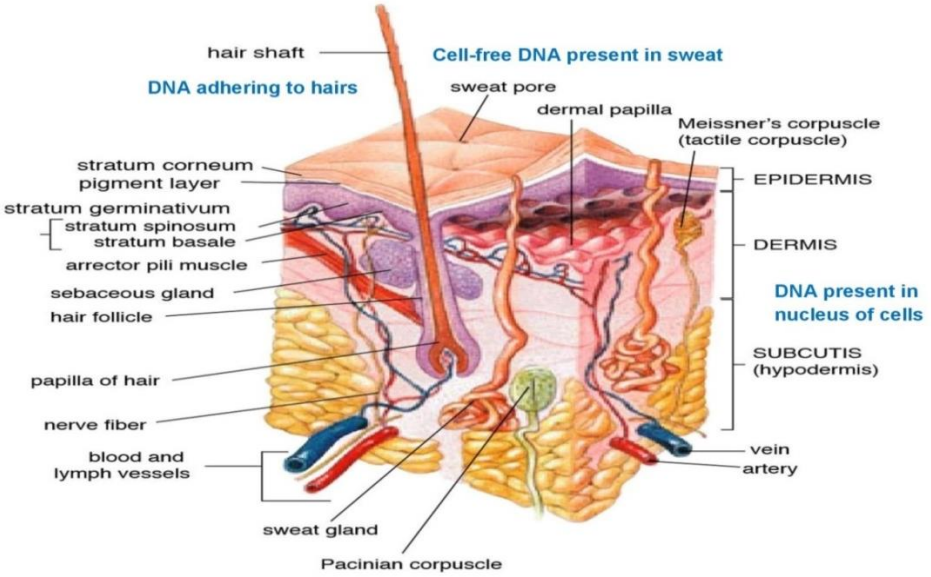


Figure 1.1 Showing skin layers, hair follicles, sweat glands, and sebaceous gland. Image reconstructed from: Human skin: https://en.wikipedia.org/wiki/Human_skin [45].

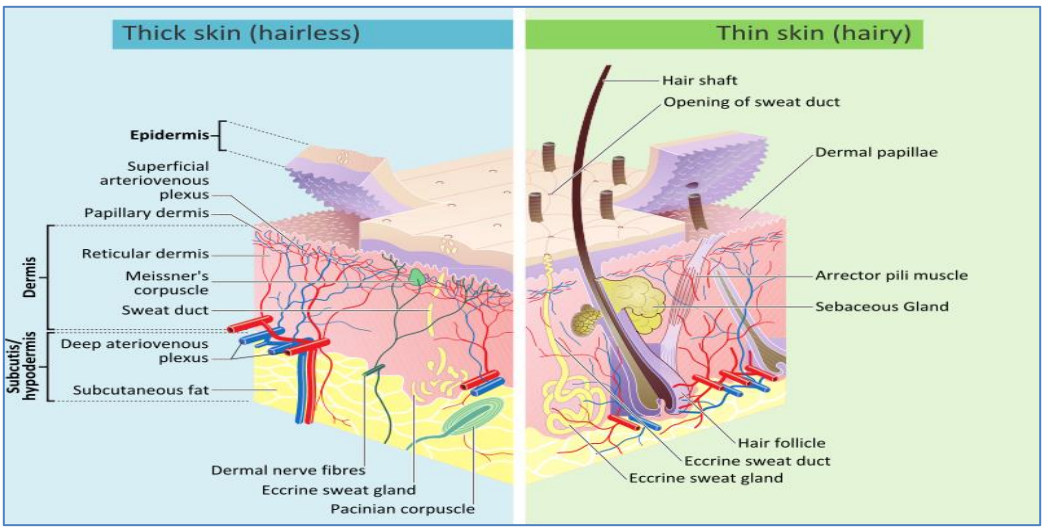


Figure 1.2 Showing layers of skin in both hairless skin (typical of fingermarks) and skin containing hairs (from other regions of the body). Image reconstructed from: Layers of the skin: https://en.wikipedia.org/wiki/Human_skin [45].

It has been demonstrated that transfer of DNA between substrates and surfaces makes it possible for forensic scientists to isolate and identify DNA present on handled or touched items [5]. This advancement in the field initially led to a wide range of exhibits being submitted (e.g. weapons, cartridge cases, clothing, personal items and drug seizures) to casework laboratories [2] for testing. Yet, Forensic Science South Australia (FSSA) has stopped accepting many 'touch' DNA items for analysis (e.g. brass cartridge cases) due to the very small likelihood of obtaining a DNA profile. A new method for recovering more DNA from touched substrates is therefore required. Addressing this need is the main aim of the work described in this thesis.

1.2.1 'Shedders'

Day-to-day activities such as brushing hair, brushing teeth, or blowing a person's nose can result in individuals shedding tens to thousands of skin cells per day [11]. It was originally believed that good and poor 'shedders' exist [40], however this concept has been challenged [41]. Cell-free DNA has been described as 'free-to-roam' DNA not encapsulated by a cell nucleus [7] and may contribute to the 'touch' DNA content. The physical/emotional state of a person and the amount of sweat they produce may influence the amount of DNA that is deposited [7, 46, 47]. The 'shedder' status of an individual has been defined as the amount of DNA deposited in the form of cell-free nucleic acid and/or skin epithelial cells [7, 40]. For example, a good 'shedder' will deposit large amounts of DNA in comparison to a poor 'shedder' who generally leaves behind less than optimal DNA for subsequent analysis [40].

The complexity of factors involved with 'touch' DNA samples makes it difficult to form a conclusion on aspects of 'shedder' status, how long DNA lasts, and the likelihood that a certain individual deposited the DNA by primary, secondary or tertiary contact. One study by Lowe *et al.* (2002) [40] determined the 'shedder' status of an individual by assessing their ability to deposit DNA on an item 15 minutes post-hand washing. All data were assessed using the low copy number (LCN) (i.e. 34 PCR cycles) process (discussed below). In this study, a person was defined as a good 'shedder' if they left behind a full DNA profile after hand washing. In contrast, a poor 'shedder' was one who produced a partial or incomplete profile. The findings conclude that

'shedder' type is not an influencing factor in determining how much DNA is left behind when the period between hand washing and depositing DNA is greater than 6 hours [40]. The results demonstrate 100% success for all individuals tested in their ability to leave behind a full STR-DNA profile 6 hours post-hand washing, regardless of 'shedder' status. Secondary transfer was observed in the study when DNA from one individual was transferred to another and subsequently to an object resulting in a mixed DNA profile [40]. Secondary and tertiary DNA transfer events have been observed in other published work [36, 37, 42, 48-53].

Other work has contradicted the existence of 'shedder' status by observing substantial differences amongst their volunteers in their ability to deposit DNA [41]. Thomasma *et al.* (2012) [28] noticed a significant difference in DNA yield depending on the digit used to deposit the fingermark. The authors observed a higher amount of DNA recovered from index fingers (compared with other fingers) and proposed that index fingers experience faster rates of dermal tissue turnover (see Figure 1.2) owing to heavier use, and this may contribute to the higher quantity of cells/DNA. Another explanation could be that index fingers are commonly used over other fingers to touch parts of the body more often, such as the hair and face. Phipps *et al.* (2006) [41] were unable to characterise the 'shedder' status of 60 volunteers, as repeated testing produced variable results. The authors observed that individuals deposit different amounts of epithelial cells through contact at different times of the day. Further work coincides with these findings by reporting variation in DNA amounts transferred by the same individual [54]. Based on this research, certain biological and environmental factors need to be assessed before commenting on the 'shedder' status of an individual. Or perhaps forensic scientists need to accept that we cannot comment on the 'shedder' status of an individual due to the large number of variables involved.

1.2.2 Low success rate for 'touch' DNA profiling

Improvements in nucleic acid purification and amplification techniques [55-57] make it easier - but not always possible - to obtain DNA profiles from touched objects. The low amounts of 'touch' DNA that scientists work with are problematic. Incorrect handling of samples at the scene may contribute to problems with contamination [49,

58, 59] and invalid results (see below, case study 2). Given its simplicity, direct Polymerase Chain Reaction (PCR) (described in detail later in this chapter) is one method to improve the sensitivity of detection and enable faster processing; a genotype can be generated quicker than usual (i.e. 2 – 3 hours) (see Figure 1.4) to exclude or confirm a match to a DNA profile.

It is common knowledge in forensic science that handled items are one of the most common forms of evidence submitted to a forensic science laboratory for testing. The study by Harbison *et al.* (2008) [2] highlights the importance of 'touch' DNA at the Institute of Environmental Science and Research Limited (ESR) in New Zealand. In 2003, the laboratory received 4447 samples crime scene samples, of which 1990 were submitted for low-level trace DNA analysis [2]. However, due to the low success rate, many handled items submitted to a forensic laboratory are not processed. Reports by Harbison *et al.* (2008) [2] and Raymond *et al.* (2009) [12] demonstrate a low success rate with 'touch' DNA in their laboratory and reported that only 5-6% of handled items generated a full DNA profile. It should be noted that older versions of STR profiling kits were used to generate the data in these studies and since then newer versions of STR kits (with enhanced buffers) have become available that optimise profiling success for low quantity/degraded DNA samples.

Two cases examples are described below that highlight the importance of 'touch' DNA evidence for forensic investigations. Both case examples utilised low copy number (LCN) methodology; a technique that has been heavily debated for credibility.

1.2.3 Two high profile case examples that utilised 'touch' DNA evidence

CASE EXAMPLE 1 – *using low copy number (LCN) technique*

One high profile case that utilised touch DNA evidence in 2005 was the murder of backpacker Peter Falconio who was ambushed in the Australian outback when driving in a campervan with girlfriend Joanne Lees. Bradley Murdoch was accused of the murder of Peter Falconio and the abduction and assault of girlfriend Joanne Lees after flagging down the camper van they drove in Barrow Creek in the Northern Territory, Australia, in 2001. Ultrasensitive DNA testing proved vital to solving the case and linked the drug runner Bradley Murdoch to the murder. The DNA profile of Bradley Murdoch was extracted from a bloodstained top the female victim had been wearing that evening, and 'touch' DNA was obtained from tape used to tie the victim's hands, and from the gear stick of the camper van. The chief prosecutor in the case (Mr Wild) described the DNA evidence as the "linchpin" in the case. "*There may be an innocent explanation for each transfer of DNA, such as, secondary contact or by another person with DNA matching Mr Murdoch, but this would be highly unlikely*". The DNA recovered from the bloodstained top was "*150 million billion times more likely to have originated from Mr Murdoch than from another individual*". The technique used for obtaining 'touch' DNA samples from the tape and gear stick was LCN (discussed later) [1]. LCN was used to amplify a low amount of DNA originating from a small number of cells present on the evidence.

CASE EXAMPLE 2 – *using low copy number (LCN) technique*

The Omagh car bombing in Northern Ireland in 1998 described as Northern Ireland's worst terrorist attack is the second example of the use of ultrasensitive DNA testing in a high profile case. Despite a DNA match being made that linked the accused to the tape found on the bomb timer and explosive wiring, the evidence was not admissible in court. The use of LCN technology to generate DNA profiles was criticized by the judge as not credible evidence due to the lack of adequate chain of custody, inadequate handling of evidence, and the method being relatively novel and having not gained 'general acceptance' in the science community. Samples were processed with inadequate LCN controls for sample handling and collecting. The accused was later acquitted based on the way the evidence was handled. In response, the Caddy report was commissioned by the UK Home Office - an expert review of the LCN process conducted by Professor Brian Caddy, Dr Graham Taylor and Professor Adrian Linacre – who ultimately concluded that low-template DNA typing was fit for purpose [60].

1.3 DNA used in Identification

The human body contains approximately 100 trillion cells. The genetic makeup of the individual is contained within the cell, in the nuclei in the form of nuclear (nu) DNA, and the mitochondria in the form of mitochondrial (mt) DNA. A person's genome is their entire set of DNA molecules in a given cell (i.e. the basic unit of life) used to code proteins and to aid cell function and division. Nuclear DNA is passed down from generation to generation as a combination from each parent and resides in every cell in the human body (with the exception of red blood cells). DNA contained within the human cell is bundled into a total of 46 chromosomes (i.e. 22-paired autosomal and two-sex determining chromosomes (XY)).

One copy of the human genome consists of approximately 3.2 billion base pairs (bp). Approximately 99.9% of the sequence of the bases are identical in all individuals, so it is the 0.1% variation that forensic biologists use to differentiate and link individuals [61]. Prior to the use of DNA technology, serum proteins and blood group antigens were used to identify samples; however, these methods lacked discrimination.

Following a forensic breakthrough in 1985, DNA profiling was developed by Sir Alec Jeffrey. DNA was isolated and compared to other samples by examining repeated sections of DNA sequence that are known to differ between individuals in length and sequence content. The analysis of mini and micro satellite variable number of tandem repeats (i.e. VNTRs) in the human genome was performed and this is discussed below. Early tests to analyse VNTRs used restriction enzymes to cut regions of interest and the technique employed was known as restriction fragment length polymorphism (RFLP). Minisatellite VNTRs that were initially used are too big for routine forensic DNA applications [62-64] that require techniques suited to degraded fragments (i.e. 100–400 bp length). Now the most suitable markers for human identification are short tandem repeats (i.e. STRs) [19, 20, 65, 66] that analyse smaller regions of nuclear DNA in the genome, or single nucleotide polymorphisms (SNPs), mentioned below.

Other DNA markers exist to aid forensic identification, such as SNPs. Regions of interest in the chromosome are called loci, and the variations of DNA sequence at a given locus are known as alleles. A SNP assay of 50 loci has similar discriminatory power to approximately 12 STR loci [67]. Fragment sizes for SNP assays favour degraded DNA samples as the amplicons targeted are shorter than standard STRs [67-81]. Additional information can be obtained from individuals by using SNP assays, such as phenotypic [82, 83] and ancestral informative [74, 78], in addition to analysing Y-chromosome markers [84]. More information on SNP genotyping can be found in the review that is enclosed at the end of this chapter, see: **Linacre, A., & Templeton, J. E. L. (2014). Forensic DNA profiling: state of the art. *Res Rep Forensic Med Sci*, *4*, 25-36.**

Additionally, mitochondrial DNA (mtDNA) can add value to a forensic investigation when STR typing fails [34] or maternal lineage information is required [85]. MtDNA typing is a useful tool for human identification owing to the fact that mtDNA exists in high copy number in the cell (i.e. approximately 100 – 10,000 copies) in comparison to nuclear DNA (i.e. 2 copies). MtDNA typing is more likely to generate a result in samples where nuclear DNA is highly compromised [34, 86-92]. Supplementary information on the use of mtDNA typing for forensic purposes is included in the

manuscript enclosed, see: **Linacre, A., & Templeton, J. E. L. (2014). Forensic DNA profiling: state of the art. *Res Rep Forensic Med Sci*, *4*, 25-36.**

The success of a DNA profile depends largely on how much DNA is retrieved from the evidence and the quality of the DNA fragments. Various collection approaches are used to recover DNA from touched items, such as self-adhesive security seals [56], minitape and other types of tape lifts [93-96], and swabs [26, 97-100]. Despite the advantages of using tape on certain substrates (e.g. clothing), the most common way to recover and retain DNA from a hard surface is the use of a sterile, DNA-free swab [98]. Double swabbing is thought to improve DNA recovery (i.e. one wet swabbing event followed by a second dry swabbing event).

Cotton swabs have been used by forensic examiners for years [26, 97, 98, 100, 101] for the retrieval of DNA from hard surfaces. The ability to recover the entire DNA adhering to the swab is a limiting step and crucial to the success of obtaining a DNA profile. Improving methods for sample collection would have a considerable effect on the overall profiling result as more DNA template would be made available to the PCR. Most importantly, the process of DNA retrieval should be carefully performed to avoid extraneous DNA contamination (see Chapter II; contamination controls for low-template DNA).

The forensic DNA community strives to achieve standardization by agreeing on a core set of STR loci and standard methodologies to generate data [65, 102]. On the other hand, standardizing methods associated with newer technologies such as massively parallel sequencing technologies will take longer to achieve. A standard workflow for DNA profiling - from extraction to analysis of data - is illustrated in Figure 1.3 and can take between 8–10 hours depending on the number of samples to be processed (i.e. ~2 hours for extraction, ~2 hours for quantification, ~2-3 hours for PCR and 1 hour for capillary electrophoresis (CE) detection). Ideally, processing samples in a quicker time frame than is currently possible, and maximizing the amount of DNA recovered, would enable a more effective forensic investigation. Isolating the DNA during the extraction and purification step is where loss of DNA is thought to occur [13, 29, 32, 101, 103, 104].

1.3.1 Standard workflow for DNA analysis (isolation to detection)

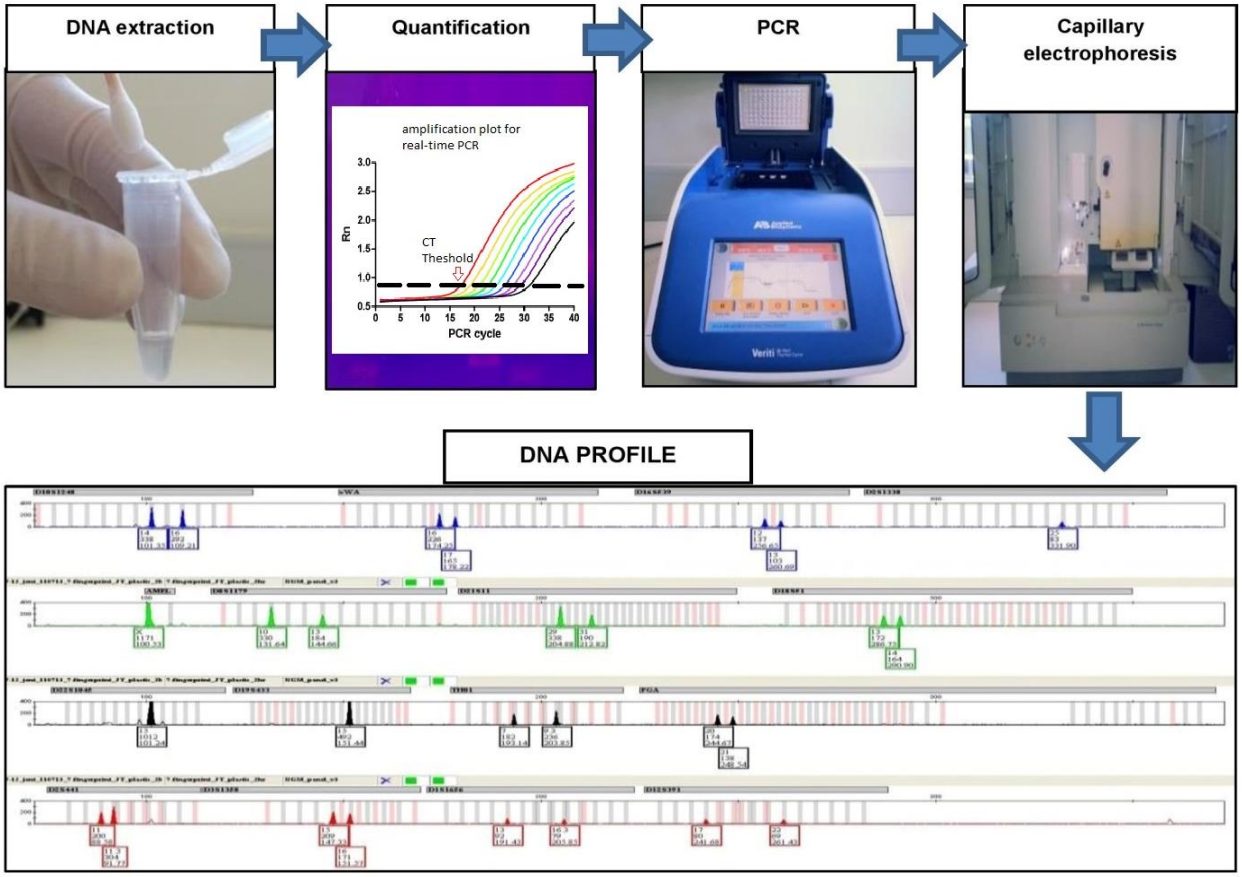


Figure 1.3 showing a standard workflow for forensic DNA profiling. Image reproduced from *'Linacre, Adrian, and J. E. L. Templeton. Forensic DNA profiling: state of the art. Res Rep Forensic Med Sci 4 (2014): 25-36*' [105]. The process starts with DNA extraction and quantification, subsequent PCR to amplify DNA fragments, and finally capillary electrophoresis (CE) detection to separate and size STR alleles for data analysis. Standard high throughput forensic laboratories use automated liquid handling systems to carry out multiple extractions and real-time PCR methods to quantify DNA prior to PCR.

1.3.1.1 DNA extraction

Depending on the method used to isolate DNA, some commonly used extraction methods utilise many reagents, tube changes, and wash steps to extract DNA. However, there is generally insufficient DNA extracted from substrates (e.g. swabs, tape lifts and cloth) in order to generate profiles, and some extraction methodologies have limitations with regards to how much DNA can bind to the columns (e.g. DNA IQ™ kit). It has been reported that between 20–70% of the DNA recovered by a cotton swab is lost when the swab is subjected to an extraction [13]. Ottens *et al.* (2013) [29] report a detrimental loss of DNA of up to 83% following a standard

extraction. Other researchers concur with these findings and describe a loss of DNA following extraction [106, 107].

DNA extraction reagents

Lysis buffer is used to break open the cell membranes and generally consists of several different components (e.g. sodium dodecyl sulphate (SDS), ethylene-diamine-tetra-acetic acid (EDTA), dithiothreitol (DTT) and proteinase K (ProK)) [108]. The nuclear envelope and histones are made of proteins. Proteinase K is known for digesting proteins, removing contaminants, and inactivating nucleases that would otherwise degrade DNA or ribonucleic acid (RNA). Nucleases are enzymes (i.e. proteins) present in the cell that degrade DNA in the presence of magnesium. The presence of a chemical known as EDTA in the extraction mix will chelate metal ions (Mg++ and Ca++) that enzymes such as DNase require to digest DNA. By chelating these ions, DNase enzymes become inactive and the DNA remains intact. DTT induces protein digestion, reduces cross links between DNA and other biomolecules, and breaks disulphide bonds (these are prevalent in keratin and the acrosome surrounding spermatozoa). Ethanol during wash steps removes impurities in the sample and separates DNA from contaminants.

In the 1990s, commercial kits became available that had been optimised for specific sample types. More recently new kits have been optimised for dealing with degraded and low-template DNA [14-17, 55, 57, 109]. Common kits still in popular demand are Promega's DNA IQ™ kit, Invitrogen's ChargeSwitch™ kit and Qiagen's QIAamp DNA Micro™ and Mini™ Kits. Other methods for sample extraction include Chelex-100® (Bio-rad) [110], organic methods (such as phenol-chloroform) [111] and FTA paper [112-114].

Organic extraction

Organic extraction (e.g. phenol-chloroform) predates commercial kits and is successful at extracting high molecular weight DNA required for previous RFLP technology [111]. The method is hazardous and time consuming compared with many commercial kits. Organic extraction used in conjunction with ethanol precipitation is commonly used to improve the rate of DNA recovery from challenging

samples, as the preliminary steps are carried out with sample material contained within the tube initiating a more effective cell lysis. The process is achieved by separation of the organic phase from the aqueous phase. Proteins are soluble in the phenol solution and DNA will remain in the aqueous layer of the mixture, along with other contaminants. The aqueous phase is removed by pipetting. Ethanol precipitation separates DNA in the aqueous phase from other contaminants (e.g. salts and carbohydrates). Spinning via centrifugation and pipetting will contribute to loss of sample DNA as there can never be 100% transfer of DNA from one tube to another. In addition, ethanol precipitation is not 100% effective and not all DNA spins down [111]. In spite of this, some laboratories still perform ethanol precipitation.

Chelex-100® extraction

Chelex is a faster method of extraction involving fewer steps than phenol-chloroform, so less opportunity for contamination, although it generates single-stranded DNA which the operator should be aware of. Chelex is an ion-exchange resin (Bio-Rad Laboratories) that comprises of styrene divinylbenzene copolymers containing paired iminodiacetate ions [108] that aid DNA extraction by chelating divalent metal ions (e.g. magnesium) which are required by nucleases to destroy DNA. Nucleases are no longer active in the presence of a chelating agent and the DNA is protected. Samples are centrifuged in a similar fashion to organic extraction where cell debris is pulled to the bottom of the tube and DNA in the upper aqueous layer can be dispensed into a PCR tube for amplification [110]. Spinning will contribute to a loss of sample DNA. Chelex-100® and organic methods have also been shown to lose up to 75% of target DNA [13, 115, 116].

Solid-phase extraction

Silica-based extraction methods are popular for high-throughput casework that involves multiple samples or ancient DNA work where you might have only one or two precious samples [117-119]. Kits utilize a silica-coated magnetic bead-based approach to capture DNA. Nucleic acids absorb to silica beads, powders, or columns, in the presence of chaotropic salt (e.g. guanidine hydrochloride) [117]. Increasing the salt concentration and decreasing pH will disrupt hydrogen ions in water allowing DNA to come out of solution and bind to silica. The sodium ions present in salt will act as cations that attract negatively charged oxygen ions in silica when pH conditions are optimal (e.g. pH ≤ 7). Strong silica-DNA interaction enables contaminants to be removed via traditional ethanol wash steps; cell debris, impurities and proteins can be separated from sample DNA by purification. DNA is then eluted from the silica under alkaline conditions, heat, and low salt concentration. This solid-phase extraction approach has been adapted to work well for single tubes or a high-throughput workflow (e.g. DNA IQ™ system, Promega) [115, 120]. Depending on the material extracted, metallic ions - common PCR inhibitors in fingerprint powders or on brass surfaces – may interfere with the silica-DNA binding complex by interacting with silica membranes common in solid phase extraction kits [39, 119]. Similar to

other extraction methods, wash steps and tubes changes contribute to loss of sample DNA.

1.3.1.2 Direct PCR vs standard extraction and quantification

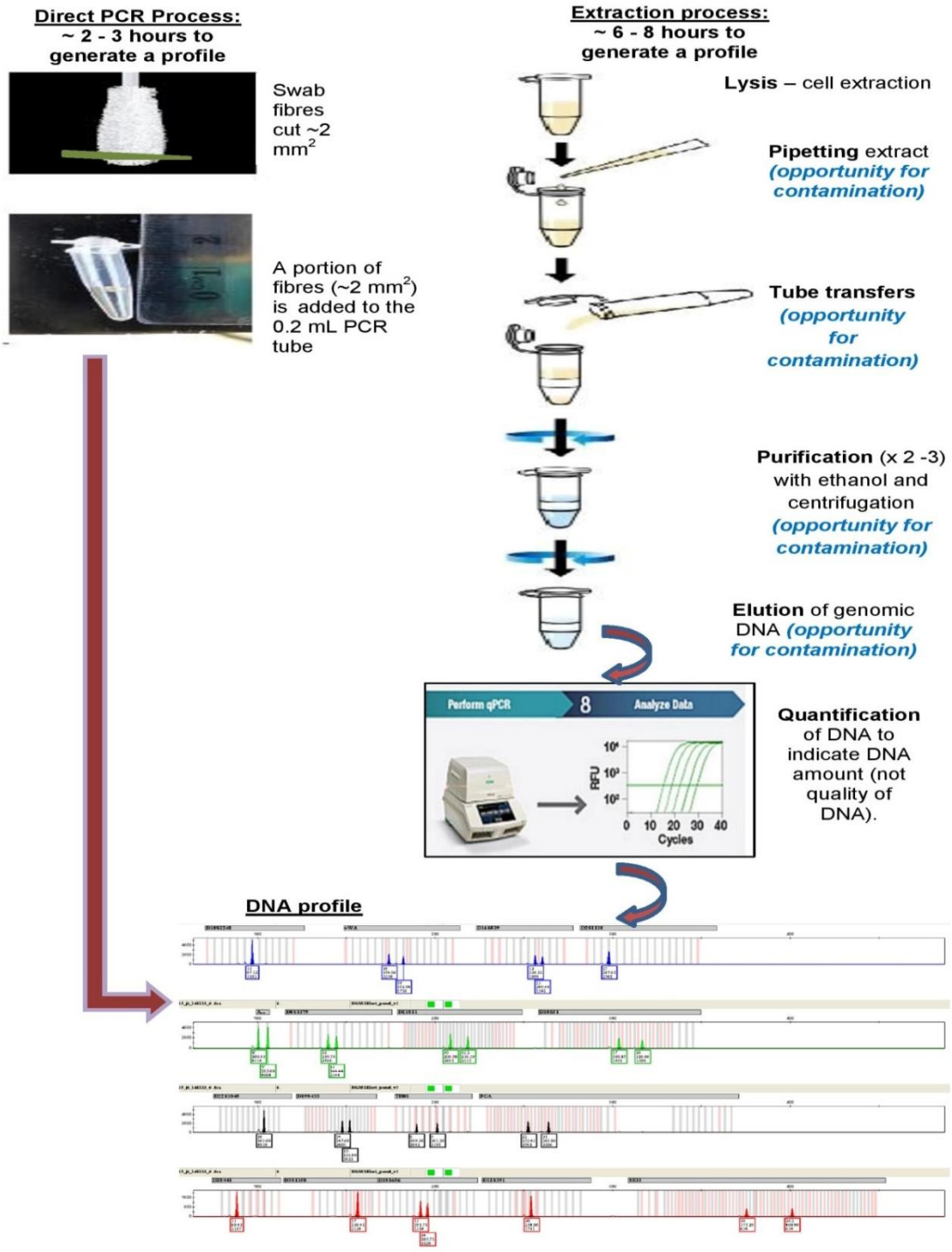


Figure 1.4 Schematic of direct PCR workflow on the left (i.e. extraction stage and quantification is omitted) and the industry standard extraction steps and quantification on the right. Direct PCR requires biological material to be placed directly into the PCR tube, and cell lysis will be achieved by heating the sample (e.g. to 94 ° C) during the initial PCR stage.

1.3.1.3 Loss of 'touch' DNA during extraction

Cell-free DNA has been proposed as a source of 'touch' DNA in forensic cases involving fabrics [33] and sweat [7]. Most forensic workflows now work on absorption methods (i.e. DNA binding to silica, as previously discussed) [121]. However, the loss of DNA is still possible through the multiple wash steps. During a standard DNA extraction process cells are centrifuged and supernatant – thought to contain cell-free DNA - is typically discarded. Organic extraction used in conjunction with ethanol precipitation would limit the loss of DNA. Cell-free DNA was first detected in the 1940s when differences in serum levels were observed amongst sick and healthy individuals [122]. Necrosis, apoptosis, and active secretion are the mechanisms behind which cell-free DNA enters the circulation [7]. The levels of cell-free DNA in an individual may vary depending on behavioural changes, such as extraneous exercise or disease states [46, 123, 124].

A study by Nakamura *et al.* (1999) [125] postulated that cell-free DNA can play an important role in the recovery of DNA from a touched item. Moreover, in 2011 Quinones *et al.* [7] recovered cell-free nucleic acid in the sweat of 80% of healthy individuals and found that nucleated cells were also noticeable in sweat. The authors report that 11.5 ng of DNA was extracted from 1 mL of cell-free sweat [7]. Likewise, another study revealed cell-free DNA as a component of the 'touch' DNA sample [126].

Analysts need to consider that current extraction methods generally discard the supernatant aqueous portion of the extract which may retain cell nucleic acid and commercial kits have limitations. Future methods aimed at improving DNA recovery from touched items need to take into account the cell-free nucleic acid portion of the sample as this could be a valuable contributing factor to generating a meaningful DNA profile.

1.3.1.4 Optimising extraction processes

Methods exist to optimise current methodology for trace DNA and include concentration of DNA using spin columns, such as Microcon® (Millipore, MA, USA), Nucleospin® (Clontech, CA, USA), or MinElute® (Qiagen, CA, USA) columns [3, 127,

128]. However spin-columns have also been shown to contribute to a loss of sample DNA [13]. The addition of carrier RNA or salmon sperm DNA to the sample tubes has the potential to prevent loss of valuable sample DNA [119, 127]. The addition of another nucleic acid to the columns will prevent sample DNA from irreversibly binding to sacrificial sites on polypropylene tubes, silica membranes, and plastic wear, and should prevent further loss. If the thermodynamics of DNA adsorption to silica is driven by an increase in entropy, in the presence of chaotropic salt, as previously suggested [117, 119, 129, 130], the addition of another nucleic acid (i.e. carrier RNA) may further drive DNA adsorption to silica by competing for water molecules that are not already bound to the chaotrope. However, in spite of this, carrier RNA may interfere with subsequent analysis of the sample as it will co-elute with DNA, so a DNA-specific detection method (such as PCR) is required.

FTA™ paper

An alternative approach for extracting DNA came about with the manufacture of FTA™ paper designed by Leigh Burgoyne at Flinders University in 1994 [112]. Absorbent cellulose-based paper containing DNA preservatives enable storage and preservation of DNA for long periods of time [113]. FTA™ paper requires a spot of blood, or other body fluid, to be placed on the surface and allowed to dry. Cells are immediately lysed and DNA is fixed within the paper matrix and protected. A section of paper can be removed using a small medical punch device to extract DNA from the substrate. Wash buffer is then added to the paper to remove inhibitors (e.g. haem in the case of blood) and the paper can be added directly to PCR for amplification [112, 131, 132] or extracted further. If future testing is required, a punch from another section of the paper may be used.

A review published by van Oorschot *et al*. (2010*)* [3] expressed how beneficial it would be to add the sample containing DNA to the same reaction vessel as the PCR reagents. The authors reported that direct PCR methods were not yet capable of amplifying template DNA adhering to swab fibres and the authors conveyed how advantageous this would be for trace DNA. Published work does not yet describe the application of direct PCR for case work samples, however operational laboratories

such as FSSA already use direct PCR on FTA reference samples, and the work described in this thesis provides additional support for the application of direct PCR.
The data chapters provided in this thesis focus experimental design around the direct PCR approach and the ability to utilise swab fibres for direct PCR amplification (see Figure 1.4).
Following on from DNA extraction, the next steps in the process are quantification and amplification to produce fluorescently labelled PCR products.

1.3.2 Quantification

Most laboratories require the DNA sample to be quantified prior to PCR to guarantee optimal input DNA for PCR to succeed and to retain sufficient DNA for secondary analysis if required. Laboratories that adhere to the Quality Assurance standards of the FBI require quantification steps to be carried out. Sub-optimal sample types, such as trace DNA, are unlikely to overload a PCR instrument with too much DNA. Rather than using a proportion of the sample to quantify the DNA extract, the whole amount of DNA could be retained for the amplification of STRs in a direct PCR approach. Quantification values predetermine the characteristics of the DNA by estimating the amount of template, but not the quality of DNA; and quality is also a major contributing factor to the success of DNA profiling. For example, a sample with high amounts of degradation may not generate a high quality DNA profile despite passing quantitative threshold values. Likewise, a sample that failed to reach adequate quantification thresholds as set by the laboratory may still be capable of generating an informative DNA profile. More recent Quantification kits such as Quantifiler Trio™ (Applied Biosystems) and Powerquant® (Promega) now have degradation indicators included to assess the fragment size and presence of inhibitors. Quantification steps remain a prerequisite to testing in many forensic DNA laboratories.

Early methods for estimating DNA quantity involved ethidium bromide staining of gels [133]. Fluorescent bands present on the gel are specific to the molecular weight of the detected fragments. Initial tests proved non-specific to human DNA [134]. Most importantly, a large volume (by forensic standards) of DNA was needed for testing and hence only good for large blood stains etc. More recent quantification

methods utilise target-specific fluorescence for specific applications, such as the Qubit™ (Applied Biosystems) and Bioanalyzer (Agilent Genomics) instruments. The most popular method for high-throughput forensic laboratories is real-time PCR, discussed below.

Slot blot quantification

A human and higher primate specific DNA test was developed in the 1990s to target a primate-specific alpha satellite DNA sequence D17Z1 on chromosome 17 using a 40 bp probe [135, 136]. Extracted DNA is used as the template for probe hybridisation on a nylon membrane and chemiluminescent or colorimetric signal intensities are recorded. Unknown sample DNA is compared against reference calibration standards of known DNA amounts. DNA can be serially diluted to prepare standards which are run alongside sample DNA for comparison [108]. Chemiluminescence is more sensitive and can detect trace levels of DNA down to 10 – 40 pg [137].

PicoGreen quantification

The UK Forensic Science Service (FSS) developed the PicoGreen assay in the 1990s to deal with high volume casework in a 96-well format. PicoGreen is a fluorescent inter-chelating dye which emits greater fluorescence when in the presence of DNA [138, 139]. Unknown sample DNA is measured by comparison to a standard curve of known DNA amounts based on the fluorescence intensity. A small volume of extracted DNA (e.g. 5 µL) is mixed with the dye (e.g. 195 µL) and examined by a fluorometer [108]. Although the assay is not specific to human DNA, it can be a useful indicator of how much DNA to add to the PCR [138].

ALUQuant™ human DNA

AluQuant probes are used to target regions in the human genome known as ALU repeats [140]. The assay is human specific and detects a wide range of DNA quantities (i.e. 0.1 ng/µL – 50 ng/µL) [108, 140] through comparison to a standard curve. The assay works when a hybrid complex is formed between probe and DNA template which triggers enzymatic reactions, ending in oxidation of luciferin and

production of light which is directly proportional to the quantity of DNA [108, 140]. The intensity of light produced is captured using a luminometer.

End-point PCR

Another human-specific method to detect and quantify a region of interest in the human genome (e.g. STR marker) is through the use of a fluorescent intercalating dye (e.g. SYBR Green) [141-143] and end-point PCR. Similar to other quantification methods, a standard curve is generated with known DNA dilutions alongside unknown sample DNA to compare signal intensities. SYBR Green dye intercalates between double-stranded DNA and the signal intensity is directly proportional to the amount of PCR product in the sample [141]. PCR inhibitors present in the sample can be monitored using this assay and the test indicates how much sample DNA is required for subsequent STR-based DNA profiling [44]. Similarly, other intercalating dyes have been used for the quantification of DNA with success (e.g. EvaGreen) [144].

Real-time PCR (or quantitative PCR)

This technique uses the same molecular mechanism as End-point PCR to determine copy number. However, real-time provides the analyst with the ability to watch the PCR proceed from cycle-to-cycle through a monitor and quantitative PCR (qPCR) equipment. Detection equipment is designed to detect the fluorescence intensity of qPCR product in a closed-tube system. Two common methods of qPCR include either the SYBR Green assay or TaqMan assay [138, 142, 145, 146]. The SYBR Green approach takes advantage of the intercalating dye properties of the SYBR Green dye, which positions itself in-between double-stranded DNA [142], and emits light confirming the presence of PCR product at each cycle. TaqMan is referred to as the 5’ Nuclease assay [146] which utilises two fluorescent dyes that emit light at different wavelengths. DNA samples that pass the quantitative PCR threshold set by the laboratory (e.g. ~ 100 pg) will continue with STR-based PCR analysis.

1.3.3 Polymerase Chain Reaction (PCR)

PCR reagents

Forward and reverse primers are required for each marker of interest in the DNA molecule. These short sequences of synthetic DNA bind to regions flanking the region of interest and should be specific to that region only, but be well conserved. Primers are designed with prior knowledge of the target sequence to be copied. PCR yield is a direct result of how well the primers anneal to the template DNA. Other components needed for PCR include template (i.e. sample DNA), deoxynucleotide triphosphates (i.e. dNTPs) - the building blocks for each of the four nucleotides, buffer (i.e. magnesium chloride ($MgCl_2$)), bovine serum albumin (BSA), potassium chloride (KCl) and Tris hydrochloride (Tris-HCL)), deionized water and thermostable DNA polymerase. The most commonly used DNA enzyme used is a thermostable enzyme *Taq* polymerase, isolated from the bacterium *Thermus aquaticus* [147] that inhabits hot springs. For PCR set up, positive and negative controls are also included for each run. A negative control uses water in place of DNA template to monitor contamination in the surrounding environment or in existing reagents. A positive control (e.g. 1 ng of commercial DNA) can be used to ensure all the reaction components and thermal cycling conditions are working effectively.

STR-based profiling kits are designed to simply add the optimum amount of DNA template (e.g. ~ 0.5 - 1 ng) to a pre-made master mix containing the amplification reagents. Commercial kits are progressing to include more enhanced buffer systems that overcome PCR inhibitors and incorporate more genetic loci for increased discrimination (such as AmpF*l*STR NGM™, PowerPlex ESX™, PowerPlex® Fusion Systems and GlobalFiler™ kits) [14, 17, 148]. Advanced multiplex kits have been re-designed to position primers closer to the repeat units in order to reduce the length of the flanking region (e.g. miniSTR loci and GlobalFiler™ kits) [3, 15-17, 109, 149, 150], and this feature is beneficial for degraded DNA [57, 80, 149-152].

PCR technique

PCR produces millions of copies of each region of interest when optimal conditions are used (e.g. ~1 ng of DNA template at 28 PCR cycles). Commercial STR profiling kits (See Table 1.1) are optimized for initial DNA starting templates between 0.5 – 2.5 ng of DNA [1, 14, 17, 148, 153]. Validation studies are performed based upon the Scientific Working Group on DNA Analysis Methods (SWGDAM) guidelines [154],

and the specific laboratories internal validation studies. Testing in the laboratory will determine the optimal amount of DNA template and PCR cycle number for the specific STR profiling kit (e.g. 28, 29 or 30 PCR cycles are standard for most STR kits). Commercial kits should amplify all loci to the same extent, that is, alleles from the same locus should be of similar heights (e.g. 60–70%) to prevent heterozygote imbalance. If there is too much DNA available in the sample, or inhibition, split peaks result from incomplete adenylation (see Figure 1.5 (a)). Also, Off-ladder peaks, locus imbalance, and a noisy background (i.e. high fluorescence), may result from high amounts of template DNA. Less than optimal DNA can cause locus-locus imbalance or stochastic effects such as heterozygote imbalance (see Figure 1.5 (b)) that result in allele ‘drop-out’ [1, 21, 155, 156] (see Figure 1.6).

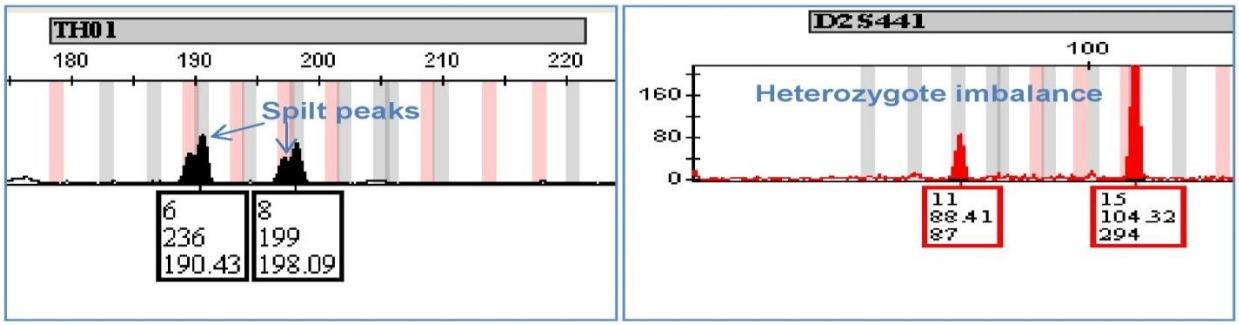


Figure 1.5 Showing (a) split peaks observed from incomplete adenylation (b) peak height imbalance where PCR efficiency of one allele is reduced.

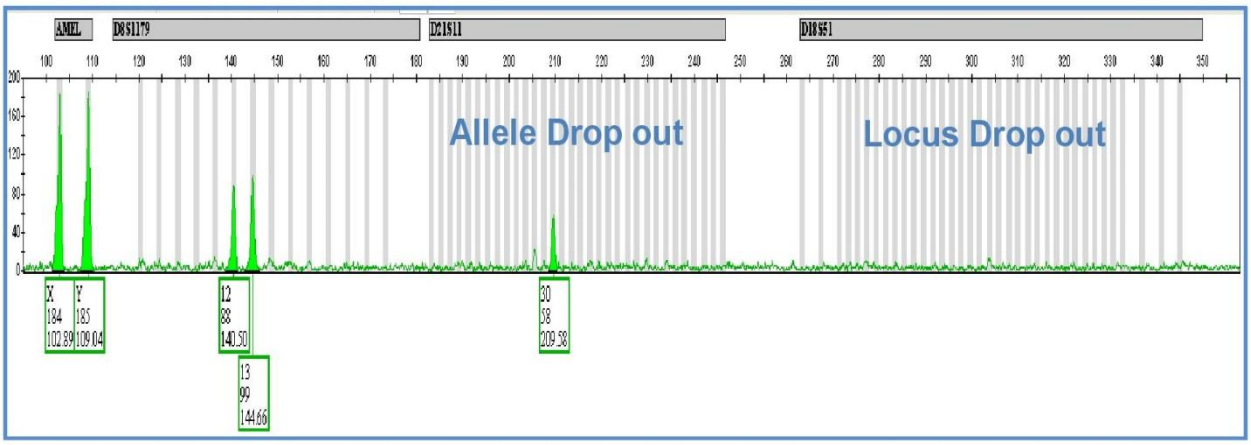


Figure 1.6 Example of allele 'drop-out' and locus 'drop-out'; commonly observed when dealing with low-template DNA. Allele 'drop-out' is the result of preferential amplification and careful interpretation is required to avoid assigning false homozygosity.

Table 1.1 Multiplex STR Systems

There are now comparable STR multiplex systems offered by Applied Biosystems (now part of ThermoFisher), Promega and Qiagen. Some of these kits are listed in the table provided to give an example of the increased power of discrimination achieved throughout the years.

Table 1.1. Multiplex STR amplifications
N.B. CODIS is the combined DNA index system, consisting of 13 STR loci, adopted by the FBI for a national DNA database in 1998.

STR multiplex	Description	Loci examined
"The Quad"; effectively the First Generation multiplex	The initial STR multiplex was used in the UK in 1995. Designed by the FSS.	4 STR loci included; match probability of ~ 1 in 10,000.
'CTT' triplex; effectively the First commercial STR kit	Promega produced the first commercial STR kit in 1994.	3 STR loci included; match probability of ~ 1 in 500.
Second Generation Multiplex, or SGM kit	Improved version of multiplex profiling combination, first used in the UK; 1996.	6 STR loci; has two loci in common with the Quad plus Amelogenin, match probability of ~ 1 in 50 million.
SGM +	Further improvement, widely used in the UK. Kit sold by Applied Biosystems; 1999.	10 STR loci; SGM loci plus 4 more STR loci (plus Amelogenin). 1 in 5 X 10^{12} average probability of identity.
AmpFLSTR Profiler Plus	Applied Biosystems; released 1997.	9 STR loci; all in CODIS. Plus Amelogenin.
AmpFLSTR COfiler	Sold by Applied Biosystems;1998.	6 STR loci; together with the Profiler plus kit will amplify 13 STR CODIS markers.
AmpFLSTR Identifiler	Applied Biosystems; 2001.	15 STR loci; all CODIS plus 2 additional markers and Amelogenin.
Promega Powerplex 2.1	Promega; 1999	9 STR loci; 8 CODIS markers plus 1 additional marker.
Promega Powerplex 16	Promega; 2000	15 STR loci; all CODIS plus 2 STRs and Amelogenin.
AmpFLSTR NGM Kit	Applied Biosystems; 2009.	15 STR loci, plus Amelogenin.
AmpFLSTR NGM SElect Kit	Applied Biosystems; 2010.	16 STR loci, plus Amelogenin. The kit generates the highest power of discrimination of any AmpFLSTR kit to date.
GlobalFiler kit	Released by Applied Biosystems/ThermoFisher 2015, first 6-dye technology.	24-locus STR kit. The inclusion of 10 mini-STRs maximizes results from degraded samples. The most common genotype results in LR of greater than 100 billion to 1[157].

N.B. Microvariant alleles or new alleles that are discovered are now added to a rapidly growing internet database known as STRBase [108, 158].

1.3.3.1 Improvements to STR profiling kits

New versions of profiling kits include optimised buffers, re-design of primers, and PCR cycling optimization [14-17, 148] in order to aid identification of degraded and low-level DNA samples.

1.3.4 Capillary Electrophoresis (CE) and profiling

Following PCR, the ability to visualise STR alleles in the DNA sample is possible by using fluorescence-based detection assays. DNA alleles are visualised as peaks on an electropherogram. Simultaneous analysis of many STR loci can be achieved using multi-colour fluorescent dyes attached to one of the primers in a pair. At present, CE is the method commonly used by the forensic community to analyse these markers. Multiple samples can be processed with CE due to full automation and the instrument requires small sample volume for injection (e.g. 1 µL of PCR product) and minimal hands-on time. The CE platform enables rapid high resolution of closely sized DNA fragments which is a requirement for distinguishing STR alleles.

Fluorescence

A dye molecule attached to a PCR primer is excited upon passing a laser in the CE instrument during electrophoresis, where the negatively charged DNA molecules move toward a positively charged anode [108]. Smaller sized DNA fragments will migrate faster than large sized amplicons under the presence of an electric field. The detection of light emission in the range of approximately 400–600 nm [108] following excitation is then measured. Conformational changes of the fluorophore (i.e. dye molecule) takes place after excitation of a photon converts it from ground energy state to an excited transition state. A resulting photon is emitted at a lower energy (i.e. longer wavelength) than the excitation photon. A stokes shift results from the difference observed between the absorption and emission spectra. Optical filters in the fluorescent detector are designed to pick up particular wavelengths of light and a separation algorithm ‘matrix’ can be used to resolve multiple fluorophores if their spectra properties are distinguishable. Multi-component analysis subtracts the fluorescent signals of neighbouring dyes. A matrix file is created by the computer

software to reflect spectral overlap of dyes by running a set of DNA fragments that are labelled with each specific dye.

If there is considerable overlap between dyes due to poor matrix a phenomenon known as ‘pull-up’ [108] can occur on the resulting electropherogram (see Figure 1.7) where there is colour overlap from one spectral channel into another. Many factors may influence the properties of fluorescent dyes (such as temperature and pH). Therefore, a matrix for a specific dye set should be run frequently to prevent dye colour overlap and subsequent ‘pull-up’.

Figure 1.7 Profile snapshot illustrating ‘pull-up’ where small blue peaks appear at the same size as the green peaks that are true STR alleles. ‘Pull-up’ can be observed if over-amplification occurs, or if the matrix file needs to be re-run due to poor resolution of the dye colours used to label the PCR products.

Charge-coupled devices (CCD) convert fluorophore light intensity signals into electronic signals when they pass by the detector and data is plotted as relative fluorescent units (RFUs) [108]. Filters are designed based on their ability to separate signals (see Figure 1.8). The ability to resolve similar sized STR products relies on the inherent property of fluorescent dyes emitting fluorescence at different wavelengths. STR alleles are distinguished from each other based on the length of the amplicons and colour of the fluorescently attached primer dyes. Electropherograms display the DNA profile results as a series of peaks plotted with fluorescence signal intensity on the Y axis - relating to quantity of DNA - and time that DNA passes the detector displayed on the X axis - relating to size of DNA. Peak morphology provides information about the STR alleles; the combination of all locus genotypes for an individual’s profile will define the sample genotype (i.e. series of numbers) that can be interpreted as a DNA profile.

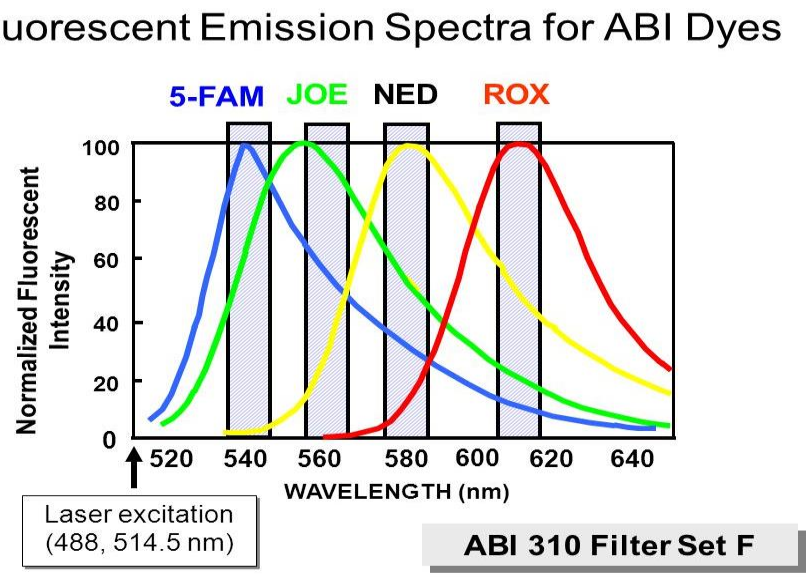


Figure 1.8 Fluorescent emission spectra of ABI dyes used with AmpF*l*STR kits. Image taken from Butler, J.M., Forensic DNA Typing, 2005 [108]. Four overlapping dyes shown need to be deconvoluted by the software to minimise 'pull-up', as illustrated above. Newer STR profiling kits (e.g. GlobalFiler™ now contains six overlapping dyes for higher powers of discrimination) [15, 17].

Genotyping

Electropherogram results are currently analysed using GeneMapper™ ID software (Life Technologies) that resolves dye colours for each marker and is used to size the DNA fragments. Genotyper software can be used to define the sample genotype by comparing the size of alleles at each locus to known size alleles in the allelic ladder. The Local Southern method [133] is an algorithm used by the computer software to measure the DNA fragment size of the alleles by measuring adjacent peaks next to the peak of interest and taking the mean value.

Relative Fluorescent Units (RFUs)

The RFU values of the peaks in an electropherogram relate to the quantity of PCR product (i.e. the greater the amount of PCR product the higher the RFU value). Peak heights and associated RFU values relate to how much fluorescence emission has been detected by the CCD. A laboratory must decide on the analytical peak detection threshold based on intra/inter-laboratory validation studies. Signal-to-noise ratio is assessed by using a series of DNA dilutions and DNA of varying quality for PCR in order to assess the average peak height of the baseline (not considering allelic peaks or stutter). The performance of the instruments used and the sensitivity of the detection equipment will determine appropriate thresholds. RFU values are essential for assigning alleles to loci - which ultimately determine genotype - and to maintain universal standards as part of a validation process. These standards may vary between laboratories from 25 – 400 RFU, depending on whether alleles are heterozygotes or homozygotes [159]. Regardless, it is crucial to separate a true allele from background and many laboratories are progressing to continous models for interpretation (e.g. +/- 10 SD above the baseline to ascertain the laboratory-associated RFU threshold) to reduce error rates [157, 160-162].

A profile is considered ‘full’ when all alleles at all loci are detected above the threshold RFU and the DNA analyst is able to detect a complete complement of genotypic information at each region tested. Exclusion is when there is no match between sample DNA and reference material in a case. If excluded, the DNA profile obtained could be entered into a DNA database (e.g. Australian National Criminal Investigation DNA Database (NCIDD)) and compared to other profiles for a potential match. On the other hand, if an inclusion is made between a crime scene DNA sample and a reference DNA sample then a population database can be used for evaluating the strength of the evidence and the likelihood of a coincidental match. A point worth considering is that a full DNA profile generated by one STR kit (e.g. SGM+ kit with 10 genetic markers plus Amelogenin) would only represent a partial DNA profile from an STR kit that examines more loci (e.g. NGM SElect™; 16 genetic loci, plus Amelogenin). It is also the quality of the profile that is crucial to the success along with the number of loci detected.

Sizing DNA fragments

Allelic ladders are used as a reference for assigning alleles and genotypes to unknown samples, by correlating PCR product (i.e. amplicon) size to the number of repeat units. Ladders are run alongside samples and contain common alleles in the human population for each specific locus. The same primers used to target the core STR markers in the sample are used to generate the artificial allelic ladder, which is synthetically made and provided alongside the STR kit. The peak height of alleles - representative of DNA quantity - is well balanced by altering the amount of input PCR product for each to produce the allelic ladder. This ensures that the DNA alleles for unknown samples will accurately line up with the known allelic ladder fragments and the genotyping software will correctly assign alleles.

Internal size standards are artificially created DNA fragments labelled with a different coloured dye to separate it from the dye-labelled amplicon. Size standards are run alongside each sample in the same reaction tube to calibrate the peak data points from unknown sample DNA to their correct size. In addition to size standard, Hi-Di Formamide (ABI, AU) is added to the PCR-size-standard mix to ensure the DNA fragments are denatured and will run inversely proportional to mass within the capillary electrophoresis machine. The formamide forms hydrogen bonds with the DNA preventing DNA hybridisation with its complementary strand.

Software exists to interpret the STR data; however, there is a human element of interpretation needed. This will involve the need to distinguish fluorescence that is representative of the DNA sequence and that which has arisen as an artefact of the DNA profiling process.

1.4 Short Tandem Repeats (STRs)

STRs are the ‘gold standard’ of evidence in the worldwide criminal justice system. They are repeating units of 2 to 6 base pairs in length [108], easily amplified without issues of differential amplification due to similarities in length; occur on average every 10,000 nucleotides [19] in the human genome; and offer high levels of discrimination due to high levels of polymorphism and multi-loci examination. The pattern of the repeat units may vary. Simple STRs contain repeat units that are all

identical in length. Compound repeats can be defined as two or more adjacent simple repeats, whilst complex repeats are variable in length and sequence. All STRs contain sequence variation that cannot be detected with traditional capillary-based analysis methods, however this is not a requirement for most DNA labs.

Genetic loci targeted for STR analysis are positioned on non-coding regions either within or between genes. The nomenclature for STR markers depends on where the repeat region is positioned within the chromosome (e.g. within a gene or outside). For example TH01 indicates the human tyrosine hydroxylase gene located on chromosome 11 [108]. The '01' denotes the repeat region situated within intron 1 of the tyrosine hydroxylase [108] gene. If a marker falls outside of a gene, they are characterised by their chromosome position. For example, D22S1045 where D = DNA, 22 = chromosome 22, S = single copy sequence and 1045 = 1045th locus on chromosome 22. The genetic loci used for identifying samples are distinguished by the number of the repeats (e.g. Dinucleotide means two nucleotides repeated; Trinucleotides contain three and so on). Tetranucleotides (i.e. four nucleotide repeats) are the most common STR markers used in commercial multiplex PCR kits as they permit a smaller overlapping allele size range for multiplexing, reduce preferential amplification of smaller alleles, target small sized amplicons, and reduce stutter product formation [108].

The history of nuclear DNA typing has progressed from typing singleplex loci of large amplicon sizes [62] to multiplexing loci of shorter amplicon lengths (i.e. many primers perform reliably together in one reaction) [19-22, 149, 156]. The smaller size of STR alleles (approximately 100–400 bp length) compared to minisatellite VNTR alleles (approximately 400-1000 bp length) make STRs ideal for degraded samples, as they target smaller intact fragments [55, 149, 152]. STR-based analysis is less labour intensive than analysing VNTRs, amenable to automation, and advantageous for high-throughput casework.

Work carried out by Lagoa *et al.* in 2007 [57] highlights the importance of mini STRs as a suitable choice of marker for genetic fingerprint analysis. The success rate can be attributed to the target of smaller amplicon sizes (e.g. < 200 bp) [57]. It is possible

that many operational forensic laboratories do not hold validated protocols for miniSTR work so they remain redundant, unless they are included in new STR profiling kits that are designed for degraded DNA.

1.4.1 Interpretation of STR alleles

Sample comparison using STR alleles relies on match probabilities and the very small likelihood of a coincidental (i.e. random) match. There is a much higher power of discrimination between unrelated and even closely related individuals if multiple STR markers are analysed that have high heterogeneity and a high level of polymorphism. The product rule uses frequency population data sets to calculate the likelihood that two DNA profiles match. The genotypes produced can be used in the assignment of a statistical weighting for the comparison of an evidence profile to a reference profile. If matching profiles are obtained then the statistical weighting provided, known as a likelihood ratio, considers the probability of obtaining the matching profiles if they have originated from the same source compared to being purely coincidental. Differences and similarities observed in the genotype between two samples are used to exclude or confirm matches. It is important that the sequences in the regions flanking the loci are not prone to mutation [108]. However, there are instances of well characterised mutations and redundant primers compensate for these events. To support the establishment of the UK National DNA database and to ensure consistency between DNA profiling results, the forensic community selected a core set of tetranucleotide tandem repeats that would be used for identifying and comparing DNA profiles between jurisdictions and would enable overlap between different commercially available kits [65, 102].

Artefacts associated with STR-based DNA Profiling

Artefacts are an inherent feature associated with PCR-based STR analysis. They have a huge impact on DNA profile interpretation and every effort should be taken to consider these artefacts before a true allele is reported. A large scale study was carried out on low-level DNA typing and the ability to characterize artefacts, and the authors of the study concluded that the most accurate method of distinguishing erroneous alleles from true STR alleles in their laboratory was the consensus method of confirming alleles that appear at least twice from four replicates [163].

This validation approach of confirming alleles is laboratory specific. A Bayesian approach to calculate likelihood ratios that factor in the probability of artefacts has been used to strength evidence [1, 164] where two alternative propositions are considered (i.e. the prosecution hypothesis vs the defence hypothesis).

Stutter

Stutters are common artefacts reported when the DNA polymerase slips [165, 166] during PCR. The resulting product is displayed on an electropherogram immediately adjacent to the main allele peak (see Figure 1.9 (a)). The majority of stutter peaks examined are one repeat unit (i.e. four base pairs) smaller than the corresponding STR allele peak (n - 4) [166]. On occasions, stutter peaks are observed that are one repeat unit larger (n + 4) than the corresponding main STR allele (see Figure 1.9 (b)). Stutter can also be observed at n – 2 and n – 8. Stutter is thought to be associated with PCR conditions and the type of polymerase used; a faster DNA polymerase would prevent the DNA-primer hybrid complex disassociating and strand-slippage occurring [165, 166].

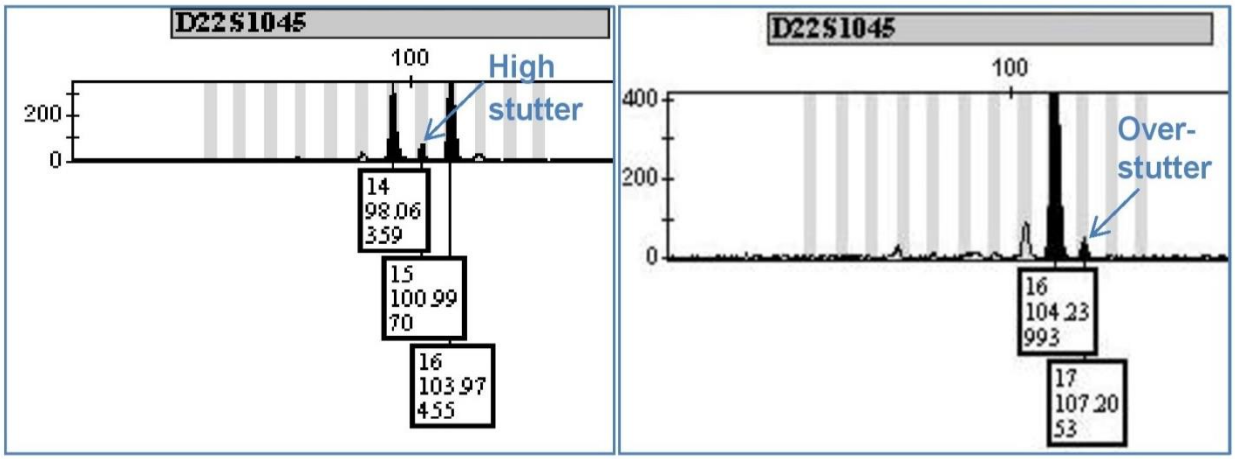


Figure 1.9 Two separate amplification reactions showing (a) high stutter at loci D22S1045 (assigned as allele 15 by the computer software). Thresholds for stutter vary between loci, however, stutter can also be characterised as less than 15% of the main peak height [167], (b) possible over-stutter observed at locus D22S1045 – this locus is more prone to stutter being a trinucleotide repeat [168].

Laboratories will determine whether a peak is a true allele or stutter by using stutter thresholds for each locus as determined by internal validation studies. The height of the potential stutter peak is divided by the height of the allele peak and a pre-determined threshold value is used to indicate stutter; STRmix software considers a range of values [161, 169]. The rate of stutter occurrence is related to the flanking sequence, the repeat unit, and the length of the allele being amplified [165, 166]. If a core repeat sequence has been interrupted (e.g. Allele 9.3 has three extra bases inserted as ATG sequence between the 6^{th} and 7^{th} repeat) [170] the quantity of stutter for this allele is also reduced to levels equivalent to the longest uninterrupted stretch (LUS) [171]. Stutter peaks are of particular concern when interpreting DNA mixtures, as minor alleles in the profile need to be distinguished from stutter peaks and other possible artefacts. One interesting study characterising stutter showed that increased stutter levels were observed with a higher A-T content [171].

Non-template addition (split peaks)

Taq DNA polymerase has both the polymerase activity but also has terminal transferase activity [147]. This latter activity adds an extra nucleotide to the 3' end of a PCR product during the extension stage [172]. This extra base is usually adenosine and is described as 'adenylation' with the resulting product one base pair longer (i.e. +A) than the target fragment [173]. Polyadenylation is not 100% effective and therefore some PCR products may not have the extra adenine (i.e. - A peak) and some do (i.e. + A peak) [174]. It is easier for profile interpretation if all PCR products are either fully adenylated or not, rather than a mix of both –A peaks and +A peaks (see Figure 1.5 (a)). To help prevent this, the extension step in PCR (e.g. ~ 60 °C) can be extended for a longer time period (e.g. 45 minutes) to allow the enzyme time to add the extra base [108].

Off-ladder (OL) alleles and Microvariants

Rare STR alleles exist that contain sequence variations in the form of insertions, deletions or single nucleotide modifications [175]. Sequence changes will not be detected on traditional CE-based platforms using standard STR profiling kits. Microvariants may differ from the common allele by only one base pair or by several base pairs [175]. Dissimilarities in the sequence content of the microvariant allele

compared to a reference ladder allele can generate true ‘off-ladder’ (OL) alleles, as they are not present in the allelic ladder.

TH01 repeat region 9.3 contains nine full repeats and an additional three base partial repeat (ATG) after repeat 6 which shorten the product length from allele 10 to allele 9.3 due to a single base deletion of adenine [170]. Mostly, complex polymorphic STRs (e.g. FGA) exhibit more microvariants than simple STRs (e.g. TPOX) due to the repeat unit being modified [108].

Allelic ladders are required to be run as reference comparison against sample DNA. Off-ladder alleles are easily detected as the base pair (bp) position of the peaks - as determined by the internal lane size standard - do not line up with the position of the allelic ladder consensus peaks and instead fall outside the designated bin for allele size calling (e.g. ± 0.5 bp). If more than one OL-allele is present in the profile, a shift in size calling across multiple alleles could result from problems with electrophoretic separation and machine maintenance. In addition, variant alleles exist that contain the same number of repeats for the microvariant as the common allele but differ in sequence content [175]. This difference would only be detected by complete sequencing of the fragments of interest and would not be detected by traditional STR analysis. Sequencing STRs is not routine, possibly due to a lack of accessible technology.

Currently, many forensic related samples do not need to be sequenced due to the high level of discrimination already achieved with standard STR kits that contain highly polymorphic loci. Occasionally new alleles will arise that have not yet been detected and are not included in the reference allelic ladder, so they fall out of range for reference alleles of specific loci. Extra peaks that are assigned to a locus need to be observed with caution, as they are not always the result of a mixture but could be due to trisomy (i.e. an extra chromosome) [175].

Allele 'drop-out'

A null allele results from a base pair change in the DNA template at the primer-binding region which results in failure to amplify the fragment of interest resulting in allele 'drop-out' [176]. Disruption of complex between primer and template DNA can result in PCR failure [176]; the allele does exist but fails to be detected. For STR loci with known problematic primer binding variants a degenerate primer can be used [177]. The degenerate primer will allow amplification of a specific region of interest in the fragment that contains the single base nucleotide change [108].

Alternatively, samples with low quantity or degraded DNA may exhibit preferential amplification of smaller sized amplicons, as larger amplicons fail to be amplified due to their size (e.g. ~ 400 bp) and typically 'drop-out' (see Figure 1.6). Stochastic effects may cause problems with mixture interpretation for low DNA template (e.g. ~ 100 pg DNA)[1]; preferential amplification of one allele over the other may result in false homozygosity (see Figure 1.6). This is the reason for using a homozygous threshold in binary interpretation systems.

Allele 'drop-in'

'Drop-in' (see Figure 1.10) is common when amplifying low amounts of DNA and can occur if proper de-contamination measures are not in place (e.g. proper sterilisation of work surfaces and equipment, and personnel protection clothing worn). Additional alleles present in a DNA profile could be the result of laboratory associated 'drop-in' rather than a true allele. FSSA have noticed a higher incidence of allele 'drop-in' and random alleles that cannot be confirmed due to the introduction of a more sensitive STR profiling kit (i.e. GlobalFiler™ amplification kit) [17]. Mixture interpretation should take place with extreme caution to prevent additional alleles being falsely assigned to a sample that are the result of poor laboratory practice or increased PCR cycles during LCN methods. Profiles of staff and personnel that have access to the laboratory need to be cross-checked against all detected alleles for exclusion purposes. Refer to Chapter II for controls used to prevent laboratory associated 'drop-in'.

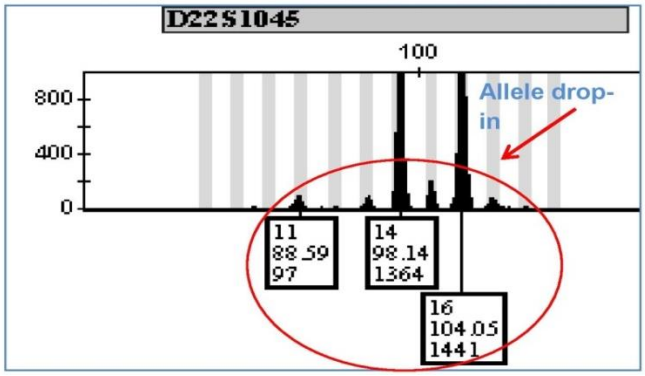


Figure 1.10 Showing allele 'drop-in' – a random event that can cause issues with profile interpretation and is not reproducible following re-PCR (i.e. the same alleles will not appear twice).

Difficulties with genotyping

Non-specific products or primer dimer

Historically there was a problem with primers annealing to template DNA at room temperature as well as the optimum temperature for *Taq* DNA polymerase [108]. Non-specific products would result from primers randomly annealing at non-specific binding sites. Primer-dimers result from primers binding together at lower temperatures causing preferential amplification of these smaller primer-dimer amplicons. These unwanted products are amplified throughout the remaining PCR cycles and may take over the reaction and consume components that would otherwise be used by the sample. To prevent these non-specific amplicons from being amplified, a 'hot-start' DNA polymerase (e.g. Ampli*Taq* Gold™ DNA polymerase) [147] can be added that requires thermal activation at a specific temperature and is rendered inactive until it reaches that temperature.

Abnormal peaks

Abnormally broad peaks (see Figure 1.11 (a)) are not characterized as true alleles as they differ in morphology. The result is a loss of DNA resolution. Dye blobs (see Figure 1.11 (b)) occur when the fluorescent dye detaches from the respective primer and migrates by itself through the capillary [108]. Dye blobs are also characterised as broad peaks that differ in morphology from true allele peaks. Spike peaks (see Figure 1.11 (c)) are displayed as sharp intense peaks that result from air bubbles, urea crystals, or voltage spike [108].

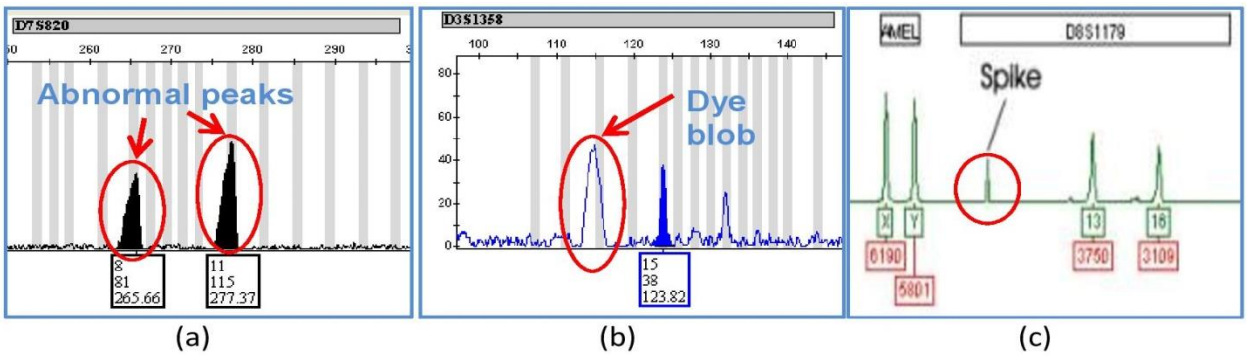


Figure 1.11 showing (a) abnormal peak morphology; b) dye blob and; (c) spike peak - unusually thinner peak compared to STR allele peaks – artefact is caused by voltage spikes [178].

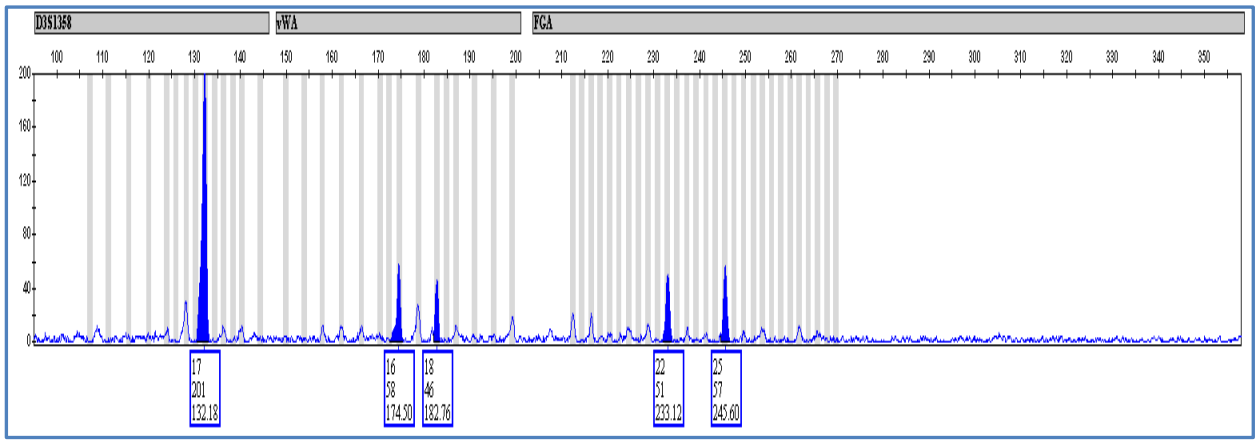


Figure 1.12 Representation of a noisy baseline in a DNA profile which can cause issues with sample interpretation. True alleles need to be separated from background and this is one reason for laboratories to move to continuous models for interpretation (e.g. + 3 SD above the baseline to determine laboratory RFU threshold values).

Advances to current methodology

Advanced methodology for trace DNA can be applied to improve the yield of PCR product and to improve the overall interpretation of profiling results. Improved methods include reduced volume amplification [179], concentration of DNA extract [127], post-PCR purification [180], increased capillary injection time and voltage [181] and increased PCR cycle numbers [153].

Low copy number (LCN) methodology was introduced by the Forensic Science Service (FSS) in the UK in the late 1990s for the analysis of low-level DNA [1]. The ability to adjust the PCR cycle number from 28 to 34 cycles known as LCN was thought to be useful for analysing low amounts of DNA by increasing the sensitivity

of the STR assay [43, 59, 182]. To improve the accuracy, the same DNA extract is used for duplicate or triplicate analysis. Increasing the PCR cycle number to 34 cycles extends the power of DNA testing as alleles can be recorded that would otherwise remain undetected. However, many operational forensic laboratories avoid implementing the LCN process due to operational constraints (e.g. lack of space for a low-level DNA testing lab, the cost associated with training staff, and the need for repetitive analysis) [181]. LCN has been challenged in court many times (see case example 2) when it has not been validated for use. Instead, improving sample collection and processing to maximise the template DNA available to the PCR would minimise stochastic effects of low-level DNA typing without the need to increase PCR cycle number.

1.5 Direct PCR

Colony or direct PCR methodology is a method of allowing more DNA to be made available to the PCR without adding steps to original forensic protocols (see Figure 1.4). This process can be easily implemented into an operational laboratory and has been around since the early 1990s [183]. Direct PCR is the process where biological material is transferred directly into the PCR tube with no prior extraction or purification steps before processing (see Figure 1.4). Direct PCR is the method used in this thesis to obtain DNA profiles from swabbed fingermarks. Lysing of cells without the extraction step will still occur during the initial stage of PCR; which involves initial heating to 95 °C for 11 minutes. The cell membranes are disrupted during the hot start cycle and the DNA is released for amplification.

In the last few years there have been more and more papers published on direct PCR in a forensic context. In FSSA and other operational laboratories the process of eliminating the extraction step is routinely used for the processing of reference samples, but not currently carried out on all sample types. Direct processing may carry over PCR inhibitors that are detrimental to the *Taq* DNA polymerase if not removed (e.g. humic acid, tannic acid and indigo dye) [184]. However, recent studies have demonstrated the great potential this method has had in analysing DNA from trace sources [30, 131, 183, 185-188], which would otherwise fail to produce a result.

Linacre *et al.* (2010) [33] report on the process of direct PCR to obtain full DNA profiles from a range of fibres taken from fabrics.

The authors summarise the key benefits of direct PCR:

- increased number of alleles obtained (less allele 'drop-out');
- reduced contamination rate (less tube changes and less chance for error);
- ability to retain valuable DNA (by eliminating spin columns and wash steps);
- increased sensitivity for low template DNA and;
- reduced cost (no commercial extraction kits are required).

Only one previous study has been published in relation to direct PCR from 'touch' DNA swabs [32]. Literature exists that highlights the benefits of the direct approach when analysing other biological materials (e.g. blood, saliva, semen and hair). Lee *et al.* (2011) [189] developed a direct PCR system for ABO typing from blood, hair and body stains to use as an initial screening test for exclusion testing and high-throughput work. The authors used a fast PCR machine and optimised polymerase (i.e. Phire hot start DNA polymerase) to target three SNPs that identify the common blood group alleles A and B in around 70 minutes. The authors report reliable profiling results from low concentrations of input DNA (i.e. 60 pg) with no allele 'drop-in' or 'drop-out' observed.

Wang *et al.* (2011) [190] used FTA paper spotted with blood or buccal swabs for direct PCR using the AmpF*l*STR® Identifiler® kit. The commercial kit includes sodium azide (0.05%), carrier protein, detergents [190] to assist with amplification. Denaturation conditions were trialled and higher peak heights were noted at 94°C, followed by 92.5 °C and then 95.5°C.

Verheij *et al.* (2012) [191] report the use of direct PCR on 149 mock casework samples that include blood, saliva, semen, hair, contact DNA and bone samples. In this study a proof-reading heat stabile enzyme (i.e. Phusion Flash DNA polymerase) was used during direct PCR. The enzyme is tolerant to PCR inhibitors [192]. In contrast to other studies, they found direct PCR profiles to demonstrate higher levels of stutter, increased base-line noise, and non-specific low-level artefact peaks [191]. However, the authors highlight the advantage of direct PCR in analysing single source samples (such as blood, saliva or semen) in a 6 hour time period that would assist in cases of high priority (e.g. national security).

Ottens *et al.* (2012) [31] report on informative DNA profiles obtained from hair follicles using the direct PCR approach. A successful profile was described as having five or more complete loci as a requirement for upload to the Australian DNA database. The authors report 100% success with anagen hairs and 33% success with telogen hairs and indicated no unusual stochastic effects (i.e. no allele 'drop-in' or heterozygote imbalance).

Sim *et al.* (2013) [193] used the direct PCR process to directly amplify DNA retrieved from buccal cells and to establish a high-throughput DNA database system using this approach. Over 4000 reference swabs were used for direct PCR and results were compared to those achieved through conventional extraction and PCR. The results from the direct PCR approach revealed perfect concordance, good quality DNA profiles and good peak height ratios [193]. The optimised buffering system included zwitterionic buffer to minimise PCR inhibition. The results from their study suggest that direct PCR minimised allele 'drop-out' in samples with a low PCR yield and the cost of analysing offender samples decreased by 35%.

A study by Swaran *et al.* (2012) [32] subjected swabs to direct PCR and compared the resulting data to extracted DNA results. The researchers used control DNA of different concentrations (i.e. 1 ng – 0.1 ng) and deposited the DNA on various surface types (e.g. plastic, glass, stainless steel and ceramic) prior to swabbing. Direct PCR out-performed conventional extraction; there was a higher yield of PCR product, and less allele 'drop-out' using direct PCR on all substrate types. Direct

PCR gave full DNA profiles (i.e. 16 loci using PowerPlex® 16 HS System, Promega) in almost all the substrates tested with starting amounts of 1 ng/µL and 0.75 ng/µL of total DNA. Incomplete or partial profiles were observed for samples that had been subjected to an extraction prior to PCR.

Further application of direct PCR was presented by Gausterer *et al.* (2012) [194] who investigated a case of suicide involving yew (i.e. *Taxus* spp.) plant poison. Phire plant PCR kit (i.e. Finnzymes, ThermoFisher) was used in the direct PCR protocol to identify the plant poison within a short time-frame of 1.5 hours.

All studies utilising direct PCR emphasised that extra tube changes involved in the extraction process may contribute to a loss of valuable sample DNA through the binding of template DNA to irreversible binding sites in polypropylene tubes and other plastic wear. Moreover, the correct choice of low bind sample tubes would be beneficial as DNA adsorption to tubes may influence the level of DNA recovered [195, 196]. The success and quality of DNA profiles obtained using the direct PCR method will be dependent on the nature of the material examined and the presence of PCR inhibitors, which are generally eliminated during the extraction process.

1.6 Expert systems

Software systems that convert electropherogram information into genotypes need to be compatible with current databases. GeneMapper® ID-*X* [197] and OSIRIS [198] are software programs commonly used by forensic laboratories to assign alleles through size comparison to allelic ladders. All software programs require overview by the operator. Quality values (PQVs) are generated that relate to confidence in allele calls [108].

The purpose of software programs such as STRess® (STR Expert System Suite) developed by the Forensic Science Service (FSS) [199] and TrueAllele® [200] are to improve STR allele calls through standardisation and to reduce manual assessment of individual profiles. Quality measures are in place to increase accuracy by reducing manual error rate. TrueAllele® assigns genotypes to peaks along with quality values that range from 0 – 1 based on the peak morphology. Additionally, TrueAllele® has

had success in defining the number of contributors in complex mixtures [201, 202]. The aim of STRess® is to reduce the sample interpretation bottle neck which generally requires two analysts to examine data. Features of the software include a high through-put search engine for hits, a cross-contamination module to prevent errors between samples and monitor staff contamination, and the pendulum module for mixture deconvolution [108].

1.7 Science and the Law

Credible evidence accepted by a court is generally in the form of witness statements, expert opinion, physical evidence and circumstantial evidence. A judge will determine whether admitted evidence is considered credible and should be accepted by a court of law, not the forensic scientist whose role is to remain impartial. The role of the expert witness is to explain how the physical evidence may relate to the case findings based on scientific fact and statistical probabilities. In court, most credible evidence is delivered through the testimony of witnesses, and physical evidence is rarely credible without an explanation of how it links to the case.

In the past, expert witness evidence was determined admissible in the United States in the 1920s if it was “generally accepted” by the wider scientific community. This arose due to the results of a polygraph test result dated to 1923. This was the *Frye* standard (*Frye v United States* 293 F 1013 (DC Cir, 1923)) [203] where a scientific technique discussed in court had to be recognised, accepted, and widely used by scientists before the results would be accepted as reliable evidence. Later, the Daubert standard *(Daubert v Merrill Dow Pharmaceuticals Inc* 509 US 579) (1993) [203] (US only) superseded the Frye standard in many jurisdictions in the United States, but not all. With the Daubert standard, criteria was set out to ensure adequate methodology and reasoning of science before evidence would be admissible in a court of law.

Requirements of the Daubert standard include:

- an understanding of basic theory and how the testing is performed;
- peer review of the science in publications;

- a knowledge of error rates associated with the method;
- use of adequate controls and standards to validate the method; and
- acceptance of the method in the field in which it is studied.

In Australia and the UK, the Bonython test (*Bonython v R* (1984) 38 SASR 45) [204] is often cited when considering the "field of expertise" test. The expert witness needs to satisfy the Judge of his knowledge, experience, or both, to qualify as an expert witness in the eyes of the court. The President's Council of Advisors on Science and Technology (PCAST) was released in 2016 with recommendations on how to strengthen forensic science in the courtroom. PCAST reported a need for 1) clarity about the scientific standards required to deem a method reliable, robust and reproducible and 2) the requirements to determine whether a method is scientifically valid before its use [205].

1.8 Laboratory validation

DNA typing needs to be performed to high quality standards following good laboratory practice (e.g. use of SOPs, qualified and trained personnel, and quality assurance and quality control measures) to ensure accuracy in results that can withstand legal scrutiny in court. Requirements for accreditation include: intense validation of all new techniques and methods to guarantee results are reliable, robust and reproducible; proficiency testing to ensure results obtained in one laboratory are generated in another accredited facility; and proper document control of all casework and evidence handling that follow a strict chain of custody. Laboratory audits of techniques, equipment and staff are all part of maintaining accreditation. All personnel need to be qualified and adequately trained to ensure day-to-day operation that meets these expectations. The PCAST report discussed above describes the need for continual improvement in certain areas of forensic science that have previously been scrutinised for scientific validity and reliability, specifically feature-comparison methods such as DNA analysis of single-sourced and simple mixed samples, DNA analysis of complex mixtures, bitemarks, latent fingerprint, firearms and footwear analysis [205].

1.9 Organisations that exist for quality

The Scientific Working Group on DNA Analysis Methods (SWGDAM) [154] is an organisation set up in the United States that consists of a team of scientists that work together to ensure adequate quality control measures are in place for DNA typing. For example, a subcommittee of the group includes STR interpretation, training, and validation. The working groups set quality assurance (QA) guidelines to aid laboratory managers in setting up their own internal QA program.

The DNA Advisory Board (DAB) organisation [206] is part of the FBI and consists of a team of scientists whose main aim is to issue standards to the forensic DNA community. For example, the DAB Standard 13.1 requires that DNA analysts perform regular external proficiency testing [108] to evaluate staff ability. The National Institute of Standards and Technology (NIST) U.S. Department of Commerce [207] provide reference materials for calibrating laboratory techniques against known standards. DAB standards [206] require annual DNA testing against a known NIST standard in order to verify results.

The European Network of Forensic Science Institutes (ENFSI) [208] was established to set up standards for data exchange between member states and to conduct audits as an accrediting body. Another European organisation that exists with a similar purpose to SWGDAM is the European DNA Profiling Group (EDNAP) [209] that conducts proficiency testing between laboratories to ensure concordant STR results are achieved. The main aim of EDNAP is to examine the reproducibility of results across various laboratories.

National Association of Testing Authorities (NATA) was established in 1947 [210] and was the world's first accreditation authority in the assurance of technical standards. Member facilities (e.g. DNA testing laboratories) comply with testing standards to uphold their status of accreditation and to be deemed competent in a range of testing that provides consistently accurate and reliable testing, through calibration, validation studies, proficiency testing, measurement, quality controls, and record keeping. Accredited forensic DNA facilities will define their own interpretation guidelines based upon their own internal lab testing and validation guidelines from

SWGDAM [154]. For example, with STR typing, each laboratory will set a minimum peak height threshold, homozygote threshold, and stutter ratios for each locus.

1.10 Databases

DNA databases are important tools to assist police investigations as the DNA profiles of repeat offenders are likely to be on the system used to link suspects with evidence, and link two or more criminal events.

Databases can be used to solve crimes by linking unknown DNA profiles obtained from evidence to known profiles in the database through 'hits' defined as a match between DNA profiles. Features important for the success of a DNA database include: a core set of STR markers, standard software used between jurisdictions, and quality standards to ensure consistency and accuracy. The forensic community strives for harmonization worldwide to ensure a core set of STR markers are used in national DNA databases that that permit an overlap of standard loci between countries, and enable statistical weighing to be applied that generate consistency between results. The FBI launched the Combined DNA Index System (CODIS) in the US in 1990s and maintains a core set of 13 STR markers that provide a random match probability of 1 in 100 trillion [108]. The National DNA database (NDNAD) launched in the UK initially used six STR markers and the sex determining marker Amelogenin to provide a random match probability of ~ 1 in 50 million [1, 65]. Later, the analysis of ten STR markers plus Amelogenin in the AmpF*l*STR® SGM Plus™ kit increased the random match probability to 1 in 3 billion [153], and has since increased further to 17 genetic loci (2016).

Improving methods to retrieve more DNA from items of forensic significance will enable more potential hits on DNA databases and more effective forensic investigations; the ability to generate high quality DNA profiles will enable more samples to be uploaded and compared to the DNA profiles of previous offenders.

SCOPE OF THESIS

This thesis is written as a series of publications that are currently published or under review in relevant forensic journals.

AIMS OF THESIS

This thesis aims to explore the use of direct PCR for amplifying 'touch' DNA deposited by fingermarks as a source of forensic human identification. In collaboration with the FSSA laboratory, aspects associated with DNA recovery from swabs, DNA extraction, PCR, CE detection, and analysis of resulting DNA profiles are reviewed. In particular, a large emphasis will be placed on issues surrounding DNA contamination and ways to minimise unwanted extraneous DNA in the reaction. The main goal of this research is to include standard operating procedures (SOPs) in the materials and methods section of this thesis, with guidance from FSSA on how to implement the research methodology into operational casework. To achieve this goal, a case study demonstrating the application of the technique is included in Chapter IV along with peer reviewed manuscripts compiled into the following chapters:

1) The first data chapter (Chapter III) compares the efficacy of cotton, nylon and foam swabs for recovering low-template DNA. Additional work highlights the most effective swabbing method prior to direct PCR. Results obtained from a standard extraction kit (DNA IQ™ System) are compared to those obtained by direct PCR when processing 'touch' DNA swabs. This chapter includes: 3 peer reviewed publications; a poster presentation; an oral presentation; and a patent application.
2) The second data chapter (Chapter IV) focuses on the application of direct PCR for processing latent 'touch' DNA samples. Results are presented in a peer reviewed publication, a journal article in *Biotechniques*, and a real case study. A further manuscript (which is in preparation) is included that explores the use of PCR facilitators (such as BSA and DMSO) on the direct PCR approach.
3) The third data chapter (Chapter V) examines the use of direct PCR for profiling visually enhanced (i.e. powdered) fingerprints. Further work

aims at exploring the effects of environmental exposure on the survival of ‘touch’ DNA, using the direct PCR approach. Work highlights the application of direct PCR on challenging substrates (e.g. metal cartridge cases). Results are published in a short peer-reviewed publication, a poster presentation, and a manuscript in preparation.

1.10.1 Statement of authorship

Title of publication: Forensic DNA profiling: state of the art

Manuscript published in the Research and Reports in Forensic Medical Science (DOVE press).

Jennifer Templeton (Candidate)

Co-wrote the manuscript.

I hereby certify that the statement of contribution is accurate.

Signed: Date: November 2016

Adrian Linacre (Supervisor)

Co-wrote the manuscript.

I hereby certify that the statement of contribution is accurate.

Signed: Date: November 2016

1.10.2 Manuscript: Forensic DNA profiling: state of the art.

Forensic DNA profiling: state of the art

ARTICLE · AUGUST 2014

DOI: 10.2147/RRFMS.S60955

2 AUTHORS:

Jennifer Templeton
Flinders University
14 PUBLICATIONS **90** CITATIONS
SEE PROFILE

Adrian Linacre
Flinders University
148 PUBLICATIONS **1,588** CITATIONS
SEE PROFILE

Research and Reports in Forensic Medical Science

Dovepress
open access to scientific and medical research

Open Access Full Text Article

REVIEW

Forensic DNA profiling: state of the art

This article was published in the following Dove Press journal:
Research and Reports in Forensic Medical Science
28 August 2014
Number of times this article has been viewed

Adrian Linacre
Jennifer EL Templeton

School of Biological Sciences, Flinders University, Adelaide, SA, Australia

Abstract: DNA profiling is now a routine test, yet it has only been 26 years since its first use. During these intervening years, the science of DNA profiling has undergone many changes, and it will continue to do so in the future. DNA profiling in the laboratory starts with DNA extractions and is followed by quantification of the DNA, amplification of up to 23 human specific short tandem repeat loci, and separation on capillary columns. The final stage is interpretation of the data leading to reporting the DNA evidence to a court. Each of these steps has seen improvements leading to higher sensitivity from minute trace samples. Genetic markers on the X and Y chromosomes are now used for both criminal and civil investigations, along with mitochondrial DNA. Mitochondrial DNA has been used in forensic science since the 1990s because of its high copy number and particular maternal inheritance. These sex chromosome markers and mitochondrial DNA were used effectively to determine the identity or ancestry of individuals, and now other genetic markers can be used to determine certain phenotypic traits to a high degree of accuracy. Genetic testing adapted from medical and pharmaceutical sciences, such as next-generation DNA sequencing, will soon be applied to mainstream forensic science, opening new avenues in criminal investigations. This review aims to cover these key events and advancements in the field from both a historical view and current practice.

Keywords: forensic science, DNA profiling, short tandem repeats, mitochondrial DNA, phenotypic testing, next-generation sequencing

Opening comments

DNA profiling has revolutionized the process of forensic human identification. It is of note that this is a relatively recent addition to the forensic science tool kit and that it is constantly undergoing developments. The aim of this review is to take the reader through the processes of DNA typing, starting with DNA isolation from biological material through novel approaches and phenotypic testing. The review is written for those in forensic medicine; therefore, certain knowledge is assumed, but equally, aspects that are specific to forensic science will be emphasized and placed into the right context. For those who wish more information than can be provided within the scope of this review, there are books ranging from introductory texts[1] to those that are more comprehensive.[2,3]

Correspondence: Adrian Linacre
School of Biological Sciences, Flinders University, Bedford Park, Adelaide 5042, SA, Australia
Tel +61 8820 12465
Email adrian.linacre@flinders.edu.au

Introduction

The first use of DNA fingerprinting was to resolve an immigration dispute in 1985.[4] A process of DNA fingerprinting had been published the year earlier in *Nature*, coauthored by Alec Jeffreys. This led to a local police force engaging Jeffreys

submit your manuscript | www.dovepress.com
Dovepress
http://dx.doi.org/10.2147/RRFMS.S60955

(now Professor Sir Alec Jeffreys) to assist with an unsolved sexual assault and murder of two young girls in Leicestershire, England. DNA fingerprinting led originally to linking the two murders to the same perpetrator, then to the exoneration of a young man, and finally to the conviction of Colin Pitchfork for the crimes in 1988.[5]

The early test performed by Alec Jeffreys to identify an individual through DNA typing is very different from modern-day DNA testing; the only commonality is the word "DNA." What had been noticed was that sections of DNA are hypervariable between people but conserved throughout the cells of an individual. DNA fingerprinting originally examined variation at many loci across the entire genome leading to multiple bands on a gel (an example of which can be seen in Figure 1), with the appearance of a barcode, and as everyone's barcode appeared to be different, the term "DNA fingerprinting" was coined. Since 1994, all DNA profiling (as it is now termed) has centered on microsatellite DNA (termed short tandem repeats, or STRs, in this review), and in particular, on repeat units of four bases (see chapter 5

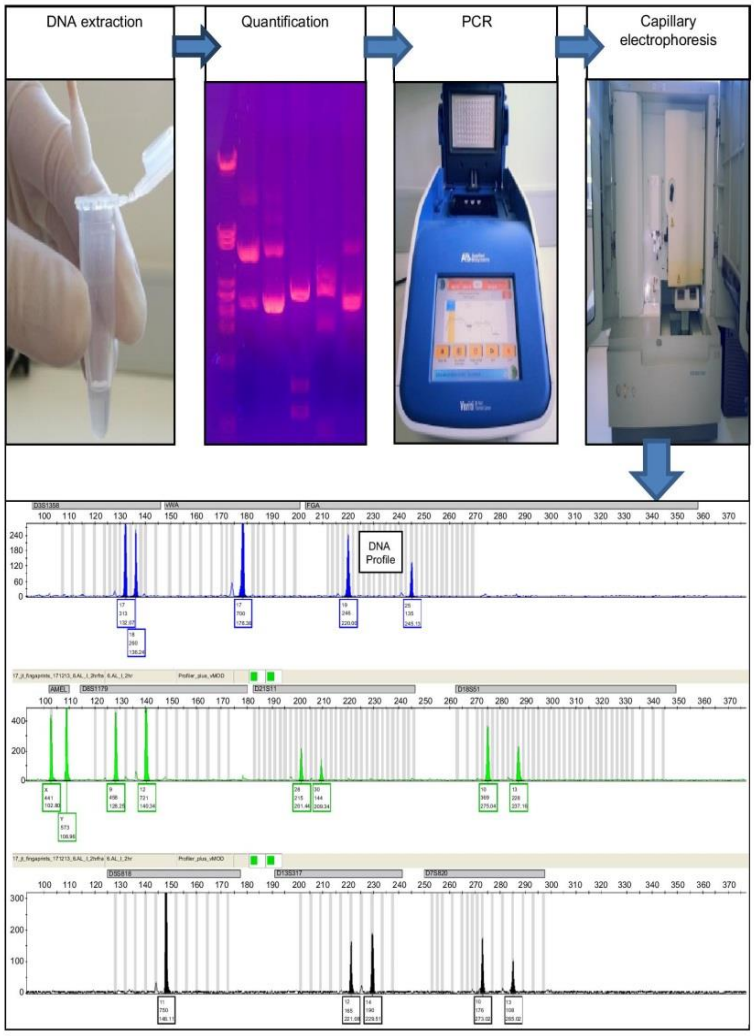


Figure 1 Steps involved in DNA profiling. Profiling starts with isolation of DNA by a process called DNA extraction, followed by quantification of the DNA in the sample, then amplification of short tandem repeat loci, separation of the short tandem repeat products, and finally, interpretation of the genotype data.
Abbreviation: PCR, polymerase chain reaction.

from Butler[2] and chapter 6 from Goodwin et al[1]). The STR loci used in standard DNA profiling are all inherited in a Mendelian manner; are not genetically linked, and therefore are inherited independently; and crucially, are thought to have no known association with a disease state.

The advent and application of polymerase chain reaction (PCR)[6] to forensic genetics was timely and advantageous. PCR is a method of amplification used to produce millions of copies of a specific region of DNA. Tiny blood spots or single hairs could generate DNA profiles, thus increasing the sensitivity of the test to subnanogram levels of DNA. Allied to this were increasing powers of discrimination: the probability that another person shares the same DNA profile by chance. DNA profiling has been heralded as a game changer because of these high powers of discrimination (resulting from an increased number of loci being examined between individuals) and has been seen to override other forensic evidence for which the powers of association are many orders of magnitude less. Validation studies and interlaboratory proficiency testing to demonstrate as comprehensively as possible the robustness and reproducibility of the tests were undertaken because of the significance of DNA evidence in any criminal case. DNA frequency databases were generated for all populations to ensure an accurate (and frequently conservative) estimate of the frequency of the DNA profile. The chances that the DNA profile came from a close genetic relative were incorporated into statistical formulae, as were mechanisms to adjust the frequencies stated for small subpopulations. The current situation is that because of the extensive validation studies and challenges in courts in many jurisdictions, DNA is now considered a gold standard[7] in forensic science, with the science underpinning the forensic application and processes of evaluation of the evidence being the standard to which other areas of forensic science should aspire.

The forensic DNA community has striven to achieve standardization to ensure the same core set of DNA loci are examined between jurisdictions. This will ensure there is commonality in the data obtained, allowing national and international sharing. This harmonization manifests itself in a number of ways, including, although far from exclusively, the basic steps in forensic DNA processing. These are illustrated in Figure 1. The process starts with DNA extraction and is followed by quantification of the DNA isolated and then amplification of the STR loci, separation of the STR alleles by capillary electrophoresis (CE), and finally, interpretation of the data.

Typically, the processes shown in Figure 1 take a minimum of 10 hours, using current standard technology. The aim of much research is to make this process quicker without compromising the quality of the data obtained. Each of these steps is described here.

DNA extraction

The ultimate aim of any DNA extraction process is for the final solution to contain DNA in sufficient purity and quantity to allow subsequent processing by PCR. The extraction process chosen may be modified depending on the biological tissue being analyzed. Blood contains DNA only within the white blood cells. Other cellular components need to be separated and purified from the DNA, such as proteins, lipids, and carbohydrates. Semen has the added complication that the DNA is encased within the spermatozoa head; this head (the acrosome) is made from a strong protein that is resistant to many naturally occurring enzymes. In contrast, the DNA from buccal and skin cells is more easily obtained, as these epithelial cells are easily lysed, exposing the DNA to the extraction buffer.

There are additional considerations for forensic science processing. For example, a forensic laboratory may process tens of thousands of samples every year. A manual procedure in which individual laboratory staff members pipette small volumes from one tube to another will be highly costly in staff time. Further, every time fluid is passed from one tube to another, it leaves open the potential for contamination or for adding fluid to a wrong or incorrectly labeled tube. The possibility of using automated extraction processes is therefore beneficial in terms of costs and ensuring that the

Table 1 Four currently used DNA extraction methods

	Phenol	Chelex® (Bio-Rad Laboratories, Hercules, CA, USA)	Silica	Magnetic beads
Safe or toxic	Toxic	Safe	Safe	Safe
Expense	~AU$3.89 per sample	~AU$1.32 per sample	~AU$10.66 per sample	~AU$6.36 per sample
Sample types	Bone, teeth/degraded biological material	Buccal swabs, epithelial cells	A wide range of biological sample types	A wide range of biological sample types
Open to automation	No	No	Yes	Yes

Abbreviation: AU$, Australian dollars.

Table 2 Inheritance and genetic sharing of autosomal STR markers, Y-chromosome STR markers, X-chromosome markers, and mitochondrial DNA

Inheritance	Autosomal STRs	Y-STR	X-STRs	Mitochondrial DNA
Father–son	50%	100%	–	–
Father–daughter	50%	–	100%	–
Mother–son	50%	–	100%	100%
Mother–daughter	50%	–	50%	100%
Brothers	25% both alleles, 50% one allele, 25% neither allele	100%	100%	100%
Sisters	25% both alleles, 50% one allele, 25% neither allele	–	50%	100%
Paternal grandmother–granddaughter	25%	–	0%	100%
Maternal grandmother–grandson	25%	–	50%	100%
Paternal grandfather–granddaughter	25%	–	–	–
Paternal grandfather–grandson	25%	100%	0%	–

Notes: Total percentage values do not account for possible mutations in the genetic material. –, no inheritance of that type of DNA (ie, females do not possess a Y chromosome; 0% is used when there is an opportunity for this marker to be passed on but none should pass on in this lineage).
Abbreviation: STR, short tandem repeat.

wrong DNA profile is not attributed to an individual because of a sample mix-up.

The principal steps in any extraction process are to lyse the cellular membrane, using a solution typically containing sodium dodecyl sulfate; break open the nuclear envelope, using a solution containing a proteinase (typically proteinase K); separate and purify the DNA molecule (ethanol), making use of its negative charge; and elute the DNA into a fresh solution (low-salt buffer or water) relatively free from inhibitors and of sufficient purity to allow enzymatic amplification of the DNA.

The original method of DNA isolation was an organic extraction process[8] in which, initially, a buffer containing sodium dodecyl sulfate, a chelating agent (ethylene diamine tetraacetic acid), and an enzyme to digest protein (proteinase K) is used to expose the native DNA molecule. The proteins are removed by addition of an equal volume of the solvent phenol (or a 1:1 mix of phenol and chloroform). Organic extractions are only used in current forensic practice for niche applications, as they are laborious, involve multiple tube changes in this process, and cannot be automated, and because phenol is a highly toxic chemical both to the user and the environment.

Chelex® (Bio-Rad Laboratories, Hercules, CA, USA) was introduced to forensic practice in the 1990s[9] and quickly became part of a popular method. The process uses a suspension of Chelex® that binds any positive ion (such as Fe^{2+}, Mg^{2+}, and Ca^{2+}), leaving the DNA behind. There are other biological materials that are negatively charged that will coextract with the DNA; hence, this method leaves the DNA dissolved finally in a fluid that is far from free of potential inhibitors. The use of Chelex® is not easy to automate for high-throughput laboratories and has been superseded by silica-based extraction methods, although Chelex® is still used for the isolation of DNA from buccal swabs, in which the DNA is in a less complex state compared with in blood or semen.

Although Chelex® uses a bead that binds positive ions, silica-based extraction processes make use of the negative charge carried by the DNA molecule. Silica-based extraction methods have become widely available as part of a commercial kit, making them more costly than either phenol or Chelex®, but they have the advantage of being open to automation. The silica-based material[10] is embedded in a layer at the base of a spin column, through which fluid containing DNA can pass. The DNA binds to the silica membrane, but all the molecules that do not possess a strong negative charge pass through the membrane and are discarded. The DNA can ultimately be released from its bound state by altering the pH of the silica membrane, allowing DNA to be eluted as a relatively pure solution.

Magnetic bead technology[11] acts in a similar way to silica-based methods, in that DNA binds to a biotin complex that is attracted to a streptavidin bound to a magnet. Other constituents of the cell that do not bind to this complex are flushed away or removed. Once this is complete, the DNA can be released from its binding and eluted into a relatively small volume of fluid. See Table 1 for a comparison of these methods.

DNA quantification

The amount of DNA isolated by the extraction step needs to be quantified to ensure the optimal amount of DNA template is added to a PCR. Typically, the amount of template

required for PCR is 1 ng human DNA. Depending on the sample type, the total DNA isolated may include bacterial, fungal, or other nonhuman DNA.

The earliest method of DNA quantification involved the measurement of the absorbance of light at 260 nm, using an optical density spectrophotometer.[8] Although this can be relatively accurate, it requires a large quantity of DNA, as the method is not sensitive, and hence most of the sample may be used. Further, this method detects all DNA and is not human-specific.

For a while in the 1990s, slot blots were used, in which DNA of known concentration was immobilized on a nylon membrane at specific locations, to act as standards. The DNA samples to be tested were immobilized near the standards. A short sequence of DNA found in humans only (actually an alpha satellite repeat sequence on chromosome 17) was used as a probe to bind to the immobilized DNA, and as the probe could be detected by chemiluminescence, the amount of human DNA present could be determined. This method was more sensitive than using an optical density spectrophotometer to measure the absorbance of light, but it took many hours to complete.

The current standard technology used in forensic science laboratories is real-time PCR.[1,2] At each cycle in the PCR, there should be a doubling of each PCR product. As these PCR products are double-stranded molecules, any dye that binds to double-stranded DNA can be used to monitor the reaction as it occurs. If the PCR primers bind to human-specific DNA sequences, then the presence of only human DNA will be detected; this is a major benefit compared with other methods of DNA quantification. Equally, the amount of template DNA present at the start of the reaction can be determined by comparison with known standards. Using this real-time approach, investigators can monitor a reaction, and if suboptimal nuclear DNA is detected in a sample, they may choose to focus on targeting other markers in the genome that are more ideal for degraded DNA. Real-time PCR has a further advantage, in that the presence of inhibitors can be noted if the reaction fails. The only disadvantage is the cost of the equipment required, as it is significantly higher than the other methods.

Amplification of STR loci by PCR

The development in the enzyme that amplifies the STRs that is relevant to DNA typing is a "hot start" activity.[12,13] Although the optimum temperature for Taq DNA polymerase is between 60°C and 72°C, it will work at room temperature, although relatively slowly. This can lead to poor amplification, and potentially amplification of the wrong fragments if there is a time delay between preparing the PCR and adding it to the thermal cycler. Hot start enzymes require heating to 95°C, during which time the pH of the buffer alters such that the enzyme becomes active. The end result is that there is more specific binding in the initial steps of the PCR, leading to more of the targeted PCR product.

STR loci

There has been an evolution in the number of STR loci amplified in one reaction. All the STRs used are noncoding, but most are close to a gene sequence or are within introns. The first STR loci were identified within introns and named after the gene with which the intron was associated. For instance, the STR locus called vWA is a repeat found in intron 40 of the von Willebrand Factor gene (named after the physician who noted a blood clotting disorder based on a mutation in this gene), which is on the short arm (p) of chromosome 12. TPOX is an STR within intron 10 of the thyroid peroxidase gene on the short arm of chromosome 2. A more simple nomenclature was adopted as more STR loci were incorporated into DNA profiling; for example, STR locus D3S1358, where D stands for DNA; 3 designates the chromosome; S indicates there is a single copy, rather than multiple copies; and the number indicates that this was the 1,358th section of chromosome 3 to be described.

Originally, there were four loci in 1994,[14,15] and then six loci 1 year later.[16–19] These reactions were assembled by the UK Forensic Science Service. Commercial companies saw there was a market and started to make their own STR multiplex kits; the main companies were Applied Biosystems (now part of Thermo Fisher Scientific, Waltham, MA, USA), and the Promega Corporation (Fitchburg, WI, USA). By 1997, there were kits that amplified nine STR loci plus a sex test, and then in 2000, a 16-STR locus test was introduced.[20] At this time, the technology has improved to the stage at which 23 STR loci plus the sex test are amplified in a single reaction.

Another advance in STR technology is that the reaction buffer has seen improvements, such as an increased tolerance to inhibitors. Minor inhibitors of the reaction, which may have previously resulted in no amplification product, have less effect, leading to more complete DNA profiles. This has also increased the sensitivity of the PCR, with the optimal starting template being 500 pg DNA rather than 1 ng; 500 pg is approximately 84 human cells.

The primers used to amplify the STR loci have, in many instances, been designed to sequence closer to the start of the

repeat unit. The aim of this was partly to ensure that these larger STRs would still amplify from highly degraded DNA, as previous empirical data had shown that larger STRs are more likely to be lost from the amplification as the amount of degradation of the DNA increases.

With Promega and Applied Biosystems offering different kits, and with occasional amplification of different STR loci, there had been little standardization of the STR data between countries. The adoption of the European Standard Set in 2009 led to 12 core STR loci being adopted by all European countries. This has opened up the exchange of DNA data. To accommodate the situation in the United States and lead to further standardization, the number of shared loci rose to 18.[2] The new STR kits include all these 18 loci plus additional markers, allowing freer exchange of DNA data, such as allele frequencies, internationally.

Separation and detection of STR alleles

Polyacrylamide gel electrophoresis was the original method for separating STR alleles. In this method, DNA can be added to a gel matrix, and in the presence of an electric field, the negatively charged DNA will be separated according to size (to distinguish the STR alleles). Shorter molecules migrate faster through the matrix, and the separation of DNA allows for sizing of fragments against a known standard (ie, DNA ladder). CE was a major breakthrough, as the separating polymer is much safer to use. The only disadvantage was that the original CE machines separated only one or up to four samples at one time, unlike a gel with many lanes. Applied Biosystems produced a single-capillary machine (310 Genetic Analyzer) but quickly developed a 16-capillary machine (3130*xl*Genetic Analyzer), in which samples are separated on all 16 capillaries simultaneously. The runs take approximately 30 minutes but can be run overnight in batches, allowing 128 samples to be run in less than 5 hours.

Detection of the STR fragments is based on laser-induced fluorescence, as a dye can be attached to the 5′ end of one of the primers used in the PCR. Originally, there were four dyes available to allow three STR alleles to overlap in size, with the fourth dye set aside for an in-lane size standard. Previous technology meant that samples that overlapped in mass had to be separated in different lanes; hence, three dyes was a major advance, allowing up to ten STRs to all fit in a size range between 100 and 400 base pairs (this size window allowed nonincorporated primers to be separated from true products and also degraded DNA would still generate alleles at most of the ten loci). More recent technological advances have allowed a fifth dye to be used (and a sixth dye will be released at the end of 2014), such that up to 23 loci can be compressed into this small size range.

Criminal justice databases

The ability to generate STR profiles quickly and with a high power of discrimination led the UK government to consider the opportunity to retain the DNA profiles of any person found guilty of an offence punishable by imprisonment. This was in 1995, just as the six-loci test was introduced. The power of discrimination of this six-loci test was close to the population of the United Kingdom at the time, and hence the chance of a coincidental match should have been extremely unlikely. At the time of writing, because of changes in legislation affecting the retention and addition of DNA data, the number of people on the UK database (NDNAD) is more than 6 million. At this time, in the United Kingdom, approximately one in 20 members of the population are on the NDNAD. As approximately 80% of those on the NDNAD are men, and given that there are 30 million men in the United Kingdom, this means that approximately 12% of the entire male UK population is on the NDNAD. The most recent figures indicate there is a 58% chance that a DNA profile from a crime scene will match a DNA profile on the NDNAD (for more information, see chapter 10 of Goodwin et al[1] and chapter 8 of Butler[2]).

In 1998, in the United States, a similar database was established that now contains more than nine million STR profiles, making it the largest such database numerically (the one in the United Kingdom contains the most DNA profiles as a percentage of the population). Indeed, almost all countries with a forensic DNA capability have established such a database.

Criminal DNA databases have been a driving force for the standardization of STR loci across countries.

Lineage markers

All the STR loci discussed so far are on one of the 44 somatic chromosomes; hence, there is a copy inherited from the biological mother and one from the biological father. There are STR loci on the X and Y chromosomes, as well as polymorphic DNA sequences within the mitochondrial DNA. All three have found niche applications in forensic science for criminal casework, identification of victims in mass disasters, or paternity/immigration disputes. These same markers have been used for evolutionary purposes to reconstruct the movement of man since leaving Africa

some 90,000 years ago.[21,22] Table 2 illustrates the mode of inheritance of these loci.

Y-STR typing

The Y chromosome should be inherited as a complete chromosome from father to son. All male siblings should therefore also share the same Y chromosome. Unlike the 22 pairs of autosomal chromosomes, the Y chromosome does not recombine with any partner (only the tips make any contact with the X chromosome), and hence all the loci remain in the same order and linked from one generation to the next, to the next, to the next, and so on. The genetically linked DNA loci on the Y chromosome are termed a haplotype.

The characterization of highly variable STR markers on the Y chromosome led to commercially available kits that currently amplify 23 such male-specific loci.[23,24] In forensic science, Y-STRs are particularly used in sexual assaults. Although there are methods to separate spermatozoa from epithelial cells, male epithelial cells will be detected, along with any female ones, leading to mixed DNA profiles. Frequently, in cases of alleged sexual assaults, samples are taken from fingernails. If taken from a female suspect, these samples will contain DNA from the epithelial cells of the female donor as well as epithelial cells from the male assailant. Any resulting DNA profile will be a mixture of the two and will be possibly difficult to interpret, but using Y-STR, markers of a single male DNA profile will be produced. The mixture of two or more men in the same sample from an alleged sexual assault, such as consensual intercourse followed a few hours later by an alleged rape, can also be problematic, using autosomal markers. This can be less of a problem using Y-STR loci if the Y-STR type from the consensual partner is known and can be removed from the mixed DNA profile.

One novel application of Y-STR loci is that some loci have been found to possess high levels of mutations, such that even close genetic relatives with a recent common paternal relative may be differentiated.[25] Brothers should have the same Y-STR type, as is the case for most STRs on the Y chromosome, but some of these rapidly mutating loci may be different, allowing even brothers to be differentiated by Y-STR typing.[23,26]

X-STR typing

STR loci on the X chromosome are very new and have a limited but valuable niche. They are most useful in paternity testing or identification of victims of a mass disaster where the parents of offspring are unavailable. A woman will pass on her intact X chromosome to her son, who will pass the same X chromosome on to his daughter. Hence a grandchild will have the same X chromosome as a paternal grandmother. Such a linkage has been used to settle cases of incest,[27] as well as help identify previously unknown individuals in cases that could be resolved only because of the particular inheritance of the X chromosome. Fourteen X-chromosome STRs have been characterized for forensic use[28–30] by the Armed Forces DNA Laboratory in the United States with the potential for identification of the deceased to complement both Y-STR typing and mitochondrial markers.

Mitochondrial DNA

In reality, many biological samples submitted for forensic analysis are not of pristine quality, and some, by their nature, have very little nuclear DNA; for example, degraded bone fragments, teeth, and hair shafts. Highly compromised samples may contain damaged, degraded, and/or low-quantity DNA and fail to be characterized by normal autosomal STR DNA markers. In such cases, when nuclear DNA is considered low template (<100 pg), using the quantification methods described earlier, the analysis of mitochondrial DNA (mtDNA) can prove to be a valuable tool. mtDNA exists in a much higher copy number in the cell compared with nuclear DNA and is more likely to produce a result.[31,32]

Mitochondrial DNA is inherited down the maternal line and is particularly useful in cases of missing persons when maternal relatives separated by several generations can act as reference samples. It has also been used to identify the skeletal remains of war casualties[33,34] and in the identification of the remains of Tzar Nicholas II of Russia.[35,36] The Federal Bureau of Investigation crime laboratory in 1990 began research into the credibility of mtDNA to be used as evidence in casework, and in 1995, it published a validation study.[37] Later, in September 1996, the first case of mtDNA evidence was used in a United States courtroom to prosecute a 27-year-old Tennessee murderer, Paul Ware. The defence claimed that the accused had been framed for the rape and murder of a 4-year-old girl, but mtDNA evidence had already linked the suspect to two hairs recovered from the victim. This mtDNA sequence had never been seen before and proved to be vital evidence in the case.

Forensic laboratories specializing in mtDNA typing generally focus on sequencing a small portion of mtDNA that is highly variable (changes occur by mutations), known as hypervariable regions I and II (HVSI/II) or the control region. The mtDNA control region random match probability (the chance that another unrelated individual will share the same sequence) is roughly one in 120 individuals, although

it has still proven to be useful evidence in certain forensic cases.[38] HVSI/II sequencing is, however, limited by the resolving power of this short sequence. Individuals with a similar maternal lineage will share the same haplotype and fail to be distinguished from each other. To maximize the information content from mtDNA, and to gain further resolution, the whole mitochondrial genome can now be sequenced with relative ease because of the advent of next-generation sequencing (NGS). Whole-mitochondrial-genome sequencing provides a high level of maternal lineage discrimination, and recent studies have shown that more than 70% of mtDNA variation lies outside HVSI/II for certain DNA sequence types (called haplogroups).[39] Other studies highlight the advantage of sequencing the entire mitochondrial genome for human identification by detecting single nucleotide polymorphisms (SNPs) in the coding region.[40–42]

To succeed in sequencing whole mitochondrial genomes using traditional CE detection methods, a large number of amplicons are needed, which requires a large quantity of extracted DNA; this work is laborious, time-consuming, and prone to sequencing errors, as shown previously.[43,44] To eliminate these problems, ancient DNA laboratories are now turning to a hybridization enrichment approach for sequencing whole mitochondrial genomes and have developed the ability to sequence samples that are thousands of years old.[45,46] For hybridization enrichment, probe DNA of a known haplotype can be used to capture and enrich target DNA of interest. The unwanted microbial contaminant DNA is then washed away, using stringency washes of increased temperature and decreased salt concentration. This allows only sequences of perfect homology (in the 20–70 bp range) to be captured and sequenced using NGS platforms.[39,47]

This approach has the potential to be adopted into mainstream forensic practice to assist with large-scale identification efforts in cases of mass disasters where the DNA is severely damaged, fragmented, and in low copy number. However, first a worldwide effort is required to produce large-scale mtDNA databases[48] compatible with such platforms.

To improve current use of mitochondrial DNA in forensic science, the quality of mtDNA sequencing and its corresponding database has been improved by the implementation of a quality assurance check of the data called the EMPOP project (http://empop.org).[49] EMPOP stands for European Mitochondrial DNA population database. Software inherent in the database screens sequencing data prone to errors, resulting in high-quality output data. Certain forensic journals now stipulate that mtDNA population studies need to be quality control checked by EMPOP first, to improve the overall quality of results before publication.

Ancestrally informative genetic markers

Genetic markers such as the Y chromosome (passed down the paternal line only), mitochondrial genome (maternally inherited only), and X chromosome (inherited from a father to his son or daughter or from a mother to her daughter only) can all be used to predict a person's ancestral lineage. In the Western developed world, up until the advent of canals in the mid-1700s and then the railways in the mid- to late 1800s, few people could travel far, and hence populations in towns and villages were made of people who shared recent common ancestors. More recently, mass transport has become available, but still many people remain close to the town or city where the rest of their family lives. The net result is that some DNA types are much more common in some local populations than in others, as these people share recent common ancestors. Sampling populations around the world has led to the observation of clustering of certain DNA types. There can be differences, however, between mitochondrial DNA types and Y-STR frequencies, as in many cultures, males and females move differently. An example is that males often inherit land from their fathers, and thus remain in the same place, but they may marry a woman from a nearby village, who will then move. Conversely, men in the historical past may have been parts of armies or colonizers, who then fathered children in the new local population. By looking at a wide range of mitochondrial, Y-chromosome, and nuclear markers from current populations, it is possible to predict from where a person most likely originated.

Variation in the genetic make-up of populations was first noticed in World War I, when the frequencies of the ABO blood group system were recorded and it was noted that soldiers from Western Europe had different frequencies than those from Eastern Europe (in fact, ABO B type is relatively common in Eastern Asia but becomes rarer moving westward to Europe, whereas the frequency of ABO A increases).[50] Leaping forward to current technology, with which many genetic markers can be examined, a more comprehensive picture of the genetic ancestry of a person can be developed.

This type of technology is particularly informative if a DNA profile from a body fluid stain (such as blood or semen) matches no known suspect or no one on any DNA database. A number of forensic cases highlight this use, including the identification of a serial rapist in London in the United

Kingdom. DNA samples from multiple rape cases were found to match each other, indicating a single rapist, and by looking at ancestrally informative DNA markers, it was proposed that the rapist most likely came from West Indian heritage, leading the police to ultimately find the suspect.[51]

Phenotypic markers

All the genetic markers discussed so far have no known phenotypic affect. The STR markers chosen originally were selected specifically such that a person's DNA profile would be a digital code that is meaningless unless compared with another digital code. However, as knowledge of the human genome has increased, so has the elucidation of certain genes that affect our appearance. At this time, these genes are limited to those for eye color and hair color, but other traits encoded genetically may soon be developed.

Hair color

The natural color of our hair is based on the DNA sequence of the melanocortin-1 receptor (*MCIR*). A number of single base changes can determine whether a person has red, black, or blonde hair. These single base changes (an example of a SNP) can be detected in a simple test.[52]

Eye color

The color of the iris is affected by SNPs within a range of genes. These SNPs are combined into a test called IrisPlex, which can be performed on any body fluid or tissue and lead to a prediction of the donor's eye color. The SNPs for eye color and hair color have been combined into one single SNP-based test called Hirisplex.[53,54] This test includes 24 SNPs from eleven different genes. In a multilaboratory trial of this assay, hair color had an accuracy of 69.5% correct calls for blonde, 78.5% for brown, 87.5% for black, and 80% for red hair.[53] In 92.5% of cases, the genetic eye color prediction test agreed with the eye color phenotype of the individual.[53] As more SNPs are found that affect either eye color or hair color, they can be added to increase the accuracy of the SNP test.

Accreditation: ISO17025

One of the greatest changes to have occurred very recently in the area of forensic science is not a new test per se but, rather, the introduction of accreditation to all operational laboratories. The ISO standard to which analytical science laboratories are assessed is ISO17025 (see http://www.iso.org/iso/home.html for more details). This standard sets in place standard operating procedures (often abbreviated to SOPs) that must be adhered to at all times, as well as competency tests. Examples of competency tests are proficiency tests where known samples are provided by a testing agency to whom the correct answers should be returned. These types of tests demonstrate confidence in the method and the ability of the staff.

Although miscarriages of justice can never be eradicated completely, as any test that involves a human is open to experimental error, such events are very rare in laboratories that operate to ISO 17025 standards. DNA profiling has been undoubtedly a spur toward accreditation, as suddenly there is a test that provides overwhelming powers of discrimination overshadowing other areas of traditional forensic science. If the DNA typing was in error, then a wrong verdict may be provided by the court, regardless of other scientific or circumstantial evidence. Not surprisingly, DNA has undergone extensive validation and verification, as recognized by the US National Academy of Sciences.[7]

NGS for forensic human identification

The use of DNA technology for the purpose of identification continues to promote interest in more powerful and novel methods of analysis. The latest tool in the field of genetic research and human identification testing is the application of NGS. Advances in genomic technology enable more and more information to be extracted from samples containing limited, degraded, and poor-quality DNA. Although CE detection methods work by distinguishing STR alleles according to the size of the amplicons, NGS can distinguish between alleles that are similar in length by providing information on the sequence content of the DNA (by typing Intra-STR SNPs) in addition to the length of the allele. NGS has the ability to target a larger number of PCR amplicons in a single assay by performing simultaneous analysis of different polymorphism types (eg, identity, ancestry, and phenotypic SNPs). The ability to barcode (individualize) samples[55] enables multiple samples to be processed in a single run, increasing the multiplexing and high-throughput capacity,[56] another major advantage to forensic labs. Small-scale bench top NGS machines[57,58] are already proving to be useful in certain forensic applications, as they are affordable, fast, easy-to-use and provide a large amount of data (10 megabases to several gigabases) in a relatively short time.

The workflow for NGS involves three fundamental steps: library preparation, sequencing, and data interpretation. The first stage in processing biological samples requires the amplification of forensically informative loci and the

creation of an amplicon library (collection of DNA fragments) of SNPs, STRs, and/or mtDNA to immortalize the extract. Library preparation involves ligating (adding on) adaptors (small sequences of synthetic DNA) to the ends of the fragments or amplicons of DNA once they are the correct sequencing length (200–400 bp). Before sequencing, library DNA is then amplified (to generate many copies), using primers (synthetic DNA of a known sequence) complementary to adaptor sequences. This library DNA can then serve as a template for the subsequent sequencing reaction. The postenriched library DNA is then sequenced to determine the order of nucleotide bases in the DNA sequence and the length of the fragments; current platforms include Illumina MiSeq/HiSeq, Life Technologies Ion Torrent PGM, and Roche GS Junior/454.[58,59] These platforms are capable of revealing the DNA sequence of thousands of individual molecules generated from library DNA. Signals received by the detection instrument are then translated into sequence, and the raw data can be analyzed using bioinformatic pipelines. Both sequence depth and coverage provided by NGS exceeds that generated by traditional CE detection methods by many orders of magnitude and, as discussed previously, is currently the best method of sequencing DNA that is highly degraded (<100 base pairs in length).[46]

NGS sequencing is the driving force behind ancient DNA research, in which the DNA may have been subjected to extreme environmental conditions for so long it appears damaged, degraded, and in low amounts compared with the unwanted microbial content. Ancient mitochondrial genomes are now being routinely sequenced using NGS,[60–62] and it is currently being used to advance metagenomic analysis of environmental samples, another important aspect in forensic science.

The potential does exist for NGS to drive forensic DNA research as the price of sequencing continues to drop. However, refinement and rigorous testing are necessary before it can be implemented in a forensic DNA testing laboratory. Robust validation processes are required before any new procedure can be adopted in an accredited forensic laboratory, and this ultimately slows down the whole process. In addition, selecting the right sequencing platform that meets the demands of sample size, coverage, cost, and accuracy is also a fundamental requirement for forensic DNA applications. In the meantime, research in this exciting new field will continue to advance.

Conclusion

As more is understood about the functioning of the human genome, then inevitably, aspects of this new knowledge will be adapted for use in forensic science. New technologies such as NGS and phenotypic typing are bound to open potential new avenues of forensic genetic typing. Over the last 2 decades, we have seen DNA typing progress from the examination of only a few genetic loci in a matter of days to sequencing the entire human genome in only a single day. A well-informed debate will be needed to determine how much knowledge law enforcement agencies should be allowed to acquire and who should have access. An individual's genome has always been seen as a very personal thing, yet both medical and forensic genetic testing can tell much about the individual. Although many would accept that there is very good reason for the police to determine information about an unknown suspect's potential ancestral background and some aspects of their complexion, many might find the potential to determine genetic dispositions to certain disorders as unacceptable. These new technologies need to be used wisely and for the reasons that they are intended. Despite these comments of concern, undoubtedly forensic DNA profiling has led to both exonerations of the innocent and implication of the guilty on a scale never seen before in the criminal justice system. It would be wrong not to embrace new technologies as they become available in the cause of justice.

Disclosure

The authors report no conflicts of interest in this work.

References

1. Goodwin W, Linacre A, Hadi S. *An Introduction to Forensic Genetics*. Chichester, West Sussex, UK: Wiley-Blackwell; 2011.
2. Butler J. *Advanced Topics in Forensic DNA Typing: Methodology*. Waltham, MA: Academic Press; 2012.
3. Liu RH. *Forensic DNA Applications*. Boca Raton, FL: CRC Press; 2014.
4. Jeffreys AJ, Brookfield JFY, Semeonoff R. Positive identification of an immigration test-case using human DNA fingerprints. *Nature*. 1985;317(6040):818–819.
5. Seton C. Life for sex killer who sent decoy to take genetic test. *The Times*. January 23, 1988.
6. Mullis KB, Faloona FA. Specific synthesis of DNA in vitro via a polymerase-catalyzed chain reaction. *Methods Enzymol*. 1987;155:335–350.
7. Academies NRCotN. *Strengthening Forensic Science in the United States: A Path Forward*. Washington, DC: National Academies Press; 2009.
8. Sambrook J, Fritsch EF, Maniatis T. *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press; 1989.
9. Walsh PS, Metzger DA, Higuchi R. Chelex 100 as a medium for simple extraction of DNA for PCR-based typing from forensic material. *Biotechniques*. 1991;10(4):506–513.
10. Castella V, Dimo-Simonin N, Brandt-Casadevall C, Mangin P. Forensic evaluation of the QIAshredder/QIAamp DNA extraction procedure. *Forensic Sci Int*. 2006;156(1):70–73.
11. Witt S, Neumann J, Zierdt H, G bel G, R scheisen C. Establishing a novel automated magnetic bead-based method for the extraction of DNA from a variety of forensic samples. *Forensic Sci Int Genet*. 2012;6(5):539–547.

12. Birch DE. Simplified hot start PCR. *Nature*. 1996;381(6581):445–446.
13. Moretti T, Koons B, Budowle B. Enhancement of PCR amplification yield and specificity using AmpliTaq Gold DNA polymerase. *Biotechniques*. 1998;25(4):716–722.
14. Kimpton C, Fisher D, Watson S, et al. Evaluation of an automated DNA profiling system employing multiplex amplification of four tetrameric STR loci. *Int J Legal Med*. 1994;106(6):302–311.
15. Lygo JE, Johnson PE, Holdaway DJ, et al. The validation of short tandem repeat (STR) loci for use in forensic casework. *Int J Legal Med*. 1994;107(2):77–89.
16. Gill P, Sparkes R, Kimpton C. Development of guidelines to designate alleles using an STR multiplex system. *Forensic Sci Int*. 1997;89(3):185–197.
17. Kimpton CP, Oldroyd NJ, Watson SK, et al. Validation of highly discriminating multiplex short tandem repeat amplification systems for individual identification. *Electrophoresis*. 1996;17(8):1283–1293.
18. Sparkes R, Kimpton C, Gilbard S, et al. The validation of a 7-locus multiplex STR test for use in forensic casework. (II), Artefacts, casework studies and success rates. *Int J Legal Med*. 1996;109(4):195–204.
19. Sparkes R, Kimpton C, Watson S, et al. The validation of a 7-locus multiplex STR test for use in forensic casework. (I). Mixtures, ageing, degradation and species studies. *Int J Legal Med*. 1996;109(4):186–194.
20. Krenke BE, Tereba A, Anderson SJ, et al. Validation of a 16-locus fluorescent multiplex system. *J Forensic Sci*. 2002;47(4):773–785.
21. Templeton A. Out of Africa again and again. *Nature*. 2002;416(6876):45–51.
22. Ingman M, Kaessmann H, P bo S, Gyllensten U. Mitochondrial genome variation and the origin of modern humans. *Nature*. 2000;408(6813):708–713.
23. Thompson JM, Ewing MM, Frank WE, et al. Developmental validation of the PowerPlex® Y23 System: a single multiplex Y-STR analysis system for casework and database samples. *Forensic Sci Int Genet*. 2013;7(2):240–250.
24. Turrina S, Ferrian M, Caratti S, De Leo D. Evaluation of genetic parameters of 22 autosomal STR loci (PowerPlex® Fusion System) in a population sample from Northern Italy. *Int J Legal Med*. 2014;128(2):281–283.
25. Ballantyne KN, Keerl V, Wollstein A, et al. A new future of forensic Y-chromosome analysis: rapidly mutating Y-STRs for differentiating male relatives and paternal lineages. *Forensic Sci Int Genet*. 2012;6(2):208–218.
26. Bosch E, Calafell F, Santos FR, et al. Variation in short tandem repeats is deeply structured by genetic background on the human Y chromosome. *Am J Hum Genet*. 1999;65(6):1623–1638.
27. Schmidtke J, K hnau W, Wand D, Edelmann J, Szibor R, Krawczak M. Prenatal exclusion without involving the putative fathers of an incestuous father-daughter parenthood. *Prenat Diagn*. 2004;24(8):662–664.
28. Diegoli TM, Coble MD. Development and characterization of two mini-X chromosomal short tandem repeat multiplexes. *Forensic Sci Int Genet*. 2011;5(5):415–421.
29. Diegoli TM, Linacre A, Coble MD. Population genetic data for 15 X chromosomal short tandem repeat markers in three US populations. *Forensic Sci Int Genet*. 2014;8(1):64–67.
30. Schanfield MS, Tiesma D, Diegoli T, Coble M, Crawford M. Preliminary report on the anthropology of 15 X STR loci. *Am J Phys Anthropol*. 2013;150:242–242.
31. Cavelier L, Jazin E, Jalonen P, Gyllensten U. MtDNA substitution rate and segregation of heteroplasmy in coding and noncoding regions. *Hum Genet*. 2000;107(1):45–50.
32. Garrido N, Griparic L, Jokitalo E, Wartiovaara J, van der Bliek AM, Spelbrink JN. Composition and dynamics of human mitochondrial nucleoids. *Mol Biol Cell*. 2003;14(4):1583–1596.
33. Holland MM, Fisher DL, Mitchell LG, et al. Mitochondrial DNA sequence analysis of human skeletal remains: identification of remains from the Vietnam War. *J Forensic Sci*. 1993;38(3):542–553.
34. Zupanič Pajnic I, Gornjak Pogorelc B, Bala ic J. Molecular genetic identification of skeletal remains from the Second World War Konfin I mass grave in Slovenia. *Int J Legal Med*. 2010;124(4):307–317.
35. Gill P, Ivanov PL, Kimpton C, et al. Identification of the remains of the Romanov family by DNA analysis. *Nat Genet*. 1994;6(2):130–135.
36. Coble MD, Loreille OM, Wadhams MJ, et al. Mystery solved: the identification of the two missing Romanov children using DNA analysis. *PLoS ONE*. 2009;4(3):e4838.
37. Wilson MR, Holland MM, Stoneking M, DiZinno JA, Budowle B. Guidelines for the use of mitochondrial DNA sequencing in forensic science. *Crime Lab Digest*. 1993;20(4):68–77.
38. Parson W, Bandelt HJ. Extended guidelines for mtDNA typing of population data in forensic science. *Forensic Sci Int Genet*. 2007;1(1):13–19.
39. Brotherton P, Haak W, Templeton J, et al; Genographic Consortium. Neolithic mitochondrial haplogroup H genomes and the genetic origins of Europeans. *Nat Commun*. 2013;4:1764.
40. Just RS, Leney MD, Barritt SM, et al. The use of mitochondrial DNA single nucleotide polymorphisms to assist in the resolution of three challenging forensic cases. *J Forensic Sci*. 2009;54(4):887–891.
41. Just RS, Loreille OM, Molto JE, et al. Titanic's unknown child: the critical role of the mitochondrial DNA coding region in a re-identification effort. *Forensic Sci Int Genet*. 2011;5(3):231–235.
42. Melton T, Clifford S, Kayser M, Nasidze I, Batzer M, Stoneking M. Diversity and heterogeneity in mitochondrial DNA of North American populations. *J Forensic Sci*. 2001;46(1):46–52.
43. Bandelt HJ, Salas A, Lutz-Bonengel S. Artificial recombination in forensic mtDNA population databases. *Int J Legal Med*. 2004;118(5):267–273.
44. Forster P. To err is human. *Ann Hum Genet*. 2003;67(Pt 1):2–4.
45. Briggs AW, Good JM, Green RE, et al. Targeted retrieval and analysis of five Neandertal mtDNA genomes. *Science*. 2009;325(5938):318–321.
46. Maricic T, Whitten M, P bo S. Multiplexed DNA sequence capture of mitochondrial genomes using PCR products. *PLoS ONE*. 2010;5(11):e14004.
47. Templeton JE, Brotherton PM, Llamas B, et al. DNA capture and next-generation sequencing can recover whole mitochondrial genomes from highly degraded samples for human identification. *Investig Genet*. 2013;4(1):26.
48. Irwin JA, Parson W, Coble MD, Just RS. mtGenome reference population databases and the future of forensic mtDNA analysis. *Forensic Sci Int Genet*. 2011;5(3):222–225.
49. Parson W, D r A. EMPOP – a forensic mtDNA database. *Forensic Sci Int Genet*. 2007;1(2):88–92.
50. Sykes B. *Blood of the Isles*. London, UK: Bantam Press; 2006.
51. DNA profiling used to track rapist. *Independent*. March 24, 2011. Available from: http://www.independent.co.uk/news/uk/crime/dna-profiling-used-to-track-rapist-2251753.html. Accessed August 8, 2014.
52. Liu F, Struchalin MV, Duijn K, et al. Detecting low frequent loss-of-function alleles in genome wide association studies with red hair color as example. *PLoS ONE*. 2011;6(11):e28145.
53. Walsh S, Liu F, Ballantyne KN, van Oven M, Lao O, Kayser M. IrisPlex: a sensitive DNA tool for accurate prediction of blue and brown eye colour in the absence of ancestry information. *Forensic Sci Int Genet*. 2011;5(3):170–180.
54. Kayser M, de Knijff P. Improving human forensics through advances in genetics, genomics and molecular biology. *Nat Rev Genet*. 2011;12(3):179–192.
55. Meyer M, Stenzel U, Myles S, Pr fer K, Hofreiter M. Targeted high-throughput sequencing of tagged nucleic acid samples. *Nucleic Acids Res*. 2007;35(15):e97.
56. Gnirke A, Melnikov A, Maguire J, et al. Solution hybrid selection with ultra-long oligonucleotides for massively parallel targeted sequencing. *Nat Biotechnol*. 2009;27(2):182–189.

57. Rothberg JM, Hinz W, Rearick TM, et al. An integrated semiconductor device enabling non-optical genome sequencing. *Nature*. 2011; 475(7356):348–352.
58. Loman NJ, Misra RV, Dallman TJ, et al. Performance comparison of benchtop high-throughput sequencing platforms. *Nat Biotechnol*. 2012; 30(5):434–439.
59. Merriman B; Ion Torrent R&D Team, Rothberg JM. Progress in ion torrent semiconductor chip based sequencing. *Electrophoresis*. 2012;33(23):3397–3417.
60. Green RE, Krause J, Briggs AW, et al. A draft sequence of the Neandertal genome. *Science*. 2010;328(5979):710–722.
61. Krause J, Fu Q, Good JM, et al. The complete mitochondrial DNA genome of an unknown hominin from southern Siberia. *Nature*. 2010;464(7290):894–897.
62. Fu Q, Rudan P, P bo S, Krause J. Complete mitochondrial genomes reveal neolithic expansion into Europe. *PLoS ONE*. 2012;7(3):e32473.

1.11 References

1. Gill, P., et al., *An investigation of the rigor of interpretation rules for STRs derived from less than 100 pg of DNA.* Forensic Science International, 2000. **112**(1): p. 17-40.
2. Harbison, S., M. Fallow, and D. Bushell, *An analysis of the success rate of 908 trace DNA samples submitted to the Crime Sample Database Unit in New Zealand.* Australian Journal of Forensic Sciences, 2008. **40**(1): p. 49-53.
3. van Oorschot, R.A., K.N. Ballantyne, and R.J. Mitchell, *Forensic trace DNA: a review.* Investigative Genetics, 2010. **1**(1): p. 1.
4. Meakin, G. and A. Jamieson, *DNA transfer: review and implications for casework.* Forensic Science International: Genetics, 2013. **7**(4): p. 434-443.
5. van Oorschot, R.A. and M.K. Jones, *DNA fingerprints from fingerprints.* Nature London, 1997: p. 767-767.
6. Saferstein, R., *Forensic Science Handbook, Volume 1*. 2002: Prentice Hall.
7. Quinones, I. and B. Daniel, *Cell free DNA as a component of forensic evidence recovered from touched surfaces.* Forensic Science International: Genetics, 2012. **6**(1): p. 26-30.
8. Findlay, I., et al., *DNA fingerprinting from single cells [10].* Nature, 1997. **389**(6651): p. 555-556.
9. Kita, T., et al., *Morphological study of fragmented DNA on touched objects.* Forensic Science International: Genetics, 2008. **3**(1): p. 32-6.
10. Taylor, N.A. and C.A. Machado-Moreira, *Regional variations in transepidermal water loss, eccrine sweat gland density, sweat secretion rates and electrolyte composition in resting and exercising humans.* Extreme Physiology & Medicine, 2013. **2**(1): p. 1.
11. Wickenheiser, R.A., *Trace DNA: a review, discussion of theory, and application of the transfer of trace quantities of DNA through skin contact.* Journal of Forensic Sciences, 2002. **47**(3): p. 442-450.
12. Raymond, J.J., et al., *Trace DNA and street robbery: a criminalistic approach to DNA evidence.* Forensic Science International: Genetics Supplement Series, 2009. **2**(1): p. 544-546.
13. van Oorschot, R.A.H., et al., *Are you collecting all the available DNA from touched objects?* International Congress Series, 2003. **1239**(0): p. 803-807.

14. Sprecher, C.J., et al., *PowerPlex® ESX and ESI Systems: A suite of new STR systems designed to meet the changing needs of the DNA-typing community.* Forensic Science International: Genetics Supplement Series, 2009. **2**(1): p. 2-4.
15. Hennessy, L.K., et al., *Developmental validation of the GlobalFiler® express kit, a 24-marker STR assay, on the RapidHIT® System.* Forensic Science International: Genetics, 2014. **13**: p. 247-258.
16. Martín, P., et al., *Improving DNA data exchange: validation studies on a single 6 dye STR kit with 24 loci.* Forensic Science International: Genetics, 2014. **13**: p. 68-78.
17. Flores, S., et al., *Internal validation of the GlobalFiler™ Express PCR Amplification Kit for the direct amplification of reference DNA samples on a high-throughput automated workflow.* Forensic Science International: Genetics, 2014. **10**: p. 33-39.
18. Montpetit, S. and P. O'Donnell, *An optimized procedure for obtaining DNA from fired and unfired ammunition.* Forensic Science International: Genetics, 2015. **17**: p. 70-74.
19. Edwards, A., et al., *DNA typing and genetic mapping with trimeric and tetrameric tandem repeats.* American Journal of Human Genetics, 1991. **49**(4): p. 746.
20. Edwards, A., et al., *Genetic variation at five trimeric and tetrameric tandem repeat loci in four human population groups.* Genomics, 1992. **12**(2): p. 241-253.
21. Fregeau, C.J. and R.M. Fourney, *DNA typing with fluorescently tagged short tandem repeats: a sensitive and accurate approach to human identification.* Biotechniques, 1993. **15**(1): p. 100-119.
22. van Oorschot, R.A., S.J. Gutowski, and S.L. Robinson, *HUMTH01: amplification, species specificity, population genetics and forensic applications.* International Journal of Legal Medicine, 1994. **107**(3): p. 121-126.
23. Alaeddini, R., S.J. Walsh, and A. Abbas, *Forensic implications of genetic analyses from degraded DNA—a review.* Forensic Science International: Genetics, 2010. **4**(3): p. 148-157.

24. Raymond, J.J., et al., *Trace DNA success rates relating to volume crime offences.* Forensic Science International: Genetics Supplement Series, 2009. **2**(1): p. 136-137.
25. Alaeddini, R., *Forensic implications of PCR inhibition - A review.* Forensic Science International: Genetics, 2012. **6**(3): p. 297-305.
26. Adamowicz, M.S., et al., *Evaluation of Methods to Improve the Extraction and Recovery of DNA from Cotton Swabs for Forensic Analysis.* PLOSONE, 2014. **9**(12): p. e116351.
27. Kishore, R., et al., *Optimization of DNA Extraction from Low-Yield and Degraded Samples Using the BioRobot® EZ1 and BioRobot® M48.* Journal of Forensic Sciences, 2006. **51**(5): p. 1055-1061.
28. Thomasma, S.M. and D.R. Foran, *The influence of swabbing solutions on DNA recovery from touch samples.* Journal of Forensic Sciences, 2013. **58**(2): p. 465-469.
29. Ottens, R., et al., *Application of direct PCR in forensic casework.* Forensic Science International: Genetics Supplement Series, 2013. **4**(1): p. e47-e48.
30. Ottens, R., D. Taylor, and A. Linacre, *DNA profiles from fingernails using direct PCR.* Forensic Science, Medicine, and Pathology, 2015. **11**(1): p. 99-103.
31. Ottens, R., et al., *Optimising direct PCR from anagen hair samples.* Forensic Science International: Genetics Supplement Series, 2013. **4**(1): p. e109-e110.
32. Swaran, Y.C. and L. Welch, *A comparison between direct PCR and extraction to generate DNA profiles from samples retrieved from various substrates.* Forensic Science International: Genetics, 2012. **6**(3): p. 407-412.
33. Linacre, A., et al., *Generation of DNA profiles from fabrics without DNA extraction.* Forensic Science International: Genetics, 2010. **4**(2): p. 137-141.
34. Templeton, J.E., et al., *DNA capture and next-generation sequencing can recover whole mitochondrial genomes from highly degraded samples for human identification.* Investigative Genetics, 2013. **4**(1): p. 1-13.
35. Horsman-Hall, K.M., et al., *Development of STR profiles from firearms and fired cartridge cases.* Forensic Science International: Genetics, 2009. **3**(4): p. 242-250.

36. Goray, M., et al., *Secondary DNA transfer of biological substances under varying test conditions.* Forensic Science International: Genetics, 2010. **4**(2): p. 62-67.
37. Goray, M. and R.A. van Oorschot, *The complexities of DNA transfer during a social setting.* Legal Medicine, 2015. **17**(2): p. 82-91.
38. Daly, D.J., C. Murphy, and S.D. McDermott, *The transfer of touch DNA from hands to glass, fabric and wood.* Forensic Science International: Genetics, 2012. **6**(1): p. 41-46.
39. Balogh, M.K., et al., *STR genotyping and mtDNA sequencing of latent fingerprint on paper.* Forensic Science International, 2003. **137**(2-3): p. 188-195.
40. Lowe, A., et al., *The propensity of individuals to deposit DNA and secondary transfer of low level DNA from individuals to inert surfaces.* Forensic Science International, 2002. **129**(1): p. 25-34.
41. Phipps, M. and S. Petricevic, *The tendency of individuals to transfer DNA to handled items.* Forensic Science International, 2007. **168**(2-3): p. 162-8.
42. Farmen, R.K., et al., *Assessment of individual shedder status and implication for secondary DNA transfer.* Forensic Science International: Genetics Supplement Series, 2008. **1**(1): p. 415-417.
43. Irwin, J.A., et al., *Application of low copy number STR typing to the identification of aged, degraded skeletal remains.* Journal of Forensic Sciences, 2007. **52**(6): p. 1322-1327.
44. Kontanis, E.J. and F.A. Reed, *Evaluation of Real-Time PCR Amplification Efficiencies to Detect PCR Inhibitors.* Journal of Forensic Sciences, 2006. **51**(4): p. 795-804.
45. *Layers of the skin*. 2016; Available from: https://en.wikipedia.org/wiki/Human_skin.
46. Atamaniuk, J., et al., *Increased concentrations of cell-free plasma DNA after exhaustive exercise.* Clinical Chemistry, 2004. **50**(9): p. 1668-1670.
47. Vandewoestyne, M., et al., *Presence and potential of cell free DNA in different types of forensic samples.* Forensic Science International: Genetics, 2013. **7**(2): p. 316-20.

48. Cale, C.M., et al., *Could Secondary DNA Transfer Falsely Place Someone at the Scene of a Crime?* Journal of Forensic Sciences, 2016. **61**(1): p. 196-203.
49. Goray, M., R.A. van Oorschot, and J.R. Mitchell, *DNA transfer within forensic exhibit packaging: Potential for DNA loss and relocation.* Forensic Science International: Genetics, 2012. **6**(2): p. 158-166.
50. Helmus, J., T. Bajanowski, and M. Poetsch, *DNA transfer—a never ending story. A study on scenarios involving a second person as carrier.* International Journal of Legal Medicine, 2016. **130**(1): p. 121-125.
51. van Oorschot, R.A., G. Glavich, and R.J. Mitchell, *Persistence of DNA deposited by the original user on objects after subsequent use by a second person.* Forensic Science International: Genetics, 2014. **8**(1): p. 219-225.
52. Goray, M., et al., *Shedder status - An analysis of self and non-self DNA in multiple handprints deposited by the same individuals over time.* Forensic Science International: Genetics, 2016. **23**: p. 190-196.
53. Ladd, C., et al., *A systematic analysis of secondary DNA transfer.* Journal of Forensic Sciences, 1999. **44**: p. 1270-1272.
54. Taylor, D., et al., *Observations of DNA transfer within an operational Forensic Biology Laboratory.* Forensic Science International: Genetics, 2016. **23**: p. 33-49.
55. Butler, J.M., Y. Shen, and B.R. McCord, *The development of reduced size STR amplicons as tools for analysis of degraded DNA.* Journal of Forensic Sciences, 2003. **48**(5): p. 1054-1064.
56. Kopka, J., et al., *New optimized DNA extraction protocol for fingerprints deposited on a special self-adhesive security seal and other latent samples used for human identification.* Journal of Forensic Sciences, 2011. **56**(5): p. 1235-40.
57. Lagoa, A.M., T. Magalhães, and M.F. Pinheiro, *Autosomic STR, Y-STR and miniSTR markers evaluation for genetic analysis of fingerprints.* Forensic Science International: Genetics Supplement Series, 2008. **1**(1): p. 435-436.
58. Poy, A.L. and R.A. van Oorschot, *Trace DNA presence, origin, and transfer within a forensic biology laboratory and its potential effect on casework.* Journal of Forensic Identification, 2006. **56**(4): p. 558.

59. Budowle, B., et al., *Low copy number—consideration and caution.* Twelfth International Symposium on Human Identification, 2001.
60. Caddy B, T.G., Linacre AMT. *A review of the science of low template DNA analysis*. 2008; Available from: http://www.homeoffice.gov.uk/publications/police/790604/Review_of_Low_Template_DNA_1.pdf.
61. Kobilinsky, L., T. Liotti, and J. Oeser-Sweat, *DNA Forensic and Legal Applications*. 2005, New Jersey: Wiley-Interscience.
62. Kasai, K., R. White, and Y. Nakamura, *Amplification of a variable number of tandem repeats (VNTR) locus (pMCT118) by the polymerase chain reaction (PCR) and its application to forensic science.* Journal of Forensic Science, 1990. **35**(5): p. 1196-1200.
63. Budowle, B., et al., *Analysis of the VNTR locus D1S80 by the PCR followed by high-resolution PAGE.* American Journal of Human Genetics, 1991. **48**(1): p. 137.
64. Sajantila, A., et al., *PCR amplification of alleles at the DIS80 locus: comparison of a Finnish and a North American Caucasian population sample, and forensic casework evaluation.* American Journal of Human Genetics, 1992. **50**(4): p. 816.
65. Gill, P., *Role of short tandem repeat DNA in forensic casework in the UK-past, present, and future perspectives.* Biotechniques, 2002. **32**(2): p. 366-385.
66. Gill, P., et al., *Report of the European DNA profiling group (EDNAP)—towards standardisation of short tandem repeat (STR) loci.* Forensic Science International, 1994. **65**(1): p. 51-59.
67. Gill, P., *An assessment of the utility of single nucleotide polymorphisms (SNPs) for forensic purposes.* International Journal of Legal Medicine, 2001. **114**(4-5): p. 204-210.
68. Sanchez, J.J., et al., *Forensic typing of autosomal SNPs with a 29 SNP-multiplex—results of a collaborative EDNAP exercise.* Forensic Science International: Genetics, 2008. **2**(3): p. 176-183.
69. Musgrave-Brown, E., et al., *Forensic validation of the SNPforID 52-plex assay.* Forensic Science International: Genetics, 2007. **1**(2): p. 186-190.

70. Budowle, B., *SNP typing strategies.* Forensic Science International, 2004. **146 Suppl**: p. S139-42.
71. Alimat, S., S. Hadi, and W. Goodwin, *SNP genotyping of forensic casework samples using the 52 SNPforID markers.* Forensic Science International: Genetics Supplement Series, 2013. **4**(1): p. e178-e179.
72. Takada, Y., et al., *Application of SNPs in forensic casework: identification of pathological and autoptical specimens due to sample mix-up.* Legal Medicine (Tokyo), 2009. **11 Suppl 1**: p. S196-7.
73. Børsting, C., H.S. Mogensen, and N. Morling, *Forensic genetic SNP typing of low-template DNA and highly degraded DNA from crime case samples.* Forensic Science International: Genetics, 2013. **7**(3): p. 345-352.
74. Phillips, C., et al., *Inferring ancestral origin using a single multiplex assay of ancestry-informative marker SNPs.* Forensic Science International: Genetics, 2007. **1**(3): p. 273-280.
75. Børsting, C., et al., *Performance of the SNPforID 52 SNP-plex assay in paternity testing.* Forensic Science International: Genetics, 2008. **2**(4): p. 292-300.
76. Brotherton, P., et al., *Preferential access to genetic information from endogenous hominin ancient DNA and accurate quantitative SNP-typing via SPEX.* Nucleic Acids Research, 2010. **38**(2).
77. Amorim, A. and L. Pereira, *Pros and cons in the use of SNPs in forensic kinship investigation: a comparative analysis with STRs.* Forensic Science International, 2005. **150**(1): p. 17-21.
78. Fondevila, M., et al., *Revision of the SNPforID 34-plex forensic ancestry test: assay enhancements, standard reference sample genotypes and extended population studies.* Forensic Science International: Genetics, 2013. **7**(1): p. 63-74.
79. Daniel, R., et al., *A SNaPshot of next generation sequencing for forensic SNP analysis.* Forensic Science International: Genetics, 2015. **14**: p. 50-60.
80. Senge, T., et al., *STRs, mini STRs and SNPs--a comparative study for typing degraded DNA.* Legal Medicine (Tokyo), 2011. **13**(2): p. 68-74.

81. Dixon, L., et al., *Validation of a 21-locus autosomal SNP multiplex for forensic identification purposes.* Forensic Science International, 2005. **154**(1): p. 62-77.
82. Walsh, S., et al., *IrisPlex: a sensitive DNA tool for accurate prediction of blue and brown eye colour in the absence of ancestry information.* Forensic Science International: Genetics, 2011. **5**(3): p. 170-180.
83. Walsh, S., et al., *The HIrisPlex system for simultaneous prediction of hair and eye colour from DNA.* Forensic Science International: Genetics, 2013. **7**(1): p. 98-115.
84. Wetton, J.H., K.W. Tsang, and H. Khan, *Inferring the population of origin of DNA evidence within the UK by allele-specific hybridization of Y-SNPs.* Forensic Science International, 2005. **152**(1): p. 45-53.
85. Butler, J.M. and B.C. Levin, *Forensic applications of mitochondrial DNA.* Trends in Biotechnology, 1998. **16**(4): p. 158-162.
86. Irwin, J.A., et al., *mtGenome reference population databases and the future of forensic mtDNA analysis.* Forensic Science International: Genetics, 2011. **5**(3): p. 222-225.
87. Piercy, R., et al., *The application of mitochondrial DNA typing to the study of white Caucasian genetic identification.* International Journal of Legal Medicine, 1993. **106**(2): p. 85-90.
88. Koehnemann, S. and H. Pfeiffer, *Application of mtDNA SNP analysis in forensic casework.* Forensic Science International: Genetics, 2011. **5**(3): p. 216-221.
89. Budowle, B., et al., *Forensics and mitochondrial DNA: Applications, debates, and foundations.* Annual Review of Genomics and Human Genetics, 2003. **4**: p. 119-141.
90. Goodwin, W., A. Linacre, and P. Vanezis, *The use of mitochondrial DNA and short tandem repeat typing in the identification of air crash victims.* Electrophoresis, 1999. **20**(8): p. 1707-1711.
91. Foran, D.R., M.E. Gehring, and S.E. Stallworth, *The recovery and analysis of mitochondrial DNA from exploded pipe bombs.* Journal of Forensic Sciences, 2009. **54**(1): p. 90-94.

92. Holland, M.M., et al., *Mitochondrial DNA sequence analysis of human skeletal remains: identification of remains from the Vietnam War.* Journal of Forensic Sciences, 1993(38): p. 542-53.
93. Hansson, O., et al., *Trace DNA collection—Performance of minitape and three different swabs.* Forensic Science International: Genetics Supplement Series, 2009. **2**(1): p. 189-190.
94. van Oorschot, R., et al., *Improving collection methods can improve the ability to obtain typings from trace amounts of DNA from touched objects.* The International Congress of Genetics, Melbourne, Proceedings, 2003.
95. Franco, M. and R. Goetz, *A new method to recover trace DNA.* Proceedings of Human Identification, 2006.
96. Barash, M., A. Reshef, and P. Brauner, *The use of adhesive tape for recovery of DNA from crime scene items.* Journal of Forensic Sciences, 2010. **55**(4): p. 1058-1064.
97. Brownlow, R.J., K.E. Dagnall, and C.E. Ames, *A comparison of DNA collection and retrieval from two swab types (cotton and nylon flocked swab) when processed using three QIAGEN extraction methods.* Journal of Forensic Sciences, 2012. **57**(3): p. 713-717.
98. Verdon, T.J., R.J. Mitchell, and R.A. Oorschot, *Swabs as DNA collection devices for sampling different biological materials from different substrates.* Journal of Forensic Sciences, 2014. **59**(4): p. 1080-1089.
99. Mawlood, S.K., M. Alrowaithi, and N. Watson, *Advantage of ForensiX Swabs in Retrieving and Preserving Biological Fluids.* Journal of Forensic Sciences, 2015. **60**(3): p. 686-689.
100. Pang, B.C. and B.K. Cheung, *Double swab technique for collecting touched evidence.* Legal Medicine (Tokyo), 2007. **9**(4): p. 181-4.
101. Colussi, A., et al., *Efficiency of DNA IQ System® in recovering semen DNA from cotton swabs.* Forensic Science International: Genetics Supplement Series, 2009. **2**(1): p. 87-88.
102. Butler, J.M., *Genetics and genomics of core short tandem repeat loci used in human identity testing.* Journal of Forensic Sciences, 2006. **51**(2): p. 253-265.

103. Altshuler, H. and R. Roy, *Evaluation of direct PCR amplification using various swabs and washing Reagents.* Journal of Forensic Sciences, 2015. **60**(6): p. 1542-1552.
104. Gaillard, C. and F. Strauss, *Eliminating DNA loss and denaturation during storage in plastic microtubes.* American Clinical Laboratory, 2001. **20**(2): p. 52-54.
105. Linacre, A. and J. Templeton, *Forensic DNA profiling: state of the art.* Research and Reports in Forensic Medical Science, 2014. **4**: p. 25-36.
106. Balogh, M.K., et al., *Fingerprints from fingerprints.* International Congress Series, 2003. **1239**(0): p. 953-957.
107. Barbaro, A., et al., *DNA profiling by different extraction methods.* International Congress Series, 2004. **1261**: p. 562-564.
108. Butler, J.M., *Forensic DNA Typing: Biology, Technology, and Genetics of STR Markers*. 2nd Edition ed. 2005, New York: Elsevier Academic Press.
109. Hill, C.R., et al., *Characterization of 26 miniSTR loci for improved analysis of degraded DNA samples.* Journal of Forensic Sciences, 2008. **53**(1): p. 73-80.
110. Walsh, P.S., D.A. Metzger, and R. Higuchi, *Chelex 100 as a medium for simple extraction of DNA for PCR-based typing from forensic material.* Biotechniques, 1991. **10**(4): p. 506-513.
111. Sambrook, J., E. Fritsch, and T. Maniatis, *Molecular Cloning: A Laboratory Reference Manual*. 1989, Cold Spring Harbor Lab. Press, Plainview, NY.
112. Burgoyne, L.A., *Dry solid medium for storage and analysis of genetic material.* 1998.
113. Smith, L. and L.A. Burgoyne, *Collecting, archiving and processing DNA from wildlife samples using FTA® databasing paper.* BMC Ecology, 2004. **4**(1): p. 4.
114. Seah, L. and L. Burgoyne, *DNA data-basing on FTA® paper: Biological assault and techniques for measuring photogenic damage.* Progress in Forensic Genetics, 2000. **8**.
115. Frégeau, C.J., C.M. Lett, and R.M. Fourney, *Validation of a DNA IQ™-based extraction method for TECAN robotic liquid handling workstations for processing casework.* Forensic Science International: Genetics, 2010. **4**(5): p. 292-304.

116. Côté, A., et al., *Automated DNA extraction from large volumes.* Forensic Science International: Genetics Supplement Series, 2008. **1**(1): p. 22-23.
117. Höss, M. and S. Pääbo, *DNA extraction from Pleistocene bones by a silica-based purification method.* Nucleic Acids Research, 1993. **21**(16): p. 3913-3914.
118. Lee, H.Y., et al., *Simple and highly effective DNA extraction methods from old skeletal remains using silica columns.* Forensic Science International-Genetics, 2010. **4**(5): p. 275-280.
119. Shaw, K.J., et al., *The use of carrier RNA to enhance DNA extraction from microfluidic-based silica monoliths.* Analytica Chimica Acta, 2009. **652**(1): p. 231-233.
120. Greenspoon, S.A., et al., *Application of the BioMek (R) 2000 laboratory automation workstation and the DNA IQ (TM) system to the extraction of forensic casework samples.* Journal of Forensic Science, 2004. **49**(1): p. 29-39.
121. Mandrekar, P.V., B.E. Krenke, and A. Tereba, *DNA IQ™: The intelligent way to purify DNA.* Profiles in DNA, 2001. **4**(3): p. 16.
122. Tan, E., et al., *Deoxybonucleic acid (DNA) and antibodies to DNA in the serum of patients with systemic lupus erythematosus.* Journal of Clinical Investigation, 1966. **45**(11): p. 1732.
123. Herber, B. and K. Herold, *DNA typing of human dandruff.* Journal of Forensic Science, 1998. **43**(3): p. 648-656.
124. Kamphausen, T., et al., *Good shedder or bad shedder—the influence of skin diseases on forensic DNA analysis from epithelial abrasions.* International Journal of Legal Medicine, 2012. **126**(1): p. 179-183.
125. Nakamura, K., et al., *The relationship between apoptosis and keratinization in human epidermis.* Acta Histochemica et Cytochemica, 1999. **32**(1): p. 77-83.
126. Lacerenza, D., et al., *A molecular exploration of human DNA/RNA co-extracted from the palmar surface of the hands and fingers.* Forensic Science International: Genetics, 2016. **22**: p. 44-53.
127. Schiffner, L.A., et al., *Optimization of a simple, automatable extraction method to recover sufficient DNA from low copy number DNA samples for generation of short tandem repeat profiles.* Croatian Medical Journal, 2005. **46**(4): p. 578.

128. Hudlow, W.R., et al., *The NucleoSpin sup® DNA Clean-up XS kit for the concentration and purification of genomic DNA extracts: An alternative to microdialysis filtration.* Forensic Science International: Genetics, 2011. **5**(3): p. 226-230.
129. Fregeau, C.J. and A. De Moors, *Competition for DNA binding sites using Promega DNA IQ (TM) paramagnetic beads.* Forensic Science International-Genetics, 2012. **6**(5): p. 511-522.
130. Rohland, N. and M. Hofreiter, *Ancient DNA extraction from bones and teeth.* Nature protocols, 2007. **2**(7): p. 1756-1762.
131. Park, S.J., et al., *Direct STR Amplification from Whole Blood and Blood-or Saliva-Spotted FTA® without DNA Purification*.* Journal of Forensic Sciences, 2008. **53**(2): p. 335-341.
132. Barbaro, A., P. Cormaci, and A. Teatino, *Use of "AnyDirect PCR buffer" for PCR amplification of washed bloodstains: A case report.* Forensic Science International: Genetics Supplement Series, 2008. **1**(1): p. 11-12.
133. Southern, E.M., *Detection of specific sequences among DNA fragments separated by gel electrophoresis.* Journal of Molecular Biology, 1975. **98**(3): p. 503-517.
134. Nicklas, J.A. and E. Buel, *Quantification of DNA in forensic samples.* Analytical and Bioanalytical Chemistry, 2003. **376**(8): p. 1160-1167.
135. Walsh, P.S., J. Varlaro, and R. Reynolds, *A rapid chemiluminescent method for quanititation of human DNA.* Nucleic Acids Research, 1992. **20**(19): p. 5061-5065.
136. Waye, J. and H. Willard, *Structure, organization, and sequence of alpha satellite DNA from human chromosome 17: evidence for evolution by unequal crossing-over and an ancestral pentamer repeat shared with the human X chromosome.* Molecular and Cellular Biology, 1986. **6**(9): p. 3156-3165.
137. Budowle, B., et al., *Using a CCD camera imaging system as a recording device to quantify human DNA by slot blot hybridization.* BioTechniques, 2001. **30**(3): p. 680-685.
138. Ahn, S.J., J. Costa, and J.R. Emanuel, *PicoGreen Quantitation of DNA: Effective Evaluation of Samples Pre-or Psost-PCR.* Nucleic Acids Research, 1996. **24**(13): p. 2623-2625.

139. Singer, V.L., et al., *Characterization of PicoGreen reagent and development of a fluorescence-based solution assay for double-stranded DNA quantitation.* Analytical Biochemistry, 1997. **249**(2): p. 228-238.
140. Mandrekar, M.N., et al., *Development of a human DNA quantitation system.* Croatian Medical Journal, 2001. **42**(3): p. 336-339.
141. Zipper, H., et al., *Investigations on DNA intercalation and surface binding by SYBR Green I, its structure determination and methodological implications.* Nucleic Acids Research, 2004. **32**(12): p. e103-e103.
142. Higuchi, R., et al., *Kinetic PCR analysis: real-time monitoring of DNA amplification reactions.* Biotechnology, 1993. **11**: p. 1026-1030.
143. Dragan, A., et al., *SYBR Green I: fluorescence properties and interaction with DNA.* Journal of Fluorescence, 2012. **22**(4): p. 1189-1199.
144. Mao, F., W.-Y. Leung, and X. Xin, *Characterization of EvaGreen and the implication of its physicochemical properties for qPCR applications.* BMC Biotechnology, 2007. **7**(1): p. 1.
145. Alonso, A., et al., *Real-time PCR designs to estimate nuclear and mitochondrial DNA copy number in forensic and ancient DNA studies.* Forensic Science International, 2004. **139**(2-3): p. 141-9.
146. Holland, P.M., et al., *Detection of specific polymerase chain reaction product by utilizing the 5'-3'exonuclease activity of Thermus aquaticus DNA polymerase.* Proceedings of the National Academy of Sciences, 1991. **88**(16): p. 7276-7280.
147. Kim, Y., et al., *Crystal structure of Thermus aquaticus DNA polymerase.* 1995.
148. Green, R.L., et al., *Developmental validation of the AmpFℓSTR® NGM SElect™ PCR Amplification Kit: a next-generation STR multiplex with the SE33 locus.* Forensic Science International: Genetics, 2013. **7**(1): p. 41-51.
149. Coble, M.D. and J.M. Butler, *Characterization of new miniSTR loci to aid analysis of degraded DNA.* Journal of Forensic Sciences, 2005. **50**(1): p. 43-53.
150. Mulero, J.J., et al., *Development and Validation of the AmpFℓSTR® MiniFilerTM PCR Amplification Kit: A MiniSTR Multiplex for the Analysis of Degraded and/or PCR Inhibited DNA*.* Journal of Forensic Sciences, 2008. **53**(4): p. 838-852.

151. Parsons, T.J., et al., *Application of novel "mini-amplicon" STR multiplexes to high volume casework on degraded skeletal remains.* Forensic Science International: Genetics, 2007. **1**(2): p. 175-179.
152. Opel, K.L., et al., *The application of miniplex primer sets in the analysis of degraded DNA from human skeletal remains.* Journal of Forensic Science, 2006. **51**(2): p. 351-6.
153. Whitaker, J., E. Cotton, and P. Gill, *A comparison of the characteristics of profiles produced with the AMPFLSTR® SGM Plus™ multiplex system for both standard and low copy number (LCN) STR DNA analysis.* Forensic Science International, 2001. **123**(2): p. 215-223.
154. *Scientific Working Group on DNA Analysis Methods*. 2017; Available from: https://www.swgdam.org/.
155. Walsh, P.S., H.A. Erlich, and R. Higuchi, *Preferential PCR amplification of alleles: mechanisms and solutions.* Genome Research, 1992. **1**(4): p. 241-250.
156. Kimpton, C., et al., *Evaluation of an automated DNA profiling system employing multiplex amplification of four tetrameric STR loci.* International Journal of Legal Medicine, 1994. **106**(6): p. 302-311.
157. Taylor, D., *Using continuous DNA interpretation methods to revisit likelihood ratio behaviour.* Forensic Science International: Genetics, 2014. **11**: p. 144-153.
158. Ruitberg, C.M., D.J. Reeder, and J.M. Butler, *STRBase: a short tandem repeat DNA database for the human identity testing community.* Nucleic Acids Research, 2001. **29**(1): p. 320-322.
159. Bregu, J., et al., *Analytical thresholds and sensitivity: establishing RFU thresholds for forensic DNA analysis.* Journal of Forensic Sciences, 2013. **58**(1): p. 120-129.
160. Evett, I.W., P.D. Gill, and J.A. Lambert, *Taking account of peak areas when interpreting mixed DNA profiles.* Journal of Forensic Science, 1998. **43**(1): p. 62-69.
161. Taylor, D., J.-A. Bright, and J. Buckleton, *The interpretation of single source and mixed DNA profiles.* Forensic Science International: Genetics, 2013. **7**(5): p. 516-528.

162. Puch-Solis, R., et al., *Evaluating forensic DNA profiles using peak heights, allowing for multiple donors, allelic dropout and stutters.* Forensic Science International: Genetics, 2013. **7**(5): p. 555-563.
163. Benschop, C.C., et al., *Low template STR typing: effect of replicate number and consensus method on genotyping reliability and DNA database search results.* Forensic Science International: Genetics, 2011. **5**(4): p. 316-328.
164. Balding, D.J. and J. Buckleton, *Interpreting low template DNA profiles.* Forensic Science International: Genetics, 2009. **4**(1): p. 1-10.
165. Hauge, X. and M. Litt, *A study of the origin of 'shadow bands' seen when typing dinucleotide repeat polymorphisms by the PCR.* Human Molecular Genetics, 1993. **2**(4): p. 411-415.
166. Walsh, P.S., N.J. Fildes, and R. Reynolds, *Sequence analysis and characterization of stutter products at the tetranucleotide repeat locus vWA.* Nucleic Acids Research, 1996. **24**(14): p. 2807-2812.
167. Gill, P., R. Sparkes, and C. Kimpton, *Development of guidelines to designate alleles using an STR multiplex system.* Forensic Science International, 1997. **89**(3): p. 185-197.
168. Goodwin, W., A. Linacre, and S. Hadi, *An introduction to forensic genetics*. Vol. 2. 2011: John Wiley & Sons.
169. Bright, J.-A., et al., *Developing allelic and stutter peak height models for a continuous method of DNA interpretation.* Forensic Science International: Genetics, 2013. **7**(2): p. 296-304.
170. Puers, C., et al., *Identification of repeat sequence heterogeneity at the polymorphic short tandem repeat locus HUMTH01 [AATG] n and reassignment of alleles in population analysis by using a locus-specific allelic ladder.* American Journal of Human Genetics, 1993. **53**(4): p. 953.
171. Brookes, C., et al., *Characterising stutter in forensic STR multiplexes.* Forensic Science International: Genetics, 2012. **6**(1): p. 58-63.
172. Magnuson, V., et al., *Substrate nucleotide-determined non-templated addition of adenine by Taq DNA polymerase: implications for PCR-based genotyping and cloning.* BioTechniques, 1996. **21**(4): p. 700-709.

173. Clark, J.M., *Novel non-templated nucleotide addition reactions catalyzed by procaryotic and eucaryotic DNA polymerases.* Nucleic Acids Research, 1988. **16**(20): p. 9677-9686.
174. Kimpton, C.P., et al., *Automated DNA profiling employing multiplex amplification of short tandem repeat loci.* Genome Research, 1993. **3**(1): p. 13-22.
175. Masibay, A., et al., *Analysis and interpretation of short tandem repeat microvariants and three-banded allele patterns using multiple allele detection systems.* Journal of Forensic Science, 1999. **44**(1): p. 87-94.
176. Walsh, S., *Commentary on Kline MC, Jenkins B, Rogers S. Non-amplification of a vWA allele.* Journal of Forensic Science, 1998 **43**(5): p. 1103-1104.
177. Holt, C.L., et al., *Twgdam Validation of Ampflstr: PCR Amplification Kits for Forensic DNA Casework.* Journal of Forensic Science, 2002. **47**(1): p. 66-96.
178. *Forensic Bioinformatics*. Available from: http://www.bioforensics.com/dna-testing-issues.
179. Leclair, B., et al., *STR DNA typing: increased sensitivity and efficient sample consumption using reduced PCR reaction volumes.* Journal of Forensic Sciences, 2003. **48**(5): p. 1001-1013.
180. Smith, P.J. and J. Ballantyne, *Simplified Low-Copy-Number DNA Analysis by Post-PCR Purification.* Journal of Forensic Sciences, 2007. **52**(4): p. 820-829.
181. Westen, A.A., et al., *Higher capillary electrophoresis injection settings as an efficient approach to increase the sensitivity of STR typing.* Journal of Forensic Sciences, 2009. **54**(3): p. 591-598.
182. Ballantyne, K.N., R.A. van Oorschot, and R.J. Mitchell, *Comparison of two whole genome amplification methods for STR genotyping of LCN and degraded DNA samples.* Forensic Science International, 2007. **166**(1): p. 35-41.
183. Mercier, B., et al., *Direct PCR from whole blood, without DNA extraction.* Nucleic Acids Research, 1990. **18**(19): p. 5908.
184. Opel, K.L., D. Chung, and B.R. McCord, *A Study of PCR Inhibition Mechanisms Using Real Time PCR.* Journal of Forensic Sciences, 2010. **55**(1): p. 25-33.

185. Gray, K., D. Crowle, and P. Scott, *Direct amplification of casework bloodstains using the Promega PowerPlex® 21 PCR Amplification System.* Forensic Science International: Genetics, 2014. **12**: p. 86-92.
186. Wang, D.Y., et al., *Direct amplification of STRs from blood or buccal cell samples.* Forensic Science International: Genetics Supplement Series, 2009. **2**(1): p. 113-114.
187. Sorensen, A., et al., *Direct-to-PCR tissue preservation for DNA profiling.* International Journal of Legal Medicine, 2016. **130**(3): p. 607-613.
188. Blackie, R., D. Taylor, and A. Linacre, *DNA profiles from clothing fibers using direct PCR.* Forensic Science, Medicine, and Pathology, 2016. **12**(3): p. 331-335.
189. Lee, H.Y., et al., *Rapid direct PCR for ABO blood typing.* Journal of Forensic Sciences, 2011. **56**(s1).
190. Wang, D.Y., et al., *Development and Validation of the AmpFℓSTR® Identifiler® Direct PCR Amplification Kit: A Multiplex Assay for the Direct Amplification of Single-Source Samples.* Journal of Forensic Sciences, 2011. **56**(4): p. 835-845.
191. Verheij, S., J. Harteveld, and T. Sijen, *A protocol for direct and rapid multiplex PCR amplification on forensically relevant samples.* Forensic Science International: Genetics, 2012. **6**(2): p. 167-175.
192. Wang, Y., et al., *A novel strategy to engineer DNA polymerases for enhanced processivity and improved performance in vitro.* Nucleic Acids Research, 2004. **32**(3): p. 1197-1207.
193. Sim, J.E., et al., *High-Throughput STR Analysis for DNA Database Using Direct PCR.* Journal of Forensic Sciences, 2013. **58**(4): p. 989-992.
194. Gausterer, C., C. Stein, and T. Stimpfl, *Application of direct PCR in a forensic case of yew poisoning.* International Journal of Legal Medicine, 2012. **126**(2): p. 315-319.
195. Belotserkovskii, B.P. and B.H. Johnston, *Denaturation and association of DNA sequences by certain polypropylene surfaces.* Analytical Biochemistry, 1997. **251**(2): p. 251-262.

196. Belotserkovskii, B.P. and B.H. Johnston, *Polypropylene tube surfaces may induce denaturation and multimerization of DNA.* Science, 1996. **271**(5246): p. 222-223.
197. Scientific, T. *GeneMapper ID-X Software*. 2017; Available from: https://www.thermofisher.com/au/en/home/industrial/human identification/genemapper-id-x-software.html.
198. *OSIRIS software, NCBI*. 2017; Available from: https://www.ncbi.nlm.nih.gov/projects/SNP/osiris/.
199. *FSS-i3™ Expert Systems*. 2005; Available from: https://www.promega.com/-/media/files/resources/profiles-in-dna/802/fssi3-expert-systems.pdf.
200. *TrueAllele software*. 2017; Available from: https://www.cybgen.com/products/casework.html.
201. Perlin, M.W. and A. Sinelnikov, *An information gap in DNA evidence interpretation.* PLOS ONE, 2009. **4**(12): p. e8327.
202. Perlin, M.W., J.B. Kadane, and R.W. Cotton, *Match likelihood ratio for uncertain genotypes.* Law, Probability and Risk, 2009: p. mgp024.
203. *Frye v. United States,* 1923; Available from: https://www.law.ufl.edu/_pdf/faculty/little/topic8.pdf.
204. *Admissibility of Expert Opinion Evidence*. 1984; Available from: http://www.aic.gov.au/media_library/conferences/medicine/sides.pdf.
205. *PCAST forensic science report*. 2016; Report to the President, Forensic Science in Criminal Courts: Ensuring Scientific Validity of Feature-Comparison Methods. Available from: https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/PCAST/pcast_forensic_science_report_final.pdf.
206. *DNA Advisory Board Quality Assurance Standards for Forensic DNA Testing Laboratories*. 2017; Available from: http://www.cstl.nist.gov/strbase/dabqas.htm.
207. *National Institute of Standards and Technology*. 2017; Available from: https://www.nist.gov/topics/forensic-science/reference-materials-standards-and-guidelines.
208. *The European Network of Forensic Science Institutes*. 1995; Available from: http://enfsi.eu/.

209. *The European DNA Profiling Group*. 1988; Available from: https://www.isfg.org/EDNAP.
210. *National Association of Testing Authorities*. 2017; Available from: https://www.nata.com.au/nata/.

Chapter II

Materials and Methods

2.1 Preface

This chapter includes a background section on the standard practice for forensic evidence collection, sample handling and processing, and describes the materials and methods section for the thesis. More specifically, the workflow involves contamination controls, sample handling, the swabbing process, direct PCR, and the protocol for capillary electrophoresis that was used to generate profiling data from 'touch' DNA swabs. The chapter concludes with recommended standard operating procedures (SOPs) that have been structured with guidance from FSSA Laboratory. All forensic DNA testing laboratories are required by 'The Quality Assurance Standards for Forensic DNA Testing Laboratories' to follow accredited protocols that have been validated [1]. Therefore, protocols have been established that will allow rapid implementation of the direct PCR approach into accredited facilities upon validation.

2.1.1 Standard practice

In order to perform STR-based DNA analysis of a biological sample, initial steps are performed that include sample collection, preservation, and adequate storage of material to minimise DNA degradation. It is vital that samples are handled properly at all stages in the workflow and follow a strict chain of custody so that DNA profiling results are legally admissible in a court of law.

Sample collection

Proper controls and checks need to be in place to prevent sample contamination or mix-up during evidence collection. If a sample becomes contaminated with extraneous DNA during early stages of handling (or at any stage) then results would be compromised. 'DNA-free' plasticware and consumables should be used to prevent the 'phantom of Heilbronn' incident [2] from happening again. In Germany and Austria, 40 cases were incorrectly linked, due to a 'sterile' swab contamination issue in the manufacturing plant. ENFSI (European Network of Forensic Science Institutes), BSAG (Biology Specialist Advisory Group) and SWGDAM (Scientific Working Group on DNA Analysis Methods) have released recommendations to the scientific community to ensure all plasticware and consumables are not only sterile

but also ‘DNA-free’ and have their own standard that is ISO 18385 [3-6]. Contamination can be introduced in the laboratory if adequate care is not taken [7-10]. A study carried out to investigate incidental DNA on sterile gloves revealed the presence of DNA on the surface of gloves taken from closed gloves boxes, emphasizing the need for ‘DNA-free’ gloves and consumables [11].

Specifics important to sample collection

The risk of contamination can be minimised with the aid of proper protective clothing, sample controls, elimination databases, strict protocols, separate work areas (see Figure 2.1) for each stage of sample handling, and the use of ‘DNA-free’, sterile consumables [12-14].

Clothing

Gloves need to be worn when evidence is collected and a separate pair of gloves worn for each handled item. Gloves should not be packaged in with evidence [15]. Facemasks, hair nets, and clothing protection, should be worn to prevent contamination (see Figures 2.2 (a) and (b)).

Packaging

Evidence should be packaged separately and exhibits (particularly those stained with biological fluids) should be air-dried prior to sealing the packaging. Evidence should be stored in paper bags that are porous to prevent condensation, as water/moisture accelerates DNA degradation. Swabs or biological stains should be air-dried prior to packaging, stored separately, and labelled correctly.

Chain of custody

Case details relating to each item should be detailed on the packaging and include: case/reference number, date and time of collection, location of evidence, and the name of the person collecting the evidence. Packages should be sealed across the package seal to ensure that samples are not tampered with.

Reference material

Reference material (e.g. buccal swabs) should be collected from the relevant individuals and the DNA isolated and compared to evidence material. Reference swabs should be stored and analysed in a separate location away from other evidence.

Storage and preservation

To minimise DNA degradation, samples should preferrably be stored in a cold, dry environment, for preservation. Upon delivery to the laboratory, DNA samples are stored as non-extracted evidence or extracted DNA. Samples should be placed at 4 °C, – 20 °C or – 70 °C depending on the required length of time in storage. Swabs at FSSA and SAPOL are currently stored at room temperature.

2.1.2 Materials and methods for thesis

2.1.3 Contamination controls for low-template DNA

To minimize contamination of sample DNA with contemporary human DNA or previously amplified PCR product, specific controls were set up for all experiments. Pre-PCR steps for swabbing took place in a dedicated fume-hood that was isolated from post-PCR activity (see Figure 2.1). Strict contamination protocols were followed, such as cleaning work benches and consumables with 70% isopropanol and 3% sodium hypochlorite before use; one-way movement of personnel (sample preparation hood > PCR laboratory > post-PCR laboratory) took place (see Figure 2.1). Additional sterilisation of gloves took place; 3% sodium hypochlorite was used to clean the surface of gloves prior to use. Gloves were changed between sampling. Each laboratory and work area had its own set of pipettes, barrier filter tips, tubes and racks. Protective clothing consisted of separate lab coats designated for each lab (and laundered every few months), double gloves, facemasks, and hair tied back to prevent skin cells from contaminating the samples (see Figure 2.2 (a) and 2.2 (b)). For research requirements it is unsure how effective laundering lab coats really is. However, there were no instances of contamination noted. In spite of this, casework laboratories would require disposable coats for one use only. PCR setup was conducted in a laminar flow hood. Aersol-resistant filter tips were used in the PCR

setup up lab and a new tip was used for each sample. Reagents were prepared with care and diluted with sterile molecular grade water. Non-template controls (PCR blanks) and extraction blanks (i.e. sterile swabs containing no DNA) were used in all PCR set up reactions to monitor contamination.

Prior to use, substrates on to which DNA was deposited were cleaned with 70% isopropanol and Milli-Q ultrapure water (Merck Millipore, VIC, AU), and air-dryed in a sterile fume hood. Substrates were further sterilised by UV light (at a wavelength of approximately 300 nm) where the substrate is situated close to the UV lamp for 15 minutes (with the exception of Chapter V). The effectiveness of the UV lamp was not tested, however, controls used in testing did not indicate any issues with contamination and the UV light could be seen to be working. Slides were swabbed prior to DNA deposition as an additional negative control. Slides were stored at 4°C post fingermark deposition and swabbing.

Using an elimination database, the STR-based DNA profiles of all staff involved in handling of samples (J.T and A.L), and all lab personnel that had access to the laboratories, were genotyped and profiles were compared to those obtained from samples. Reference buccal swabs provided by the donors were isolated in a separate DNA extraction laboratory and processed after fingermark swabs had been taken.

Positive controls, containing 1 ng/µL of control 2800M DNA, (Promega, US) were included for each PCR set up to ensure that all PCR components were working effectively. Care was taken in all aspects of sample handling. PCR was carried out in a designated PCR-set up laboratory (see Figure 2.1) as one small drop of PCR product may contain as many as a billion copies of DNA [12].

2.1.3.1 DNA laboratory setup

The DNA Linacre lab 330 setup, Physical Sciences Building, Flinders University

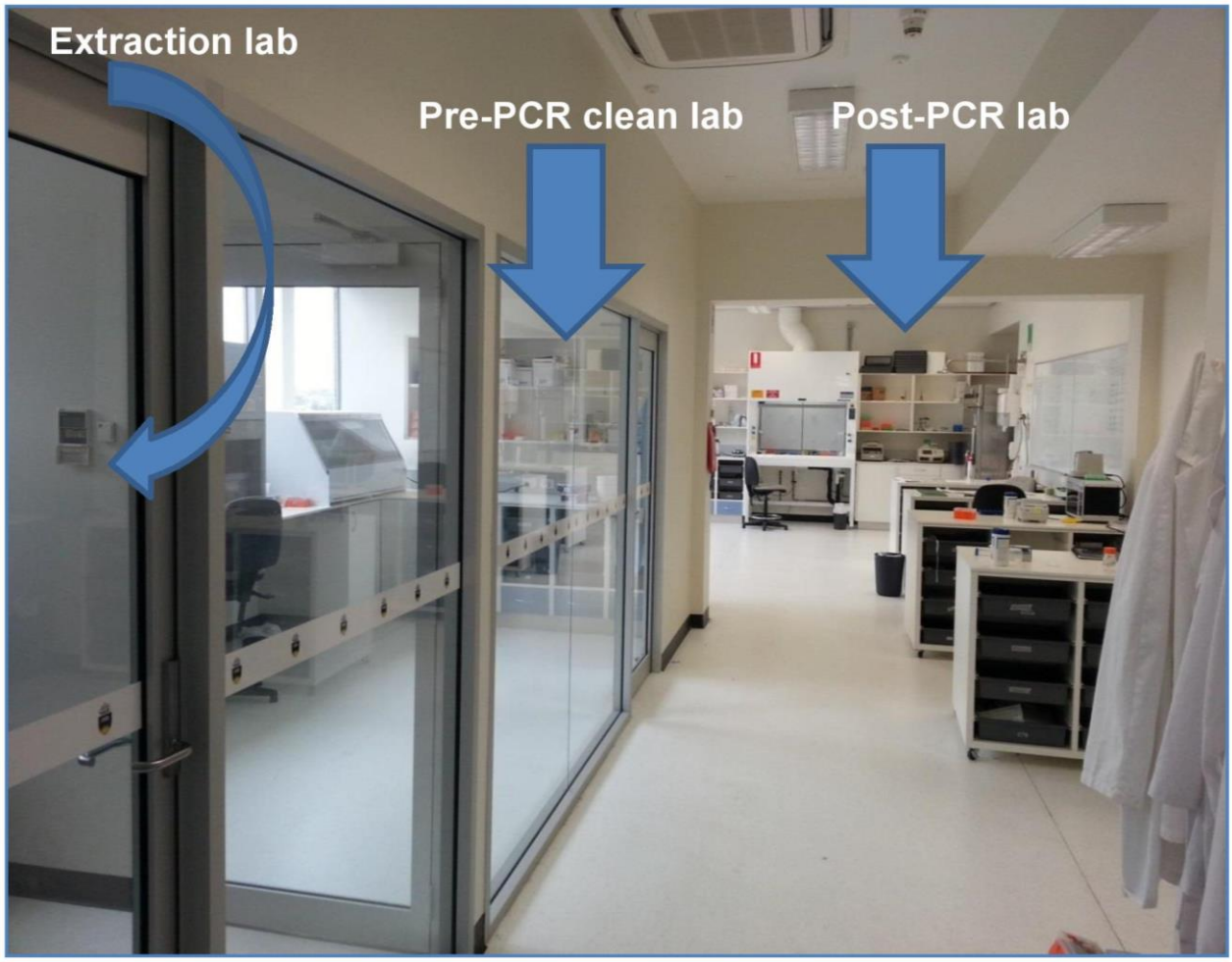


Figure 2.1 Three DNA testing laboratories at Flinders University are physically isolated from each other. Pre and post-PCR laboratory work is restricted to designated areas only.

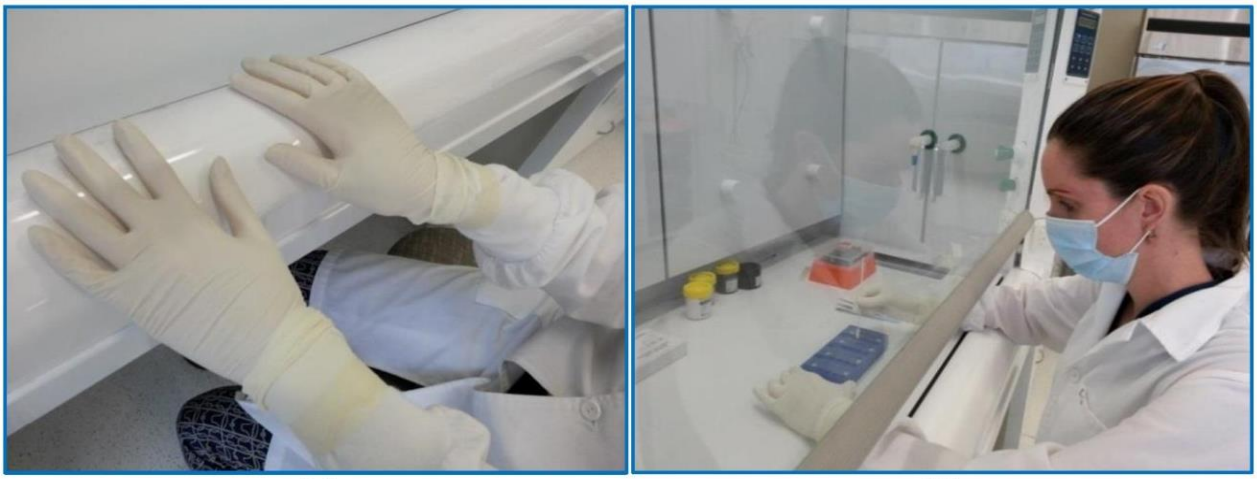


Figure 2.2 personal protective clothing; (a) double gloving; (b) facemask, lab coat and gloves worn at all times.

Ethics

Ethical approval was obtained from Southern Adelaide Clinical Human Research Ethics Committee prior to starting lab work. Ethics approval number: 461.13. Informed consent was provided by all the volunteers in the study prior to starting.

DNA reference material

Buccal swabs (Cotton, Copan, Brescia, Italy) were rubbed inside the inner cheeks of volunteers and used as reference material.

DNA extraction

Cells were lysed using the DNA IQ™ extraction Kit (Promega, US) following manufacturer's recommendations.

Sample preparation

Prior to starting, Triton-X™ 100 (Sigma-Aldrich, Castle Hill, NSW, Sydney) was diluted to 0.1% using Milli-Q ultrapure water (Merck Millipore, VIC, AU) and 2 – 20 µL of Triton-X™ 100 was dispensed onto the tip of a swab prior to swabbing. The amount of swab media varied per experimental set up (see data chapters for specific details). A larger volume of swab media is recommended for swabbing large surface areas and this was used in early experimental design with 'whole swabbing'. Conversely, a 'targeted approach' to swabbing, using pre-cut fibre clumps, required

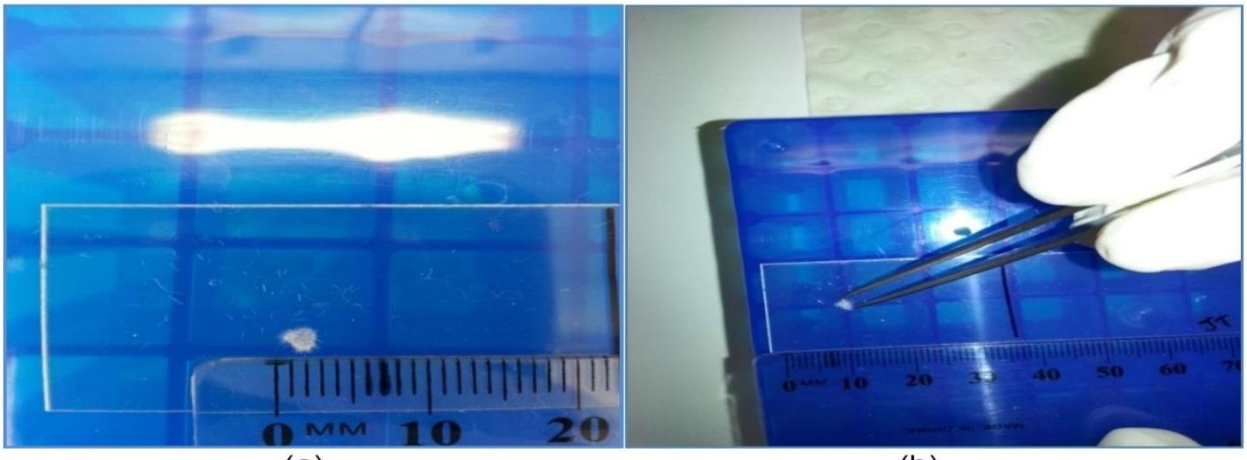
a smaller volume of swab media. Control DNA (2800M, Promega, US, at 10 ng/µL was diluted to varying concentrations (depending on the experimental design) using Milli-Q ultrapure water (Merck Millipore, VIC, AU) or TE buffer (i.e. 10 mM Tris, 0.1 mM EDTA, pH 8.0). It is recommended that this swabbing protocol is carried out in the laboratory only until a field device becomes available for direct PCR use.

Swabbing

Swabbing consisted of 8-20 strokes horizontally (left-to-right direction) and 8-20 strokes vertically (top-to-bottom) over the substrate using medium pressure to ensure consistency. The number of strokes used varied depending on the experimental design and substrate that was swabbed (see data chapters for specific details).

Targeted swabbing

Swab fibre clumps (~ 2 mm^2) were cut from the tips of pre-moistened DNA-free Ethylene Oxide (EO) treated sterile swabs: a) nylon FLOQswabs™ (Copan, Brescia, Italy), b) foam swabs (Whatman, USA), and c) cotton swabs (Livingstone, NSW). Swabs were further sterilised under UV light for 15 minutes. Sterile forceps were used to hold the fibre clumps in place over the slide (see Figure 2.3). Glue fragments from the inner core of the nylon swab held flocked fibres neatly in place and moisture from the swabbing media prevented fibres from becoming static and dislodging from the forceps.



(a) (b)

Figure 2.3 (a) fibre clump used for swabbing (~ 2mm^2) (b) sterile forceps used for swabbing to target a small surface area where DNA had been deposited.

Polymerase Chain Reaction

The specifics for each experiment are detailed in each data chapter.

STR typing for all samples was performed using the AmpFℓSTR® ProfilerPlus™ kit (10 genetic STR loci typed, including Amelogenin; Applied Biosystems, AU) or the NGMSElect kit™ (16 genetic STR loci typed, plus Amelogenin; Applied Biosystems, AU) following manufacturer's recommendations for sub-optimal DNA (i.e. < 1 ng). For ProfilerPlus™, the final reaction consisted of 21 µL ProfilerPlus™ reaction mix, 11 µL of ProfilerPlus™ primer mix, 1 µL (5U) Ampli*Taq* Gold™ (Applied Biosystems), DNA template (2 – 4 mm^2 fibres or 1 ng of extracted/control DNA) and the remainder of the solution was made up to 50 µL with TE buffer (i.e. 10 mM Tris, 0.1 mM EDTA, pH 8.0). Half volume PCRs were carried out as indicated in specific data chapters by halving all the reaction components but keeping the amount of enzyme consistent (i.e. 1 µL (5U) Ampli*Taq* Gold™ DNA polymerase). Additional work highlights the importance of extra Ampli*Taq* Gold™ DNA polymerase for processing low-level DNA samples [16]. For ProfilerPlus™, cycling was performed on a 9700 GeneAmp thermal cycler (Applied Biosystems) and consisted of an initial denaturation at 95 ° C for 11 min followed by 28 or 29 cycles of 94 ° C for 1 min, 59 ° C for 1 min and 72 ° C for 1 min, then a final extension at 60 ° C for 45 min, and 25 °C hold. Additional PCR facilitators were added in later experiments (see chapter for specific details).

For PCR with the AmpFℓSTR® NGM SElect™ kit (Life Technologies), the reaction set up followed manufacturer's recommendations with additional Ampli*Taq* Gold™ DNA polymerase added. The final PCR reaction volume consisted of 10 µL NGM SElect™ reaction mix, 5 µL of NGM™ primer mix, 1 µL (5U) Ampli*Taq* Gold™ DNA polymerase (Life Technologies) and DNA template (either 2 – 4 mm^2 swabbed fibres for fingermark samples or 1 ng of DNA extract/control sample), the remainder of the solution was made up to a final volume of 25 µL with TE buffer (i.e. 10 mM Tris, 0.1 mM EDTA, pH 8.0). Cycling was performed on a 9700 GeneAmp thermal cycler (Life Technologies) and consisted of an initial denaturation at 95 ° C for 11 min followed by 30 cycles of 94 ° C for 20 s, 59 ° C for 3 min, then a final extension at 60 ° C for 10 min, and 4 °C hold. PCR products were spun in a microcentrifuge for 20 seconds at maximum speed to centrifuge the powders to the bottom of the tube.

Capillary Electrophoresis

For the AmpFℓSTR® NGM SElect™ kit, amplicons were analysed on a 3130*xl* Genetic Analyser (Applied Biosystems). The final reaction volume (11-11.5 µL) consisted of 1-1.5 µL PCR product, 9.5 µL HiDi™ Formamide (Life Technologies) and 0.5 µL GeneScan™ 600 LIZ™ Size Standard v2.0 (Life Technologies).

For ProfilerPlus™ kit, amplicons were analysed on a 3130*xl* Genetic Analyser (Applied Biosystems) in a final set up volume of 14 µL that consisted of between 1-2 µL PCR product (as recommended by the ProfilerPlus™ kit manufacturer), 12 µL HiDi™ Formamide (Applied Biosystems) and either 0.5 µL GeneScan-400 [ROX] size standard or GeneScan-500 [ROX] size standard. Profiles were analysed using GeneMapper ID software (v3.2.1) (Applied Biosystems), and a threshold value of 30-50 relative fluorescence units (RFU) was used to assign alleles (specific threshold values for each experiment are provided in the chapter specific methods section).

Data Analysis

Internal validation studies were used to assess analytical threshold values. Allelic 'drop-out' was reported if associated peak heights were less than 50 RFU. A threshold was set at 150 RFU for potential homozygote genotypes; if only one allele was detected at a locus then this threshold value was used to account for potential 'allele drop-out'. Stutter was considered for peak heights less than 15% of the main true STR allele. A profile was considered 'full' if all alleles, for all loci, were detected above the threshold RFU. An informative DNA profile (i.e. one considered 'up-loadable' to the Australian National Criminal Investigation DNA Database (NCIDD)) was defined as ≥ 12 alleles (plus Amelogenin).

Profiling success (%) was assessed by dividing the number of detected alleles by the total number of expected alleles and multiplying this value by 100. The average peak height for each profile was calculated by dividing the summed peak heights of all alleles by the total number of peaks observed in that sample.

2.2 Standard operating procedure (SOP) for recovering 'touch' DNA from casework exhibits

This protocol has been designed for the purpose of targeting small intricate surface areas for latent DNA or to retrieve DNA from fingerprints that have been dusted with fingerprint powders prior to DNA analysis. A targeted swabbing method can be used which involves moistening the tip of nylon flocked swabs with 2-20 μL of Triton-X™ (0.1%), pre-heated to 50 °C, cutting moistened swab material (2 mm^2) using a sterile scapel blade, and using the swab fibre clump and sterile forceps to rub over the surface thought to contain DNA. A swab designed for the purpose of direct PCR would prove more beneficial. See protocol described below for using a whole swab approach for direct PCR. Further testing would be required to determine if heating the detergent is necessary in order to improve DNA yield, as this may not be practical in the field.

Template sample handling sheet

Case ID:
Sample ID:
Date and time:
Name of operator:
Signature of operator:
Laboratory location:
Kit name:
Expiry date:
Lot number:

1. Preface and Scope

Wet and dry nylon flocked swabs (Copan or similar) are used to collect DNA from crime scene items.

2. Definitions

'Touch' DNA = DNA left on an object by touching i.e. no obvious body fluid/tissue stained areas are observed.

3. Principle

Skin (and other) cells are lysed in the presence of 0.1% Triton-X™ and heated to isolate the DNA.

4. Reagents

4.1.1 Triton-X™ (0.1%)

4.1.2 Bleach; sodium hypochlorite (3%), (in squeeze bottles, for cleaning)

4.1.3 MilliQ water or sterile (DNA-free, RNA-free distilled water – molecular grade)

4.1.4 Kleralcohol wipes (70% Isopropanol, for cleaning)

5. Equipment

5.1.1 Biohazard hood

5.1.2 Sterile disposable forceps

5.1.3 P10, P200 and P1000 pipettes

5.1.4 Heating block set at 50 ^{0}C

5.1.5 Thermometer calibrated to 50 ^{0}C

5.1.6 Racks to hold 0.2 mL tubes

6. Consumables

6.1.1 0.2 mL sterile low-bind DNA tubes (Eppendorf or similar)

6.1.2 Filtered tips for P10, P200 and P1000 pipettes

6.1.3 Nylon flocked swabs (Copan or similar)

6.1.4 Sterile scalpel blades

6.1.5 Kimwipes

6.1.6 Dental bibs

6.1.7 Disposable gloves

6.1.8 Disposable face masks

6.1.9 Plastic bags for waste and bin

6.1.10 Clean separate lab coat designated to each lab (pre and post-PCR)

7. Procedure

7.1.1 Quality Control:

i. Do not place samples from the same case number adjacent to each other;

ii. Include a reagent blank (RB);

iii. Thoroughly clean forceps, equipment and work areas with 3% sodium hypochlorite and 70% Ispropanol wipes prior to starting work and between sample preparation;

iv. Wear a fresh pair of gloves for handling each sample.

7.1.2 Prepare a sample handling sheet for each case

7.1.3 Pre-heat an aliquot of Triton-X™ (0.1%) to 50 ^{0}C in a heat block in the pre-PCR extraction lab

7.1.4 Retrieve samples from 4°C and place on the bench in the pre-PCR extraction lab

7.1.5 Ensure the case number and item codes match form details

7.1.6 Prepare the hood by exposing to UV light for 15 minutes prior to starting

7.1.7 In an extraction rack in the hood, label the appropriate number of sterile 0.2 mL tubes with the case number and sample description on the lid

7.1.8 Witness step:

i. Witness checks that the sample case number and item code match the sample handling sheet;

ii. Witness passes each sample/item to be swabbed to the

extractor in the order of the sample handling sheet, verbalizing sample details and which tube number the sample is to be placed into;

iii. Witness continues with all the samples in the batch and;

iv. Witness signs the sample handling sheet

7.1.9 Moisten the tip of a nylon flocked swab with 2 μL of Triton-X™ (0.1%), pre-heated to 50 ^{0}C.

7.1.10 Hold the swab vertically over the item, so that only the tip of the swab is in contact with the item. Swab 10 strokes vertically (top-to-bottom direction) and 10 strokes horizontally (left-to-right direction) over the substrate.

7.1.11 Cut fibres from the tip of the swab (~ 2 mm^2) using a sterile scalpel blade and using sterile forceps place fibres directly into a 0.2 mL PCR tube (pre-labelled with case details).

7.1.12 Repeat step 7.10 – 7.11 with a second dry swab and add fibres from the second swabbing event into the same pre-labelled PCR tube or a separate PCR tube if a backup sample is required. Double swabbing was shown to be more effective than single swabbing.

7.1.13 Seal PCR tubes and send tubes to the pre-PCR set up lab for processing. Clean work areas (as detailed previously), change gloves and proceed to next sample.

8. Validation

Reference for validation: “DNA profiles from fingermarks”, J E.L Templeton and Adrian Linacre. Biotechniques 57: 259-266 (November 2014) doi 10.2144/000114227.

2.3 Standard operating procedure (SOP) for processing ‘touch’ DNA swabs

Template sample handling sheet

Case ID:
Sample ID:
Date and time:
Name of operator:
Signature of operator:
Laboratory location:
Kit name:
Expiry date:
Lot number:

1. Preface and Scope

Wet and dry nylon flocked swabs (Copan or similar) can be used to collect DNA from evidence items. All swabs arriving into the lab should be intact and sealed prior to processing. Swabs have been pre-moistened with Triton-X™ prior to swabbing and can be left to air-dry or stored in the freezer post-swabbing. Additional testing would be required to determine the most effective storage conditions, as both methods have been found to be sufficient.

2. Definitions

Touch DNA = DNA left on an object by touching i.e. no obvious body fluid/tissue stained areas are observed.

3. Principle

Skin (and other) cells are lysed in the presence of 0.1% Triton-X™ and heat.

4. Reagents

MilliQ water or sterile (DNA-free, RNA-free distilled water – molecular grade)

Sodium hypochlorite (3%), (in squeeze bottles, for cleaning)

Kleralcohol wipes (70% Isopropanol, for cleaning)

5. Equipment

5.1 Biohazard hood

5.2 Sterile disposable forceps

5.3 Racks to hold 0.2 mL tubes

6. Consumables

6.1 0.2 mL sterile low-bind DNA tubes (Eppendorf or similar)

6.2 Sterile scalpel blades

6.3 Kimwipes

6.4 Dental bibs

6.5 Disposable gloves

6.6 Disposable face masks

6.7 Plastic bags for waste and bin

6.8 Clean separate lab coat designated to each lab (pre and post-PCR)

7. Procedure

7.1 Quality Control:

i. Swabs should arrive into the lab intact and sealed following the chain of custody procedure;

ii. Do not place samples from the same case number adjacent to each other;

iii. Include a reagent blank (RB);

iv. Thoroughly clean forceps, equipment and work areas with 3% sodium hypochlorite and 70% Isopropanol wipes prior to starting work and between sample preparation and;

v. Wear a fresh pair of gloves for each sample.

7.2 Prepare a sample handling sheet for each case

7.3 Retrieve swabs from the 4 ^{0}C fridge, and place on the bench in the pre-PCR extraction lab

7.4 Ensure the case number and item codes match form details

7.5 Prepare the hood by exposing to UV light for 15 minutes prior to starting

7.6 In an extraction rack in the hood, label the appropriate number of sterile 0.2 mL tubes with the case number and sample description on the lid.

i. Witness step: Witness checks that the sample case number and item code match the sample handling sheet;

ii. Witness passes each sample/item to be swabbed to the

extractor in the order of the sample handling sheet, verbalizing sample details and which tube number the sample is to be placed into;

iii. Witness continues with all the samples in the batch and;

iv. Witness signs the sample handling sheet.

7.7 Cut fibres from the tip of the swab (~ 2 mm^2) using a sterile scalpel blade and using sterile forceps place fibres directly into a 0.2 mL PCR tube (pre-labelled with case details). Repeat if a second swab has been supplied for the sample and add fibres from the second swab into the same pre-labelled PCR tube or a separate PCR tube if a backup sample is required.

7.8 Seal PCR tubes and send to the pre-PCR set up lab for processing. Clean work areas (as detailed previously), change gloves and proceed to next sample.

8. Validation

Reference: Templeton, J. E., and Adrian Linacre. "DNA profiles from fingermarks." BioTechniques 57.5 (2014): 259.

2.4 References

1. *National Association of Testing Authorities*. 2017; Available from: https://www.nata.com.au/nata/.
2. *Time Magazine Online, Germany's Phantom Serial Killer: A DNA Blunder*. 2009; Available from: www.time.com/time/world/article/0,8599,1888125,00.html.
3. Gill, P., et al., *Manufacturer contamination of disposable plastic-ware and other reagents—an agreed position statement by ENFSI, SWGDAM and BSAG.* Forensic Science International: Genetics, 2010. **4**(4): p. 269-270.
4. SWGDAM. *SWGDAM Guidelines for STR Enhanced Detection Methods*. 2014 November 19, 2014; Available from: http://swgdam.org/SWGDAM%20Guidelines%20for%20STR%20Enhanced%20Detection%20Methods%20FINAL%20100614.pdf.
5. *Scientific Working Group on DNA Analysis Methods*. 2017; Available from: https://www.swgdam.org/.
6. *ISO 18385:2016*. 2016; Available from: https://www.iso.org/standard/62341.html.
7. Schmidt, T., S. Hummel, and B. Herrmann, *Evidence of contamination in PCR laboratory disposables.* Naturwissenschaften, 1995. **82**(9): p. 423-431.
8. Rutty, G., A. Hopwood, and V. Tucker, *The effectiveness of protective clothing in the reduction of potential DNA contamination of the scene of crime.* International Journal of Legal Medicine, 2003. **117**(3): p. 170-174.
9. Raymond, J.J., et al., *Trace DNA analysis: Do you know what your neighbour is doing?* Forensic Science International: Genetics, 2008. **2**(1): p. 19-28.
10. Szkuta, B., et al., *DNA transfer by examination tools–a risk for forensic casework?* Forensic Science International: Genetics, 2015. **16**: p. 246-254.
11. Daniel, R. and R.A. van Oorschot, *An investigation of the presence of DNA on unused laboratory gloves.* Forensic Science International: Genetics Supplement Series, 2011. **3**(1): p. e45-e46.
12. Butler, J.M., *Forensic DNA Typing: Biology, Technology, and Genetics of STR Markers*. 2nd Edition ed. 2005, New York: Elsevier Academic Press.

13. Howitt, T., *Ensuring the integrity of results: a continuing challenge in forensic DNA analysis.* Proceedings of the 14th International Symposium on Human Identification, Promega Genetic Identity Conference, 2003.
14. Gill, P. and A. Kirkham, *Development of a simulation model to assess the impact of contamination in casework using STRs.* Journal of Forensic Sciences, 2004. **49**(3): p. 485-491.
15. Goray, M., R.A. van Oorschot, and J.R. Mitchell, *DNA transfer within forensic exhibit packaging: Potential for DNA loss and relocation.* Forensic Science International: Genetics, 2012. **6**(2): p. 158-166.
16. McNevin, D., et al., *Reduced reaction volumes and increased Taq DNA polymerase concentration improve STR profiling outcomes from a real-world low template DNA source: telogen hairs.* Forensic Science, Medicine, and Pathology, 2015. **11**(3): p. 326-338.

Chapter III
Improving the collection of ‘touch’ DNA

Publications included in Chapter III

1. **Templeton, J.,** Ottens, R., Paradiso, V., Handt, O., Taylor, D., & Linacre, A. (2013). *Genetic profiling from challenging samples: direct PCR of 'touch' DNA.* Forensic Science International: Genetics, Supplement Series, 4(1), e224-e225.
2. **Templeton, J.,** Taylor, D., Handt, O., Skuza, P., & Linacre, A. (2015). *Direct PCR Improves the Recovery of DNA from Various Substrates*. Journal of Forensic Sciences, 60(6), 1558-1562.
3. Ottens, R., **Templeton, J.,** Paradiso, V., Taylor, D., Abarno, D., & Linacre, A. (2013). *Application of direct PCR in forensic casework*. Forensic Science International: Genetics, Supplement Series, 4(1), e47-e48.

Appendix

a) Supplementary Information – Results – **Experiment 3.1)** Comparison of swab types for direct PCR.

b) Supplementary Information – Results – **Experiment 3.2)** Comparison of swab media and swab technique.

c) Supplementary Information – Results – **Experiment 3.3)** Direct PCR vs Extraction.

d) Poster: **Templeton. J,** Renée Blackie, O Handt, D Taylor and A Linacre. *Direct PCR improves the recovery of DNA from various substrates*. Presented at the 26th "International Society for Forensic Genetics" congress, Melbourne, Australia, 2013.

e) Provisional patent information - Invention Title: NUCLEIC ACID COLLECTION DEVICE, 2016 [1]. Image of microswab included.

f) Oral presentation: **Templeton. J.** Forensic Science South Australia (FSSA), 2013, *Improving methods for the recovery and analysis of 'touch' DNA.*

3.1 Preface

The ability to acquire genetic information from 'touch' DNA aids criminal investigations and is now an integral part of a forensic laboratory's workflow. In particular, collection is a critical stage in the recovery of DNA from handled or touched items. It is now possible to generate interpretable DNA profiles from picogram amounts of DNA [2]. This is important for 'touch' DNA samples that may yield limited DNA as there is the possibility to generate DNA profiles from sub-optimal samples. The vast improvement in the sensitivity of STR profiling kits is partly due to the enhanced buffers and enzymes present in the multiplex systems [3-5]. Despite this advancement, DNA recovered from touched substrates is often close to the limit of sensitivity for most STR-based DNA profiling kits [6, 7]. Part of the reason for this could be the direct result of poor sampling efficiency and/or the low extraction efficiency of the swabs that may contribute to a loss of sample DNA [8, 9]. Direct PCR has the potential to improve DNA recovery, but a consequence of this is the detection of complex mixtures (i.e. additional alleles) which require resolving due to the increased sensitivity of this approach.

This data chapter is split into three phases, each of which examines aspects relating to DNA recovery when swabs are used as the collection device. Raw data are presented in the Appendix - Supplementary Information (SI) section at the end of the chapter.

1) The first phase of the study compares the sampling efficiency of three commonly used swab types (i.e. wound, pad and flocked swabs) when direct PCR is used, and explores the use of detergent-based swabbing media and swabbing technique. The main result for direct PCR showed that fibres present in the PCR tube, with 1 ng of DNA, did not affect the ability to generate a DNA profile and results were compared to a positive control (containing no fibres).

2) Detergent-based nylon FLOQSwabs™ recovered a higher yield of DNA (as assessed by RFU values of peak heights) compared to foam and cotton water-based collection devices. Preliminary data are reported in supplementary information (SI) Tables 3.1 and 3.2, with results extending into

two short peer reviewed manuscripts published in the journal Forensic Science International: Genetics, Supplement Series.

3) The second phase of the study explores the efficacy of direct PCR compared to a standard DNA extraction method (i.e. DNA IQ™ System) when using control DNA of known mass as the template for amplification. The DNA IQ™ System included the use of spin baskets as recommended in the protocol for swab substrates. Various substrates on which DNA was deposited were examined (i.e. brass, plastic and glass). The nature of substrates is explored. Results are reported in a peer reviewed publication in the Journal of Forensic Sciences.

4) The third, and final, phase of the work describes the application for a patent for a novel swab device designed as a direct result of this work.

3.2 Aims of study

1) To compare a) the efficacy of swab types for recovering DNA; b) to assess the effectiveness of detergent-based swabbing media in comparison to water-based swabs and; c) to determine the most effective method of swabbing for direct PCR.

2) To determine if a meaningful DNA profile could be obtained using direct PCR when swab fibres are present in the PCR tube.

3) To assess the profiling results generated by direct PCR in comparison to those obtained by a standard DNA extraction method.

3.3 Introduction

Swabs

DNA collection using swabs is currently one of the most widely used techniques for collecting biological evidence [10]. Recent reports [10-13] are suggesting that an alternative fibre material to cotton should be considered for processing 'touch' DNA samples. A comprehensive study by Verdon *et al.* (2014) [10] compared swab types of different compositions for the efficient recovery of DNA from various materials. It was reported that swabs with a high sampling efficiency do not necessarily have high extraction efficiency [10]. A significant contributing factor is the elution capability of the swab to release DNA which has been collected. A previous study [14] reports that the success of DNA profiling from swabs is dependent on the extraction method used. However, the overall DNA yield will be reduced if a swab is not efficient at releasing DNA during extraction [14]. Research conducted by Prinz *et al.* (2006) examined 109 touched items (including 30 deposited fingermarks) [15] and found that most of the collected samples contained less than 100 pg of DNA, post extraction. Most commercial STR DNA profiling kits recommend optimal starting templates between 0.5 – 2.5 ng of DNA [3-5, 16-18], and would not be suitable for analysing highly degraded or low-template DNA samples. However, laboratories also perform internal validation studies to determine their own optimal starting DNA amounts and this may be lower than manufacturer guidelines.

Selection of an appropriate sampling device - suitable for the nature of the substrate that is swabbed - would have great impact upon recovery, extraction, and profiling of DNA [10, 11, 19, 20]. The ideal collection device should be: inexpensive; easy to transport; sterile; DNA and RNA-free; and have a wide-range application for use on many different substrates.

There are many characteristics of a swab that may influence the ability to collect and preserve DNA from a substrate, such as:

1) tip composition and design;
2) thickness and length of the swab tip;
3) nature of how the fibres are attached to the shaft (tightly woven round

or glued and projected from the shaft);

4) the material of the shaft; and
5) the shape and design of the storage tube (e.g. air-vents to facilitate air-drying, anti-microbial properties/agents present to preserve DNA).

Three swab designs commonly used are: wound swabs – fibres are tightly woven around the shaft of the swab (e.g. cotton); flocked swabs – fibres are sprayed onto the shaft and project from the shaft with a brush effect (e.g. nylon flocked); and pad swabs – porous material (e.g. foam) wrapped around the core (see Figure 3.1).

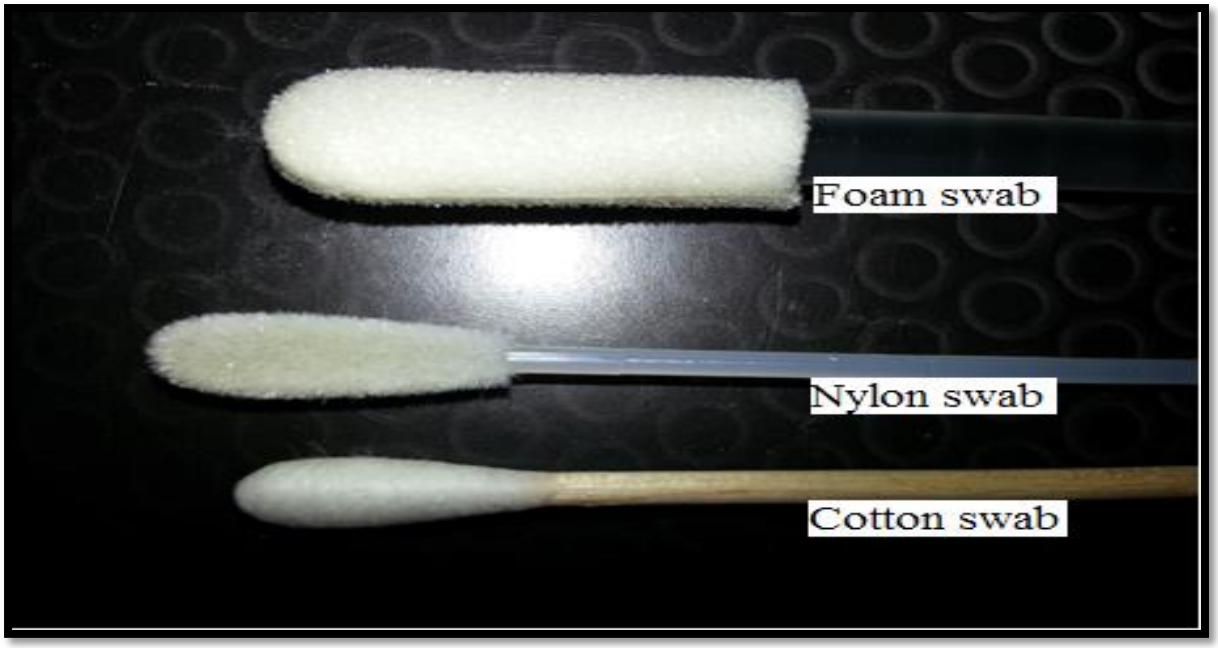


Figure 3.1 Swabs used in a comparative study (foam, nylon and cotton) for chapter III.

Nature of cotton wound swabs

Cotton swabs have a woven design of cotton fibres tightly wrapped around the tip of a wooden or plastic shaft (see Figure 3.1). A hollow internal shaft holding the swab matrix together draws solution up along the fibres through capillary action. The swabs feature a dense inner core that traps cellular material within its fibres. These swabs are popular for use in forensic DNA recovery as the OH groups of cotton fibres (see Figure 3.2) form strong hydrogen bonds with DNA [21]. Similarly, cotton swabs are highly permeable to water and other polar solvents owing to their hydrophilic functional groups [22]. Hydrogen bonds hold long chains of glucose molecules together and account for the 99% cellulose structure [23] (see Figure 3.2).

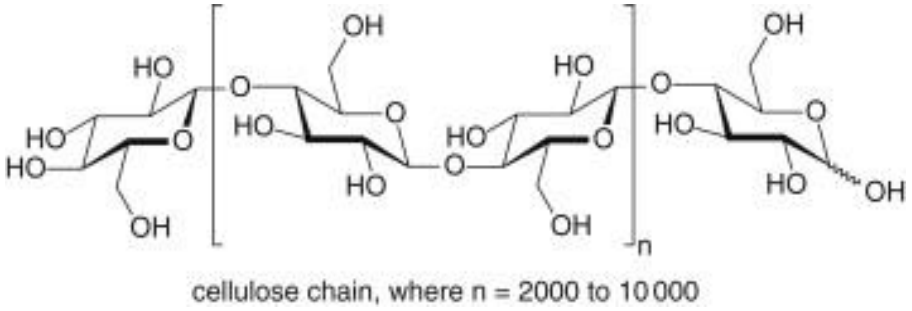


Figure 3.2 Chemical structure of cellulose that comprises 99% of the cotton swab. Image generated from Rudnik, E., Handbook of Biopolymers and Biodegradable Plastics: Properties, Processing and Applications [24].

Nature of pad foam swabs

Foam swabs are formed from gas pockets trapped in a solid matrix (see Figure 3.1) Padded foam swabs designed for forensic purposes are examples of open-cell foam; water flows through the entire matrix displacing the air. Most foam swabs used for forensic purposes are designed to efficiently absorb liquid from the surface from which it makes contact. Isopropanol foam popules (i.e. Mini-popule) (Puritan Medical Products, USA) are currently used by Forensic Science South Australia (FSSA) to recover 'touch' DNA evidence. The swab heads have a larger surface area compared to standard cotton swabs (see Figure 3.1). The inner shafts of Mini-popules are filled with isopropanol which is released onto the substrate when the swab is used.

Nature of nylon flocked swabs

A swab called the Nylon FLOQSwab™ (Copan Flock Technologies, Brescia, Italy) should be considered as an alternative for swabbing forensic material (see Figure 3.1 and 3.4). Medical research conducted reports that nylon swabs are more effective than rayon swabs at collecting epithelial cells from the nasal passage [25]. Similarly in a microbiological study, nylon swabs were proven to be extremely effective at collecting microbes from swabbed surfaces compared to other swab types and there was an improvement in microbe yield of 20 – 60% [26].

Recently a comprehensive study was published comparing the effectiveness of three swab types for recovering DNA from saliva and data reported on the superior performance of nylon-flocked swabs (i.e. forensicX) [12]. The average yield of DNA

recovered was between two-to-three times higher with the nylon swab [12], and nylon swabs yielded more DNA 40 days after storage at room temperature.

Nylon FLOQSwab™ fibres are positively charged and attract the negative charge on the sugar-phosphate backbone of DNA; the NH groups of nylons (see Figure 3.3) form strong hydrogen bonds with DNA or the hydrophobic cell membrane [21]. By design of a flocked swab technology, the swabs are made up of thousands of parallel nylon fibres that have been sprayed onto a plastic stick while it is held in an electrostatic field (see Figures 3.1 and 3.4) [27]. The short fibre strands are arranged in a perpendicular fashion creating an absorbent thin layer with an open structure. Unlike traditional cotton swabs that contain an internal absorbent core, nylon FLOQSwabs™ have no internal mattress or cushion to entrap the sample. This design strategy allows sample DNA to remain close to the surface of the fibre and increases the surface area for sample collection and release from the hydrophilic layer of nylon pile [27]. This design feature is especially important when working with low-level or 'touch' DNA, as a dense core may trap, and not release, the small amount of cells captured. In addition, nylon FLOQSwabs™ initially designed for medical purposes [25, 26] are sterile, PCR inhibitor-free, RNase-free and DNase-free, so ideally suited to forensic science.

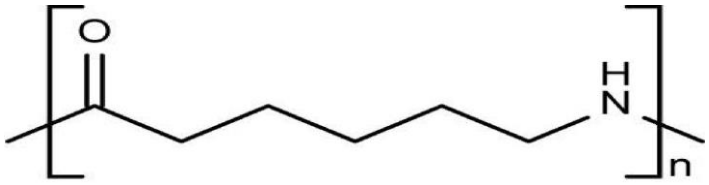


Figure 3.3 Chemical structure of nylon demonstrating NH groups that will form strong hydrogen bonds with DNA. Image adapted from Sigma-Aldrich [28].

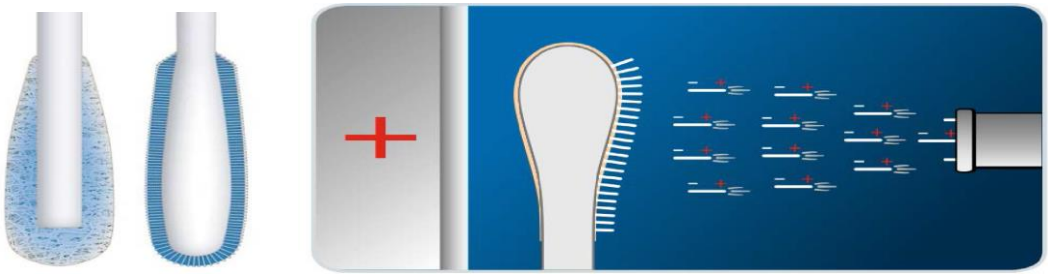


Figure 3.4 Diagram of the flocking process used by Copan Diagnostics, where strands of nylon are applied to the plastic core swab head. Image generated from Copan, 2012, nylon FLOQSwabs™ by Copan, California, USA [27].

Double swabbing method

The double swabbing technique [29, 30] involves using a water absorbed swab on the substrate to recover DNA, followed by a second dry/wet swab. A study by Pang *et al.* (2007) [31] demonstrate the advantage of a double swabbing technique for recovering DNA from touched items. The authors report that a single wet swab did not efficiently capture DNA and they observe an improvement in profiling results after pooling the first wet swab and second dry swab together prior to extraction [31]. Their findings suggest that the second dry swab alone yielded sufficient DNA to generate a profile in 60% of cases [31]. Another study by Sweet *et al.* (1997) [29] indicates a significant improvement in profiling results when the double swabbing technique was used to recover DNA from saliva in a human bite mark case. Likewise, the double swabbing technique has been used in many other studies aiming to retrieve 'touch' DNA from different sources [32-35].

Swab media

In addition to the type of swab used, the mechanism of DNA recovery will depend on the nature of the substrate swabbed [31] and the rehydrated state of the cells [29]. When epithelial cells are rehydrated by moisture they loosen and the natural capillary movement of liquid between the fibre strands facilitates strong hydraulic uptake of samples. Mechanical swabbing with the aid of hydrophilic swab fibres will aid the absorption of rehydrated cells. Devices such as FTA® paper, used for collection of body fluid, contain DNA preserving reagents that help to lyse cells. Most swabs used for collecting 'touch' DNA do not contain lysing agents; however, they can be pre-treated with buffers prior to sampling.

Many forensic laboratories currently use swabs moistened with distilled sterile water as a media for cell recovery from touched items. DNA is a polar hydrophilic molecule (i.e. dissolves in water) owing to the negative charge on the sugar-phosphate backbone. However, water cannot solubilize all membrane proteins and fats easily and this may prevent total cell lysis; hydrophobic components of the cell that lack polar or charged functional groups will not dissolve readily in water. For this reason, other swab media should be considered that will further aid cell lysis.

Minimal research has been conducted into comparing different swabbing media for the retrieval of 'touch' DNA. A short study by Phetpeng *et al.* (2013) [36] was published comparing swab moistening agents and different brands of cotton swabs, and findings concluded that various combinations of swabs and buffers yielded different amounts of DNA. The authors report that the EO cotton swab (Thai Gauze, Thailand) combined with isopropanol buffer generated the highest DNA yield [36]. Other buffers tested included phosphate buffered saline (PBS), sterile water, ethanol, SDS, and lysis buffer.

A recent study by Thomasma *et al.* (2013) [37] reported that swabs moistened with detergents, compared to water-based swabs, led to a greater rate of DNA recovery from hand prints deposited on glass. In particular, detergents such as Triton™-X-100 (0.1%) and SDS provided significantly higher yields of DNA.

Detergents are amphiphilic in nature (i.e. allow solubility in polar and non-polar solvents) and have hydrophobic-associated properties as a result of their non-polar groups but also have hydrophilic-associated properties making them soluble in water (see Figure 3.5). Triton™-X is an example of a non-ionic detergent that would help to aid cell lysis and extraction of DNA, RNA and soluble protein [37, 38]. Detergent-based swab media should, therefore, be considered to help aid 'touch' DNA recovery as it should, in theory, improve cellular DNA recovery during swabbing and subsequently increase the likelihood of obtaining a DNA profile. However, it should be noted that lysing cells may not be ideal in some circumstances that require laser micro-dissection or XY FISH of specific cell types. If a lysing agent is used then a validation study would be required to ensure that there are no downstream implications for analysis (e.g. detection of body fluid such as saliva).

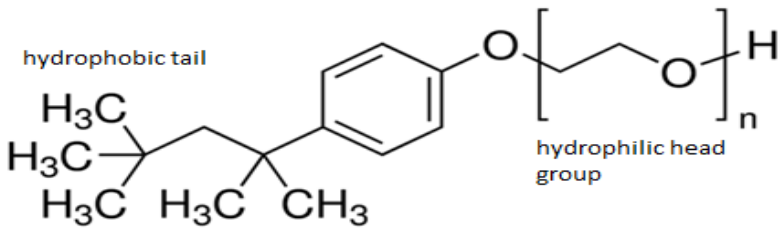


Figure 3.5 Complex chemical structure of Triton™-X-100 showing hydrophobic and hydrophilic groups. Structure adapted from Sigma-Aldrich [28].

Substrates

Surface materials examined in this thesis for the recovery of DNA include plastic, wood, glass and metal substrates. The binding affinity of DNA to substrates may influence the level of DNA recovery and hinder the subsequent profiling ability of the sample. If the surface that is swabbed is porous (e.g. paper or wood) then the DNA will become less accessible for recovery and will be harder to retrieve by swabbing alone [31], compared to surfaces that are non-porous (e.g. glass). Additional moistened swabs or other methods of collection may be required to collect DNA from large and/or porous surface areas [29]. On the other hand, substrates that have no grooves for DNA/skin cells to adhere/become trapped (e.g. smooth surfaces) may not be effective at retaining DNA for long periods of time. A study by Nanassy *et al.* (2007) report that glass has an affinity to DNA at the molecular level [39]. Equally, glass beads have been used to concentrate and bind DNA for many years.

It is important to consider the nature of the substrate that is being targeted for DNA prior to swabbing in order to determine the most effective collection device and method to use [20, 32]. Recent developments in DNA testing have opened up new avenues of research, some of which have great potential in forensic science. These include using the direct PCR approach for amplifying low quantity DNA from biological sources [40-44].

Direct PCR

Most DNA extraction methods are tailored towards high molecular weight DNA and cellular material that is intact and not highly degraded [45]. Current methodology for 'touch' DNA extraction encompasses multiple tube transfers and a sample purification stage that contributes to a significant loss of target DNA [8, 46, 47]. Large quantities of DNA may be retained on the swab surface [48] and subsequently lost during a routine extraction [42]. A recent study by Adamowicz *et al.* (2014), reported a loss of 50% of DNA, as assayed by quantitative PCR, from a swab subjected to an extraction [48]. The binding of DNA to columns or silica particles and the removal of PCR inhibitors - through purification wash steps – may prevent efficient recovery of DNA [49, 50]. Previously, published work reports a significant loss of DNA following an extraction with the highest loss reported at 80% [21, 40-42, 50-55].

3.3.1 Comparison of swab types for direct PCR

3.3.1.1 Materials and Methods

Preliminary work to supplement manuscript data

All contamination measures and controls are indicated in Chapter II. Sampling was repeated in triplicate. Control DNA (10 ng of 2800M, Promega) was diluted to 1 ng / µL, and 1 µL was deposited onto 3 x sterile plastic microscope slides (Rinzl plastic, ProSci Tech, QLD, AU) and slides air-dried in a sterile fume hood for 10 minutes prior to swabbing. The tips of three swabs were then moistened with 20 µL 0.1% Triton™-X (Sigma-Aldrich, AU) and swab types consisted of: i) DNA-free nylon FLOQswabs™ (Copan Brescia, Italy); ii) foam tipped DNA-free applicator swabs (Whatman, USA) and; iii) sterile cotton swabs (Livingstone, NSW, AU). All slides were subjected to 'targeted' swabbing using fibre clumps (see Chapter II). Post-swabbing, the fibre clumps were added directly to a 0.2 mL PCR tube for amplification. To examine the inhibition effect that fibres may have on the PCR, diluted DNA (2800M, Promega, AU) at 1 ng was dispensed directly onto ~ 2 mm^2 swab fibres labelled: i) nylon; ii) cotton; and iii) foam, prior to placing the fibres into a 0.2 mL PCR tube, and results compared to the positive control with no fibres present.

Controls

A positive control was set up where 1 ng of control DNA (2800M, Promega) was added directly to a PCR tube. Negative controls were set up consisting of a PCR blank (i.e. no DNA), a swab of a slide containing no DNA (to monitor slide and swab contamination), and fibres added to a PCR tube containing no DNA (to monitor fibre contamination).

Amplification

Amplification of all reactions took place following the AmpFℓSTR® ProfilerPlus™ kit or NGM SElect kit™ guidelines (see Chapter II).

3.3.1.2 Results and Discussion

Preliminary work to supplement manuscript data

Optimum method of recovery

Swab type

Nylon FLOQ™ swab fibres were the only type of fibre to produce a DNA profile upon swabbing and direct PCR (see S.I. Table 3.1 – data example). However, controls set up to examine the inhibitory effect of fibres on the direct PCR approach - without swabbing - showed that foam and nylon FLOQ™ swab fibres present in the PCR tube (with 1 ng of template DNA) did not affect the ability to generate a DNA profile (see SI Table 3.2 – data example). In this experiment, foam fibres present in the PCR tube did result in DNA alleles; however, the RFU value of peak heights for foam was less than the RFU values generated from nylon swabs (see SI Table 3.2). On the contrary, cotton fibres appeared to have a detrimental effect, resulting in failed amplification (see SI Table 3.2).

Replicate testing

Testing was repeated in triplicate and repeat testing demonstrated the inability to obtain a DNA profile when foam or cotton swab fibres were present in the PCR tube. This inconsistency in results – from previously generating a profile using foam fibres when no swabbing took place - may be due to the practical difficulties in obtaining the right amount of fibre material for foam and cotton fibres to work well under direct PCR conditions. In comparison, nylon FLOQ™ swab fibres consistently generated DNA alleles for every sample tested.

Nylon FLOQ™ swabs aid DNA collection due to the nylon fibres being positively charged (i.e. they have an affinity for negatively charged molecules such as DNA), and the sample is more easily absorbed by capillary action [12]. Secondly, the nature of the nylon flocking process enables sample to be in close proximity to the surface of the swab rather than trapped within the inner core matrix. This inherent feature of the swab (i.e. brush-like) is thought to reduce sample retention [26] and improve the release of sample material [31]. On the contrary, for foam and cotton swabs, the sample is entrapped within the core matrix (see Figure 3.1). This dense matrix may

contribute to the sample not fully eluting from the swab during the 95 °C heat step in PCR. A combination of these factors may influence the success of direct PCR profiling of swab material. A smaller section of swab material (i.e. < 2 mm^2) may be required for future success with foam and cotton swabs.

Additional data for the comparison of swab types for the purpose of direct PCR are published in the manuscript enclosed: "**Templeton, J.,** Ottens, R., Paradiso, V., Handt, O., Taylor, D., & Linacre, A. (2013). Genetic profiling from challenging samples: direct PCR of 'touch' DNA. Forensic Science International: Genetics, Supplement Series, 4(1), e224-e225".

3.3.1.3 Statement of authorship

Manuscript: Genetic profiling from challenging samples: Direct PCR of touch DNA

Manuscript published in Forensic Science International: Genetics Supplement Series

Jennifer Templeton (Candidate)

Carried out laboratory work, analysis, and wrote the paper.

I hereby certify that the statement of contribution is accurate.

Signed: Date: November 2016

Renée Blackie (nee Ottens)

Helped to edit the manuscript

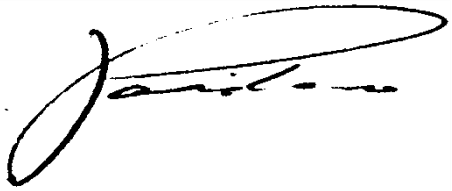
I hereby certify that the statement of contribution is accurate.

Signed: Date: November 2016

Viviana Paradiso

Helped to edit the manuscript

I hereby certify that the statement of contribution is accurate.



Signed: Date: November 2016

Oliva Handt

Helped to edit the manuscript

I hereby certify that the statement of contribution is accurate.

Signed: Date: November 2016

Duncan Taylor

Helped to edit the manuscript

I hereby certify that the statement of contribution is accurate

Signed: Date: November 2016

Adrian Linacre (Supervisor)

Helped to edit paper and assisted with experimental design.

I hereby certify that the statement of contribution is accurate.

Signed: Date: November 2016

Manuscript: Genetic profiling from challenging samples: Direct PCR of touch DNA

Forensic Science International: Genetics Supplement Series 4 (2013) e224–e225

Contents lists available at ScienceDirect

Forensic Science International: Genetics Supplement Series

journal homepage: www.elsevier.com/locate/FSIGSS

Genetic profiling from challenging samples: Direct PCR of touch DNA

Jennifer Templeton [a,*], Renée Ottens [a], Viviana Paradiso [a], Oliva Handt [b], Duncan Taylor [b], Adrian Linacre [b]

[a] School of Biological Sciences, Flinders University, Australia
[b] Forensic Science South Australia, Adelaide, Australia

ARTICLE INFO

Article history:
Received 26 August 2013
Accepted 2 October 2013

Keywords:
Direct PCR
Touch DNA
Identification
Short tandem repeat
Profiling

ABSTRACT

We report on a novel direct PCR approach for retrieving touch DNA from handled items. Trace levels of DNA on handled items can be difficult to recover and analyse using standard methodology. The lack of a full nuclear STR DNA profile may be related to insufficient DNA being present initially, the presence of PCR inhibitors, or the loss of nucleic acid during sample collection and processing. The aim of this study was to optimise the swabbing method used to collect DNA and to improve the success rate of STR-based DNA profiling from touched items. A comparison was carried out to evaluate the use of three common swab types, nylon, cotton and foam at retrieving and releasing template DNA for amplification by direct PCR. To test the effectiveness of our method at processing low levels of DNA, complete STR DNA profiles were obtained from fingerprints deposited onto sterile plastic slides, which were swabbed and subjected to direct PCR. Our findings indicate the potential use of direct PCR when analysing samples that contain limited DNA.

1. Introduction

Low levels of DNA recovered from touched items can be close to the limit of detection for current STR-based DNA profiling kits. The success of a DNA profile depends largely on how much DNA is retrieved from the evidence. Cotton swabs, pre-moistened with distilled water, are used routinely in forensic laboratories. However, other swab types should also be considered due to the low success rate of current methods. Following DNA collection, current methodology subjects the swab to a DNA extraction to isolate and purify the target DNA, typically using a solid phase substrate, prior to amplification. The DNA extraction process contributes to a significant loss (20–70%) of sample DNA [1] and has the potential to introduce extraneous DNA into the reaction [2]. Direct PCR circumvents the need for a DNA extraction. For direct PCR, the biological material is transferred directly into the PCR tube with no prior extraction or purification steps. Previous studies report full STR DNA profiles obtained from trace evidence using direct PCR [3,4]. The aim of this study was to examine aspects of sample collection, apply the method of direct PCR to low level DNA templates, and to improve the success of STR DNA profiling. In this study we report the ability of direct PCR, using 29 cycles, to generate full STR DNA profiles from fingerprints deposited on plastic substrates.

2. Materials and methods

Control DNA (2800 M, Promega) of a known concentration (1 ng/μL) was deposited onto 3× sterile plastic microscope slides and swabbed using either foam (Whatman, USA), DNA-free nylon FLOQswabs™ (Copan Industries, Vic) or cotton swabs (Livingstone, NSW). For direct PCR, a 2 mm^2 portion of each swab tip was cut using a sterile scalpel blade and placed directly into a 0.2 mL PCR tube. Amplification conditions followed the Profiler Plus® STR kit (ABI) guidelines in a final 25 μL reaction volume and standard 28 cycles. For fingerprint deposition, three donors placed the index finger of their dominant hand onto a sterile plastic slide 2 h after hand washing. For direct PCR, a 2 mm^2 portion of the swab tip was cut using a sterile scalpel blade and placed directly into a 0.2 mL PCR tube. Amplification followed the NGM™ STR kit (ABI) guidelines in a final 25 μL reaction volume and 29 cycles. PCR products were run on the 3130*xl* (ABI) and profiles analysed using GeneMapper™ ID v3.2 software and a peak amplitude threshold of 30 RFU was used to assign alleles.

3. Results and discussion

Nylon swabs generated the highest DNA yield, as assayed by the relative fluorescence units (RFU) of peak heights, followed by foam and then cotton swabs (Fig. 1). Full STR DNA profiles were

* Corresponding author at: School of Biological Sciences, Flinders University, Bedford Park, South Australia 5001, Australia. Tel.: +61 882 015 003.
E-mail addresses: jennifer.templeton@flinders.edu.au, jenny_templeton27@hotmail.com (J. Templeton).

1875-1768/$ – see front matter © 2013 Elsevier Ireland Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.fsigss.2013.10.115

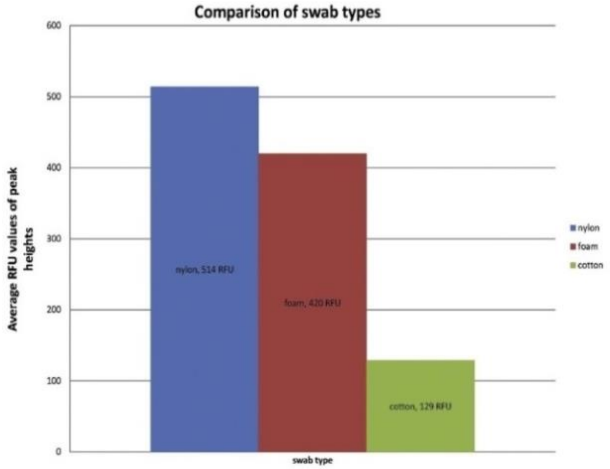


Fig. 1. A comparison between swab types, using control DNA deposited on a plastic slide, prior to swabbing, PCR and STR DNA profiling.

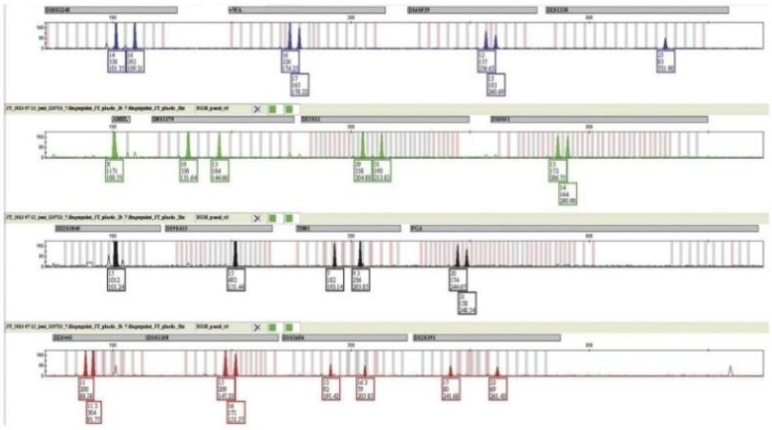


Fig. 2. Example of an STR DNA profile generated, using direct PCR, from a fingerprint deposited on a plastic slide prior to swabbing and processing (29 cycles of PCR, using NGM™ STR kit (ABI)).

obtained, using nylon flocked swabs and direct PCR, from all three fingerprints deposited on plastic substrates 2 h after the donor washed their hands. No allele drop-in or allele drop-out was observed (Fig. 2). Laboratories may benefit from using direct PCR on handled items due to its high-throughput potential and increased sensitivity. However, the success and quality of DNA profiles obtained using direct PCR will be dependent on the nature of the material examined and the presence of PCR inhibitors; which are generally eliminated during the extraction process.

4. Conclusion

Nylon flocked swabs (Copan) provided the optimum method of DNA recovery from plastic slides. Nylon flocked swabs and direct PCR was used to generate full STR DNA profiles from three fingerprints deposited on plastic substrates 2 h after the donor washed their hands. No allelic drop-in was detected and only 29 cycles were used in the PCR process. The method of direct PCR should be considered as an alternative method for analysing samples that contain low amounts of DNA for use in a high-throughput environment. A validated study is required to assess the limitations of the direct PCR approach. However, we anticipate the method to have future niche applications in analysing latent DNA recovered from touched items that face the limits of detection when using standard protocols.

Role of funding

Funding was provided by the Department of Justice, South Australia.

Conflict of interest

None.

Acknowledgements

We thank members of Forensic Science South Australia and the School of Biological Sciences at Flinders University for assisting with reagents and providing helpful advice.

References

[1] R.A.H. van Oorschot, D.G. Phelan, S. Furlong, G.M. Scarfo, N.L. Holding, M.J. Cummins, Are you collecting all the available DNA from touched objects, International Congress Series 1239 (2003) 803–807.
[2] J.M. Butler, Forensic DNA Typing: Biology, Technology, and Genetics of STR Markers, 2nd ed., Elsevier Academic Press, New York, 2005.
[3] A. Linacre, V. Pekarek, Y.C. Swaran, S.S. Tobe, Generation of DNA profiles from fabrics without DNA extraction, Forensic Science International: Genetics 4 (2010) 137–141.
[4] R. Ottens, D. Taylor, D. Abarno, A. Linacre, Successful direct amplification of nuclear markers from a single hair follicle, Forensic Science, Medicine, and Pathology (2012) 1–6.

3.4 Comparison of swab media and swab technique

3.4.1 Materials and Methods

Preliminary work to supplement manuscript data

Repeat testing was performed to determine the most effective swab technique and swab media for direct PCR. Prior to swabbing, control DNA (10 ng of 2800M, Promega) was diluted to 1 ng / µL, and 1 µL was deposited onto 4 x sterile plastic microscope slides and left to air-dry in a sterile fume hood for 10 minutes.

Triton™-X swabbing media

Two DNA-free nylon FLOQswabs™ were moistened with 20 µL 0.1% Triton™-X prior to swabbing. One slide of the set was subjected to 'targeted' swabbing using fibre clumps (see Chapter II). Post-swabbing, the fibre clump was added directly to a 0.2 mL PCR tube. One slide of the set was subjected to 'whole' swabbing (see Chapter II). After 'whole' swabbing took place, a sterile scalpel blade was used to section a clump of fibres from the tip of the swab (i.e. ~ 2 mm^2) and fibre material was placed into a 0.2 mL PCR tube using sterile forceps.

Water swab media

Two DNA-free nylon FLOQswabs™ were moistened with 20 µL sterile Milli-Q ultrapure water (Merck Millipore, VIC, AU) prior to swabbing. One slide of the set was subjected to 'targeted' swabbing (see Chapter II) and one slide of the set was subjected to 'whole' swabbing (see Chapter II).

Amplification

Amplification of all reactions took place following the AmpFℓSTR® ProfilerPlus™ kit or NGM SElect kit™ guidelines (see Chapter II).

3.4.2 Results and Discussion

Preliminary work to supplement manuscript data

Swab media

Repeat testing revealed that Nylon FLOQ™ swab fibres, pre-moistened with Triton™-X 0.1%, were more effective than water-based Nylon FLOQ™ swab fibres at collecting DNA from a plastic substrate and then releasing the DNA into the amplification vessel (see SI Table 3.3 and SI Table 3.4). For example, a full STR-based DNA profile using AmpF*ℓ*STR® ProfilerPlus™ (i.e. 20 alleles out of a possible 20 alleles) was generated from Triton™-X swabs used to retrieve 1 ng of DNA, in comparison to a partial DNA profile (i.e. 9 alleles out of a possible 20 alleles) obtained from water-based swabs (see SI Table 3.3). Likewise, Triton™-X swabs outperformed water-based swab media when 'whole' swabbing was performed (see SI Table 3.4). For example, when swabbing 1 ng of DNA, Triton™-X swabs generated a partial profile of 8 alleles (out of a possible 20 alleles), compared to only one marker (i.e. Amelogenin) amplified from water-based swabs (see SI Table 3.4). Results are consistent with previously published data stating that detergents are superior to water-based swab media [37].

The addition of Triton™-X appears to aid the lysis of cellular material by solubilizing the membrane proteins and phospholipids more readily than water. Another inherent feature of the Triton™-X surfactant is its high viscosity in preventing DNA from binding to the walls of the reaction tubes. It is known that ancient DNA laboratories add dilute non-denaturing detergents to their limited DNA extracts in order to prevent this irreversible binding effect and to minimise loss of crucial DNA [56].

Swab technique

Due to the flexible component of the nylon swab plastic shaft, the physical act of 'whole' swabbing proved difficult. Hence, the current design of the nylon FLOQSwab™ is not ideally suited for the collection of biological material at crime scenes [14].

A re-design of the standard swab shaft and swab material dimensions suitable for direct PCR is illustrated in a provisional patent at the end of this chapter. Repeat testing showed that a 'targeted' approach to swabbing, using sterile forceps to hold the pre-cut fibre clump in place over the slide and apply pressure resulted in a higher yield of DNA, compared to when the whole swab head was used (see SI Tables 3.3 and 3.4 and 3.5). A targeted approach to swabbing (i.e. fibres pre-cut) ensured that all DNA-saturated fibres in contact with the substrate were added to the PCR tube. This would increase DNA yield as more template DNA is made available to the amplification reaction. A whole swab head used for swabbing is a less targeted approach, resulting in a reduced yield (see SI Table 3.3, 3.4 and 3.5). For example, a full STR-based DNA profile (i.e. 20 alleles out of a possible 20 alleles) was generated from swabbing 1 ng of control DNA using a 'targeted' approach with Triton™-X nylon FLOQSwabs™ fibres (see SI Table 3.3), compared to 8 alleles (out of a possible 20 alleles) when using a 'whole' swab approach (see SI Table 3.4). Likewise, when water-based nylon FLOQSwab™ fibres were used with a 'targeted' approach there was a higher profiling success than with a water-based 'whole' swabbing approach. For example, nine alleles (out of a possible 20 alleles) were generated from swabbing 1 ng of control DNA using a 'targeted' approach (see SI Table 3.3), compared to only 1 allele generated when using a 'whole' swab approach (see SI Table 3.4). Further testing with a new STR kit (i.e NGM SElect kit™) was carried out to assess the swabbing technique, and full DNA profiles were obtained from both 'whole' swabbing and the 'targeted' approach, but substantially lower yield of DNA obtained by using 'whole' swabbing (as assessed by RFU values of peak heights) (see SI Table 3.5).

Complications observed in the experimental design relate to the application of swab media to the surface of the nylon FLOQSwabs™ fibres. The nylon swab did not appear to be highly absorbent when a large amount of swab media was dispensed onto the tip of the swab; however, by reducing the amount of swab media added to the tip of the swab from 20 µL to 2 µL, absorption was more effective. Further results using reduced swab media are published in the manuscript: "***Templeton, J. E.,*** *Taylor, D., Handt, O., Skuza, P., & Linacre, A. (2015). Direct PCR Improves the*

Recovery of DNA from Various Substrates. Journal of Forensic Sciences, 60(6), 1558-1562".

3.5 Direct PCR vs Extraction

3.5.1 Materials and Methods

Preliminary work to supplement manuscript data

For Preliminary testing for direct PCR, control DNA (10 ng, 2800M, Promega) was diluted to two different concentrations: 1 ng / µL and 0.5 ng / µL. DNA (1 µL) of each concentration was deposited onto a separate sterile plastic microscope slide (Rinzl plastic, ProSci Tech) labelled either: Direct i) 1 ng / µL or Direct ii) 0.5 ng / µL. Prior to swabbing, slides were air-dried in a sterile fume hood for 10 minutes. Two DNA-free nylon FLOQswabs™ were pre-moistened with 20 µL 0.1% Triton™-X (Sigma-Aldrich, AU) prior to swabbing (see Figure 3.6 (a)). For each concentration of DNA tested, swabbing consisted of 20 strokes horizontally (i.e. left to right) and 20 strokes vertically (i.e. top to bottom) over the slide using medium pressure to ensure consistency (see Figure 3.6 (b)). Later studies using control DNA involved swabbing with 10 strokes over the substrate of interest with no effect on the ability to recover adequate DNA for profiling (see manuscript detailed below: Direct PCR improves the recovery of DNA from various substrates). For each sample, a sterile scalpel blade was used to cut a small portion of fibres (~ 2 mm^2) from the tip of a DNA-free nylon FLOQswab™ (see Figure 3.6 (c)). Swab fibres (~ 2 mm^2) were added directly to a 0.2 mL PCR tube (see Figure 3.6 (d and e)).

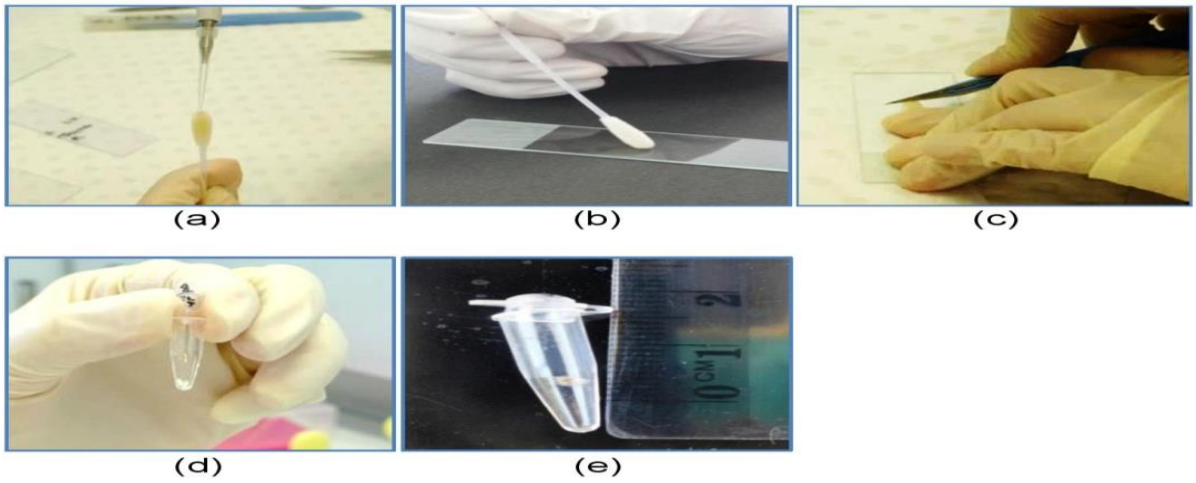


Figure 3.6 Sample preparation for swabbing: a) 20 µL of Triton™-X or sterile 'DNA-free' water was deposited onto nylon FLOQswab™ fibres; b) swabbing using the tip of the nylon FLOQswab™ c) sterile scalpel blade and forceps were used to transfer ~ 2 mm^2 fibre section to the PCR tube. d) and e) 0.2 mL PCR tube containing ~ 2 mm^2 section of swab fibres ready for direct PCR.

For extraction, Control DNA (10 ng, 2800M, Promega) was diluted to two different concentrations: 1 ng / µL and 0.5 ng / µL. DNA (1 µL) of each concentration was deposited onto a separate sterile plastic microscope slide (Rinzl plastic, ProSci Tech) labelled either: Extraction a) 1 ng / µL or Extraction b) 0.5 ng / µL. Prior to swabbing, slides were left to air-dry in a sterile fume hood for 10 minutes. DNA-free nylon FLOQswabs™ were moistened with 20 µL 0.1% Triton™-X prior to swabbing. For each concentration of DNA tested, swabbing consisted of 20 strokes horizontally (i.e. left to right) and 20 strokes vertically (i.e. top to bottom) over the slide using medium pressure to ensure consistency. For each concentration of DNA tested, the whole swab head was added to a 1.5 mL Eppendorf tube and the DNA was extracted following manufacturer's recommendations for the DNA IQ™ System (Promega).

Amplification

Amplification followed using the AmpFℓSTR® ProfilerPlus™ kit guidelines.

For Capillary Electrophoresis and Data Analysis of all experiments - See Chapter II

For the data analysis steps, a 'full' DNA profile was defined when all the expected alleles, in all loci, were observed. Full DNA profiles did not show 'allele drop-out' at any loci.

3.5.2 Results and Discussion

Preliminary work to supplement manuscript data

Replicate testing revealed that the presence of nylon FLOQSwabs™ fibres and glue fragments in the PCR tube (with PCR volume = 25 µL) did not affect the ability to generate amplified product (see SI Table 3.6, 3.7 and 3.8). Full STR-based DNA profiles were generated from swabbing 1 ng of control DNA deposited on plastic and subjecting the swab fibres to direct PCR at standard PCR cycling conditions (Example of data is illustrated in SI Table 3.7) (n=3). When a smaller mass of DNA was deposited on the slide for swabbing (i.e. 0.5 ng), only 4 alleles were generated using the direct PCR approach, compared to 17 alleles for positive control DNA (i.e. no swabbing) (see SI Table 3.8). Direct PCR improved DNA profiling results over conventional DNA extraction (i.e. DNA IQ™ System) and the results are consistent with a recent published study [42]. To expand on these results and to explore the limit of detection for the direct PCR approach, results were published in the manuscript: "***Templeton, J. E.,*** *Taylor, D., Handt, O., Skuza, P., & Linacre, A. (2015). Direct PCR Improves the Recovery of DNA from Various Substrates. Journal of Forensic Sciences, 60(6), 1558-1562".*

Controls

There was one case where the Amelogenin marker (XY) was amplified from a negative control swab of a slide (see SI Table 3.6). All other negative controls in this study were blank. All samples produced concordant profiles with no apparent contamination detected. Amelogenin is a common marker to be amplified in low-level DNA forensic laboratories [2], and it was thought to be from background low-level DNA in the laboratory, however, extreme caution should be taken to minimise extraneous contamination.

All sampling was replicated in an extensive study and data published in the manuscript enclosed: **Templeton, J. E.,** Taylor, D., Handt, O., Skuza, P., & Linacre, A. (2015). Direct PCR Improves the Recovery of DNA from Various Substrates. Journal of Forensic Sciences, 60(6), 1558-1562".

3.5.2.1 Statement of authorship

Title of publication: Direct PCR improves the recovery of DNA from various substrates

Manuscript published in Journal of the Forensic Sciences

Jennifer Templeton (Candidate)

Carried out all the laboratory work, analysis, and wrote the paper

I hereby certify that the statement of contribution is accurate.

Signed:  Date: November 2016

Oliva Handt

Helped to edit the manuscript

I hereby certify that the statement of contribution is accurate.

Signed:  Date: November 2016

Duncan Taylor

Helped to edit the manuscript

I hereby certify that the statement of contribution is accurate.

Signed:  Date: November 2016

Pawel Skuza

Helped with the statistical analysis and helped to edit the manuscript

I hereby certify that the statement of contribution is accurate.

Signed: Date: November 2016

Adrian Linacre (Supervisor)

Assisted with experimental design and helped to edit paper

I hereby certify that the statement of contribution is accurate.

Signed: Date: November 2016

Manuscript: Direct PCR improves the recovery of DNA from various substrates.

JOURNAL OF FORENSIC SCIENCES

J Forensic Sci, 2015
doi: 10.1111/1556-4029.12843
Available online at: onlinelibrary.wiley.com

TECHNICAL NOTE

PATHOLOGY/BIOLOGY

Jennifer E. L.Templeton,[1] *M.Sc.; Duncan Taylor,*[1,2] *Ph.D.; Oliva Handt,*[1,2] *Ph.D.; Pawel Skuza,*[3] *M.Ed.; and Adrian Linacre,*[1] *D.Phil.*

Direct PCR Improves the Recovery of DNA from Various Substrates*

ABSTRACT: This study reports on the comparison of a standard extraction process with the direct PCR approach of processing low-level DNA swabs typical in forensic investigations. Varying concentrations of control DNA were deposited onto three commonly encountered substrates, brass, plastic, and glass, left to dry, and swabbed using premoistened DNA-free nylon FLOQswabs™. Swabs (n = 90) were either processed using the DNA IQ™ kit or, for direct PCR, swab fibers (~2 mm^2) were added directly to the PCR with no prior extraction. A significant increase in the height of the alleles ($p < 0.005$) was observed when using the direct PCR approach over the extraction methodology when controlling for surface type and mass of DNA deposited. The findings indicate the potential use of direct PCR for increasing the PCR product obtained from low-template DNA samples in addition to minimizing contamination and saving resources.

KEYWORDS: forensic science, DNA typing, trace DNA, direct Polymerase Chain Reaction, short tandem repeats, human identification

The recovery of latent DNA from touched items may establish a clear association between the perpetrator and the item of evidence, or link a suspect to the scene, particularly in cases where secondary or tertiary transfer is not present. "Touch" DNA is a form of trace evidence that refers to DNA recovered from skin (epithelial) cells (1) or cell-free nucleic acid (CNA) (2) left behind when a person comes into contact with an item. A human can shed on average 400,000 skin cells per day (3). However, not every contact transferred by "touch" via finger marks will leave behind sufficient DNA to generate a STR-based DNA profile for identification purposes. In contrast, many transfer events do, and can, result in a meaningful DNA profile if appropriate measures are taken to adequately recover trace levels of DNA. Recent advancements to DNA recovery methodology include new swab types, swabbing reagents, and methods of processing swabs prior to analysis.

The first stage in the recovery of "touch" DNA involves swabbing the area of interest at the scene or in the laboratory. Many forensic laboratories currently recover DNA from substrates using sterile cotton swabs premoistened with distilled water. More recently, it has been reported that detergents offer an advantage for capturing DNA (4) over water-based swabbing media. Nondenaturing detergents, such as Triton X, have the ability to solubilize the hydrophobic components of cells such as membrane fats and proteins more readily than water, thus having the potential to capture more DNA from a substrate (4).

[1]School of Biological Sciences, Flinders University, Bedford Park, Adelaide, SA 5042, Australia.
[2]Forensic Science South Australia, Adelaide, SA 5000, Australia.
[3]eResearch, Flinders University, Bedford Park, Adelaide, SA 5042, Australia.

*Funding provided by the Department of Justice, South Australia.

Received 1 July 2014; and in revised form 19 Oct. 2014; accepted 2 Nov. 2014.

Following DNA retrieval, swabs are subjected to an extraction process using Chelex, or a commercial kit such as DNA IQ™ (Promega, Alexandria, NSW, Sydney, Australia), PrepFiler® (Life Technologies, Mulgrave, Vic., Australia), or QIAamp DNA Investigator® kit (Qiagen, Chadstone, Vic., Australia) to lyse and purify endogenous DNA. Despite advances in extraction kit optimization for dealing with low-template DNA samples, protocols utilize multiple tube changes and several wash steps that may contribute to a loss of valuable sample DNA (5) and hence valued information.

In a recent study, a DNA loss of 50–90% was reported following an extraction from 150 nanograms of control DNA deposited on paper using the phenol–chloroform method, Nucleospin® (Macherey-Nagel, Bethlehem, PA) or Invisorb® forensic kit (Invitek, Hayward, CA) (6). Likewise, silica-based extraction protocols have irreversible binding sites on the silica matrix that permanently bind nucleic acids (7) and may contribute to a loss of sample DNA if carrier RNA is not present in the extraction (8). Recent findings report cell-free DNA, derived from apoptotic skin cells, in the supernatant of 90% of biological samples (9); this supernatant portion is usually discarded during a routine extraction (2) and could prove to be a valuable source of target DNA. Hence, altering the extraction methodology to utilize cell-free nucleic acid has the potential to increase the amount of template DNA available for PCR (10).

Eliminating the DNA extraction stage altogether provides an additional source of DNA to the PCR in the form of cell-free DNA. In addition, the risk of contamination is minimized by side-stepping multiple tube changes and wash steps. Recent studies highlight the benefits of omitting the extraction step, also known as direct PCR, in a forensic context (11–15). Direct PCR has previously been used to amplify DNA adhering to cotton swabs that were premoistened with water (16). However, in light of recent advances, new swab types (17) and new swab media

(4) can now be considered as alternative methods for DNA recovery. Novel DNA capture methods in conjunction with direct PCR can be used to explore the potential for improvement in information obtained from forensic samples (15).

In this study, we report the ability of detergent-based nylon swabbing and direct PCR, using standard PCR conditions, to generate full STR-based DNA profiles from a low mass of control DNA deposited on plastic, glass, and brass substrates. In order to investigate properly the potential for a new method to be considered for the analysis of "touch" DNA, a known amount of control DNA is required to standardize DNA deposition. Control DNA was used to circumvent DNA variation levels naturally deposited via finger marks and to mimic the cell-free DNA component of a biological sample where the cell wall is no longer intact. If fingerprint residue is used to compare two methodologies for DNA recovery, then significant sample to sample variability would exist. Recent research has shown that variable amounts of DNA can be recovered from the hands of individuals depending on a multitude of factors (18), making it difficult to directly compare methods of analysis. In this study, profiles obtained by traditional PCR, which includes an extraction step ($n = 45$), were compared to those obtained by direct PCR ($n = 45$). We describe a one-step protocol for the amplification of low-template DNA adhering to swab fibers by eliminating the extraction step in an attempt to prevent the loss of valuable sample DNA in addition to saving resources.

Materials and Methods

DNA Extraction, Quantification, and STR Profiling

Control DNA (single-source male human genomic DNA, 2800 M; Promega) of a known concentration (10 ng/μL) was diluted to 5 different concentrations using Milli-Q ultrapure water (Merck Millipore, Bayswater, Vic., Australia): 1, 0.75, 0.5, 0.2, and 0.1 ng/μL. The concentration of each DNA dilution was confirmed by quantification using the Qubit® 2.0 Fluorometer and dsDNA high-sensitivity reagents (Invitrogen, Carlsbad, CA) following manufacturer's recommendations. Diluted DNA of each concentration (1 μL) was deposited onto 6 sterile surface materials: 2 × brass cartridge cases (Winchester, Australia), 2 × plastic microscope slides (Rinzl plastic; ProSci Tech, Kirwan, QLD, Australia), and 2 × glass microscope slides (Livingstone, Rosebery, NSW, Sydney, Australia) and left to air-dry in a fume hood for 10 min prior to swabbing. DNA-free nylon mini FLOQswabs™ (Copan, Brescia, Italy) premoistened with 2 μL 0.1% Triton X (Sigma-Aldrich, Castle Hill, NSW, Sydney, Australia) (preheated to 50°C) were used to swab the area where the control DNA was deposited onto each substrate. The swab was held perpendicular to the substrate to ensure that only the swab tip had direct contact with the substrate. Swabbing consisted of 10 strokes horizontally (left to right direction) and 10 strokes vertically (top to bottom) using medium pressure to ensure consistency. The first set of 15 swabs were subjected to a DNA extraction using the DNA IQ™ System (Promega) following manufacturer's recommendations and eluted in a final volume of 30 μL. For extraction, the stem of the swab was cut using a sterile scalpel blade, and the whole swab head was subsequently processed. The second set of 15 swabs was processed by direct PCR. Fibers (2 mm^2) were cut from the tip of the swab using sterile scalpel blades and placed directly into a 0.2-mL PCR tube. All sampling was carried out in triplicate.

Positive controls, containing 1 ng/μL of control 2800 M DNA, were included in each PCR setup. Negative controls included a swab of each substrate prior to DNA being deposited and distilled water controls in the PCR setup to monitor contamination. The eluted DNA was quantified prior to PCR using the Qubit® 2.0 Fluorometer and dsDNA high-sensitivity reagents (Invitrogen) following manufacturer's recommendations. STR typing for all samples was performed using the AmpFℓSTR ProfilerPlus™ kit (10 genetic STR loci typed, including Amelogenin; Applied Biosystems) following manufacturer's recommendations for suboptimal DNA (<1 ng) but at half the volume. The final reaction consisted of 10 μL ProfilerPlus™ reaction mix, 5 μL of ProfilerPlus™ primer mix, 1 μL (5 U) AmpliTaq Gold™ Applied Biosystems), and DNA template (2-mm^2 fibers or 20 μL eluted DNA). Cycling was performed on a 9700 GeneAmp thermal cycler (Applied Biosystems) and consisted of an initial denaturation at 95°C for 11 min followed by 29 cycles of 94°C for 1 min, 59°C for 1 min, and 72°C for 1 min, then a final extension at 60°C for 45 min. PCR products were analyzed on a 3130*xl* Genetic Analyzer (Applied Biosystems) in a 14 μL final volume that consisted of 1.5 μL PCR product (as recommended by ProfilerPlus™ kit manufacturer), 12 μL HiDi™ Formamide (Applied Biosystems), and 0.5 μL ROX-500 Size Standard (Applied Biosystems). Profiles were analyzed using GeneMapper ID software (v3.2.1) (Applied Biosystems), and a peak detection threshold of 30 relative fluorescence units (RFU) was used to assign alleles. A wildcard designation was used for potential homozygotes with peak heights <150 RFU to account for potential allele dropout (e.g., "11, F" instead of "11,11"). "F" was used to denote dropout of a specific allele. A profile was considered "full" when all alleles were detected above the threshold RFU and matched the reference profile of the control DNA. An informative DNA profile (one considered uploadable to the Australian DNA databases) was defined as at least 12 alleles (plus Amelogenin) that matched the reference DNA profile. Allelic dropout was reported as alleles with peak heights <30 RFU. The profiling success (%) was measured by dividing the number of alleles successfully called by the total number of expected alleles and multiplying this value by 100. The average peak height for each profile was calculated by dividing the summed peak heights of all detected peaks by the total number of peaks observed in that sample.

Statistical Analysis of Data

Data were compiled in Excel (Microsoft, 2010), and statistical analyses of the data were performed using Statistical Package for Social Sciences (SPSS, IBM, Armonk, NY) (19) to assess the difference in peak heights between the DNA profiles. A general linear model (GLM) was utilized with between factors of (a) surface (three levels: brass, glass, and plastic), (b) method (two levels: direct and extraction), and (c) DNA mass (five levels: 1 ng, 0.75 ng, 0.5 ng, 0.2 ng, and 0.1 ng). All factors were entered into GLM as fixed factors. Given that visual inspection of the data as well as Levene's test of equality of error variances indicated that GLM assumption of homogeneity of the variances is violated, all statistically significant interactions and pairwise comparisons were investigated using biased corrected accelerated bootstrapped significance testing (20,21). The results were considered statistically significant if p-values <0.05 were obtained. p-values were obtained after Bonferroni adjustment for multiple comparisons was applied.

Results

Effect of Method

The results of the GLM analysis are reported in Supporting information (See Table S1). The method used for DNA recovery (direct vs. extraction) had a statistically significant main effect on results ($p = 0.001$). When controlling for surface and mass of DNA as a factor, the direct PCR approach resulted in a significant increase in peak height of the alleles by 119 RFU (95% CI: 76–164 RFU) ($p = 0.02$) compared to the extraction method. As expected, the absolute increase in average peak heights from direct PCR was dependent on input DNA ($p = 0.006$ for interaction term in GLM) so that the peak height difference in profiles generated using the two methods increase with input DNA.

Figure 1 which reports estimated marginal RFU means from GLM procedure along with bootstrapped 95% confidence intervals (CI) visualize this interaction effect. This increase in PCR product for direct PCR samples (see Fig. 1) resulted in a higher % profiling success compared to the extracted swabs (see Fig. 2). For direct PCR, the lowest mass of control DNA deposited which yielded a full profile was as follows: 0.2 ng for plastic and 0.5 ng for glass and brass. The lowest mass of control DNA deposited which yielded a partial (and informative) profile by direct PCR was as follows: 0.1 ng for plastic, 0.2 ng for glass, and 0.5 ng for brass. In contrast, extracted DNA swabs required more initial template DNA to achieve partial or full DNA profiles. For extracted swabs, the lowest mass of control DNA deposited which yielded a full profile was as follows: 0.75 ng for brass, and 0.5 ng for plastic and glass. The lowest mass of control DNA deposited which yielded a partial (and informative) profile by extraction was 0.5 ng for all three substrates. All negative controls were blank as expected.

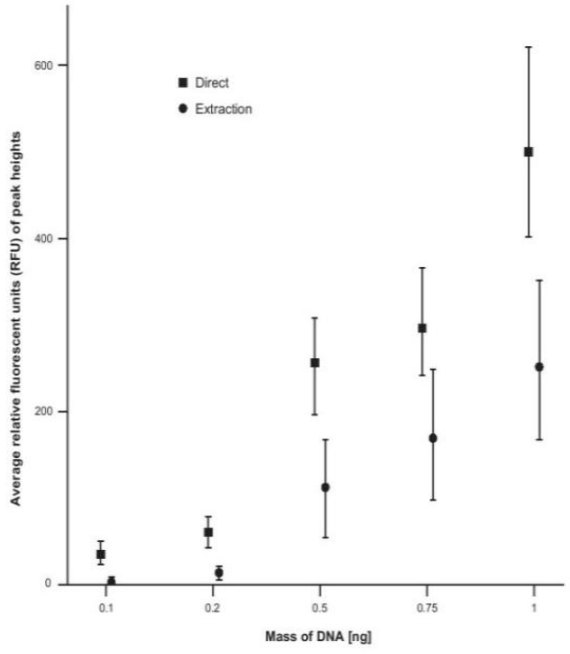


FIG. 1—*Estimated marginal RFU means from GLM procedure along with bootstrapped 95% confidence intervals (CI) displaying the interaction between method and mass of DNA, across three substrates (plastic, glass, and brass).*

Effect of Substrate

The nature of substrate (brass, glass, or plastic) on which the DNA was deposited prior to swabbing had a statistically significant effect on the results ($p = 0.003$). The average RFU value of peak heights obtained from glass was 125 RFU (95% CI: 85–159 RFU) which was lower than for plastic 211 RFU (95% CI: 167–248 RFU) ($p = 0.01$), but not significantly lower as compared to brass 172 RFU (95% CI: 122–225 RFU).

Discussion

The results indicate that a direct approach has the potential to significantly improve the profiling results from "low-level" DNA samples by increasing the yield of PCR product, negating the need for increased PCR cycle number. Our data indicate that direct PCR has the ability to effectively amplify low quantities of control DNA recovered from brass, glass, and plastic substrates. Samples and substrates used in this study did not reflect all possible scenarios of contact DNA left behind in forensic investigations (e.g., damaged, degraded, and intact cellular DNA). We have previously reported that the direct PCR method can be used to generate full DNA profiles from fingermarks deposited 2 h after individuals washed their hands (15). The direct PCR approach in this study led to significantly higher peak heights of alleles ($p = 0.001$) compared to the extraction experiment when controlling for DNA mass and substrate (See Table S1). Our results indicate on average a lower DNA profiling success for all extracted swabs (see Fig. 2) in comparison with the direct PCR swabs for DNA amounts <0.75 ng, which may account for a loss of DNA during the extraction process (6). A previous study has reported that direct PCR can minimize allele dropout in samples with low PCR yield (13).

During swabbing, the tip of a swab can be used to collect material adhering to a surface, allowing a targeted area to become saturated in DNA. Our direct PCR approach utilizes the tip of the swab, where residual DNA remains entrapped, by placing a 2-mm^2 fiber clump directly into a PCR tube to maximize the amount of template DNA available in the reaction. Multiple fiber clumps taken from the same swab will enable repeat analyses or the number of genetic marker tests to be maximized. However, the efficacy of swab fibers for recovering DNA may largely be dependent on the nature of the substrate that is swabbed (e.g., porous or nonporous) (17). All substrates tested in this study were nonporous in nature; however, brass was chosen as it is also known to be a difficult substrate for DNA recovery (22–24). Nylon-flocked fibers demonstrate the ability to capture more DNA, as assayed by average RFU of peak heights, from plastic compared to glass and brass substrates (data not shown). This highlights the need for other collection devices to be considered and in-house validation studies carried out to assess the efficacy of certain swab types for recovering DNA from a wide range of substrates. Inhibitors present in brass (such as copper and zinc ions) may also contribute to the overall lower yields of PCR product seen from this substrate (see Fig. 2). However, direct PCR offered an improvement in profiling success for brass (see Fig. 2), emphasizing that specific direct PCR multiplex kits or a standard extraction process is not necessarily required to remove the inhibitors prior to PCR.

Sacrificial binding of DNA to polypropylene tubes (25,26), and in our opinion multiple wash steps during the extraction process, may result in a reduced amount of template DNA available for the subsequent PCR. This irreversible binding of low-template

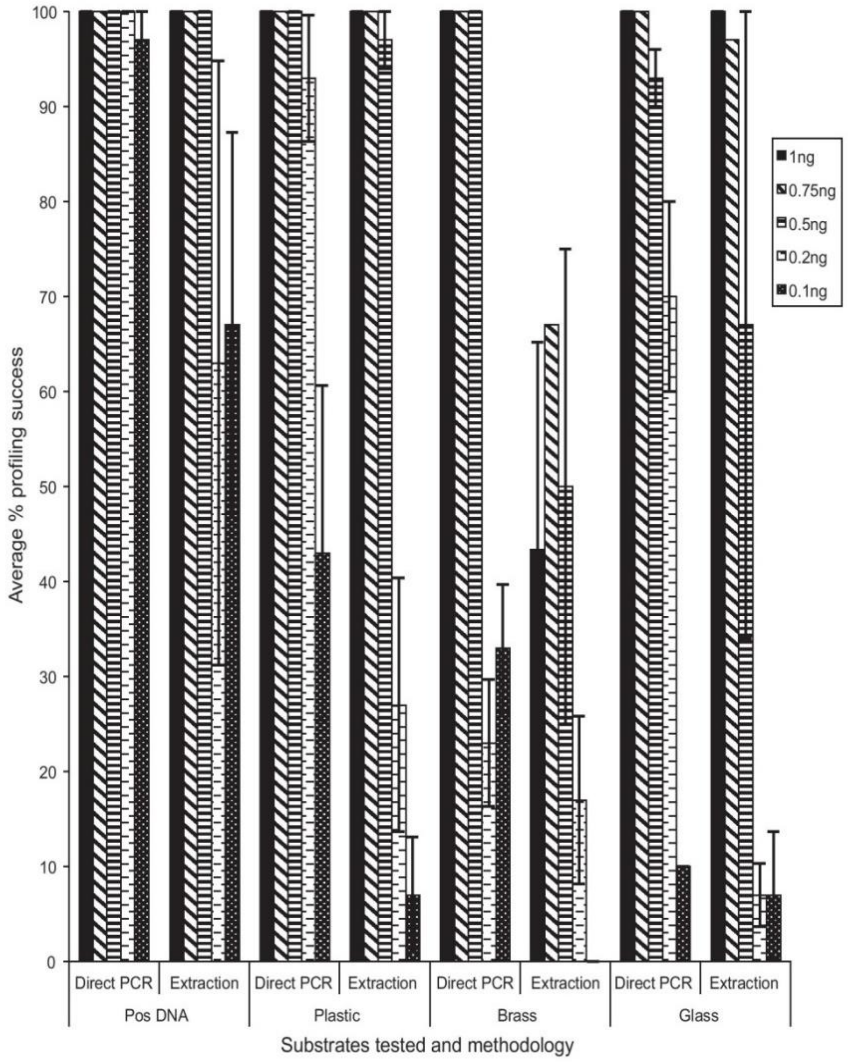


FIG. 2—*Comparison of STR-based DNA profiling success between direct PCR swabs* (n = *45) and DNA IQ™ extracted swabs* (n = *45). Bars indicate the difference in profiling success in relation to the initial mass of DNA deposited. Error bars showing ± standard error of mean.*

DNA to plasticware can be minimized by circumventing the extraction step and the subsequent multitube changes that follow.

The presence of a nondenaturing detergent in the PCR buffer and the inclusion of a 95°C hold step during the initial stages of PCR are hypothesized to lyse intact cells (11) and release DNA into solution without the need for a prior extraction process. Triton X can aid cell lysis due to its hydrophobic and hydrophilic-associated properties (4), and because of its high viscosity, it may prevent DNA from adhering to the walls of the reaction vessel. Preliminary data (unpublished) indicate a higher success rate of swabs moistened with Triton X in comparison with water. Ancient and low-level DNA laboratories are now routinely adding dilute nondenaturing detergents such as Triton X or Tween 20 to their precious DNA extracts or investing in expensive lo-bind DNA tubes in order to prevent sample DNA from binding to the walls of the reaction vessels (27,28).

Additionally, the presence of cell-free DNA as a component of "touch" DNA deserves further consideration when processing biological samples of this nature in order to maximize profiling success. Most extraction protocols do not utilize the cell-free nucleic acid of a sample; the aqueous portion of the DNA extract is generally discarded in the supernatant after centrifugation. Direct PCR does not discard the cell-free DNA of a sample and takes advantage of this component of "touch" DNA by making it available to the PCR.

Two common concerns when dealing with trace amounts of DNA in casework samples have the potential to be eliminated or reduced by implementing the direct PCR approach: (i) typical extraction protocols utilize a multistep, multitube approach, introducing the potential for sample mix-up during processing; and (ii) the potential to introduce extraneous DNA into the reaction is minimized. The ability to proceed directly to the PCR when processing low-level DNA swabs typical of "touch" DNA samples has the potential to improve greatly the overall profiling success.

References

1. Van Oorschot RA, Jones MK. DNA fingerprints from fingerprints. Nature 1997;387:767–7.
2. Quinones I, Daniel B. Cell free DNA as a component of forensic evidence recovered from touched surfaces. Forensic Sci Int Genet 2012;6 (1):26–30.

3. Wickenheiser RA. Trace DNA: a review, discussion of theory, and application of the transfer of trace quantities of DNA through skin contact. J Forensic Sci 2002;47(3):442–50.
4. Thomasma SM, Foran DR. The influence of swabbing solutions on DNA recovery from touch samples. J Forensic Sci 2013;58(2):465–9.
5. Van Oorschot RAH, Phelan DG, Furlong S, Scarfo GM, Holding NL, Cummins MJ. Are you collecting all the available DNA from touched objects? Int Congress Series 2003;1239:803–7.
6. Balogh MK, Burger J, Bender K, Schneider PM, Alt KW. STR genotyping and mtDNA sequencing of latent fingerprint on paper. Forensic Sci Int 2003;137(2–3):188–95.
7. Shaw KJ, Thain L, Docker PT, Dyer CE, Greenman J, Greenway GM, et al. The use of carrier RNA to enhance DNA extraction from microfluidic-based silica monoliths. Anal Chim Acta 2009;652(1):231–3.
8. Kishore R, Reef Hardy W, Anderson VJ, Sanchez NA, Buoncristiani MR. Optimization of DNA Extraction from low-yield and degraded samples using the BioRobot® EZ1 and BioRobot® M48. J Forensic Sci 2006;51(5):1055–61.
9. Vandewoestyne M, Van Hoofstat D, Franssen A, Van Nieuwerburgh F, Deforce D. Presence and potential of cell free DNA in different types of forensic samples. Forensic Sci Int Genet 2013;7(2):316–20.
10. Kita T, Yamaguchi H, Yokoyama M, Tanaka T, Tanaka N. Morphological study of fragmented DNA on touched objects. Forensic Sci Int Genet 2008;3(1):32–6.
11. Linacre A, Pekarek V, Swaran YC, Tobe SS. Generation of DNA profiles from fabrics without DNA extraction. Forensic Sci Int Genet 2010;4(2):137–41.
12. Park SJ, Kim JY, Yang YG, Lee SH. Direct STR amplification from whole blood and blood-or saliva-spotted FTA® without DNA purification. J Forensic Sci 2008;53(2):335–41.
13. Sim JE, Park SJ, Lee HC, Kim SY, Kim JY, Lee SH. High-throughput STR analysis for DNA database using direct PCR. J Forensic Sci 2013;58(4):989–92.
14. Ottens R, Taylor D, Abarno D, Linacre A. Optimising direct PCR from anagen hair samples. Forensic Sci Int Genet Suppl Series 2013;4(1):e109–10.
15. Templeton J, Ottens R, Paradiso V, Handt O, Taylor D, Linacre A. Genetic profiling from challenging samples: direct PCR of touch DNA. Forensic Sci Int Genet Suppl Series 2013;4(1):e224–5.
16. Swaran YC, Welch L. A comparison between direct PCR and extraction to generate DNA profiles from samples retrieved from various substrates. Forensic Sci Int Genet 2012;6(3):407–12.
17. Verdon TJ, Mitchell RJ, van Oorschot RA. The influence of substrate on DNA transfer and extraction efficiency. Forensic Sci Int Genet 2013;7(1):167–75.
18. Lowe A, Murray C, Whitaker J, Tully G, Gill P. The propensity of individuals to deposit DNA and secondary transfer of low level DNA from individuals to inert surfaces. Forensic Sci Int 2002;129(1):25–34.
19. IBM_Corp R. IBM SPSS statistics for windows. Armonk, NY: IBM Corp, 2010.
20. Wehrens R, Putter H, Buydens L. The bootstrap: a tutorial. Chemometr Intell Lab Syst 2000;54(1):35–52.
21. Manly BF. Randomization, bootstrap and Monte Carlo methods in biology, 3rd edn. Boca Raton, FL: CRC Press, 2006.
22. Gashi B, Edwards M, Sermon P, Courtney L, Harrison D, Xu Y. Measurement of 9 mm cartridge case external temperatures and its forensic application. Forensic Sci Int 2010;200(1):21–7.
23. Xu Y, Harrison D, Hay G, Hopwood A, Gashi B, Edwards M, et al. Design and manufacture of surface textures on gun cartridge cases to trap DNA material. Proc Inst Mech Eng Part B J Eng Manuf 2010;224(8):1229–38.
24. Horsman-Hall KM, Orihuela Y, Karczynski SL, Davis AL, Ban JD, Greenspoon SA. Development of STR profiles from firearms and fired cartridge cases. Forensic Sci Int Genet 2009;3(4):242–50.
25. Belotserkovskii BP, Johnston BH. Denaturation and association of DNA sequences by certain polypropylene surfaces. Anal Biochem 1997;251(2):251–62.
26. Gaillard C, Strauss F. Eliminating DNA loss and denaturation during storage in plastic microtubes. Am Clin Lab 2001;20(2):52–4.
27. Brotherton P, Haak W, Templeton J, Brandt G, Soubrier J, Adler CJ, et al. Neolithic mitochondrial haplogroup H genomes and the genetic origins of Europeans. Nat Commun 2013;4:1764.
28. Rohland N, Hofreiter M. Comparison and optimization of ancient DNA extraction. Biotechniques 2007;42(3):343.

Additional information and reprint requests:
Jennifer E. L. Templeton, M.Sc.
The School of Biological Sciences
Flinders University
Bedford Park
Adelaide
SA, Australia, 5042
E-mail: jennifer.templeton@flinders.edu.au

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Table S1. Three-way analysis of variance (ANOVA) for average relative fluorescent units (RFU) as a function of concentration, substrate and method.

3.6 Extraction efficiency

A recent study was published investigating the ability of cotton and nylon FLOQSwabs™ to recover DNA after using three different extraction methodologies (i.e. QIAamp DNA investigator kit, BioRobot EZ1 and the QIAcube) [14]. Nylon FLOQSwabs™ retrieved significantly more DNA than cotton swabs, although the success of swabbing was based on the extraction platform used [14]. By eliminating the extraction protocol variable results are thought to be more consistent. Additional results that reveal a loss of DNA through standard extraction protocols are published in the manuscript enclosed: “Ottens, R., **Templeton, J.,** Paradiso, V., Taylor, D., Abarno, D., & Linacre, A. (2013). Application of direct PCR in forensic casework. Forensic Science International: Genetics, Supplement Series, 4(1), e47-e48”.

3.6.1 Statement of authorship

Title of publication: Application of direct PCR in forensic casework

Manuscript published in the Journal of the Forensic Sciences

Renée Ottens

Carried out lab work, data analysis, and wrote the paper

I hereby certify that the statement of contribution is accurate.

Signed: Date: November 2016

Jennifer Templeton (Candidate)

Contributed to the experimental design and lab work and helped to edit the paper.

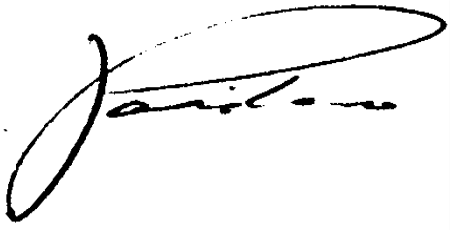
I hereby certify that the statement of contribution is accurate.

Signed: Date: November 2016

Viviana Paradiso

Helped to edit the manuscript

I hereby certify that the statement of contribution is accurate.



Signed: Date: November 2016

Duncan Taylor

Helped to edit the manuscript

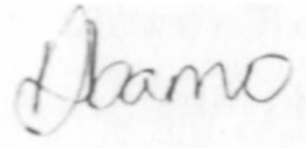
I hereby certify that the statement of contribution is accurate.

Signed: Date: November 2016

Damien Abarno

Helped to edit the manuscript

I hereby certify that the statement of contribution is accurate.



Signed: Date: November 2016

Adrian Linacre

Contributed to experimental design and helped to edit the manuscript

I hereby certify that the statement of contribution is accurate.

Signed: Date: November 2016

3.6.2 Manuscript: Application of direct PCR in forensic casework.

Forensic Science International: Genetics Supplement Series 4 (2013) e47–e48

Contents lists available at ScienceDirect

Forensic Science International: Genetics Supplement Series

journal homepage: www.elsevier.com/locate/FSIGSS

Application of direct PCR in forensic casework

Renée Ottens [a,*], Jennifer Templeton [a], Viviana Paradiso [a], Duncan Taylor [b], Damien Abarno [b], Adrian Linacre [a]

[a] School of Biological Sciences, Flinders University, Australia
[b] Forensic Science South Australia, Adelaide, Australia

ARTICLE INFO

Article history:
Received 30 August 2013
Accepted 2 October 2013

Keywords:
Direct PCR
Latent DNA
Human identification
Short tandem repeat
Profiling

ABSTRACT

Direct PCR is fast becoming a popular method in forensic science due to the advantages of saving time and money in the lab while increasing the probability of obtaining substantial results has a positive rippling effect. A laboratory is able to reduce the time spent on processing trace DNA samples, which can lead to investigators receiving important information in a timely manner that may not have been possible using standard methods. This study highlights the benefits of direct PCR in forensic casework by analysing trace and touch DNA on a range of substrates and exploring the loss of initial DNA due to extraction.

1. Introduction

Direct PCR can generate full STR DNA profiles from trace samples without using extraction or quantification processes [1]. A previous study reports a loss of up to 76% of initial DNA due to the multiple wash steps and tube changes involved in the extraction process [2], ultimately affecting the quality of the DNA profile obtained. Samples containing less than 100 pg are considered to be trace DNA [3], and generally require an increase to 34 cycles to obtain meaningful data. This study highlights direct PCR sensitivity and the ability for trace DNA to be amplified without the need to increase cycle number or modify current protocols to obtain meaningful data. A direct PCR approach is a viable option for the future of trace DNA recovery and analysis for forensic science purposes, vastly improving efficiency, sensitivity and the quality of results.

2. Materials and method

2.1. Determining loss of DNA via extraction

QIAamp DNA Micro Kit (QIAGEN) and DNA IQ™ System (Promega) were used to determine the average loss of DNA. A total of 20 ng of control DNA (2800 M, Promega) was used as the starting concentration for each extraction. Extractions were performed followed the manufacturer's instructions for each kit and were eluted in a final volume of 30 μL for maximum DNA yield. Extractions were repeated five times for each kit used. Each extraction sample was quantified on a Qubit® 2.0 Fluorometer (Invitrogen™) following manufacturer's instructions, using 5 μL from each sample for quantification. The quantification results will be used to determine the loss of DNA. The quantification results will be compared against the initial input DNA to determine percentage lost.

2.2. Direct PCR

DNA extraction and quantification are bypassed. The range of samples tested include: single hairs in anagen and telogen growth phases, fibres from swabs used on plastic and glass surfaces, single fibres from clothing, latex gloves, and a range of plastic tapes. Each sample is placed directly into a prepared 0.2 mL thin walled tube containing 25 μL of reagents from either the AmpFℓSTR® Profiler Plus® or NGM™ kit (Life Technologies, Victoria, Australia). Sample sizes were approximately 5 mm in length for hair and fibres, and 5 mm × 5 mm for tapes, swabs and gloves. Amplification conditions for all samples followed the manufacturer's instructions at 29 cycles. PCR was performed on a GeneAmp® 9700 96-well thermal cycler (Applied Biosystems) following manufacturer's instructions. PCR products were analysed using a 3130*xl* (ABI) and GeneMapper® ID v3.2 with a threshold of 50 RFU for allele assignment.

2.3. Limit of detection for PCR

A series of positive control (DNA 2800M) samples were created to determine the lowest concentration of DNA needed to produce a

* Corresponding author at: School of Biological Sciences, Flinders University, Bedford Park, South Australia 5001, Australia. Tel.: +61 8 82015003.
E-mail address: renee.ottens@flinders.edu.au (R. Ottens).

1875-1768/$ – see front matter © 2013 Elsevier Ireland Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.fsigss.2013.10.024

Table 1
DNA lost from standard extraction.

DNA extraction kit	Starting concentration (ng)	Average final concentration (ng/30 μL)	Average percentage lost (%)
Promega IQ	20	3.3	83.5
QIAGEN Micro	20	5.7	71.5

full DNA profile. Positive control PCRs were set up in triplicates in the following DNA concentrations: 0.05 0.1, 0.2, 0.3, 0.4, 0.5, 0.7, 0.8, 0.9, and 1 ng. Each sample was amplified and analysed in the same way as the direct PCR method, except 1 μL of each DNA concentration was used.

3. Results and discussion

3.1. Loss of DNA

The QIAGEN Micro DNA extraction kit yielded a higher DNA concentration overall, retaining on average 28.5% of the original amount. Although the QIAGEN kit performed better, both methods lost a substantial amount of DNA ranging from 71.5% to 83.5% (Table 1). Validated STR kits recommend using 1 ng of input DNA to optimise PCR, thus from a 30 μL elution, 9.1 μL would be required from the Promega kit and 5.3 μL from the QIAGEN kit. The average DNA lost when combining both kits is 77.5%.

3.2. Direct PCR

Profiles 'up-loadable' (≥12 alleles) to the Australian National DNA Database were obtained from the following substrates:anagen and telogen hairs, single fibres from worn or touched fabric, plastic tape, and single fibres from a range of swabs used on glass and plastic surfaces. Less than 12 alleles were observed from latex glove samples.

3.3. Limit of detection

Using standard PCR cycling conditions, the lowest initial template concentration of DNA at which a full DNA profile could be obtained was 100 pg. With a reduction to 50 pg, up-loadable profiles were obtained, with some allelic drop-out occurring. A single diploid human cell contains ~6 pg of DNA [4], therefore 100 pg equates to ~17 cells (assuming there is no contribution from cell-free DNA). If a sample lost 80% of its DNA via an extraction method (based on a 30 μL elution), a starting concentration of 500 pg (~84 cells) is needed in order to obtain a template of 100 pg for the PCR, and hence a full profile. However, most kits allow up to a maximum of 10 μL of input DNA for a standard 25 μL reaction. Therefore, to maintain this reaction volume size, the true starting DNA concentration using standard extraction procedures needs to be approximately 1.5 ng (250 cells), with an extraction loss no greater than 80%.

4. Conclusion

As DNA is not lost during an extraction process when using direct PCR methods, only 17 cells are required to obtain a full STR profile compared to 250 cells when using standard DNA extraction methods; thus making direct PCR up to ~15 times more sensitive than conventional extraction methods. No modifications are made to STR kit protocols allowing for quick and easy implementation into forensic laboratories. The results obtained by this laboratory are encouraging when typing trace amounts of DNA, and potentially are a vast improvement when compared to current industry methods.

Role of funding

Funding was provided by the Department of Justice, South Australia.

Conflict of interest

None.

Acknowledgements

We thank members of Forensic Science South Australia and the School of Biological Sciences at Flinders University for assisting with reagents and providing helpful advice.

References

[1] R. Ottens, D. Taylor, D. Abarno, A. Linacre, Successful direct amplification of nuclear markers from a single hair follicle, Forensic Science, Medicine, and Pathology 9 (2013) 238–243.
[2] R.A.H. van Oorschot, D.G. Phelan, S. Furlong, G.M. Scarfo, N.L. Holding, M.J. Cummins, Are you collecting all the available DNA from touched objects, International Congress Series 1239 (2003) 803–807.
[3] P. Gill, J. Whitaker, C. Flaxman, N. Brown, J. Buckleton, An investigation of the rigor of interpretation rules for STRs derived from less than 100 pg of DNA, Forensic Science International 112 (2000) 17–40.
[4] J.M. Butler, Forensic DNA Typing, second ed., Elsevier, 2005.

3.7 Conclusion

Current swabs used for collecting and extracting ‘touch’ DNA are inefficient; template DNA most likely becomes trapped within the inner core of the swab matrix. This residually retained DNA is not always efficiently released following an extraction step. Direct PCR ensures that all DNA template would be made available to the reaction if the fibres are added directly to the PCR tube. Improving the ability to capture more ‘touch’ DNA from a substrate would increase the chance of providing probative value to forensic investigations. By using a ‘targeted’ swabbing approach and positively charged fibres that are saturated in Triton™-X (0.1%) detergent there is the potential to increase the amount of DNA recovered. This is particularly important if there is a small amount of DNA available for collection in the first place. Detergent-based nylon FLOQSwabs™ offer a potential solution for processing low-template DNA samples that may otherwise yield insufficient DNA for conventional STR typing protocols. The results presented in this study report the use of optimised swabbing and direct PCR when using control DNA deposited on various substrates. The research to follow in Chapter IV explores the use of direct PCR for processing latent DNA from touched items. The method has the potential to be highly informative if there is an unknown mark found on an item to which an immediate match is required on a DNA database. Additionally, direct PCR has the advantage of speeding up the processing of swabs for high-throughput casework.

3.8 References

1. *Nucleic acid collection device and method - AU2016902269*. 2016; Patent application. Available from: http://www.ipaustralia.com.au/applicant/flinders-university-of-south-australia/patents/AU2016902269/.
2. Gill, P., et al., *An investigation of the rigor of interpretation rules for STRs derived from less than 100 pg of DNA*. Forensic Science International, 2000. **112**(1): p. 17-40.
3. Hennessy, L.K., et al., *Developmental validation of the GlobalFiler® express kit, a 24-marker STR assay, on the RapidHIT® System.* Forensic Science International: Genetics, 2014. **13**: p. 247-258.
4. Martín, P., et al., *Improving DNA data exchange: validation studies on a single 6 dye STR kit with 24 loci.* Forensic Science International: Genetics, 2014. **13**: p. 68-78.
5. Green, R.L., et al., *Developmental validation of the AmpFℓSTR® NGM SElect™ PCR Amplification Kit: a next-generation STR multiplex with the SE33 locus.* Forensic Science International: Genetics, 2013. **7**(1): p. 41-51.
6. Harbison, S., M. Fallow, and D. Bushell, *An analysis of the success rate of 908 trace DNA samples submitted to the Crime Sample Database Unit in New Zealand.* Australian Journal of Forensic Sciences, 2008. **40**(1): p. 49-53.
7. Raymond, J.J., et al., Trace DNA success rates relating to volume crime offences. Forensic Science International: Genetics Supplement Series, 2009. **2**(1): p. 136-137.
8. van Oorschot, R.A., K.N. Ballantyne, and R.J. Mitchell, *Forensic trace DNA: a review.* Investigative Genetics, 2010. **1**(1): p. 1.
9. Budowle, B. and A. van Daal, *Extracting evidence from forensic DNA analyses: Future molecular biology directions.* Biotechniques, 2009. **46**(5): p. 339.
10. Verdon, T.J., R.J. Mitchell, and R.A. Oorschot, *Swabs as DNA collection devices for sampling different biological materials from different substrates.* Journal of Forensic Sciences, 2014. **59**(4): p. 1080-1089.
11. Verdon, T.J., R.J. Mitchell, and R.A. van Oorschot, *The influence of substrate on DNA transfer and extraction efficiency.* Forensic Science International: Genetics, 2013. **7**(1): p. 167-175.

12. Mawlood, S.K., M. Alrowaithi, and N. Watson, *Advantage of ForensiX Swabs in Retrieving and Preserving Biological Fluids.* Journal of Forensic Sciences, 2015. **60**(3): p. 686-689.
13. Marshall, P.L., et al., *Evaluation of a novel material, Diomics X-Swab™, for collection of DNA.* Forensic Science International: Genetics, 2014. **12**: p. 192-198.
14. Brownlow, R.J., K.E. Dagnall, and C.E. Ames, *A comparison of DNA collection and retrieval from two swab types (cotton and nylon flocked swab) when processed using three QIAGEN extraction methods.* Journal of Forensic Sciences, 2012. **57**(3): p. 713-717.
15. Prinz, M., et al., *Maximization of STR DNA typing success for touched objects.* International Congress Series, 2006. **1288**: p. 651-653.
16. Sprecher, C.J., et al., *PowerPlex® ESX and ESI Systems: A suite of new STR systems designed to meet the changing needs of the DNA-typing community.* Forensic Science International: Genetics Supplement Series, 2009. **2**(1): p. 2-4.
17. Wetton, J.H., et al., *Analysis and interpretation of mixed profiles generated by 34 cycle SGM Plus® amplification.* Forensic Science International: Genetics, 2010. **5**(5): p. 376-380.
18. Whitaker, J., E. Cotton, and P. Gill, *A comparison of the characteristics of profiles produced with the AMPFISTR® SGM Plus™ multiplex system for both standard and low copy number (LCN) STR DNA analysis.* Forensic Science International, 2001. **123**(2): p. 215-223.
19. Mulligan, C.M., S.R. Kaufman, and L. Quarino, *The Utility of Polyester and Cotton as Swabbing Substrates for the Removal of Cellular Material from Surfaces.* Journal of Forensic Sciences, 2011. **56**(2): p. 485-490.
20. Hansson, O., et al., *Trace DNA collection—Performance of minitape and three different swabs.* Forensic Science International: Genetics Supplement Series, 2009. **2**(1): p. 189-190.
21. Linacre, A., et al., *Generation of DNA profiles from fabrics without DNA extraction.* Forensic Science International: Genetics, 2010. **4**(2): p. 137-141.
22. Hansora, D., N. Shimpi, and S. Mishra, *Performance of hybrid nanostructured conductive cotton materials as wearable devices: an overview of materials,*

fabrication, properties and applications. RSC Advances, 2015. **5**(130): p. 107716-107770.

23. McCall, E.R. and J.F. Jurgens, *Chemical composition of cotton.* Textile Research Journal, 1951. **21**(1): p. 19-21.
24. Rudnik, E., *10 Compostable Polymer Materials: Definitions, Structures, and Methods of Preparation.* Handbook of Biopolymers and Biodegradable Plastics: Properties, Processing and Applications, 2012: p. 189.
25. Daley, P., et al., *Comparison of flocked and rayon swabs for collection of respiratory epithelial cells from uninfected volunteers and symptomatic patients.* Journal of Clinical Microbiology, 2006. **44**(6): p. 2265-2267.
26. Dalmaso, G., et al., *Qualification of high-recovery, flocked swabs as compared to traditional rayon swabs for microbiological environmental monitoring of surfaces.* PDA Journal of Pharmaceutical Science and Technology, 2008. **62**(3): p. 191.
27. *Copan flock technologies*. 2016; Available from: http://www.copanflocktech.com/index.php/prod/flockedswabs/.
28. Sigma-Aldrich. 2016; Available from: http://www.sigmaaldrich.com/catalog/product/sigma/t8787?lang=en&region=AU.
29. Sweet, D., et al., *An improved method to recover saliva from human skin: the double swab technique.* Journal of Forensic Sciences, 1997. **42**: p. 320-322.
30. Valenzuela, A., et al., *PCR-based DNA typing of saliva stains recovered from human skin.* Journal of Forensic Science, 1997. **42**(3): p. 447-451.
31. Pang, B.C. and B.K. Cheung, *Double swab technique for collecting touched evidence.* Legal Medicine (Tokyo), 2007. **9**(4): p. 181-4.
32. van Oorschot, R.A.H., et al., *Are you collecting all the available DNA from touched objects?* International Congress Series, 2003. **1239**(0): p. 803-807.
33. Goray, M., R.A. van Oorschot, and J.R. Mitchell, *DNA transfer within forensic exhibit packaging: Potential for DNA loss and relocation.* Forensic Science International: Genetics, 2012. **6**(2): p. 158-166.
34. Raymond, J.J., et al., *Trace DNA and street robbery: a criminalistic approach to DNA evidence.* Forensic Science International: Genetics Supplement Series, 2009. **2**(1): p. 544-546.

35. Graham, E.A.M. and G.N. Rutty, *Investigation into "normal" background DNA on adult necks: implications for DNA profiling of manual strangulation victims.* Journal of Forensic Sciences, 2008. **53**(5): p. 1074-1082.
36. Phetpeng, S., et al., *Touch DNA collection from improvised explosive devices: A comprehensive study of swabs and moistening agents.* Forensic Science International: Genetics Supplement Series, 2013. **4**(1): p. e29-e30.
37. Thomasma, S.M. and D.R. Foran, *The influence of swabbing solutions on DNA recovery from touch samples.* Journal of Forensic Sciences, 2013. **58**(2): p. 465-469.
38. Giorgio, G., et al., *The impact of alkanes on the structure of Triton X100 micelles.* RSC Advances, 2016. **6**(1): p. 825-836.
39. Nanassy, O.Z., P.V. Haydock, and M.W. Reed, *Capture of genomic DNA on glass microscope slides.* Analytical Biochemistry, 2007. **365**(2): p. 240-245.
40. Park, S.J., et al., *Direct STR Amplification from Whole Blood and Blood-or Saliva-Spotted FTA® without DNA Purification*. Journal of Forensic Sciences, 2008. **53**(2): p. 335-341.
41. Ottens, R., et al., *Optimising direct PCR from anagen hair samples.* Forensic Science International: Genetics Supplement Series, 2013. **4**(1): p. e109-e110.
42. Swaran, Y.C. and L. Welch, *A comparison between direct PCR and extraction to generate DNA profiles from samples retrieved from various substrates*. Forensic Science International: Genetics, 2012. **6**(3): p. 407-412.
43. Sim, J.E., et al., *High-Throughput STR Analysis for DNA Database Using Direct PCR.* Journal of Forensic Sciences, 2013. **58**(4): p. 989-992.
44. Altshuler, H. and R. Roy, *Evaluation of direct PCR amplification using various swabs and washing Reagents.* Journal of Forensic Sciences, 2015. **60**(6): p. 1542-1552.
45. Rohland, N. and M. Hofreiter, *Ancient DNA extraction from bones and teeth.* Nature protocols, 2007. **2**(7): p. 1756-1762.
46. Quinones, I. and B. Daniel, *Cell free DNA as a component of forensic evidence recovered from touched surfaces.* Forensic Science International: Genetics, 2012. **6**(1): p. 26-30.

47. Vandewoestyne, M., et al., *Presence and potential of cell free DNA in different types of forensic samples.* Forensic Science International Genetics, 2013. **7**(2): p. 316-20.
48. Adamowicz, M.S., et al., *Evaluation of Methods to Improve the Extraction and Recovery of DNA from Cotton Swabs for Forensic Analysis*. PLOS ONE, 2014. **9**(12): p. e116351.
49. Phillips, K., N. McCallum, and L. Welch, *A comparison of methods for forensic DNA extraction: Chelex-100® and the QIAGEN DNA Investigator Kit (manual and automated).* Forensic Science International: Genetics, 2012. **6**(2): p. 282-285.
50. Barbaro, A., et al., *DNA profiling by different extraction methods.* International Congress Series, 2004. **1261**: p. 562-564.
51. Ottens, R., et al., *Application of direct PCR in forensic casework.* Forensic Science International: Genetics Supplement Series, 2013. **4**(1): p. e47-e48.
52. Hoff-Olsen, P., et al., *Extraction of DNA from decomposed human tissue: an evaluation of five extraction methods for short tandem repeat typing.* Forensic Science International, 1999. **105**(3): p. 171-183.
53. van Oorschot, R.A. and M.K. Jones, *DNA fingerprints from fingerprints.* Nature London, 1997: p. 767-767.
54. Höss, M. and S. Pääbo, *DNA extraction from Pleistocene bones by a silica-based purification method.* Nucleic Acids Research, 1993. **21**(16): p. 3913-3914.
55. Idris, B. and W. Goodwin, *Comparison of Chelex®-100 with two solid phase DNA extraction techniques.* Forensic Science International: Genetics Supplement Series, 2015. **5**: p. e274-e275.
56. Brotherton, P., et al., *Neolithic mitochondrial haplogroup H genomes and the genetic origins of Europeans.* Nature Communications, 2013. **4**: p. 1764.
57. IBM_Corp, R., *IBM SPSS Statistics for Windows.* IBM Corp, Armonk, NY, 2010.

Appendices - Chapter III

Appendix A - Supplementary Information

Experiment 3.1) Comparison of swab types for direct PCR

The threshold for assigning alleles was 30 RFU in this preliminary study that compared swab types, swab media and technique. This threshold is considered lower than what is generally used in case work laboratories; a threshold used for later studies in this thesis was 3 standard deviations above the baseline to account for background noise.

Results - preliminary work to supplement manuscript data

SI Table 3.1 Example of STR DNA profiling results for swab comparison test, using control DNA at 1 ng as template for swabbing and amplification using direct PCR (n=3).

GENOTYPE PROFILE			**GENOTYPE PROFILE**			**GENOTYPE PROFILE**		
Sample ID: NYLON			**Sample ID: COTTON**			**Sample ID: FOAM**		
Genetic loci	Allele (and RFU)	Allele (and RFU)	**Genetic loci**	Allele (and RFU)	Allele (and RFU)	**Genetic loci**	Allele (and RFU)	Allele (and RFU)
D3S1358	17 (90)	18 (51)	**D3S1358**	F	F	**D3S1358**	F	F
VWA	16 (71)	19 (44)	**VWA**	F	F	**VWA**	F	F
FGA	20 (33)	23 (46)	**FGA**	F	F	**FGA**	F	F
Amel	X (111)	Y (81)	**Amel**	F	F	**Amel**	F	F
D8S1179	14 (50)	15 (70)	**D8S1179**	F	F	**D8S1179**	F	F
D21S11	29 (59)	31.2 (30)	**D21S11**	F	F	**D21S11**	F	F
D18S51	F	F	**D18S51**	F	F	**D18S51**	F	F
D5S818	12 (94)	F	**D5S818**	F	F	**D5S818**	F	F
D13S317	9 (30)	11 (27)	**D13S317**	F	F	**D13S317**	F	F
D7S820	F	F	**D7S820**	F	F	**D7S820**	F	F

N.B. **Relative fluorescent units highlighted in brackets in blue indicate the highest RFU values for all samples tested. Allele drop-out is denoted by 'F'.**

SI Table 3.2 Example of STR DNA profiling results for swab fibre inhibition test, using control DNA at 1 ng deposited directly onto swab fibres and subjected to direct PCR. **No swabbing.**

GENOTYPE PROFILE			GENOTYPE PROFILE			GENOTYPE PROFILE		
Sample ID: NYLON			Sample ID: COTTON			Sample ID: FOAM		
Genetic loci	Allele (and RFU)	Allele (and RFU)	**Genetic loci**	Allele (and RFU)	Allele (and RFU)	**Genetic loci**	Allele (and RFU)	Allele (and RFU)
D3S1358	17 (662)	18 (597)	**D3S1358**	F	F	**D3S1358**	17 (490)	18 (578)
VWA	16 (434)	19 (695)	**VWA**	F	F	**VWA**	16 (533)	19 (494)
FGA	20 (663)	23 (535)	**FGA**	F	F	**FGA**	20 (423)	23 (519)
Amel	X (574)	Y (556)	**Amel**	F	F	**Amel**	X (451)	Y (409)
D8S1179	14 (424)	15 (528)	**D8S1179**	F	F	**D8S1179**	14 (380)	15 (439)
D21S11	29 (464)	31.2 (518)	**D21S11**	F	F	**D21S11**	29 (565)	31.2 (428)
D18S51	16 (415)	18 (426)	**D18S51**	F	F	**D18S51**	16 (329)	18 (324)
D5S818	12 (903)	12	**D5S818**	F	F	**D5S818**	12 (661)	12
D13S317	9 (392)	11 (387)	**D13S317**	F	F	**D13S317**	9 (300)	11 (224)
D7S820	8 (277)	11 (317)	**D7S820**	F	F	**D7S820**	8 (215)	11 (234)

N.B. **Relative fluorescent units highlighted in brackets in blue indicate the highest RFU values for all samples tested. Allele drop-out is denoted by 'F'.**

Appendix B - Supplementary Information

Experiment 3.2) Comparison of swab media and swab technique

Results - preliminary work to supplement manuscript data

SI Table 3.3 Example of STR DNA profiling results for swab media comparison test, using control DNA at 1 ng as template for 'targeted' swabbing and amplification with direct PCR.

GENOTYPE PROFILE			GENOTYPE PROFILE		
Sample ID: TRITON X			**Sample ID: WATER**		
Genetic loci	Allele (and RFU)	Allele (and RFU)	**Genetic loci**	Allele (and RFU)	Allele (and RFU)
D3S1358	17 (332)	18 (376)	**D3S1358**	17 (115)	18 (115)
VWA	16 (333)	19 (309)	**VWA**	F	F
FGA	20 (228)	23 (349)	**FGA**	20 (100)	F
Amel	X (338)	Y (360)	**Amel**	X (167)	Y (113)
D8S1179	14 (263)	15 (363)	**D8S1179**	14 (108)	15 (64)
D21S11	29 (182)	31.2 (170)	**D21S11**	F	31.2 (59)
D18S51	16 (301)	18 (102)	**D18S51**	F	F
D5S818	12 (457)	12	**D5S818**	12 (122)	F
D13S317	9 (116)	11 (145)	**D13S317**	F	F
D7S820	8 (161)	11 (122)	**D7S820**	F	F

N.B. **Relative fluorescent units highlighted in brackets in blue indicate the highest RFU values for all samples tested. Allele drop-out is denoted by 'F'.**

SI Table 3.4 Example of STR DNA profiling results for swab media comparison test, using control DNA at 1 ng as template for 'whole' swabbing and amplification with direct PCR.

GENOTYPE PROFILE			**GENOTYPE PROFILE**		
Sample ID: TRITON-X			**Sample ID: WATER**		
Genetic loci	Allele (and RFU)	Allele (and RFU)	**Genetic loci**	Allele (and RFU)	Allele (and RFU)
D3S1358	17 (63)	18 (94)	**D3S1358**	F	F
VWA	F	F	**VWA**	F	F
FGA	20 (53)	23 (75)	**FGA**	F	F
Amel	X (273	Y (149)	**Amel**	X (117)	F
D8S1179	14 (91)	F	**D8S1179**	F	F
D21S11	29 (61)	F	**D21S11**	F	F
D18S51	F	F	**D18S51**	F	F
D5S818	F	F	**D5S818**	F	F
D13S317	F	F	**D13S317**	F	F
D7S820	F	F	**D7S820**	F	F

N.B. **Relative fluorescent units highlighted in brackets in blue indicate the highest RFU values for all samples tested. Allele drop-out is denoted by 'F'.**

SI Table 3.5 Example of STR DNA profiling results for swab technique test, using control DNA at 1 ng as template using Triton-X swabbing media, and amplification with direct PCR, following the NGM SElect kit™ guidelines.

GENOTYPE PROFILE			GENOTYPE PROFILE		
Sample ID: 'WHOLE' SWABBING			Sample ID: 'TARGETED' SWABBING		
Genetic loci	Allele (and RFU)	Allele (and RFU)	**Genetic loci**	Allele (and RFU)	Allele (and RFU)
D10	13 (260)	15 (297)	**D10**	13 (1649)	15 (1257)
vWA	16 (303)	19 (218)	**vWA**	16 (1175)	19 (1193)
D16	9 (183)	13 (313)	**D16**	9 (1006)	13 (905)
D2S1	22 (190)	25 (159)	**D2S1**	22 (1400)	25 (937)
Amel	X (545)	Y (426)	**Amel**	X (2247)	Y (1407)
D8	14 (297)	15 (346)	**D8**	14 (1417)	15 (2013)
D21	29 (180)	31.2 (348)	**D21**	29 (1691)	31.2 (1939)
D18	16 (115)	18 (307)	**D18**	16 (1304)	18 (1343)
D22	16 (1039)		**D22**	16 (3734)	
D19	13 (415)	14 (540)	**D19**	13 (1849)	14 (2448)
THO1	6 (426)	9.3 (495)	**THO1**	6 (2313)	9.3 (1583)
FGA	20 (265)	23 (260)	**FGA**	20 (1365)	23 (1147)
D2S4	10 (202)	14 (320)	**D2S4**	10 (898)	14 (670)
D3	17 (163)	18 (220)	**D3**	17 (831)	18 (624)
D1	12 (206)	13 (146)	**D1**	12 (750)	13 (626)
D12	18 (190)	23 (61)	**D12**	18 (513)	23 (557)
SE33	15 (77)	16 (106)	**SE33**	15 (462)	16 (410)

N.B. **Relative fluorescent units highlighted in brackets in blue indicate the highest RFU values for all samples tested. Allele drop-out is denoted by 'F'.**

Appendix C - Supplementary Information

Experiment 3.3) Direct PCR vs Extraction

Results - preliminary work to supplement manuscript data

SI Table 3.6 Preliminary results to report on DNA profiling success for direct PCR nylon FLOQswabs™ and extracted nylon FLOQswabs™, using 'whole' swabbing.

Date of run: 15.04.13 and 22.04.13		
Sample run name: JT_15.04.13 and JT_22.04.13		
Sample description	**Direct PCR Result**	**Extraction Result**
1) Neg control slide (i.e. No DNA)	Amelogenin only (XY)	No result
2) PCR blank (no DNA)	No result	No result
3) Slide nylon swab, 1ng / µL DNA	Full STR profile	No result
4) Slide nylon swab, 0.5 ng / µL DNA	2 loci (including Amelogenin)	No result
5) pos control, 1 ng / µL DNA added to PCR	Full STR profile	Full STR profile
6) pos control, 0.5 ng / µL DNA added to PCR	Full STR profile	Full STR profile
7) DNA swab pos control (buccal)	Full STR profile	Full STR profile

SI Table 3.7 Example of STR DNA profiling results for direct PCR, using control DNA at 1 ng as template for 'targeted' swabbing and direct PCR. Positive control indicates no swabbing.

GENOTYPE PROFILE			**GENOTYPE PROFILE**		
Sample ID: SWABBING – DIRECT PCR			**Sample ID: POSITIVE CONTROL DNA (i.e. No Fibres)**		
Genetic loci	Allele (and RFU)	Allele (and RFU)	**Genetic loci**	Allele (and RFU)	Allele (and RFU)
D3S1358	17 (158)	18 (173)	**D3S1358**	17 (803)	18 (748)
VWA	16 (145)	19 (202)	**VWA**	16 (940)	19 (828)
FGA	20 (98)	23 (100)	**FGA**	20 (830)	23 (679)
Amel	X (174)	Y (92)	**Amel**	X (2446)	Y (887)
D8S1179	14 (135)	15 (145)	**D8S1179**	14 (709)	15 (587)
D21S11	29 (131)	31.2 (138)	**D21S11**	29 (576)	31.2 (744)
D18S51	16 (64)	18 (147)	**D18S51**	16 (778)	18 (781)
D5S818	12 (211)	12	**D5S818**	12 (968)	12
D13S317	9 (104)	11 (53)	**D13S317**	9 (471)	11 (545)
D7S820	8 (93)	11 (89)	**D7S820**	8 (432)	11 (371)

N.B. **Relative fluorescent units highlighted in brackets in blue indicate the highest RFU values for all samples tested. Allele drop-out is denoted by 'F'.**

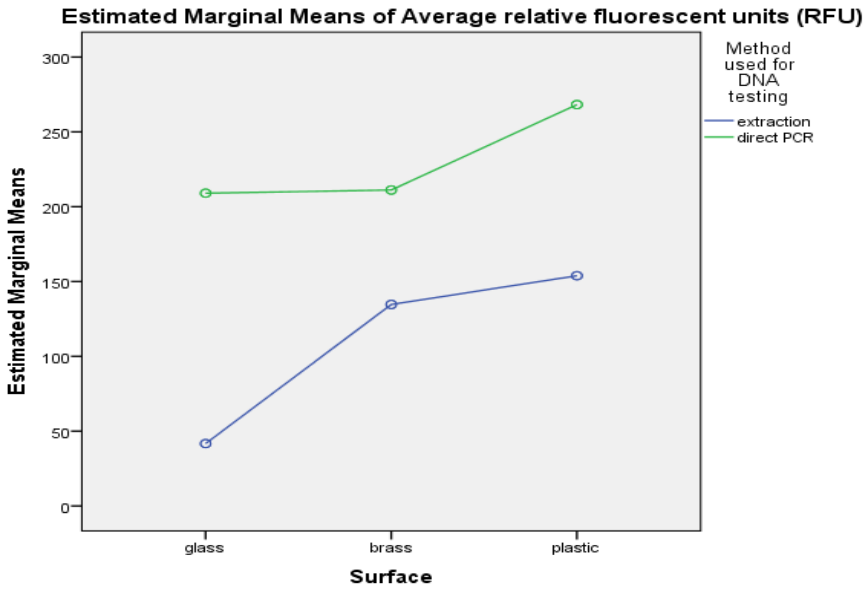
SI Table 3.8 Example of STR DNA profiling results for direct PCR, using control DNA at 0.5 ng as template for 'targeted' swabbing and direct PCR. Positive control indicates no swabbing.

GENOTYPE PROFILE			GENOTYPE PROFILE		
Sample ID: SWABBING (0.5 ng) DNA – DIRECT PCR			**Sample ID: POSITIVE CONTROL (0.5 ng) (i.e. No fibres)**		
Genetic loci	Allele (and RFU)	Allele (and RFU)	**Genetic loci**	Allele (and RFU)	Allele (and RFU)
D3S1358	17 (72)	18 (52)	**D3S1358**	17 (153)	18 (152)
VWA			**VWA**	16 (139)	19 (154)
FGA			**FGA**	F	23 (182)
Amelogenin	X (87)	Y (56)	**Amel**	X (120)	Y (69)
D8S1179			**D8S1179**	14 (96)	15 (88)
D21S11			**D21S11**	F	31.2 (84)
D18S51			**D18S51**	16 (69)	18 (59)
D5S818			**D5S818**	12 (144)	F
D13S317			**D13S317**	9 (91)	11 (62)
D7S820			**D7S820**	8 (68)	11 (51)

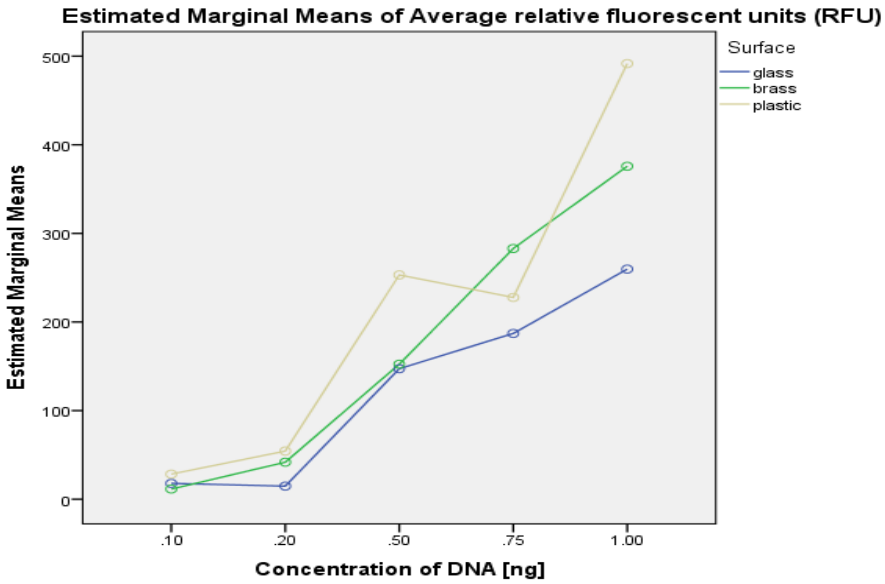
N.B. **Relative fluorescent units highlighted in brackets in blue indicate the highest RFU values for all samples tested. Allele drop-out is denoted by 'F'.**

Supplementary Information - Results - Experiment 3.3) Direct PCR vs Extraction

Data for direct PCR vs extraction were analysed using Statistical Package for Social Sciences (SPSS, IBM, Armonk, NY) [57]. The average relative fluorescent unit (RFU) for allele peak heights was used to assess the effectiveness of the direct PCR method of amplification in comparison to a standard extraction (i.e. DNA IQ™ Systems).



SI Figure 3.7 Comparison between extraction and direct PCR DNA profiling results, as assayed by RFU values of peak heights (n=90 samples).



SI Figure 3.8 Comparison between substrates used for DNA deposition, as assayed by RFU values of peak heights for (n=90 samples).

Supplementary Information - Results - Experiment 3.3) Direct PCR vs Extraction

SI Table 3. 9 Mean and standard deviation results with upper and lower confidence Intervals [57] for method of DNA testing (Extraction vs Direct PCR).

Concentration of DNA [ng] * Surface * Method used for DNA testing

Dependent Variable: Average relative fluorescent units (RFU)

Concentration of DNA [ng]	Surface	Method used for DNA testing	Mean	Std. Error	95% Confidence Interval	
					Lower Bound	Upper Bound
.10	glass	extraction	1.137E-013	53.652	-107.319	107.319
		direct PCR	35.667	53.652	-71.653	142.986
	brass	extraction	7.000	53.652	-100.319	114.319
		direct PCR	16.000	53.652	-91.319	123.319
	plastic	extraction	3.000	53.652	-104.319	110.319
		direct PCR	53.667	53.652	-53.653	160.986
.20	glass	extraction	11.667	53.652	-95.653	118.986
		direct PCR	18.000	53.652	-89.319	125.319
	brass	extraction	12.000	53.652	-95.319	119.319
		direct PCR	71.667	53.652	-35.653	178.986
	plastic	extraction	18.000	53.652	-89.319	125.319
		direct PCR	90.667	53.652	-16.653	197.986
.50	glass	extraction	54.667	53.652	-52.653	161.986
		direct PCR	239.912	53.652	132.593	347.231
	brass	extraction	140.667	53.652	33.347	247.986
		direct PCR	164.140	53.652	56.821	271.460
	plastic	extraction	141.579	53.652	34.260	248.898
		direct PCR	364.667	53.652	257.347	471.986
.75	glass	extraction	105.386	53.652	-1.933	212.705
		direct PCR	268.807	53.652	161.488	376.126
	brass	extraction	232.632	53.652	125.312	339.951
		direct PCR	333.263	53.652	225.944	440.582
	plastic	extraction	169.965	53.652	62.646	277.284
		direct PCR	285.439	53.652	178.119	392.758
1.00	glass	extraction	36.667	53.652	-70.653	143.986
		direct PCR	482.684	53.652	375.365	590.003
	brass	extraction	280.912	53.652	173.593	388.231
		direct PCR	470.667	53.652	363.347	577.986
	plastic	extraction	436.772	53.652	329.453	544.091
		direct PCR	546.509	53.652	439.190	653.828

Supplementary Information - Results
Experiment 3.3) *Direct PCR vs Extraction*

SI Table 3.10 Table of estimates for direct PCR vs Extraction
Dependent Variable: Average relative fluorescent units (RFU)

Surface	Method used for DNA testing	Mean	Std. Error	95% Confidence Interval	
				Lower Bound	Upper Bound
glass	extraction	41.677	23.994	-6.317	89.672
	direct PCR	209.014	23.994	161.019	257.009
brass	extraction	134.642	23.994	86.647	182.637
	direct PCR	211.147	23.994	163.153	259.142
plastic	extraction	153.863	23.994	105.869	201.858
	direct PCR	268.189	23.994	220.195	316.184

Appendix D - Poster: Direct PCR improves the recovery of DNA from various substrates.

Presented at "The International Society for Forensic Genetics" conference, Melbourne, Australia, 2013.

Direct PCR improves the recovery of DNA from various substrates

Jennifer E.L Templeton[1], Renée Blackie[1], Oliva Handt[2], Duncan Taylor[2], Adrian Linacre[1]
[1]School of Biological Sciences, Flinders University, Adelaide, Australia
[2]Forensic Science South Australia, Adelaide, Australia

Forensic Biology Lab South Australia

Funding was provided by the Department of Justice, South Australia and a John Stocker Postgraduate Scholarship.

Background – 'Touch' DNA

Trace DNA evidence recovered from handled items can be close to the limit of detection for most STR-based DNA profiling kits. The DNA may be:

- low mass (< 100 pg);
- chemically damaged;
- degraded / fragmented (< 200 bp) and/or;
- contaminated (environmental or human DNA).

Therefore, improved methods are required for the collection and processing of 'touch' DNA samples to increase the success rate of DNA profiling.

Introduction

Handled/touched items submitted to a forensic laboratory for analysis are:

- one of the most common forms of evidence and;
- one of the least successful in terms of generating a full DNA profile.

Current methodology subjects the sample to a DNA extraction to isolate and purify the target DNA, typically using a solid phase substrate, prior to amplification. DNA extraction encompasses multiple tube transfers that contribute to a significant loss (20-76%) of sample DNA [1] and is where the sample is more susceptible to contamination than any other stage in the analysis [2]. Omitting the DNA extraction stage, known as direct PCR, has the potential to minimise loss of target DNA in a critical sample. Further, removing steps involved in DNA extraction will reduce both labour time and costs, while increasing the likelihood of obtaining a meaningful DNA profile for interpretation.

Aim

The aim of this study was to:

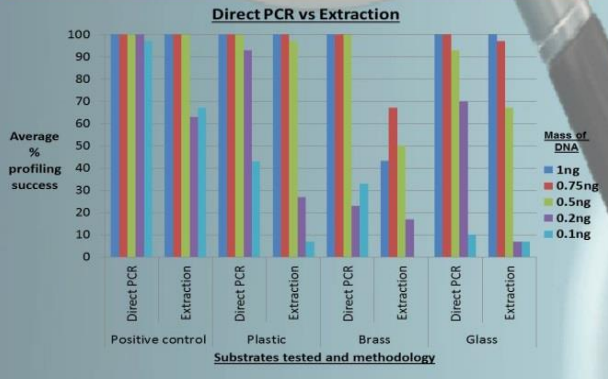
- Compare the DNA IQ™ kit extraction method to the direct PCR approach for processing low-level DNA swabs typical of crime scene samples.

Method

- Varying mass (1 ng, 0.75 ng, 0.5 ng, 0.2 ng and 0.1 ng) of control DNA (2800M, Promega) was deposited onto three sterile substrates:
  - Brass;
  - Plastic and;
  - Glass

- DNA was left to dry and swabbed using pre-moistened (Triton X 0.1 %) DNA-free nylon FLOQswabs ™
- For direct PCR, 2mm² fibres from the tip of the swab were added directly to PCR tube.
- For extracted swabs, whole swabs were subjected to the DNA IQ™ methodology following manufacturer's recommendations.
- Amplification conditions followed the Profiler Plus™ kit (ABI) guidelines in a final 25 µL reaction volume and 29 cycles.
- 3130xl capillary electrophoresis, all PCR products were run on the 3130xl (ABI) genetic analyser. Profiles were analysed using the GeneMapper ID v3.2 software and a peak amplitude threshold of 30 RFU was used to assign alleles.
- Fingerprint samples, 34 individuals provided fingerprints on plastic substrates as a source of latent DNA and samples were processed using direct PCR (as described above) to test the sensitivity of the technique.

Results – Direct PCR vs Extraction



Results – Direct PCR vs Extraction

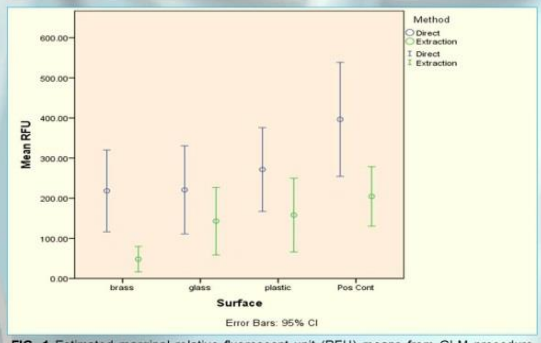


FIG. 1 Estimated marginal relative fluorescent unit (RFU) means from GLM procedure along with bootstrapped 95% Confidence Intervals (CI) displaying the interaction between method and substrate (plastic, glass and brass).

Results – DNA profile from a fingerprint

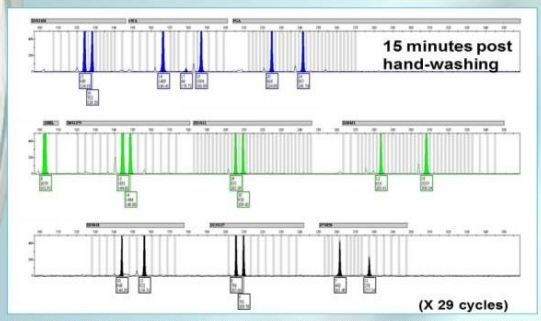


FIG. 2 Full STR DNA profile generated, using direct PCR and 29 amplification cycles, from a fingerprint deposited onto a sterile plastic slide 15 minutes after the donor washed their hands. Amplification kit: **Profiler Plus™ (ABI).**

Summary

Based on these data, direct PCR has the potential to:

- amplify low quantities of DNA recovered from brass, glass and plastic substrates;
- significantly increase the peak height of alleles **($p < 0.005$)** compared to the extraction experiment;
- generate **full STR DNA profiles** from fingerprints deposited on plastic substrates, using only 29 cycles of PCR;
- reduce the chance of contamination (allele drop-in) by minimising tube transfers;
- reduce cost and labour time of processing samples and;
- minimise stochastic effects inherent in analysing low-level DNA samples by removing the need to employ profile enhancement techniques (such as increased PCR cycle number).

Laboratories may benefit from using direct PCR on handled items due to its high-throughput potential. However, the success and quality of DNA profiles obtained using direct PCR is dependent on the nature of the material examined and the presence of PCR inhibitors; which are generally eliminated during the extraction process.

Future samples

References:
- [1] van Oorschot RAH, Phelan DG, Furlong S, Scarfo GM, Holding NL, Cummins MJ. Are you collecting all the available DNA from touched objects? International Congress Series. 2003;1239:803-7.
- [2] Butler JM. Forensic DNA Typing: Biology, Technology, and Genetics of STR Markers. 2nd Edition ed. New York: Elsevier Academic Press; 2005.

Contact details: jennifer.templeton@flinders.edu.au Telephone: +61 8 82015003

Appendix **E – Patent application**

Invention Title: NUCLEIC ACID COLLECTION DEVICE

Applicant: THE FLINDERS UNIVERSITY OF SOUTH AUSTRALIA

Patents Act 1990

PROVISIONAL SPECIFICATION

Regulation 3.2

AUSTRALIA

Patents Act 1990

PROVISIONAL SPECIFICATION

Invention Title: **NUCLEIC ACID COLLECTION DEVICE AND METHOD**

Applicant: **THE FLINDERS UNIVERSITY OF SOUTH AUSTRALIA**

The invention is described in the following statement:

Description of patent detailed below, wording is taken from the patent application [1].

NUCLEIC ACID COLLECTION DEVICE AND METHOD

FIELD

The present invention relates to devices for the collection of nucleic acid samples and methods for their use. Particularly, in some embodiments, the present invention relates to devices and methods for the collection of forensic nucleic acid samples, such as touch DNA samples.

BACKGROUND

Fingermarks are essential forensic evidence in numerous criminal investigations. Generating a DNA profile from a fingermark for the purpose of human identification would be beneficial in resolving a broad spectrum of criminal investigations, ranging from theft to crimes of violence.

DNA retrieved from fingermarks deposited by touch (referred to as “touch DNA”) is often degraded, limited in quantity and may contain elements that co-extract with the DNA, which can hinder subsequent analysis.

Although forensic genetics has seen substantial improvements in DNA profiling sensitivity typically the use of less than 100 pg of DNA template (equating to ~16 human somatic cells) can result in poor-quality profiles using standard forensic techniques. This limit of sensitivity still precludes many items that have been touched at a crime scene from generating a usable DNA profile, despite their potential importance in a criminal investigation; these samples can include triggers, steering wheels, bullet cartridges, and handles of knives.

In many criminal investigations, the ability to retrieve the maximum amount of DNA from

touch DNA samples is of paramount importance in resolving the case. The first DNA profile generated from a fingerprint was reported more than a decade ago and revolutionized forensic science. Despite this advance in the field, research has found an extremely low success rate (5%–6%) when using the standard methodology in generating touch DNA profiles, highlighting the need for improvements to the touch DNA process.

The standard workflow for touch DNA samples includes an extraction step before amplification. However, many current extraction processes are thought to result in a loss of DNA; hence a touch DNA swab initially containing less than 100 pg of DNA may result in insufficient template at the PCR step to generate an informative profile. Many touch DNA samples are therefore either not submitted for DNA typing or fail to generate any data without further enhancement. Previous attempts to generate DNA profiles from fingermarks used a low copy number (LCN) methodology. However, any enhancement of the amplification process from limited and low-level DNA has the inherent risk of introducing stochastic events such as allelic drop-out, allelic drop-in, or an increase in stutter heights or allelic imbalance for a heterozygote.

In light of the above extraction-independent techniques for DNA analysis have been developed and these include “direct PCR” type methods such as described in Ottens *et al*. (*Forensic Science International Genetics Supplement Series* 4: e47-e48, 2013).

Regardless of the choice of analytical methodology, optimizing the yield of DNA collected from an environment is critical when dealing with touch samples where limited quantities of DNA may be left behind. The number of cells transferred to touched objects is highly variable, and often results in less than 300 picograms of DNA. Most forensic short tandem repeat (STR) kits call for 0.5–1.0 nanograms of DNA (about 200 cells) for full profiles to be generated, thus it is crucial to maximize DNA recovery from handled objects.

As set out above, the success of a DNA profile depends largely on how much DNA is retrieved from the evidence. Cotton swabs, pre-moistened with distilled water, are used routinely in forensic laboratories. However, it has been recognized by the present inventors that these relatively large cotton swabs are not optimal for the collection of touch DNA samples.

DESCRIPTION

It has been recognised by the present inventors that relatively large cotton swabs as used in routine forensic practice are not optimised for the collection of touch DNA samples. In particular, it has been recognised by the inventors that the relatively large working volumes of these swabs, typically in the 100's of microlitre range, may dilute touch DNA samples to the point that they are no longer suitable for forensic analysis. Moreover, it has also been recognised by the inventors that direct amplification of touch DNA off a swab is an optimal way of analysing touch DNA samples and, as such, the relatively large size of routine cotton swabs impedes the direct analysis of DNA as the swab head is generally too large and/or absorbent to be directly included in a reaction to detect a nucleic acid.

Accordingly, in a first aspect, the present invention provides a device for collecting a nucleic acid sample from an environment, the device comprising:

a collection head comprising a positively charged collection surface, wherein the collection surface comprises a saturation volume of between 1ul and 100ul; and

a handle connected to the collection head.

The device of the present invention is adapted to the collection of nucleic acid samples from an environment. Typically, nucleic acid samples collected from an environment are for later analysis to determine the source of the nucleic acid, such as is done in forensic analysis cases.

In light of the above, the "environment" referred to herein may be any environment in which a nucleic acid sample of interest may be found. For example, the environment may be a surface that has come into contact with a nucleic acid of interest such as a surface that has been touched by a person of interest or where a body fluid has been deposited on a surface.

In some embodiments, the "environment" may also comprise an organism, or a cell, tissue or organ thereof, such as an animal, plant or microorganism; an organic product including a plant product such as timber or other plant material or an animal-derived product such as animal horns or ivory; a device used to collect a nucleic acid or organic sample; and the like.

In some embodiments, the device of the present invention has particular application in the collection of nucleic acid samples from an environment that has been touched by a person of interest and, as such, the device has particular application in the collection of nucleic acids from fingerprints left on a surface by a person of interest. In some embodiments, the present invention has particular application in the collection of a nucleic acid from environments such as adhesive tape, plastic bags, plastic covering of wires, plastics and metals associated with electric or electronic devices, metal surfaces of firearms or ammunition, natural substrates including plants and outer covering of animals.

The device of the present invention is particularly adapted to working with very low volumes of liquid as are typically used when attempting to collect DNA from a touched surface, referred to herein as 'Touch DNA'.

As such, in some embodiments, the collection surface comprises a saturation volume of less than 100ul, less than 50ul, less than 20ul or less than 10ul. In some preferred embodiments, the collection surface comprises a saturation volume of: between 1ul and 100ul, between 1ul and 50ul or between 2ul and 20ul.

A "saturation volume" as referred to herein refers to the volume of liquid that the collection surface can absorb or adsorb before reaching a point of saturation, ie. the point at which no more liquid can be absorbed or adsorbed by the collection surface.

Moreover, it is has been recognized by the inventors that it is highly beneficial for the collection surface (after collection of a nucleic acid) to be directly incorporated into a DNA analysis technique without pre-processing.

In some embodiments, it has been recognized that polymerase chain reaction ('PCR') is involved in many suitable analysis techniques for touch DNA samples. As such, in some embodiments of the invention, it is desirable for at least the collection surface of the collection head of the device to be sufficiently small such that it can be fully immersed in the working volume of a PCR reaction in a PCR reaction vessel.

As such, in some embodiments, at least the collection surface of the collection head, when

ejected from the handle, is sufficiently small such that it can be fully immersed in no more than 100ul of liquid in a 200ul PCR tube; no more than 50ul of liquid in a 200ul PCR tube; no more than 30ul of liquid in a 200ul PCR tube; or about 25ul of liquid in a 200ul PCR tube.

A person skilled in the art would readily recognize what is meant by a "200ul PCR tube". However, by way of non-limiting example, a "200ul PCR tube" includes the Applied Biosystems 200ul PCR tube illustrated in Figure 10. However, as would be appreciated by those of skill in the art, a range of other 200ul PCR tubes possibly having different dimensions are also contemplated by the present invention. For example, in some embodiments a "200ul PCR tube" may include a 200ul PCR tube comprising a screw cap closure rather than a friction fit closure as illustrated in Figure 10.

As set out above, the collection surface of the device of the present invention comprises a positively charged surface. In use, the collection surface of the device is contacted with a surface in the environment to be sampled in order to adhere nucleic acids from the environment to collection surface. Without being bound by any particular mode of action, it is considered that negatively charged nucleic acids such as DNA or RNA are attracted to, and adhere to, the positively charged collection surface.

In light of the above, any suitable positively charged surface may be used for the collection surface. However, in some embodiments, the collection surface comprises a positively charged polymer or cationic polymer. In some embodiments, the collection surface may also comprise a natural or synthetic polymer comprising an amide linkage or an amine or thiol group in the polymer backbone or side-chain.

In some embodiments, the collection surface comprises Nylon. "Nylon" as used herein refers to a family of synthetic polymers that are typically made by reacting monomers such as lactams, acid/amines or stoichiometric mixtures of diamines ($-NH_2$) and diacids (-COOH). Mixtures of these can be polymerized together to make copolymers. Nylon polymers can be mixed with a wide variety of additives to achieve many different property variations. "Nylon" as referred to herein should be understood to not refer to any one particular form of nylon, but refer to any member of the nylon polymer family that carries a net positive charge.

The collection surface may be any suitable configuration to promote transfer of a nucleic acid from an environment to the collection surface.

In some embodiments, the collection surface comprises a “flocked” surface. A flocked surface as referred to herein refers to a surface, which comprises a coating of particles or fibres attached to a substrate. In some embodiments a flocked surface is made by applying a charge to the particles or fibres to be attached to the surface (ie. the flock) whilst the substrate is earthed. Flock material flies vertically onto the substrate attaching to previously applied glue. A number of different substrates can be flocked including textiles, fabric, woven fabric, paper, PVC, sponge, toys, and automotive plastic.

In some embodiments, the flocked surface of the present invention comprises positively charged polymer flock or nylon flock.

In some embodiments, the flocked surface comprises a flock fibre length of less than 0.6 mm. In some embodiments the flocked surface comprises a flock fibre length of less than 0.4 mm.

In some embodiments, the collection surface may also comprise a smooth or textured surface, a brush configuration, a microfibre cloth, or a woven configuration.

As set out above, the device of the present invention comprises a handle. In use, the handle of the device is gripped by a user to manipulate the device such that at least the collection head contacts the environment to be sampled in order to effect transfer of nucleic acids from the environment to collection head. The handle may be in any suitable configuration to promote effective sampling. For example, the handle may be straight or may include one or more bends to place the collection head at a suitable angle for effective sampling relative to the handle.

In some embodiments, the handle may be substantially rigid that is, in use, the handle does not substantially bend or deflect from its original configuration.

However, in some embodiments, the handle may comprise a flexible portion. In some embodiments, the handle may comprise a resilient portion.

A “flexible portion” as referred to herein contemplates a portion of the handle that may be bent or positioned to a desired angle in order to generate and hold a desired angle between the handle and the collection head of the device.

A “resilient portion” as referred to herein contemplates a portion of the handle that may flex under force in order to transiently allow an angle to be formed between handle and collection head. However, in this case, when the force causing flexion of the resilient portion is removed, the resilient portion will return to its original resting position.

In some embodiments the flexible or resilient portion may comprise a part of the handle. In some embodiments the flexible or resilient portion may comprise substantially all of the handle.

The handle, or a portion of the handle, may be manufactured from any suitable material or materials to achieve the desired level or rigidity, flexibility or resilience. For example, in some embodiments, flexible polystyrene (PS), copolymer polypropylene (PP), may be used in order to produce flexible handle portions. In some embodiments thermoplastic elastomer (TPE/TPU) may be used in order to produce a resilient handle portion. Where it is desired for the handle or a portion thereof to be substantially rigid, materials such as metals, wood or substantially rigid polymers (such as polycarbonate or polystyrene), carbon fibre or ceramic, among others, may be used.

In some embodiments, the device of the present invention comprises an ejection means for ejecting the at least the collection surface from the device.

Ejection of at least the collection surface from the device is useful where the collection surface needs to be transferred to another vessel, such as a vessel for analysis of a nucleic acid present on the collection surface after sampling an environment. In some embodiments, ejection of at least the collection surface may include ejection into a PCR reaction vessel for analysis of a nucleic acid on the collection surface.

Ejection of “at least the collection surface” may include ejection of all or part of the collection surface alone from the device or, in some embodiments, include co-ejection of

another part of the device such as all or part of the collection head and/or a portion of the handle.

In some embodiments, the ejection means comprises an ejection actuator that is actuatable by a user to eject at least the collection surface from the device.

Actuation of the ejection actuator may be via any suitable means such as pushing, pulling twisting, bending, squeezing, slicing and the like to effect ejection of at least the collection head from the device.

Actuation of the ejection actuator may effect ejection of at least the collection head in any suitable way, which would be readily ascertained by one skilled in the art. For example, actuation of the ejection actuator may displace a friction fit between the collection head or surface and the handle; may sever a join between the collection head or surface and the handle; may unscrew a threaded join between the collection head or surface and the handle; and the like.

The ejection actuator may be located at any convenient location on the device. For example, in some embodiments, the ejection actuator is positioned on the device proximate to the collection head. In some embodiments the ejection actuator is positioned on the device distal to the ejection head.

In some embodiments, the device further comprises an engaging means to engage with a reaction vessel, wherein when the device and the reaction vessel are engaged via the engaging means, at least the collection surface of the device is positioned within the reaction vessel.

In some embodiments, the engaging means may be configured as a collar that engages with the opening of a reaction vessel to form a substantially liquid tight seal between the engaging means and the opening of the reaction vessel. Typically, the engaging means or collar on the device is positioned such that when a reaction vessel is engaged with the engaging means, at least the collection surface of the device is positioned within the reaction vessel.

Formation of a seal between the engaging means and the opening of a reaction vessel may be via any convenient mechanism including, for example, a friction fit engagement, a threaded engagement or the like.

A “reaction vessel” as referred to herein may be any vessel in which a nucleic acid analysis process or method may be conducted, such as a vessel in which a PCR reaction may be conducted. One skilled in the art would readily identify vessels suitable for the conduct of a PCR reaction or other nucleic acid analysis method. However, by way of example, a suitable reaction vessel for a PCR reaction would include a typical 200ul volume PCR reaction tube, an example of which is illustrated in Figure 10.

The device of the present invention is useful for the collection of a nucleic acid from an environment.

A “nucleic acid” as referred to herein may include any type of nucleic acid that might be found in an environment. As such, in some embodiments, the nucleic acid may comprise DNA or RNA. Furthermore, the subject nucleic acid may be in any suitable form such as single stranded or double stranded.

In some embodiments, the nucleic acid comprises DNA.

The sampled nucleic acid may come from any suitable source such as an animal, a plant or a microorganism such as a bacterium or fungus. However, in some embodiments, the nucleic acid comprises a nucleic acid from an animal, including for example human DNA or RNA.

The nucleic acid may also be derived from any part of a cellular organism. As such, a nucleic acid as referred to herein may include genomic DNA, mitochondrial DNA, chloroplast DNA, viral DNA, DNA or RNA from an extrachromosomal element and the like. In some preferred embodiments, the nucleic acid comprises genomic DNA.

In a second aspect, the present invention also provides a method for determining the presence of a nucleic acid in an environment, the method comprising:

providing a device according to the first aspect of the invention;

Micro-swab concept [1]

SI Figure 3.9 Microswab concept design.

Micro-swab concept [1]

SI Figure 3.10 Microswab in 0.2 mL PCR tube.

SI Figure 3.11 Microswab showing swab fibre tip.

Appendix F - Oral presentation given to FSSA, 2013.

Improving methods for the recovery and analysis of "Touch DNA"

Jennifer Templeton, PhD Forensic Biology, Flinders University

Principle supervisor: Professor Adrian Linacre
Co-supervisors: Dr Oliva Handt and Dr Duncan Taylor

Funding provided by the Department of Justice, South Australia

70% 'touch' DNA samples fail!

Direct PCR – "Touch DNA"

- Advantages to avoiding DNA extraction:
  - Minimise loss of target DNA;
  - Omit steps involved in standard practice;
  - Reduce labour time and cost;
  - Increase the likelihood of obtaining a DNA profile and;
  - Reduces risk of contamination

Published papers:

- *Ottens, R., Taylor, D., Abarno, D. and Linacre, A. (2013).* ***Successful direct amplification of nuclear markers from a single hair follicle.*** *Forensic Science, Medicine, and Pathology, 9(2) pp. 238-243.*
- *Linacre, A., Pekarek, V., Swaran, Y. C., & Tobe, S. S. (2010).* ***Generation of DNA profiles from fabrics without DNA extraction.*** *Forensic science international: genetics, 4(2), 137-141.*

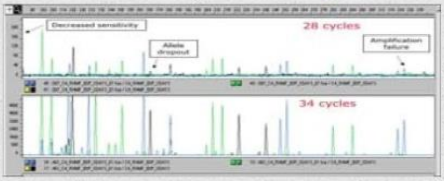
Low success rate for "Touch DNA"

- Industry standard process – limited success
  - low concentration or structurally modified
    →primers might not consistently hybridise
- Contamination issues
- Other complications:
  - amplification failure (locus drop-out);
  - preferential amplification (allele drop-out) and;
  - artefacts (stutter, allele drop-in and heterozygote imbalance)

Low Copy Number (LCN) analysis

- Forensic Science Service (FSS) UK, 1999
  - increases cycle number 28 → 34 cycles
- Forensic labs avoid LCN analysis due to:
  - stochastic effects
  - increased occurrence of allele drop-in

DNA profile comparison



Touch DNA

- Papers published on touch DNA:
  - *Van Oorschot, Roland AH, and Maxwell K. Jones. "DNA fingerprints from fingerprints." NATURE-LONDON- (1997): 767-767.* ***x34 CYCLES***
  - *Saito, Yahachi, et al. "DNA fingerprinting from single cells." Nature 389 (1997): 555.* ***x34 CYLCES***
  - *Vandewoestyne, Mado, et al. "Presence and potential of cell free DNA in different types of forensic samples." Forensic Science International: Genetics 7.2 (2013): 316-320.* ***x34 CYCLES***
  - *Gashi, B., et al. "Measurement of 9mm cartridge case external temperatures and its forensic application." Forensic science international 200.1 (2010): 21-27.* ***x34 CYCLES***
  - *Balogh, M., et al. "Fingerprints from fingerprints". International Congress Series 1239 (2003): 953-957* ***x38 CYCLES***
- STR kits – mostly validated for 28, 29 or 30 cycles!

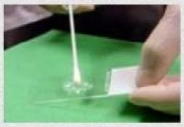
Methodology

- **1st experiment - optimising swabbing method**
- Control DNA (2800M, Promega, 1ng/uL) - slides
  - types of swabs
  - swabbing technique
  - detergent-based medium

- Direct PCR (AmpFℓSTR ProfilerPlus™ kit, 28 cycles)
- Amplification products run on 3130xl (ABI)
- Results analysed – Genemapper ID software (v3.2.1)
  - peak amplitude threshold 50 RFU
- Optimum swabbing technique
  - nylon fibres, pre-cut and pre-moistened with triton X-100 (0.1%) – double swabbing technique

Compared method of direct PCR to extraction (DNA IQ)

- Control DNA used to control DNA variation from fingerprints
  - varying amounts of DNA recovered from hands depending on:
    - body temperature and stress
- **METHOD:**
  - Varying concentrations of control DNA (1ng, 0.75ng, 0.5ng, 0.2ng and 0.1ng) deposited:
    - brass, glass and plastic
- DNA left to dry and swabbed using pre-moistened DNA-free nylon flocked swabs (Copan)
- For Direct PCR:
  - 2mm² fibres added directly to 0.2mL PCR tube
- DNA IQ Extraction:
  - whole swab used
- Three replicates were carried out for each
  - 45 samples processed in total



Pairwise Comparisons

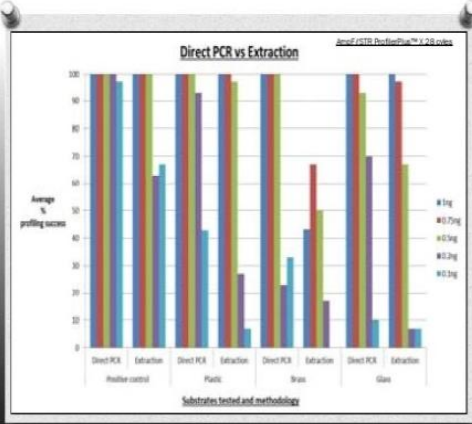
Dependent Variable: Average relative fluorescent units (RFU)

(I) Method used for DNA testing	(J) Method used for DNA testing	Mean Difference (I-J)	Std. Error	Sig.[b]	95% Confidence Interval for Difference[b]	
					Lower Bound	Upper Bound
extraction	direct PCR	-119.389*	21.080	.000	-161.356	-77.423
direct PCR	extraction	119.389*	21.080	.000	77.423	161.356

Based on estimated marginal means

*. The mean difference is significant at the .05 level.

b. Adjustment for multiple comparisons: Bonferroni.



Fingerprint test

- Ten volunteers – latent DNA:
  → 5 females and 5 males
- Time:
  → 15 minutes, 1 hour and 2 hours
- Contact:
  → 15 seconds, medium pressure
- Storage:
  → slides stored 4°C over-night

(2mm² fibres)

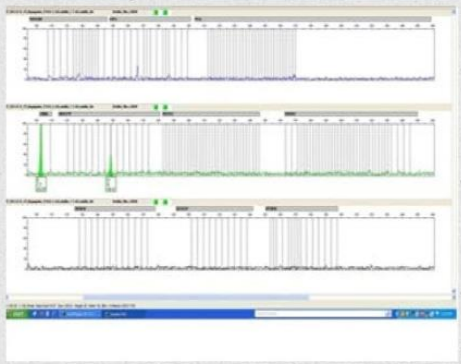
- Following day: prints swabbed
  - double swabbing technique
  - fibres (2mm²) cut from tip – placed in 0.2mL PCR tube
  - amplification: AmpFℓSTR ProfilerPlus™ with BSA and DMSO

DNA Profile – fingerprint after 15 minutes

DNA profile – fingerprint after 1 hour

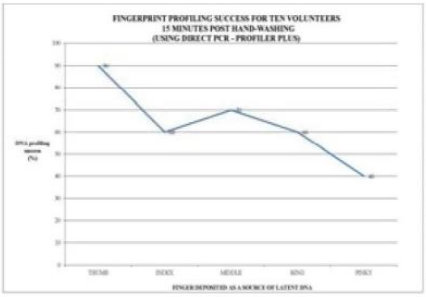
DNA profile – fingerprint after 2 hours

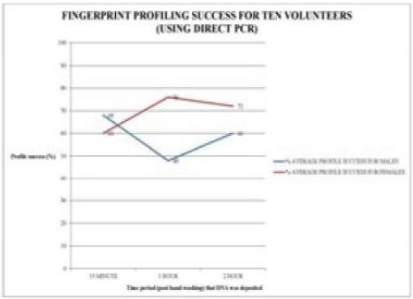
DNA profile – no result after 2 hours!



Latent DNA extraction

- Two individuals picked – consistently good results
- Repeated fingerprint test
  - whole swab extracted - DNA IQ kit
  - amplification kit: AmpFℓSTR ProfilerPlus™
- 10 swabs processed in total
- Variation in amount of DNA deposited
  → No result from extracted swabs
- Direct PCR does have potential for latent DNA!





Criminal case – drug seizure (led by SA police)

- Smudged fingermark on tape – direct PCR
- Containers - methamphetamine

Drug case - profile

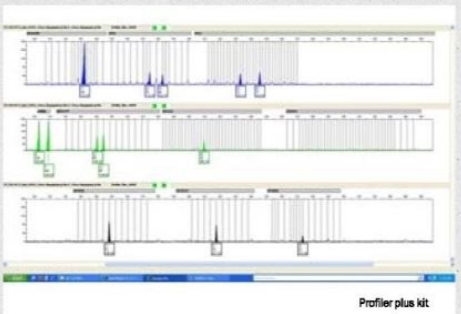
CARTRIDGE CASES

Cartridge - Aluminium

Cartridge - Nickel

Cartridge - Brass

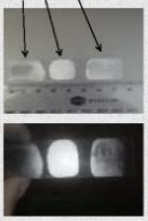
Cartridge – Brass (profiler plus)



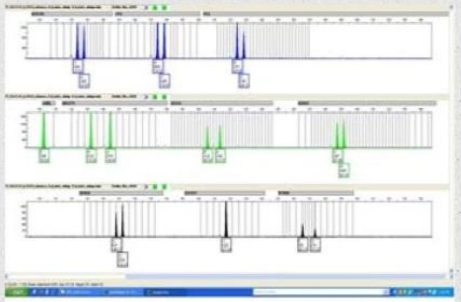
Fingerprint powders – Direct PCR

- Powders:
  - Silver Aluminium (stearic acid and aluminium powder)
  - Black carbon
  - White (titanium dioxide, zinc serate, lycopodium)

Latent DNA deposited
15 minutes after hand-washing



Powder results -white



Profiler plus kit

Powder - black

Profiler plus kit

Powder - silver

Profiler plus kit

Future experiments

- Quantitative PCR (qPCR) – quantify amount of nuclear DNA in fingerprint
- Trial different enzymes e.g. Phusion
- STR kits eg. Globalfiler and NGM
- Target mtDNA to generate haplotypes
- Ancestry Informative Markers - SNPs
- Y STRs

Chapter IV
DNA profiles from fingermarks

Publication included in Chapter IV

1. **Templeton, J.E**. and Linacre, A., 2014. *DNA profiles from fingermarks.* BioTechniques, 57(5), p.25

Manuscript (in preparation) in Chapter IV

1. R. Blackie, **J. Templeton**, D.A. Taylor, A. Linacre. *PCR buffer enhancement of STR kits used for human identification.* Letter to the Editor. International Journal of Legal Medicine.

Case work included in Chapter IV

Case 1) Application of case work: seizure of Methamphetamine in South Australia. The application of direct PCR to generate a DNA profile.

Appendix

a) Poster: **Templeton, J.E.,** Handt, O., Taylor, D., and Linacre, A. *Genetic profiling from challenging samples.* Presented at "The International Society for Forensic Genetics" conference, Melbourne, Australia, 2013.

b) Oral presentation: **Templeton, J.E**. 3MT (Three minute thesis). Runner up, static presentation slide – Flinders University, 2015,

c) Oral presentation: **Templeton, J.E.,** Handt, O., Taylor, D., and Linacre, A. *Improving methods for the recovery and analysis of 'touch' DNA from fingerprints.* Presented at "The American Academy of Forensic Sciences" conference, Orlando, Florida, America. 2015.

d) Award: Ross Vining memorial student scholarship.

e) Award: 3M thesis competition.

4.1 Preface

Despite the various interpretations of the term ‘trace DNA’, there is a common association that it is derived from the smallest samples encountered in forensic biology and has the potential to provide crucial evidence to a case. PCR-based methodologies have enabled DNA profiles to be generated from touched objects – a common form of trace evidence [1-4]. The term ‘touch’ DNA can be used to describe DNA deposited by touch that is not attributed to a biological fluid. However, not all DNA is a result of direct contact (e.g. secondary or tertiary transfer can occur) and the terms ‘contact/trace/transfer DNA’ are used interchangeably to describe samples of this nature. Consequently, items of ‘touch/contact’ DNA content are routinely submitted to a forensic laboratory for analysis.

Latent DNA derived from fingermarks or touched evidence can be difficult to recover routinely using standard methodology. This may be due, in part, to the way that the samples are collected and processed. Low-levels of DNA recovered from swabs [5] and the extraction methodology may result in low success [5-7]. As discussed in Chapter III, by eliminating the extraction steps (i.e. direct PCR) there is the potential to increase the amount of DNA template available to the reaction.

Direct PCR has the potential to be applied to ‘touch’ DNA samples that commonly don’t yield sufficient information to allow evaluation of the DNA profile [8, 9]. Chapter III previously reported the use of direct PCR to amplify control DNA deposited on various substrates, and highlighted the potential gains for swabs to be processed in this manner. Accordingly, this chapter proceeds to using direct PCR on latent DNA obtained from fingermarks, also referred to as ‘touch’ DNA.

This data chapter is split into four phases, each of which examines aspects relating to the direct PCR of ‘touch’ DNA swabs:

- **Buffer enhancement:**

The first phase of the study examines the presence of PCR facilitators in the PCR tube and their ability to improve the yield of DNA available for PCR.

- **Direct PCR of ‘touch’ DNA:**

The second phase of the study investigates the amount of DNA transferred to a substrate 15 minutes, 1 hour and 2 hours after hand washing. The presence of mixed profiles and secondary transfer is observed. Additional work explores the effect of depositing DNA immediately after hand washing.

- **Extracting DNA from fingermarks using the DNA IQ™ System:**

The third phase of the study investigates the difference between direct PCR and a standard DNA extraction on the profiling ability of ‘touch’ DNA swabs.

- **Application of case work: seizure of Methamphetamine in South Australia:**

The fourth, and final, phase of the study investigates the application of the direct PCR approach in a real case study (in collaboration with FSSA).

4.2 Aims of study

1) To determine if PCR facilitators present in the PCR vessel improve the DNA profiling success.

2) To determine if informative DNA profiles can be generated from touched substrates at T0 (0 minutes), T15 (15 minutes), T1hr (1 hour) and T2hr (2 hour) post hand washing, and to investigate the profiling success of extracted DNA swabs in comparison to direct PCR swabs.

3) To determine the success of optimised swabbing and direct PCR at amplifying 'touch' DNA 15 minutes after hand washing, by using two common STR-based DNA profiling kits.

4) To apply the methodology to a real case study and highlight the potential benefits and limitations of the direct approach.

4.3 Introduction

4.3.1 'Touch' DNA

In 1997, the first DNA profile was generated from a touched object using the LCN technique [1]. Further studies by van Hoofstat *et al.* (1999) [10], van Renterghem *et al.* (2000) [11] and Zamir *et al.* (2000) [12] corroborate the findings of van Oorschot *et al.* (1997) [1] with further success in generating DNA profiles from touched items. This new avenue of research opened up the potential to collect DNA from a wide range of exhibits, such as: personal effects, clothing, tools, vehicles, weapons, firearms, doors, glass, paper, plastic, cosmetics and many more, using increased PCR cycle numbers [1, 10, 13-18]. However, the application of increased PCR cycle number has its own challenges, most noticeably an increase of stochastic effects [19]. Interpretation becomes more complex when stochastic effects, such as peak height imbalance, 'drop-in', and 'drop-out' are evident. Mixtures as a result of secondary and tertiary DNA transfer add another level of complexity to the analysis of 'touch' DNA and need to be considered [20].

FSSA success rate for contact/'touch' DNA

Forensic laboratories often need to deal with 'touch' DNA samples that come from a less-than-pristine environment [21] and/or have a limited quantity of DNA. FSSA report a low success rate in obtaining DNA profiles from a range of touched objects when using current methodology (see Figure 4.1). A summary of results concluded that many items submitted for contact DNA analysis yielded mixed DNA profiles and all sample types that were examined included DNA samples that yielded insufficient data for profile interpretation (see Figure 4.1). Firearms (under grip) when analysed for contact DNA resulted in the lowest success rate with 88% of samples submitted generating insufficient results using Chelex® 100, and 60% of samples yielding insufficient results with the DNA IQ™ System (see Figure 4.1). The highest profiling success, in terms of obtaining 'up-loadable' DNA profiles for the Australian DNA database (i.e. > 12 alleles, plus Amelogenin), was obtained from drug balloons using the DNA IQ™ System which generated 'up-loadable' profiles in 56% of cases (see Figure 4.1). Other casework laboratories report low success rates from swabbing case work exhibits. The Institute of Environmental Science and Research Limited (ESR) in New Zealand reported that only 5 % of 'handled items' generated a full

DNA profile and 69 % of samples gave no results at all using the AmpF/STR SGM Plus™ PCR amplification kit (Applied Biosystems) [6]. Similarly, Raymond *et al* (2009) interpreted trace DNA results from 252 casework samples and stated that 44 % of 'touch' DNA samples failing to generate DNA alleles [22]. Likewise, Castella *et al* (2008) examined 1739 'touch' DNA samples from case work exhibits and noted a success rate of 26 % in terms of generating DNA profiles that were considered unloadable to the Swiss DNA database [18].

One reason that FSSA, and many other laboratories, do not use the direct PCR approach for analysing contact/'touch' DNA swabs is the laboratory's requirement to quantify samples prior to analysis. If accredited laboratories validate methods for processing 'touch' DNA swabs by direct PCR, there is the potential to implement these protocols into standard practice, and potentially improve the profiling results for sub-optimal samples. Direct PCR could be used on samples of high importance to increase the success rate of DNA profiling.

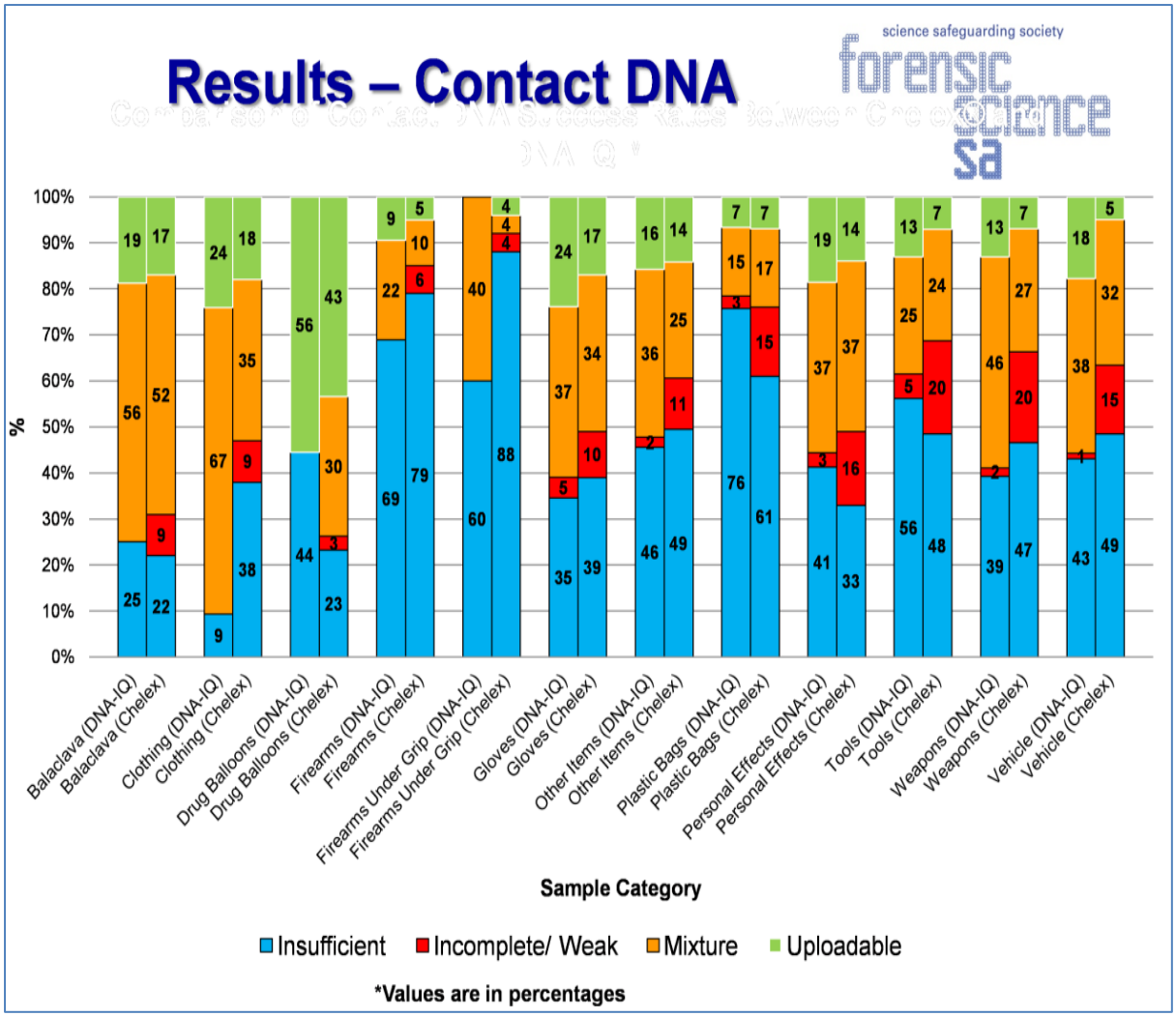


Figure 4.1. Image generated at FSSA by Oliva Handt and Nicol Sly. DNA profiling success for contact/'touch' DNA exhibits submitted to FSSA for analysis in 2013.

Is quantification of 'touch' DNA necessary?

Low-template DNA has been characterised as less than 100 picograms (pg) of genomic DNA in a sample [23], which equates to approximately 15 diploid copies of nuclear DNA. One study estimated the quantity of DNA recovered from touched substrates and reported the average amount of DNA recovered from fingermarks was less than 100 pg [24]. If individuals washed their hands prior to depositing a print the amount recovered was substantially lower than the detection threshold of 40 pg [24].

A laboratory may choose to quantify their DNA extract prior to amplification and decide not to proceed with the sample if the total amount of DNA measured is below a set laboratory-specific threshold (e.g. 200 pg) [19, 25]. Quantification cut-off values

represent a resource-based decision. Aspects such as time, resources, sample type, and the likelihood of generating a result will help to determine whether to process a sample. A review published by van Oorschot *et al.* (2010) defined trace DNA evidence as "*any sample that falls below recommended thresholds at any stage in the analysis, from sample detection through to profile interpretation, and cannot be defined by a precise picogram amount*" [21]. This is an important consideration, as a good quality DNA profile can be generated from as little as 100 pg of DNA and a poor quality DNA profile obtained from a sample containing 500 pg of DNA [26]. In addition, Kamphausen *et al.* (2012) [27] found no correlation between the mass of starting template and the completeness of a DNA profile.

A low quantification result should not be the sole determining factor in deciding whether a sample should undergo downstream analysis. For example, one study reported that a nine STR locus profile was obtained from a sample that did not meet the threshold for quantification using the Quantifiler Human DNA Quantification kit (Applied Biosystems, Foster City, CA, USA) [28], yet this sample may never have been analysed if the laboratory was strict with threshold limits. More sensitive quantification kits are now commercially available (e.g. Quantifiler® Trio DNA Quantification Kit, ABI) that have a lower limit of detection for trace DNA. However, quantification methods do not provide absolute measurements with real-time PCR assays [21]; an indication of concentration can be useful, however, it is not always a true reflection of sample quality.

In addition to the initial mass of DNA present in the sample, other factors will influence how well a sample performs downstream, such as the presence of PCR inhibitors and/or fragmentation of the DNA template (i.e. degradation). The greater the quality and quantity of the initial DNA template the better the overall profiling result.

Improving amplification of low-level DNA

Sub-optimal profiles are generated if less than adequate DNA template is present in the amplification vessel. Methods to improve profiling results include concentrating the DNA using spin columns [21, 29, 30]. However, a recent study reports a loss of DNA using filtration columns [31]. Methods exist to optimise PCR conditions and include altering the magnesium concentration, buffer pH, cycle numbers, annealing temperatures and time. However, all commercial STR typing kits produced for forensic DNA profiling are pre-optimised for targeting multiplex loci; it is therefore not recommended to alter these parameters. An alternative strategy to enhance amplification for direct PCR is the addition of PCR facilitators to overcome inhibitors.

Dimethyl sulfoxide (DMSO) and Bovine Serum Albumin (BSA) have been shown to increase the quality of DNA profiles when a DNA extraction is used [32-36]. BSA is a well-known amplification facilitator which inactivates inhibitors through the binding efficiency of albumin [32]. DMSO has been shown to increase PCR yields [37, 38] by stabilizing nucleic acid complexes, improving strand separation to facilitate enzyme function, and improve primer annealing efficiency. By altering the PCR master mix to accommodate these additives there is the potential to improve primer binding and stabilize nucleic acids [39]. Analogous to DMSO, Betaine has been linked to its ability to destabilize regions of high G-C content in the DNA sequence in order to assist with strand separation [37, 40]. Other additives, such as Triton-X™, may prevent sample DNA from binding to the walls of reaction tubes and assist with lysing cells. These reagents may prove to be beneficial for processing direct PCR swabs that contain low-template DNA.

It is misleading to assume that all ‘touch’ DNA samples contain low quantities of DNA. High amounts of DNA may be shed onto a touched item, depending on the characteristics of the individual, the item that was touched, and/or the nature of the contact [1, 17, 18, 41].

Structure of skin cells

The majority of surface skin cells are known as keratinocytes which are thought to be 'dead' [42]. They become compressed in the granular layer of the epidermis and lose their nuclei (i.e. become anucleate) as they progress through layers of the skin towards the upper surface [42]. Apoptosis (i.e. cell death) is the mechanism behind which cells lose their nuclei and according to Kita *et al.* (2008) "DNA is degraded by several enzymes during keratinization (i.e. differentiation process)" [42]. Nakamura *et al.* (1999) reported that during this differentiation process "large organelles, including the nucleus, are eventually lost and the cells are filled with keratohyalin and keratin filaments" [43]. Quinones *et al.* (2012) assumed that despite the cells being anucleate, keratinocytes may still contain residual DNA [44].

A study by Balogh *et al.* (2003) indicates that 'touch' DNA consists of nucleated cells that have been transferred to the hands through touching other areas of the body, and also originates from sweat glands [45]. Similar to previous findings, Alessandrini *et al.* (2003) [46] state that "DNA in the keratinized cornified layer of epidermal cells may be derived from stripped nuclei and show evidence of apoptotic degradation". Later in 2008, Toshiro Kita *et al.* (2008) [42] carried out morphological and immunohistochemical analysis on the skin to determine where the DNA on the surface of the skin originates. Their findings concluded that small amounts of degraded DNA fragments were detected and thought to have originated from the cornified layers of skin and that sweat may also consist of cell-free DNA [42].

Balogh *et al.* (2003) [45] stated that "the differentiation process regulating the epidermal growth is accompanied by changes in cellular biochemistry involving the activation of catabolic enzymes". Sen *et al.* (1992) described the endonuclease responsible as a "Ca^{2+}/Mg^{2+} dependent enzyme that fragments the DNA into 200 bp fragments" [47, 48]. This may be one reason why mini STRs – that target smaller sized amplicons – have had good success at amplifying 'touch' DNA samples [49]. A more recent publication by Lacerenza *et al.* (2016) [50] performed DNA/RNA co-analysis of the palmar surface of hands and fingers and found that tissues other than skin cells occasionally contributed (i.e. in 15% of cases) to the 'touch' DNA content.

In addition to the complex nature of ‘touch’ DNA, various factors play a part in determining whether cellular material deposited by touch will transfer to an object.

4.3.2 Factors influencing DNA transfer

Profiling success of ‘touch’ DNA is dependent on a number of variables, such as:

Shedder status

One study defines the ‘shedder’ status of an individual based on their ability to deposit fingermarks and generate a DNA profile 15 minutes after hand washing [51]. It was found that shedder type may not be relevant if a longer time had passed since hand washing (e.g. between 2-6 hours). Two hours after hand washing, the majority of individuals tested deposited enough DNA for a profile to be generated by LCN [51]. The Lowe *et al.* (2002) study [51] has been used to categorize individuals as ‘shedders’ or ‘non-shedders’ and to provide information in court regarding DNA transfer [26]. This was the reason for choosing 15 minutes as an initial time period of research in this study. The research conducted by Lowe *et al.* (2002) [51] noted a measurable difference between individuals in their tendency to shed DNA. Additional research groups have shown that the amount of DNA transferred to an object by primary transfer is dependent on the donor [1, 3]. Goray *et al.* (2016) examined 10 volunteers’ handprints at different times of the day, and on different days, to determine the degree of DNA variation observed [52]. Observations included considerable variation in the amount of DNA that people deposit at different times of the day depending individuals activities. However, it was also noted that some individuals consistently deposited more DNA that others, with non-self DNA being deposited alongside self-DNA in the majority of cases. A conflicting study failed to support ‘shedder’ status and the authors report that individuals deposit different amounts of DNA on different days and the success of profiling varies [53]. Phipps *et al.* (2006) [53] found that between 50 – 70% of volunteers in their study failed to generate a DNA profile after a 10 second touch on plastic. Additionally, work by Raymond *et al.* (2009) [22] report that 44% of ‘touch’ DNA samples failed to produce an informative profile. Further research conducted by Castella *et al.* (2008) [18] report 26% profiling success from ‘touch’ DNA exhibits, in terms of generating informative DNA profiles that are suitable for upload to the Swiss DNA database.

The difference in success rates are due to the individuals depositing DNA and also, in part, to the way that samples are processed and the inability to capture enough DNA for PCR.

Physical nature of the surface

It is believed that cellular material is more likely to adhere to 'rough' or porous surfaces (such as wood, concrete and grooved surfaces) more easily than smooth surfaces. In theory, rough objects that have cracks or grooves would be expected to collect and retain more skin cells compared to smoother surfaces.

Type of contact

The length of time that an object is handled by the donor may influence the number of cells that are transferred [54]. In addition, the force of pressure and friction may affect the amount of DNA transferred, as it does with transfer of body fluids.

Hand washing

Fingers have small quantities of oil and perspiration which are secreted from microscopic pores on the ridges of the finger tips. There may be less DNA transferred by an individual who washes their hands more frequently [51, 53], as more cells and DNA will be removed from the skin via washing.

Personal habits

Individuals who have an increased tendency to touch their hair and face often may increase the likelihood of transferring DNA onto objects via touch [50, 55]. Skin conditions have been linked to the quantity of DNA left behind at crime scenes [27]. People with atopic dermatitis or psoriasis conditions are found to contribute more DNA left by skin contact than individuals without these conditions [27].

Emotional response of the individual

Perspiration may contribute to a portion of the cell-free DNA left behind by touch [44, 56]. Quinones *et al.* (2011) [44] reported up to 11.5 ng of DNA in 1 mL of cell-free sweat. Wickenheiser *et al.* (2002) [55] report that additional cells are transferred to the surface of the skin as sweat passes through an individual's pores. The authors

hypothesize that a person who tends to sweat more has the potential to transfer more DNA to an object through contact compared to an individual at rest state.

Environmental conditions

If DNA has been exposed to prolonged elements of heat, humidity, UV light and bacterial growth (i.e. typical of an outdoor environment) there may be no surviving endogenous DNA left to capture [57].

Potential for mixture

Many everyday objects that are targeted for 'touch' DNA have been handled by multiple users (e.g. door handles), and the presence of a mixed DNA profile may cause an issue with profile interpretation [58]. A study by Daly *et al.* (2010) [59] report that out of 300 samples analysed for 'touch' DNA content approximately 10% exhibit mixed DNA profiles as a result of secondary transfer.

Preservation of 'touch' DNA

DNA may not survive long outside of its natural environment in the cell. If a long time has passed between depositing DNA and swabbing taking place there may be less surviving DNA to recover. Raymond *et al.* (2009) investigated the persistence of DNA that had been transferred onto objects by touch [41]. DNA was preserved for up to 2 weeks in most cases where DNA was exposed to an outdoor environment or 6 weeks if samples were preserved indoors. The authors reported that a full DNA profile could be obtained from a drug balloon 55 days after the offense was committed and a full (but mixed) DNA profile from a laptop seized 62 days after the crime [41]. Persistence of DNA appears to be influenced by the amount of DNA initially transferred. A second study published by Linacre *et al.* (2010) [60] used direct PCR on fabric to yield 'touch' DNA after volunteers handled fibres briefly for 5 seconds, and results demonstrate near complete (PowerPlex® 16 System) DNA profiles 36 days after transfer. It is clear from the literature that additional studies are required in order to investigate the persistence of 'touch' DNA on substrates which may assist future criminal investigations.

Secondary transfer

A common question asked in court is, “Is it possible that my client’s DNA transferred onto the item via secondary transfer?” Currently it is not possible to determine the last person that handled an item, as there are many possible scenarios.

Secondary transfer is a result of:

1) ***Person to person to object transfer***. For example, two people shake hands briefly, person 1 then picks up an object (e.g. weapon) and transfers DNA from person 2 onto the weapon. However, person 2 did not physically touch the object. For example, Samie *et al.* (2016) [20] conducted a study to examine the presence of mixed DNA profiles on a knife used in stabbing simulations and concluded that mixed DNA was present in 61 out of 64 cases, despite cleaning the weapon prior to handling.
2) ***Person to object to person transfer***. For example, one individual (person 1) transfers their DNA onto an object by handling it (e.g. weapon). A second individual then handles the same object and consequently picks up DNA from the first individual who handled the object initially. One study providing insight into the complexity of DNA transfer during a social setting highlight that simple everyday activities can lead to the innocent transfer of DNA without the individuals having direct contact [61].
3) ***Person to object to object transfer***. For example, an individual may have used a bath towel (transferring their DNA to the towel). The towel is then used to wrap around a weapon, and the DNA from the individual in contact with the towel was then transferred to the weapon. Alternate scenarios of how DNA is transferred need to be considered.

Studies exploring DNA transfer have raised more questions than they have answered [26], due to the complex nature of ‘touch’ DNA. A study conducted by Ladd *et al.* (1999) [62] carried out controlled experimental conditions in their laboratory to monitor the possibility of primary and secondary transfer. They concluded that secondary transfer was not observed. A closer look reveals that secondary transfer was observed, as low-level peaks (i.e. < 75 RFU) from a second individual had been detected (i.e. additional minor peaks). Following the increased

sensitivity of STR profiling kits, it is now more likely that mixtures and secondary transfer events will be detected.

Lowe *et al.* (2002) [51] reported that secondary transfer was evident in a case when one individual held hands with another individual and then handled a plastic tube - transferring DNA from the second individual only. Goray *et al.* (2010) [63] performed a detailed study of secondary DNA transfer and found that porous surfaces permit more DNA to be transferred than non-porous surfaces. Transfer rates approximately doubled when increased pressure was placed on the surface, and transfer rates increased even more when friction was used [63]. Later, a second study by Goray *et al.* (2012) [64] discussed the likelihood of DNA transfer during the collection and handling of items relating to crime exhibits. The authors set up a 'mock' case to package and handle items at a crime scene in order to mimic a real case. Not surprisingly, their results showed that, for items packaged together, DNA transferred between items of evidence (e.g. cigarette butts and blood stains) and to the interior of the packaging itself, leading to inaccurate DNA profiling results. Similar findings were observed by Poy *et al*. (2006) [65].

Currently, it is not possible to determine how long 'touch' DNA persists on an object or to probabilistically determine the mode of transfer (i.e. primary, secondary or tertiary contact) as there are multiple factors to consider. However, further studies may indicate how long 'touch' DNA will survive under controlled laboratory conditions, and this may prove to be useful for interpreting casework results and assisting in court.

The main aim for the current chapter is to investigate the application of direct PCR on the profiling ability of 'touch' DNA swabs, to observe DNA transfer, and to apply the methodology to a real case exhibit.

4.4 Buffer enhancement

The effect of PCR buffer enhancers on PCR was investigated in an attempt to design a robust system that limits stochastic effects (see manuscript enclosed).

In the following study, the well-known PCR facilitators DMSO, BSA and Triton™-X-100 were shown to increase the amplification yield of STR-based DNA profiling results by increasing the RFU value of peak heights. Results were included in a manuscript and submitted to a journal and rejected: ***"R. Blackie, J. Templeton, D.A. Taylor, A. Linacre. PCR buffer enhancement of STR kits used for human identification. International Journal of Legal Medicine,"*** *(see below).* It is assumed from the results that the addition of PCR facilitators and extra Ampli*Taq* Gold® DNA polymerase will help to boost the PCR and overcome inhibitors that could otherwise be detrimental to the DNA and prevent amplification. For AmpFℓSTR® ProfilerPlus™ kit at 29 cycles, a combination of DMSO and BSA or an alternative combination of DMSO and Triton X-100™ led to an increase in the quantity of STR products. Similarly, for NGM SElect™ kit at 29 cycles, the addition of BSA enhanced the profiling results by increasing the quantity of STR products. A more in-depth study (with a larger sample size and assessment of different sample types) would be required in order to determine the effect of these buffers on the direct PCR approach, however, preliminary data is promising and are enclosed in the manuscript provided.

4.4.1 Statement of authorship

Manuscript: PCR buffer enhancement of STR kits used for human identification

Manuscript prepared as: a letter to the editor, International Journal of Legal Medicine.

Jennifer Templeton (Candidate)

Assisted with experimental design, provided laboratory results for the use of Triton X-100™ as a buffer enhancer, and edited the paper

I hereby certify that the statement of contribution is accurate.

Signed: Date: November 2016

Renée Blackie (nee Ottens)

Performed laboratory work and wrote the paper

I hereby certify that the statement of contribution is accurate.

Signed: Date: November 2016

Duncan Taylor

Assisted with experimental design, commented on data, and edited the paper

I hereby certify that the statement of contribution is accurate

Signed: Date: November 2016

Adrian Linacre (Supervisor)

Supervised the project, assisted with experimental design, and commented on data

I hereby certify that the statement of contribution is accurate

Signed: Date: November 2016

4.4.2 Manuscript: PCR buffer enhancement of STR kits used for human identification

PCR buffer enhancement of STR kits used for human identification

--Manuscript Draft--

Manuscript Number:	IJLM-D-14-00097
Full Title:	PCR buffer enhancement of STR kits used for human identification
Article Type:	Letter to the Editor
Corresponding Author:	Renee Ottens Flinders University Bedford Park, South Australia AUSTRALIA
Corresponding Author Secondary Information:	
Corresponding Author's Institution:	Flinders University
Corresponding Author's Secondary Institution:	
First Author:	Renee Ottens
First Author Secondary Information:	
Order of Authors:	Renee Ottens
	Jennifer Templeton
	Duncan Taylor
	Adrian Linacre
Order of Authors Secondary Information:	
Abstract:	We report on the significant improvement of DNA profiles by the addition of standard chemicals to the polymerase chain reaction (PCR) buffer provided in commercially available human identification kits. Triton® X-100, BSA and DMSO were added to the PCR buffer, both singly and in combinations, provided in the AmpFλSTR® Profiler Plus® or NGM SElect™ kit (Life Technologies, Victoria, Australia). PCRs were performed in 0.2 mL thin walled tubes following the standard protocol for a 25 L volume reaction, following all manufacturers' conditions for amplification. PCR products were analysed using a 3130xl (ABI) and GeneMapper ID v3.2 (Life Technologies) with a threshold of 50 relative fluorescence units (RFU) for allele assignment. The average RFU was calculated for each profile obtained where the buffer was adjusted. The values were compared against positive and negative controls where no buffer adjustments were made. A 70 % increase in RFU value was observed with the addition of a combination of DMSO with either BSA or Triton X when using Profiler Plus®, and a 35.7 % increase was observed with the singular addition of BSA for NGM SElect™. Our data indicate the clear improvement in the quantity of profiles obtained across both kits, offering a fast and cost effective way to boost the results. This enhancement may be beneficial in cases where limited or degraded DNA is present, offering further assistance in difficult investigations.
Author Comments:	Dear Editor, We submit a Letter to the Editor entitled 'PCR buffer enhancement of STR kits used for human identification'. We confirm that these data have not been presented as yet at any conference nor been published in any scientific journal. We also confirm that all four authors have contributed to the manuscript and their in-put justifies their inclusion as an author. Funding for the research that supported this work was provided by the Department of Justice South the Australia. This is acknowledged as a footnote on the title page. The authors are not aware of any conflict of interest, financial or scientific. We thank you in advance for your assistance and look forward to receiving comments from any reviewers.

Dear Editor,

We show how the addition of standard chemicals to the PCR buffer provided in commercially available STR kits improve the quality of DNA profiles by significantly increasing the RFU value. The factors that affect the quantity of PCR product include the initial quantity of DNA template and the presence of inhibitors of the amplification and the buffer components. Triton® X-100, BSA and DMSO have been shown to increase the amount of PCR product produced and overcome inhibitors to the amplification [1, 2] and hence these standard materials were added to the PCR buffer, both singly and in combinations, provided in two commercially supplied STR kits, and the resulting data analysed. Although this experiment uses control DNA, in which there are no inhibitors present, the effect of each buffer adjustment should be documented first to indicate which addition will potentially have the greatest impact for the use of more difficult sample sets such as trace DNA or overcoming inhibitors in direct PCR.

PCRs were performed using either the AmpFλSTR® Profiler Plus® or NGM SElect™ kit (Life Technologies, Victoria, Australia) in 0.2 mL thin walled tubes following the standard protocol for a 25 µL volume reaction. Positive and negative controls acted as the standards in this experiment with no additions made to their buffers. Final concentrations of Triton X (0.004 %), DMSO (4 %), BSA (0.008 µg) were added to PCR buffers in the following combinations: Triton X (T), DMSO (D), BSA (B), B/D, T/D, and T/B/D. All PCRs contained 1 ng of control DNA (2800M, Promega, Victoria, Australia). Amplification conditions for all samples followed the manufacturer's instructions at 29 cycles using a GeneAmp® 9700 96-well thermal cycler (Life Technologies). PCR products were analysed using a 3130*xl* (ABI) and GeneMapper® ID v3.2 (Life Technologies) with a threshold of 50 RFU for allele assignment. The average RFU value of a profile (9 loci Profiler Plus® or 15 loci NGM SElect™) for each modified PCR was compared against the average RFU value of the standard (positive control) to determine the effect of the buffer adjustments. For each kit, each buffer addition (six combinations), and positive and negative controls was set up three

times and run in triplicate on the 3130*xl*, resulting in a total of 144 samples used in this experiment. No other combinations were trialled in this experiment.

The effect of the buffer adjustments is illustrated in Figure 1, which shows the increase in RFU values for both STR kits using the additions to the buffer. PCR products obtained using the Profiler Plus® kit showed the greatest increase in RFU value across all buffer adjustments compared to the NGM SElect™ kit. The combinations of BSA/DMSO and Triton X/DMSO resulted in an RFU increase of approximately 70 % across all nine loci. All RFU increases using this kit were statistically significant ($p = < 0.001$). Using NGM SElect™, the average RFU increase ranged from 1.4 % to 35.7 %. The only statistically significant increase was observed with the singular addition of BSA ($p = < 0.05$). Overall, an increase in RFU value was seen for each addition and combination across both kits.

Our data indicate the clear improvement in the quantity of STR products produced with the addition of a combination of DMSO with either BSA or Triton X when using Profiler Plus®, and the addition of BSA for NGM SElect™. This offers a fast and cost effective way to boost the quality of results obtained for these kits. Such an enhancement may be crucial in cases where there is trace DNA. Direct PCR incorporates limited DNA template into the buffer to enhance the profiles obtained [3, 4], coupled with the buffer additions outlined in this study could further aid difficult investigations.

Acknowledgements

We thank members of Forensic Science South Australia and the School of Biological Sciences at Flinders University for assisting with reagents and providing helpful advice. Funding was provided by the Department of Justice, South Australia.

[1] Strien J, Sanft J, Mall G. Enhancement of PCR amplification of moderate GC-containing and highly GC-rich DNA sequences. Mol Biotechnol. 2013;54:1048-54.

[2] Liu YS, Thomas RJS, Phillips WA. SINGLE-STEP DIRECT PCR AMPLIFICATION FROM SOLID TISSUES. Nucleic Acids Res. 1995;23:1640-.

[3] Ottens R, Templeton J, Paradiso V, Taylor D, Abarno D, Linacre A. Application of direct PCR in forensic casework. Forensic Science International: Genetics Supplement Series. 2013;4:e47-e8.

[4] Templeton J, Ottens R, Paradiso V, Handt O, Taylor D, Linacre A. Genetic profiling from challenging samples: Direct PCR of touch DNA. Forensic Science International: Genetics Supplement Series. 2013;4:e224-e5.

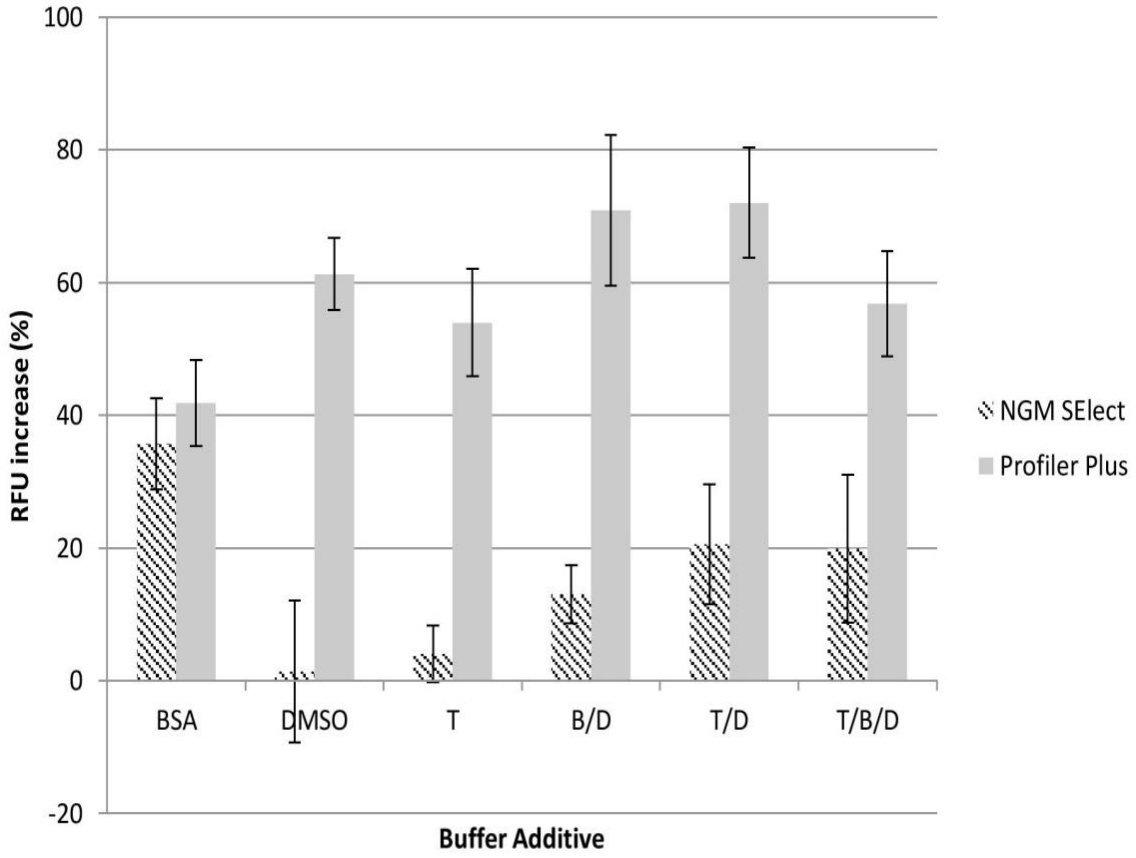


Figure 1. Known buffer enhancers BSA, DMSO and Triton-X were added individually and in combination to two standard STR kit buffers. The average RFU value across all loci for a sample was compared against the RFU value of the positive control sample (no buffer additions). The increase in RFU value from the positive control was recorded in per cent (%) and shown above across all enhancement types for both STR kits. Standard Error (SE) bars are displayed for n = 9 replicates of each buffer additive.

4.5 Direct PCR from fingermarks

4.5.1 Materials and Methods

Materials and methods that are not described in the published manuscripts for Chapter IV are detailed below.

Contamination controls

All contamination measures and controls are indicated in Chapter II. Latent DNA on handled or touched evidentiary items and collection devices (e.g. swabs) were processed and analyzed in a similar manner to samples processed in a low-template DNA/ancient DNA environment (see Chapter II). Strict procedures were followed to avoid the potential risk of contamination with contemporary human DNA or previously amplified PCR products. Extreme caution and care was adhered to at all stages.

Sample controls

Positive PCR controls consisted of 1 ng of control DNA (2800M, Promega) or control DNA supplied with the amplification kit in use. Negative controls were set up consisting of a PCR blank (i.e. no DNA); and a swab of a plastic slide (Rinzl plastic, ProSci Tech, QLD, AU) containing no DNA (to monitor slide and swab contamination). Amplification followed using the AmpFℓSTR® ProfilerPlus™ or NGM SElect kit™ guidelines.

Sampling

Ten individuals were selected to provide fingermarks (all five digits of their dominant hand) on sterile plastic microscope slides (Rinzl plastic, ProSci Tech, QLD, AU) 15 minutes, 1 hour and 2 hours after hand washing, all marks were created on the same day. The donor was asked to wash and then dry their hands using fresh paper towel. A fingermark was created on the slide using the individual's dominant hand at the designated time period (i.e. T15 mins = 15 minutes, T1hr = 1 hour, or T2hr = 2 hours post hand washing). The mark was created on the slide using medium pressure for 15 seconds to ensure consistency between testing. Activities that the donor took part in between hand washing and touching slides were recorded during the study.

The slides were subjected to 'targeted' swabbing using pre-cut fibre clumps of DNA-free nylon FLOQswabs™ (see Chapter II). Post-swabbing, the pre-cut ~ 2 mm^2 fibre clump was added directly to a 0.2 mL PCR tube using sterile forceps. Double swabbing was performed using a second moistened DNA-free nylon FLOQswab™. Post swabbing, the pre-cut ~ 2 mm^2 fibre clump was added directly to the same 0.2 mL PCR tube using sterile forceps. The PCR tube consisted of 2 x 2 mm^2 fibre clumps. Amplification was carried out using the AmpFℓSTR® ProfilerPlus™ kit with an additional 1 µL of DMSO (5%) (Expand Long Range DNTPACK; Roche, Vic, AU), 1 µL of Molecular Biology Grade BSA (0.1 µg, New England Biolabs, NSW, AU) and 1 µL Ampli*Taq* Gold® (5 units) (Life Technologies) included in the PCR set up.

Direct PCR from fingermarks immediately after hand washing

This study involved selecting two individuals from the same sampling set based on their ability to deposit enough DNA consistently to generate profiles. The donors were asked to wash their hands and dry them using fresh paper towel. The donor deposited fingermarks (all five digits of their dominant hand) onto sterile plastic microscope slides (Rinzl plastic, ProSci Tech, QLD, AU) immediately after hand washing. The process of applying fingermarks to slides was the same as detailed above. The slide was subjected to 'targeted' swabbing and swab fibres subjected to direct PCR using the conditions described above.

Extracting DNA from fingermarks using the DNA IQ™ System

Two individuals were selected for this study based on their ability to deposit enough DNA consistently to generate profiles. The donors were asked to wash their hands and dry them using paper towel. The donor deposited fingermarks (all five digits of their dominant hand) onto sterile plastic microscope slides (Rinzl plastic, ProSci Tech, QLD, AU) 15 minutes after hand washing. The process of applying fingermarks to slides was the same as detailed above. The slide was subjected to 'double' swabbing (see Chapter II) using DNA-free nylon FLOQswabs™. The whole swabs were extracted following the DNA IQ™ System (Promega Corporation, AU) manufacturer's recommendations. Amplification was carried out using the AmpFℓSTR® ProfilerPlus™ kit (using the conditions described above).

Capillary Electrophoresis and Data Analysis - See Chapter II

For the data analysis steps, a ‘full’ DNA profile was defined when all the expected alleles, in all loci, were observed. Full DNA profiles did not show ‘allele drop-out’ in any loci.

4.5.2 Results

Direct PCR from fingermarks, 15 minutes, 1 hour and 2 hours post hand washing

Single amplification results generated from fingermarks using direct PCR are reported below (see Tables 4.1 – 4.10). Ten volunteers deposited fingermarks onto plastic substrates 15 minutes, 1 hour and 2 hours post hand washing. Swabs were rubbed on the surface that made contact with the fingermark deposit, and the swab fibres were processed by direct PCR using AmpFℓSTR® ProfilerPlus™ kit at 29 cycles.

Success rate is measured by the number of donor alleles detected (out of 20 alleles for AmpFℓSTR® ProfilerPlus™ kit) and the quality of the profiles. Columns highlighted in blue indicate the number of samples that would be considered ‘up-loadable’ to the Australian DNA database (i.e. > 12 alleles, plus Amelogenin).

DNA profiling results for volunteers – direct PCR

Volunteer 1 – Donor DNA profiling results

Table 4.1 STR-based DNA results obtained for volunteer 1, showing success rate using AmpFℓSTR® ProfilerPlus™ kit (of a possible 20 STR alleles, 10 total loci). Profiles obtained using direct PCR after the donor's fingermarks (all five digits) were deposited on a substrate 15 minutes, 1 hour and 2 hours post hand washing. The activity of the donor was recorded during this time period.

Profiling success	Activity during 2 hour time period for volunteer 1: Typing at keyboard and reading			
Time (post hand washing):	Finger used for depositing DNA:	Number of donor alleles detected (out of 20):	% donor alleles detected:	Profile up-loadable to database (i.e. > 12 alleles):
15 minutes	THUMB	20	100%	Yes
	INDEX	20	100%	Yes
	MIDDLE	20	100%	Yes
	RING	20	100%	Yes
	PINKY	20	100%	Yes
1 hour	THUMB	20	100%	Yes
	INDEX	20	100%	Yes
	MIDDLE	20	100%	Yes
	RING	20	100%	Yes
	PINKY	20	100%	Yes
2 hour	THUMB	20	100%	Yes
	INDEX	20	100%	Yes
	MIDDLE	20	100%	Yes
	RING	20	100%	Yes
	PINKY	20	100%	Yes

A maximum of 20 alleles (column 3) indicates a full DNA profile. Numbers less than 20 indicate a partial or failed profile. All fingers depositing a fingermark in this study generated a full DNA profile using the direct PCR approach 15 minutes, 1 hour and 2 hour post hand washing. **All profiles highlighted in blue are considered 'up-loadable' to the Australian DNA Database (i.e. > 12 alleles, plus Amelogenin).**

Volunteer 2 – Donor DNA profiling results

Table 4.2 STR-based DNA results obtained for volunteer 2, showing success rate using AmpFℓSTR® ProfilerPlus™ kit (of a possible 20 STR alleles, 10 total loci). Profiles obtained using direct PCR after the donor's fingermarks (all five digits) were deposited on a substrate 15 minutes, 1 hour and 2 hours post hand washing. The activity of the donor was recorded during this time period.

Profiling success	**Activity during 2 hour time period for volunteer 2: Watching TV**			
Time (post hand washing):	**Finger used for DNA deposition:**	**Number of donor alleles detected (out of 20):**	**% donor alleles detected:**	**Profile up-loadable to database (i.e. > 12 alleles):**
15 minutes	THUMB	20	100%	Yes
	INDEX	3	15%	No
	MIDDLE	0	0%	No
	RING	20	100%	Yes
	PINKY	20	100%	Yes
1 hour	THUMB	20	100%	Yes
	INDEX	6	30%	No
	MIDDLE	20	100%	Yes
	RING	20	100%	Yes
	PINKY	5	25%	No
2 hour	THUMB	20	100%	Yes
	INDEX	0	0%	No
	MIDDLE	20	100%	Yes
	RING	20	100%	Yes
	PINKY	7	35%	No

A maximum of 20 alleles (column 3) indicates a full DNA profile. Numbers less than 20 indicate a partial or failed profile. Nine fingers out of 15 depositing a fingermark in this study generated a full DNA profile using the direct PCR approach. **Nine profiles (60%) highlighted in blue are considered 'up-loadable' to the Australian DNA Database (i.e. > 12 alleles, plus Amelogenin).**

Volunteer 3 – Donor DNA profiling results

Table 4.3 STR-based DNA results obtained for volunteer 3, showing success rate using AmpFℓSTR® ProfilerPlus™ kit (of a possible 20 STR alleles, 10 total loci). Profiles obtained using direct PCR after the donor's fingermarks (all five digits) were deposited on a substrate 15 minutes, 1 hour and 2 hours post hand washing. The activity of the donor was recorded during this time period.

Profiling success	**Activity during 2 hour time period for volunteer 3: Watching TV**			
Time (post hand washing):	**Finger used for DNA deposition:**	**Number of donor alleles detected (out of 20):**	**% donor alleles detected:**	**Profile up-loadable to database (i.e. > 12 alleles):**
15 minutes	THUMB	20	100%	Yes
	INDEX	12	60%	Yes
	MIDDLE	19	95%	Yes
	RING	18	90%	Yes
	PINKY	11	55%	No
1 hour	THUMB	0	0%	No
	INDEX	0	0%	No
	MIDDLE	0	0%	No
	RING	6	30%	No
	PINKY	6	30%	No
2 hour	THUMB	6	30%	No
	INDEX	0	0%	No
	MIDDLE	6	30%	No
	RING	15	75%	Yes
	PINKY	20	100%	Yes

A maximum of 20 alleles (column 3) indicates a full DNA profile. Numbers less than 20 indicate a partial or failed profile. Two fingers out of 15 depositing a fingermark in this study generated a full DNA profile using the direct PCR approach. **Six profiles (40%) highlighted in blue are considered 'up-loadable' to the Australian DNA Database (i.e. > 12 alleles, plus Amelogenin).**

Volunteer 4 – Donor DNA profiling results

Table 4.4 STR-based DNA results obtained for volunteer 4, showing success rate using AmpFℓSTR® ProfilerPlus™ kit (of a possible 20 STR alleles, 10 total loci). Profiles obtained using direct PCR after the donor's fingermarks (all five digits) were deposited on a substrate 15 minutes, 1 hour and 2 hours post hand washing. The activity of the donor was recorded during this time period.

Profiling success	**Activity during 2 hour time period for volunteer 4: Writing and reading**			
Time (post hand washing):	**Finger used for DNA deposition:**	**Number of donor alleles detected (out of 20):**	**% donor alleles detected:**	**Profile up-loadable to database (i.e. > 12 alleles):**
15 minutes	THUMB	12	60%	Yes
	INDEX	20	100%	Yes
	MIDDLE	20	100%	Yes
	RING	0	0%	No
	PINKY	0	0%	No
1 hour	THUMB	5	25%	No
	INDEX	15	75%	Yes
	MIDDLE	2	10%	No
	RING	16	80%	Yes
	PINKY	2	10%	No
2 hour	THUMB	19	95%	Yes
	INDEX	20	100%	Yes
	MIDDLE	4	20%	No
	RING	2	10%	No
	PINKY	9	45%	No

A maximum of 20 alleles (column 3) indicates a full DNA profile. Numbers less than 20 indicate a partial or failed profile. Three fingers out of 15 depositing a fingermark in this study generated a full DNA profile using the direct PCR approach. **Seven profiles (46%) highlighted in blue are considered 'up-loadable' to the Australian DNA Database (i.e. > 12 alleles, plus Amelogenin).**

Volunteer 5 – Donor DNA profiling results

Table 4.5 STR-based DNA results obtained for volunteer 5, showing success rate using AmpFℓSTR® ProfilerPlus™ kit (of a possible 20 STR alleles, 10 total loci). Profiles obtained using direct PCR after the donor's fingermarks (all five digits) were deposited on a substrate 15 minutes, 1 hour and 2 hours post hand washing. The activity of the donor was recorded during this time period.

Profiling success	**Activity during 2 hour time period for volunteer 5: Eating and drinking**			
Time (post hand washing):	**Finger used for DNA deposition:**	**Number of donor alleles detected (out of 20):**	**% donor alleles detected:**	**Profile up-loadable to database (i.e. > 12 alleles):**
15 minutes	THUMB	17	85%	Yes
	INDEX	15	75%	Yes
	MIDDLE	8	40%	No
	RING	1	5%	No
	PINKY	6	30%	No
1 hour	THUMB	12	60%	Yes
	INDEX	4	20%	No
	MIDDLE	0	0%	No
	RING	17	85%	Yes
	PINKY	2	10%	No
2 hour	THUMB	13	65%	Yes
	INDEX	2	10%	No
	MIDDLE	4	20%	No
	RING	18	90%	Yes
	PINKY	19	95%	Yes

A maximum of 20 alleles (column 3) indicates a full DNA profile. Numbers less than 20 indicate a partial or failed profile. There were no full DNA profiles generated in this study. **Seven profiles (46%) highlighted in blue are considered 'up-loadable' to the Australian DNA Database (i.e. > 12 alleles, plus Amelogenin).**

Volunteer 6 – Donor DNA profiling results

Table 4.6 STR-based DNA results obtained for volunteer 6, showing success rate using AmpFℓSTR® ProfilerPlus™ kit (of a possible 20 STR alleles, 10 total loci). Profiles obtained using direct PCR after the donor's fingermarks (all five digits) were deposited on a substrate 15 minutes, 1 hour and 2 hours post hand washing. The activity of the donor was recorded during this time period.

Profiling success	**Activity during 2 hour time period for volunteer 6: Typing on keyboard and reading**			
Time (post hand washing):	**Finger used for DNA deposition:**	**Number of donor alleles detected (out of 20):**	**% donor alleles detected:**	**Profile up-loadable to database (i.e. > 12 alleles):**
15 minutes	THUMB	20	100%	Yes
	INDEX	10	50%	No
	MIDDLE	17	85%	Yes
	RING	6	30%	No
	PINKY	20	100%	Yes
1 hour	THUMB	17	85%	Yes
	INDEX	10	50%	No
	MIDDLE	1	5%	No
	RING	19	95%	Yes
	PINKY	4	20%	No
2 hour	THUMB	18	90%	Yes
	INDEX	2	10%	No
	MIDDLE	1	5%	No
	RING	4	20%	No
	PINKY	1	5%	No

A maximum of 20 alleles (column 3) indicates a full DNA profile. Numbers less than 20 indicate a partial or failed profile. Two fingers out of 15 depositing a fingermark in this study generated a full DNA profile using the direct PCR approach. **Six profiles (40%) highlighted in blue are considered 'up-loadable' to the Australian DNA Database (i.e. > 12 alleles, plus Amelogenin).**

Volunteer 7 – Donor DNA profiling results

Table 4.7 STR-based DNA results obtained for volunteer 7, showing success rate using AmpFℓSTR® ProfilerPlus™ kit (of a possible 20 STR alleles, 10 total loci). Profiles obtained using direct PCR after the donor's fingermarks (all five digits) were deposited on a substrate 15 minutes, 1 hour and 2 hours post hand washing. The activity of the donor was recorded during this time period.

Profiling success	**Activity during 2 hour time period for volunteer 7: Typing on computer and touching mobile phone**			
Time (post hand washing):	Finger used for DNA deposition:	Number of donor alleles detected (out of 20):	% donor alleles detected:	Profile up-loadable to database (i.e. > 12 alleles):
15 minutes	THUMB	19	95%	Yes
	INDEX	20	100%	Yes
	MIDDLE	20	100%	Yes
	RING	13	65%	Yes
	PINKY	4	20%	No
1 hour	THUMB	20	100%	Yes
	INDEX	20	100%	Yes
	MIDDLE	8	40%	No
	RING	18	90%	Yes
	PINKY	19	95%	Yes
2 hour	THUMB	20	100%	Yes
	INDEX	19	95%	Yes
	MIDDLE	2	10%	No
	RING	20	100%	Yes
	PINKY	11	55%	No

A maximum of 20 alleles (column 3) indicates a full DNA profile. Numbers less than 20 indicate a partial or failed profile. Six fingers out of 15 depositing a fingermark in this study generated a full DNA profile using the direct PCR approach. **Eleven profiles (73%) highlighted in blue are considered 'up-loadable' to the Australian DNA Database (i.e. > 12 alleles, plus Amelogenin).**

Volunteer 8 – Donor DNA profiling results

Table 4.8 STR-based DNA results obtained for volunteer 8, showing success rate using AmpFℓSTR® ProfilerPlus™ kit (of a possible 20 STR alleles, 10 total loci). Profiles obtained using direct PCR after the donor's fingermarks (all five digits) were deposited on a substrate 15 minutes, 1 hour and 2 hours post hand washing. The activity of the donor was recorded during this time period.

Profiling success	Activity during 2 hour time period for volunteer 8: Typing on keyboard and touching mobile phone			
Time (post hand washing):	**Finger used for DNA deposition:**	**Number of donor alleles detected (out of 20):**	**% donor alleles detected:**	**Profile up-loadable to database (i.e. > 12 alleles):**
15 minutes	THUMB	7	35%	No
	INDEX	9	45%	No
	MIDDLE	8	40%	No
	RING	2	10%	No
	PINKY	10	50%	No
1 hour	THUMB	20	100%	Yes
	INDEX	20	100%	Yes
	MIDDLE	20	100%	Yes
	RING	14	70%	Yes
	PINKY	12	60%	Yes
2 hour	THUMB	20	100%	Yes
	INDEX	19	95%	Yes
	MIDDLE	20	100%	Yes
	RING	20	100%	Yes
	PINKY	* CAUTION - MIXED! 14	75%	Yes

A maximum of 20 alleles (column 3) indicates a full DNA profile. Numbers less than 20 indicate a partial or failed profile. Six fingers out of 15 depositing a fingermark in this study generated a full DNA profile - that matched the donor - using the direct PCR approach. *One mixed DNA profile was obtained where the major contributor did not match the donor (alleles matched their partner's DNA) and was a result of secondary transfer from a mobile phone. **Ten profiles (66%) highlighted in blue are considered 'up-loadable' to the Australian DNA Database (i.e. > 12 alleles, plus Amelogenin).**

Volunteer 9 – Donor DNA profiling results

Table 4.9 STR-based DNA results obtained for volunteer 9, showing success rate using AmpFℓSTR® ProfilerPlus™ kit (of a possible 20 STR alleles, 10 total loci). Profiles obtained using direct PCR after the donor's fingermarks (all five digits) were deposited on a substrate 15 minutes, 1 hour and 2 hours post hand washing. The activity of the donor was recorded during this time period.

Profiling success	**Activity during 2 hour time period for volunteer 9: Exercise and reading**			
Time (post hand washing):	**Finger used for DNA deposition:**	**Number of donor alleles detected (out of 20):**	**% donor alleles detected:**	**Profile up-loadable to database (i.e. > 12 alleles):**
15 minutes	THUMB	20	100%	Yes
	INDEX	20	100%	Yes
	MIDDLE	20	100%	Yes
	RING	19	94%	Yes
	PINKY	20	100%	Yes
1 hour	THUMB	19	100%	Yes
	INDEX	20	100%	Yes
	MIDDLE	18	94%	Yes
	RING	20	100%	Yes
	PINKY	20	100%	Yes
2 hour	THUMB	20	100%	Yes
	INDEX	18	94%	Yes
	MIDDLE	20	100%	Yes
	RING	17	94%	Yes
	PINKY	17	88%	Yes

A maximum of 20 alleles (column 3) indicates a full DNA profile. Numbers less than 20 indicate a partial or failed profile. Ten fingers out of 15 depositing a fingermark in this study generated a full DNA profile using the direct PCR approach. **Fifteen profiles (100%) highlighted in blue are considered 'up-loadable' to the Australian DNA Database (i.e. > 12 alleles, plus Amelogenin).**

Volunteer 10 – Donor DNA profiling results

Table 4.10 STR-based DNA results obtained for volunteer 10, showing success rate using AmpFℓSTR® ProfilerPlus™ kit (of a possible 20 STR alleles, 10 total loci). Profiles obtained using direct PCR after the donor's fingermarks (all five digits) were deposited on a substrate 15 minutes, 1 hour and 2 hours post hand washing. The activity of the donor was recorded during this time period.

Profiling success	**Activity during 2 hour time period for volunteer 10: Typing on keyboard, answering communal phone used by others, and reading**			
Time (post hand washing):	Finger used for DNA deposition:	Number of donor alleles detected (out of 20):	% donor alleles detected:	Profile up-loadable to database (i.e. > 12 alleles):
15 minutes	THUMB	17	85%	Yes
	INDEX	5	25%	No
	MIDDLE	17	85%	Yes
	RING	16	80%	Yes
	PINKY	3	15%	No
1 hour	THUMB	9	45%	No
	INDEX	1	5%	No
	MIDDLE	15	75%	Yes
	RING	6	30%	No
	PINKY	18	90%	Yes
2 hour	THUMB	19	95%	Yes
	INDEX	12	60%	No
	MIDDLE	20	100%	Yes
	RING	19	95%	Yes
	PINKY	16	80%	Yes

A maximum of 20 alleles (column 3) indicates a full DNA profile. Numbers less than 20 indicate a partial or failed profile. One finger out of 15 depositing a fingermark in this study generated a full DNA profile using the direct PCR approach. **Ten profiles (66%) highlighted in blue are considered 'up-loadable' to the Australian DNA Database (i.e. > 12 alleles, plus Amelogenin).**

Summary of profiling results for ten volunteers depositing fingermarks onto plastic substrates

Table 4.11 Summary of STR-based DNA profiling success for 10 volunteers. Results obtained by direct PCR after the donor's fingermarks (all five digits) were deposited onto a plastic substrate 15 minutes, 1 hour and 2 hours post hand washing. Results in the table are a summary of those provided in tables 4.1 – 4.10.

Overall DNA profiling success for fingermarks generated by direct PCR							
Donor depositing DNA:	**Thumb profiling success (%):**	**Index finger profiling success (%):**	**Middle finger profiling success (%):**	**Ring finger profiling success (%):**	**Pinky finger profiling success (%):**	**Profiles considered 'up-loadable' to the Australian DNA database (out of 15):**	**Failed DNA profiles out of 15 (i.e. no alleles detected):**
Volunteer 1	**100**	**100**	**100**	**100**	**100**	**15 (100%)**	**0 (0%)**
Volunteer 2	**100**	**15**	**66**	**100**	**53**	**9 (60%)**	**2 (13%)**
Volunteer 3	**43**	**20**	**42**	**65**	**62**	**6 (40%)**	**4 (26%)**
Volunteer 4	**60**	**92**	**43**	**30**	**18**	**7 (46%)**	**2 (13%)**
Volunteer 5	**70**	**35**	**20**	**60**	**45**	**7 (46%)**	**1 (6%)**
Volunteer 6	**92**	**37**	**32**	**48**	**42**	**6 (40%)**	**0 (0%)**
Volunteer 7	**98**	**98**	**50**	**85**	**57**	**11 (73%)**	**0 (0%)**
Volunteer 8	**78**	**80**	**80**	**60**	**62**	**10 (66%)**	**0 (0%)**
Volunteer 9	**98**	**97**	**97**	**93**	**95**	**15 (100%)**	**0 (0%)**
Volunteer 10	**75**	**30**	**87**	**68**	**62**	**10 (66%)**	**0 (0%)**
Average:	***81***	***60***	***61***	***71***	***59***	***Total (out of 150): 96 (64%)***	***Total (out of 150): 9 (6%)***

Results for the number of profiles considered 'up-loadable' to the Australian DNA database (i.e. > 12 alleles, plus Amelogenin), full DNA profiles and failed DNA profiles, for the AmpFℓSTR® ProfilerPlus™ kit directly amplified samples. Overall profiling success ranged from 40% (see volunteer 3 and 6) to 100% (see volunteer 1).

4.5.2.2 Direct PCR profiling success – comparison between fingermarks depositing DNA

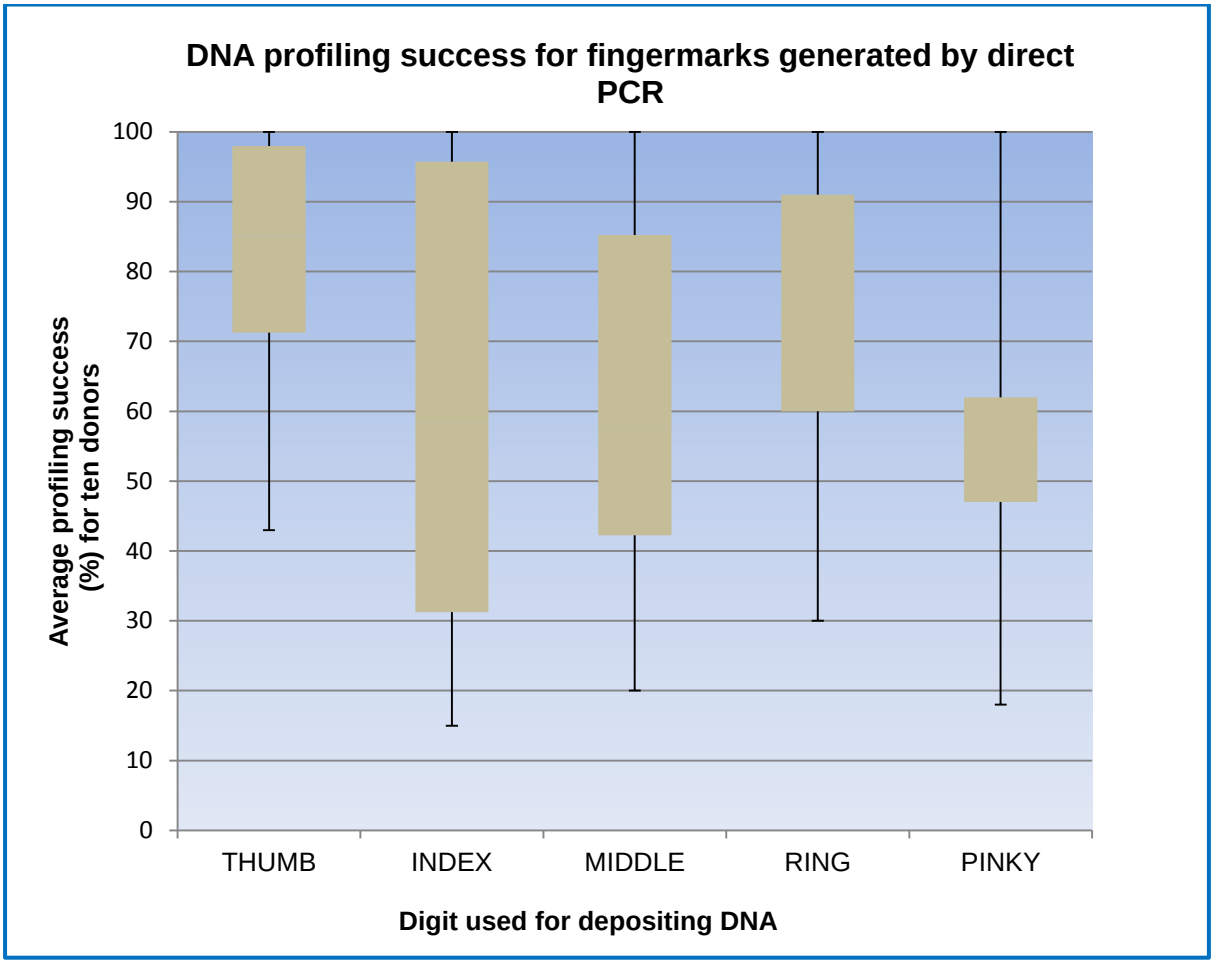


Figure 4.2 Box-and-whisker plot showing the relative distribution of DNA data generated from all five fingers depositing DNA. The average percentage profiling success is a result of testing 5 females and 5 males across three time periods after hand washing: 15 minutes, 1 hour and 2 hours, and subjecting swab fibres to direct PCR using the AmpFℓSTR® ProfilerPlus™ kit at 29 cycles.

4.5.2.3 Comparison of data obtained from fingermarks at three different time periods post hand washing: 15 minutes, 1 hour and 2 hour

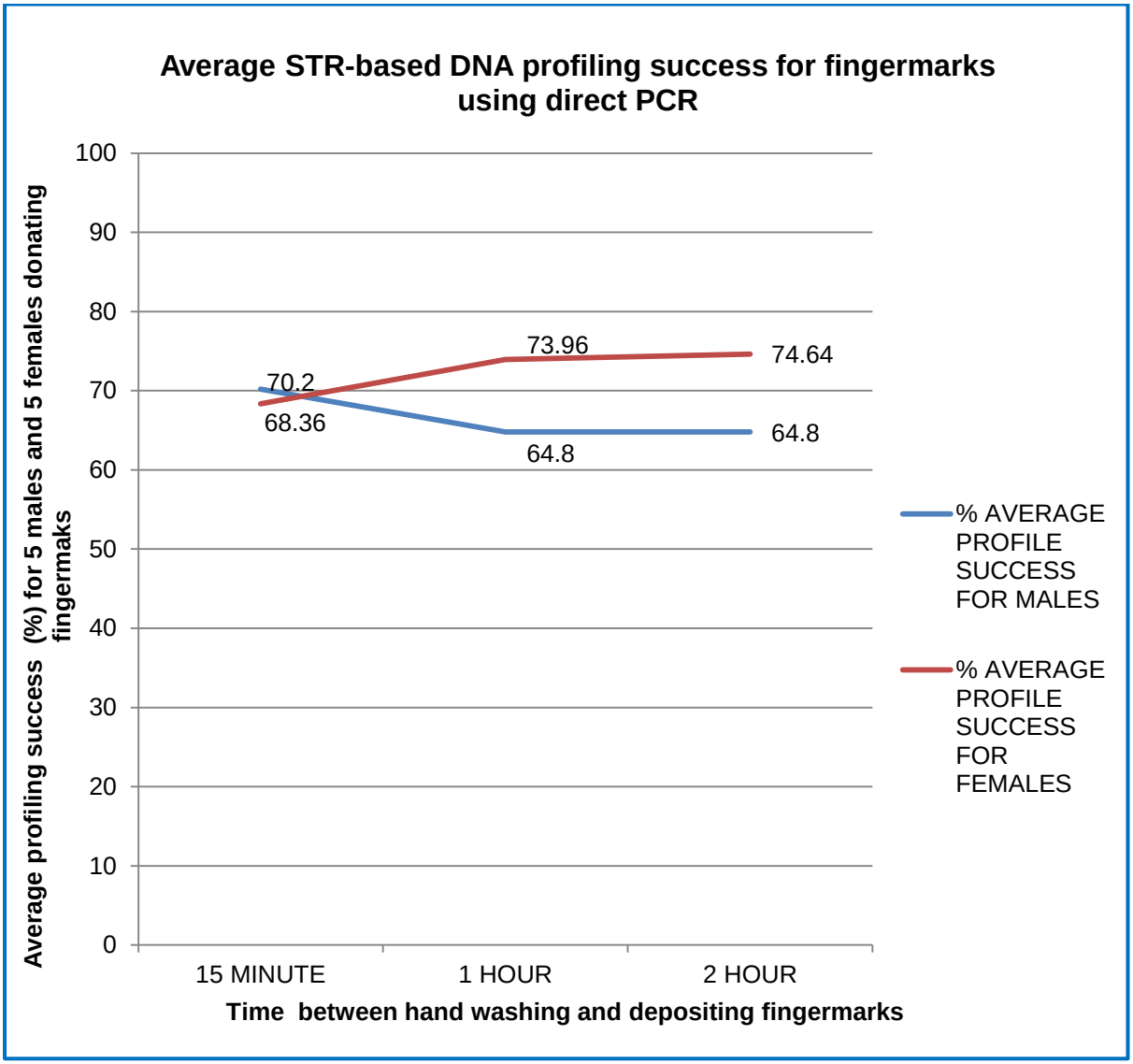


Figure 4.3 Showing the average DNA profiling success obtained from 5 males and 5 females fingermarks (from all five digits). Results were generated by using direct PCR and the AmpFℓSTR® ProfilerPlus™ kit at 29 cycles. For females, there is a moderate increase in profile success when a longer time passed since hand washing. In contrast, males exhibit a decrease in profiling success when a longer time passed since hand washing.

4.5.2.4 Direct PCR DNA profiles

DNA profile obtained from a fingermark deposit, using swab fibres and direct PCR with the AmpFℓSTR® ProfilerPlus™ kit at 29 cycles.

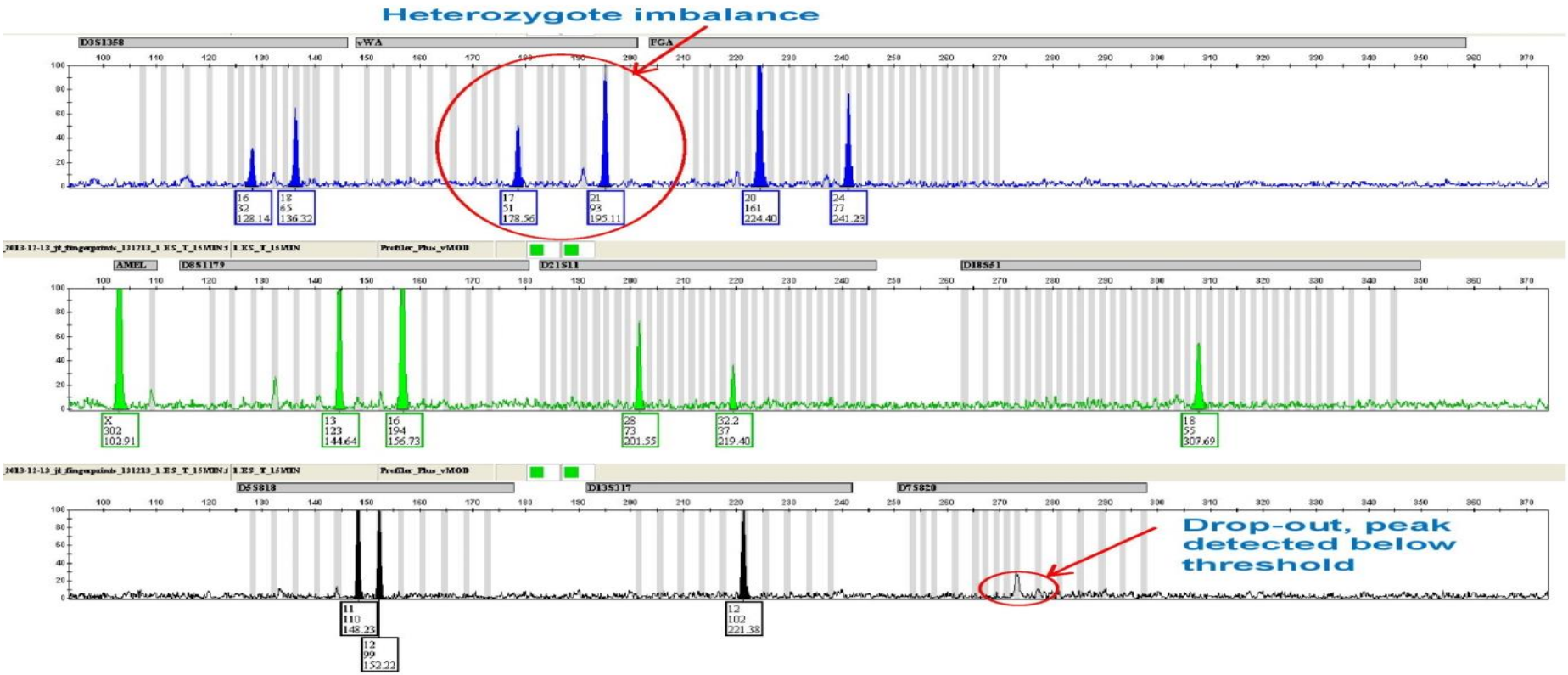


Figure 4.4 Example of a DNA profile obtained to illustrate heterozygote imbalance that was observed. DNA was recovered from a touched substrate (i.e. plastic slide) 15 minutes after the individual had washed their hands and deposited a thumb print. A partial DNA profile was obtained (i.e. 16 out of 20 possible alleles) with the AmpFℓSTR® ProfilerPlus™ kit; allele drop-out is reported at loci D7S820. Heterozygote imbalance was noted at most loci observed in this profile, particulary loci vWA. This profile would be considered 'up-loadable' to the Australian DNA database.

DNA profile obtained from a fingermark deposit, using swab fibres and direct PCR with the AmpFℓSTR® ProfilerPlus™ kit at 29 cycles.

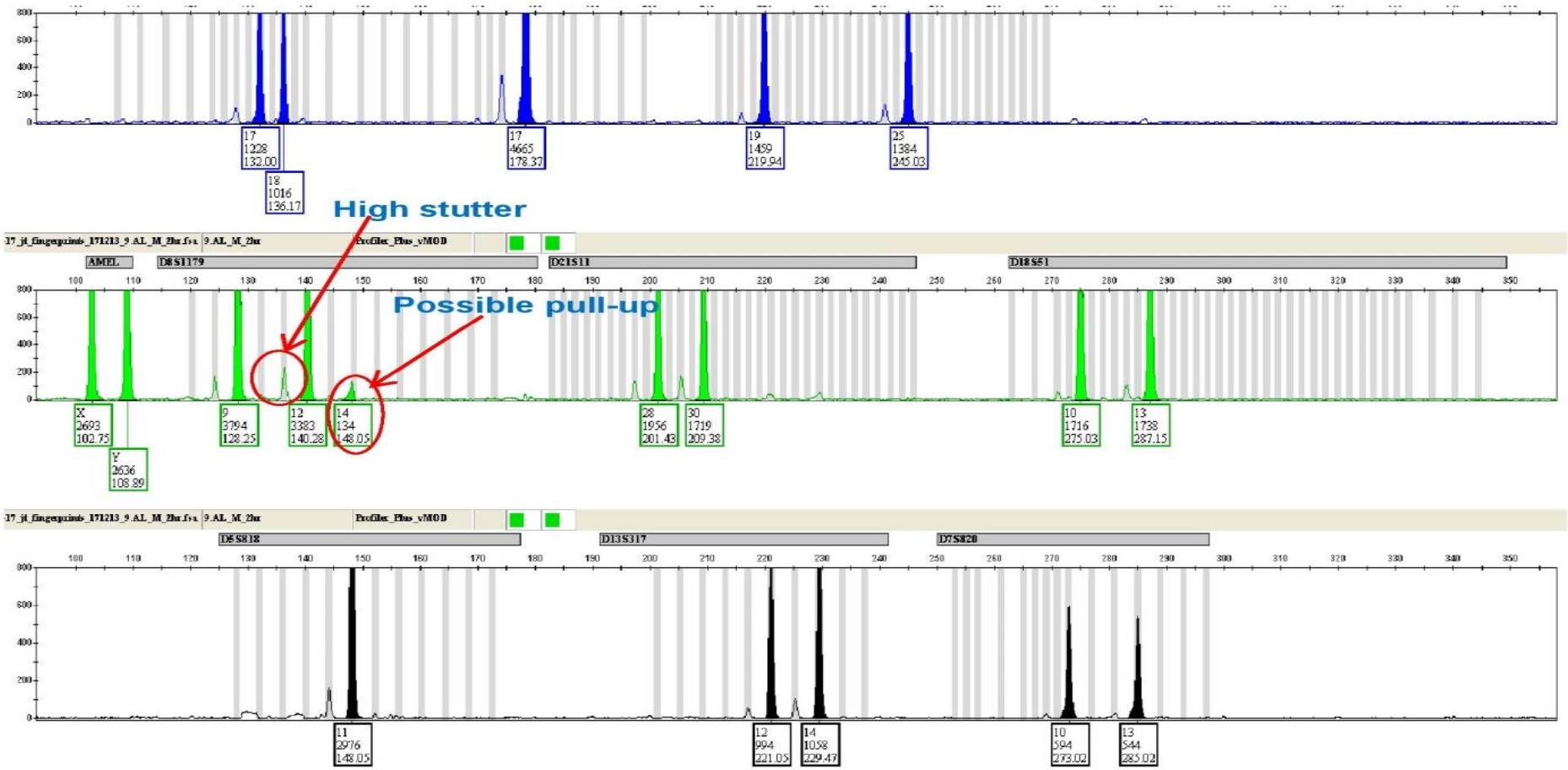


Figure 4.5 Example of a DNA profile recovered from a touched substrate (i.e. plastic slide) 2 hours after an individual had washed their hands and deposited a fingermark. A full DNA profile was obtained (i.e. 20 out of 20 possible alleles) with the AmpFℓSTR® ProfilerPlus™ kit; pull-up (possibly) and high stutter was observed at loci D8S1179.

DNA profile obtained from a fingermark deposit, using swab fibres and direct PCR with the AmpFℓSTR® ProfilerPlus™ kit at 29 cycles.

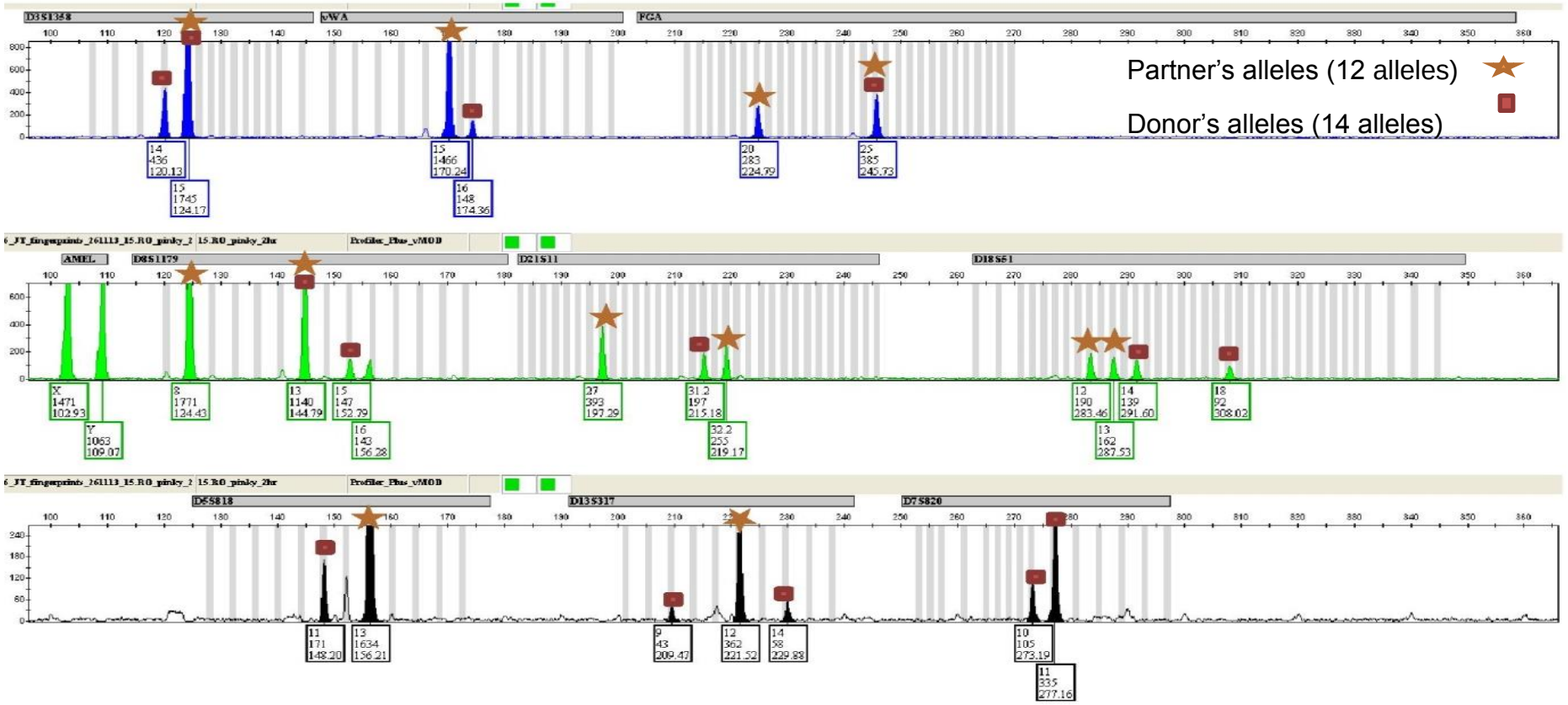


Figure 4.6. Example of a complex STR DNA profile produced as a result of secondary DNA transfer. DNA was recovered from a touched substrate (i.e. plastic slide) 2 hours after the individual had washed their hands and deposited a fingermark (i.e. pinky finger). A mixed DNA profile was observed, with minor alleles matching the donor of the print and major alleles matching the donor's partner - who had not been in contact with the substrate at all, or in contact with the donor for > 3 hours. Secondary transfer (i.e. person 1 to object to person 2 to object) was observed.

4.5.3 Discussion

Donor profiling success

A trend was apparent with some individuals consistently depositing more DNA than others (see Table 4.1 and Table 4.9). These results are consistent with other published findings by Lowe *et al.* (2002) [51] who found differences between individuals in their tendency to leave behind DNA when touching an object, and hence coined the term 'shedder', as previously discussed.

From the data results of the AmpFℓSTR® ProfilerPlus™ kit there was a slightly higher profiling success for females compared to males for fingermark deposits (see Figure 4.3). However, a larger sample size would be required to examine this finding in more detail and to investigate statistical significance. Investigating the difference in recovered DNA levels between fingers used to deposit DNA led to the observation that wide inter- and intra-individual differences were observed (see Figure 4.2). One apparent observation was that the thumb sheds more DNA than the pinky finger as suggested by the average profiling success generated from 10 donors (see Figure 4.2); this result is not surprising given the larger and rougher surface area of the thumb and the higher use of this digit compared to others. There was no apparent difference in DNA results obtained from swabbing fingermarks that had been created by different fingers (i.e. there was no impact on results).

In the summary data described (see Table 4.11), 96 out of 150 profiles (64%) were considered 'up-loadable' to the Australian DNA database. Individuals were asked to record their activity during the entire study prior to depositing fingermarks. However, due to the small sample set (i.e. ten volunteers) it is difficult to determine if profiling success is activity-dependent. From observation it appeared that those individuals that undertook keyboard work or handled personal items were more inclined to leave behind amplifiable DNA (see Tables 4.1, 4.7. 4.8 and 4.10). For the volunteers that state that they were typing on their keyboards during the study, only one individual (out of 10) produced a lower success rate (6 out of 15 profiles were 'up-loadable') (see Table 4.6) compared to the other donors that took part in the same activity and achieved greater profiling success. Overall, it can be assumed from these results and

previous research that a higher yield of amplifiable DNA from fingermarks can be attributed to the phenomenon known as ‘loading’ the fingers with DNA [55], as discussed previously.

Donor 9 generated 15 profiles (out of 15) that were all considered ‘up-loadable’ to the Australian DNA database (see Table 4.9). This individual was the only volunteer that took part in intense exercise directly before fingermarks were deposited, and this activity may contribute to the high success rate obtained; cell-free DNA originating from sweat is thought to be a vital component of the ‘touch’ DNA sample [44, 56], as discussed previously.

Time of DNA deposition

It has been assumed previously that a long time period is needed between hand washing and depositing DNA in order to generate a profile from a touched item [26]. For example, Lowe *et al.* (2002) [51] report that if a longer time has passed since hand washing (i.e. 2 – 6 hours) shedder type may not be relevant. Results in this chapter contradict some of these findings published by Lowe *et al.* (2002) [51]. This difference could be due to the way in which samples were processed in their study (i.e. extracted) compared to this study (i.e. direct PCR). Full DNA profiles in this chapter were obtained in as little as 15 minutes post hand washing (see Tables 4.1 and 4.9). Donor 1 (see Table 4.1) deposited enough fingermark DNA to yield interpretable profiles for all fingers tested and all time periods. Overall, for females there was a slight increase in profiling success when a longer time passed since handwashing, from 68% (at 15 minutes post hand washing) to 74% (at 2 hours post hand washing) (see Figure 4.3). On the contrary, there was a slight decrease in profiling success for males when a longer time passed since hand washing, from 70% success (at 15 minutes post hand washing) to 64% (at 2 hours post hand washing) (see Figure 4.3).

Two individuals (out of ten) appeared to yield more amplifiable ‘touch’ DNA when a longer time elapsed between hand washing (i.e. 2 hours) and depositing fingermarks (see Table 4.8 and Table 4.10). On the contrary, donor 3 generated four (out of five) ‘up-loadable’ DNA profiles 15 minutes after hand washing (see Table 4.3), with

success decreasing when a longer time had passed since hand washing (i.e. 1 and 2 hours later). Samples analysed immediately after hand washing (T=0) failed to generate any DNA alleles. It can hypothesized that DNA builds up on the surface of the skin at some point within the first 15 minutes after hand washing, and a contributing factor could be 'loading' the fingers with DNA from other parts of the body (e.g. hair and face), and/or touching personal items that comprise of donors DNA. It can be assumed from these results that the profiling ability of the individual's fingermark deposit is more dependent on the donor's activities prior to depositing DNA. However, it was also noted in preliminary testing that certain individuals were consistent at generating good quality DNA profiles recovered from fingermark traces; although the term 'shedder' has not been used to classify these individuals. Further testing could involve swabbing fingermarks created by individuals post hand washing with a larger number of time intervals assessed (e.g. every 5 minutes, between 0 min and 1 hour) to understand the loss of DNA associated with donor activities.

DNA transfer

There was one case of tertiary DNA transfer observed in the data set (i.e. person to object to person to object transfer) (see Figure 4.6) where the major peaks (i.e. 12 alleles) obtained from swabbing the plastic substrate matched the alleles of the donor's partner, and the minor peaks (i.e. 14 alleles) obtained matched the individual who left the mark (see Table 4.8). The major profile may appear misleading and suggest that the last person who touched the substrate was the donor's partner. Characterizing the factors that contribute to the transfer and persistence of 'touch' DNA is complex, and further research would be required before commenting on these issues. Results agree with those of Djuric *et al.* (2008) [66] in which secondary DNA transfer was observed. It is well-known that DNA could be deposited by "accidental" transfer or from a separate contamination event. Additional research conducted by Lowe *et al.* (2002) [51] have shown that secondary transfer – where the major alleles do not match the donor - is observed under controlled laboratory conditions. Many other research papers have been published in the area of "non-self" DNA and secondary/tertiary DNA transfer events [52, 61-63, 65, 67-70]. Considerable caution must be taken by DNA analysts to address the issues of DNA transfer and persistence. One high profile case that is a perfect example of

misinterpretation of the DNA evidence is the high profile case against Amanda Knox and Raffaele Sollecito [71].

High profile case example – DNA transfer in the case of Amanda Knox

One high profile case that received intense media scrutiny and a clear lack of knowledge in the area of DNA transfer was the murder of Meredith Kercher, and the implications of the miscarriage of justice of Amanda Knox and Raffaele Sollecito. Amanda Knox was an American exchange student accused of murdering her room-mate in Italy in 2007. Amanda Knox's DNA was not found in the murder room, and she was separately implicated in the murder by a knife found in a different location from the crime scene and traces of her DNA linked to the bathroom in the flat that she shared with the victim. In the first court hearing there was no indication of possible innocent DNA transfer. In addition, an extensive review published by Peter Gill in 2016 [71] to the forensic community highlights the misinterpretation of the low-level DNA evidence in the case and the use of improper contamination controls. For example, gloves used to collect evidence at the crime scene were stored with other items of evidence that would contribute to cross-contamination. It was heard during court proceedings that the distribution of DNA alleles could infer activity level proposition; the presence of DNA on the knife indicated that it had been used in an upward stabbing motion, rather than a cutting motion (e.g. to cut food). It is not possible to infer how, when or why DNA transfer occurs based on the STR profile obtained. This case is a perfect example of how important correct interpretation of low-level DNA evidence is and the importance of contamination controls. Other published data examining the transfer of DNA from examination gloves to case work exhibits emphasize the huge potential for cross-contamination and the risk of DNA transfer compromising investigations [65].

Artefacts of direct PCR

Stochastic effects, associated with the PCR amplification of low-template DNA, were occasionally observed in the data set in this chapter. Examples of profiles illustrate an incidence where the peak heights of alleles were imbalanced (see Figure 4.4), high stutter was observed (see Figure 4.5), 'pull-up' (possibly) (see Figure 4.5) and 'drop-out' was apparent (see Figure 4.4). A more in depth examination of stochastic effects observed in data generated after using direct PCR are detailed in the proceeding Chapter V.

Direct PCR from fingermarks immediately after hand washing

All samples that were deposited by individuals immediately after hand washing failed to yield informative DNA profiles by the direct PCR approach. It is assumed that DNA needs to accumulate on the surface of the skin in order to generate a DNA profile, and cellular material or cell-free DNA is potentially washed away immediately after hand washing. From these results it is estimated that 15 minutes is needed post hand washing in order for DNA to accumulate on the surface of the skin. However, one published DNA transfer study contradicts this hypothesis [72] stating that cellular material survived the hand washing process. Despite cells being transferred immediately after hand washing, no DNA was detected. Surprisingly, the authors observed a higher level of cellular transfer from individuals who had previously washed their hands compared to individuals who had not washed their hands. The authors postulate that the physical act of hand washing would loosen corneocytes [72] and these cells would then transfer more easily through contact. A possible variable would be how hands are dried (e.g. if hands are dried with the same towel or a fresh towel) and detergents used in the hand washing process. Consistent with Locard's principle, while 'touch' DNA samples may not always transfer DNA in sufficient amounts for nuclear detection to generate a probative STR-based DNA profile, there is the possibility to examine other markers (i.e. mtDNA) [45, 73] that could provide genetic information on the sample.

Extracting DNA from fingermarks using the DNA IQ™ System

DNA loss by extraction has been emphasized in previous research [5] and data results from Chapter III corroborate earlier findings that report a loss of extracted DNA. All extracted DNA swabs (i.e. DNA IQ™ System) utilised in this chapter failed to generate a DNA profile following the standard protocol, however, positive control DNA generated a full concordant DNA profile. It can be proposed that individuals who donated fingermark samples did not deposit enough DNA for detection at the time of sampling. However, this notion seems implausible given that the individuals chosen for the extraction study were found to consistently generate DNA profiles when direct PCR was employed. It is not unreasonable to assume that DNA has been washed away during the extraction steps and multiple tube changes and binding of low-level DNA to plastic ware may contribute to a further loss. Future work to optimise DNA

extraction protocols could involve concentrating the DNA extract before PCR to maximise the amount of DNA template available to PCR, however this work would require additional steps, more tube changes, and further validation to be carried out.

Loss of DNA by extraction

Previously it was believed that DNA transferred by 'touch' to an object was primarily based on the quantity of cells that people shed [51]. Recent studies have shown that these sample types also comprise of 'cell-free' DNA [44, 56, 74] and this valuable portion should also be taken into consideration when processing the samples. A recent study stresses the importance of cell-free DNA present in the supernatant of extracted material [72]. It was discovered that the majority of nuclear DNA (i.e. ~ 84 – 100%) [72] recovered from 'touch' DNA samples was extracellular. The authors stipulate that direct PCR methodology is essential to avoid loss of extracellular DNA. Furthermore, a recent survey of case work samples reported the presence of extracellular DNA in 70% of samples comprising of 'touch' DNA content [74]. These findings are compatible with medical research studies revealing a lack of genomic DNA in the upper epidermal skin layer (i.e. stratum corneum) that is comprised of corneocytes [75]. In other published work, the loss of 'touch' DNA is attributed to the extraction methodology employed and the use of filtration columns [31]. Addressing these issues has significant implications for optimising 'touch' DNA methodology and the future success of 'touch' DNA swabs.

Future work

Further studies examining DNA deposition could look at time points immediately after hand washing and within 15 minutes of hand washing; as data in this chapter indicate that DNA accumulates on the surface of the skin within this short time frame. The difference in shedder status between fingers could also be explored further. For example, one study examining the anatomical location of the skin surface observed a difference in the amount of DNA deposited between the palm and digits [76]; the quantity of DNA shed from the palmar surface was significantly less than from two fingers.

Direct PCR of 'touch' DNA – an in-depth study investigating the use of 2 common STR profiling kits: NGM SElect™ kit and AmpFℓSTR® ProfilerPlus™ kit.

Results for the third phase of this study were to expand on the preliminary results observed in tables 4.1 – 4.11. A detailed study was carried out on a larger sample set (n = 34) at only one chosen time period (i.e. 15 minutes post hand washing). Results are published in the manuscript: ***"Templeton, J.E. and Linacre, A., 2014. DNA profiles from fingermarks. BioTechniques, 57(5), p.259" [77] (see manuscript enclosed).***

4.5.3.1 Statement of authorship

Title of publication: DNA profiles from fingermarks

Manuscript published in the journal Biotechniques.

Jennifer Templeton (Candidate)

Performed all the laboratory work, data analysis, and wrote the paper

I hereby certify that the statement of contribution is accurate.

Signed: Date: November 2016

Adrian Linacre (Supervisor)

Supervised the project, commented on data, and helped to edit the paper

I hereby certify that the statement of contribution is accurate.

Signed: Date: November 2016

4.5.3.2 Manuscript: DNA profiles from fingermarks.

Reports

DNA profiles from fingermarks

Jennifer E.L. Templeton and Adrian Linacre
School of Biological Sciences, Flinders University, Adelaide, Australia

BioTechniques 57: 259-266 (November 2014) doi 10.2144/000114227
Keywords: direct PCR; DNA profiling; fingermarks; NGM SElect; Profiler Plus; short tandem repeats

Criminal investigations would be considerably improved if DNA profiles could be routinely generated from single fingermarks. Here we report a direct DNA profiling method that was able to generate interpretable profiles from 71% of 170 fingermarks. The data are based on fingermarks from all 5 digits of 34 individuals. DNA was obtained from the fingermarks using a swab moistened with Triton-X, and the fibers were added directly to one of two commercial DNA profiling kits. All profiles were obtained without increasing the number of amplification cycles; therefore, our method is ideally suited for adoption by the forensic science community. We indicate the use of the technique in a criminal case in which a DNA profile was generated from a fingermark on tape that was wrapped around a drug seizure. Our direct DNA profiling approach is rapid and able to generate profiles from touched items when current forensic practices have little chance of success.

Fingermarks are essential forensic evidence in numerous criminal investigations. Generating a DNA profile from a fingermark for the purpose of human identification would be beneficial in resolving a broad spectrum of criminal investigations, ranging from theft to crimes of violence. DNA retrieved from fingermarks deposited by touch (referred to as "touch" DNA) is often degraded (1) and limited in quantity (2) and may contain elements that co-extract with the DNA (3), which can hinder subsequent amplification. Although forensic genetics has seen substantial improvements in DNA profiling sensitivity (4–6), typically the use of less than 100 pg of DNA template (equating to ~16 human somatic cells) can result in poor-quality profiles (7,8). This limit of sensitivity still precludes many items that have been touched at a crime scene from generating a usable DNA profile, despite their potential importance in a criminal investigation; these samples can include triggers, steering wheels, bullet cartridges, and handles of knives. In many criminal investigations, the ability to retrieve the maximum amount of DNA from touch DNA samples is of paramount importance in resolving the case.

The first DNA profile generated from a fingerprint was reported more than a decade ago (9) and revolutionized forensic science. Despite this advance in the field, research has found an extremely low success rate (5%–6%) when using the standard methodology in generating touch DNA profiles (10), highlighting the need for an improved methodology. The standard workflow for touch DNA samples includes an extraction step before amplification. However, many current extraction processes are thought to result in a loss of DNA (11,12); hence a touch DNA swab initially containing less than 100 pg of DNA may result in insufficient template at the PCR step to generate an informative profile. Many touch DNA samples are therefore either not submitted for DNA typing or fail to generate any data without further enhancement. Previous attempts to generate DNA profiles from fingermarks used a low copy number (LCN) methodology (9,13,14). However, any enhancement of the amplification process from limited and low-level DNA has the inherent risk of introducing stochastic events such as allelic drop-out, allelic drop-in, or an increase in stutter heights or allelic imbalance for a heterozygote. Many factors affect the deposition of DNA by touch at crime scenes, with previous reports classifying individuals as "shedders" or "nonshedders" (15–19) depending on the profiling success from fingermarks 15 minutes after handwashing. The issue of whether or not a person is a shedder and how this may or may not affect transfer of DNA is a concern in forensic science (16,20–22); however, in reality, there is no knowledge of a person's shedder status during a criminal act.

Here we report a procedure that eliminates the need for LCN technology by omitting the extraction process, therefore minimizing the opportunity for sample mix-up or introduction of extraneous DNA into the reaction vessel. If DNA exists as a free molecule, then it should bind readily to positively charged fibers (23). Furthermore, the use of a surfactant such as Triton-X is thought to assist in the binding of DNA to a swab (24) and hence, these two features were used to capture DNA from touched substrates. By placing the fibers that made contact with the swabbed surface directly into the PCR tube, there is no potential loss of DNA during an extraction process. Direct PCR has been used successfully by our laboratory in a forensic context to generate informative profiles from single hairs (25–27), and a similar approach to direct PCR is used here to generate DNA profiles from individual fingermarks using two different commercially available kits, Profiler Plus, which is designed to amplify 9 short terminal repeat (STR) loci plus the sex determining marker amelogenin (making a total of 20 alleles), and NGM SElect,

METHOD SUMMARY

A novel DNA profiling method using a swab moistened with Triton-X to obtain sample DNA from fingermarks followed by direct polymerase chain reaction (PCR) amplification to generate a forensic profile is described.

which targets 16 STR loci plus amelogenin (making a total of 34 alleles).

Materials and methods

Contamination controls for low-template DNA

All laboratory steps preceding DNA amplification were carried out in a dedicated fume hood separate from post-PCR activity. Strict decontamination procedures were followed, including cleaning of equipment and work areas with 3% sodium hypochlorite and 70% isopropanol before and after use. Nontemplate controls (PCR blanks) were included in each experiment to monitor potential contamination from human DNA sources and cross-contamination from other samples. Before use, sterile and DNA-free plastic slides (Rinzl plastic; ProSci Tech, Kirwan, QLD, Australia) were cleaned with 70% isopropanol and Milli-Q ultrapure water (Merck Millipore, Victoria, Australia) and left to dry in a sterile fume hood before ultraviolet exposure for 15 min. As an additional negative control, slides were swabbed before individuals deposited fingermarks/DNA. The STR DNA profiles of both staff members (J.E.L.T. and A.L.) involved in the handling of the samples were genotyped. All buccal reference swabs provided by donors were processed after fingermark deposition and analysis.

DNA reference swab collection

Cotton buccal swabs (Copan Industries, Victoria, Australia) taken from the inner lining of the cheeks of 34 volunteers were collected. Ethical approval was obtained from Southern Adelaide Clinical Human Research Ethics Committee before starting. DNA was extracted following the DNA IQ Kit (Promega Corporation, Victoria, Australia) manufacturer's recommendations.

DNA deposition 15 minutes after handwashing

Thirty-four volunteers washed their hands with warm water (no soap or detergent was used) before depositing DNA. Fifteen minutes after handwashing, each volunteer deposited a fingermark using four fingers and thumb from his or her dominant hand onto sterile and DNA-free plastic slides. Contact was for 15 s using medium pressure (to ensure consistency) such that a fingermark was created. All slides were stored at 4°C before processing. All marks made by touch in this experiment were clearly visible, and these fingermarks were removed using a double-swab technique.

Targeted swabbing

A small portion of fibers (~2 mm^2) was cut from the tip of a sterile nylon FLOQswab (Copan Industries) using a new sterile scalpel blade. Fibers were pre-moistened with 2 µL of 0.1% Triton X (Sigma, Victoria, Australia) (preheated to 50°C) before swabbing. Sterile forceps were used to gather the 2-mm^2 portion of swab and apply pressure over the mark. Glue fragments of the swab held the fibers neatly in place. Swabbing consisted of 2 sets of 10 strokes horizontally (left to right direction) and 2 sets of 8 strokes vertically (top to bottom) over the mark using heavy pressure to ensure consistency. The procedure was repeated using a second dry swab. Sterile forceps were used to place swab fibers (4 mm^2) from both swabbing events into the same 0.2-mL PCR tube.

DNA amplification using Profiler Plus

STR typing was performed using the AmpFLSTR Profiler Plus kit (Life Technologies, Victoria, Australia) following the manufacturer's recommendations but at half the volume. The final reaction consisted of 10 µL Profiler Plus reaction mix, 5 µL Profiler Plus primer mix, 1 µL (5 U) AmpliTaq Gold (Life Technologies), and DNA template (4 mm^2 swabbed fibers for samples or 1 ng of eluted DNA extract from a buccal swab). Additional PCR facilitators were added to the PCR due to previous studies suggesting an increase in performance. Purified molecular biology grade BSA (0.1 µg) (New England Biolabs, New South Wales, Australia) and DMSO (dimethyl sulfoxide) (5%) (Expand Long Range dNTPACK; Roche, Victoria, Australia) were added to the PCR tube, and cycling was performed on a 9700 GeneAmp thermal cycler (Life Technologies). Cycling consisted of an initial denaturation at 95°C for 11 min followed by 29 cycles of 94°C for 1 min, 59°C for 1 min, 72°C for 1 min, and then a final extension at 60°C for 45 min followed by a hold at 4°C. PCR products were analyzed on a 3130*xl* Genetic Analyzer (Life Technologies) in a 14-µL final volume that consisted of 1.5 µL PCR product (as recommended by the Profiler Plus kit manufacturer), 12 µL HiDi formamide (Life Technologies), and 0.5 µL ROX-500 Size Standard (Life Technologies).

DNA amplification using NGM SElect

STR typing was performed using the AmpFLSTR NGM SElect PCR Amplification Kit (Life Technologies) following the manufacturer's recommendations but with additional AmpliTaq Gold. The final reaction consisted of 10 µL NGM SElect reaction mix, 5 µL NGM primer mix, 1 µL (5U) AmpliTaq Gold, and DNA template (4 mm^2 swabbed fibers for samples or 1 ng of eluted DNA extract from a buccal swab). Cycling was performed on a 9700 GeneAmp thermal cycler and consisted of an initial denaturation at 95°C for 11 min followed by 30 cycles of 94°C for 20 s, 59°C for 3 min, and then a final extension at 60°C for 10 min. PCR products were analyzed on a 3130*xl* Genetic Analyzer in an 11.5-µL final volume that consisted of 1.5 µL PCR product, 9.5 µL HiDi Formamide, and 0.5 µL of GeneScan 600 LIZ Size Standard v2.0 (Life Technologies).

DNA deposition immediately after handwashing

Two individuals who consistently generated informative DNA profiles from fingermarks deposited 15 min after handwashing (using direct PCR) were asked to provide additional prints immediately after handwashing. Volunteers deposited a fingermark using 4 fingers and thumb of their dominant hand (as detailed above) to create 10 fingermarks. Sterile nylon FLOQswabs were used to swab the 10 fingermarks (in the same manner as detailed above), and the DNA was amplified using the Profiler Plus kit following the manufacturer's recommendations (as detailed above).

DNA extraction from a fingermark 15 minutes after handwashing

Two individuals who consistently generated informative DNA profiles from fingermarks

deposited 15 min after handwashing (using direct PCR) were asked to provide additional prints for an extraction experiment. Volunteers deposited a fingermark using 4 fingers and thumb of their dominant hand (as detailed earlier) to create 10 fingermarks. Sterile nylon FLOQswabs were used to swab the 10 fingermarks (in the same manner as detailed earlier, except that the whole swab was used for an extraction). Extraction followed the DNA IQ Kit manufacturer's (Promega Corporation) recommendations. PCR amplification followed the Profiler Plus recommendations (as detailed earlier).

Data analysis

All DNA profiles were analyzed using GeneMapper ID software (v3.2.1) (Life Technologies), and a peak detection threshold of 30 relative fluorescence units (RFU) (3 standard deviations above the baseline) was used to assign alleles. A wildcard designation was used for potential homozygotes with peak heights less than 150 RFU to account for potential allelic drop-out (e.g., "11, F" instead of "11,11"). A profile was considered to be full when all alleles were detected above the threshold RFU and matched the reference profile of the donor. Allelic drop-out was reported as alleles with peak heights less than 30 RFU. Additional alleles were reported if peak heights were greater than 30 RFU (3 standard deviations above the baseline) and did not match the donor providing the sample. An informative DNA profile (one considered uploadable to the Australian DNA databases) was defined as having at least 12 alleles (plus Amelogenin) that matched the reference DNA profile. The profiling success (%) was measured by dividing the number of alleles successfully called by the total number of expected alleles and multiplying this value by 100.

Results and discussion

DNA profiles were generated from swab fibers collected from 170 fingermarks created by 34 individuals. Each fingermark was created after a period of only 15 minutes after handwashing. From the total sample size (170 fingermarks donated by each of the 5 digits of 34 donors), 122 DNA profiles (71%) were recorded with 12 or more alleles (sufficient for uploading to the Australian DNA Database), with only 4 samples (<2%) failing to yield any DNA (see Tables 1 and 2). The success rate was 66% to 74%, depending on the STR DNA profiling kit used, and was obtained without the need for increased PCR cycle number.

Five of the 10 thumbprints produced full DNA profiles using Profiler Plus with a further 3 producing 17 or more alleles such that only a single thumb print produced a very partial profile with 7 of the possible 20 alleles obtained. Similar results were obtained for the index and middle fingers (Table 1).

Full DNA profiles were observed from 38% (19 of 50) of the fingermarks using Profiler Plus. Of these 50 fingermarks using all 10 donors and all 5 digits, 66% (33 of 50) generated 12 or more alleles with a further 28% (14/50) of the profiles generating at least 1 allele. A total of 696 alleles from a possible 1000 were observed.

Twelve of the 24 thumbprints analyzed using the 17 loci NGM SElect kit produced 25 or more alleles (characterized as full DNA profiles when using other STR typing kits) (Table 2). Only 4 of the 24 donors gave poor results, with fewer than 12 alleles for each of the 5 digits, and only 1 of the 120 fingermarks failed to yield DNA.

Of a possible 4080 alleles from all 5 digits and all 24 donors, 2266 matching alleles were produced. By contrast, all 10 swabs from which DNA was extracted by standard forensic processes failed to result in any DNA profile, as expected.

Table 1. DNA profiles obtained from 10 donors showing success rates using the Profiler Plus STR kit (of a possible 20 total STR alleles, 10 total loci).

	Finger used for depositing DNA					
	Thumb total alleles (out of 20)	Index total alleles (out of 20)	Middle total alleles (out of 20)	Ring total alleles (out of 20)	Pinky total alleles (out of 20)	Informative profiles (out of 5 digits)‡
Donor 1 ♂	20	20	20	20	20	5 (100%)
Donor 2 ♂	20	3	0†	20	20	3 (60%)
Donor 3 ♂	20	12	19	18	11	4 (80%)
Donor 4 ♂	12	20	20	0†	0†	3 (60%)
Donor 5 ♂	17	15	8	1	6	2 (40%)
Donor 6 ♀	20	10	17	6	20	3 (60%)
Donor 7 ♀	19	20	20	13	4	4 (80%)
Donor 8 ♀	7	9	8	2	10	0 (0%)
Donor 9 ♀	20	20	20	19	20	5 (100%)
Donor 10 ♀	17	5	17	16	15	4 (80%)

The first column designates the donor and the sex of the individual. Columns 2 to 6 report the number of alleles recorded from each donor. A maximum of 20 alleles indicates a full DNA profile. Numbers less than 20 indicate a partial or failed DNA profile. Three DNA profiles resulted in no alleles (indicated by †).
‡Informative profiles indicate the total number of profiles obtained from each individual that can be considered uploadable to the Australian DNA database (≥12 alleles) and are displayed in parentheses as the percentage profiling success for that individual.

Negative controls were performed using the same swabs and slides as well as the same solutions to detect any alleles from these sources, and no alleles were detected from all 34 negative controls. Any alleles detected in the fingermark DNA profiles, either matching the donor or otherwise, were most likely attributable to transfer to the slide at the time of deposition of the fingermark.

Examples of the DNA profiles obtained by direct PCR using either the Profiler Plus kit (Figure 1) or NGM SElect kit (Figure 2) illustrate that many profiles exhibited no increase in stutter products, a stochastic effect that may be observed in low-template DNA typing. Allelic drop-out did occur in samples that did not generate full DNA profiles, and additional alleles were recorded as unknown alleles that did not match the donor. Single-source DNA profiles were generated from 40% of the samples amplified using the Profiler Plus kit and 36% of the samples amplified using the NGM SElect kit. The surface touched by the individuals had been cleaned before use, and negative controls from areas on which the fingermark was placed failed to yield DNA profiles. The substrate on which the fingermark was deposited was open to the environment for approximately 2 minutes during sample collection, allowing potential deposition of DNA from other sources. Additionally, donors were active during the 15 minutes after handwashing such that DNA from other persons may have been deposited onto the surface of the finger by secondary or tertiary transfer, such as holding a shared item. This low-level secondary transfer has been observed under laboratory conditions (17–19, 21) and would be expected in real life when examining a substrate at a crime scene.

It remains unknown if the DNA from a fingermark is found within cellular material or as a free molecule. Regardless of the type of DNA present on the skin and available for transfer, it has been reported that the amount of DNA on the skin increases with time since handwashing (15). The 15 minute time period between handwashing and fingermark deposition was chosen deliberately as the minimum time point to provide consistency between this study and the original study into transfer of DNA to touched items (15). To confirm this, we asked two volunteers (who had consistently deposited sufficient DNA for informative profiles to be generated) to create fingermarks from each finger and thumb of the dominant hand in the usual way but immediately after handwashing. No alleles were detected in any of the resulting DNA profiles, indicating that rinsing in water had removed any DNA on the skin of the individuals tested. Our data indicate that sufficient DNA had accumulated on the surface of the skin to generate DNA profiles within 15 minutes after all detectable DNA had been removed. Based on the data from 15 minutes after handwashing, there is an approximately 66% to 74% chance of gaining a DNA profile with 12 or more alleles, with many full DNA profiles being observed.

Using our optimized technique, we were requested by the South Australia Police Service to generate a DNA profile from a smeared fingermark on tape that was wrapped around a drug seizure. The resulting NGM SElect profile contained 31 dominant alleles (of a possible 34). A further 10 low-level alleles were also recorded, indicating a low-level mixture of persons

Table 2. DNA profiles obtained from 14 donors showing the success rate using the NGM SElect STR kit (of a possible 34 total STR alleles, 17 total loci).

	Finger used for depositing DNA					
	Thumb total alleles (out of 34)	Index total alleles (out of 34)	Middle total alleles (out of 34)	Ring total alleles (out of 34)	Pinky total alleles (out of 34)	Informative profiles‡ (out of 5 digits)
Donor 11 ♂	33	34	34	32	32	5 (100%)
Donor 12 ♂	26	26	19	22	16	5 (100%)
Donor 13 ♂	18	4	8	14	24	3 (60%)
Donor 14 ♂	33	34	25	31	4	4 (80%)
Donor 15 ♂	17	16	9	1	16	3 (60%)
Donor 16 ♂	23	16	11	13	24	4 (80%)
Donor 17 ♂	3	3	2	1	1	0 (0%)
Donor 18 ♂	12	7	15	15	8	3 (60%)
Donor 19 ♂	25	28	28	24	26	5 (100%)
Donor 20 ♀	23	22	19	20	26	5 (100%)
Donor 21 ♀	34	34	31	34	32	5 (100%)
Donor 22 ♀	5	2	3	2	1	0 (0%)
Donor 23 ♀	20	23	21	22	19	5 (100%)
Donor 24 ♀	24	26	25	25	27	5 (100%)
Donor 25 ♀	14	18	19	4	5	3 (60%)
Donor 26 ♀	7	10	10	8	12	1 (20%)
Donor 27 ♀	32	32	17	23	20	5 (100%)
Donor 28 ♀	25	27	23	8	16	4 (80%)
Donor 29 ♀	27	21	22	15	17	5 (100%)
Donor 30 ♀	11	7	4	0†	7	0 (0%)
Donor 31 ♀	33	33	5	27	33	4 (80%)
Donor 32 ♀	26	28	29	26	30	5 (100%)
Donor 33 ♀	28	17	19	12	12	5 (100%)
Donor 34 ♀	28	16	17	27	26	5 (100%)

The first column designates the donor and the sex of the individual. Columns 2 to 6 report the number of alleles recorded from each donor. A maximum of 34 alleles indicates a full DNA profile. Numbers less than 34 indicate a partial or failed DNA profile. Only one DNA profile resulted in no alleles (indicated by †).
‡Informative profiles indicate the total number of profiles obtained from each individual that can be considered uploadable to the Australian DNA database (≥12 alleles) and are displayed in parentheses as the percentage profiling success for that individual.

depositing DNA onto the drug packaging (data not shown).

In this study, we demonstrate the ability to generate informative DNA profiles from latent fingermarks deposited by touch. Our method is rapid (23) because there is no need for a DNA extraction step, and there is a reduction in associated costs because there is no need to purchase DNA extraction kits. Most importantly for forensic science laboratories, by eliminating the need to increase PCR cycle number or concentrate the amplified products, the procedure is easily adapted into working practices, allowing laboratories accredited to ISO17025 to use the method. Still, future work is required to assess the limitations of direct PCR and to determine what effect chemical enhancement methods will have on the DNA template and resulting DNA profile.

Author contributions

J.E.L.T. and A.L. developed the experimental design and coauthored the paper. J.E.L.T. performed the laboratory work.

Acknowledgments

Funding was provided by the office of the Attorney General South Australia and a John Stocker Postgraduate Scholarship from the Science and Industry Endowment Fund. We would like to thank all volunteers that participated in the study and are grateful to Duncan Taylor and Oliva Handt for providing feedback on an earlier version of the manuscript. We are also grateful to Dr Bastien Llamas for help and valuable input.

Competing interests

The authors declare no competing interests.

References

1. **Kita, T., H. Yamaguchi, M. Yokoyama, and T. Tanaka.** 2008. Morphological study of fragmented DNA on touched objects. Forensic Sci. Int. Genet. *3*:32-36.

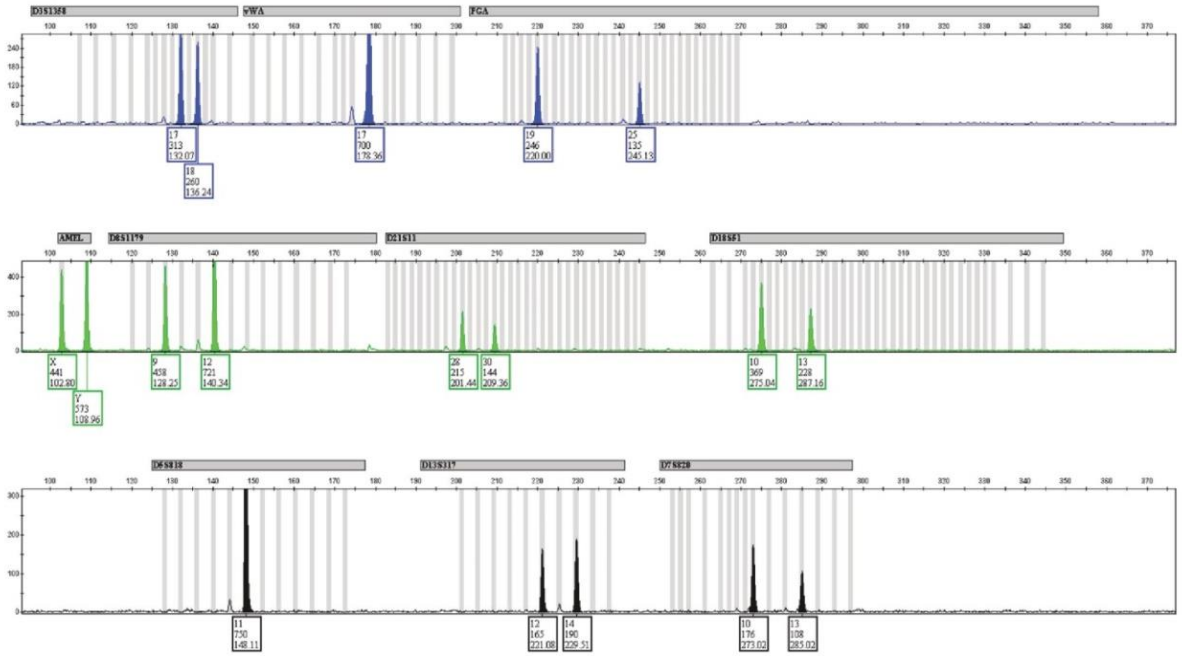


Figure 1. STR DNA profile generated using the Profiler Plus kit. The full DNA profile displayed was generated from a fingermark deposited 15 minutes after handwashing using direct PCR at 29 cycles. STR: short terminal repeat.

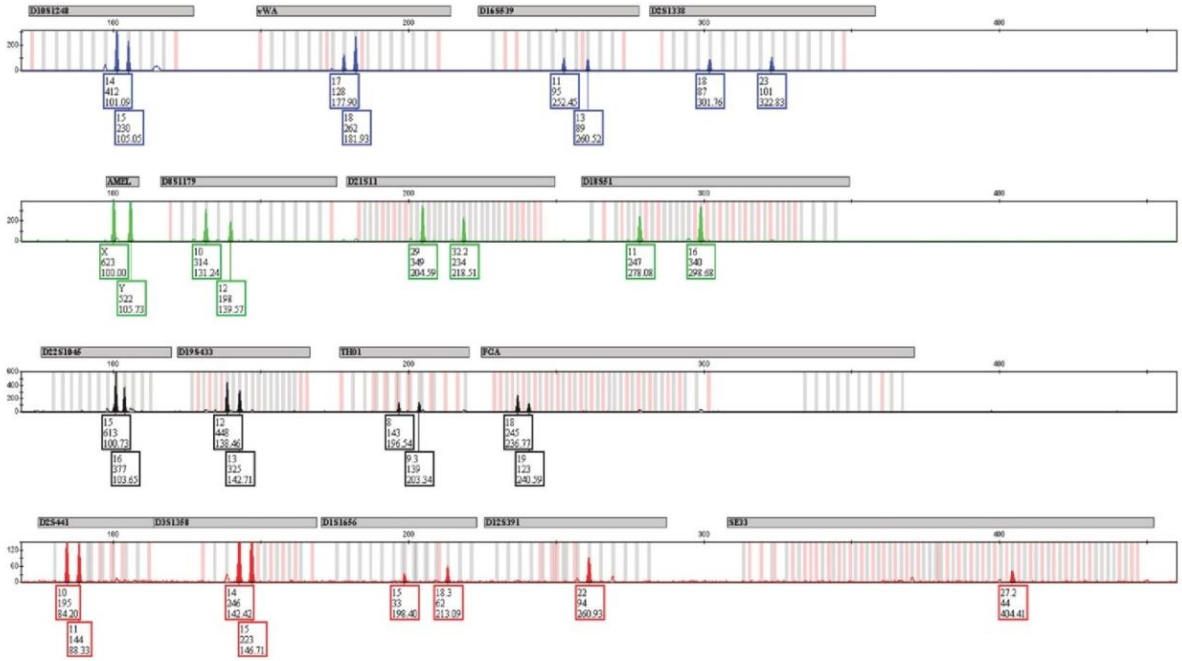


Figure 2. STR DNA profile generated using the NGM SElect kit. The DNA profile displayed was generated from a fingermark deposited 15 minutes after handwashing using direct PCR at 30 cycles. STR: short terminal repeat.

4.5.3.3 Journal article

BioTechniques
The International Journal of Life Science Methods

Generating DNA Profiles from "Touch DNA" in Fingermarks

11/13/2014
Amy R. Volpert

A new method increases the success rate of DNA profile generation from fingermarks/fingerprints. How will this improve forensic analysis? Find out...

ADVERT FROM JOURNAL REMOVED

Fingermarks (or fingerprints) and DNA are two pillars of modern forensic analysis, each extremely useful for identifying persons of interest in criminal investigations. In addition to the characteristic ridges and whorls used in fingerprint analysis, fingermarks often contain "touch DNA" shed by the individual who made it. DNA has a higher specificity for identification and can also be useful in cases where only a partial or smudge fingermark is present.

Unfortunately, the amount of DNA that can be isolated from a fingermark is usually very small, resulting in poor success rates (5% – 6%) when attempting to generate DNA profiles using current methods.

STR DNA profile generated using the NGM SElect kit.

Now, in an article in the November issue of *BioTechniques*, Jennifer Templeton and Adrian Linacre at Flinders University in Australia describe a method with a much higher rate of success for profiling DNA from fingermarks using commercially available short tandem repeat (STR) typing kits. "We have shown a dramatic increase in success of touched items and that for the first time, a DNA profile can be obtained from a fingermark without the need for enhanced amplification," Linacre said.

To collect negatively charged DNA from fingermarks, the authors used swabs made of positively charged nylon fibers that had been moistened with Triton-X, a detergent capable of dissolving membrane lipids, proteins, and other cellular components that could sequester DNA from PCR reagents. After swabbing the fingermark, the swab fibers were placed directly into the PCR tubes, bypassing the usual DNA extraction step with its potential for sample loss, and reducing the time and cost of the analysis.

ADVERT FROM JOURNAL REMOVED

According to Linacre, "The initial 95°C heating step in PCR and the addition of Triton-X has the potential to degrade the nuclear envelope of corneocytes and keratinocytes and to release the DNA into solution. We believe there is every opportunity for both cell-free DNA and intact cells to be made available to the amplification reaction."

With their approach, Templeton and Linacre achieved a 66%–74% success rate using DNA isolated from fingermarks for detecting the minimum number of alleles required to meet the standard for upload to the Australian DNA Database, a significant improvement over standard protocols.

The authors further tested whether the interval between hand washing and fingermark generation affected the deposition of DNA onto fingermarks and the likelihood of generating DNA profiles. For samples collected immediately after hand washing (with warm water and no soap), no alleles were detected; however, after 15 minutes, DNA profiles were generated from most fingermarks. Similar results were obtained for fingermarks deposited 1 or 2 hours after hand washing, although there was a greater incidence of allele drop-in (i.e., contamination).

However, the amount of DNA isolated from a fingermark is not strictly dependent on when the individual last washed his or her hands, according to Linacre, who said that "individuals do not necessarily deposit more DNA when a longer time has passed since hand washing; it appears to have more to do with the activities that the individuals are taking part in during that time period that will ultimately affect the profiling results." In addition, individuals can be classified as either "shedders" or "nonshedders" based on whether they deposit sufficient DNA with their fingermarks for successful analysis.

"The method has every opportunity of being used by forensic laboratories due to the application of standard PCR conditions," said Linacre, noting that the direct PCR approach eliminates the need for additional PCR cycles and that the DNA profiles they generated in this study showed no low-template DNA typing issues. "We are working towards simplifying the process to make the technical side easier and more adaptable to testing in the field," he added.

But until then, if you're planning any criminal mischief, wash your hands frequently and hope that you're a nonshedder.

Reference

1. J.E.L. Templeton and A. Linacre. 2014. DNA profiles from fingermarks. BioTechniques, 57:259–266.

Keywords: DNA isolation PCR forensic analysis

Journal article 2

BioTechniques
The International Journal of Life Science Methods

Fingerprinting the Fingerprints: The Past, Present, and Future of Molecular Forensics

05/06/2015

Nicholas Miliaras, PhD

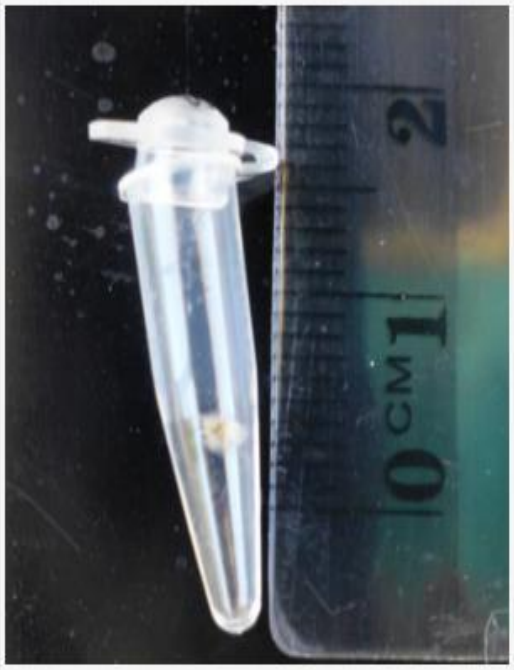
Recent advances in PCR and sequencing technologies have significantly advanced forensic science, but new questions have been raised in the process.

Since its first use as a forensic tool in the 1980s, DNA analysis has become the gold standard of evidence in criminal justice. Early forensic scientists used Restriction Fragment Length Polymorphism (RFLP) analysis to compare profiles of likely suspects. More recently, improvements in nucleic acid purification and amplification techniques made it much easier to obtain reliable DNA profiles from minimal sample material And next generation sequencing (NGS) technologies are identifying new targets that will yield markers more specific for individual suspects, as well as allow the collection and sharing of these data through centralized databases.

Fingerprinting the Fingerprints

Given its simplicity, PCR is ideally suited for forensics since it provides exquisite sensitivity of detection and enables a genotype to be generated quickly, allowing investigators to rule suspects in or out within the few hours it takes to perform and analyze the reaction. However, sample preparation protocols and contamination by others who may have handled the sample or were present at the crime scene can make it difficult to interpret the results. The low amounts of DNA that investigators must work with are also problematic.

Jennifer Templeton and Adrian Linacre at Flinders University in Adelaide, Australia recently addressed these issues in a *BioTechniques* paper, "DNA Profiles from fingermarks" (1). In this study, they used PCR to profile DNA collected directly from swabs of fingerprints by analyzing short tandem repeats (STRs). "There is latent DNA on things that you can't see such as door knobs, knife handles, etc. Currently, you need to do a DNA extraction where you lose 80 percent, even with the best methods, so you end up with insufficient template to do anything. The only way to get it to work is if you boost the number of cycles, which a lot of labs don't like to do," said Linacre.



Swab fibers in a PCR tube.

Others have been able to amplify fingerprint DNA using PCR with increased numbers of cycles for low copy number detection, but this can easily introduce errors that include the misdetection of targets (allelic drop out or drop in) and over-amplification of a target in a heterozygote. Templeton and Linacre decided to do direct PCR on DNA from swabs of fingerprints collected from all 5 fingers of 34 volunteers and were able to generate interpretable STR profiles for 71 percent of the 170 swabs using the ProfilerPlus and NGM Select STR typing kits.

Linacre came up with the idea to try direct PCR on swabs from earlier work where he and his colleagues generated STR profiles from human hairs. "If you take 2 mm of your hair from the proximal end, every hair will generate a full profile. This made me think that we could also get it from fingermarks," he said.

Linacre's group used a positively charged swab to collect negatively-charged DNA. The approach was so effective that the researchers did not need to extract DNA. They simply placed the fibers from the swab into a PCR tube and performed the recommended number of PCR cycles to generate the profiles.

"We wanted to do it as forensic labs currently do with commercial kits so that it will be easy to implement," Linacre said. "This also reduces the chances of swapping things up and greatly reduces cost."

"When we started a few years ago, I thought that this was already done; I am not going to take any great praise for this because it is patently obvious," he added.

Jennifer Templeton.

Less Is More

Although Templeton and Linacre's technique is not necessarily novel, other molecular forensic experts also see advantages in their approach. For example, Bruce Budowle at the University of North Texas Health Science Center mentioned that, "earlier studies didn't have good amounts of DNA. Usually you perform fewer cycles on something, with direct PCR and a blood spot, for example, because there is more blood DNA. [Linacre] is using the same principle. With low-level material such as fingerprints, once you put it into solution in a very small volume, it ends up much more concentrated." Concentrated samples reduce errors.

Ron Fourney of the Royal Canadian Mounted Police described Templeton and Linacre's work as "teaching an old dog new tricks. But they weren't afraid to take risks and try something different. They treated the swabbed fingerprints with Triton-X 100, which has an affinity for DNA. But the real secret to their success was that the technology has evolved so rapidly. Ten or even five years ago, they probably wouldn't have even had a result. Their approach was elegant in its simplicity because they didn't need to clean it up, so there is a reduced risk in loosing valuable DNA or introducing extraneous DNA during the analysis itself." He added, "I suspect direct PCR will be the way of the future for many routine forensic DNA processes."

The Wave of the Future?

PCR is essential for NGS since most platforms depend on an amplification step. As NGS becomes more widespread, it will certainly become a central part of forensic laboratories, and direct PCR would be the only step between a crime scene sample and a suspect's DNA sequence.

According to Budowle, clinical sequencing is unlikely to surpass PCR as a forensic tool. "NGS has advantages but the turnaround time is slower." However, NGS is valuable for identifying additional markers, which will improve the potential of PCR for individualization.

Fourney sees NGS and PCR as displacement technologies. "For the first time, rapid genome analysis is converging with automated easy-to-use technology at a relatively inexpensive cost," he said. He expects that soon it will be easier, faster, and less expensive to sequence the entire genome and reverse engineer specific forensic markers than it is to sequence the mitochondrial DNA markers currently used or to carry out conventional STR analysis.

4.6 Application of case work: seizure of Methamphetamine in South Australia.

In March 2013 South Australia Police uncovered several containers buried under ground in a secluded suburb of Adelaide, the case was later known as ‘Operation Divulge’. The operation was brought to light when police officers noticed two men acting suspiciously at a roadside in South Australia when they were digging up containers pictured below (see Figure 4.7). Over 7 kg of methamphetamine was discovered in these pots and valued at more than AU $20 million.

FSSA received the containers and masking tape (pictured below) that was used to seal the tops of the containers, and were asked to obtain DNA profiles and/or fingerprints from the evidence. Standard protocols of extracting swabs used to collect DNA were unsuccessful. The samples were thought to be buried under ground for several months and the DNA was potentially degraded and subjected to environmental PCR inhibitors such as humic substances in soil, bacteria, humidity and varying temperatures. Standard fingerprint powders were used by South Australia Police to dust for prints. The samples that were submitted for DNA profiling were saturated in black fingerprint powders as a result of dusting, and resulted in the majority of swabs pre-tarnished with black fingerprint powder. Samples taken by FSSA for DNA profiling failed to generate any DNA profiles using standard technology.

Figure 4.7 Containers discovered in March 2013, during operation divulge, that had been buried under ground and found to contain 7.33 kg of pure methamphetamine.

The Forensic Biology research group at Flinders University were contacted to assist with the case, due to the recent success that Flinders University achieved with

obtaining DNA profiles from touched items using the direct PCR approach. Jennifer Templeton, Renée Blackie and Professor Adrian Linacre were called to the DNA laboratory, Flinders University to examine over 100 exhibits relating to operation divulge and 24 items were selected for sampling and direct PCR. Drug containers and masking tape were collected (see Figures: 4.8 – 4.10) – some of which held fibres and single hairs.

Hairs and fibres were processed by direct PCR (using the amplification conditions listed below), but failed to yield any DNA results. The optimum swabbing technique – fundamental to this thesis (see Chapter II) - was used to process the containers and masking tape (see Figures: 4.8-4.10) thought to contain trace DNA. DNA-free nylon FLOQswabs™ pre-moistened with 2 µL Triton™-X-100 (Sigma-Aldrich) at 0.1% concentration were used. Post-swabbing, the pre-cut ~ 2 mm^2 fibre clump was added directly to a 0.2 mL PCR tube using sterile forceps. Double swabbing was performed using a second moistened DNA-free nylon FLOQswab™. Post swabbing, the pre-cut ~ 2 mm^2 fibre clump was added directly to the same 0.2 mL PCR tube using sterile forceps. The PCR tube consisted of 2 x 2 mm^2 fibre clumps. Amplification was carried out using the NGM SElect kit™ (Life Technologies, Victoria, AU) at standard 25 µL PCR volume, with additional Ampli*Taq* Gold® DNA polymerase (1 µL, containing 5 units) added to help overcome PCR inhibitors. Cycling conditions consisted of 29 PCR cycles using a GeneAmp® 9600 thermal cycler (Life Technologies). PCR products were analysed on the ABI 3130*xl* Genetic Analyser (Life Technologies) and GeneMapper® v3.2. software.

All 24 samples analysed resulted in DNA alleles and 17 of these samples produced 10 or more alleles. Sampling success ranged from 4 – 42 alleles per sample. Mixtures (i.e. DNA originating from more than one individual) were observed in 58% of samples analysed. Data files were sent to FSSA for interpretation and further analysis. Despite the success of obtaining DNA profiles, heterozygote imbalance was observed (see Figure 4.11) and this could be due to the non-removal of PCR inhibitors when the direct PCR approach was used, or the presence of a mixed DNA profile causing imbalance. Another negative aspect of direct PCR is that samples are completely consumed so there is only one chance of achieving a result from precious

sample material and this limitation may prevent many laboratories from exploring the potential of direct PCR.

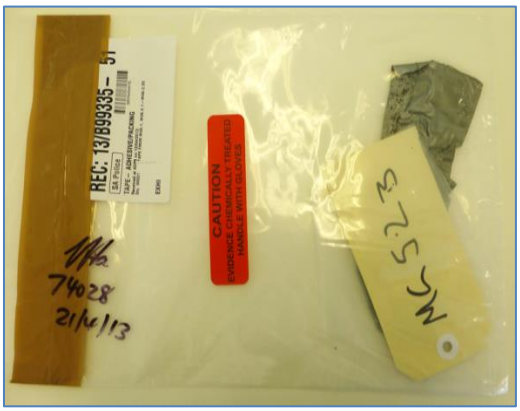


Figure 4.8 Case exhibit MG523.B, grey duct tape found wrapped around a drug seizure.

Figure 4.9 Case exhibit MG523.B, grey duct tape highlighting the area of cutting where the fingerprint was identified by eye. The tape was later subjected to direct PCR.

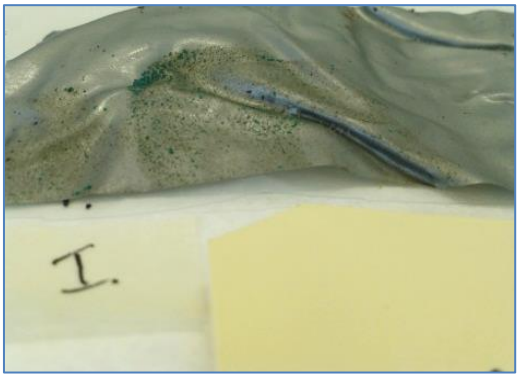


Figure 4.10 Case exhibit MG523.B, grey duct tape at higher magnification to illustrate potential smudged fingermark, ridges of fingermark were observed and drug debris was present.

Application of case work: seizure of Methamphetamine in South Australia

DNA Profiles obtained from case work

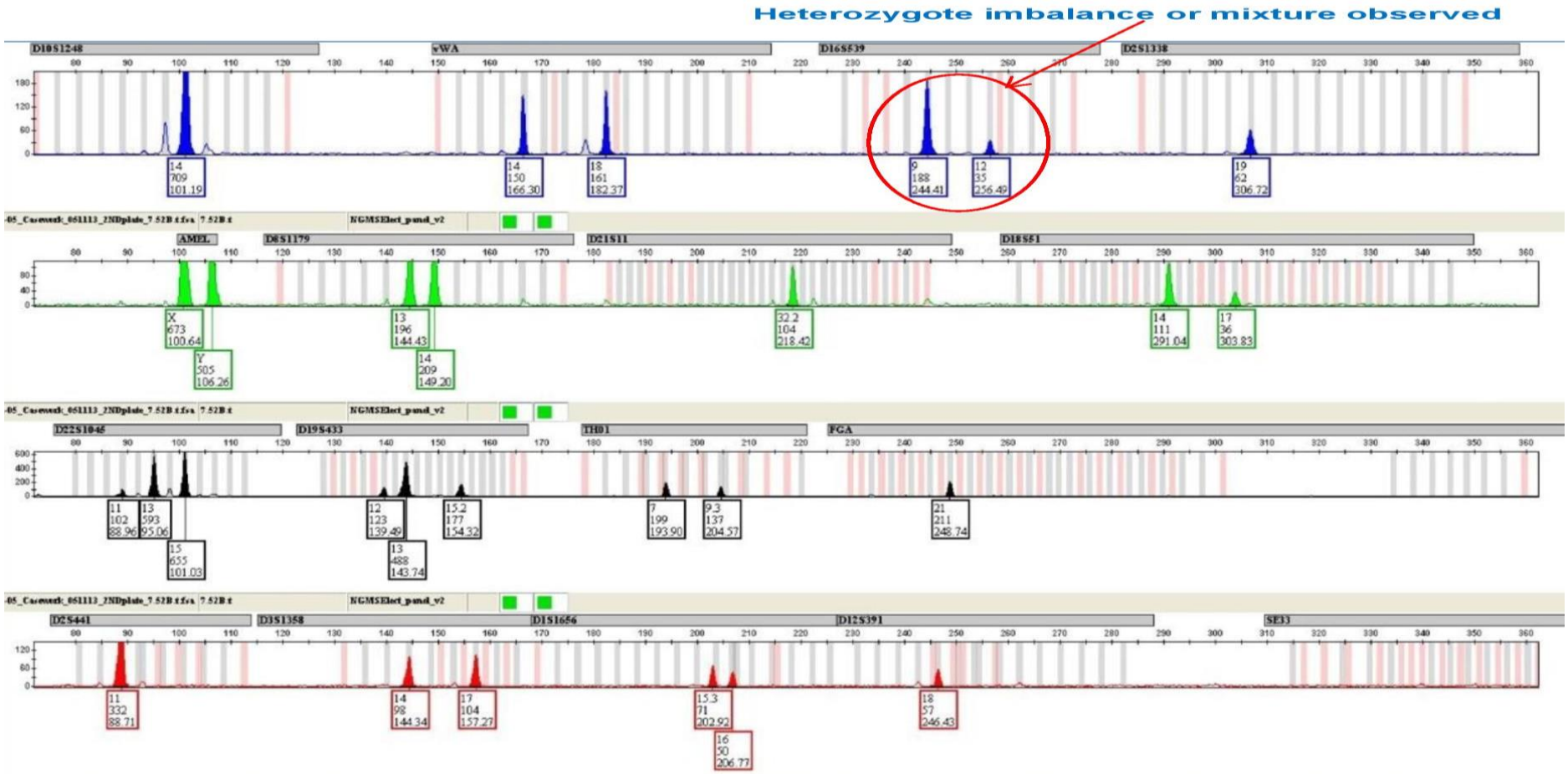


Figure 4.11 NGM SElect™ kit DNA profile obtained from tape, labelled PCR tube t52B. Case exhibit: MG523.B. Masking tape cutting seen in Fig 4.8, 4.9 and 4.10. was used for direct PCR at 29 cycles. A mixed DNA profile was obtained, with a major profile observed.

Application of case work: seizure of Methamphetamine in South Australia

DNA Profiles obtained from case work

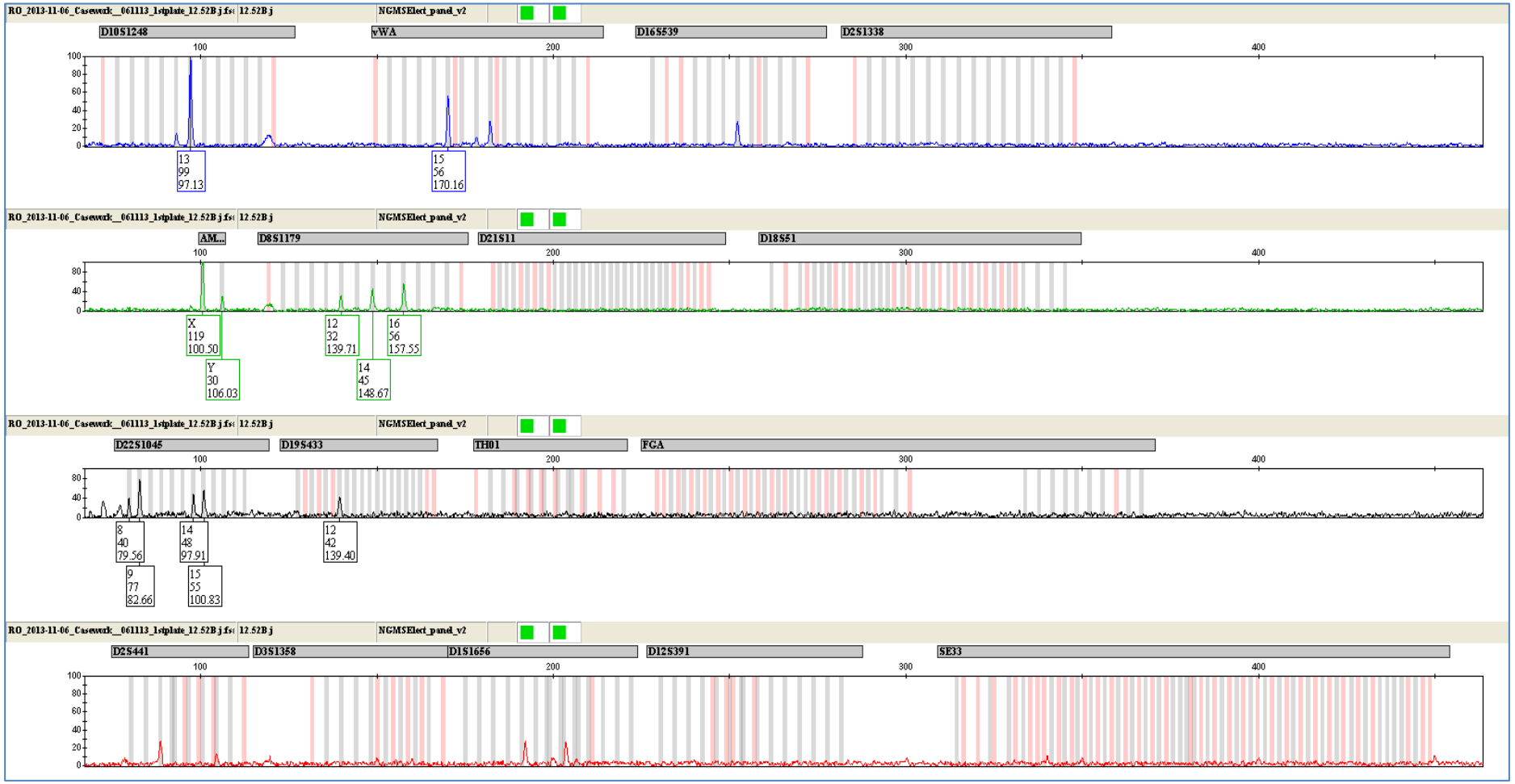


Figure 4.12 NGM SElect™ kit DNA profile obtained from tape, PCR tube label j52B, and case exhibit: MG523.B. Direct PCR at 29 cycles was used to obtain the profile.

Application of case work: seizure of Methamphetamine in South Australia

DNA Profile obtained from case work

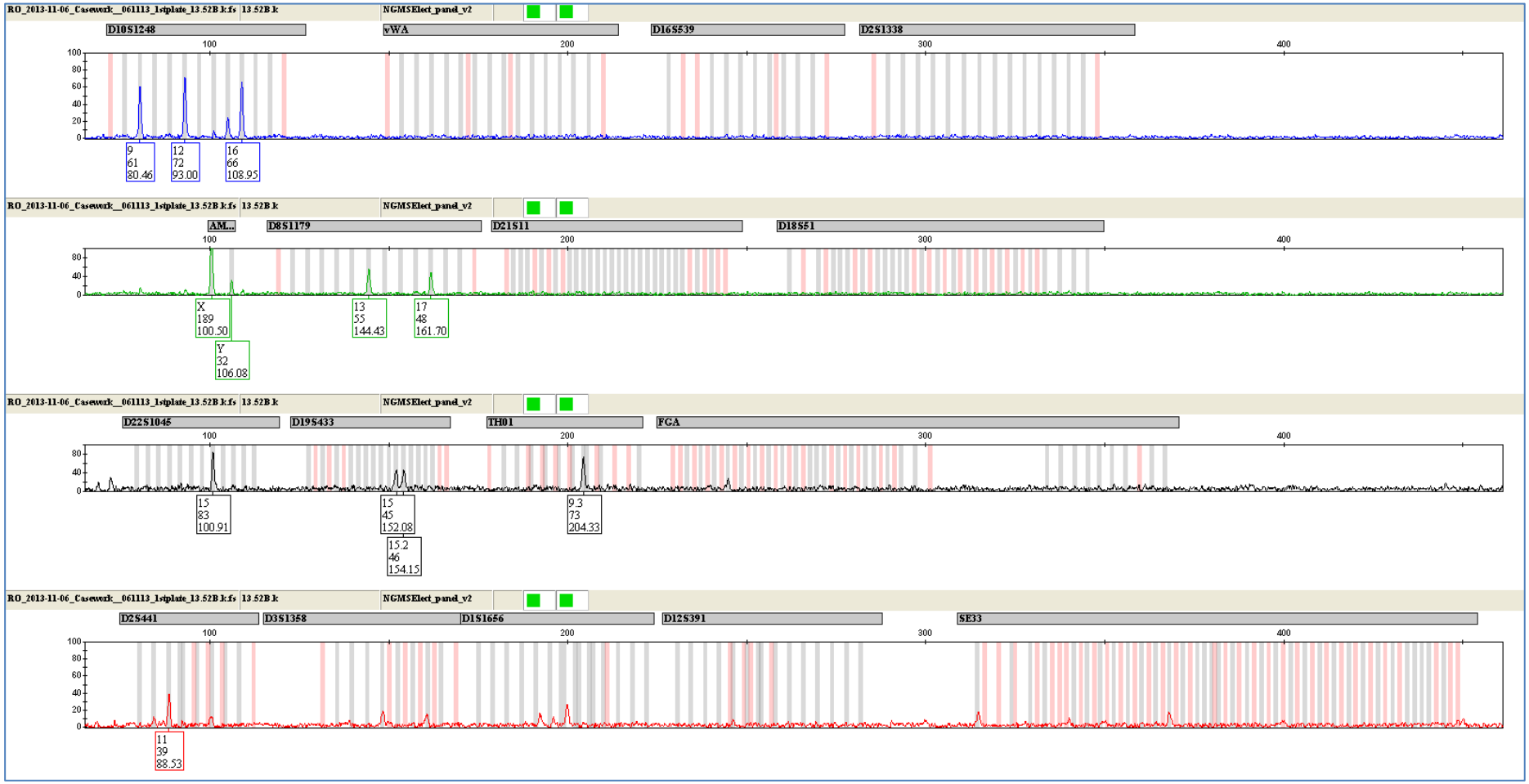


Figure 4.13 NGM SElect™ kit DNA profile obtained from case exhibit: MG523.B tape, PCR tube label: k52B. Direct PCR at 29 cycles was used to obtain the profile.

Application of case work: seizure of Methamphetamine in South Australia

DNA Profile obtained from case work

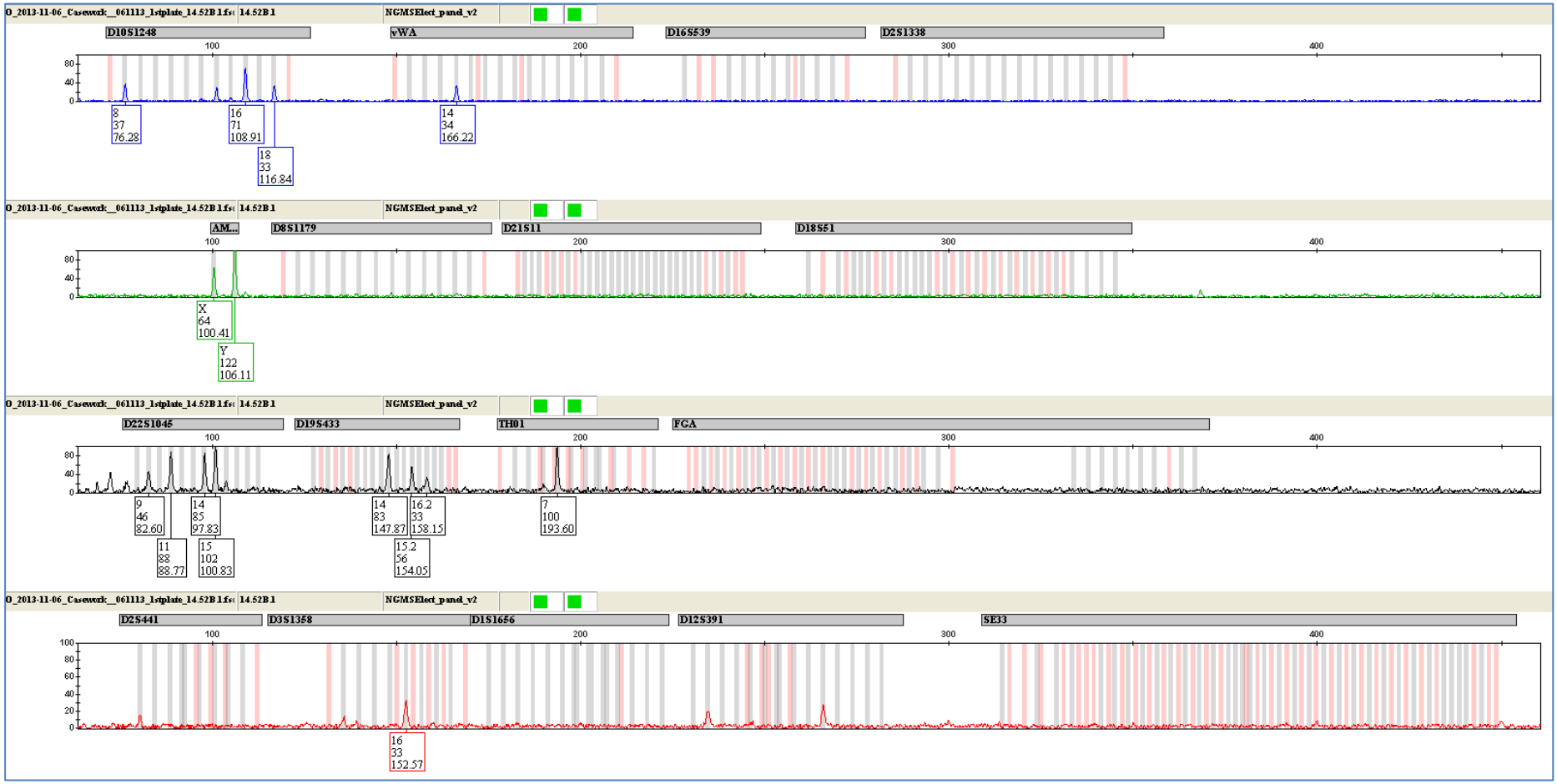


Figure 4.14. NGM SElect™ kit DNA profile obtained from case exhibit: MG523.B tape, PCR tube label: 152B. Direct PCR at 29 cycles was used to obtain the profile.

Application of case work: seizure of Methamphetamine in South Australia

DNA Profile obtained from case work

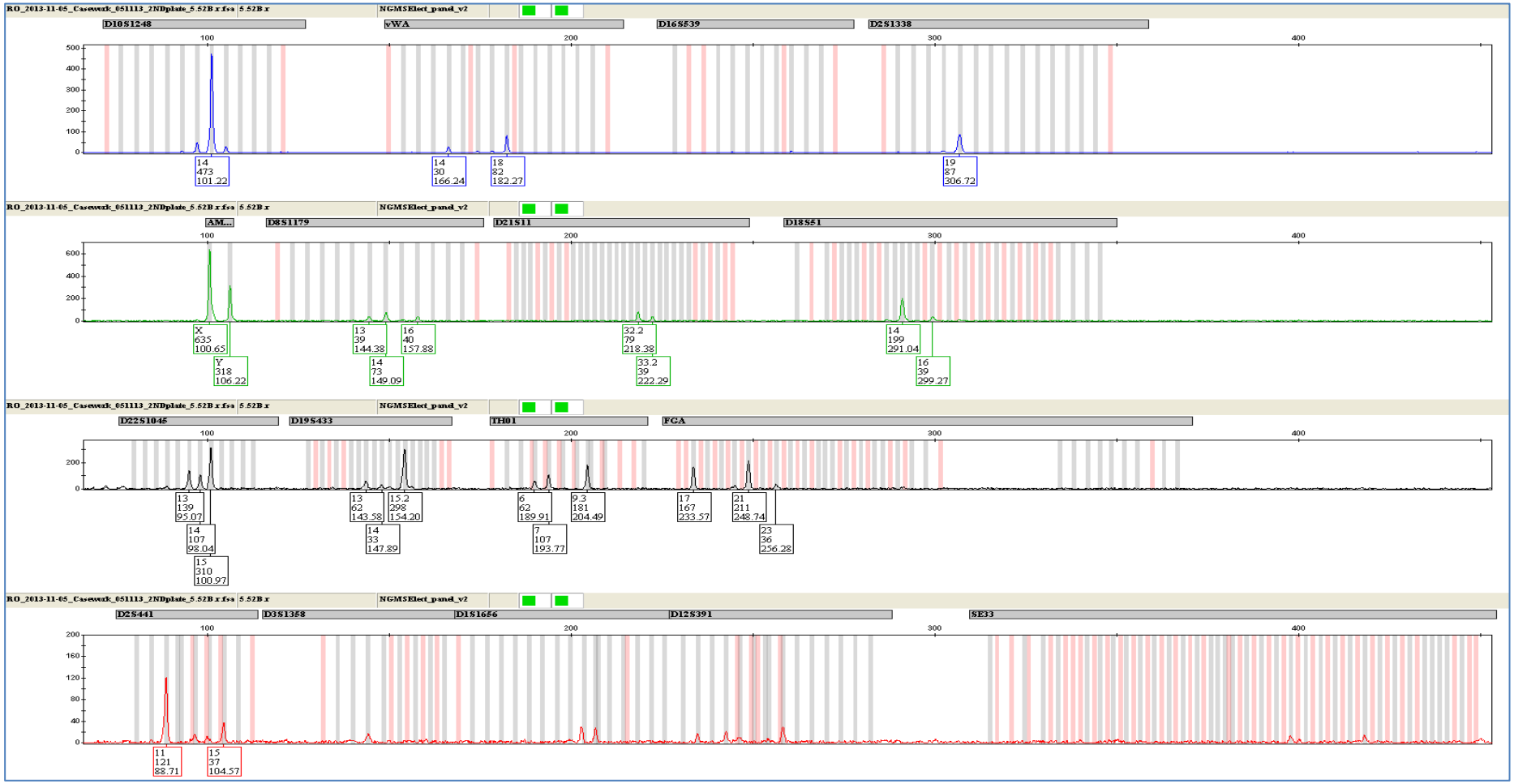


Figure 4.15. NGM SElect™ kit DNA profile obtained from case exhibit: MG523.B tape, PCR tube label: r52B. Direct PCR at 29 cycles was used to obtain the profile.

4.7 Conclusion

The results discussed in this chapter help to clarify the biological context for 'touch' DNA samples when direct PCR is used. The goal of this work was to determine if direct PCR is capable of generating STR-based DNA profiles from swab fibres that were used to recover fingermark residue. To accomplish this, individuals deposited fingermarks onto plastic substrates immediately (T0 min), 15 minutes (T15 min), 1 hour (T1 hr) and 2 hours (T2 hr) post hand washing, and nylon swabs were used to recover DNA and fibres utilised for direct PCR. It is believed from these results that an individual's inherent or circumstantial susceptibility to deposit DNA from fingermarks is partly reliant on the donor's activities (e.g. what they are doing with their hands) and emotional state (e.g. exercising or at rest) prior to handling the item. Wide inter- and intra-individual differences in profiling success were observed in the data set. Overall, the observations from this study suggest that 'touch' DNA samples may need to be processed differently to other forensically relevant biological material. There should be a fundamental shift in the way that forensic laboratories process 'touch' DNA samples in order to utilise the cell-free DNA portion of the sample which is generally discarded during a routine extraction. The extracellular component of 'touch' DNA that is washed away may provide an added value to the STR profile. Direct PCR is one viable option to explore with the presence of PCR facilitators to boost the amplification yield. Developing other methods to maximise the quantity of DNA recovered from contact surfaces would help steer police investigations in a more informative direction. Further work is needed to expand the forensic communities understanding of 'touch' DNA regarding its persistence, mechanism of transfer, and the effect of environmental inhibitors.

4.8 References

1. van Oorschot, R.A. and M.K. Jones, *DNA fingerprints from fingerprints.* Nature London, 1997: p. 767-767.
2. van Hoofstat, D.E., et al., *DNA typing of fingerprints and skin debris: sensitivity of capillary electrophoresis in forensic applications using multiplex PCR.* Proceedings of the 2nd European Symposium of Human Identification, 1998: p. 131-137.
3. Wiegand, P. and M. Kleiber, *DNA typing of epithelial cells after strangulation.* International Journal of Legal Medicine, 1997. **110**(4): p. 181-183.
4. Schulz, M. and W. Reichert, *A strategy for STR-analysis of cryptic epithelial cells on several textiles in practical casework.* Excerpta Medica International Congress Series, 2000. **1193**(1): p. 514-516.
5. van Oorschot, R.A.H., et al., *Are you collecting all the available DNA from touched objects?* International Congress Series, 2003. **1239**(0): p. 803-807.
6. Harbison, S., M. Fallow, and D. Bushell, *An analysis of the success rate of 908 trace DNA samples submitted to the Crime Sample Database Unit in New Zealand.* Australian Journal of Forensic Sciences, 2008. **40**(1): p. 49-53.
7. Ottens, R., et al., *Application of direct PCR in forensic casework.* Forensic Science International: Genetics Supplement Series, 2013. **4**(1): p. e47-e48.
8. Swaran, Y.C. and L. Welch, *A comparison between direct PCR and extraction to generate DNA profiles from samples retrieved from various substrates.* Forensic Science International: Genetics, 2012. **6**(3): p. 407-412.
9. Altshuler, H. and R. Roy, *Evaluation of direct PCR amplification using various swabs and washing Reagents.* Journal of Forensic Sciences, 2015. **60**(6): p. 1542-1552.
10. van Hoofstat, D.E.O., et al., *DNA typing of fingerprints using capillary electrophoresis: Effect of dactyloscopic powders.* Electrophoresis, 1999. **20**(14): p. 2870-2876.
11. van Renterghem, P., D. Leonard, and C. De Greef, *Use of latent fingerprints as a source of DNA for genetic identification*. Progress in Forensic Genetics 8, G.F. Sensabaugh, P.J. Lincoln, and B. Olaisen, Editors. 2000. p. 501-503.

12. Zamir, A., E. Springer, and B. Glattstein, *Fingerprints and DNA: STR typing of DNA extracted from adhesive tape after processing for fingerprints.* Journal of Forensic Sciences, 2000. **45**(3): p. 687-688.
13. Herber, B. and K. Herold, *DNA typing of human dandruff.* Journal of Forensic Science, 1998. **43**(3): p. 648-656.
14. van Renterghem, P., D. Leonard, and C. De Greef, *Use of latent fingerprints as a source of DNA for genetic identification.* Excerpta Medica International Congress Series, 2000. **1193**(1): p. 501-503.
15. Bright, J.-A. and S.F. Petricevic, *Recovery of trace DNA and its application to DNA profiling of shoe insoles.* Forensic Science International, 2004. **145**(1): p. 7-12.
16. Polley, D., et al., *An investigation of DNA recovery from firearms and cartridge cases.* Canadian Society of Forensic Science Journal, 2006. **39**(4): p. 217-228.
17. Sewell, J., et al., *Recovery of DNA and fingerprints from touched documents.* Forensic Science International: Genetics, 2008. **2**(4): p. 281-285.
18. Castella, V. and P. Mangin, *DNA profiling success and relevance of 1739 contact stains from caseworks.* Forensic Science International: Genetics Supplement Series, 2008. **1**(1): p. 405-407.
19. Caddy B, T.G., Linacre AMT. *A review of the science of low template DNA analysis*. 2008; Available from: http://www.homeoffice.gov.uk/publications/police/790604/Review_of_Low_Template_DNA_1.pdf.
20. Samie, L., et al., *Stabbing simulations and DNA transfer.* Forensic Science International: Genetics, 2016. **22**: p. 73-80.
21. van Oorschot, R.A., K.N. Ballantyne, and R.J. Mitchell, *Forensic trace DNA: a review.* Investigative Genetics, 2010. **1**(1): p. 1.
22. Raymond, J.J., et al., *Trace DNA success rates relating to volume crime offences.* Forensic Science International: Genetics Supplement Series, 2009. **2**(1): p. 136-137.
23. Gill, P., et al., *An investigation of the rigor of interpretation rules for STRs derived from less than 100 pg of DNA.* Forensic Science International, 2000. **112**(1): p. 17-40.

24. Kumar, P., et al., *Effects of latent fingerprint development reagents on subsequent forensic DNA typing: A review.* Journal of Forensic and Legal Medicine, 2015. **32**: p. 64-69.
25. Budowle, B., A.J. Eisenberg, and A.v. Daal, *Validity of low copy number typing and applications to forensic science.* Croatian Medical Journal, 2009. **50**(3): p. 207-217.
26. Meakin, G. and A. Jamieson, *DNA transfer: review and implications for casework.* Forensic Science International: Genetics, 2013. **7**(4): p. 434-443.
27. Kamphausen, T., et al., *Good shedder or bad shedder—the influence of skin diseases on forensic DNA analysis from epithelial abrasions.* International Journal of Legal Medicine, 2012. **126**(1): p. 179-183.
28. Cupples, C.M., et al., *STR Profiles from DNA Samples with "Undetected" or Low Quantifiler (TM) Results.* Journal of Forensic Sciences, 2009. **54**(1): p. 103-107.
29. Schiffner, L.A., et al., *Optimization of a simple, automatable extraction method to recover sufficient DNA from low copy number DNA samples for generation of short tandem repeat profiles.* Croatian Medical Journal, 2005. **46**(4): p. 578.
30. Hudlow, W.R., et al., *The NucleoSpin® DNA Clean-up XS kit for the concentration and purification of genomic DNA extracts: An alternative to microdialysis filtration.* Forensic Science International: Genetics, 2011. **5**(3): p. 226-230.
31. Doran, A.E. and D.R. Foran, *Assessment and mitigation of DNA loss utilizing centrifugal filtration devices.* Forensic Science International: Genetics, 2014. **13**: p. 187-190.
32. Kreader, C.A., *Relief of amplification inhibition in PCR with bovine serum albumin or T4 gene 32 protein.* Applied and Environmental Microbiology, 1996. **62**(3): p. 1102-1106.
33. Akane, A., et al., *Identification of the heme compound copurified with deoxyribonucleic acid (DNA) from bloodstains, a major inhibitor of polymerase chain reaction (PCR) amplification.* Journal of Forensic Sciences, 1994. **39**: p. 362-362.
34. Xin, Z., et al., *High-throughput DNA extraction method suitable for PCR.* Biotechniques, 2003. **34**(4): p. 820-827.

35. Bachmann, B., W. Lüke, and G. Hunsmann, *Improvement of PCR amplified DNA sequencing with the aid of detergents.* Nucleic Acids Research, 1990. **18**(5): p. 1309.
36. Pagan, F., et al., *Comparison of DNA extraction methods for identification of human remains.* Australian Journal of Forensic Sciences, 2012. **44**(2): p. 117-127.
37. Kang, J., M.S. Lee, and D.G. Gorenstein, *The enhancement of PCR amplification of a random sequence DNA library by DMSO and betaine: application to in vitro combinatorial selection of aptamers.* Journal of Biochemical and Biophysical Methods, 2005. **64**(2): p. 147-151.
38. Marshall, P.L., J.L. King, and B. Budowle, *Utility of amplification enhancers in low copy number DNA analysis.* International Journal of Legal Medicine, 2014: p. 1-10.
39. Nilsson, M., et al., *Comparison of DNA polymerases for improved forensic analysis of challenging samples.* Forensic Science International: Genetics, 2016.
40. Musso, M., et al., *Betaine, dimethyl sulfoxide, and 7-deaza-dGTP, a powerful mixture for amplification of GC-rich DNA sequences.* The Journal of Molecular Diagnostics, 2006. **8**(5): p. 544-550.
41. Raymond, J.J., et al., *Trace DNA and street robbery: a criminalistic approach to DNA evidence.* Forensic Science International: Genetics Supplement Series, 2009. **2**(1): p. 544-546.
42. Kita, T., et al., *Morphological study of fragmented DNA on touched objects.* Forensic Science International Genetics, 2008. **3**(1): p. 32-6.
43. Nakamura, K., et al., *The relationship between apoptosis and keratinization in human epidermis.* Acta Histochemica et Cytochemica, 1999. **32**(1): p. 77-83.
44. Quinones, I. and B. Daniel, *Cell free DNA as a component of forensic evidence recovered from touched surfaces.* Forensic Science International Genetics, 2012. **6**(1): p. 26-30.
45. Balogh, M.K., et al., *STR genotyping and mtDNA sequencing of latent fingerprint on paper.* Forensic Science International, 2003. **137**(2-3): p. 188-195.

46. Alessandrini, F., et al., *Fingerprints as evidence for a genetic profile: morphological study on fingerprints and analysis of exogenous and individual factors affecting DNA typing.* Journal of Forensic Sciences, 2003. **48**(3): p. 586-592.
47. Sen, S., *Programmed cell death: concept, mechanism and control.* Biological Reviews, 1992. **67**(3): p. 287-319.
48. Teraki, Y. and T. Shiohara, *Apoptosis and the skin.* European Journal of Dermatology, 1999. **9**(5): p. 413-26.
49. Lagoa, A.M., T. Magalhães, and M.F. Pinheiro, *Autosomic STR, Y-STR and miniSTR markers evaluation for genetic analysis of fingerprints.* Forensic Science International: Genetics Supplement Series, 2008. **1**(1): p. 435-436.
50. Lacerenza, D., et al., *A molecular exploration of human DNA/RNA co-extracted from the palmar surface of the hands and fingers.* Forensic Science International: Genetics, 2016. **22**: p. 44-53.
51. Lowe, A., et al., *The propensity of individuals to deposit DNA and secondary transfer of low level DNA from individuals to inert surfaces.* Forensic Science International, 2002. **129**(1): p. 25-34.
52. Goray, M., et al., *Shedder status—An analysis of self and non-self DNA in multiple handprints deposited by the same individuals over time.* Forensic Science International: Genetics, 2016. **23**: p. 190-196.
53. Phipps, M. and S. Petricevic, *The tendency of individuals to transfer DNA to handled items.* Forensic Science International, 2007. **168**(2-3): p. 162-8.
54. Goray, M., R.J. Mitchell, and R.A. van Oorschot, *Investigation of secondary DNA transfer of skin cells under controlled test conditions.* Legal medicine, 2010. **12**(3): p. 117-120.
55. Wickenheiser, R.A., *Trace DNA: a review, discussion of theory, and application of the transfer of trace quantities of DNA through skin contact.* Journal of Forensic Sciences, 2002. **47**(3): p. 442-450.
56. Atamaniuk, J., et al., *Increased concentrations of cell-free plasma DNA after exhaustive exercise.* Clinical Chemistry, 2004. **50**(9): p. 1668-1670.
57. Raymond, J.J., et al., *Trace evidence characteristics of DNA: a preliminary investigation of the persistence of DNA at crime scenes.* Forensic Science International: Genetics, 2009. **4**(1): p. 26-33.

58. van den Berge, M., et al., *Prevalence of human cell material: DNA and RNA profiling of public and private objects and after activity scenarios.* Forensic Science International: Genetics, 2016. **21**: p. 81-89.
59. Daly, D.J., C. Murphy, and S.D. McDermott, *The transfer of touch DNA from hands to glass, fabric and wood.* Forensic Science International: Genetics, 2012. **6**(1): p. 41-46.
60. Linacre, A., et al., *Generation of DNA profiles from fabrics without DNA extraction.* Forensic Science International: Genetics, 2010. **4**(2): p. 137-141.
61. Goray, M. and R.A. van Oorschot, *The complexities of DNA transfer during a social setting.* Legal Medicine, 2015. **17**(2): p. 82-91.
62. Ladd, C., et al., *A systematic analysis of secondary DNA transfer.* Journal of Forensic Sciences, 1999. **44**: p. 1270-1272.
63. Goray, M., et al., *Secondary DNA transfer of biological substances under varying test conditions.* Forensic Science International: Genetics, 2010. **4**(2): p. 62-67.
64. Goray, M., R.A. van Oorschot, and J.R. Mitchell, *DNA transfer within forensic exhibit packaging: Potential for DNA loss and relocation.* Forensic Science International: Genetics, 2012. **6**(2): p. 158-166.
65. Poy, A.L. and R.A. van Oorschot, *Trace DNA presence, origin, and transfer within a forensic biology laboratory and its potential effect on casework.* Journal of Forensic Identification, 2006. **56**(4): p. 558.
66. Djuric, M., et al., *DNA typing from handled items.* Forensic Science International: Genetics Supplement Series, 2008. **1**(1): p. 411-412.
67. van Oorschot, R.A., G. Glavich, and R.J. Mitchell, *Persistence of DNA deposited by the original user on objects after subsequent use by a second person.* Forensic Science International: Genetics, 2014. **8**(1): p. 219-225.
68. Cale, C.M., et al., *Could Secondary DNA Transfer Falsely Place Someone at the Scene of a Crime?* Journal of Forensic Sciences, 2016. **61**(1): p. 196-203.
69. Farmen, R.K., et al., *Assessment of individual shedder status and implication for secondary DNA transfer.* Forensic Science International: Genetics Supplement Series, 2008. **1**(1): p. 415-417.

70. Helmus, J., T. Bajanowski, and M. Poetsch, *DNA transfer—a never ending story. A study on scenarios involving a second person as carrier.* International Journal of Legal Medicine, 2016. **130**(1): p. 121-125.
71. Gill, P., *Analysis and implications of the miscarriages of justice of Amanda Knox and Raffaele Sollecito.* Forensic Science International: Genetics, 2016. **23**: p. 9-18.
72. Stanciu, C.E., et al., *Optical characterization of epidermal cells and their relationship to DNA recovery from touch samples.* Research, 2015. **4**.
73. Templeton, J.E., et al., *DNA capture and next-generation sequencing can recover whole mitochondrial genomes from highly degraded samples for human identification.* Investigative Genetics, 2013. **4**(1): p. 1-13.
74. Vandewoestyne, M., et al., *Presence and potential of cell free DNA in different types of forensic samples.* Forensic Science International Genetics, 2013. **7**(2): p. 316-20.
75. Eckhart, L., H. Fischer, and E. Tschachler, *Mechanisms and emerging functions of DNA degradation in the epidermis.* Frontiers in Bioscience (Landmark edition), 2011. **17**: p. 2461-2475.
76. Oleiwi, A., et al., *The relative DNA-shedding propensity of the palm and finger surfaces.* Science & Justice, 2015. **55**(5): p. 329-334.
77. Templeton, J. and A. Linacre, *DNA profiles from fingermarks.* BioTechniques, 2014. **57**(5): p. 259.

Appendices - Chapter IV

Appendix A – Poster: Genetic profiling from challenging samples: direct PCR of touch DNA.

Presented at “The International Society for Forensic Genetics” conference, Melbourne, Australia, 2013.

Genetic profiling from challenging samples: direct PCR of touch DNA

Jennifer E.L Templeton[1], Renée Ottens[1], Viviana Paradiso[1], Oliva Handt[2], Duncan Taylor[2] , Adrian. Linacre[1]
[1]School of Biological Sciences, Flinders University, Adelaide, Australia
[2]Forensic Science South Australia, Adelaide, Australia

Forensic Biology Lab South Australia

Presented at the 2013 ISFG Conference Melbourne, Australia

Funding was provided by the Department of Justice, South Australia

Background – Touch DNA

Trace DNA evidence recovered from handled items can be close to the limit of detection for most STR-based DNA profiling kits. The DNA may be:

- low concentration (< 100 pg);
- chemically damaged;
- degraded / fragmented (<200 bp) and/or;
- contaminated (environmental or human DNA)

Therefore, improved methods are required for collection and processing of DNA from handled items to improve the success rate of DNA profiling.

Introduction

Handled items are one of the most common forms of evidence submitted to a forensic laboratory and one of the least successful in terms of generating a full DNA profile. Current methodology subjects the sample to a DNA extraction to isolate and purify the target DNA, typically using a solid phase substrate, prior to amplification. DNA extraction encompasses multiple tube transfers that contribute to a significant loss (20-76%) of sample DNA [1] and is where the sample is more susceptible to contamination than any other stage in the analysis [2]. Omitting the DNA extraction stage, known as direct PCR, has the potential to minimise loss of target DNA in a critical sample. Further, removing steps involved in DNA extraction will reduce both labour time and costs, while increasing the likelihood of obtaining a meaningful DNA profile for interpretation.

Aim

The aim of this study was to improve the success rate of STR-based DNA profiling from touch DNA samples by:

- improving methods of DNA recovery;
- omitting the extraction process and;
- minimising stochastic effects inherent in analysing low DNA templates.

Method – direct PCR

❑ Improving methods of recovery – comparing the efficiency of swab types

1 ng of control DNA (2800M,Promega) was deposited onto sterile plastic slides, swabbed using either foam (Whatman), nylon (Copan) or cotton (Livingstone) swabs, and subjected to direct PCR. Amplification conditions followed the **Profiler Plus™ kit (ABI)** guidelines in a final 25 µL reaction volume and standard **28 cycles**.

❑ Fingerprint profiling

A fingerprint was deposited onto a sterile plastic microscope slide 2 hours after the donor washed their hands. The slide was swabbed using the optimum method of DNA recovery and the swab subjected to direct PCR. Amplification: **NGM™ kit (ABI)** in a final 25 µL reaction and 29 cycles.

A fingerprint was deposited onto a brass bullet cartridge 8 hours after the donor washed their hands. The cartridge was swabbed using the optimum method of DNA recovery and subjected to direct PCR. Amplification: **Profiler Plus™ kit (ABI)** in a final 25 µL reaction and 29 cycles.

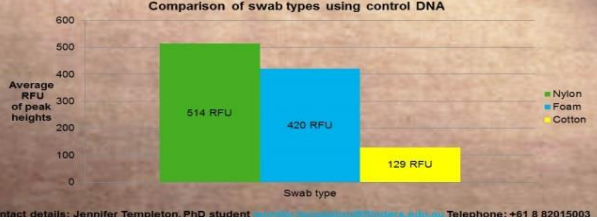
- 3130xl capillary electrophoresis

All PCR products were run on the 3130xl (ABI) genetic analyser. Profiles were analysed using the GeneMapper ID v3.2 software and a peak amplitude threshold of 30 RFU was used to assign alleles.

Results

Optimum method of recovery:

Nylon flocked swabs (Copan) pre-moistened with water and fibres cut from the tip prior to PCR



Contact details: Jennifer Templeton, PhD student Telephone: +61 8 82015003

Results - cont

DNA PROFILE FROM A FINGERPRINT DEPOSIT – plastic

Figure 1. Full STR DNA profile generated, using direct PCR and 29 cycles, from a fingerprint deposited onto a sterile plastic slide 2 hours after the donor washed their hands. Amplification kit: NGM™ (ABI).

DNA PROFILE FROM A FINGERPRINT DEPOSIT - brass

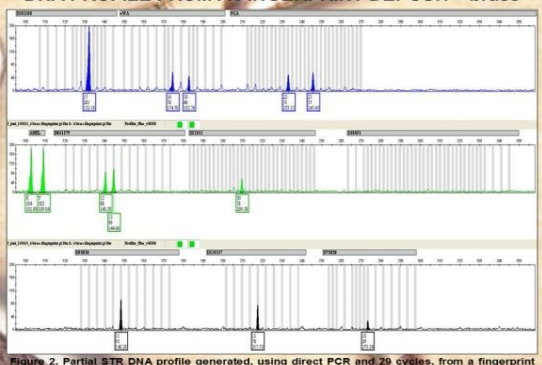


Figure 2. Partial STR DNA profile generated, using direct PCR and 29 cycles, from a fingerprint deposited onto a brass bullet cartridge 8 hours after the donor washed their hands. Amplification kit: Profiler Plus ™ (ABI).

Concluding remarks

Based on these data, direct PCR has the potential to:

- generate a full STR DNA profile from a fingerprint deposited on a plastic slide, using only 29 cycles of PCR;
- generate a partial DNA profile from a fingerprint deposited on a brass bullet cartridge, using only 29 cycles of PCR
- reduce the chance of contamination (allele drop-in);
- minimise stochastic effects (allele and locus drop-out and heterozygote imbalance) inherent in analysing low DNA templates and;
- circumvent the need for increased cycle number when analysing samples that contain sub-optimal concentrations of DNA

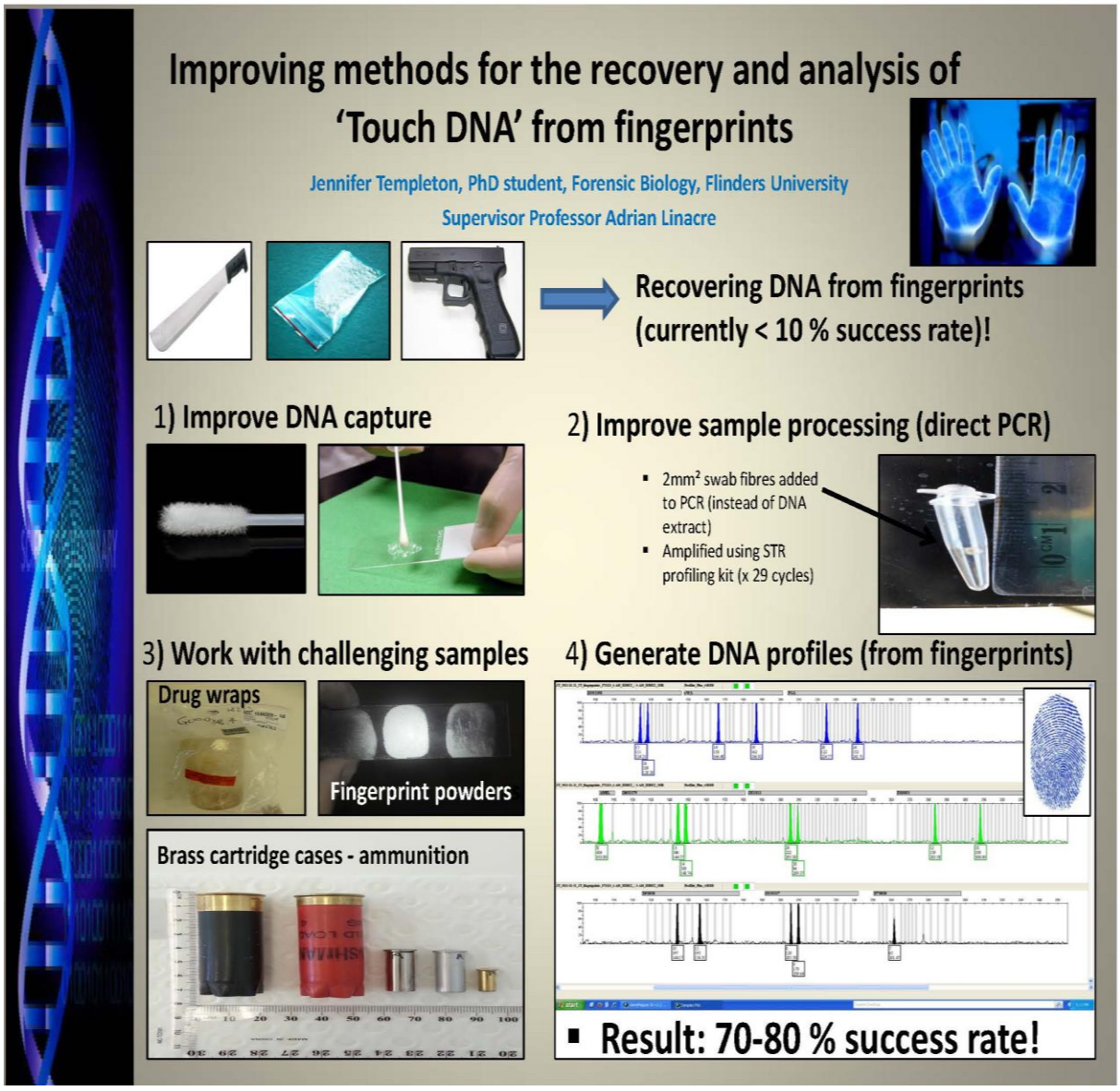
Laboratories may benefit from using direct PCR on handled items due to its high-throughput potential and increased sensitivity. However, the success and quality of DNA profiles obtained using direct PCR is dependent on the nature of the material examined and the presence of PCR inhibitors; which are generally eliminated during the extraction

Future samples

References:
- [1] van Oorschot RAH, Phelan DG, Furlong S, Scarfo GM, Holding NL, Cummins MJ. Are you collecting all the available DNA from touched objects? International Congress Series. 2003;1239:603-7.
- [2] Butler JM. Forensic DNA Typing: Biology, Technology, and Genetics of STR Markers. 2nd Edition ed. New York: Elsevier Academic Press; 2005.
- [3] Linacre A, Pekarek V, Swaran YC, Tobe SS. Generation of DNA profiles from fabrics without DNA extraction. Forensic science international genetics. 2010;4:137-41.
- [5] Ottens R, Taylor D, Abarno D, Linacre A. Successful direct amplification of nuclear markers from a single hair follicle. Forensic science,

Appendix B – Three minute (3M) thesis competition static slide.

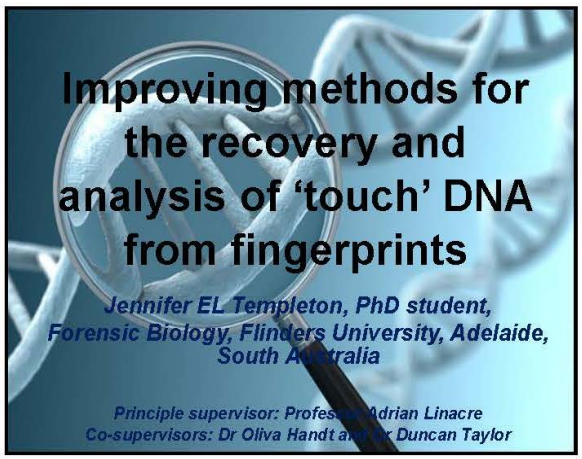
One static slide presented to the school of Biology, Flinders University, 2014.

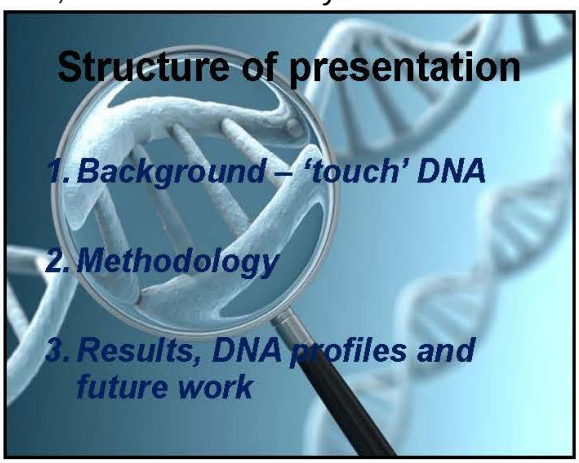


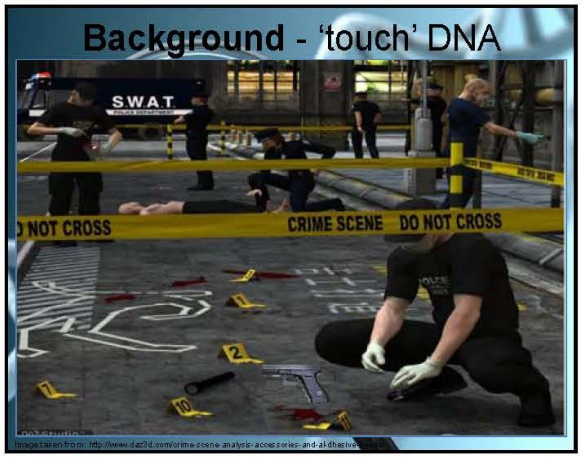
Generating informative DNA profiles from 'touch/contact' swabs has been reported as less than 10 % in terms of profiles considered 'uploadable' to the DNA database. Therefore, success rates in this thesis relate to samples being considered 'uploadable' to the Australian DNA database.

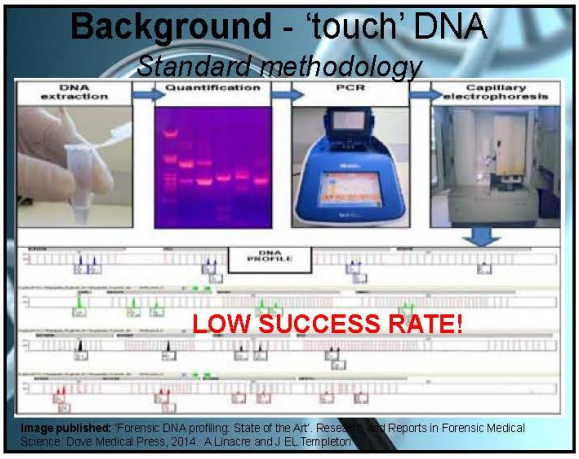
Appendix C – Oral presentation.

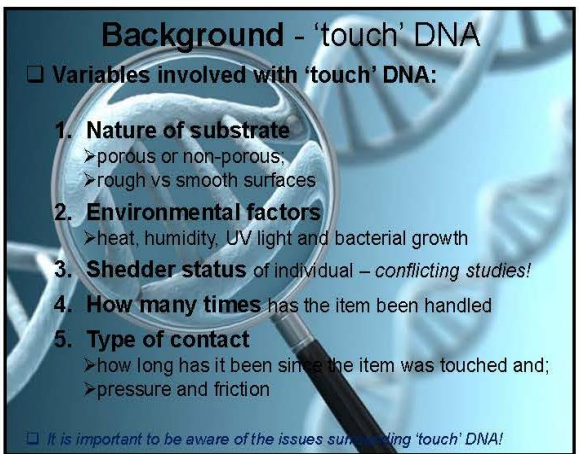
Templeton, J.E., Handt, O., Taylor, D., and Linacre, A. "The American Academy of Forensic Sciences" conference, Orlando, Florida, America. February 2015.











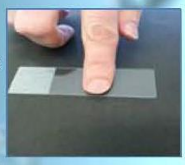
Methodology

❑ **Optimising swabbing method**

- Swab type;
- Swabbing technique and;
- Swabbing media

❑ **270 fingermarks – source of latent DNA:**

- Samples:
  → 25 females and 10 males
- Time:
  → 15 minutes, 1 hour and 2 hours
- Contact:
  → 15 seconds, medium pressure
- Storage:
  → slides stored 4°C until processed



Methodology

❑ **Optimised recovery and direct PCR**

➢ *Amplification kit: AmpFℓSTR ProfilerPlus™*

(2mm² nylon fibres in PCR tube)

Advantages to direct PCR:

➢ no wash steps;
➢ no tube changes;
➢ reduce labour time and cost;
➢ minimise risk of contamination

❑ **Success rate is 65 – 70 %**

Methodology

❖ Published papers – most use *Low copy number (LCN) methodology*

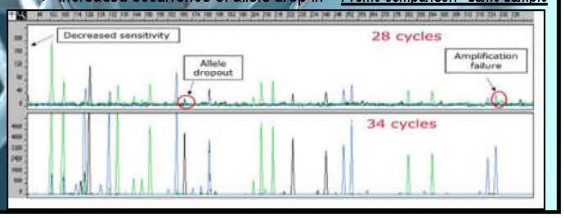
❑ Forensic Science Service (FSS) UK, *1990s*

➢ increases cycle number 28 → 34 cycles

❑ Forensic labs avoid LCN analysis due to:

➢ stochastic effects and;
➢ increased occurrence of allele drop-in

Profile comparison - same sample



Results

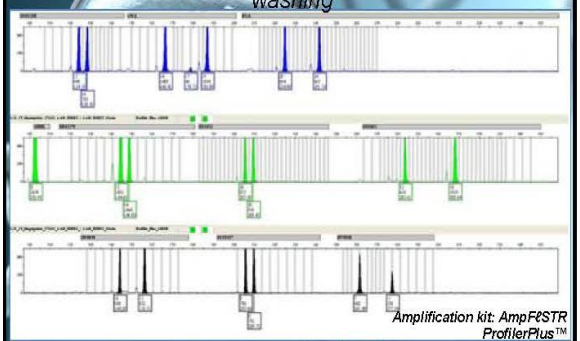
❑ Optimum swabbing technique[1]:

➢ nylon swabs, pre-moistened with detergent - triton X (0.1 %)
➢ double swabbing more effective than a single swab

[1] **J. Templeton, et al.,** Genetic profiling from chall[illegible] samples: Direct PCR of touch DNA, Forensic Sci. Int. Gene. Suppl. (2013)

Results

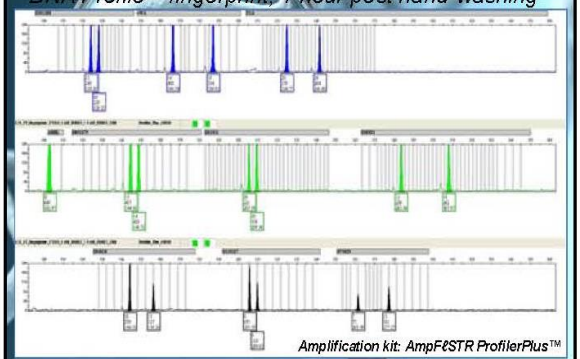
DNA Profile – fingerprint, 15 minutes post hand-washing



Manuscript published Biotechniques: **'DNA profiles from fingerm[illegible]** EL Templeton and A Linacre, 2014

Results

DNA Profile – fingerprint, 1 hour post hand-washing



Results

DNA Profile – fingerprint, 2 hours post hand-washing

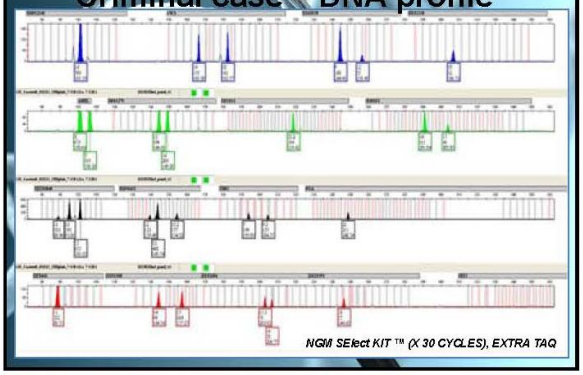
Results

DNA extraction

- ❑ Individuals selected – generated full DNA profiles from direct PCR
- ❑ Repeated fingerprint test
  - whole swab extracted - DNA IQ™ kit
  - amplification kit: AmpFℓSTR ProfilerPlus™
  - 10 samples processed
- ❑ Variation in amount of DNA deposited
  - → *No result from extracted swabs!*
- ✓ Direct PCR does have potential for latent DNA!

Results

Criminal case – DNA profile

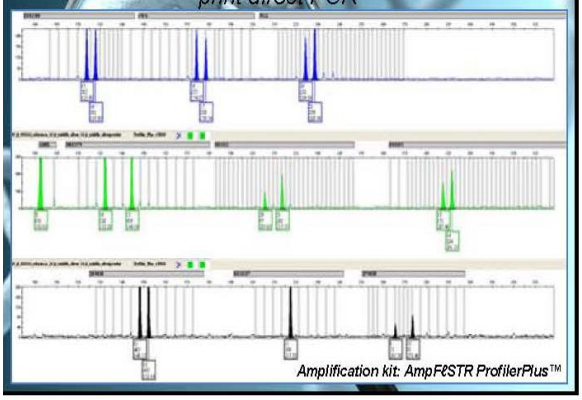


Results and future work

- ❑ **Fingerprint powders:**
  - Silver Aluminium (stearic acid and aluminium powder);
  - White (titanium dioxide, zinc stearate and lycopodium) and;
  - Black (carbon)

Results

Fingerprint powders *– Aluminium treated print direct PCR*



To conclude

- ❑ **Variables will affect the success rate**
  - Including swab type, swabbing reagent and method of processing
- ❑ **Direct PCR generates full DNA profiles from fingerprints**
  - 15 minutes post hand-washing
  - No increased PCR cycle number!
- ❑ **Success rate is 65 – 70 %** (with controlled sampling)
- ❑ **Technique was applied in a criminal case – success!**
- ❑ **Future work includes:**
  - PCR inhibitors e.g. fingerprint powders
  - Other chemical enhancement methods

Appendix D – Award: Ross Vining memorial student scholarship.

Appendix E – Award, 3M thesis competition.

3MT® THREE MINUTE THESIS
FOUNDED BY THE UNIVERSITY OF QUEENSLAND

Flinders UNIVERSITY
inspiring achievement

School Winner

is awarded to

Jennifer Templeton

Research higher degree student
who participated in the

Three Minute Thesis Competition 2014

OCTOBER 2014

David Day

Professor David Day
Deputy Vice-Chancellor (Research)
Flinders University

Jeri Kroll

Professor Jeri Kroll
Dean of Graduate Research
Deputy Vice-Chancellor (Research) Office
Flinders University

The Three Minute Thesis (3MT™) is an academic competition developed by
The University of Queensland (UQ), Australia for research students.

Chapter V
Direct PCR of ‘touch’ DNA:
Effect of PCR inhibitors and environmental exposure

Publications included in Chapter V

1. **Templeton, Jennifer EL**, *et al*. *DNA profiles from fingermarks*: *A mock case study.* Forensic Science International: Genetics Supplement Series 5 (2015): e154-e155.
2. **Templeton, Jennifer EL**, *et al. Direct PCR DNA profiling of dactyloscopic powdered fingermarks* (Accepted, May 2017): http://dx.doi.org/10.1016/j.fsigen.2017.05.006.

Oral presentation

1. **Templeton, J.EL.,** and Linacre, A. *DNA profiles from fingermarks*. Presented at "The Australian and New Zealand Forensic Science Society", ANZFSS conference, Auckland, NZ. 2016. *Presented by Adrian Linacre (on my behalf).*

Appendix

a) Supplementary Information – preliminary results – Effect of fingerprint powders on direct PCR.

b) Supplementary Information - MinElute® PCR purification test results.

c) Supplementary Information - Measuring the reliability, robustness and reproducibility of direct PCR.

d) Supplementary Information - Pre-soak lysis method used on masking tape and cartridge cases.

e) Poster: Renée Blackie, **Jennifer EL Templeton**, Duncan Taylor, Emily Rowe, Oliva Handt, Adrian Linacre. *Direct PCR: Successes and Limitation*s. Presented at "The Australian and New Zealand Forensic Science Society", ANZFSS conference, Auckland, NZ. 2016.

5.1 Preface

The recovery of ‘touch’ DNA is highly variable and dependent on many factors. Two main parameters that affect the robustness of ‘touch’ DNA profiling are the quantity of DNA template [1, 2], and quality of DNA recovered [3] (i.e. damaged or degraded). Many variables will influence these parameters such as:

- the characteristics of the donor that deposits the DNA (e.g. condition of skin);
- the environmental conditions of the fingermark deposit (e.g. wet/humid/dry/hot);
- presence of PCR inhibitors (e.g. fingerprint powders or other chemical enhancers, soil, microbes);
- nature of the substrate that was touched (e.g. porous/non-porous, rough/ smooth);
- pressure/friction used during the transfer of DNA to the item;
- time between fingermark deposition and collecting DNA (i.e. DNA will degrade over time); and
- method of DNA recovery used and analysis.

Research into the ‘shedder’ status of individuals has been investigated to varying degrees [4-7] and discussed previously. Further studies have explored the transfer and persistence of ‘touch’ DNA [8-10]. This data chapter focuses mainly on the effect of crime scene inhibitors (e.g. samples tarnished with fingerprint powders) and environmental inhibitors (e.g. UV light, bacteria and soil) on the resulting DNA profiles obtained by direct PCR. The ability to type autosomal DNA from a previously enhanced fingerprint or any ‘touch’ DNA deposit that has been subjected to harsh conditions could provide new avenues for forensic investigations.

This data chapter is split into 3 phases, each of which examines aspects relating to the use of direct PCR on challenging substrates:

1) The first phase of the study explores the effect of fingerprint powders on the direct PCR approach. Preliminary data show the ability of direct PCR to generate DNA profiles in the presence of fingerprint powders. Results are extended into a full length manuscript (submitted to FSI:G in October 2016)

where 160 powdered fingerprint samples are processed using the direct PCR approach.

2) The second phase of the study examines a mock case scenario and the potential for using direct PCR on fingermarks deposited on various forensic exhibits (e.g. tape, plastic, cartridge cases; brass, nickel and aluminium, glass and wood). Results were extended into a peer reviewed publication in the journal Forensic Science International: Genetics Supplement Series. One case study examines the profiling ability of fingermarks left exposed to the environment for up to 8 days. Additional supplements were added to the PCR (i.e Prep-n-Go™ buffer (ABI) and Ampli*Taq* Gold® DNA polymerase (ABI)) to boost amplification yield.

3) The third phase of the study explores the use of optimised swabbing and direct PCR on a suite of forensic substrates under controlled laboratory conditions, and the use of PCR purification columns (i.e. MinElute® PCR purification, Qiagen) was investigated. The appendix provides additional data on preliminary experiments, including the use of a pre-soak buffer lysis method for specific substrate types, such as cartridge cases and tape cuttings.

5.2 Aims of study:

1. To determine the effect of fingerprint powders on the direct PCR approach;
2. To explore the success and limitations of direct PCR; substrates containing 'touch' DNA were exposed to varying degrees of out-door environmental factors for a set period of time. Various surface types and materials were examined;
3. To investigate the success of direct PCR using swab fibres that retrieved DNA from a suite of forensic related substrates and;
4. The final aim investigates the quantity and quality of DNA profiles obtained by post-PCR purification. Further work explores the potential for cartridge cases and tape cuttings to be processed by a pre-soak lysis method prior to amplification.

5.3 Introduction

A 'touch' DNA sample may undergo structural and chemical modifications which can lead to damage and degradation [3], hampering the chances of obtaining a good quality DNA profile.

5.3.1 Environmental effect on DNA

In natural conditions - protected within the cell - DNA is a hydrated macromolecule with between 8-10 tightly bound water molecules per nucleotide residue [11]. Outside of its natural environment, DNA is subjected to environmental conditions that may adversely affect the rate of DNA degradation [11]. A dry environment is thought to reduce the rate of degradation, in contrast to a humid environment which can promote the growth of bacteria [12] and facilitate hydrolytic enzymes. More specifically, fingermarks deposited on objects outdoors may be exposed to harsh environmental surroundings (e.g. humidity, soil, micro-organisms and UV exposure from the sun) that can lead to further DNA degradation. It is for these reasons that forensic laboratories choose to store swabs containing trace DNA in a cold, dry, dark environment in order to preserve the DNA content.

Damaged/degraded DNA

Cells deposited by fingermarks, frequently found at crime scenes, develop an apoptotic phenomenon due to endogenous endonucleases [12] that may cause DNA fragmentation [3]. Cells are known to undergo one of two programmed cell death patterns (i.e. apoptosis or necrosis) depending on the biochemical events that lead to characteristic cell changes and DNA degradation. Hydrolysis (i.e. addition of water) and oxidation (i.e. loss of an electron) processes can further modify the DNA by degradation at a much slower rate. Jennings *et al.* (1975) [12] characterised apoptosis as "energy dependent programmed cell death accompanied by condensation of cytoplasm, loss of plasma membrane, segmentation of the nucleus, and extensive degradation of chromosomal DNA into oligomers around 180 bp in length". The second cell death pattern that can affect DNA degradation is necrosis. According to Alaeddini *et al*. (2010) [13] necrosis is characterised by "increased cell volume, swelling of cytoplasmic organelles, chromatin condensation, and a random pattern of DNA degradation".

In extreme cases of DNA degradation, there may be no surviving endogenous DNA above 100 bp in length [14]. More common in ancient DNA studies, miscoding lesions are known to prevent the DNA polymerase from extending the growing DNA chain [15]. It is important to be aware of the extent of DNA degradation and the inability to recover DNA that is severely damaged and degraded. A sample may undergo preferential or complete amplification failure as a result of DNA degradation or low-template DNA. Consequently, this would lead to difficulties with analysis as the primers that are used to target loci may not consistently find and hybridise to the entire set of DNA molecules that are available [15].

5.3.2 Interpretation issues

Preferential amplification is due to unequal sampling of alleles at heterozygote loci, and can result in the failure to detect one or both alleles (as discussed in Chapter I). Random sampling means that different DNA alleles may be detected after re-amplification of the same DNA extract [2]. Generally shorter amplicons are preferentially amplified over longer amplicons [16]. A study by Briggs *et al.* (2009) [17] demonstrates that complete amplification failure can be a result of PCR

inhibition, low-template DNA, or degradation of DNA template below target amplicon sizes (i.e. 100 – 400 bp). The LCN method of increased PCR cycle number is a highly sensitive application to detect low-level DNA (as discussed in Chapter I). However, most laboratories choose not to use the LCN process given the challenges faced in court when presenting evidence with this approach. An extensive review on detecting incidental DNA with LCN methodology was published and reports on the benefits and limitations of this approach [18]. Allele 'drop-in' is a major concern when interpreting profiles as extra peaks in the profile may appear to be true but are actually a result of contamination. The extraction step is said to be where the sample is more prone to contamination that at any other stage in the sample handling process [19]. Eliminating the extraction stage will therefore minimise tube transfers and the chance of introducing extraneous DNA into the reaction.

The recent approach to understanding why DNA samples fail to yield informative results has focussed on examining DNA degradation [13]. Other studies have focussed on understanding tissue preservation [20] and DNA repair mechanisms [21]. However, one area of research that has been explored to a lesser extent is the co-existence of PCR inhibitors in a sample and the effect they may have on the resulting DNA profile.

5.3.3 PCR inhibitors

DNA samples often contain impurities which can inhibit PCR. Co-existing impurities may also prevent sufficient cell lysis during an extraction [22, 23]. Most fingerprint enhancement methods will not interfere with the ability to perform DNA profiling [24-29]. However, some reagents and powders used to enhance fingerprints have been shown to have a negative effect on the subsequent profiling ability of the print [26, 30]. In particular, magnetic-based powders that are thought to interfere with the binding of the DNA-bead-complex extraction methodology have failed to yield measureable DNA results [31].

Inhibitors may interfere with and compromise all aspects of the PCR including DNA template, primers, Mg^{2+} concentrations and the polymerase enzyme [32]. Research has focussed on the effect inhibitors have on the polymerase enzyme. Inhibitors may

chelate with Mg^{2+} that is required by the *Taq* polymerase to function. The inhibitor may act as a blocking agent of the active site to prevent the enzyme from extending the growing chain. In addition, proteases, phenol, and detergents are known to degrade or denature the polymerase enzyme [32-34]. A study by Al-Soud *et al.* (1998) [35] reported that Ampli*Taq* Gold® DNA polymerase – routinely used in forensic amplifications - was sensitive to Ca^{2+} and concentrations above 3 mM in the PCR were found to have an adverse effect. It was found that increased Mg^{2+} concentration would compensate for the reduced activity [35]. Young *et al.* (1993) reported that inhibitors may also contain "*oligomeric compounds with free phenolic groups that oxidise to form Quinones that bind to and inactivate the DNA polymerase*" [36]. The result is a decrease in V_{max} (i.e. maximum velocity that the enzyme can reach).

Quantitative PCR can be used to detect PCR inhibitors by using control DNA of a known concentration and calculating PCR efficiency [37]. A study by Opel *et al.* (2010) [38] used quantitative PCR (qPCR) to determine how primer sequences, amplicon lengths and melting temperatures affect DNA in the presence of inhibitors, and found that primers with a high melting temperature are less affected by inhibition. Designing a robust PCR buffer system to inactivate as many different types of inhibitors as possible would be a major advantage for the analysis of compromised samples.

5.3.4 Improving the quality of DNA profiles

Trouble-shooting strategies should be implemented to minimise PCR inhibition. New versions of STR profiling kits have enhanced buffers to boost amplification. Some laboratories already choose to include extra components in the reaction mix such as BSA [39] or DMSO [40], to overcome PCR inhibitors. A study by Bourke *et al.* (1999) [41] used 0.4 mM NaOH treatment to neutralize inhibitors of *Taq* DNA polymerase, and then passed the sample through microcon®-100 columns (Millipore) to separate the DNA from smaller sized inhibitors; however, this carries an increased risk of contamination. A later study by Kemp *et al.* (2006) [42] supported the use of microcon® columns for DNA purification and reported an increase in success rate when using microcon®-30 size columns for fragmented DNA template. Primorac *et*

al. (2004) [42] attempted to use the NaOH method to purify DNA recovered from skeletal remains found in mass graves, and reported a significant loss of DNA using spin columns. The authors advised that size separation using microcons is not advised when a sample contains limited DNA. On the other hand, MinElute® (Qiagen) purification columns removes negative ions (e.g. Cl^-) from the sample and prevents DNA competing with these negative ions for capillary electrophoresis injection, resulting in the maximum amount of DNA injected [43]. Furthermore, performing post-PCR purification will enable a sample to be concentrated by eluting in a smaller final volume. One study reported a fourfold increase in peak height of alleles when amplicons were purified using the MinElute® (Qiagen) purification columns [44].

Silica-based extraction methods and silica membrane columns are effective methods of reducing inhibitors [45]; however, they endure a loss of valuable DNA. Diluting the DNA extract is another method to overcome inhibition for samples with large DNA mass [46]. The dilution method is effective for processing mitochondrial DNA samples containing higher copy numbers (cf. to only two copies for autosomal nuclear DNA). However, due to the limited quantity of DNA available in a degraded sample, the method is not well suited for analysing low-template DNA [47].

Other prevention measures used to reduce PCR inhibition include adding extra Ampli*Taq* Gold® DNA polymerase to the reaction [9, 48]. If the enzyme is the target of the inhibitor then adding extra Ampli*Taq* Gold® DNA polymerase to the reaction will help compensate for the reduced enzyme activity [10]. However, adding too much Ampli*Taq* Gold® DNA polymerase may have a negative effect by increasing the likelihood of amplifying non-specific products [11]. Different enzymes exhibit different properties in their ability to tolerate inhibitors [6]. Inhibitor tolerant DNA polymerases [35, 49, 50] would be ideally suited to a direct PCR approach. A new reagent for processing buccal swabs using direct PCR has recently come to the market to help overcome inhibitors, known as Prep-n-Go™ Buffer (ABI), and can be trialled on other substrate types.

5.3.5 Substrates

In addition to the presence of PCR inhibitors, the impact of substrates has a profound effect on the quality of the resulting DNA profile [24].

Casework results

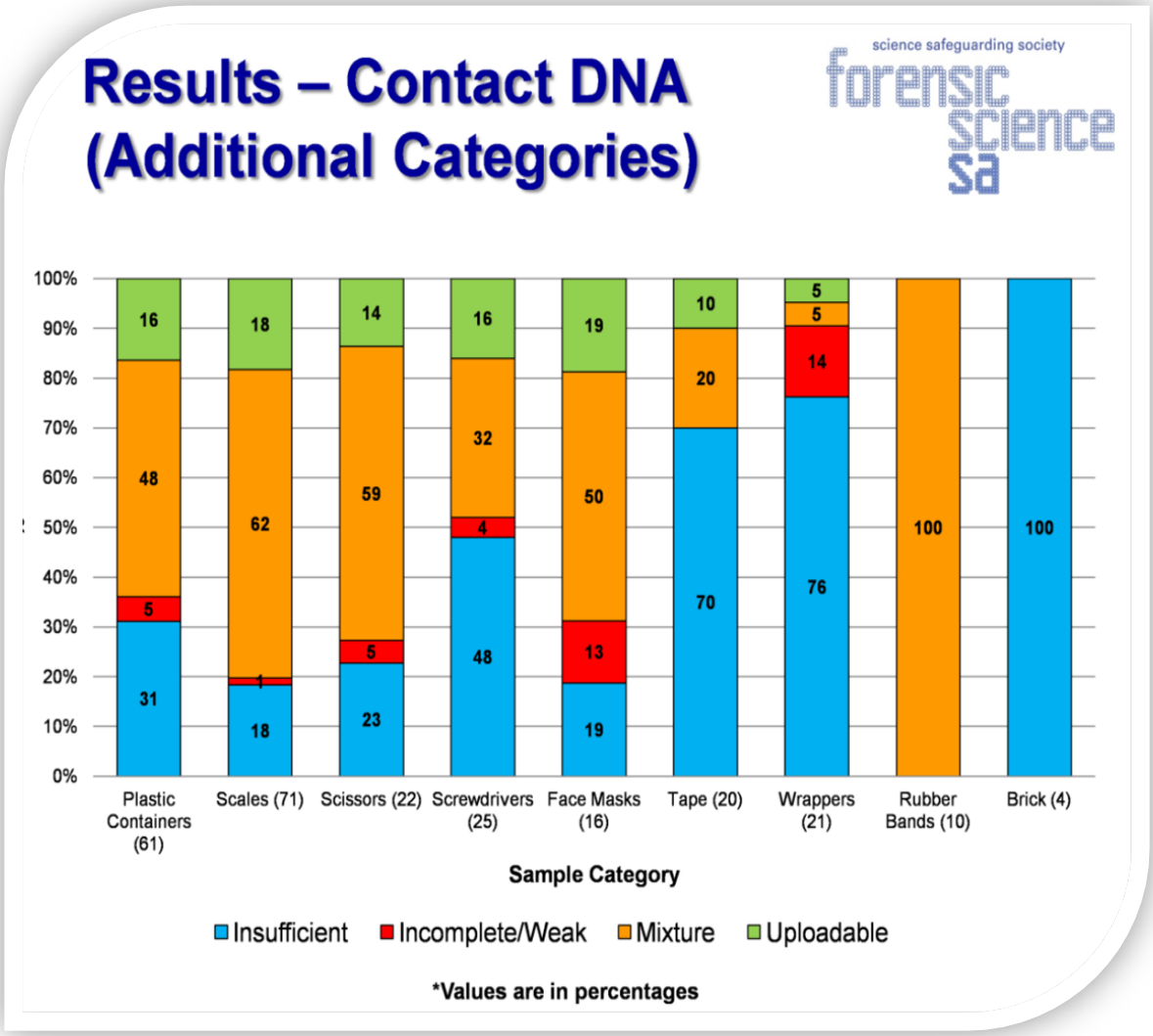


Figure 5.1. DNA profiling success rates for 'touch' DNA items submitted to FSSA in 2013, bar chart constructed at FSSA by Oliva Handt and Nicol Sly. A small number of DNA profiling results are considered 'up-loadable' to the Australian National Criminal Investigation DNA Database (NCIDD).

From the FSSA casework results it can be seen that one category of sample type (i.e. brick) consistently failed to yield STR DNA alleles (see Figure 5.1). All other categories (i.e. plastic, scales, scissors, screwdrivers, face masks, tape, wrappers and rubber bands) produced mixed DNA profiles (i.e. more than one contributor present); which is expected for the complex nature of 'touch' DNA items. Success rates ranged from 5% (for wrappers generating 'up-loadable' profiles) to 19% (for face masks producing 'up-loadable' profiles), (see Figure 5.1). Samples are considered 'up-loadable' when ≥ 12 alleles are assigned above the laboratory-

specific threshold limits. The results from this study are likely to be due to the differing amounts of DNA deposited onto the items rather than any inhibitory feature of the substrate. Results highlight the need for increased sensitivity low-template DNA analysis, in conjunction with software programs like TrueAllele® [51] and STRmix™ [52, 53] for mixture analysis.

Other research investigating the effect of substrate on the resulting DNA profile (i.e. number of detectable loci) recorded a higher yield of DNA from glass, followed by plastic and paper [54]. Metal substrates did not yield enough DNA to generate interpretable profiles. This could be due to the Cu, Zn and Ni-ions causing DNA degradation [55], or the binding of metals to DNA that inhibit amplification [56]. Similarly, other research reports that more DNA is recovered from glass substrates compared to metal [57]. The value of detecting latent prints and 'touch' DNA on the surface of ammunition is well recognised. FSSA reported that 100% of ammunition samples submitted for contact DNA analysis failed to yield measurable results following a standard DNA extraction (see Figure 5.2).

Casework results

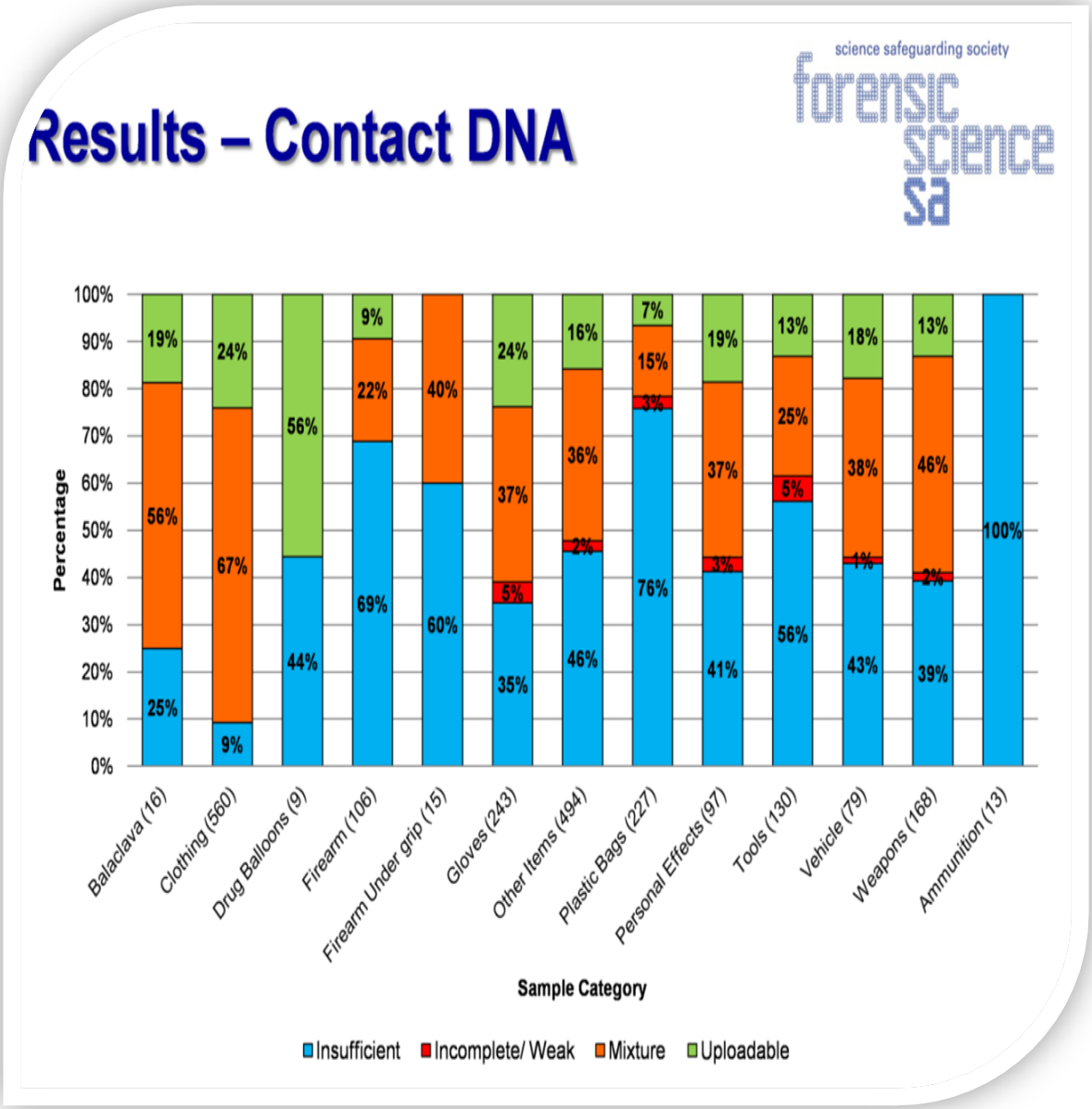


Figure 5.2. DNA profiling success for 'touch' DNA items submitted to FSSA in 2013. Image generated at FSSA by Oliva Handt and Nicol Sly.

The main goal of the research in this chapter is to profile DNA using direct PCR from powdered fingerprints thought to be potential PCR inhibitors. Supplementary data is included that explores the limitations of direct PCR when analysing a suite of forensic related material and varying environmental stresses. Further work highlights the benefits of adding PCR facilitators to the amplification vessel (i.e. Prep-n-Go™

Buffer (ABI) and Ampli*Taq* Gold® DNA polymerase), and measures the effect of post-PCR purification on the quality of DNA profiles.

5.4 Materials and methods – for preliminary work

Additional preliminary studies to the published data are detailed below.

All contamination measures and controls are indicated in Chapter II. Latent DNA on handled or touched evidentiary items and collection devices (e.g. swabs) were processed and analyzed in a similar manner to samples processed in a low-template/ancient DNA environment. Extreme caution and care was adhered to at all stages.

Controls

Positive PCR controls consisted of 1 ng of control DNA (2800M, Promega) or control DNA supplied with the amplification kit in use. Negative controls included a PCR blank (i.e. no DNA); and a swab of a slide containing no DNA (to monitor slide and swab contamination). Amplification followed using the AmpFℓSTR® ProfilerPlus™ or NGM SElect kit™ guidelines.

Capillary Electrophoresis and Data Analysis - See Chapter II

For the data analysis steps, a ‘full’ DNA profile was defined when all the expected alleles, in all loci, were observed. Full DNA profiles did not show ‘allele drop-out’ in any loci.

5.4.1 Effect of fingerprint powders on direct PCR

5.4.1.1 Materials and Methods

Fingerprint powders

The swab head of a sterile DNA-free nylon FLOQswab™ (Copan Industries, Italy) was immersed in one of the following 8 fingerprint powders (similar in amount to a fingerprint powdered swab):

- Black (HiFi Volcano silk black, Sirchie, NC, USA);
- White powder (Hadonite powder, ACE chemical company, SA, AU);
- Silver powder (Aluminium powder uncoated, Merck, Vic, AU);
- Magnetic black (Supranano black magnetic, ARRO SupraNano Ltd, UK);
- Supranano magnetic red fluorescent (Supranano red fluorescent, ARRO SupraNano Ltd, UK);
- Supranano red (Supranano red, ARRO SupraNano Ltd, UK);
- Supranano green fluorescent (Supranano green fluorescent, ARRO SupraNano Ltd, UK); and
- Supranano green magnetic fluorescent (Supranano green magnetic fluorescent, ARRO SupraNano Ltd, UK).

Figure 5.3 Showing fingerprint powders trialled for direct PCR (black, silver, white, fluorescent and magnetic powders), and swab fibres tarnished in powder prior to direct PCR.

After immersing each swab in one specific fingerprint powder, the tip of the swab was cut (i.e. 2 mm^2 fibre clump) using a sterile scalpel blade and added directly to a 0.2 mL PCR tube using sterile forceps. Control DNA (i.e. 1ng, 2800M Promega) was added to the PCR tube and amplification was carried out using the NGM SElect kit™ or AmpFℓSTR® ProfilerPlus™ kit guidelines.

5.4.1.2 Results and Discussion

As previously established in earlier data chapters, direct PCR is a proven effective method for processing 'touch' DNA swabs. The earlier work reported in this thesis demonstrates a significant improvement in DNA profiling results when direct PCR was employed over the traditional extraction methodology. In this study, fingerprint powders were added to the PCR tube in order to determine the inhibitory effect of powders on the resulting DNA profile. All eight powders tested permitted the interpretation of STR amplicons and did not affect the ability to generate full DNA profiles when control DNA was used as template for amplification. A comparison of reportable loci reports a full DNA profile obtained (i.e. 34 alleles out of 34 alleles) for all powders trialled (see SI Table 5.1 – 5.9) using the AmpFℓSTR® NGM SElect™ PCR kit. Additionally, peak heights of alleles in DNA profiles generated from fingerprint powdered samples were similar in height to control DNA samples without powders present. The average RFU values of peak heights ranged from 1366 RFU for magnetic red powder (see SI Table 5.5) to 3127 RFU for red powder (see SI Table 5.6). All RFU values of peak heights were well above the detection threshold limit of 50 RFU.

Data indicate no detrimental inhibition as a result of fingerprint powders being present in the PCR tube. Peak height ratios are an important consideration when analysing STR-based DNA profiles. Heterozygote loci must be well balanced. Hence all data was calculated for peak height ratios. The average heterozygote peak balance (i.e. lower peak height divided by higher peak height) was 88% for powdered samples. The average heterozygote peak balance for the control DNA sample (i.e. no powder present) was 89%.

This promising result led to a more in-depth study being carried out on fingermark samples (n=160) to mimic real life scenarios and to investigate the possible interference of powders on direct PCR amplification ***(see Manuscript accepted in FSI:G, May 2017: Typing DNA profiles from previously enhanced fingerprints using direct PCR).*** In the manuscript enclosed DNA profiling was carried out on previously enhanced fingerprints (i.e. fingermarks dusted immediately before sampling). DNA was left behind on plastic substrates 15 minutes after individuals

had washed their hands. To establish the ability to perform both techniques simultaneously and highlight eventual problems, fingermarks were dusted with four commonly used dactyloscopic powders (i.e. black, silver aluminium, white and magnetic black) and swabbed immediately prior to direct PCR; fingermarks were exposed to powders for a short time prior to swabbing (i.e. 10 – 15 minutes)).

5.4.1.2.1 Statement of authorship

Manuscript: Typing DNA profiles from previously enhanced fingerprints using direct PCR.

Jennifer Templeton (Candidate)

Designed the experiment, performed all the laboratory work, data analysis, and wrote the paper

I hereby certify that the statement of contribution is accurate.

Signed:  Date: November 2016

Duncan Taylor (Supervisor)

Commented on data, and helped to edit the paper

I hereby certify that the statement of contribution is accurate.

Signed:  Date: November 2016

Oliva Handt (Supervisor)

Commented on data, and helped to edit the paper

I hereby certify that the statement of contribution is accurate.

Signed:  Date: November 2016

Adrian Linacre (Primary supervisor)

Supervised the project, assisted with experimental design, commented on data, and edited the paper

I hereby certify that the statement of contribution is accurate.

Signed: Date: November 2016

5.4.1.2.2 Manuscript: Typing DNA profiles from previously enhanced fingerprints using direct PCR

Forensic Science International: Genetics 29 (2017) 276–282

Contents lists available at ScienceDirect

Forensic Science International: Genetics

journal homepage: www.elsevier.com/locate/fsig

Research paper

Typing DNA profiles from previously enhanced fingerprints using direct PCR

Jennifer E.L. Templeton[a,*], Duncan Taylor[a,b], Oliva Handt[a,b], Adrian Linacre[a]

[a] School of Biological Sciences, Flinders University, Bedford Park, Adelaide, South Australia, 5042, Australia
[b] Forensic Science South Australia, Adelaide, South Australia, 5000, Australia

ARTICLE INFO

Article history:
Received 12 October 2016
Received in revised form 27 March 2017
Accepted 19 May 2017
Available online 21 May 2017

Keywords:
'Touch' DNA
Short tandem repeat
Human identification
Trace DNA
Fingermarks
Fingerprint powder

ABSTRACT

Fingermarks are a source of human identification both through the ridge patterns and DNA profiling. Typing nuclear STR DNA markers from previously enhanced fingermarks provides an alternative method of utilising the limited fingermark deposit that can be left behind during a criminal act. Dusting with fingerprint powders is a standard method used in classical fingermark enhancement and can affect DNA data. The ability to generate informative DNA profiles from powdered fingerprints using direct PCR swabs was investigated. Direct PCR was used as the opportunity to generate usable DNA profiles after performing any of the standard DNA extraction processes is minimal. Omitting the extraction step will, for many samples, be the key to success if there is limited sample DNA. DNA profiles were generated by direct PCR from 160 fingermarks after treatment with one of the following dactyloscopic fingerprint powders: white hadonite; silver aluminium; HiFi Volcano silk black; or black magnetic fingerprint powder. This was achieved by a combination of an optimised double-swabbing technique and swab media, omission of the extraction step to minimise loss of critical low-template DNA, and additional AmpliTaq Gold® DNA polymerase to boost the PCR. Ninety eight out of 160 samples (61%) were considered 'up-loadable' to the Australian National Criminal Investigation DNA Database (NCIDD). The method described required a minimum of working steps, equipment and reagents, and was completed within 4 h. Direct PCR allows the generation of DNA profiles from enhanced prints without the need to increase PCR cycle numbers beyond manufacturer's recommendations. Particular emphasis was placed on preventing contamination by applying strict protocols and avoiding the use of previously used fingerprint brushes. Based on this extensive survey, the data provided indicate minimal effects of any of these four powders on the chance of obtaining DNA profiles from enhanced fingermarks.

1. Introduction

DNA profiling and fingerprint analysis both have significant discriminatory power for the purpose of human identification and are an integral part of a forensic laboratory's workflow. Classical fingerprinting techniques (e.g. chemical, electronic or physical) are used to match a fingerprint from a crime scene or casework exhibit to the fingerprint of the donor [1,2]. Depending on the nature of the substrate examined (e.g. porous or non-porous), the powdering technique is a common method used to physically develop a fingerprint on a non-absorbent surface [1,3]. Dactyloscopic powders are traditionally used by forensic investigators to visualise and characterize the ridges and minutiae of a latent fingerprint left behind on a substrate. Powders adhere to the moisture and oil secretion that are deposited by skin. Moisture in aged prints evaporates leaving behind a viscous print that is difficult to develop using powders [3]. The same is true for fingerprints exposed to warm climates [3]. Other methods of enhancement may be considered; cyanoacrylate fuming can be used on non-porous surface types [4,5] and DFO/ninhydrin reagents are typically used on porous surfaces [6–8]. More recently intercalating dyes such as SYBR® Green I and GelGreen™ have been considered for the detection of DNA within latent fingermarks [9,10]. Overall, dusting with powders is a successful method to enhance fingerprints as it is quick, easy to use, relatively inexpensive, and provides immediate results [11]. The shape and size of the powder particles will determine how well the powder adheres to the fingermark. Small, fine particles are more commonly used by scene investigators as they will adhere more easily to the fingermark compared to large coarse powders [3].

* Corresponding author.
E-mail address: jennifer.templeton@flinders.edu.au (J.E.L. Templeton).

http://dx.doi.org/10.1016/j.fsigen.2017.05.006
1872-4973/

Powder selection is also dependent on the colour contrast of the surface being investigated. Silver, aluminium and white powders are commonly applied to dark coloured surface areas, whereas black carbon powder is routinely used to highlight a fingerprint on a lighter coloured surface. In cases where traditional fingermark enhancement methods do not result in a useable profile, such as the case of a partial, smudged or distorted fingermark, investigators can resort to sampling the 'touch' DNA content of the fingermark shed by the individual who made it.

Since the first DNA profile was generated from a touched object in 1997 [12] further studies have supplemented this finding [2,11,13–15]. Despite this significant advancement in the field, many operational forensic laboratories report low success rates from 'touch/contact' DNA samples in casework exhibits [16]. Swabs taken from touched or handled items are increasingly being submitted to forensic police laboratories for analysis, therefore current methodology should be evaluated to improve the success rates. Research conducted on latent DNA samples has proven difficult, as any preserved DNA is often limited in quantity [16] and in various states of degradation [17]. Most commercial DNA extraction kits are designed to deal with samples containing intact cells and high molecular weight DNA [18] that will yield optimum template DNA (e.g. 1 ng) for PCR; atypical of 'touch' or trace DNA sources [19]. Previous studies have reported a loss in DNA following a standard extraction [20,21] as there is the potential to discard endogenous DNA within the sample during the multiple wash steps and tube changes that follow. Cell-free DNA is known to be a valuable source of template DNA [22] and a direct PCR approach would take advantage of this component of 'touch' DNA by making all template DNA available to the reaction. On the contrary, a standard DNA extraction involves discarding the supernatant; the aqueous portion of a DNA extract may contain valuable sample DNA.

Detergents and reducing agents have been recommended when processing fresh samples where intact proteins and/or cells are present [18]. Extraction and purification is not fundamental to all sample types as demonstrated previously [23–28]. Samples pre-treated with specific chemicals may not adversely affect the ability to generate a DNA profile [13,29–32]. A previous research study reports that DNA typing of saliva and blood stains is possible following pre-treatment with cyanoacrylate, gentian violet, ninhydrin or amorphous carbon [33]. Additionally, it has proven possible to recover DNA from fingerprints after archiving [13]. Some reagents and powders used to enhance fingerprints have been shown to prevent successful DNA amplification [34,35]. In particular, magnetic fingerprint powders used in conjunction with magnetic bead-based DNA extraction protocols – now commonplace in many forensic laboratories (e.g. DNA IQ™ from the Promega Corporation) – have been shown to fail [29]. The potential interference of these magnetic powders with the magnetic bead-based chemistry may adversely affect the ability to recover DNA. Many samples of 'touch' DNA content contain limited DNA [16,19,36,37], so it is crucial to maximise the amount of template DNA available for subsequent PCR and avoid loss during extraction.

DNA retrieved from latent fingermark residue is known to be a source of template DNA using direct PCR [38]. In this previous study we demonstrate the ability to optimise the amount of DNA template available for subsequent enzymatic manipulation without the need for extraction and purification. In comparison to other published work, we found direct PCR to yield the highest values of amplifiable DNA when tested on a single fingermark deposit at standard PCR cycling conditions [38]. This is not to claim that for other sample types other protocols may not generate higher quality profiles (e.g. fingermarks combined with other body fluids or if certain inhibitors are present). However, given the superior performance of direct PCR on latent DNA, it can be seen as a promising tool to amplify 'touch' DNA fingermarks that have been previously enhanced with dactyloscopic powders. Direct PCR also offers a benefit by reducing contamination (i.e. fewer tube changes and manual handling of samples), reducing labour time and cost (as there is no need to carry out a DNA extraction), and generating a higher yield of PCR product.

Swabbing large surface areas for the presence of 'touch' DNA may not generate informative results without a more targeted approach being used to indicate where DNA is present prior to swabbing (i.e. visualisation of fingerprints). In this study, the method has been further developed to maximise the recovery of PCR-amplifiable DNA from enhanced fingermarks and to minimise carry-over of fingerprint powder to the 3130*xl* capillary electrophoresis machine prior to analysis. Results verify that DNA profiling can be used in conjunction with the classical powdering technique to provide additional evidence if required. The main goal of this research was to investigate the effect of various dactyloscopic powders on subsequent DNA profiles following traditional fingermark enhancement and direct PCR ($n = 160$). Four commonly used fingerprint powders: white (Hadonite powder, ACE Chemical Company, SA, Australia); silver (Aluminium powder uncoated, Merck, Vic, Australia); black (HiFi Volcano silk black, Sirchie, NC, USA) and; black magnetic powder (Supranano black magnetic, ARRO SupraNano Ltd, UK) were assessed in order to determine which powder (if any) was more likely to yield DNA profiles.

2. Materials and methods

2.1. DNA reference swab collection

Cotton buccal swabs (Copan Industries, Italy) taken from the inner lining of the cheeks of 10 volunteers were collected. Ethical approval was obtained from Southern Adelaide Clinical Human Research Ethics Committee prior to starting, and informed consent was obtained from individuals taking part in the study. DNA was extracted following the DNA IQ™ System (Promega Corporation, VIC, Australia) manufacturer's recommendations.

2.2. Authentication criteria

As is generally the case when working with 'touch' or low-template DNA samples, a number of precautions have to be taken during sampling and processing steps; these have been extensively described elsewhere [38]. Due to the high risk of cross-contamination when dealing with low-template DNA samples, numerous precautions were taken. (i) The samples were stored securely in separate sterile Petri dishes in a fridge (i.e. 4 °C) dedicated to pre-PCR activity and processed the following day. (ii) All preparatory steps prior to DNA amplification were performed in a clean-room area solely dedicated to pre-PCR work and physically isolated from post-PCR activity. (iii) Amplification and typing of PCR products were carried out in a post-PCR laboratory. (iv) No more than three individuals were sampled per day to limit any potential cross-contamination. If batch sampling is required for case work, adequate controls would be in place to ensure all tools and lab equipment are cleaned and that gloves are changed between sampling. (v) Contamination was further minimised by adopting strict protocols. The researcher involved in the handling of the samples was required to wear a lab coat (i.e. designed for pre-PCR work), facemask and gloves that were changed regularly. A separate disposable fingerprint brush was used for each application to avoid cross-contamination. Sterile, disposable fingerprint brushes were subjected to UV irradiation for 15 min prior to use. All tools, benches, consumables, disposables and instruments used for fingerprint sampling were cleaned with bleach (3%), isopropanol (70%) and consumables were irradiated with ultraviolet (UV) light

for 15 min prior to use. (vi) Contamination was monitored by including non-template controls in parallel, reagent blanks (i.e. PCR negative controls), swab fibres, swabbing of surfaces without fingermarks present, and cross-comparison of all profiles obtained to the profiles of staff members. All alleles in all loci generated were compared to previous genotypes to assess the possibility of cross-contamination events and to monitor potential allele drop-in from previous samples and runs. For future processing of case work samples, a recommendation follows that hairnets should be applied to minimise contamination. Samples should be stored at −20 °C upon air-drying if long term storage is required.

2.3. Measuring the effect of fingerprint powders on direct PCR and DNA profiling

The swab head of a sterile DNA-free nylon FLOQswab™ (Copan Industries, Italy) was immersed in an aliquot containing one of following fingerprint powders: white (Hadonite powder, ACE chemical company, SA, Australia); silver (Aluminium powder uncoated, Merck, VIC, Australia); black (HiFi Volcano silk black, Sirchie, NC, USA) and; black magnetic powder (Supranano black magnetic, ARRO SupraNano Ltd, UK). The tip of the swab was cut (i.e. 2 mm^2 fibre clump) using a sterile scalpel blade and added directly to a 0.2 mL PCR tube using sterile forceps. Control DNA (1 ng, 2800 M Promega) was added to the PCR tube and amplification was carried out using the AmpFℓSTR® NGM SElect™ PCR amplification kit (Life Technologies, VIC, Australia) following manufacturer's recommendations. Positive control DNA was set up with no fingerprint powders present in the PCR tube. Cycling was performed on a 9700 GeneAmp thermal cycler (Life Technologies) and consisted of an initial denaturation at 95 °C for 11 min followed by 30 cycles of 94 °C for 20 s, 59 °C for 3 min, then a final extension at 60 °C for 10 min, and 4 °C for indefinite. PCR products were spun in a microcentrifuge for 20 s at maximum speed to centrifuge the powders to the bottom of the tube. PCR product was aliquoted with care (to avoid powder carry over). PCR products were analyzed on a 3130*xl* Genetic Analyzer (Life Technologies) in a 11.5 μL final volume that consisted of 1.5 μL PCR product, 9.5 μL HiDi™ Formamide (Life Technologies) and 0.5 μL GeneScan™ 600 LIZ™ Size Standard v2.0 (Life Technologies).

2.4. DNA deposition – 15 min post hand washing

Ten volunteers (comprised of 5 males and 5 females) washed their hands prior to depositing DNA. To determine the sensitivity of the technique, a short time period was chosen between hand washing and fingermark deposition. All donors were asked to continue with daily activities but to refrain from washing their hands within the 15 min time frame. Donor activities were recorded at the time of sampling and included: watching TV; reading; writing; touching mobile phones and personal items; and typing at keyboards. Fifteen minutes post hand washing each volunteer deposited a fingermark using the first four fingers of their dominant hand onto a sterile and DNA-free plastic slide (Rinzl plastic, ProSci Tech). All samples taken from the same individual were processed on the same day. Contact was for 15 s using medium pressure (to ensure consistency) such that a fingermark was created. Testing for each volunteer was repeated four times to allow four different powders to be tested for each individual. All slides were stored in separate sterile Petri dishes with lids secure when not in use. Volunteers that processed samples away from the laboratory stored the Petri dishes at 4 °C until they were transported back to the laboratory on ice. All slides were stored at 4 °C prior to processing.

2.5. Dusting prints – fingerprint powders

Slides were taken from the fridge the following day (~16 h) after DNA deposition took place and allowed to air-dry in a sterilised fume hood prior to dusting the slides with fingerprint powders. Prints were clearly visible post-dusting and this enabled a targeted swabbing approach to be carried out that focused on swabbing the area where the fingerprint was identified. Fingermarks were lightly dusted using each of the fingerprint powders mentioned above and fibre glass disposable brushes (DNA free fiberglass brush, Sirchie, ACT, Australia) prior to swabbing for DNA. A sterile DNA-free separate disposable fingerprint brush was used for each application to avoid cross-contamination. In total, 16 tests were carried out for each individual (i.e. four fingermarks created for each powder tested). Swabbing of the prints took place immediately after dusting.

2.6. Swabbing and direct amplification – post-fingerprint enhancement

All donor samples (i.e. 16 swabs) were processed at the same time. A small portion of fibres (i.e. ~2 mm^2) was cut from the tip of a sterile nylon FLOQswab™ (Copan) using a new sterile scalpel blade. Fibres were pre-moistened with 2 μL 0.1% Triton™ X-100 (Sigma, VIC, Australia) (i.e. aliquot of solution pre-heated to 50 °C) prior to swabbing. Sterile forceps were used to gather the 2 mm^2 portion of swab and apply pressure over the mark in a targeted swabbing approach, as previously demonstrated [38]. Glue fragments of the swab held the fibres neatly in place. Swabbing consisted of two sets of 10 strokes horizontally (i.e. left to right) and two sets of 8 strokes vertically (i.e. top to bottom) over the mark using medium pressure to ensure consistency. Sterile forceps were used to place swab fibres (i.e. ~2 mm^2) into the 0.2 mL PCR tube and the PCR tube was closed until the second swabbing event took place. The procedure was repeated using a second moistened swab (i.e. 2 μL 0.1% Triton™ X-100 (Sigma, VIC, Australia) pre-heated to 50 °C). Preliminary data (unpublished) in our laboratory indicate that a second moist swab is more effective than using a second dry swab. Swab fibres (i.e. ~2 mm^2) from the second swabbing event were placed into the same 0.2 mL PCR tube as the first swabbing event and the PCR tube sealed prior to PCR. STR typing was performed using the AmpFℓSTR® NGM SElect™ PCR amplification kit (Life Technologies, VIC, Australia) following manufacturer's recommendations, but with additional Ampli*Taq* Gold® 360 DNA polymerase 1 μL (5 U). The final reaction consisted of 10 μL NGM SElect™ reaction mix, 5 μL of NGM™ primer mix, 1 μL (5 U) Ampli*Taq* Gold® 360 (Life Technologies), DNA template (i.e. 4 mm^2 swabbed fibres for samples or 1 ng of eluted DNA extract from buccal swab), the remainder of the solution was made up to a final volume of 25 μL with TE buffer (i.e. 10 mM Tris, 0.1 mM EDTA, pH 8.0). Cycling was performed on a 9700 GeneAmp thermal cycler (Life Technologies) and consisted of an initial denaturation at 95 °C for 11 min followed by 30 cycles of 94 °C for 20 s, 59 °C for 3 min, then a final extension at 60 °C for 10 min, and a 4 °C hold. PCR products were spun in a microcentrifuge for 20 s at maximum speed to centrifuge the powders to the bottom of the tube. PCR product was aliquoted with care (to avoid powder carry over). PCR products were analyzed on a 3130*xl* Genetic Analyzer (Life Technologies) in a 11.5 μL final volume that consisted of 1.5 μL PCR product, 9.5 μL HiDi™ Formamide (Life Technologies) and 0.5 μL GeneScan™ 600 LIZ® Size Standard v2.0 (Life Technologies). Preliminary data (unpublished) in our laboratory indicate a higher yield of PCR product with good quality profiles being generated when additional PCR product (i.e. 1.5 μL) is added to the CE run.

2.7. Data analysis

All DNA profiles were analyzed using GeneMapper ID software (v3.2.1) (Life Technologies) and a peak detection threshold of 50 relative fluorescence units (RFU) was used to assign alleles. A wildcard designation was used for potential homozygotes with peak heights <150 RFU to account for potential allelic 'drop-out' (e.g. "11, F" instead of "11, 11"). A profile was considered 'full' when all alleles, of all loci, were detected above the threshold RFU. Allelic 'drop-out' was reported as alleles with peak heights <50 RFU. Additional alleles were reported if associated peak heights were >50 RFU. Previous studies report DNA profiling success differently and all laboratories have different thresholds for genotyping. To eliminate this inconsistency we report our typing success based on previous publications (i.e. the number of detected alleles that match the donor) [38,39]. An informative DNA profile (i.e. one considered 'up-loadable' to the Australian National Criminal Investigation DNA Database (NCIDD) was defined as ≥12 alleles (plus Amelogenin). The profiling success (%) was measured by dividing the number of alleles successfully called by the total number of expected alleles and multiplying this value by 100.

3. Results and discussion

Results were carried out to determine the inhibitory effect of powders on the resulting DNA profile using control DNA. All powders tested permitted the interpretation of STR amplicons and did not prevent the ability to generate full DNA profiles when control DNA was used as a template for amplification. A comparison of reportable loci reports a full DNA profile obtained (i.e. 34 alleles out of 34 alleles) for all powders trialled using the AmpF*l*STR® NGM SElect™ PCR kit at standard cycles. Stochastic effects (i.e. peak height imbalance, allele/loci 'drop-out', split peaks and increased baseline noise) are important considerations when analyzing STR-based DNA profiles, and were not over-represented in the resulting data. The average heterozygote peak balance (i.e. lower peak height divided by higher peak height) was 88% for powdered samples. The average heterozygote peak balance for the control DNA sample (i.e. no powder present) was 89%; indicating little to no difference in heterozygote imbalance between powdered and un-powdered samples. Based on inter-locus balance and general profile peak intensities, none of the powders tested appeared to have a deleterious effect on PCR amplification. These results progressed into the second phase of the study.

Results were generated from 160 dusted fingerprints created by 10 individuals, without the need for increased PCR cycle number beyond manufacturer's recommendations. Qualitative and quantitative analysis of DNA recovered from fingerprints already demonstrates the ability to obtain profiles from enhanced fingermarks [8,40,41]. Published data indicate a significant increase in PCR yield when direct PCR was used to process swab fibres in comparison to extracted DNA swabs [42,43]. Direct PCR has since been applied to the recovery of latent DNA from fingermarks [38] and the main aim of the current work was to investigate the interference of powders on the direct PCR approach. Data indicate that DNA accumulates on the surface of the skin as time progress after hand washing [44]. We demonstrate the sensitivity of the direct PCR approach in its ability to pick up DNA after such a short time has passed since hand washing (i.e. 15 min), see Fig. 1. Additional electropherogram data generated by direct PCR are included in the supplementary material.

DNA profiles obtained were compared to the profiles of the donors and evaluated for the matching number of alleles and complete loci (i.e. measuring success rate). The average heterozygote peak balance for the fingerprint dusted samples, using the direct PCR approach, was 76%. Statistical analysis was carried out using One-way ANOVA in Excel. A p-value of 0.49 (i.e. >0.05) indicated that there was no significant difference in the mean values of the DNA data generated from the four powders tested. Based on inter-locus balance and general profile peak intensities, none of the powders tested exhibited a deleterious effect on PCR amplification. Ninety eight out of 160 samples (61%) were considered 'up-loadable' to the Australian National Criminal Investigation DNA Database (NCIDD), see Table 1. Results achieved in this study reveal a higher success rate compared with other published findings; one study carrying out similar work with dactyloscopic powdered fingerprints and extracted DNA swabs (QIAamp DNA Micro kit, Qiagen) reveal a lower DNA success rate of 55% despite enhancing the profile by LCN [45].

Only one sample displayed a mixed DNA profile where the donor was not the major contributor, see Fig. 2; a further two digits tested at the same time period for this individual displayed a mixed DNA profile consisting of more than one contributor (i.e. more than 2 alleles detected at one locus) with no clear major/minor contributor. This individual reported entering a communal area and handling a door handle used by many individuals, and this activity may have contributed to the detection of additional alleles.

Samples subjected to the silver fingerprint dusting powder prior to swabbing had the highest success rate for direct PCR and subsequent DNA profiling; 30 out 40 samples tested (75%) yielded 'up-loadable' DNA profiles. White fingerprint powder was the second best performing powder; 26 'up-loadable' profiles were generated out of 40 samples tested (65%). Black magnetic powder was the third best performing powder producing 23 'up-loadable' profiles out of 40 samples tested (57%). Black powder had the lowest success rate out of all the four powders tested producing 19 'up-loadable' profiles out of 40 samples tested (47%).

A full DNA profile, comprising of all 16 STR loci plus the amelogenin marker, was obtained from 23 samples (profile example given in Fig. 1). The presence of additional alleles (i.e. mixed profiles) and the incidence of allele 'drop-out' and discordant genotypes were examined, see Fig. 2. Thirteen out of the 160 samples (8%) failed to yield any alleles. Mixed profiles were observed in 49 out of the 160 samples (31% of cases), and 98 samples out of the 160 (61%) produced what appeared to be single source results (i.e. no more than 2 visible peaks per loci), see Fig. 2. Locus 'drop-out' was observed mostly at larger sized amplicons, as was expected.

Additional alleles noted as mixed profiles could be a result of exogenous DNA present on the substrate that was touched, or from the fingerprint brushes/swabs/consumables used for testing, or as a result of secondary/tertiary transfer. Negative controls of swabs and swabbed substrates prior to fingerprint deposition did not yield any DNA. For the mixtures observed, ≤2 minor alleles were detected in the majority of cases (i.e. in 28 out of 49 samples). In samples where more than two additional alleles were recorded, other digits from the same individual tested displayed similar findings indicating the likelihood of secondary/tertiary DNA transfer. Strict contamination procedures were adhered to in order to prevent laboratory associated allele 'drop-in' (see authentication criteria) and we anticipate additional alleles are more likely a result of secondary or tertiary transfer, similar to other studies [46–49]. From observation, individuals that undertook keyboard work or handled personal items were more inclined to leave behind amplifiable DNA. It can be assumed from previous research that a higher yield of amplifiable DNA from fingermarks can be attributed to the phenomenon known as 'loading' the fingers with DNA (e.g. individuals transferring DNA to their hands) [50]. To mimic real life scenarios, donors did not refrain from activities such as touching objects handled by other individuals (e.g. door handles) before depositing fingerprints and the

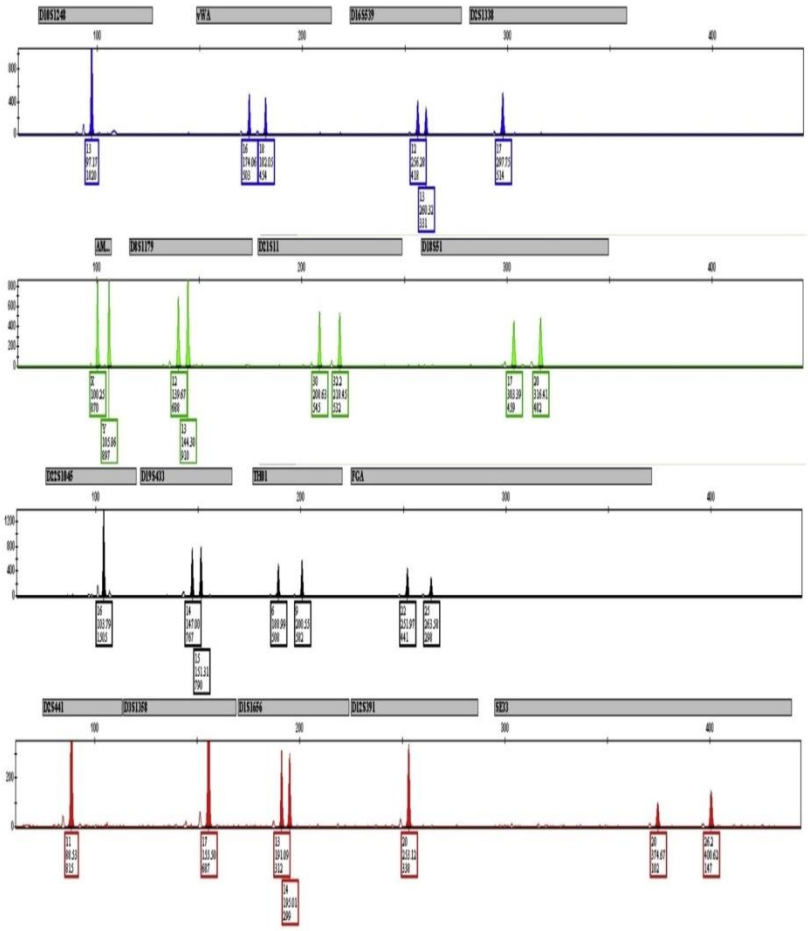


Fig. 1. Autosomal STR-based DNA profile generated from an enhanced fingerprint using nylon FLOQswab™ fibres and heated Triton™ X-100 solution. Direct amplification was carried out using NGM SElect™ kit at 30 PCR cycles. The full DNA profile displayed (i.e. 34 out of 34 alleles) was obtained from a dusted fingerprint deposited 15 min after the individual had washed their hands.

likelihood of generating additional alleles by secondary/tertiary transfer is therefore likely.

Classical fingerprint enhancement and DNA profiling can both be performed on the same sample if one of the aforementioned powders examined in this study is used and with proper measures in place to prevent contamination. It is important that fingerprint examination takes place first, as swabbing an area where DNA is present will result in distortion of the mark and destruction of the evidence. Cross-contamination can easily occur if a fingerprint brush has been used multiple times. A previous study by Montpetit et al. [39] reported that 31% of ammunition samples tested for DNA profiling produced mixed profiles even though only one known individual had handled the cartridges prior to swabbing. It is possible that individuals deposit not only their own DNA upon touching an object but also a second or third person's DNA (i.e. 'secondary/tertiary' transfer) if they have been in close contact with them previously [47,51]. This is one possible explanation for the occurrence of multiple DNA types on items that have been handled by a single individual. If proper contamination controls are in place during sample collection and processing then there can be greater confidence that any additional DNA detected is the result of secondary or tertiary transfer events.

If samples are subjected to classical fingerprint techniques prior to DNA typing then additional precautions should be taken to

Table 1
STR-based DNA profiling results for 160 NGM SElect™ directly amplified fingerprint samples. Results are generated from enhanced fingerprints post swabbing and direct PCR at 30 cycles.

Category	White powdered prints	Silver powdered prints	Black powdered prints	Black magnetic powdered prints	All powdered prints
Samples failed (i.e. 0 alleles)	4 (10%)	0 (0%)	4 (10%)	5 (12%)	13 (8%)
Donor not identified (i.e. 1–11 alleles)	10 (25%)	10 (25%)	17 (42%)	12 (30%)	49 (30%)
Samples 'up-loadable' to NCIDD (i.e. ≥12 alleles, plus amelogenin)	26 (65%)	30 (75%)	19 (47%)	23 (57%)	98 (61%)
Total	40	40	40	40	160

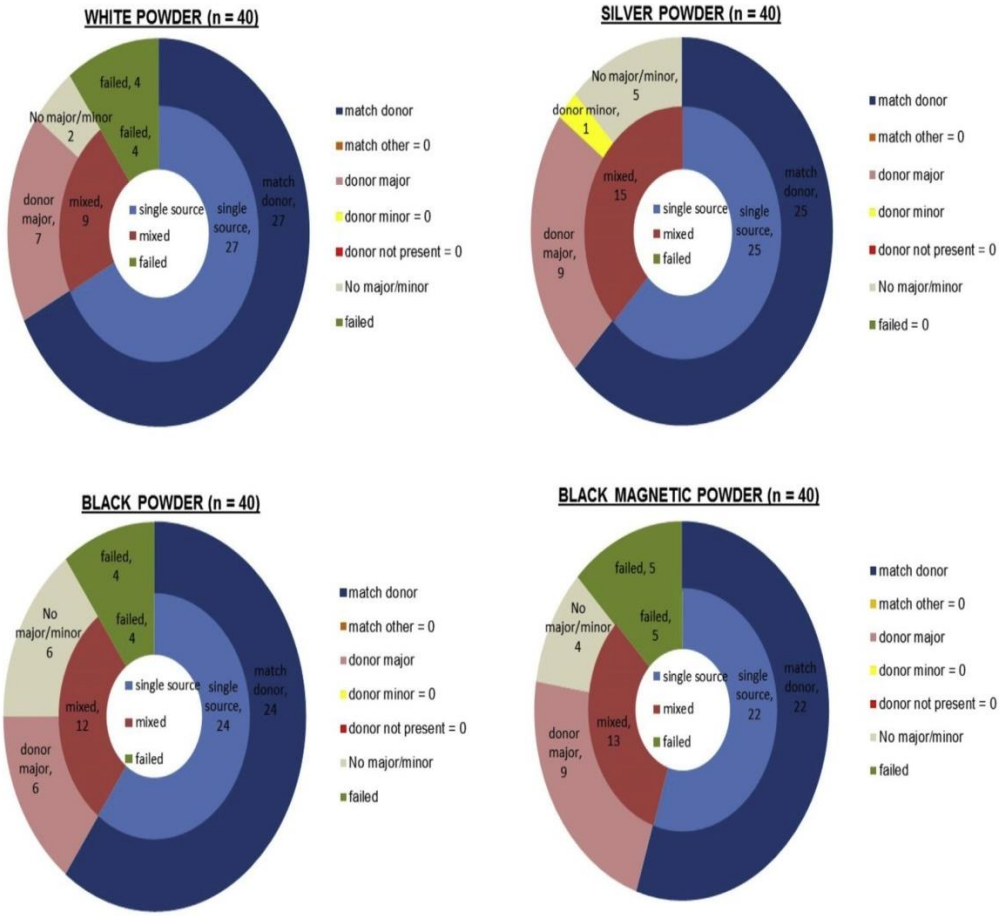


Fig. 2. STR-based DNA profiling results for 160 NGM SElect™ directly amplified fingerprint samples at 30 PCR cycles. Results display the proportion of single source and mixed DNA profiles.

prevent cross-contamination of DNA between samples. The use of disposable gloves, separate fingerprint brushes and face masks between applications will minimise this problem. An alternative solution is to apply fingerprint powder with an electrostatic depositor, atomizer or aerosol spray instead of using a fingerprint brush [52].

4. Conclusion

A study was conducted to investigate the effect of dactyloscopic powders on the direct PCR approach of sampled fingermarks. The results from this study assess the ability of direct PCR to amplify the DNA from fingerprints after dactyloscopic enhancement. A clear advantage of direct PCR is its broad applicability to 'touch' DNA samples – both latent and enhanced – and the relatively short time necessary to generate the profiles. A limitation lies in the fact that DNA quantification of the sample cannot take place and that there is no opportunity to remove potential PCR inhibitors. Sufficient DNA can be isolated from enhanced fingermark deposits using direct PCR provided precautions are taken to reduce cross-contamination and minimise the loss of DNA. The majority of samples (i.e. 61%) were considered 'up-loadable' to the NCIDD. In this study, DNA typing of enhanced fingermarks displayed similar profiling success as when using direct PCR on fingermarks without enhancement [38]. We conclude that the dactyloscopic powders tested in this study do not affect the success rate for producing profiles using direct PCR, or the profile quality.

Conflicts of interest

None.

Acknowledgments

Funding was provided by the office of the Attorney General South Australia and a John Stocker Postgraduate Scholarship from the Science and Industry Endowment Fund (SIEF). We would like to thank all volunteers that participated in the study and Sturt Police station in Adelaide for providing fingerprint powders. Additional thanks goes to Brendan R Chapman and Lauren Rask-Nielsen for providing helpful input and advice.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.fsigen.2017.05.006.

References

[1] A. Knowles, Aspects of physicochemical methods for the detection of latent fingerprints, J. Phys. E: Sci. Instrum. 11 (8) (1978) 713.
[2] J. Sewell, et al., Recovery of DNA and fingerprints from touched documents, Forensic Sci. Int. Genet. 2 (4) (2008) 281–285.
[3] G. Sodhi, J. Kaur, Powder method for detecting latent fingerprints: a review, Forensic Sci. Int. 120 (3) (2001) 172–176.

[4] B.J. Reedy, et al., The detection and enhancement of latent fingermarks using infrared chemical imaging, J. Forensic Sci. 50 (1) (2005) JFS2004213–JFS2004219.
[5] C. McLaren, C. Lennard, M. Stoilovic, Methylamine pretreatment of dry latent fingermarks on polyethylene for enhanced detection by cyanoacrylate fuming, J. Forensic Identif. 60 (2) (2010) 199.
[6] R. Jelly, et al., The detection of latent fingermarks on porous surfaces using amino acid sensitive reagents: a review, Anal. Chim. Acta 652 (1) (2009) 128–142.
[7] C. Lennard, et al., Synthesis and evaluation of ninhydrin analogues as reagents for the development of latent fingerprints on paper surfaces, J. Forensic Sci. Soc. 28 (1) (1988) 3–23.
[8] M. Schulz, et al., Ninhydrin-dyed latent fingerprints as a DNA source in a murder case, J. Clin. Forensic Med. 11 (4) (2004) 202–204.
[9] A.M. Haines, et al., Finding DNA: using fluorescent in situ detection, Forensic Sci. Int. Genet. Suppl. Ser. 5 (2015) e501–e502.
[10] A.M. Haines, et al., Detection of DNA within fingermarks, Forensic Sci. Int. Genet. Suppl. Ser. 4 (1) (2013) e290–e291.
[11] D.E. van Hoofstat, et al., DNA typing of fingerprints using capillary electrophoresis: effect of dactyloscopic powders, Electrophoresis 20 (14) (1999) 2870–2876.
[12] R.A. van Oorschot, M.K. Jones, DNA fingerprints from fingerprints, Nat. Lond. (1997) 767.
[13] M. Schulz, W. Reichert, Archived or directly swabbed latent fingerprints as a DNA source for STR typing, Forensic Sci. Int. 127 (1) (2002) 128–130.
[14] M.K. Balogh, et al., Fingerprints from fingerprints, Int. Congr. Ser. 1239 (0) (2003) 953–957.
[15] M.K. Balogh, et al., STR genotyping and mtDNA sequencing of latent fingerprint on paper, Forensic Sci. Int. 137 (2–3) (2003) 188–195.
[16] S. Harbison, M. Fallow, D. Bushell, An analysis of the success rate of 908 trace DNA samples submitted to the Crime Sample Database Unit in New Zealand, Aust. J. Forensic Sci. 40 (1) (2008) 49–53.
[17] T. Kita, et al., Morphological study of fragmented DNA on touched objects, Forensic Sci. Int. Genet. 3 (1) (2008) 32–36.
[18] N. Rohland, M. Hofreiter, Ancient DNA extraction from bones and teeth, Nat. Protoc. 2 (7) (2007) 1756–1762.
[19] R.A. van Oorschot, K.N. Ballantyne, R.J. Mitchell, Forensic trace DNA: a review, Investig. Genet. 1 (1) (2010) 1.
[20] R.A.H. van Oorschot, et al., Are you collecting all the available DNA from touched objects? Int. Congr. Ser. 1239 (0) (2003) 803–807.
[21] F. Alessandrini, et al., Fingerprints as evidence for a genetic profile: morphological study on fingerprints and analysis of exogenous and individual factors affecting DNA typing, J. Forensic Sci. 48 (3) (2003) 586–592.
[22] I. Quinones, B. Daniel, Cell free DNA as a component of forensic evidence recovered from touched surfaces, Forensic Sci. Int. Genet. 6 (1) (2012) 26–30.
[23] B. Mercier, et al., Direct PCR from whole blood, without DNA extraction, Nucleic Acids Res. 18 (19) (1990) 5908.
[24] R. Ottens, et al., Successful direct amplification of nuclear markers from a single hair follicle, Forensic Sci. Med. Pathol. 9 (2) (2013) 238–243.
[25] A. Linacre, et al., Generation of DNA profiles from fabrics without DNA extraction, Forensic Sci. Int. Genet. 4 (2) (2010) 137–141.
[26] R. Ottens, et al., Application of direct PCR in forensic casework, Forensic Sci. Int. Genet. Suppl. Ser. 4 (1) (2013) e47–e48.
[27] J.E. Sim, et al., High-throughput STR analysis for DNA database using direct PCR, J. Forensic Sci. 58 (4) (2013) 989–992.
[28] D.Y. Wang, et al., Development and validation of the AmpF/STR® Identifiler® direct PCR amplification kit: a multiplex assay for the direct amplification of single-source samples, J. Forensic Sci. 56 (4) (2011) 835–845.
[29] C. Roux, et al., A further study to investigate the effect of fingerprint enhancement techniques on the DNA analysis of bloodstains, J. Forensic Identif. 49 (4) (1999) 357.
[30] J.J. Raymond, et al., The effect of common fingerprint detection techniques on the DNA typing of fingerprints deposited on different surfaces, J. Forensic Identif. 54 (1) (2004) 22.
[31] P. Kumar, et al., Effects of latent fingerprint development reagents on subsequent forensic DNA typing: a review, J. Forensic Legal Med. 32 (2015) 64–69.
[32] C. Thamnurak, et al., DNA typing from fluorescent powder dusted latent fingerprints, Forensic Sci. Int. Genet. Suppl. Ser. 3 (1) (2011) e524–e525.
[33] C. Stein, C. Henssge, S.H. Kyeck, DNA typing of fingerprint reagent treated biological stains, J. Forensic Sci. 41 (6) (1996) 1012–1017.
[34] C.J. Frégeau, O. Germain, R.M. Fourney, Fingerprint enhancement revisited and the effects of blood enhancement chemicals on subsequent Profiler Plus™ fluorescent short tandem repeat DNA analysis of fresh and aged bloody fingerprints, J. Forensic Sci. 45 (2) (2000) 354–380.
[35] A. Barbaro, et al., Anonymous letters? DNA and fingerprints technologies combined to solve a case, Forensic Sci. Int. 146 (2004) S133–S134.
[36] J.J. Raymond, et al., Trace DNA success rates relating to volume crime offences, Forensic Sci. Int. Genet. Suppl. Ser. 2 (1) (2009) 136–137.
[37] L. Ostojic, E. Wurmbach, Analysis of fingerprint samples, testing various conditions, for forensic DNA identification, Sci. Just. (2016).
[38] J. Templeton, A. Linacre, DNA profiles from fingermarks, Biotechniques 57 (5) (2014) 259.
[39] S. Montpetit, P. O'Donnell, An optimized procedure for obtaining DNA from fired and unfired ammunition, Forensic Sci. Int. Genet. 17 (2015) 70–74.
[40] D.W. Steadman, S.A. Andersen, Latent print processing of human bones, J. Forensic Identif. 53 (5) (2003) 556.
[41] M. Pesaresi, et al., Qualitative and quantitative analysis of DNA recovered from fingerprints, Int. Congr. Ser. 1239 (2003) 947–951.
[42] Y.C. Swaran, L. Welch, A comparison between direct PCR and extraction to generate DNA profiles from samples retrieved from various substrates, Forensic Sci. Int. Genet. 6 (3) (2012) 407–412.
[43] J.E. Templeton, et al., Direct PCR improves the recovery of DNA from various substrates, J. Forensic Sci. 60 (6) (2015) 1558–1562.
[44] A. Lowe, et al., The propensity of individuals to deposit DNA and secondary transfer of low level DNA from individuals to inert surfaces, Forensic Sci. Int. 129 (1) (2002) 25–34.
[45] P. Tozzo, et al., Effect of dactyloscopic powders on DNA profiling from enhanced fingerprints: results from an experimental study, Am. J. Forensic Med. Pathol. 35 (1) (2014) 68–72.
[46] R.K. Farmen, et al., Assessment of individual shedder status and implication for secondary DNA transfer, Forensic Sci. Int. Genet. Suppl. Ser. 1 (1) (2008) 415–417.
[47] M. Goray, et al., Secondary DNA transfer of biological substances under varying test conditions, Forensic Sci. Int. Genet. 4 (2) (2010) 62–67.
[48] M. Goray, R.A. van Oorschot, J.R. Mitchell, DNA transfer within forensic exhibit packaging: potential for DNA loss and relocation, Forensic Sci. IntGenet. 6 (2) (2012) 158–166.
[49] C. Ladd, et al., A systematic analysis of secondary DNA transfer, J. Forensic Sci. 44 (1999) 1270–1272.
[50] R.A. Wickenheiser, Trace DNA: a review, discussion of theory, and application of the transfer of trace quantities of DNA through skin contact, J. Forensic Sci. 47 (3) (2002) 442–450.
[51] M. Goray, J.R. Mitchell, R.A. van Oorschot, Evaluation of multiple transfer of DNA using mock case scenarios, Leg. Med. 14 (1) (2012) 40–46.
[52] H.C. Lee, R. Ramotowski, R. Gaensslen, Advances in Fingerprint Technology, CRC Press, 2001.

Further discussion of manuscript

Results were generated from 160 previously enhanced fingermarks created by 10 individuals, without the need to increase PCR cycle numbers beyond manufacturer's guidelines. Resulting DNA profiles were cross examined against buccal swab reference profiles obtained from the donors and evaluated for the matching number of STR alleles. Out of 160 samples tested by direct PCR, 98 samples (i.e. 61%) were considered informative profiles (i.e. 'up-loadable' to the Australian National Criminal Investigation DNA Database (NCIDD)). Mixtures were evident in 31% of samples (i.e. 49 out of 160 samples), single source profiles were observed in 61% of cases (i.e. 98 out of 160 samples) and only 8% of the profiles completely failed to yield DNA alleles (i.e. 13 out of 160 samples). No unusual levels of stochastic effects were observed and the average heterozygote peak balance for the fingerprint dusted samples using direct PCR was 76%.

The interference of fingerprint powders on the ability to generate a DNA profile from fingerprints has been investigated previously when DNA extraction was employed [58]; 14 out of 48 samples were successfully profiled (i.e. 30%) after powder enhancement, and for scotch tape-archived prints, 9 out of 48 samples (i.e. 18%) generated meaningful profiles [58]. In comparison, the data in this chapter report a higher success rate with 61% of direct PCR samples generating informative profiles (i.e. ≥ 12 alleles, plus Amelogenin). Further work, following on from this thesis, could involve a comparison between extracted DNA swabs and direct PCR swabs post fingerprint enhancement. Traditional practice would involve extraction and quantification of swabs that are saturated in fingerprint powders.

Additional studies highlight a variation in DNA profiling results after fingerprint development techniques [24-29]. Future work could measure the effect of other fingerprint enhancement techniques on the direct PCR approach (e.g. superglue fuming and ninhydrin). One research group examined the recovery of DNA obtained from latex gloves, grip areas of tools, drinking glasses, and clothes, and concluded that DNA yield varied between 33%-100% of donor DNA and the success rate was highly dependent on the chemicals used for enhancement [59]. The authors

observed that if DNA typing was performed < 7 days post-fingerprint enhancement there was a higher rate of DNA recovery [59]. Moreover, fluorescent powders were reported to have no apparent effect on the quality of DNA profiles [29], and the authors claim that variation in results seen would be due to the physical act of dusting cells away from the surface as opposed to the inhibitory effect of powders. A loss of cells is possible when using fingerprint brushes as they are potentially wiped away during the process.

One study examining fingerprints left behind on corpses showed that magnetic-based powdered fingerprints were superior to black powdered fingerprints for classical fingerprinting techniques [60]. There was a low DNA recovery rate for powdered prints overall, however, black powdered prints were more successful than magnetic-based fingerprint powders; 2.2% of black powdered prints generated meaningful profiles compared to 1.8% of magnetic-based powders [60]. Other research indicates problems with DNA recovery and low success rates in obtaining meaningful profiles after fingerprint powder enhancement or from sprayed fingerprints [61]. Data from this chapter highlight that direct PCR is a viable option for future processing of fingerprints as the technique vastly improves sensitivity; circumvents the need for an extraction; generates results in a timely manner; and compared to other published studies has an extremely good success rate at obtaining good quality DNA profiles.

Example of DNA profile obtained by direct PCR, using AmpFℓSTR® ProfilerPlus™ kit, from a previously enhanced fingermark dusted with silver fingerprint powder

Silver powder DNA profile – direct PCR

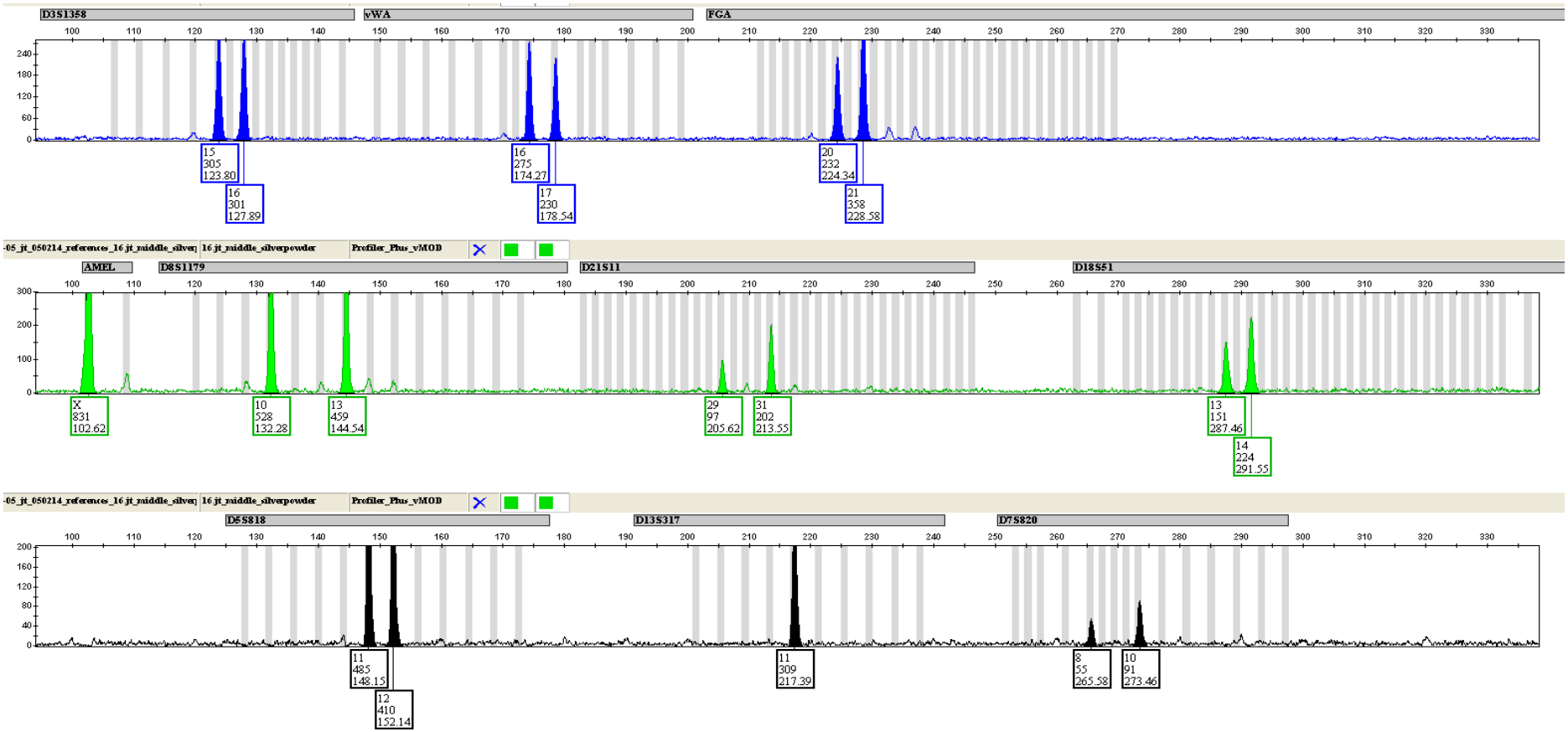


Figure 5.4 DNA profile obtained from a previously enhanced fingermark deposited 15 minutes after the donor washed their hands. A full DNA profile was obtained (i.e. 20 out of 20 possible alleles) with the AmpFℓSTR® ProfilerPlus™ kit and all DNA alleles

matched the reference profile of the donor. This profile would be considered ‘up-loadable’ to the Australian National Criminal DNA Database (NCIDD).

Example of DNA profile obtained by direct PCR, using AmpFℓSTR® ProfilerPlus™ kit, from a previously enhanced fingermark dusted with white fingerprint powder

White powder DNA profile – direct PCR

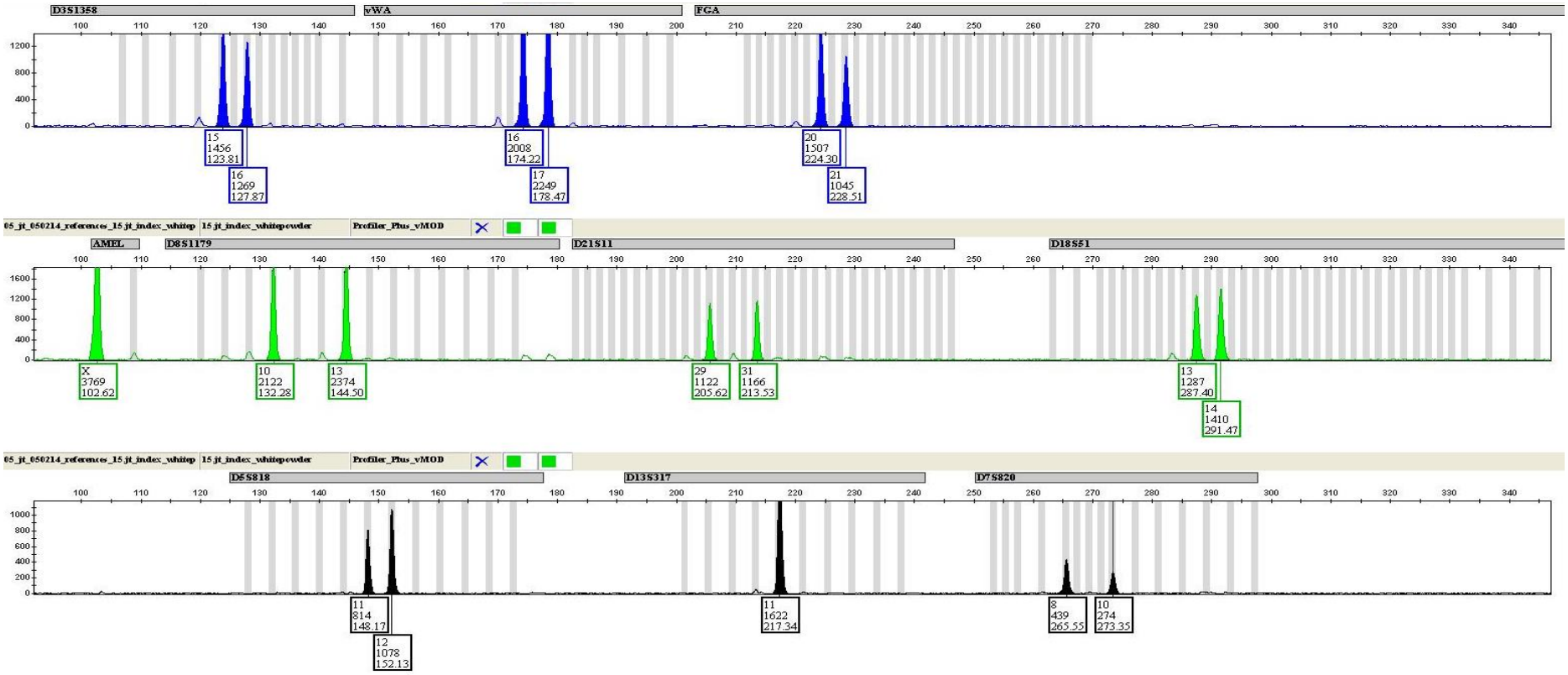


Figure 5.5 DNA profile obtained from a previously enhanced fingermark deposited 15 minutes after the donor washed their hands. A full DNA profile was obtained (i.e. 20 out of 20 possible alleles) with the AmpFℓSTR® ProfilerPlus™ kit, and all DNA alleles matched the reference profile of the donor. This profile would be considered ‘up-loadable’ to NCIDD.

Example of DNA profile obtained by direct PCR, using AmpFℓSTR® ProfilerPlus™ kit, from a previously enhanced fingermark dusted with black fingerprint powder

Black powder DNA profile – direct PCR

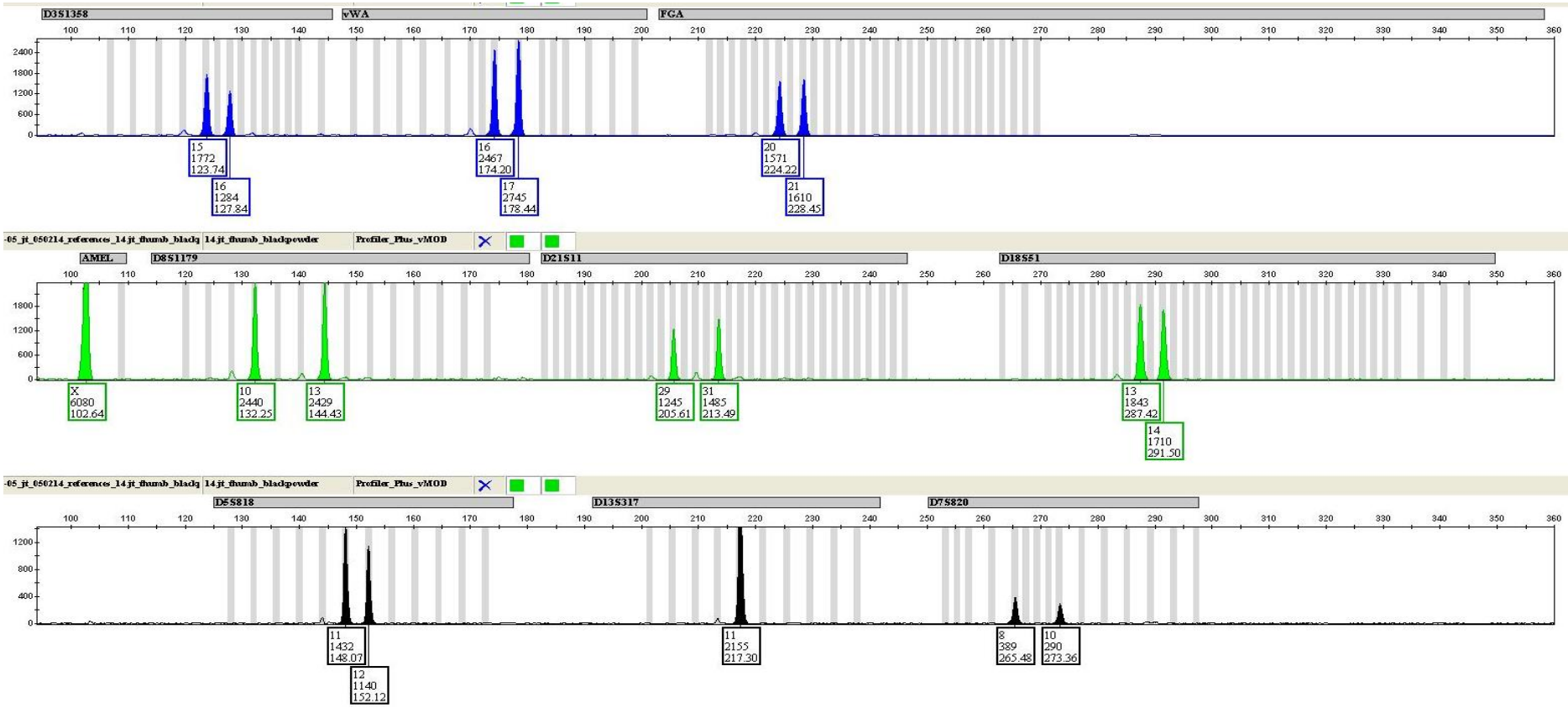


Figure 5.6 DNA profile obtained from a previously enhanced fingermark deposited 15 minutes after the donor washed their hands. A full DNA profile was obtained (i.e. 20 out of 20 possible alleles) with the AmpFℓSTR® ProfilerPlus™ kit, and all DNA alleles matched the reference profile of the donor. This profile would be considered 'up-loadable' to NCIDD.

5.4.2 Substrates and surfaces

5.4.2.1 Materials and Methods

Preliminary studies to supplement manuscript data

Substrates

The 'targeted' swabbing method (detailed in Chapter II) was used in conjunction with direct PCR to test the effectiveness of the technique on a range of touched substrates (see below).

Individuals were asked to wash their hands and briefly (15 seconds) touch an item 15 minutes after hand washing. Previous results from Chapter IV indicate that DNA can accumulate on the skin within 15 minutes after hand washing [62]. Each substrate was swabbed and swab fibres utilised for direct PCR using either: AmpFℓSTR® NGM SElect™ PCR kit or AmpFℓSTR® ProfilerPlus™ kit (Life Technologies). For ProfilerPlus™, half reaction volume was used with 1 µL Ampli*Taq* Gold® DNA polymerase (5 units) (Life Technologies), and additional 1 µL of DMSO (5%) (Expand Long Range DNTPACK; Roche, Vic, AU), and 1 µL of Molecular Biology Grade BSA (0.1 µg, New England Biolabs, NSW, AU) in the PCR set up. For AmpFℓSTR® NGM SElect™, manufacturer's recommendations were followed and extra Ampli*Taq* Gold® DNA polymerase 360 (Life Technologies) (1 µL, 5 units) was added to boost amplification. All PCR products were analyzed on a 3130*xl* Genetic Analyzer (Life Technologies) following kit recommendations for the kit-specific size standard and ladder. Each test was carried out in triplicate.

For the mock case study, AmpFℓSTR® NGM SElect™ kit was used with extra Ampli*Taq* Gold® DNA polymerase 360 (Life Technologies) (1 µL, 5 units) and Prep-n-Go™ Buffer (ABI) (1 µL). The remainder of the reaction was made up to 25 µL with TE buffer (i.e. 10 mM Tris, 0.1 mM EDTA, pH 8.0).

Images of substrates examined

Samples analysed by direct PCR for the mock case study

Figure 5.7 Ammunition: nickel (38 Special, Winchester “Super X” centre fire ammunition, nickel plated case), aluminium (38 Special, CCI “Blazer” centre fire ammunition) and brass (Winchester, Australia) cartridge cases.

Figure 5.8 Cartridge cases: nickel, aluminium and brass casings (same as above) left exposed to external environmental inhibitors.

Samples analysed by direct PCR for the mock case study

Figure 5.9 Samples analysed for the mock case study: knifes; tape; and cartridge cases, left exposed to external environmental inhibitors.

5.4.2.2 Results and Discussion

Direct PCR from swab fibres was used to generate full and partial DNA profiles from a range of touched substrates (see substrates highlighted in Figures 5.7, 5.8 and 5.9) using standard amplification conditions and the AmpFℓSTR® NGM SElect™ PCR amplification kit (Life Technologies, AU) or AmpFℓSTR® ProfilerPlus™ PCR profiling kit. Results were replicated (X 3) for each individual depositing DNA (n=4), and for each substrate tested (i.e. wood, plastic, glass, nickel, brass and aluminium), and data published in the manuscript enclosed: "***Templeton, Jennifer EL, et al.*** **DNA profiles from fingermarks: A mock case study. Forensic Science International: Genetics Supplement Series 5 (2015): e154-e155".**

5.4.2.2.1 Statement of authorship

Manuscript: DNA profiles from fingermarks: A mock case study.

Manuscript submitted to the journal Forensic Science International: Genetics

Jennifer Templeton (Candidate)

Designed the experiment, performed all the laboratory work, data analysis, and wrote the paper

I hereby certify that the statement of contribution is accurate.

Signed:  Date: November 2016

Duncan Taylor (Supervisor)

Commented on data, and helped to edit the paper

I hereby certify that the statement of contribution is accurate.

Signed:  Date: November 2016

Oliva Handt (Supervisor)

Commented on data, and helped to edit the paper

I hereby certify that the statement of contribution is accurate.

Signed:  Date: November 2016

Adrian Linacre (Supervisor)

Supervised the project, assisted with experimental design, commented on data, and edited the paper

I hereby certify that the statement of contribution is accurate.

Signed: Date: November 2016

5.4.2.2.2 Manuscript: DNA profiles from fingermarks: A mock case study.

Forensic Science International: Genetics Supplement Series 5 (2015) e154–e155

Contents lists available at ScienceDirect

Forensic Science International: Genetics Supplement Series

journal homepage: www.elsevier.com/locate/FSIGSS

DNA profiles from fingermarks: A mock case study

Jennifer E.L. Templeton[a], Duncan Taylor[a,b], Oliva Handt[a,b], Adrian Linacre[a,*]

[a] School of Biological Sciences, Flinders University, Adelaide, Australia
[b] Forensic Science South Australia, Adelaide, Australia

ARTICLE INFO

Article history:
Received 17 August 2015
Accepted 14 September 2015
Available online 28 September 2015

Keywords:
Direct PCR
Touch DNA
Identification
Short tandem repeat
Profiling

ABSTRACT

Environmental factors and subsequent PCR inhibitors may affect the ability to retrieve DNA from handled objects. A poor quality DNA profile results from low DNA template or modifications imposed on the template (damage or degradation) prior to amplification. In this study, we show that direct PCR has the ability to amplify 'touch' DNA from uncleaned, handled, substrates following exposure to various environmental conditions. A variety of items ($n = 90$), ranging from metal cartridge cases, glass, tape and wood were touched briefly (less than 15 s) by 4 volunteers, exposed to the environment overnight, and in one case for 8 days, and tested for the presence of surviving DNA. Informative DNA profiles were obtained from 49 out of the 90 items (a success rate of 54%), with glass having the highest rate of DNA recovery.

1. Introduction

Since its inception, research on 'touch' DNA samples has suffered from the problem that preserved endogenous DNA is often limited in quantity [1] and in various states of degradation [2]. It has been found that a significant number of crime-related surfaces, even though they should have DNA present, do not necessarily yield analytically useful material [3]. In natural conditions DNA is a hydrated macromolecule but outside its natural environment DNA becomes the target of several physical and chemical reactions [4]. There may be no surviving DNA left to capture on a handled object if the DNA is exposed to elements of heat, humidity, ultra violet (UV) light and bacterial growth, typical of an outdoor environment. Under these circumstances, standard protocols for swabbing and extracting 'touch' DNA from these items often recover sub-optimal levels of trace nuclear DNA [3] that can result in a poor quality STR-based DNA profile or no profile at all. The aim of this study is to determine the value of using the optimised swabbing technique [5] and direct PCR at amplifying 'touch' DNA from fingermarks as part of a mock case study.

2. Materials and methods

A number of precautions were taken for the handling and processing of 'touch' DNA samples; these have been extensively described elsewhere [5]. Four volunteers touched various substrates (wooden knife handles, glass, masking tape, brass cartridge cases, aluminium cartridge cases and nickel cartridge cases) for less than 15 s, before the items were left out-side in a shrub bed area. An indoor glass door was touched for less than 15 s to create fingerprint; these samples were not subjected to external conditions but were left for the same time period as other samples in the study. Substrates were not cleaned prior to touching to ensure results were as realistic as possible and to determine the extent of mixed DNA profiles. Environmental conditions were monitored. Sampling was repeated in triplicate for each volunteer. Twenty four hours later, or 8 days later (sample study 5), items were collected/processed ($n = 90$) and for wood, glass and cartridge cases a targeted swabbing approach was used [5] subsequent to direct PCR. For masking tape, a 2 cm^2 section of tape was placed into a 1.5 mL Eppendorf tube containing Triton X buffer at 0.1% (Sigma, VIC, AU) and heated at 50 °C for 1 h, vortexed and 10 μL of this lysis buffer was added directly to the PCR tube in place of water/DNA. Amplification conditions followed the NGM™ Select kit (ABI, VIC, AU) guidelines in a final 25 μL reaction volume, plus 1 μL Prep-N-Go buffer™ (ABI) and 1 μL AmpliTaq Gold 360 Polymerase (ABI) per reaction. Amplification was for 30 PCR cycles. PCR products were run on the 3130*xl* (ABI) genetic analyser. Profiles were analysed using the GeneMapper ID v3.2 software and a peak amplitude threshold of 50 relative fluorescent units (RFU) was used to assign alleles. Informative profiles were recorded as >12 alleles called (plus Amelogenin) that match the donor.

3. Results and discussion

Certain substrates, such as brass cartridge cases (alloy of copper and zinc), present difficulties for DNA recovery (see Table 1). From

* Corresponding author.
E-mail address: adrian.linacre@flinders.edu.au (A. Linacre).

http://dx.doi.org/10.1016/j.fsigss.2015.09.062
1875-1768/

Table 1
Results for the number of single source profiles, mixtures, and the number of interpretable DNA profiles in each category for the NGM SElect™ directly amplified samples.

DNA types detected	Substrates handled					
	Glass	Wooden knife handles	Masking tape	Brass cartridge cases	Nickel cartridge cases	Aluminium cartridge cases
Total interpretable profiles 15 = no. of alleles	13/15 (87%)	8/15 (53%)	11/15 (73%)	0/15 (0%)	9/15 (60%)	8/15 (53%)
Informative single source profiles	38%	50%	64%	0%	89%	75%
Informative mixed profiles	62%	50%	36%	0%	11%	25%
Donor identified	87%	53%	73%	0%	60%	53%

Table 2
Summary of the mock case work conditions and overall profiling success rates for substrates exposed to varying environmental conditions over-night and in one case 8 days (study 5*).

Category	Environmental conditions during the 24 h time period or 8 days (study 5*)					Total number of samples (out of 15)	Interpretable profiles (out of 15)	Profiling success (%)
	Temp. range (°C)	Relative maximum humidity (%)	Max. wind speed (km/h)	Average rain fall (mm)	UV index			
Volunteer 1 study	11–17 Av: 15	88	30	0	4	15	14	93
Volunteer 2 study	12–17 Av: 14.5	71	39	5.8	2	15	9	60
Volunteer 3 study	5–16 Av: 10.5	58	24	0.6	3	15	8	53
Volunteer 4 study	8–16 Av: 12	61	39	4.8	2	15	11	73
Volunteer 1 study 5*	1–19 Av: 10	94	30	3.6	3	15	7	46

our results, glass had the highest rate of DNA recovery, followed by masking tape, nickel cartridge cases, wood and aluminium cartridge cases (see Table 1). Mixed DNA profiles were observed as expected, as substrates were not cleaned prior to volunteers depositing fingermarks. However, the major profile was identified as the donor in all cases. Good quality DNA profiles were observed in all cases with minimum observed stochastic effects or artefacts (i.e. heterozygote imbalance, increased stutter and allele/laboratory associated drop-in). Substrates left outside and exposed to environmental conditions for up to 8 days may reduce the overall profiling success (see Table 2, sample study 5*). Short exposure to UV light and rain do not appear to adversely affect DNA recovery in this study as informative profiles were obtained. However a larger sample study is required before conclusions can be drawn as to the effect of long term environmental exposure.

4. Conclusion

Optimising the swabbing procedure and eliminating the extraction stage allowed informative DNA profiles to be generated from touched substrates left exposed to the environment for a fixed length of time. The method of direct PCR should be considered as an alternative method for analysing samples that contain low amounts of DNA. A validated study is required to assess the limitations of the direct PCR approach.

Conflict of interest

None.

Acknowledgements

Funding was provided by the Attorney-General's Office, South Australia and a John Stocker Postgraduate Scholarship from the Science and Industry Endowment Fund.

References

[1] J.M. Butler, Advanced Topics in Forensic DNA Typing: Methodology, Academic Press, San Diego, 2011.
[2] K. Toshiro, H. Yamaguchi, M. Yokoyama, et al., Morphological study of fragmented DNA on touched objects, Forensic Sci. Int. Genet. 3 (1) (2008) 32–36.
[3] S. Harbison, M. Fallow, D. Bushell, An analysis of the success rate of 908 trace DNA samples submitted to the Crime Sample Database Unit in New Zealand, Aust. J. Forensic Sci. 40 (1) (2008) 49–53.
[4] R. Alaeddini, Forensic implications of PCR inhibition—a review, Forensic Sci. Int. Genet. 6 (3) (2012) 297–305.
[5] J.E. Templeton, A. Linacre, DNA profiles from fingermarks, BioTechniques 57 (5) (2014) 259.

Gun swabbing

On the 24th October 2013, I visited Thebarton Police Academy to swab three guns that had been used by a police officer to fire ammunition. All guns had been cleaned prior to handling with gun oil (i.e. mineral oil), as part of standard practice. All guns were handled for as long as required to load and shoot bullets (i.e. ~ 30 – 45 seconds). 'Whole' swabbing was used (see Chapter II) with nylon flocked swabs and 4 µL Triton-X™ (0.1%) to swab the surface of the guns immediately after firing, and swabs stored at 4 °C prior to direct PCR. Two swabs were taken from each gun and fibres subjected to direct PCR.

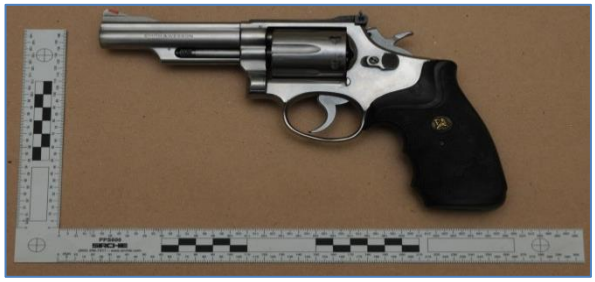
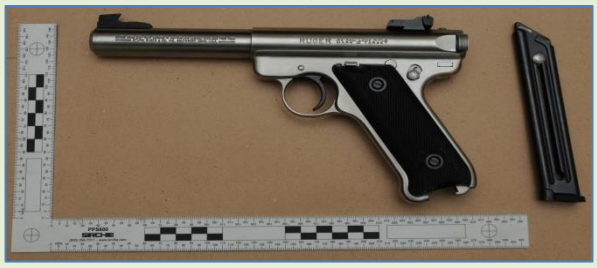
Guns swabbed:	DNA profile obtained (loci):
i) **22 Long Rifle semi-automatic handgun** 	Partial profile from shoulder of rifle (5 loci); mixed Partial profile from handle (3 loci); mixed
ii) **38 Special revolver** 	Partial profile obtained from hand-rod (14 loci); mixed - **see figure 5.10 below** Partial profile obtained from trigger (5 loci); mixed
iii) **12 Guage pump action shotgun** 	Partial profile obtained from trigger (5 loci); mixed No result from handle

Table 5.1 Guns swabbed at Thebarton Police Academy on 24th October 2013, post-handling and firing.

Gun swabbing profiling results

Example of DNA profile obtained from swabbing the hand-rod of a revolver post-handling and firing

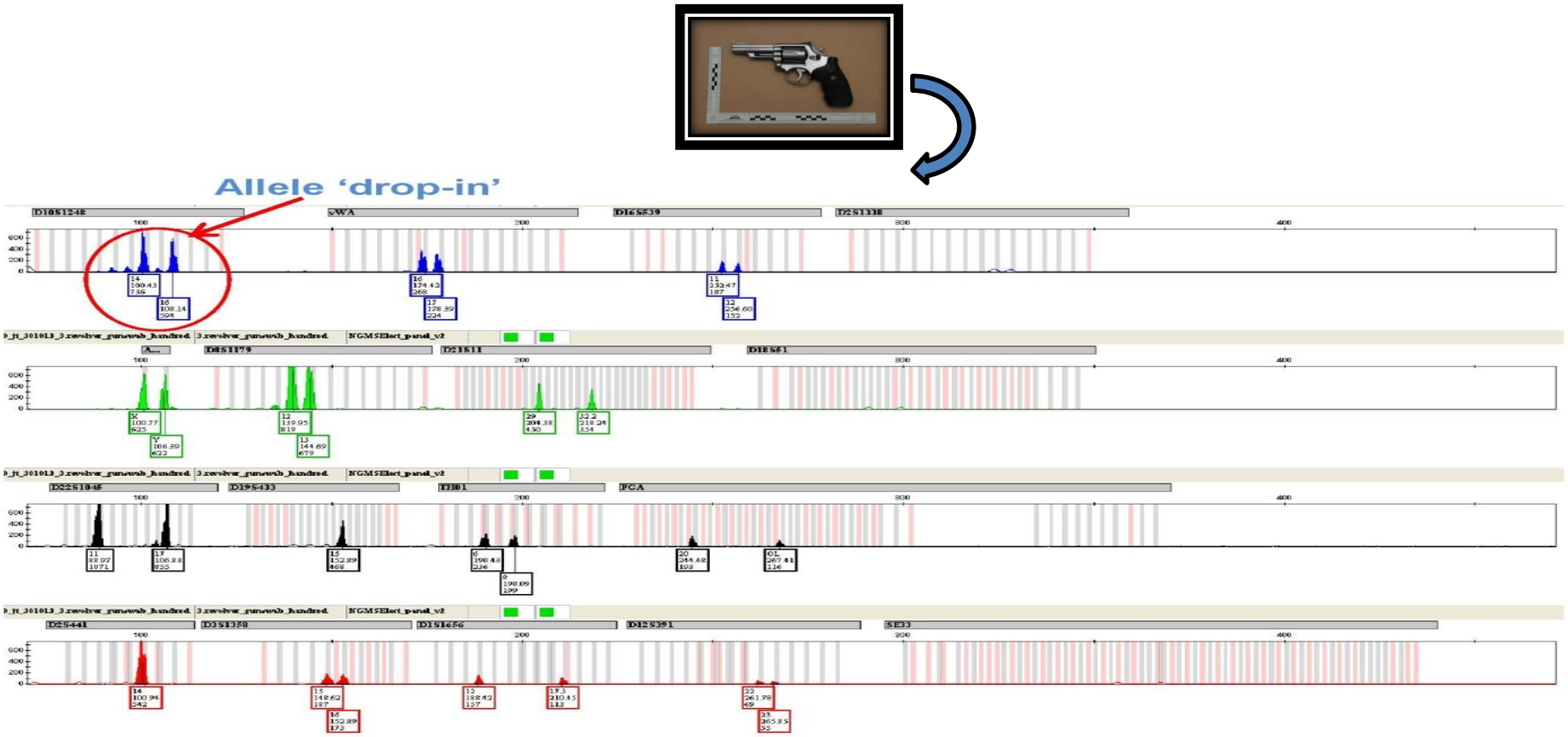


Figure 5.10 Showing a poor quality STR-based DNA profile obtained from swabbing the ejector rod of a revolver and subjecting the swab fibres to direct PCR. Ideally this sample would be re-run as peak morphology is not correct, most likely a size standard or run issue. A mixed DNA profile was generating by AmpFℓSTR® NGM SElect™ PCR. Additional alleles (i.e. allele 'drop-in') were detected, indicating a mixed DNA profile (i.e. more than one contributor). There are two off-ladder peaks at D10S1248 and alleles that have fallen off ladder at marker FGA.

DNA profiling results for other substrate types

Example of DNA profile obtained from fingermark residue deposited on masking tape

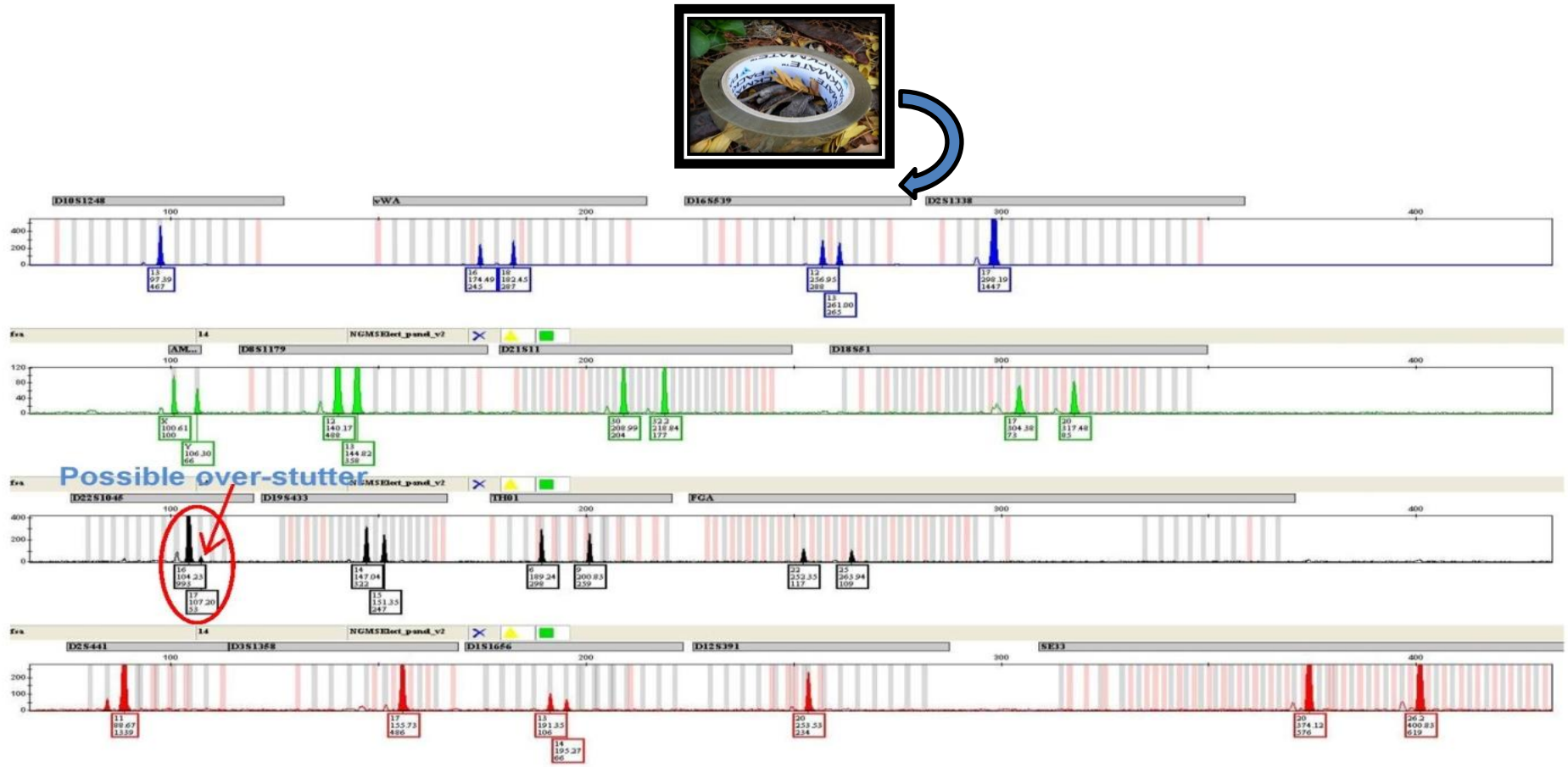


Figure 5.11 STR-based DNA profile obtained using direct PCR from handled masking tape. A full DNA profile was obtained (i.e. 34 out of 34 possible alleles) with the AmpFℓSTR® NGM SElect™ PCR amplification kit, and all DNA alleles matched the reference profile of the donor. This profile would be considered 'up-loadable' to NCIDD. A small peak is observed at loci D22S1045; possibly over-stutter, as D22S1045 has a higher rate for stutter being a tri-nucleotide.

DNA profiling results for other substrate types

Example of DNA profile obtained from fingermark residue deposited on a nickel cartridge case

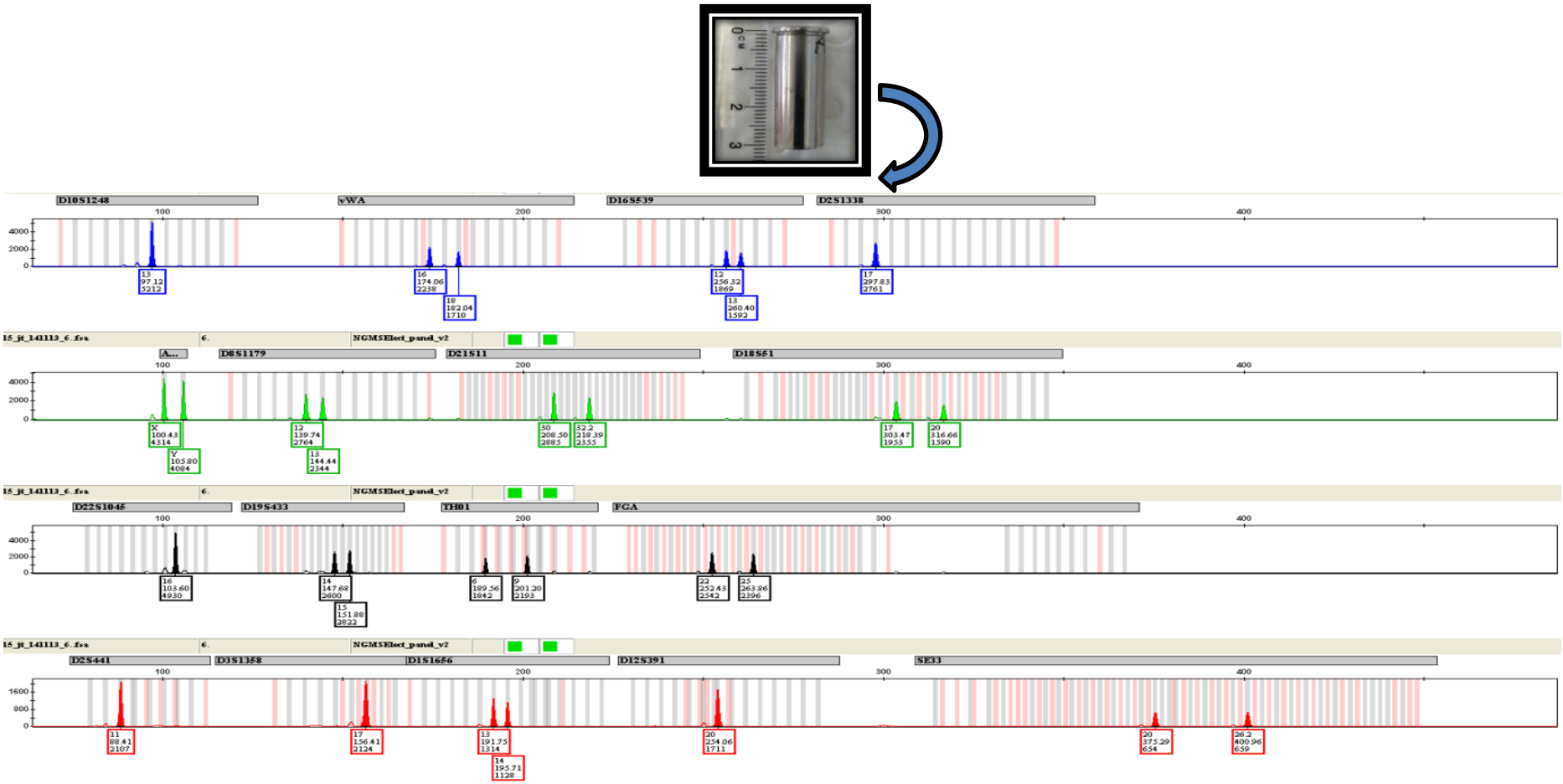


Figure 5.12 STR-based DNA profile obtained from a nickel cartridge case post-handling. A full DNA profile was obtained (i.e. 34 out of 34 possible alleles) with the AmpFℓSTR® NGM SElect™ PCR amplification kit, and all DNA alleles matched the reference profile of the donor. This profile would be considered ‘up-loadable’ to NCIDD.

DNA profiling results for other substrate types

Example of DNA profile obtained from fingermark residue deposited on an aluminium cartridge case

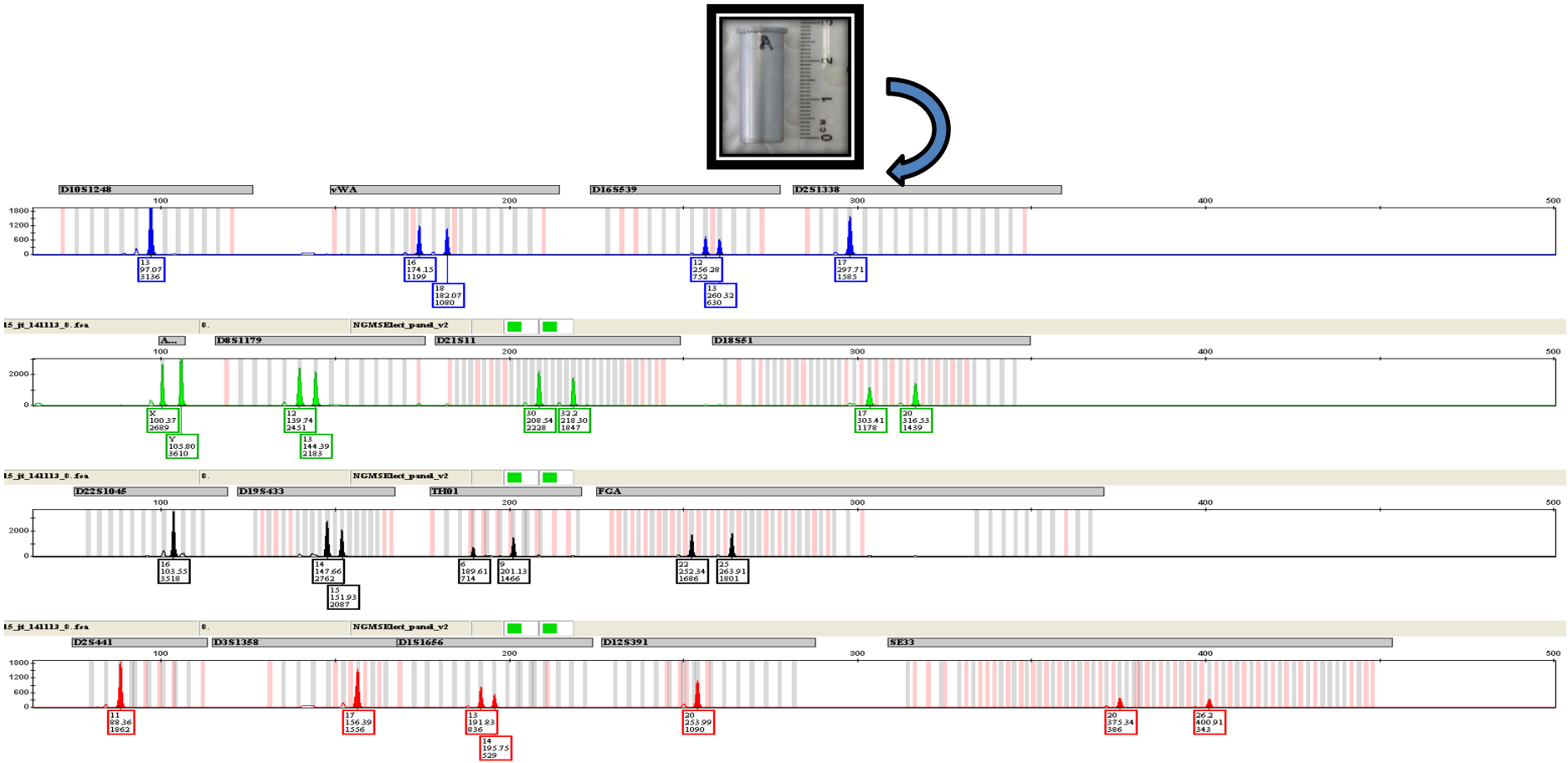


Figure 5.13 STR-based DNA profile obtained from an aluminium cartridge case post-handling. A full DNA profile was obtained (i.e. 34 out of 34 possible alleles) with the AmpFℓSTR® NGM SElect™ PCR amplification kit, and all DNA alleles matched the reference profile of the donor. This profile would be considered 'up-loadable' to NCIDD.

DNA profiling results for other substrate types

Example of DNA profile obtained from fingermark residue deposited on a brass cartridge case

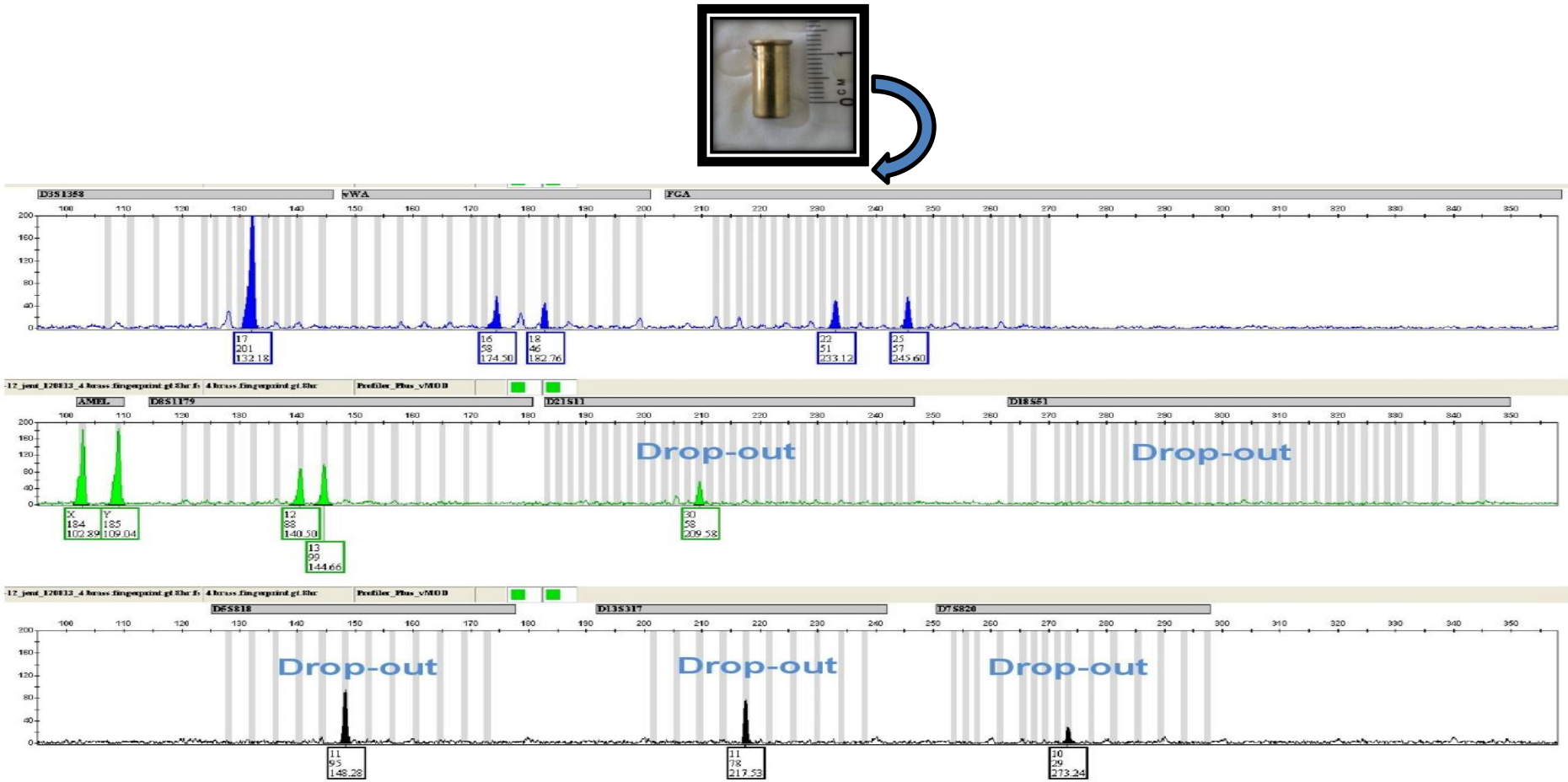


Figure 5.14 STR-based DNA profile obtained from a brass cartridge case. A partial DNA profile was obtained (i.e. 12 out of 20 possible alleles) with the AmpFℓSTR® ProfilerPlus™ PCR amplification kit, and all DNA alleles matched the reference profile of the donor. This profile would be considered 'up-loadable' to NCIDD.

DNA profiling results for other substrate types

Example of DNA profile obtained from fingermark residue deposited on plastic ziplock bag

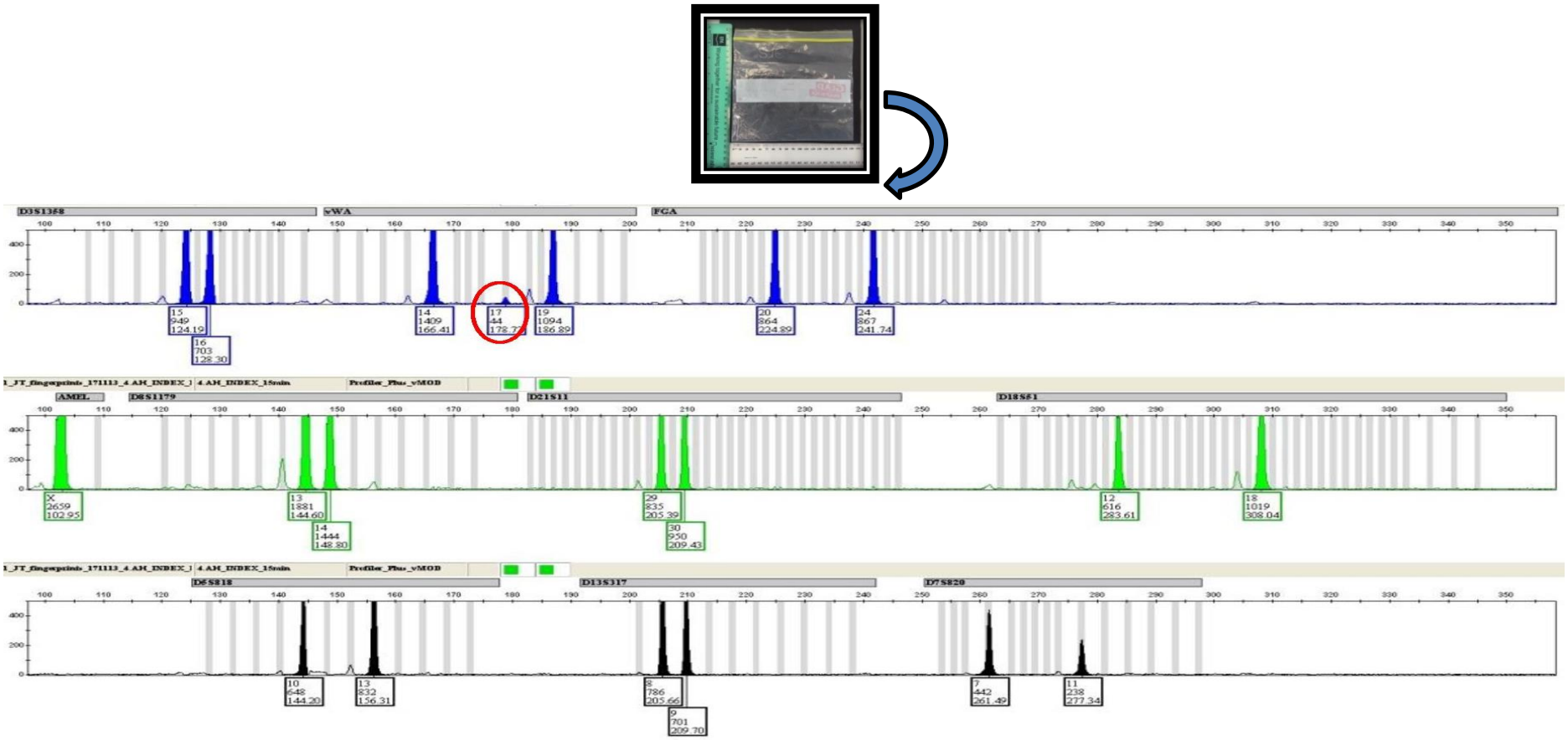


Figure 5.15 STR-based DNA profile obtained by direct PCR from swabbing a handled plastic ziplock bag. A full DNA profile was obtained (i.e. 20 out of 20 possible alleles) with the AmpFℓSTR® ProfilerPlus™ PCR amplification kit, and all DNA alleles matched the reference profile of the donor. This profile would be considered 'up-loadable' to NCIDD. An additional allele at marker vWA was observed that was not attributed to stutter.

DNA profiling results for other substrate types

Environmental exposure

Example of DNA profile generated from a hand print 8 days after print deposition

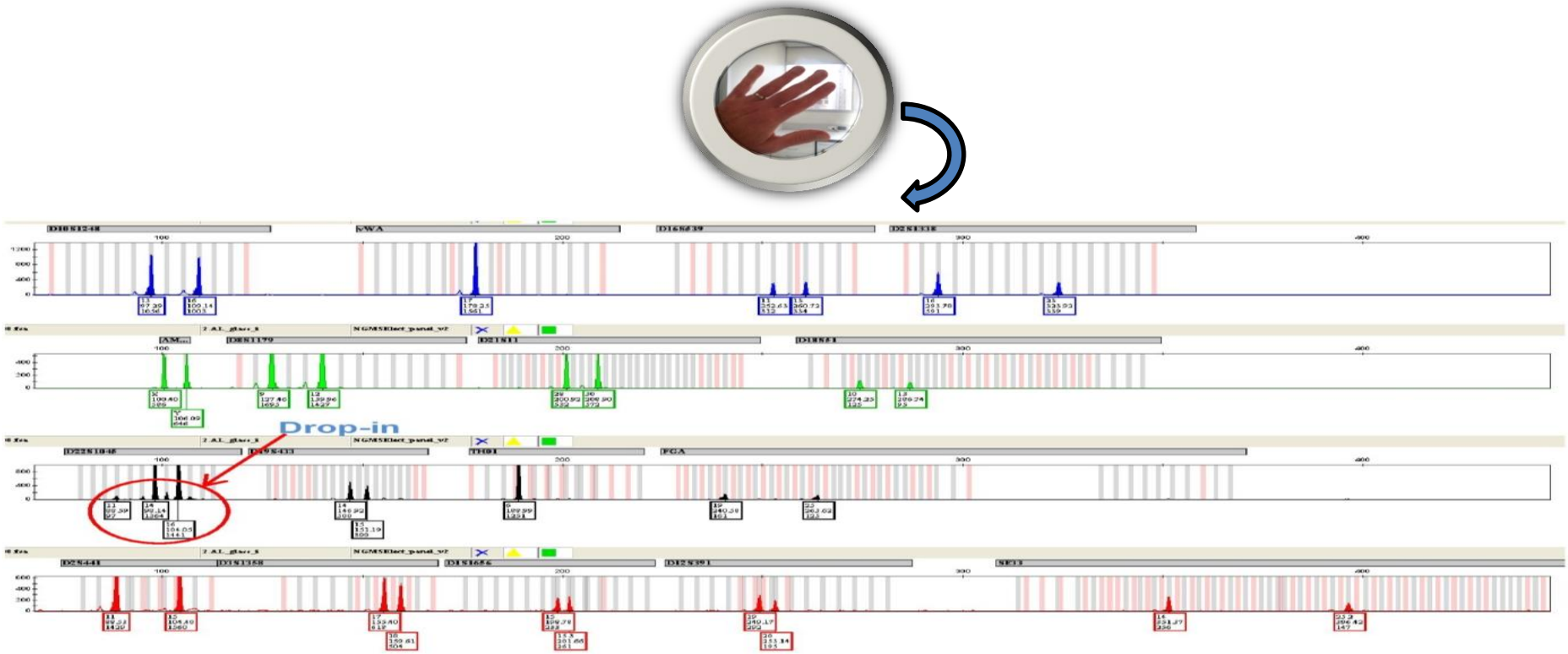


Figure 5.16 STR-based DNA profile obtained by direct PCR from a glass handprint exposed to the environment for 8 days prior to swabbing. A full DNA profile was obtained with the AmpFℓSTR® NGM SElect™ PCR amplification kit where DNA alleles matched the reference profile of the donor (i.e. 34 out of 34 possible alleles). This profile would be considered 'up-loadable' to NCIDD. Additional alleles (i.e. 'drop-in') were also detected and mixture analysis is required when more than two alleles per loci are detected.

DNA profiling results for other substrate types

Example of DNA profile generated from a handled nickel cartridge case 8 days after print deposition

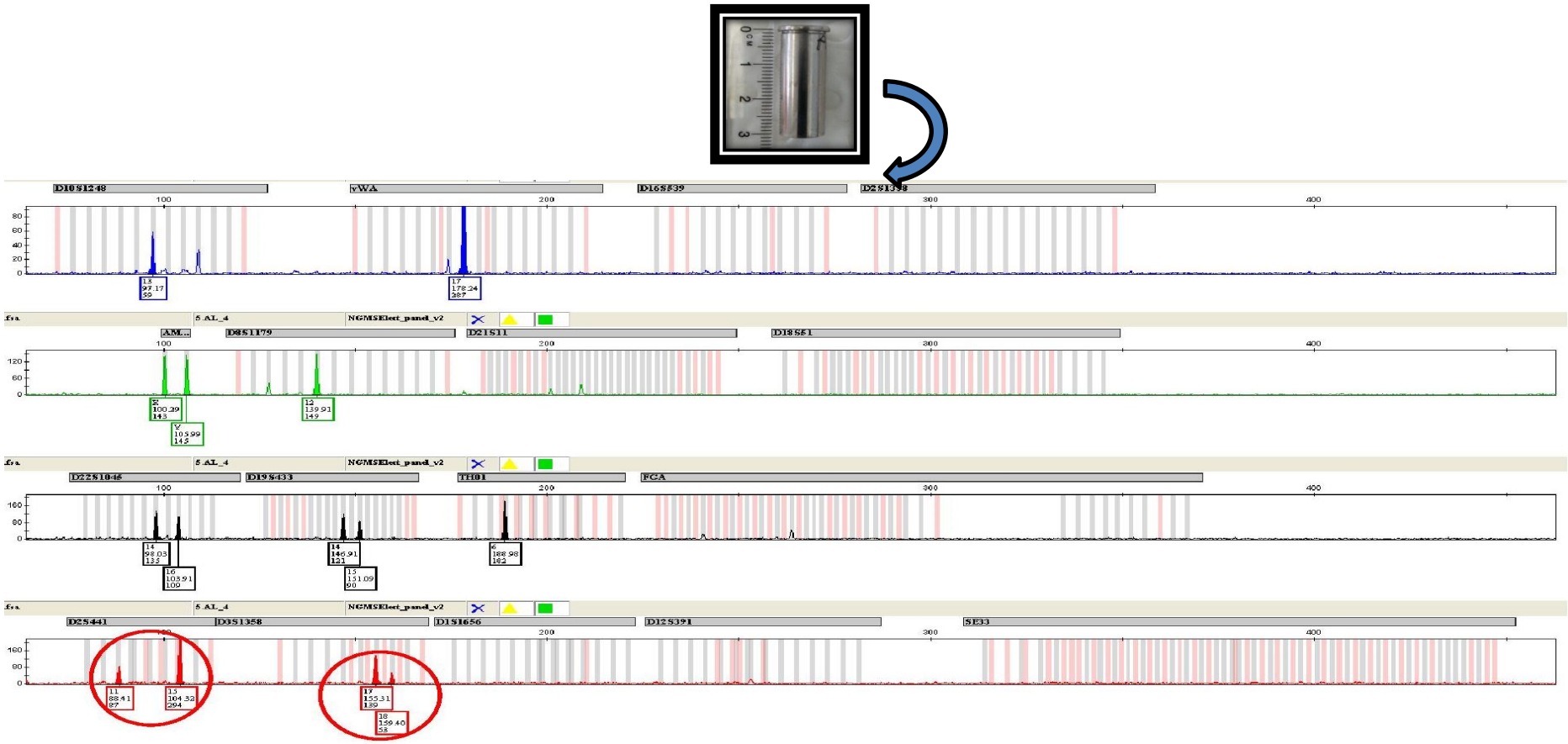


Figure 5.17 STR-based DNA profile obtained by direct PCR from a swabbed nickel cartridge case that was subjected to the environment for 8 days. A partial DNA profile was obtained with the AmpFℓSTR® NGM SElect™ PCR amplification kit (i.e. 15 alleles out of 34) where DNA alleles matched the reference profile of the donor. This profile would be considered 'up-loadable' to NCIDD. Heterozygote imbalance is observed - a sign of PCR inhibition (alleles highlighted in red).

DNA profiling results for other substrate types

Example of DNA profile generated from an aluminium cartridge case 8 days after print deposition

Figure 5.18 STR-based DNA profile obtained by direct PCR from a handled aluminium cartridge case that was subjected to the environment for 8 days prior to swabbing. A partial DNA profile was obtained (i.e. 21 out of 34 possible alleles) with the AmpFℓSTR® NGM SElect™ PCR amplification kit, where DNA alleles matched the reference profile of the donor. This profile would be considered 'up-loadable' to NCIDD. Peaks were observed that were under the detection threshold and heterozygote imbalance observed at the Amelogenin marker.

DNA profiling results for other substrate types

Example of DNA profile generated from a wooden knife handle 8 days after print deposition

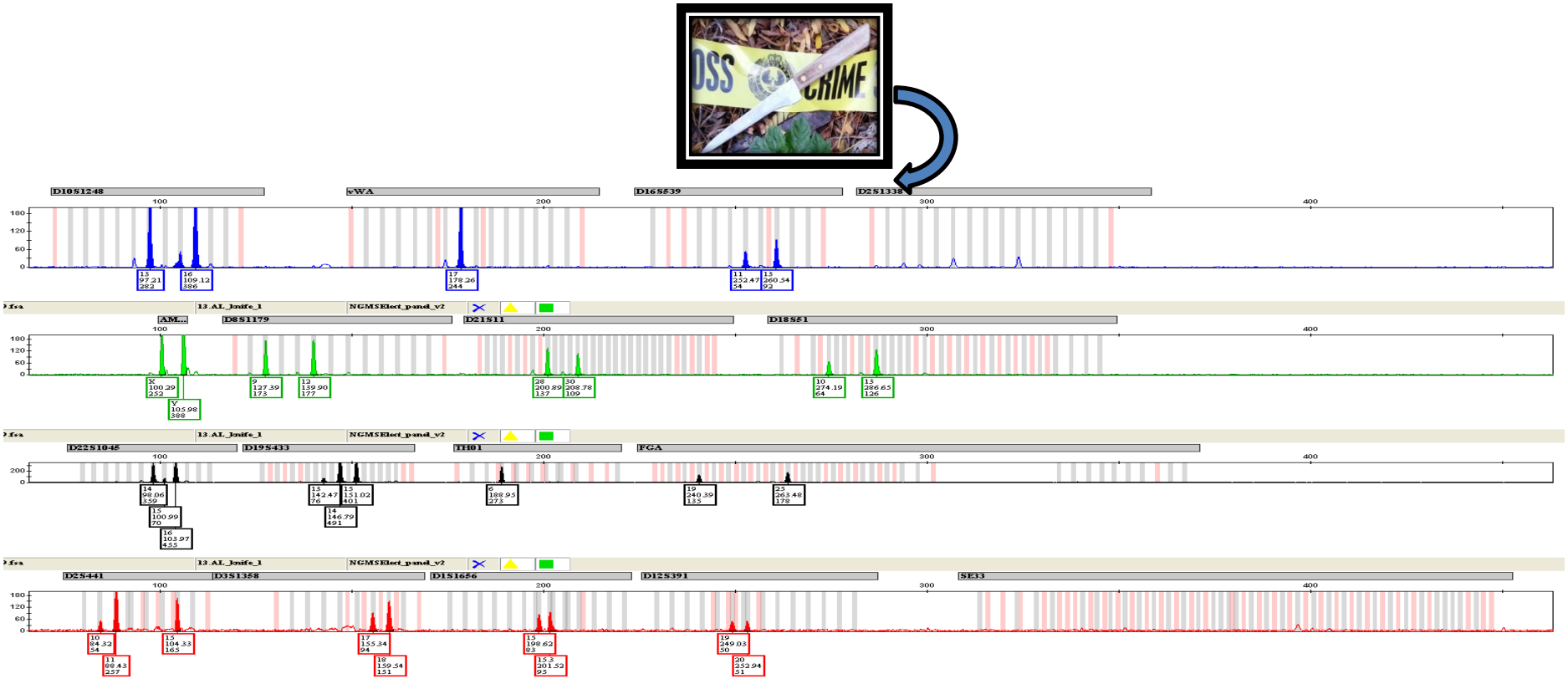


Figure 5.19 STR-based mixed DNA profile obtained by direct PCR after swabbing a wooden knife handle that was subjected to the environment for 8 days after it had been handled. Major alleles were identified as the donor (i.e. 28 DNA alleles that matched with the reference profile) and minor alleles indicated a second contributor. The profile was generating using the AmpFℓSTR® NGM SElect™ PCR amplification kit. This profile would be considered 'up-loadable' to NCIDD and deconvoluted with a mixture analysis software program [53]. 'Drop-in' was observed at loci: D10S1248, D22S1045, D19S433 and D2S441.

DNA profiling results for other substrate types

Example of DNA profile generated from a metal door handle 24 hour after print deposition

(no additional Ampli*Taq* Gold® DNA polymerase was used in the PCR)

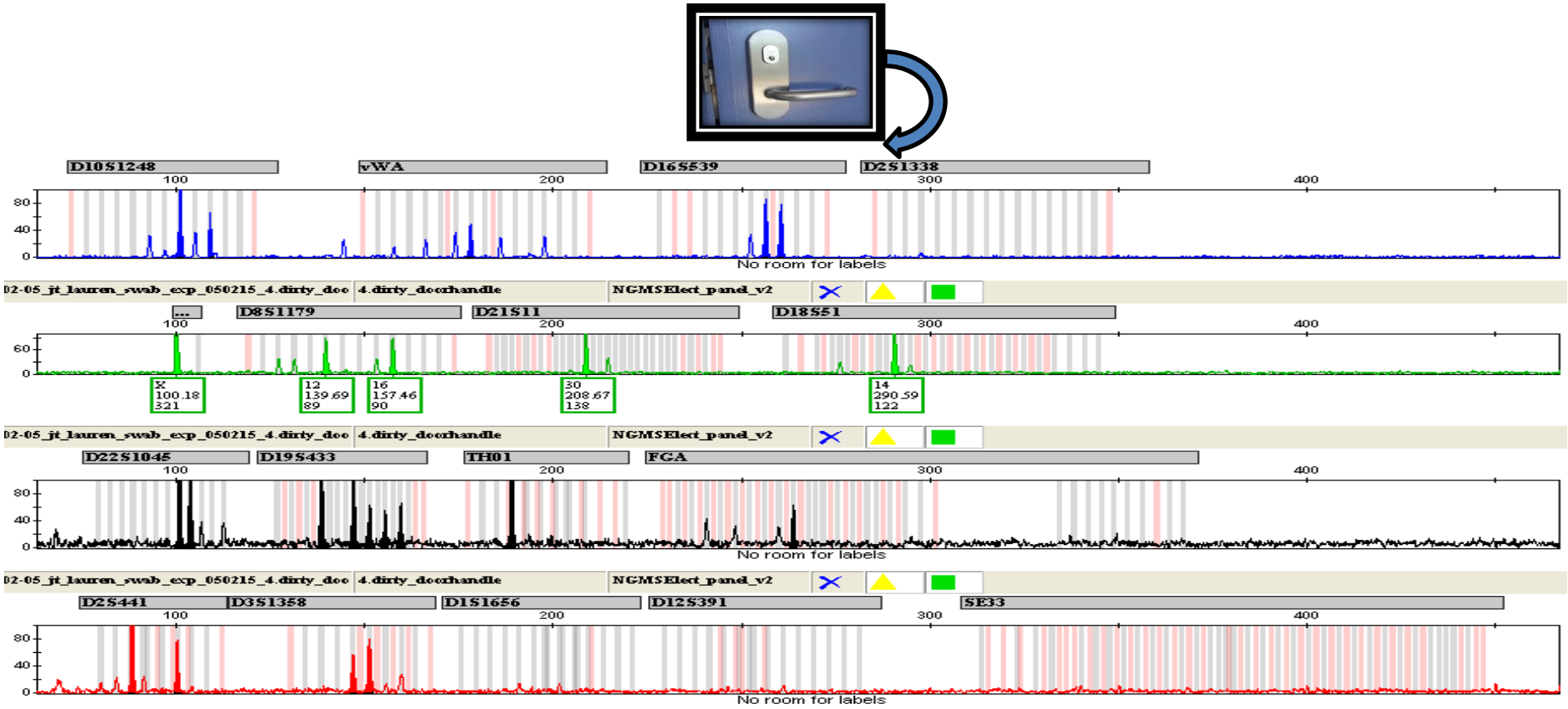


Figure 5.20 Poor quality STR-based DNA profile obtained from swabbing a previously used door handle and subjecting swab fibres to direct amplification with the AmpFℓSTR® NGM SElect™ PCR kit. A mixed DNA profile was observed, as expected. Stochastic effects were observed, such as, allele 'drop-out', heterozygote imbalance, and increased baseline noise.

5.4.3 MinElute® PCR purification

5.4.3.1 Materials and methods

The 'targeted' swabbing method (detailed in Chapter II) was used in conjunction with direct PCR. Individuals were asked to wash their hands and briefly (15 seconds) deposit a fingermark onto a sterile plastic microscope slide (Rinzl plastic, ProSci Tech, QLD, AU) 15 minutes or 30 minutes post hand washing. Fingermarks were dusted with fingermark powders prior to swabbing and the AmpFℓSTR® NGM SElect™ PCR amplification kit (Life Technologies, AU) was used for direct PCR, following manufacturer's recommendations and extra Ampli*Taq* Gold® DNA polymerase 360 (Life Technologies) (1 µL, 5 units) was added to boost PCR. All PCR products were analyzed on a 3130*xl* Genetic Analyzer (Life Technologies) (see Chapter II).

The remaining PCR product was subjected to PCR purification using MinElute® PCR purification columns (Qiagen), following manufacturer's recommendations and DNA eluted in a final 10 µL volume. STR products were analyzed on a 3130*xl* Genetic Analyzer (Life Technologies) in the same conditions described above, with 2 µL of post-purification PCR product run on the CE machine.

5.4.3.2 Results and Discussion

Alternative methods to the LCN process have been explored previously to increase sensitivity and detection, and include post-PCR purification and increased Ampli*Taq* Gold® DNA polymerase [44, 63]. Data in this chapter compare post-PCR purification DNA profiling results for 7 fingerprint samples previously dusted with various fingerprint powders prior to direct PCR. The MinElute® purification method generated the highest yield of PCR product for all powdered fingermarks tested (see SI Table 5.10 – 5.16). However post-PCR purification also produced the highest incidence of disconcordant (i.e. dropped in) alleles (see SI Table 5.11 – 5.16).

5.5 Chapter V Results and Discussion

Direct PCR

The main benefit of direct PCR is its increase in sensitivity, but this approach may lead to the detection of additional alleles. The second part of this data chapter focussed on examining handled substrates that had not been cleaned prior to handling in order to mimic real scenarios and to gain a real understanding of how well samples perform under direct PCR conditions. Published literature observing DNA transfer report that only small amounts of DNA will be transferred immediately after hand washing [46, 64], and in a high percentage of experiments the amount of DNA detected was under the laboratory-specific detection threshold of 40 pg [46]. The reality is that loss of DNA by extraction can be detrimental to 'touch' DNA samples that contain limited DNA to begin with. However, results from this data chapter highlight the ability to generate meaningful DNA profiles (see figures 5.10 – 5.20) using direct PCR from a range of touched substrates after only a short time has passed since hand washing (i.e. 15 minutes); indicating a highly sensitive technique. Regardless, a multitude of factors will contribute to the overall success of DNA profiling.

One criticism of direct PCR is the inability to quantify DNA prior to PCR, preventing the optimised amount of DNA being added to the PCR matrix. The shortcoming of not having a quantification value is of limited practical consequence; sub-optimal 'touch' DNA samples with limited DNA are unlikely to require dilution prior to PCR. Samples of this nature are unlikely to overload the PCR or electrophoresis instruments with too much DNA, and peak heights on the electropherogram can be compared to calibration curves to extrapolate back to starting DNA amounts. Profiles that exhibit saturation can be diluted post capillary electrophoresis and re-run to obtain a higher quality profile if needed.

Interpreting results

Discordant genotypes were observed in this study where the major contributor did not match the profile of the donor (see substrates in Table 5.1); mixed DNA profiles were observed from the gun swabbing events (i.e. more than 2 alleles at loci

detected). The profile example shown (see Figure 5.10) was obtained from the hand-rod area of a revolver where previous handlers DNA would easily adhere and become trapped. Additional data in this chapter report the observation of a mixed DNA profile, with at least 3 contributors, from a door handle that was not cleaned prior to swabbing (see Figure 5.20). Other explanations for the presence of a second or third persons DNA relate to secondary or tertiary transfer events. Secondary and tertiary DNA transfer has been reported in other studies of this nature and the observation of mixed profiles ranged from 10% to 85% [10, 45, 65-68]. Interpreting mixed DNA samples from fingermarks is a challenge, particularly if the DNA is present in trace amounts [69, 70]. Mixed or sub-optimal DNA profiles would benefit from using continuous software programs such as STRmix™ [53] or TrueAllele® [51, 71] to assign alleles that are borderline for detection and help deconvolute mixtures where more than 2 alleles are detected per loci. These experts systems will calculate Bayesian likelihood ratios when comparisons with reference samples are carried out [72], and this is extremely important for evidence evaluation.

Additional Ampli*Taq* Gold® DNA polymerase was added to help overcome PCR inhibition. In one sample that did not have additional Ampli*Taq* Gold® DNA polymerase added stochastic effects could be observed (i.e. heterozygote imbalance/imbalanced loci, increased baseline noise, and split peaks) (see profile example, Figure 5.20). Other samples displayed split peaks (see Figure 5.10) possibly as a result of PCR inhibition as the profiles did not appear overloaded with DNA. If profiles exhibit split peaks as a result of saturation (i.e. too much DNA), the quality of the profile could be enhanced by re-running the sample on the CE machine with diluted PCR product to prevent overloading the capillary with amplified product, or altering the CE injection parameters [73, 74].

Time between DNA deposition and recovery

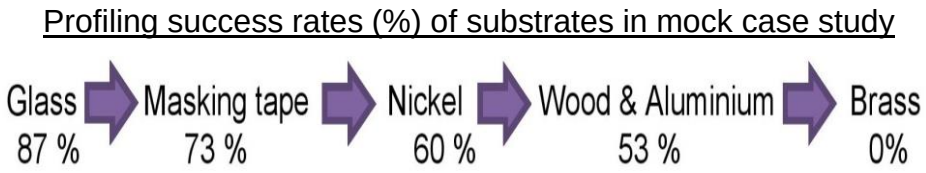
It is still not known how long DNA lasts outside of its protected environment in the cell. In previous research studies, DNA has been recovered from fingerprint samples under various conditions [54, 65]. The persistence of DNA on objects at crime scenes has been investigated, and the amount and quality of DNA has been shown to deteriorate over time [75]. Work carried out to recover DNA using cotton swabs

report 50% reduction in DNA yield 24 hour after print deposition [76]. Proper storage conditions are imperative. In one study, items taken to the laboratory for 'touch' DNA analysis were protected by in-house storage and found to generate DNA profiles 40 days after fingermarks were found [54]. The number of database eligible profiles decreased over time; periods exceeding 10 days significantly reduced the DNA yield [54]. Ideally swabs would be taken from evidence as soon as possible and processed at the time; however, operational laboratory constraints prevent this from happening.

It has been hypothesized that environmental conditions have more of an effect on DNA preservation than time since DNA deposition [77].

Environmental factors

Results from the mock case study in this chapter investigate the potential for direct PCR to amplify DNA recovered from a range of touched objects after exposure to environmental stresses.



One case involved subjecting the items (n=6; glass, masking tape, nickel, aluminium, brass and wood) to the environment for 8 days (see profile examples given in Figures 5.16 – 5.19); DNA was exposed to prolonged rain and UV prior to swabbing and direct PCR. Leaving substrates open to the environment for this period of time appeared to affect the overall success rate for obtaining DNA. For example, informative profiling results were generated for 7 out of 15 sample types tested (i.e. 46 %) - a lower success rate compared to samples left for only 24 hours (i.e. 93 % success) (see manuscript enclosed: **"Templeton, Jennifer EL, *et al*. DNA profiles from fingermarks: A mock case study. *Forensic Science International: Genetics Supplement Series* 5 (2015): e154-e155)"**. The reduced success rate observed from leaving samples for longer is not surprising, as it is well-known that samples exposed to a harsh outdoor environment (e.g. UV light, humidity, heat, water and bacterial growth) are less likely to have surviving endogenous DNA compared to

samples protected within an indoor environment [20, 78]. Moisture accelerates the growth of bacteria and will drastically reduce the ability to retrieve DNA [20]. As demonstrated in other published work, fingerprints that had been submerged in water failed to generate DNA profiles [64]. Results from the mock case study also support the theory that substrate type contributes substantially to the amount of DNA that can be recovered [10, 79-81]. For example, the items chosen to study the substrates nickel, aluminium and brass were cartridge cases. This factor may have an influence on the DNA success rates due to the smooth surface nature and small area available for testing. Other surface materials made of nickel, aluminium and brass may generate different results.

Substrates

Glass demonstrated the highest number of detectable loci recorded (i.e. 87% of profiles were considered 'up-loadable' to the NCIDD, followed by masking tape (73%), nickel (60%), wood and aluminium (53%) and brass (0%). In other preliminary work for handled substrates, brass cartridge casings produced a low success rate with 3 out 16 samples producing partial DNA profiles and the rest failing to produce any DNA alleles. Two profiles obtained from brass casings contained < 12 alleles, and one profile was considered informative (i.e. 12 alleles, plus Amelogenin) (see Figure 5.14). It was observed that the brass cartridge casings were smooth with no pits or grooves for DNA to trap or for skin cells to adhere; objects of a rough nature tend to collect and preserve more DNA as there are more grooves for DNA/skin cells to adhere. On close microscopic examination, aluminium and nickel cartridge cases appeared more textured (i.e. pitted surface) in comparison to brass casings. This could be one reason for the lower success rates seen with brass casings.

It is well-known that cartridge cases in general are difficult surface types to obtain DNA from [82-87]. In particular, it is thought that the chance of obtaining a DNA profile is reduced post-firing [24]. It is also known that metal ions present in brass (i.e. copper and zinc) may inhibit PCR [56, 85]. This could be due in part to the anions in sweat residue being found to corrode the surface of the cartridge casing where the fingermark deposit is present [86]. One study in particular noted that the high temperatures that casings are exposed to during firing will accelerate the corrosion

process, increasing the amount of free metal ions, which is a product of corrosion [85].

One study examining surface texture report a higher yield of DNA retrieved from porous substrates such as wood and fabric compared with glass [10]. Other research investigating the nature of different substrates (i.e. aluminium foils, polythene bags, glass, adhesive tapes and glass) agree that DNA recovery is largely dependent on substrate type [88]. Sub-optimal samples (e.g. fired and unfired cartridge cases) may benefit from protocol enhancement methods (e.g. newer STR profiling kits with more sensitive buffers, or PCR purification using MinElute® columns to concentrate DNA) in order to generate meaningful DNA results.

5.6 Conclusion

The experimental data in this chapter highlight the ability to obtain substantial results from fingermarks - both latent and prints that have been enhanced with common fingerprint powders. The data support the hypothesis that direct PCR can be used on a variety of substrates and surface types. Success can be attributed to the nature of the substrate where DNA was deposited, time of deposition, and the way that swabs are processed and handled (i.e. direct PCR with additional Ampli*Taq* Gold® DNA polymerase), with care to avoid issues with contamination. Data continue to support the theory that cell-free DNA on the surface of skin is a major contributing factor to the PCR matrix. Performing direct PCR on swab fibres tarnished in fingerprint powder indicates a real positive outcome for future analysis; 61% of samples produced meaningful profiles after a short time from hand washing. Post PCR purification increased yield of PCR product (see Appendix – data, SI Table 5.10 – 5.16) which could prove extremely valuable for challenging sample types (e.g. brass or sub-optimal samples that yield few alleles). The only substrate examined in this study that underperformed was the brass cartridge casing, this is most likely due to the small surface area of the casing, the smooth nature of the surface, and potential inhibitors that are present (e.g. zinc and copper).

Projecting forward, future work to target mtDNA and SNPs (ancestral informative and phenotypic) from cartridge cases - that fail to yield nuclear STRs - would benefit the

forensic community. A pre-soak method should be explored further (see Appendix D – data, SI Figures 5.22 – 5.25) where cartridge cases are placed into a 5 mL tube and lysed with Triton-X™ prior to PCR.

For direct PCR, further studies with a larger sample size should be conducted to examine the perseverance of DNA over time. Additional experiments should explore DNA survival in various outdoor conditions and trial other metals and surfaces that bare different chemical reactivity. Furthermore, direct PCR applications would benefit from research into enzyme properties, such as thermostability, processivity, fidelity, specificity and their resistance to inhibitors. The addition of Prep-n-Go™ buffer (ABI) may contribute to a higher yield of DNA when contaminants are present; however, further work is required in order to compare DNA data in the presence and absence of this buffer.

The overall simplicity of the direct PCR technique should enable fast validation and implementation into routine casework. Operational laboratories may change their policy for accepting exhibits if DNA can be routinely obtained from bullet cartridges, steering wheels, triggers, incendiary devices and knife handles that currently have little success.

5.7 References

1. Marshall, P.L., J.L. King, and B. Budowle, *Utility of amplification enhancers in low copy number DNA analysis.* International Journal of Legal Medicine, 2014: p. 1-10.
2. Gill, P., et al., *An investigation of the rigor of interpretation rules for STRs derived from less than 100 pg of DNA.* Forensic Science International, 2000. **112**(1): p. 17-40.
3. Kita, T., et al., *Morphological study of fragmented DNA on touched objects.* Forensic Science International Genetics, 2008. **3**(1): p. 32-6.
4. Lowe, A., et al., *The propensity of individuals to deposit DNA and secondary transfer of low level DNA from individuals to inert surfaces.* Forensic Science International, 2002. **129**(1): p. 25-34.
5. Farmen, R.K., et al., *Assessment of individual shedder status and implication for secondary DNA transfer.* Forensic Science International: Genetics Supplement Series, 2008. **1**(1): p. 415-417.
6. Goray, M., et al., *Secondary DNA transfer of biological substances under varying test conditions.* Forensic Science International: Genetics, 2010. **4**(2): p. 62-67.
7. Phipps, M. and S. Petricevic, *The tendency of individuals to transfer DNA to handled items.* Forensic Science International, 2007. **168**(2-3): p. 162-8.
8. Goray, M., R.A. van Oorschot, and J.R. Mitchell, *DNA transfer within forensic exhibit packaging: Potential for DNA loss and relocation.* Forensic Science International: Genetics, 2012. **6**(2): p. 158-166.
9. Ladd, C., et al., *A systematic analysis of secondary DNA transfer.* Journal of Forensic Sciences, 1999. **44**: p. 1270-1272.
10. Daly, D.J., C. Murphy, and S.D. McDermott, *The transfer of touch DNA from hands to glass, fabric and wood.* Forensic Science International: Genetics, 2012. **6**(1): p. 41-46.
11. Alaeddini, R., *Forensic implications of PCR inhibition - A review.* Forensic Science International Genetics, 2012. **6**(3): p. 297-305.
12. Jennings, R.B., C.E. Ganote, and K.A. Reimer, *Ischemic Tissue Injury.* The American Journal of Pathology, 1975. **81**(1): p. 179.

13. Alaeddini, R., S.J. Walsh, and A. Abbas, *Forensic implications of genetic analyses from degraded DNA - a review.* Forensic Science International: Genetics, 2010. **4**(3): p. 148-157.
14. Allentoft, M.E., et al., *The half-life of DNA in bone: measuring decay kinetics in 158 dated fossils.* Proceedings of the Royal Society B: Biological Sciences, 2012. **279**(1748): p. 4724-4733.
15. Brotherton, P., et al., *Preferential access to genetic information from endogenous hominin ancient DNA and accurate quantitative SNP-typing via SPEX.* Nucleic Acids Research, 2010. **38**(2).
16. van Oven, M. and M. Kayser, *Updated Comprehensive Phylogenetic Tree of Global Human Mitochondrial DNA Variation.* Human Mutation, 2009. **30**(2): p. E386-E394.
17. Briggs, A.W., et al., *Targeted Retrieval and Analysis of Five Neandertal mtDNA Genomes.* Science, 2009. **325**(5938): p. 318-321.
18. Caddy, B., G. Taylor, and A. Linacre, *A review of the science of low template DNA analysis.* Office of the Forensic Regulator, 2008: p. 23-24.
19. Butler, J.M., *Forensic DNA Typing: Biology, Technology, and Genetics of STR Markers*. 2nd Edition ed. 2005, New York: Elsevier Academic Press.
20. Gill-King, H., *Chemical and ultrastructural aspects of decomposition.* Forensic Taphonomy: The Postmortem Fate of Human Remains. CRC, Boca Raton, FL, 1997: p. 93-108.
21. Longo, M.C., M.S. Berninger, and J.L. Hartley, *Use of uracil DNA glycosylase to control carry-over contamination in polymerase chain reactions.* Gene, 1990. **93**(1): p. 125-128.
22. Jacobsen, C.S. and O.F. Rasmussen, *Development and application of a new method to extract bacterial DNA from soil based on separation of bacteria from soil with cation-exchange resin.* Applied and Environmental Microbiology, 1992. **58**(8): p. 2458-2462.
23. Simon, M.C., D.I. Gray, and N. Cook, *DNA Extraction and PCR Methods for the Detection of Listeria monocytogenes in Cold-Smoked Salmon.* Applied and Environmental Microbiology, 1996. **62**(3): p. 822-824.

24. Kumar, P., et al., *Effects of latent fingerprint development reagents on subsequent forensic DNA typing: A review.* Journal of Forensic and Legal Medicine, 2015. **32**: p. 64-69.
25. Shalhoub, R., et al., *The recovery of latent fingermarks and DNA using a silicone-based casting material.* Forensic Science International, 2008. **178**(2): p. 199-203.
26. Frégeau, C.J., O. Germain, and R.M. Fourney, *Fingerprint enhancement revisited and the effects of blood enhancement chemicals on subsequent Profiler Plus™ fluorescent short tandem repeat DNA analysis of fresh and aged bloody fingerprints.* Journal of Forensic Science, 2000. **45**(2): p. 354-380.
27. Yu, P.-H. and M.M. Wallace, *Effect of 1, 2-indanedione on PCR-STR typing of fingerprints deposited on thermal and carbonless paper.* Forensic Science International, 2007. **168**(2): p. 112-118.
28. Lee, H.C., R. Ramotowski, and R. Gaensslen, *Advances in fingerprint technology*. 2001: CRC press.
29. Thamnurak, C., et al., *DNA typing from fluorescent powder dusted latent fingerprints.* Forensic Science International: Genetics Supplement Series, 2011. **3**(1): p. e524-e525.
30. Barbaro, A., et al., *Anonymous letters? DNA and fingerprints technologies combined to solve a case.* Forensic Science International, 2004. **146**: p. S133-S134.
31. Roux, C., et al., *A further study to investigate the effect of fingerprint enhancement techniques on the DNA analysis of bloodstains.* Journal of Forensic Identification, 1999. **49**(4): p. 357.
32. Powell, H., et al., *Proteinase inhibition of the detection of Listeria monocytogenes in milk using the polymerase chain reaction.* Letters in Applied Microbiology, 1994. **18**(1): p. 59-61.
33. Katcher, H. and I. Schwartz, *A distinctive property of Tth DNA polymerase: enzymatic amplification in the presence of phenol.* Biotechniques, 1994. **16**(1): p. 84-92.

34. Rossen, L., et al., *Inhibition of PCR by components of food samples, microbial diagnostic assays and DNA-extraction solutions.* International Journal of Food Microbiology, 1992. **17**(1): p. 37-45.
35. Al-Soud, W.A. and P. Rådström, *Capacity of nine thermostable DNA polymerases to mediate DNA amplification in the presence of PCR-inhibiting samples.* Applied and Environmental Microbiology, 1998. **64**(10): p. 3748-3753.
36. Young, C.C., et al., *Polyvinylpyrrolidone-agarose gel electrophoresis purification of polymerase chain reaction-amplifiable DNA from soils.* Applied and Environmental Microbiology, 1993. **59**(6): p. 1972-1974.
37. Kontanis, E.J. and F.A. Reed, *Evaluation of Real-Time PCR Amplification Efficiencies to Detect PCR Inhibitors.* Journal of Forensic Sciences, 2006. **51**(4): p. 795-804.
38. Opel, K.L., D. Chung, and B.R. McCord, *A Study of PCR Inhibition Mechanisms Using Real Time PCR.* Journal of Forensic Sciences, 2010. **55**(1): p. 25-33.
39. Akane, A., et al., *Identification of the heme compound copurified with deoxyribonucleic acid (DNA) from bloodstains, a major inhibitor of polymerase chain reaction (PCR) amplification.* Journal of Forensic Sciences, 1994. **39**: p. 362-362.
40. Shen, W.-H. and B. Hohn, *DMSO improves PCR amplification of DNA with complex secondary structure.* Trends in Genetics, 1992. **8**(7): p. 228.
41. Park, S.J., et al., *Direct STR Amplification from Whole Blood and Blood-or Saliva-Spotted FTA® without DNA Purification.* Journal of Forensic Sciences, 2008. **53**(2): p. 335-341.
42. Harbison, S., M. Fallow, and D. Bushell, *An analysis of the success rate of 908 trace DNA samples submitted to the Crime Sample Database Unit in New Zealand.* Australian Journal of Forensic Sciences, 2008. **40**(1): p. 49-53.
43. Butler, J.M., et al., *Forensic DNA typing by capillary electrophoresis using the ABI Prism 310 and 3100 genetic analyzers for STR analysis.* Electrophoresis, 2004. **25**(10-11): p. 1397-1412.
44. Smith, P.J. and J. Ballantyne, *Simplified Low-Copy-Number DNA Analysis by Post-PCR Purification.* Journal of Forensic Sciences, 2007. **52**(4): p. 820-829.

45. Wickenheiser, R.A., *Trace DNA: a review, discussion of theory, and application of the transfer of trace quantities of DNA through skin contact.* Journal of Forensic Sciences, 2002. **47**(3): p. 442-450.
46. Alessandrini, F., et al., *Fingerprints as evidence for a genetic profile: morphological study on fingerprints and analysis of exogenous and individual factors affecting DNA typing.* Journal of Forensic Sciences, 2003. **48**(3): p. 586-592.
47. Merriman, B., I. Torrent, and J.M. Rothberg, *Progress in Ion Torrent semiconductor chip based sequencing.* Electrophoresis, 2012. **33**(23): p. 3397-3417.
48. Yeo, Z.X., et al., *Improving Indel Detection Specificity of the Ion Torrent PGM Benchtop Sequencer.* PLOS ONE, 2012. **7**(9): p. e45798.
49. Hedman, J., et al., *Improved forensic DNA analysis through the use of alternative DNA polymerases and statistical modeling of DNA profiles.* Biotechniques, 2009. **47**(5): p. 951-958.
50. Hedman, J., et al., *Synergy between DNA polymerases increases polymerase chain reaction inhibitor tolerance in forensic DNA analysis.* Analytical Biochemistry, 2010. **405**(2): p. 192-200.
51. Duceman, B.W., et al., *Validation Study of the True Allele® Automated Data Review System.* Journal of Forensic Science, 2004. **49**(4): p. 1-8.
52. Taylor, D., *Using continuous DNA interpretation methods to revisit likelihood ratio behaviour.* Forensic Science International: Genetics, 2014. **11**: p. 144-153.
53. Taylor, D., J. A. Bright, and J. Buckleton, *The interpretation of single source and mixed DNA profiles.* Forensic Science International: Genetics, 2013. **7**(5): p. 516-528.
54. Ostojic, L. and E. Wurmbach, *Analysis of fingerprint samples, testing various conditions, for forensic DNA identification.* Science and Justice, 2016.
55. Henle, E.S. and S. Linn, *Formation, prevention, and repair of DNA damage by iron/hydrogen peroxide.* Journal of Biological Chemistry, 1997. **272**(31): p. 19095-19098.

56. Sagripanti, J.-L., P.L. Goering, and A. Lamanna, *Interaction of copper with DNA and antagonism by other metals.* Toxicology and Applied Pharmacology, 1991. **110**(3): p. 477-485.
57. Spinello, A., A. Terenzi, and G. Barone, *Metal complex - DNA binding: Insights from molecular dynamics and DFT/MM calculations.* Journal of Inorganic Biochemistry, 2013. **124**: p. 63-69.
58. Schulz, M. and W. Reichert, *Archived or directly swabbed latent fingerprints as a DNA source for STR typing.* Forensic Science International, 2002. **127**(1): p. 128-130.
59. Lowe, A., et al., *Use of low copy number DNA in forensic inference.* International Congress Series, 2003. **1239**: p. 799-801.
60. Färber, D., et al., *Recovery of latent fingerprints and DNA on human skin.* Journal of Forensic Sciences, 2010. **55**(6): p. 1457-1461.
61. van Hoofstat, D.E., et al., *DNA typing of fingerprints and skin debris: sensitivity of capillary electrophoresis in forensic applications using multiplex PCR.* Proceedings of the 2nd European Symposium of Human Identification, 1998: p. 131-137.
62. Templeton, J. and A. Linacre, *DNA profiles from fingermarks.* BioTechniques, 2014. **57**(5): p. 259.
63. McNevin, D., et al., *Reduced reaction volumes and increased Taq DNA polymerase concentration improve STR profiling outcomes from a real-world low template DNA source: telogen hairs.* Forensic Science, Medicine, and Pathology, 2015. **11**(3): p. 326-338.
64. Jastrzebowska, S., W. Makulec, and J. Jrarnica, *Fingerprint detection and DNA typing on objects recovered from water.* Journal of Forensic Identification, 2007. **57**(5): p. 681.
65. Djuric, M., et al., *DNA typing from handled items.* Forensic Science International: Genetics Supplement Series, 2008. **1**(1): p. 411-412.
66. Cale, C.M., et al., *Could Secondary DNA Transfer Falsely Place Someone at the Scene of a Crime?* Journal of Forensic Sciences, 2016. **61**(1): p. 196-203.
67. Helmus, J., T. Bajanowski, and M. Poetsch, *DNA transfer - a never ending story. A study on scenarios involving a second person as carrier.* International Journal of Legal Medicine, 2016. **130**(1): p. 121-125.

68. van Oorschot, R.A., G. Glavich, and R.J. Mitchell, *Persistence of DNA deposited by the original user on objects after subsequent use by a second person.* Forensic Science International: Genetics, 2014. **8**(1): p. 219-225.
69. Haned, H., K. Slooten, and P. Gill, *Exploratory data analysis for the interpretation of low template DNA mixtures.* Forensic Science International: Genetics, 2012. **6**(6): p. 762-774.
70. Gill, P., et al., *Genotyping and interpretation of STR-DNA: Low-template, mixtures and database matches—Twenty years of research and development.* Forensic Science International: Genetics, 2015.
71. Ballantyne, J., E.K. Hanson, and M.W. Perlin, *DNA mixture genotyping by probabilistic computer interpretation of binomially-sampled laser captured cell populations: combining quantitative data for greater identification information.* Science & Justice, 2013. **53**(2): p. 103-114.
72. Tvedebrink, T., et al., *Estimating the probability of allelic drop-out of STR alleles in forensic genetics.* Forensic Science International: Genetics, 2009. **3**(4): p. 222-226.
73. Budowle, B., et al., *Low copy number—consideration and caution.* Twelfth International Symposium on Human Identification, 2001. **2001**.
74. Forster, L., J. Thomson, and S. Kutranov, *Direct comparison of post-28-cycle PCR purification and modified capillary electrophoresis methods with the 34-cycle "low copy number" (LCN) method for analysis of trace forensic DNA samples.* Forensic Science International: Genetics, 2008. **2**(4): p. 318-328.
75. Raymond, J.J., et al., *Trace evidence characteristics of DNA: a preliminary investigation of the persistence of DNA at crime scenes.* Forensic Science International: Genetics, 2009. **4**(1): p. 26-33.
76. Goray, M., R.J. Mitchell, and R.A. van Oorschot, *Investigation of secondary DNA transfer of skin cells under controlled test conditions.* Legal Medicine, 2010. **12**(3): p. 117-120.
77. Lindahl, T., *Instability and decay of the primary structure of DNA.* Nature, 1993. **362**(6422): p. 709-715.
78. Meakin, G. and A. Jamieson, *DNA transfer: review and implications for casework.* Forensic Science International: Genetics, 2013. **7**(4): p. 434-443.

79. Verdon, T.J., R.J. Mitchell, and R.A.H. van Oorschot, *Evaluating the efficiency of DNA extraction methods from different substrates.* Forensic Science International: Genetics Supplement Series, 2011. **3**(1): p. e93-e94.
80. Swaran, Y.C. and L. Welch, *A comparison between direct PCR and extraction to generate DNA profiles from samples retrieved from various substrates.* Forensic Science International: Genetics, 2012. **6**(3): p. 407-412.
81. Verdon, T.J., R.J. Mitchell, and R.A.H. van Oorschot, *The influence of substrate on DNA transfer and extraction efficiency.* Forensic Science International: Genetics, 2013. **7**(1): p. 167-175.
82. Gashi, B., et al., *Measurement of 9mm cartridge case external temperatures and its forensic application.* Forensic Science International, 2010. **200**(1): p. 21-27.
83. Horsman-Hall, K.M., et al., *Development of STR profiles from firearms and fired cartridge cases.* Forensic Science International: Genetics, 2009. **3**(4): p. 242-250.
84. Polley, D., et al., *An investigation of DNA recovery from firearms and cartridge cases.* Canadian Society of Forensic Science Journal, 2006. **39**(4): p. 217-228.
85. Bond, J.W., *Visualization of latent fingerprint corrosion of metallic surfaces.* Journal of Forensic Sciences, 2008. **53**(4): p. 812-822.
86. Hong, S. and A. Han, *Development of fingermark on the surface of fired cartridge casing using amino acid sensitive reagents: Change of viewpoint.* Forensic Science International, 2016. **266**: p. 86-90.
87. Montpetit, S. and P. O'Donnell, *An optimized procedure for obtaining DNA from fired and unfired ammunition.* Forensic Science International: Genetics, 2015. **17**: p. 70-74.
88. Balogh, M.K., et al., *STR genotyping and mtDNA sequencing of latent fingerprint on paper.* Forensic Science International, 2003. **137**(2-3): p. 188-195.

Appendices - Chapter V

Appendix A - Supplementary Information – preliminary results

Effect of fingerprint powders on direct PCR

STR-based DNA profiling results - preliminary work to supplement manuscript data

Effect of white fingerprint powder on the direct PCR approach

SI Table 5.1 STR-based DNA profiling results for fingerprint powder inhibition test, using control DNA at 1 ng as template for amplification in the presence of white fingerprint powder. Swab fibres (~ 2 mm^2) were also present in the PCR tube to measure the inhibitory effect.

GENOTYPE PROFILE				
Sample ID: WHITE_POWDER_1ng_DNA				
Date of run: 27-1-2015				
Sample run name: JT_POWDER_INHIBITION TEST_270115				
Genetic loci	Allele	RFU	Allele	RFU
D10	13	2496	15	2336
VWA	16	1400	19	2144
D16	9	1878	13	1866
D2S1	22	1669	25	1732
Amelogenin	X	2790	Y	2468
D8	14	3144	15	3231
D21	29	2460	31.2	2779
D18	16	2098	18	2367
D22	16	3998	16	3998
D19	13	2739	14	2886
THO1	6	2206	9.3	1971
FGA	20	1778	23	2685
D2S4	10	831	14	1017
D3	17	1066	18	949
D1	12	1037	13	1070
D12	18	803	23	735
SE33	15	1065	16	1082

Full DNA profile generated (i.e. 34 alleles out of 34 alleles), for AmpFℓSTR® NGM SElect™ PCR amplification kit.Average relative fluorescent unit of peak heights is 1963 RFU.

Effect of silver fingerprint powder on the direct PCR approach

SI Table 5.2 STR-based DNA profiling results for fingerprint powder inhibition test, using control DNA at 1 ng as template for amplification in the presence of silver aluminium powder. Swab fibres (~ 2 mm^2) were also present in the PCR tube to measure the inhibitory effect.

GENOTYPE PROFILE				
Sample ID: SILVER_ALUMINIUM _POWDER_1ng_DNA				
Date of run: 27-1-2015				
Sample run name: JT_POWDER_INHIBITION TEST_270115				
Genetic loci	Allele	RFU	Allele	RFU
D10	13	2729	15	2867
VWA	16	2172	19	2176
D16	9	3049	13	2884
D2S1	22	1908	25	2129
Amelogenin	X	4150	Y	2717
D8	14	3147	15	3791
D21	29	4038	31.2	2982
D18	16	3377	18	3596
D22	16	6298	16	6298
D19	13	3752	14	2501
THO1	6	4271	9.3	3248
FGA	20	2230	23	2254
D2S4	10	1173	14	1485
D3	17	1733	18	1356
D1	12	1467	13	1239
D12	18	1061	23	783
SE33	15	1347	16	1139

Full DNA profile generated (i.e. 34 alleles out of 34 alleles), for AmpFℓSTR® NGM SElect™ PCR amplification kit.

Average relative fluorescent unit of peak heights is 2577 RFU.

Effect of black fingerprint powder on the direct PCR approach

SI Table 5.3 STR-based DNA profiling results for fingerprint powder inhibition test, using control DNA at 1 ng as template for amplification in the presence of black powder. Swab fibres (~ 2 mm^2) were also present in the PCR tube to measure the inhibitory effect.

GENOTYPE PROFILE				
Sample ID: BLACK_POWDER_1ng_DNA				
Date of run: 27-1-2015				
Sample run name: JT_POWDER_INHIBITION TEST_270115				
Genetic loci	Allele	RFU	Allele	RFU
D10	13	1909	15	1887
VWA	16	2481	19	2109
D16	9	2269	13	2466
D2S1	22	1748	25	1735
Amelogenin	X	2834	Y	2848
D8	14	3058	15	2331
D21	29	2319	31.2	2194
D18	16	2227	18	2047
D22	16	5021	16	5021
D19	13	2429	14	1661
THO1	6	2812	9.3	3010
FGA	20	2044	23	1675
D2S4	10	978	14	830
D3	17	889	18	888
D1	12	1035	13	632
D12	18	640	23	590
SE33	15	701	16	627

Full DNA profile generated (i.e. 34 alleles out of 34 alleles), for AmpFℓSTR® NGM SElect™ PCR amplification kit.

Average relative fluorescent unit of peak heights is 1906 RFU.

Effect of black magnetic fingerprint powder on the direct PCR approach

SI Table 5.4 STR-based DNA profiling results for fingerprint powder inhibition test, using control DNA at 1 ng as template for amplification in the presence of black magnetic powder. Swab fibres (~ 2 mm^2) were also present in the PCR tube to measure the inhibitory effect.

GENOTYPE PROFILE				
Sample ID: MAGNETIC_BLACK_POWDER_1ng_DNA				
Date of run: 27-1-2015				
Sample run name: JT_POWDER_INHIBITION TEST_270115				
Genetic loci	Allele	RFU	Allele	RFU
D10	13	2303	15	2116
VWA	16	2290	19	2314
D16	9	2078	13	2330
D2S1	22	1797	25	1604
Amelogenin	X	2284	Y	2472
D8	14	3217	15	3083
D21	29	2241	31.2	2329
D18	16	2222	18	2050
D22	16	5520	16	5520
D19	13	2293	14	2197
THO1	6	2969	9.3	2427
FGA	20	1675	23	1382
D2S4	10	1200	14	1209
D3	17	854	18	898
D1	12	883	13	875
D12	18	728	23	725
SE33	15	764	16	876

Full DNA profile generated (i.e. 34 alleles out of 34 alleles), for AmpFℓSTR® NGM SElect™ PCR amplification kit.

Average relative fluorescent unit of peak heights is 1945 RFU.

Effect of magnetic red fingerprint powder on the direct PCR approach

SI Table 5.5 STR-based DNA profiling results for fingerprint powder inhibition test, using control DNA at 1 ng as template for amplification in the presence of red magnetic powder. Swab fibres (~ 2 mm^2) were also present in the PCR tube to measure the inhibitory effect.

GENOTYPE PROFILE				
Sample ID: MAGNETIC_RED _POWDER_1ng_DNA				
Date of run: 27-1-2015				
Sample run name: JT_POWDER_INHIBITION TEST_270115				
Genetic loci	Allele	RFU	Allele	RFU
D10	13	1758	15	1350
VWA	16	1802	19	1231
D16	9	1330	13	1397
D2S1	22	1091	25	940
Amelogenin	X	2268	Y	2313
D8	14	2284	15	2385
D21	29	1457	31.2	1589
D18	16	1595	18	1353
D22	16	3245	16	3245
D19	13	1813	14	1485
THO1	6	1932	9.3	1758
FGA	20	1130	23	1347
D2S4	10	795	14	682
D3	17	742	18	541
D1	12	804	13	717
D12	18	540	23	606
SE33	15	430	16	390

Full DNA profile generated (i.e. 34 alleles out of 34 alleles), for AmpFℓSTR® NGM SElect™ PCR amplification kit.

Average relative fluorescent unit of peak heights is 1366 RFU.

Effect of red fingerprint powder on the direct PCR approach

SI Table 5.6 STR-based DNA profiling results for fingerprint powder inhibition test, using control DNA at 1 ng as template for amplification in the presence of red powder. Swab fibres (~ 2 mm^2) were also present in the PCR tube to measure the inhibitory effect.

GENOTYPE PROFILE				
Sample ID: RED_POWDER_1ng_DNA				
Date of run: 27-1-2015				
Sample run name: JT_POWDER_INHIBITION TEST_270115				
Genetic loci	Allele	RFU	Allele	RFU
D10	13	2619	15	2858
VWA	16	3280	19	2763
D16	9	3376	13	2782
D2S1	22	2220	25	2573
Amelogenin	X	5815	Y	4870
D8	14	5609	15	5031
D21	29	4452	31.2	4291
D18	16	3367	18	3764
D22	16	5893	16	5893
D19	13	3600	14	3410
THO1	6	5002	9.3	5377
FGA	20	1981	23	2174
D2S4	10	1679	14	2078
D3	17	1824	18	1792
D1	12	1743	13	1947
D12	18	1469	23	1163
SE33	15	1283	16	1112

Full DNA profile generated (i.e. 34 alleles out of 34 alleles), for AmpFℓSTR® NGM SElect™ PCR amplification kit.

Average relative fluorescent unit of peak heights is 3127 RFU.

Effect of green fingerprint powder on the direct PCR approach

SI Table 5.7 STR-based DNA profiling results for fingerprint powder inhibition test, using control DNA at 1 ng as template for amplification in the presence of green powder. Swab fibres (~ 2 mm^2) were also present in the PCR tube to measure the inhibitory effect.

GENOTYPE PROFILE				
Sample ID: GREEN_POWDER_1ng_DNA				
Date of run: 27-1-2015				
Sample run name: JT_POWDER_INHIBITION TEST_270115				
Genetic loci	Allele	RFU	Allele	RFU
D10	13	2402	15	1999
VWA	16	2699	19	2140
D16	9	2719	13	2220
D2S1	22	1994	25	2067
Amelogenin	X	4553	Y	2773
D8	14	3977	15	3406
D21	29	3348	31.2	2850
D18	16	3574	18	3165
D22	16	6547	16	6547
D19	13	4312	14	3773
THO1	6	5250	9.3	4330
FGA	20	2612	23	2575
D2S4	10	1322	14	1592
D3	17	1285	18	1626
D1	12	1470	13	1413
D12	18	1178	23	966
SE33	15	934	16	884

Full DNA profile generated (i.e. 34 alleles out of 34 alleles), for AmpFℓSTR® NGM SElect™ PCR amplification kit.

Average relative fluorescent unit of peak heights is 2665 RFU.

Effect of magnetic green fingerprint powder on the direct PCR approach

SI Table 5.8 STR-based DNA profiling results for fingerprint powder inhibition test, using control DNA at 1 ng as template for amplification in the presence of green magnetic powder. Swab fibres (~ 2 mm^2) were also present in the PCR tube to measure the inhibitory effect.

GENOTYPE PROFILE				
Sample ID: MAGNETIC_GREEN _POWDER_1ng_DNA				
Date of run: 27-1-2015				
Sample run name: JT_POWDER_INHIBITION TEST_270115				
Genetic loci	Allele	RFU	Allele	RFU
D10	13	2479	15	2198
VWA	16	2673	19	1704
D16	9	2332	13	2319
D2S1	22	2436	25	1470
Amelogenin	X	3010	Y	2883
D8	14	3310	15	3009
D21	29	3245	31.2	2975
D18	16	2334	18	2505
D22	16	4890	16	4890
D19	13	2685	14	2051
THO1	6	3254	9.3	3111
FGA	20	1977	23	2142
D2S4	10	889	14	982
D3	17	1154	18	936
D1	12	1097	13	1089
D12	18	742	23	909
SE33	15	673	16	821

Full DNA profile generated (i.e. 34 alleles out of 34 alleles), for AmpFℓSTR® NGM SElect™ PCR amplification kit.

Average relative fluorescent unit of peak heights is 2129 RFU.

Positive control DNA results

SI Table 5.9 Control STR-based DNA profiling results from 1 ng of control DNA. Swab fibres (~ 2 mm^2) were also present in the PCR tube to measure the inhibitory effect.

GENOTYPE PROFILE				
Sample ID: POSITIVE CONTROL (NO POWDER)				
Date of run: 27-1-2015				
Sample run name: JT_POWDER_INHIBITION TEST_270115				
Genetic loci	Allele	RFU	Allele	RFU
D10	13	2437	15	2419
VWA	16	3766	19	2273
D16	9	3305	13	3112
D2S1	22	3131	25	2201
Amelogenin	X	3623	Y	3293
D8	14	4632	15	4193
D21	29	3486	31.2	3269
D18	16	2947	18	4405
D22	16	6914	16	6914
D19	13	3801	14	3332
THO1	6	3884	9.3	3526
FGA	20	2813	23	2910
D2S4	10	1405	14	1820
D3	17	1465	18	1188
D1	12	1420	13	1374
D12	18	1176	23	1163
SE33	15	1587	16	1234

Full DNA profile generated (i.e. 34 alleles out of 34 alleles), for AmpFℓSTR® NGM SElect™ PCR amplification kit.

Average relative fluorescent unit of peak heights is 2833 RFU.

Appendix B - Supplementary Information

MinElute® PCR purification test results

SI Table 5.10 STR-based DNA profiling results for **MinElute® PCR purification tests**. Fingermarks were deposited and dusted with fingerprint powder prior to swabbing and direct PCR. Swab fibres (~ 2 mm^2) were present in the PCR tube. PCR product was analysed pre- and post-MinElute® PCR purification to assess the difference in RFU value of peak heights and the quality of associated profiles.

GENOTYPE PROFILE				**GENOTYPE PROFILE**			
Sample ID: JT_Powders_18/11/14				**Sample ID: JT_Powders_18/11/14**			
Date of run: 18/11/14				**Date of run: 18/11/14**			
Sample run name: Thumb_15 min_BLACK_powder PRE-MINELUTE PURIFICATION				**Sample run name: Thumb_15 min_BLACK_powder POST-MINELUTE PURIFICATION**			
Genetic loci	Allele (RFU)	Allele (RFU)	Additional alleles (RFU)	**Genetic loci**	Allele (RFU)	Allele (RFU)	Additional alleles (RFU)
D10	F	16 (55)	-	**D10**	14 (66)	16 (288)	
VWA	F	F	-	**VWA**	F	F	
D16	F	F	-	**D16**	F	F	
D2S1	F	F	-	**D2S1**	F	F	
Amelogenin	F	F	-	**Amelogenin**	X (107)	F	
D8	F	13 (54)	-	**D8**	10 (109)	13 (164)	
D21	F	F	-	**D21**	29 (109)	F	
D18	F	F	-	**D18**	13 (166)	F	
D22	15 (79)	F	-	**D22**	15 (349)	15	
D19	15 (50)	F	-	**D19**	15 (211)	15	
THO1	F	F	-	**THO1**	F	9.3 (124)	
FGA	F	F	-	**FGA**	20 (109)	21 (126)	
D2S4	F	F	-	**D2S4**	F	F	
D3	F	F	-	**D3**	15 (82)	F	
D1	F	F	-	**D1**	13 (65)	F	
D12	F	F	-	**D12**	F	F	
SE33	F	F	-	**SE33**	F	F	

N.B. Relative fluorescent units highlighted in brackets in blue indicate the highest RFU values for all samples tested. Allele drop-out is denoted by 'F'.

SI Table 5.11 STR-based DNA profiling results for **MinElute® PCR purification tests**. Fingermarks were deposited and dusted with fingerprint powder, prior to swabbing and direct PCR. Swab fibres (~ 2 mm^2) were present in the PCR tube. PCR product was analysed pre- and post-MinElute® PCR purification to assess the difference in RFU value of peak heights and the quality of associated profiles.

GENOTYPE PROFILE				**GENOTYPE PROFILE**			
Sample ID: JT_Powders_18/11/14				**Sample ID: JT_Powders_18/11/14**			
Date of run: 18/11/14				**Date of run: 18/11/14**			
Sample run name: Index_15 min_White_powder **PRE-MINELUTE PURIFICATION**				**Sample run name: Index_15 min_White_powder** **POST-MINELUTE PURIFICATION**			
Genetic loci	Allele (RFU)	Allele (RFU)	Additional alleles (RFU)	**Genetic loci**	Allele (RFU)	Allele (RFU)	Additional alleles (RFU)
D10	14 (91)	16 (134)		**D10**	14 (804)	16 (1187)	
VWA	F	17 (56)		**VWA**	16 (142)	17 (519)	18 (198)
D16	F	13 (96)		**D16**	12 (422)	13 (889)	
D2S1	25 (72)	F		**D2S1**	25 (700)	F	
Amel	X (194)	X		**Amel**	X (1477)	X	
D8	10 (143)	13 (152)		**D8**	10 (1192)	13 (1182)	9 (388), 12 (346), 14 (309)
D21	29 (84)	31 (105)	25 (51)	**D21**	29 (667)	31 (864)	23.2 (339), 25 (362)
D18	F	14 (77)		**D18**	13 (224)	14 (642)	
D22	15 (301)	15		**D22**	15 (2602)	15	
D19	15 (442)	15		**D19**	15 (3713)	15	
THO1	7 (66)	F		**THO1**	7 (517)	9.3 (142)	12 (93)
FGA	20 (90)	21 (96)		**FGA**	20 (682)	21 (775)	
D2S4	F	F		**D2S4**	11 (384)	F	
D3	F	F		**D3**	15 (353)	16 (182)	
D1	F	F		**D1**	13 (400)	16.3 (271)	
D12	F	F		**D12**	17 (210)	22 (328)	
SE33	F	F		**SE33**	16 (164)	F	

N.B. Relative fluorescent units highlighted in brackets in blue indicate the highest RFU values for all samples tested. Allele drop-out is denoted by 'F'.

SI Table 5.12 STR-based DNA profiling results for **MinElute® PCR purification tests**. Fingermarks were deposited and dusted with fingerprint powder, prior to swabbing and direct PCR. Swab fibres (~ 2 mm^2) were present in the PCR tube. PCR product was analysed pre- and post-MinElute® PCR purification to assess the difference in RFU value of peak heights and the quality of associated profiles.

GENOTYPE PROFILE				**GENOTYPE PROFILE**			
Sample ID: JT_Powders_18/11/14				**Sample ID: JT_Powders_18/11/14**			
Date of run: 18/11/14				**Date of run: 18/11/14**			
Sample run name: Middle_15min_SILVER **PRE-MINELUTE PURIFICATION**				**Sample run name: Middle_15min_SILVER** **POST-MINELUTE PURIFICATION**			
Genetic loci	Allele (RFU)	Allele (RFU)	Additional alleles (RFU)	**Genetic loci**	Allele (RFU)	Allele (RFU)	Additional alleles (RFU)
D10	14 (134)	16 (162)	-	**D10**	14 (782)	16 (939)	-
VWA	16 (142)	17 (114)	-	**VWA**	16 (808)	17 (662)	-
D16	12 (51)	F	-	**D16**	12 (314)	13 (167)	-
D2S1	F	F	-	**D2S1**	25 (42)	F	-
Amel	X (537)	X	-	**Amel**	X (2543)	X	Y (126)
D8	10 (141)	13 (140)	-	**D8**	10 (700)	13 (714)	-
D21	F	F	-	**D21**	29 (123)	31 (80)	-
D18	13 (174)	14 (191)	-	**D18**	13 (929)	14 (1021)	-
D22	15 (342)	15	-	**D22**	15 (1766)	15	-
D19	15 (61)	F	-	**D19**	15 (399)	15	-
THO1	F	F	-	**THO1**	9.3 (106)	F	-
FGA	F	21 (125)	-	**FGA**	20 (209)	21 (599)	-
D2S4	11 (119)	F	-	**D2S4**	F	F	-
D3	15 (66)	16 (67)	-	**D3**	15 (265)	16 (251)	13 (74)
D1	13 (77)	16.3 (80)	-	**D1**	13 (305)	16.3 (312)	-
D12	17 (78)	F	-	**D12**	17 (320)	F	-
SE33	16 (72)	F	-	**SE33**	16 (364)	F	20 (63), 27.2 (113)

N.B. Relative fluorescent units highlighted in brackets in blue indicate the highest RFU values for all samples tested. Allele drop-out is denoted by 'F'.

SI Table 5.13 STR-based DNA profiling results for **MinElute® PCR purification tests**. Fingermarks were deposited and dusted with fingerprint powder, prior to swabbing and direct PCR. Swab fibres (~ 2 mm^2) were present in the PCR tube. PCR product was analysed pre- and post-MinElute® PCR purification to assess the difference in RFU value of peak heights and the quality of associated profiles.

GENOTYPE PROFILE				**GENOTYPE PROFILE**			
Sample ID: JT_Powders_18/11/14				**Sample ID: JT_Powders_18/11/14**			
Date of run: 18/11/14				**Date of run: 18/11/14**			
Sample run name: Thumb_30min_black **PRE-MINELUTE PURIFICATION**				**Sample run name: Thumb_30min_black** **POST-MINELUTE PURIFICATION**			
Genetic loci	Allele (RFU)	Allele (RFU)	Additional alleles (RFU)	**Genetic loci**	Allele (RFU)	Allele (RFU)	Additional alleles (RFU)
D10	14 (204)	16 (107)		**D10**	14 (1059)	16 (572)	
VWA	16 (101)	17 (98)		**VWA**	16 (507)	17 (492)	
D16	12 (58)	F		**D16**	12 (295)	F	
D2S1	25 (87)	F		**D2S1**	25 (479)	25	
Amel	X (418)	X		**Amel**	X (1867)	X	
D8	10 (251)	13 (148)		**D8**	10 (1193)	13 (669)	
D21	29 (79)	F		**D21**	29 (409)	31 (197)	
D18	13 (88)	14 (57)		**D18**	13 (446)	14 (274)	
D22	15 (256)	15		**D22**	15 (1383)	15	16 (232)
D19	15 (181)	15		**D19**	15 (924)	15	
THO1	F	9.3 (71)		**THO1**	7 (144)	9.3 (328)	
FGA	20 (191)	21 (74)		**FGA**	20 (920)	21 (334)	
D2S4	11 (126)	F		**D2S4**	11 (204)	F	
D3	15 (69)	16 (51)		**D3**	15 (237)	16 (167)	
D1	F	F		**D1**	13 (178)	F	
D12	F	F		**D12**	17 (171)	22 (81)	
SE33	F	F		**SE33**	16 (78)	27.2 (81)	

N.B. Relative fluorescent units highlighted in brackets in blue indicate the highest RFU values for all samples tested. Allele drop-out is denoted by 'F'.

SI Table 5.14 STR-based DNA profiling results for **MinElute® PCR purification tests**. Fingermarks were deposited and dusted with fingerprint powder, prior to swabbing and direct PCR. Swab fibres (~ 2 mm^2) were present in the PCR tube. PCR product was analysed pre- and post-MinElute® PCR purification to assess the difference in RFU value of peak heights and the quality of associated profiles.

GENOTYPE PROFILE				**GENOTYPE PROFILE**			
Sample ID: JT_Powders_18/11/14				**Sample ID: JT_Powders_18/11/14**			
Date of run: 18/11/14				**Date of run: 18/11/14**			
Sample run name: Index_30min_white **PRE-MINELUTE PURIFICATION**				**Sample run name: Index_30min_white** **POST-MINELUTE PURIFICATION**			
Genetic loci	Allele (RFU)	Allele (RFU)	Additional alleles (RFU)	**Genetic loci**	Allele (RFU)	Allele (RFU)	Additional alleles (RFU)
D10	14 (314)	16 (281)	13 (118)	**D10**	14 (1431)	16 (1269)	
VWA	16 (374)	17 (533)		**VWA**	16 (1688)	17 (2398)	
D16	12 (368)	13 (349)		**D16**	12 (1857)	13 (1781)	14 (268)
D2S1	25 (211)	25		**D2S1**	25 (1151)	25	
Amelogenin	X (1351)	X		**Amelogenin**	X (5662)	X	
D8	10 (523)	13 (267)		**D8**	10 (2266)	13 (1170)	
D21	29 (492)	31 (385)	30 (80), 33 (81)	**D21**	29 (2205)	31 (1778)	30 (368), 32.2(183), 33 (373)
D18	13 (289)	14 (343)		**D18**	13 (1438)	14 (1692)	17 (213)
D22	15 (881)	15		**D22**	15 (3841)	15	
D19	15 (963)	15		**D19**	15 (3677)	15	
THO1	7 (299)	9.3 (262)		**THO1**	7 (1227)	9.3 (1149)	
FGA	20 (324)	21 (388)		**FGA**	20 (1263)	21 (1491)	
D2S4	11 (274)	11.3 (149)		**D2S4**	11 (662)	F	
D3	15 (204)	16 (173)		**D3**	15 (679)	16 (642)	
D1	13 (173)	16.3 (118)		**D1**	13 (541)	16.3 (363)	
D12	17 (165)	22 (110)		**D12**	17 (527)	22 (356)	
SE33	16 (65)	27.2 (82)		**SE33**	16 (301)	27.2 (349)	

N.B. Relative fluorescent units highlighted in brackets in blue indicate the highest RFU values for all samples tested. Allele drop-out is denoted by 'F'.

SI Table 5.15 STR-based DNA profiling results for **MinElute® PCR purification tests**. Fingermarks were deposited and dusted with fingerprint powder, prior to swabbing and direct PCR. Swab fibres (~ 2 mm^2) were present in the PCR tube. PCR product was analysed pre- and post-MinElute® PCR purification to assess the difference in RFU value of peak heights and the quality of associated profiles.

GENOTYPE PROFILE				**GENOTYPE PROFILE**			
Sample ID: JT_Powders_18/11/14				**Sample ID: JT_Powders_18/11/14**			
Date of run: 18/11/14				**Date of run: 18/11/14**			
Sample run name: Middle_30min_silver **PRE-MINELUTE PURIFICATION**				**Sample run name: Middle_30min_silver** **POST-MINELUTE PURIFICATION**			
Genetic loci	Allele (RFU)	Allele (RFU)	Additional alleles (RFU)	**Genetic loci**	Allele (RFU)	Allele (RFU)	Additional alleles (RFU)
D10	14 (188)	16 (99)		**D10**	14 (1823)	16 (1000)	17 (389)
VWA	F	F		**VWA**	16 (189)	17 (453)	14 (67), 19 (106)
D16	12 (57)	F		**D16**	12 (646)	13 (208)	
D2S1	F	F		**D2S1**	25 (439)	F	24 (109)
Amel	X (345)	X		**Amel**	X (2463)	X	
D8	10 (119)	13 (237)		**D8**	10 (860)	13 (1786)	
D21	F	F		**D21**	29 (147)	31 (224)	30 (53)
D18	13 (85)	F		**D18**	13 (750)	14 (363)	23 (138)
D22	15 (209)	15	16 (51)	**D22**	15 (1679)	15	16 (354)
D19	15 (86)	F		**D19**	15 (772)	15	14.2 (258)
THO1	F	F		**THO1**	7 (342)	9.3 (93)	4 (209), 8 (246)
FGA	20 (52)	21 (89)		**FGA**	20 (359)	21 (756)	
D2S4	11 (73)	11.3 (57)	10 (60)	**D2S4**	11 (327)	F	
D3	15 (103)	16 (67)		**D3**	15 (556)	16 (315)	
D1	F	16.3 (55)		**D1**	13 (242)	16.3 (297)	
D12	F	F		**D12**	17 (116)	22 (124)	
SE33	F	F		**SE33**	16 (164)	27.2 (101)	

N.B. Relative fluorescent units highlighted in brackets in blue indicate the highest RFU values for all samples tested. Allele drop-out is denoted by 'F'.

SI Table 5.16 STR-based DNA profiling results for **MinElute® PCR purification tests**. Fingermarks were deposited and dusted with fingerprint powder, prior to swabbing and direct PCR. Swab fibres (~ 2 mm^2) were present in the PCR tube. PCR product was analysed pre and post-MinElute® PCR purification to assess the difference in RFU value of peak heights and the quality of associated profiles.

GENOTYPE PROFILE				**GENOTYPE PROFILE**			
Sample ID: JT_Powders_18/11/14				**Sample ID: JT_Powders_18/11/14**			
Date of run: 18/11/14				**Date of run: 18/11/14**			
Sample run name: Ring_30min_MagneticBlack **PRE-MINELUTE PURIFICATION**				**Sample run name: Ring_30min_MagneticBlack** **POST-MINELUTE PURIFICATION**			
Genetic loci	Allele (RFU)	Allele (RFU)	Additional alleles (RFU)	**Genetic loci**	Allele (RFU)	Allele (RFU)	Additional alleles (RFU)
D10	F	F		**D10**	14 (123)	16 (154)	
VWA	F	F		**VWA**	16 (112)	17 (50)	19 (53)
D16	F	F		**D16**	F	F	
D2S1	F	F		**D2S1**	F	F	
Amel	X (94)	F		**Amel**	X (334)	X	
D8	10 (86)	13 (61)		**D8**	10 (370)	13 (225)	
D21	F	F		**D21**	29 (165)	31 (132)	
D18	F	F		**D18**	F	F	
D22	15 (98)	F		**D22**	15 (370)	F	
D19	F	F		**D19**	15 (96)	F	
THO1	F	F		**THO1**	7 (146)	F	
FGA	F	F		**FGA**	20 (79)	F	
D2S4	F	F		**D2S4**	11 (96)	F	
D3	F	F		**D3**	F	F	
D1	F	F		**D1**	F	F	
D12	F	F		**D12**	F	F	
SE33	F	F		**SE33**	F	F	

N.B. Relative fluorescent units highlighted in brackets in blue indicate the highest RFU values for all samples tested. Allele drop-out is denoted by 'F'.

Appendix C - Supplementary Information

In order for a new forensic DNA test to be accepted in an accredited laboratory a series of validation tests are required in order to assess:

1) Reliability - a measure of consistency to ensure accuracy and results reflect the sample being tested;

2) Robustness - quality assurance methodology to measure precision. A method is considered robust when there is high success from a large number of samples, and only a few samples need to be repeated due to limitations in the method (e.g. low success rates from degraded/low-level DNA).

3) Reproducibility - to determine if the same, or very similar, results would be obtained each time the sample is tested.

For the peer-reviewed publications included in this thesis the reliability, robustness and reproducibility of the direct PCR approach was assessed. The direct PCR methodology proved to be reliable at consistently generating DNA profiles from glass, plastic, and metal substrates as demonstrated in the publication in Chapter III: "*"Direct PCR improves the recovery of DNA from various substrates"*. Reliability was measured by repeated sampling of the same individuals and the use of donor's reference profiles to ensure concordant genotypes were obtained and limited stochastic effects. The limit of sensitivity of the direct PCR approach was assessed to measure the robustness of the technique. A series of dilution experiments can define the limitations and sensitivity of a test and determine the optimum quantity of DNA that is required to generate a profile. From the data, the lowest mass of DNA deposited that yielded an informative DNA profile using direct PCR was 0.1 ng for plastic, 0.2 ng for glass, and 0.5 ng for brass substrates. From data chapters IV and V, reproducibility was assessed by repeated sampling of the same individuals to measure overall success rates with the direct PCR approach. Variables consisted of: a) time since hand washing; b) substrates used to deposit DNA; c) environmental exposure and; d) fingers used to deposit DNA.

Proficiency testing is a requirement for all ISO17025 accredited laboratories and in order to test the reliability of the optimised protocol designed in Chapter II, a form of proficiency testing was carried out where a second DNA analyst was asked to follow a standard operating procedure (see Chapter II) that had previously generated data for published studies. The laboratory analyst used optimised swabbing to recover DNA from a hand print deposited on 'unclean' glass and from a fingermark deposit on a 'clean' plastic substrate. One criticism with direct PCR is that the entire sample, in most cases, will be used up in one reaction in an attempt to generate a DNA profile. In response to this, an additional experiment was set up to recover DNA from fingermark residue (that had previously been swabbed using a wet swab) to explore the potential to generate additional alleles. By applying a second wet swab to the same targeted area the fibres from the second swabbing event were utilized for direct PCR.

Sampling

One individual was asked to wash their hands and briefly (15 seconds) create a handprint on a glass window and to deposit a thumb print on a sterile plastic microscope slide (Rinzl plastic, ProSci Tech, QLD, AU) 15 minutes post hand washing.

Samples examined:

Sample 1) Glass window handprint 'dirty' - 5 digits swabbed;

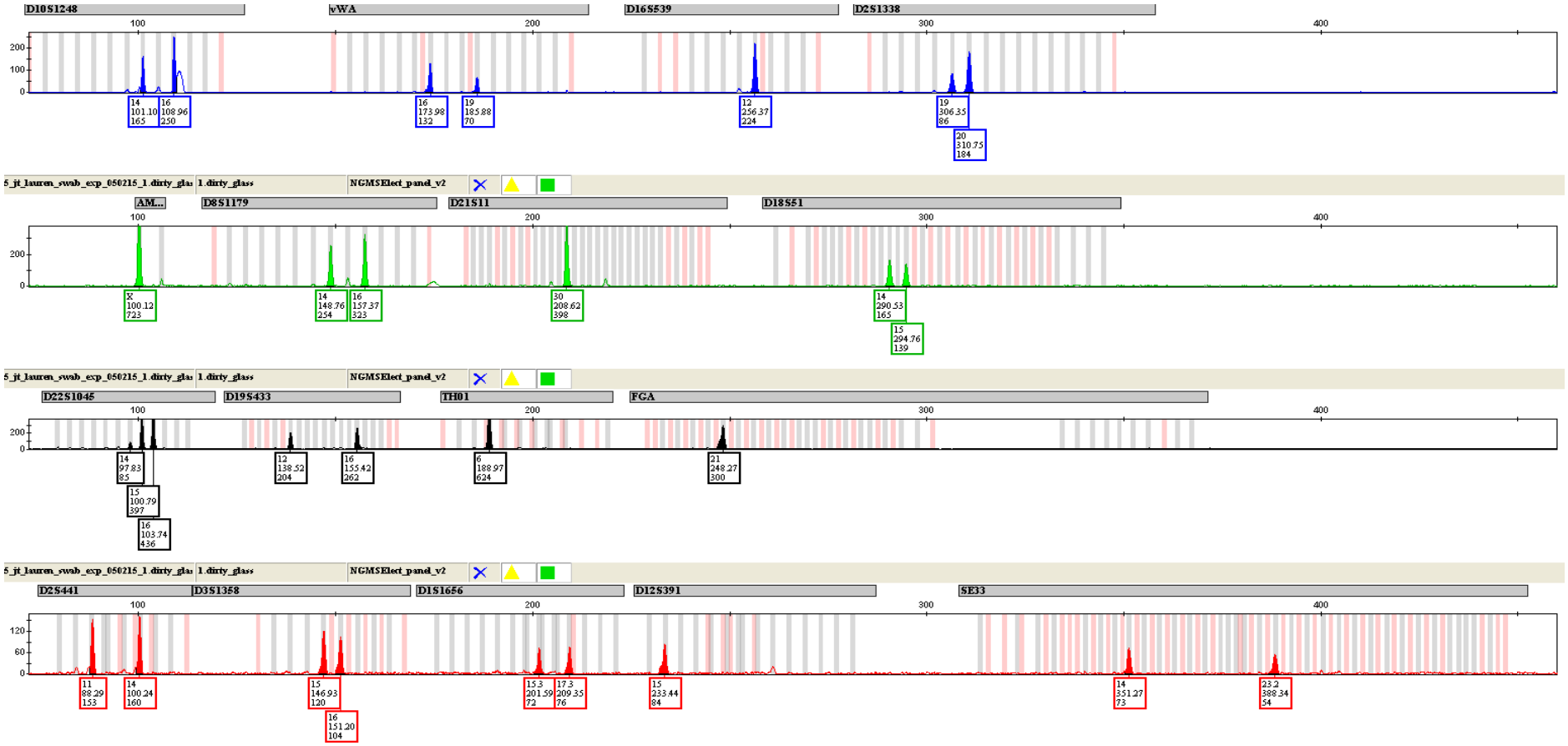
Sample 2) Plastic 'clean' substrate – thumb print swabbed and;

Sample 3) Second wet swab from sample 2 – thumb print repeat swabbing.

Post-swabbing, the AmpF*ℓ*STR® NGM SElect™ PCR amplification kit (Life Technologies, AU) was used for direct PCR following manufacturer's recommendations. All PCR products were analyzed on a 3130*xl* Genetic Analyzer (Life Technologies).

Results

Sample 1: DNA profile obtained by direct PCR from a glass hand print



SI Figure 5.21 STR-based DNA profile obtained by direct PCR from swabbing a hand-print on glass. A partial (near complete) DNA profile - that matched the donor - was obtained (i.e. 33 out of 34 possible alleles) with the AmpFℓSTR® NGM SElect™ PCR amplification kit, where DNA alleles matched the reference profile of the donor. An additional allele was observed at loci D22S1045.

No additional Ampli*Taq* Gold® DNA polymerase was added to the PCR and this may be the reason for observing some degree of heterozygote imbalance; an indication of PCR inhibition (see SI Figure 5.21). The analyst was capable of following an SOP (detailed in Chapter II) in order to generate a meaningful DNA profile using the direct PCR approach and standard amplification conditions that were recommended by the STR kit manufacturer.

Sample 2 Result: The number of alleles generated from the first swabbing event = 12 alleles, plus the Amelogenin marker. The profile was considered informative (i.e. 'up-loadable' to NCIDD).

Sample 3 Result: Number of alleles generated from the second swabbing event = 11 alleles, plus the Amelogenin marker. Profile not considered informative (i.e. fewer than 12 alleles generated). While there is an artificial threshold of 12 alleles being used to consider a profile 'up-loadable', it is worth noting a difference of only one allele in the profiles obtained from sample 2 and sample 3.

Combining the DNA profiling results from both swabbing events allowed a compilation of a composite profile that is more informative (i.e. 15 alleles, plus Amelogenin) (see SI Table 5.17). The method of double or triple swabbing may provide additional alleles that could be used in a similar manner. Further investigative work is required to measure the reliability of this method. Additionally, recording the same allele twice - from at least two separate amplification events – minimises the risk of profiling spurious alleles (i.e. 'drop-in') via contamination.

SI Table 5.17 STR-based DNA profiling results for composite profile building from two swabbing events using direct PCR. A fingermark was deposited on a plastic substrate and subjected to 'targeted' swabbing (twice) in order to generate two profiles from two separate direct PCR applications.

GENOTYPE PROFILE						
Sample ID: COMPOSITE PROFILE						
Combined result from Sample 2 and Sample 3 (swab events)						
Genetic loci	Allele	RFU	Allele	RFU	Additional alleles	RFU
D10	14	224	F	F		
VWA	16	64	19	50		
D16	F	F	F	F		
D2S1	F	F	20	51		
Amelogenin	X	198	F	F		
D8	14	99	16	124		
D21	30	133	F	F		
D18	14	51	F	F		
D22	15	133	16	253		
D19	12	56	F	F		
THO1	6	89	F	F		
FGA	21	110	F	F		
D2S4	11	56	14	53		
D3	F	F	F	F		
D1	F	F	F	F		
D12	F	F	F	F		
SE33	F	F	F	F		

N.B. Alleles highlighted in yellow are additional alleles obtained from the second wet swab that were combined with alleles from the first swabbing event in order to generate a composite profile. Four alleles dropped out from the second swabbing event that were obtained from the first swabbing event. Allele 'drop-out' is denoted by 'F'.

Appendix D - Supplementary Information

Pre-soak lysis method for masking tape and cartridge cases

AIM: A non-direct method is explored for use on larger substrates.

Samples were placed in a buffer containing only Triton-X™ to lyse the cells, and the buffer was then added directly to PCR. This was a non-direct PCR approach to see how well Triton-X™ performed at extracting DNA without additional reagents, swabs, or purification wash steps involved.

Soaking method

Masking tape (Scotch general purpose brown tape, 48 mm x 75 M, OrderMax, AU) and cartridge cases: nickel (38 Special, Winchester “Super X” centre fire ammunition, nickel plated case); aluminium (38 Special, CCI “Blazer” centre fire ammunition, aluminium case); and brass (Winchester, Australia), were handled by an individual for 15 seconds, 15 minutes after the individual had washed their hands.

Masking tape

The masking tape was handled by the donor for 15 seconds and placed adhesive side down on a plastic ziplock bag for storage. For DNA recovery, a section of tape was cut (i.e. 2 cm^2 portion) and added to a 1.5 mL sterile ‘DNA-free’ Eppendorf tube using sterile forceps and 150 µL of lysis buffer (Triton-X™, 0.1%) pre-heated to 50 ° C, to lyse cells. The sample was left for 1 hour at 50 °C prior to aliquoting 10 µL into the PCR (in place of DNA extract). Replicates for this study are described in the mock case study where 15 tape samples were examined for DNA post handling, see mock case manuscript: ***“Templeton, Jennifer EL, et al. DNA profiles from fingermarks: A mock case study. Forensic Science International: Genetics Supplement Series 5 (2015): e154-e155”.***

Cartridge cases

Nickel, brass and aluminium cartridge cases (detailed above) were handled briefly (15 seconds) by a donor and the cartridge cases were transferred by sterile forceps to a 5 mL sterile DNA-free Eppendorf containing 200 µL of Triton-X™, 0.1%, pre-heated to 50 °C, to lyse cells. Samples were placed on a heat block for 1 hour at

50 °C, and vortexed every 10 minutes prior to aliquoting 10 µL into the PCR (in place of DNA extract).

Amplification conditions

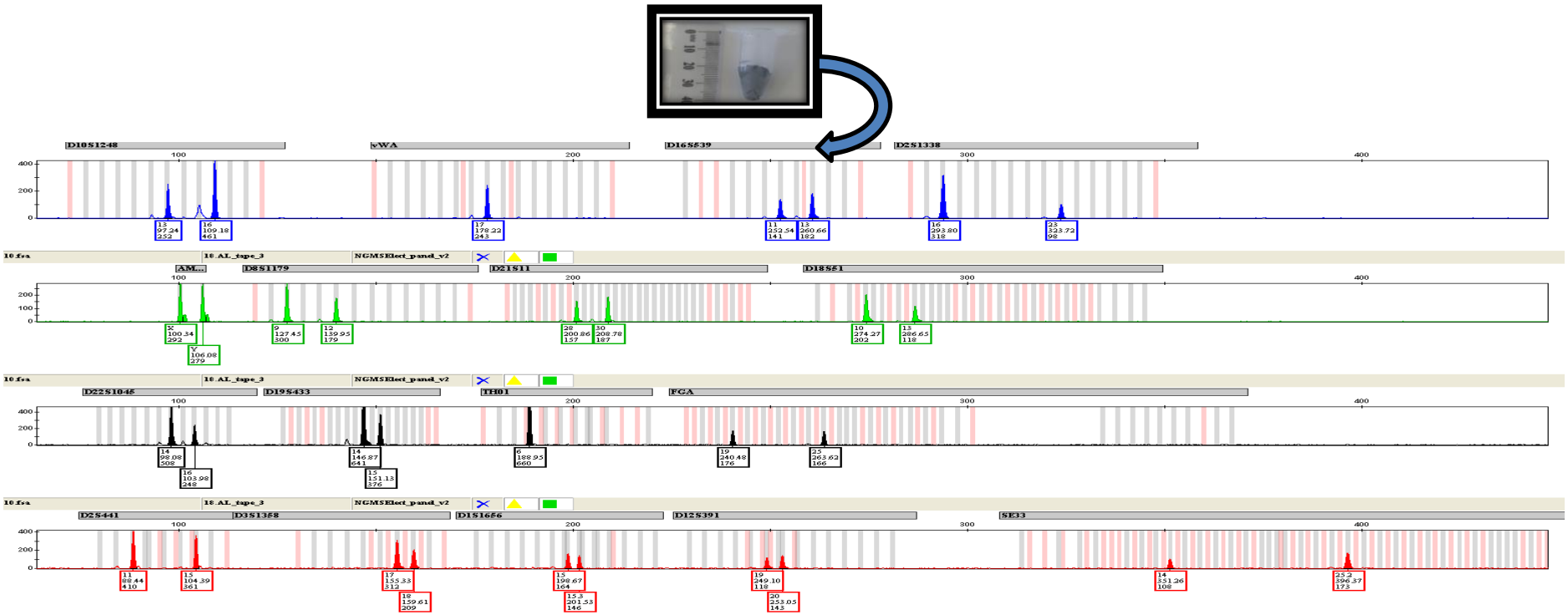
PCR was carried out using the AmpFℓSTR® NGM SElect™ PCR amplification kit (Life Technologies, AU), following manufacturer's recommendations. All sampling was carried out in triplicate.

Results - Pre-soak lysis method for masking tape and cartridge cases

AIM: A non-direct method is explored for use on larger substrates.

Example of DNA profile obtained from a fingermark - masking tape

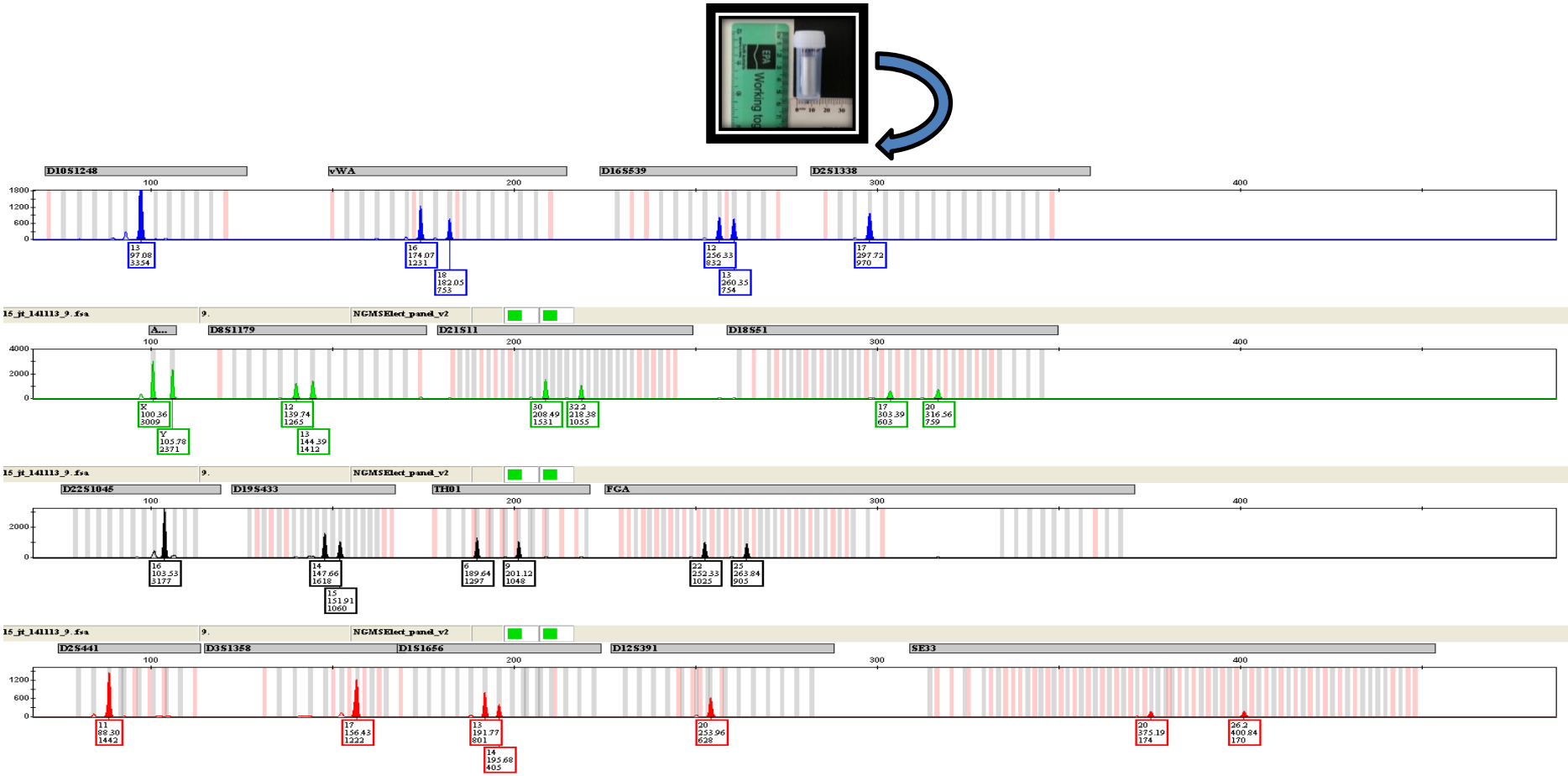
1.5 mL tube with masking tape and 150 µL of lysis buffer (Triton-X™)



SI Figure 5.22 STR-based DNA profile obtained by a pre-soak method where masking tape was placed into Triton-X™ buffer and buffer added directly to PCR. A full DNA profile was obtained (i.e. 34 out of 34 possible alleles) with the AmpFℓSTR® NGM SElect™ PCR amplification kit, where DNA alleles matched the reference profile of the donor. This profile would be considered 'up-loadable' to NCIDD.

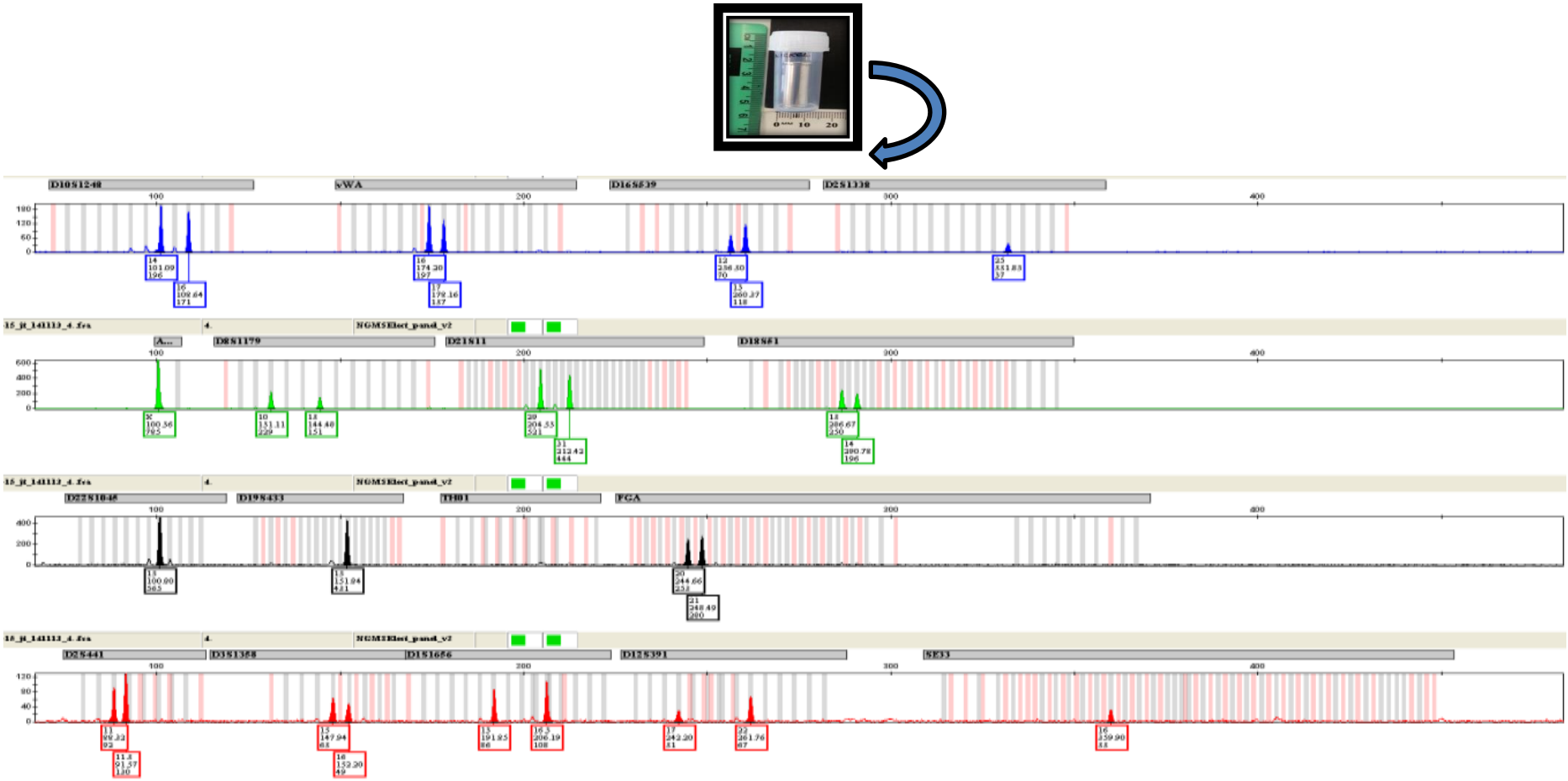
Further results for tape soaking in lysis buffer are included in the mock case study (see manuscript); 11 out 15 masking tape profiles were considered ‘up-loadable’ to NCIDD (i.e. 73% success rate).

Example of DNA profile obtained from a fingermark - Aluminium cartridge case



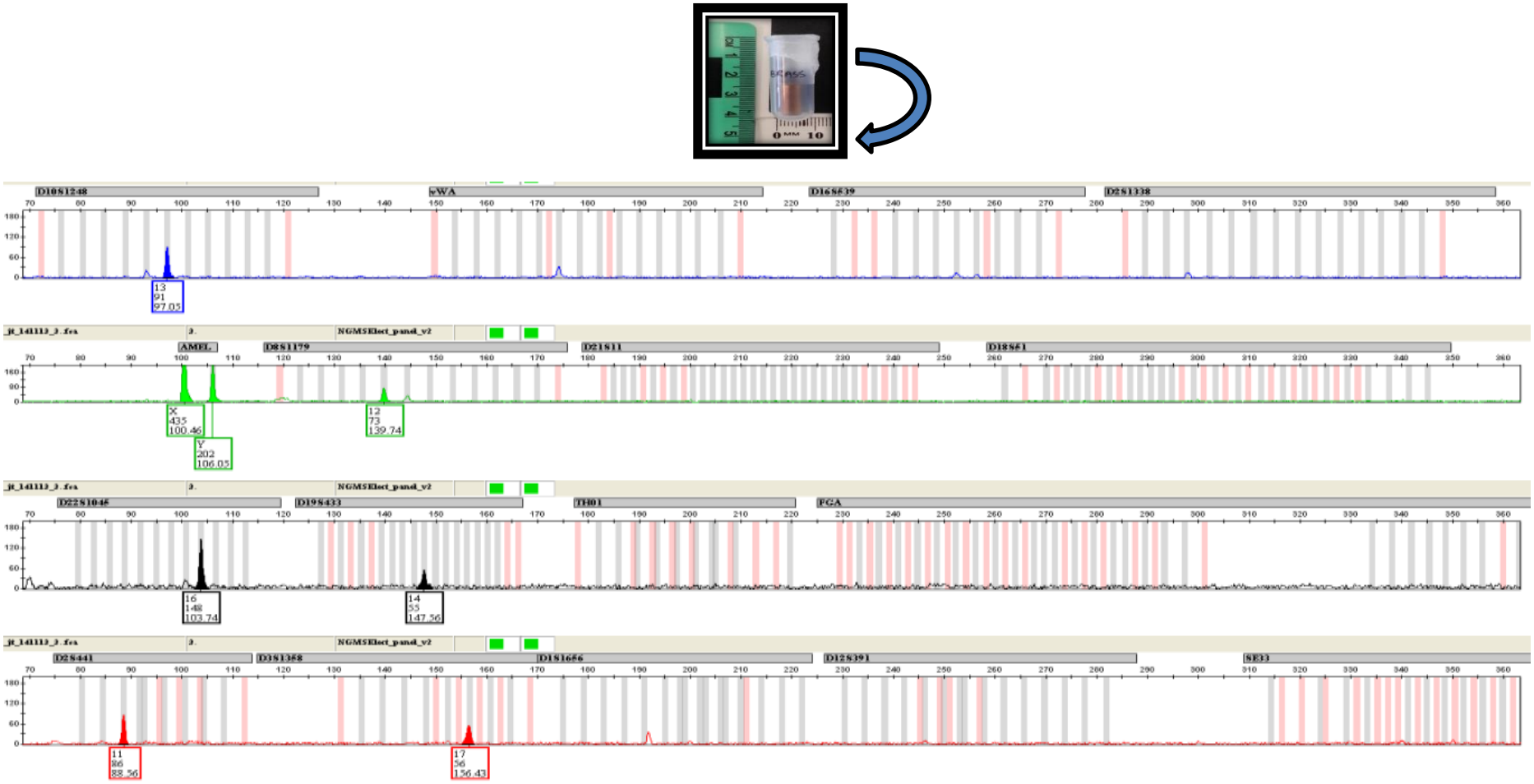
SI Figure 5.23 STR-based DNA profile obtained by a pre-soak method where the aluminium cartridge case was placed into Triton-X™ buffer, and buffer added directly to PCR. A full DNA profile was obtained (i.e. 34 out of 34 possible alleles, including Amelogenin) with the AmpFℓSTR® NGM SElect™ PCR amplification kit, where DNA alleles matched the reference profile of the donor. This profile would be considered 'up-loadable' to NCIDD. On zooming in further, 'pull-up' could be observed.

Example of DNA profile obtained from a fingermark - Nickel cartridge case



SI Figure 5.24 STR-based DNA profile obtained by a pre-soak method where the nickel cartridge case was placed into Triton-X™ buffer and buffer added directly to PCR. A partial DNA profile was obtained (i.e. 30 out of 34 possible alleles, including Amelogenin) with the AmpFℓSTR® NGM SElect™ PCR amplification kit, where DNA alleles matched the reference profile of the donor. This profile would be considered 'up-loadable' to NCIDD.

Example of DNA profile obtained from a fingermark - brass cartridge case



SI Figure 5.25 STR-based DNA profile obtained by a pre-soak method where the brass cartridge case was placed into Triton-X™ buffer and buffer added directly to PCR. A partial DNA profile was obtained (i.e. 8 out of 34 possible alleles, including Amelogenin) with the AmpFℓSTR® NGM SElect™ PCR amplification kit, where DNA alleles matched the reference profile of the donor.

SI Table 5.18 STR-based DNA profiling results for handled cartridge cases processed by the pre-soak method and AmpFℓSTR® NGM SElect™ amplification kit at standard PCR cycles.

Cartridge casing:	STR alleles obtained that match donor:	Additional minor alleles obtained:
Sample1_Nickel	30 alleles	-
Sample2_Nickel	6 alleles	-
Sample3_Nickel	34 alleles	-
Sample1_Aluminium	34 alleles	-
Sample2_Aluminium	27 alleles	4 alleles
Sample3_Aluminium	34 alleles	-
Sample1_Brass	0 alleles	-
Sample2_Brass	1 allele	-
Sample3_Brass	8 allele	-

Table showing alleles generated that match the donor of the fingermark created on the cartridge case. One mixed DNA profile was observed (i.e. more than 2 alleles detected at loci) for sample 2 - aluminium cartridge casing; four additional minor alleles were recorded indicating a second contributor. Results highlighted in blue are profiles considered 'up-loadable' to NCIDD (i.e. 5 out of 9 profiles).

Appendix E - Poster: Direct PCR: Successes and Limitations.

Presented at "The Australian and New Zealand Forensic Science Society", ANZFSS conference, Auckland, NZ. 2016.

Direct PCR: Successes & Limitations

Renée Blackie[1], Jennifer Templeton[1], Duncan Taylor[1,2], Emily Rowe[2], Oliva Handt[1,2], Adrian Linacre[1]

[1]School of Biological Sciences, Flinders University, Adelaide, Australia

[2]Forensic Science South Australia, Adelaide, Australia

Presented at the ANZFSS Symposium, Auckland, 2016

Funding was provided by the Attorney General's Department, South Australia

Introduction

Direct PCR has the attraction of omitting any extraction process and placing the item of interest directly into the PCR tube. Amplification of DNA profiles from single hairs, fingernails, fingermarks and non-human substrates have all be successfully achieved [1-5]. There are associated limitations as outlined in this poster.

BENEFITS

No loss of sample
No need for DNA extraction kits
Increased speed of analyses

LIMITATIONS

Not reproducible as all the sample is used in one reaction
Quantification of initial template DNA cannot be performed

Method

Direct PCR

Hairs & Nails: ~ 3mm of hair shaft is placed directly into the 0.2 mL PCR tube

Fingermarks: nylon swabs moistened with 0.1% Triton-X collect the DNA. Fibres are removed and placed directly into the PCR tube

STR amplification

Amplification of fingermarks and nails used the NGM SElect™ kit (Life Technologies, Vic, AUS) in a volume of 25 µL following standard conditions except an extra 1 µL of AmpliTaq Gold® DNA polymerase was added. A standard cycle number of 29 was used.

DogFiler loci were amplified as in reference 2.

SNP Testing

SNaPShot was performed following the manufacturer's recommendations.

Capillary electrophoresis was performed on an ABI 3130x*L* Genetic Analyser (Life Technologies) using POP-4 polymer (Life Technologies). An aliquot of 2 µL of the amplified samples was added to a solution containing 0.5 µL of Liz600 LIZ Size Standard and 9.5 µL of Hi-Di Formamide. Samples were denatured at 95 °C for 3 min. Electrophoresis was conducted at 3 kV with a 10 s injection. The data were analyzed using GeneMapper v3.2.

Figure 1: example of a section of hair (left) and fibres from a fingermark (right) prior to the addition of the PCR reagents

Results

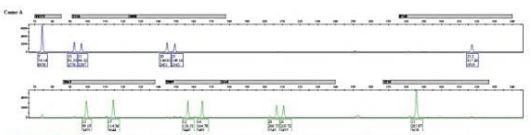


Figure 2: example of a shed telogenic hair after direct PCR using loci from Dogfiler [2]

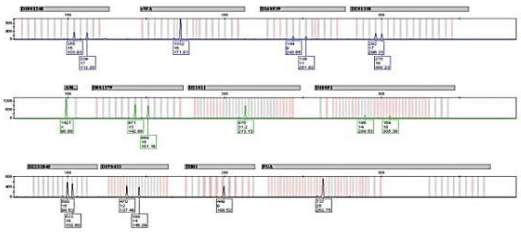


Figure 3: example of a ~2 mm section of a fingernail after direct PCR [3]

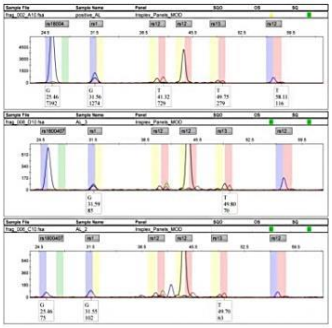


Figure 4: example of the Irisplex SNP obtained from a fingermark made on packaging tape

Discussion & Comments

Direct PCR has a niche application and need only be applied when there is a low chance of successful STR amplification following the DNA extraction process.

Full or near full DNA profiles were obtained from smudged fingermarks when no alleles were obtained post DNA extraction.

Phenotypic testing was possible from touched items

In the case of fingernail clippings only a small amount (~2 mm) was sacrificed such that the rest of the nail could be subjected to standard DNA extraction.

Shed dog hairs generated full profiles routinely.

References

1. JEL Templeton, D Taylor, O Handt, P Skuza, A Linacre. *Direct PCR improves the recovery of DNA from various substrates.* Journal of Forensic Sciences (2015) 60(6):1558-1562.
2. R Blackie, D Taylor, A Linacre. Successful direct amplification of nuclear markers from single dog hairs using DogFiler multiplex. Electrophoresis, (2015) 36(17) 2082-2085.
3. R Ottens, D Taylor, A Linacre. *DNA profiles from fingernails using Direct PCR.* Forensic Science Medicine & Pathology (2015) 11:99-103.
4. JEL Templeton, A Linacre. *DNA profiles from fingermarks.* BioTechniques (2014) 57(5):259-266.
5. R Ottens, D Taylor, D Abarno, A Linacre. *Successful direct amplification of nuclear markers from a single hair follicle.* Forensic Science Medicine and Pathology (2013) 9:2 238-243.

Correspondence: A. Linacre, Flinders University, South Australia, adrian.linacre@flinders.edu.au

Chapter VI
Concluding remarks and further applications

6.1 Preface

This final chapter begins with a summary of the data presented in this thesis. Each data chapter comprises a detailed results and discussion section relating to specific research questions. The aim of this present chapter is to bring together all the results and emphasise how the findings from each chapter come together as a whole and contribute to the overall aims of the thesis. Significance of the research is highlighted. Equally, limitations, concerns, and potential restrictions that forensic laboratories may face when trying to implement direct PCR are expressed. Future perspectives for direct PCR, and collaborative projects undertaken at Flinders University that complement the data herein, are discussed. Lastly, a novel swab device used to collect DNA at crime scenes is described and future research studies for testing the swab with multiple genetic markers and new sequencing platforms are considered.

The future application of massive parallel sequencing (MPS) and the analysis of multiple discriminatory markers for identifying the source of origin of fingermark traces are discussed. As DNA profiling moves forward at an exciting rapid pace there are new directions for future research. The ability to sequence the entire human genome in a single day provides scope for investigating new and existing genetic markers in trace material by new means of analysis.

Additional manuscripts to which I have contributed during this candidature are included in the appendix and relate to other areas of low-level (i.e. ancient) DNA typing. The publications are included to highlight additional work that has been carried out with alternative methods of analysis (i.e. DNA-capture using hybridisation/enrichment and MPS), and the potential to sequence DNA that is severely damaged, degraded and in low quantity from samples that are thousands of years old.

6.2 Relevance and significance of research project

The relevance and value of ‘touch’ DNA to the forensic community has become more evident as time goes on. The ability to generate DNA profiles from fingermarks has been with us since approximately 1994 and impacted greatly on the criminal justice

system; data was not reported in the literature until 1997. At present, sites of significance or legal importance may require multiple swabs to be taken from areas where DNA is potentially present. DNA samples have to be taken at a crime scene as soon as possible, from possible murder weapons, other items of forensic interest, points of entry and exit, or areas with a high likelihood of returning a usable DNA profile. Hence, data chapters in this thesis focus on this type of trace material as evidence and the ability to retrieve DNA using swab fibres for the purpose of direct PCR. New means to collect and process 'touch' DNA samples have led to vast improvements in the ability to profile samples of this nature. By circumventing the extraction process more DNA is retained for amplification, there are no tube changes and wash steps; and therefore less opportunity for contamination.

More recently, direct PCR has become a worldwide technique in many applications of forensic science [1-10] enabling fast and effective identification. In particular, the successful use of direct PCR has been highlighted in areas specific to wildlife crime [11, 12] and recovery of mtDNA from explosive devices [13]. Wildlife trade is estimated at a devastating $54 billion (USD) [14]; seized material often consists of hairs, bones, ivory, tissues and horns [12]. Many cases of illegal wildlife trade go undetected as there is a lack of funding in this area of research. More recently, a rapid direct PCR assay has been developed to aid the identification of animal species [11, 12] and enables a more cost effective investigation due to its increased sensitivity and reduced time and cost of processing samples.

In this thesis, the suitability of swab fibres for direct PCR was assessed through several experiments. Ultimately the work herein offers a solution to current methodology which deals with loss of DNA from swab material when a standard extraction is performed. Specifically, this thesis comprises of three data chapters that investigate the ability to profile DNA using the direct PCR approach. It is clear from the results (see Chapters III, IV and V) that direct PCR has the potential to be an alternative method for processing swabs. Results demonstrate that direct PCR enabled significant improvement in DNA yield from swabs used on brass, plastic and glass substrates, over conventional extraction. FSSA have since validated the method of direct PCR for processing biological material (i.e. FTA cards used for

reference material and individual hairs). Validation studies at FSSA involved replicate testing of over 200 reference samples to ensure the technique was reliable, robust, reproducible, and fit for purpose. FSSA have emphasized that direct PCR is at times the only way to generate an informative DNA profile from evidence.

6.3 Summary of thesis chapters

Chapter I – Introduction

The introductory chapter (Chapter I) includes a published review introducing the broad application of forensic DNA technology and covers key events and progression in the field from both a historical view and current practice [15].

1) Review paper: Linacre, A. and **Templeton, J.E.L.,** 2014. Forensic DNA profiling: state of the art. **Res Rep Forensic Med Sci**, 4, pp.25-36.

Genetic markers used in both criminal and civil investigations are discussed, such as STRs, X and Y chromosomes, along with mitochondrial DNA. The review paper progresses to more advanced methodology such as next-generation DNA sequencing, and highlights the enormous capacity in the generation of sequence data. The bottleneck may lie in the development of adequate software and bioinformatics tools to handle the large amounts of data that can be generated.

Chapter II – Materials and Methods

The methods section of the thesis describes an alternative approach to processing swabs that were used to retrieve DNA from various substrates. Emphasis was placed on contamination controls in order to prevent cross-over contamination and laboratory associated allele 'drop-in'. Preventative measures - similar to ancient DNA practice - were in place to minimise extraneous DNA contamination. Standard operating procedures have been drafted to assist with future validation.

Chapter III – Improving the collection of ‘touch’ DNA

In search for means to improve the recovery of DNA from touched substrates, a nylon flocked swab was used in conjunction with a detergent-based swab medium, and saturated fibres were utilised for direct PCR. The most effective method of swabbing for direct PCR was examined, and data generated using conventional extraction was compared to data obtained using direct PCR from swab material.

Previous studies report that detergents are more effective swabbing reagents [16] and Triton-X™ and SDS [17] have been shown to work well under laboratory conditions. The first data chapter (i.e. Chapter III) is divided into 3 manuscripts:

1) **Templeton, J.,** Ottens, R., Paradiso, V., Handt, O., Taylor, D. and Linacre, A., 2013. Genetic profiling from challenging samples: direct PCR of touch DNA. **Forensic Science International: Genetics Supplement Series**, 4(1), pp.e224-e225**.**
2) **Templeton, J.E.,** Taylor, D., Handt, O., Skuza, P. and Linacre, A., 2015. Direct PCR Improves the Recovery of DNA from Various Substrates. **Journal of forensic sciences**, 60(6), pp.1558-1562.
3) Ottens, R., **Templeton, J.,** Paradiso, V., Taylor, D., Abarno, D. and Linacre, A., 2013. Application of direct PCR in forensic casework. **Forensic Science International: Genetics Supplement Series**, 4(1), pp.e47-e48.

Manuscript 1 - *a short supplementary series* - compared the efficacy of three swab types (i.e. cotton, nylon and foam) for direct PCR processing. Nylon flocked swabs moistened with Triton-X™ generated the highest yield of PCR product and an optimised protocol was initiated for future experimental work.

Manuscript 2 - *a technical note* - compared results from the direct PCR approach to a standard extraction when DNA was deposited on various substrates (i.e. brass, glass and plastic) and swabbed. As expected, direct PCR produced a higher yield of DNA for all substrates tested. Brass generated the lowest yield of PCR product and is thought to be a difficult surface type to obtain DNA from. Encouragingly, swab fibres present in the reaction vessel did not prevent the amplification of DNA and full profiles could be obtained using direct PCR.

Manuscript 3 - *a short supplementary series* - highlights the benefits of direct PCR by analysing trace and 'touch' DNA on a range of substrates and explores the loss of DNA due to extraction. It was observed that by following a standard routine extraction protocol loss of DNA can be up to 80%.

Chapter IV – DNA profiles from fingermarks

The capacity to use direct PCR on fingermark deposits left behind on plastic substrates was explored in Chapter IV. The addition of PCR facilitators for the purpose of direct PCR was explored and the time of DNA deposition after hand

washing was investigated (i.e. immediately, 15 minutes, 1 hour and 2 hours post hand washing). A comparison of profiling results obtained by direct PCR swabs and extracted DNA swabs were carried out. The ability to profile DNA using direct PCR from two common STR profiling kits was determined. Results were split into a full length manuscript, a second manuscript (under review), and a real case study:

1) **Templeton, J.E.** and Linacre, A., 2014. DNA profiles from fingermarks. **BioTechniques**, 57(5), p.259 – 266.
2) Blackie, **J. Templeton,** D.A. Taylor, A. Linacre. PCR buffer enhancement of STR kits used for human identification. **(In preparation).**

   Case work:

   Case 1) Application of case work: Seizure of Methamphetamine in South Australia.

Volunteers were asked to deposit fingermarks onto substrates at different time intervals post hand washing. Fingermark residue was swabbed using optimised methodology and fibres were utilised for direct PCR. The most important finding was that direct PCR could generate informative DNA profiles after a short time had passed since hand washing (i.e. 15 minutes). Further application of direct PCR in a real case study highlights the potential for swabs to be processed in operational high-throughput laboratories, and the benefits and limitations of this approach. Such observations led to work carried out in Chapter V.

Chapter V – Direct PCR of 'touch' DNA: Effect of PCR inhibitors and environmental exposure

The final data chapter is divided into two manuscripts:

1) **Templeton, J.E.** and Linacre, A. Direct PCR DNA profiling of dactyloscopic powdered fingermarks. **Forensic Science International: Genetics,** DOI: http://dx.doi.org/10.1016/j.fsigen.2017.05.006
2) **Templeton, J.E.,** Taylor, D., Handt, O. and Linacre, A., 2015. DNA profiles from fingermarks: A mock case study. **Forensic Science International: Genetics Supplement Series**, 5, pp.e154-e155.

To refine our understanding of 'touch' DNA and explore the limitations of direct PCR, the effect of common fingerprint powders on direct PCR was investigated. Fingerprint powders present in the direct PCR vessel do not appear to influence the

ability to generate a DNA profile. Surprisingly, the presence of black magnetic fingerprint powder still permitted the amplification of DNA and these powders are known to fail during routine extractions when magnetic-based racks are used. Post-PCR purification (i.e. MinElute® purification columns) generated a higher yield of PCR product and this finding was in agreement with other relevant literature [18]. The method of concentrating DNA amplicons and removing negative ions may be of value to challenging sample types (e.g. brass casings or sub-optimal samples) when DNA alleles are borderline for detection. Further work investigated the use of direct PCR on a suite of forensic related substrates. Glass recorded the highest number of detectable loci: 87% of direct PCR swabs from glass were considered ‘up-loadable’ to the NCIDD, followed by masking tape (73%), swabs from nickel (60%), wood and aluminium (53%) and brass (0%). A final experiment was carried out to evaluate the likelihood of being able to obtain DNA by direct PCR from touched substrates left exposed to outdoor conditions overnight (i.e. wind, UV light, rain), and in one case for 8 days.

The spectrum of sample types and variables to explore with direct PCR methodology are beyond the capabilities of one research candidature. Further work would be required with a large sample set to more accurately determine how long DNA survives under harsh conditions. The variety of environmental conditions that 'touch' DNA samples are exposed to may affect the quality and quantity of surviving DNA [19-21]. Future work could aim to explore DNA transfer, time of DNA deposition, and persistence/preservation of DNA. A large amount of work is currently being undertaken by other research groups in these areas, in particular the occurrence of “non-self” DNA and the potential of secondary/tertiary DNA transfer in reference to fingermark traces [22-33].

6.4 Limitations, concerns, and restrictions of direct PCR

One criticism of direct PCR is that the entire sample, in most cases, is completely consumed. For direct PCR of swab material, it has been found that a second separate swabbing event from the same targeted area is capable of generating DNA alleles by direct PCR, and this result is promising for future validation studies (see Appendix in Chapter V). On occasions, multiple swabs are taken from areas of

evidence and this will allow fibres from one swabbing event to be carefully removed for direct PCR in addition to a second or third swab to be retained as backup. Another area of concern is how to preserve swabs that have been moistened with a swab lysis media (i.e. Triton-X™ buffer). It can be assumed that without air-drying DNA degradation would occur at a faster rate as moisture would promote the growth of bacteria. Cells would no longer remain intact in the presence of a lysis reagent and it is thought that swabs would need to be air-dried prior to storage or processed immediately. Fox *et al*. [34] reflects upon swab moistening agents and the effects of lysis buffer on DNA preservation. In this study, swabs used to collect blood stains were moistened with lysis buffer (20 %) in place of sterile water, with no adverse effect on DNA profiling results. Further DNA preserving agents, similar to those found in Whatman FTA® paper (e.g. EDTA, SDS and Tris solution) should be considered for stabilizing DNA and further work in this area is required to measure the effect of these additives on direct PCR profiling.

Direct PCR provides a trade-off between eliminating the extraction methodology to prevent loss of sample DNA and not being able to quantify a sample prior to PCR. The inability to quantify DNA is a limiting factor for many forensic DNA laboratories; STR profiling kits recommend optimum DNA template of 0.5-2 ng prior to PCR. Importantly, fingermark samples are unlikely to contain so much DNA as to overload PCR or electrophoresis instruments, and peak heights on the electropherogram can be compared to calibration curves to extrapolate back to starting DNA amounts. Many laboratories, however, use quantification to indicate whether a sample should be processed further and this is one limitation that direct PCR faces. In spite of this, quantification should not be the sole determining factor in deciding whether to proceed with a sample.

It can be mis-leading to regard quantification as the only way to predetermine the characteristics of a DNA profile; the ability to obtain a good quality DNA profile is dependent on both quality and quantity of template. For example, a degraded DNA sample may exhibit high amounts of DNA and generate only a sub-optimal profile due to damaged or degraded template. It is important to note the quantitation value obtained indicates the amount of DNA only and is not a measure of the quality of the

DNA. Despite this, quantification is a still a requirement for many DNA testing laboratories.

One other limitation with the direct PCR approach is the non-removal of PCR inhibitors. It would be interesting to measure the effect of different contaminants on the ability to process direct PCR swabs by mixing biological materials together (e.g. blood, saliva, and fingermark residue) and evaluate the success.

6.5 Future perspectives

6.5.1 Targeted approach for detecting fingermark deposits

To complement the work in this thesis, candidature currently undertaken by Alicia Haines at Flinders University explores the use of intercalating fluorescent binding dyes as a source of latent DNA fingermark detection. A Polilight® has been used to observe specific areas where DNA may have been deposited by fingermarks [35-37]. Fluorescence indicates an area with potential DNA that can be swabbed and later profiled for DNA typing. At present, no other studies have examined the use of intercalating dyes that do not pose a risk to the resulting DNA profile. Current practice involves swabbing large areas of substrates and the use of multiple swabs with no real indication of where DNA can be found. By using a targeted approach of dye-fluorescence to detect DNA there is an increased likelihood of targeting the correct area rather than the blind swabbing approach. The optimum concentration of DNA-binding dyes enable the ridge detail within the fingermark to be studied and the dyes examined do not interfere with the ability to amplify STR markers. The method was further developed to be suitable as a screening tool for hair analysis to ascertain if sufficient DNA is present for STR profiling [38]. In the future it would be beneficial to use the optimum protocol developed by Alicia Haines to screen for cellular material, in conjunction with direct PCR, to improve the overall success rate of obtaining DNA from previously touched substrates.

6.5.2 Other applications for direct PCR

The application of direct PCR has been tailored to suit other forensic related material with profound success. Candidature carried out by Renée Blackie at Flinders University highlights the broad application of direct PCR to the forensic community.

Renée Blackie and co-authors have reported on successful STR DNA profiling from single human hairs [39, 40], dog hairs [41], fibres from worn clothing [42], and fingernail clippings [43], using direct PCR. By eliminating the extraction step and proceeding to direct PCR the authors report 100% success with anagen hairs, and further success with a 5 year-old hair generating a full DNA profile [39]. For fingernail clippings (i.e. 4 mm^2 sample), direct PCR resulted in 95% of the samples yielding a profile containing 5 or more STR loci [43], considered 'up-loadable' to the Australian National Criminal Investigation DNA Database (NCIDD). Higher quality DNA profiles (i.e. less heterozygote imbalance and limited stochastic effects) were observed from samples subjected to direct PCR compared to extracted samples. The authors recommend a dilution of PCR product for fingernail clippings prior to electrophoresis to prevent overloaded DNA profiles from being observed [43]. Further research conducted on worn clothing fibres and the use of direct PCR highlights the potential for tape-lift samples to be processed directly without the need to extract DNA [42]. Samples produced DNA profiles of 15 or more alleles. Performing an extraction of fibres taken from the same garment resulted in the failure to produce DNA alleles. Work undertaken to process canine hair samples by direct PCR resulted in a significantly higher success rate for guard hairs compared to undercoat hairs [41]. The researchers recommended guard hairs only for direct PCR processing.

In all the aforementioned studies, the authors used standard DNA profiling kits with no increase in PCR cycles in order to minimise stochastic effects. By following manufacturer's recommendations with standard STR profiling kits the direct PCR approach would enable a much faster implementation into operational laboratories upon validation. This is an important consideration for case work laboratories (see case example given below).

Case example – Reed & Reed vs R

One landmark case based substantially on DNA evidence surrounding the use of the LCN process was *Reed & Reed vs R* [44].

Terrence and David Reed where found guilty of the murder of Peter Hoe on the 12th October 2006 with key evidence linking a knife and a broken plastic handle found at the scene to the DNA profiles of the accused. Testing was carried out at the Forensic Science Service (FSS) in the UK where DNA samples were pre-assessed by quantitative PCR (qPCR) prior to standard STR typing or the LCN method (for low-template DNA). When the LCN technique was first implemented in the FSS in 1999 a prior quantification step was not necessarily carried out; as it was then a difficult task to accurately measure low-levels of DNA. In the case of *Reed & Reed vs R*, senior scientists choose not to quantify swabs taken from plastic sections of the knife and instead preceded with LCN, based on the assumption that swabs contained low-level DNA.

Interpretation of results

A key scientist from the FSS proposed that direct transfer was the most likely explanation for the presence of the accused's DNA on the knife handle and that "*the DNA was transferred at the time the handles broke*". This was adding activity level to source when using LCN. This was highly contentious and led to an appeal. There were also concerns regarding the validation of science and methodology used to generate the DNA profiles using LCN. Three judges at the Royal Court of Justiciary rejected the appeal and found that the DNA evidence was reliable and that the scientist should provide to the court the most likely scenarios to account for the presence of the DNA profile(s) [44].

Given that quantification was not a prerequisite to testing in this case, direct PCR could have been implemented in order to maximise the amount of template DNA available to the PCR without the need to increase the PCR cycle number beyond manufacturer's recommendations.

6.5.3 Inventing a ‘touch’ DNA swab for direct PCR

It is my opinion that direct PCR would not replace standard methods of recovery (e.g. swabbing large blood stained areas at crime scenes); instead, swabbing of small intricate surface areas where DNA is likely to adhere to but standard protocols have difficulty with recovery (e.g. keypads and grooves/triggers/handles of weapons). A swab designed for the purpose of direct PCR would be used on case exhibits that generally yield sub-optimal DNA (e.g. cartridge casings used as ammunition).

The cost of collecting and processing DNA samples which return a poor success rate presents a significant opportunity for improvement. In response to this need, a single use field device is currently being designed as a result of work accomplished in this thesis. The use of a novel swab co-designed by Jennifer Templeton, Adrian Linacre and Flinders Partners will be tested on a range of substrates on which 'touch' DNA will be deposited. The swab includes a pre-attached 0.2 mL PCR tube to the tip of the swab to enable ejection/release of swab fibres into the PCR vessel without further manipulation. Direct PCR processing of nylon flocked fibres is required to facilitate the efficient, reliable, and rapid recovery process. Alteration of the fibres - both how they are arranged and chemical composition - has the potential to collect even more DNA. By implementing the direct PCR approach and alleviating the ‘DNA extraction bottleneck’ created in the laboratory there would be a more effective forensic investigation.

Device features for the novel swab device (taken from the patent description) are detailed below:

- simple for a forensic technician to use in ’field conditions’, while wearing appropriate personal protective equipment (PPE) and in particular, while wearing sterile gloves;
- compatible with existing collection practices and standards;
- easy to label and record sample information;
- maintains reliable precision grip and maintains control with the swabs and PCR tubes used;
- provides good visibility to the surfaces being sampled;

- allows the micro-swab to contact the surface being sampled at the optimum fibre angle of attack, with an appropriate sampling force;
- simple to reliably eject a micro-swab into a 0.2 mL PCR tube;
- simple to handle and reliably manipulate micro-swabs and PCR tubes;
- provides good access to internal sampling surfaces (for example, the inside surface of a toilet roll or the rear surface of a door handle);
- minimises the risk of contact with unintended surfaces; and
- minimises the risk of cross contamination.

These requirements for the swab device have now been addressed and a provisional patent has been submitted “Nucleic acid collection device and method, 2016” [45]. An important area of research that will be investigated will involve testing the field device on weapons, cartridge cases (both fired and non-fired), components of timing devices, drug seizures and other items of significance to a forensic investigation.

6.5.3.1 Explosive devices

Terrorism crime usually involves explosive devices that consist of postal or pipe bombs composed of plastic, paper, cables, adhesive tapes and electrical components [46]. Current studies attribute the low success rates of DNA recovery from explosive devices and cartridge cases to the high temperatures that are reached during the firing process, causing subsequent DNA damage [13, 46-49]. On the contrary, one extensive study published data on the use of a thermal imaging camera to measure the exact temperatures reached due firing and the temperatures recorded were not high enough to affect DNA degradation [50]. Authors concluded that DNA survival is more dependent on the surface roughness pattern [50], also noted by Xu *et al.* (2010) [51]. In addition, metal inhibitors present in cartridge cases, pipe bomb fragments, or primer components of gunshot residue may affect the ability to successfully recover DNA.

Future work will involve collection of DNA from a range of forensically informative substrates, and transition of fibre collection device will go from laboratory-based to an operationally validated process. Using the collected DNA, a full suite of DNA

markers will be targeted that include all the standard STR loci (autosomal and lineage markers), mitochondrial DNA sequence data, together with ancestry and phenotypic markers. The platform used will be capillary-based electrophoresis separation and will transition into MPS platforms. These ideas are listed in a grant application to progress the work further.

6.5.3.2 Future application of MPS

SNP typing allows an increasing number of polymorphic loci to be examined in a single assay (i.e. 30 – 50 SNPs) [52-54], providing the opportunity to obtain more information from a single sample. The IrisPlex [55] and HIrisPlex panels [56] that explore hair and eye colour, along with the SNP*for*ID 52-plex ([52, 53, 57, 58] panel all expand on the number and type of polymorphisms that can be examined.

Other new exciting areas of research explore skin-associated lifestyle chemistries found on personal belongings that provide an additional form of trace evidence. One interesting study published by Bouslimani *et al*. (2016) [59] examined fingermark deposits specifically to shed light on the lifestyle and human habits of the person leaving behind the fingermark. Scientists obtained molecular signatures that provided insights into the dietary and medical status of the individual, type of hygiene/beauty products used, and even the places the person has been prior to depositing a fingermark. In spite of this modernistic approach for examining trace material, DNA profiling and classical fingerprinting remain the two main methods for identification of fingermark traces.

Massive parallel sequencing is the driving force behind ancient DNA research as other PCR-based sequencing methods are not effective at analysing DNA that is severely damaged, degraded and limited in quantity [60, 61]. Current Next-generation platforms recommend input DNA amounts similar to CE-based methods [62-64]. Increase in sensitivity means that opportunities exist for sampling trace material of limited DNA. An interesting approach - following on from the work described in this thesis - would be to attempt MPS from fingerprint DNA swabs and examine the bacterial composition, SNP markers for ancestry and phenotype, and STR markers used for identity, all in one assay.

Multiple different polymorphism types (i.e. phenotype, ancestry, identity SNPs) can be examined simultaneously in one assay using the capabilities of MPS. Multiple samples can now be run together through the use of barcoding (i.e. individualizing) to increase high-throughput capacity. The sequence depth and coverage generated by MPS exceeds CE-based detection and is at times the only way to sequence ancient DNA samples that are highly damaged and degraded (i.e. < 100 base pairs in length) [61]. A natural progression in forensic science is to follow ancient DNA research and its ability to target multiple existing loci in a single tube using MPS capabilities and to explore new SNPs that may arise. One interesting area to explore further is the application of MPS using detergent-saturated swab fibres as a template for creating DNA libraries; as currently it is not known if direct PCR using library primers would be an effective means of generating sufficient template DNA for MPS.

6.6 Concluding remarks

The forensic community as a whole strives to improve methods in order to gain more information from valuable samples that will assist forensic investigations. Direct PCR does so with ease, and not only by improving the yield of DNA that can be recovered from fingermark residue, but also by speeding up the way that samples are processed, minimizing the risk of contamination, and reducing cost; all major benefits to high-throughput laboratories. Encouragingly, the presence of multiple swab fibres in the PCR tube does not affect the ability to generate an informative DNA profile. In fact, the yield of DNA increases significantly when direct PCR is employed over the conventional extraction methodology. It is expected that a routine extraction will

greatly hinder a sub-optimal DNA sample from generating a useable profile. Direct PCR takes advantage of the cell-free DNA component on the surface of skin that is normally discarded in a routine extraction. This finding is fundamental to forensic DNA research as the field is most entirely built on the need to extract DNA from biological material. Optimising the amount of template DNA available for subsequent enzymatic manipulation will enable greater success for DNA profiling.

In this thesis, it has been shown that direct PCR offers tremendous value for amplifying DNA directly from fingermark traces that are either latent, or enhanced with powders, without the need to extract or purify the sample. This is not to claim that for other sample types other extraction methods may not generate higher quality profiles (e.g. fingermarks contaminated with body fluids or PCR inhibitors). Exploring new buffers and swab fibre compositions that complement direct PCR will pave the way for future analysis of 'touch' DNA samples. An important finding was that DNA could be recovered and typed from fingermarks in as little as 15 minutes post hand washing.

To summarise, operational laboratories may benefit from adopting the direct approach, especially when trace material is limited. Omitting the multi-step process enables a more efficient, simple, and cost effective work flow. Direct PCR is a low-technology, rapid and robust method that can be easily incorporated into forensic laboratories and should be considered as an alternative method for future use.

6.7 References

1. Kitpipit, T., K. Sittichan, and P. Thanakiatkrai, *Are these food products fraudulent? Rapid and novel triplex-direct PCR assay for meat identification.* Forensic Science International: Genetics Supplement Series, 2013. **4**(1): p. e33-e34.
2. Wong, W.H., et al., *Direct PCR optimization yields a rapid, cost-effective, nondestructive and efficient method for obtaining DNA barcodes without DNA extraction.* Molecular Ecology Resources, 2014. **14**(6): p. 1271-1280.
3. Biswas, C., P. Dey, and S. Satpathy, *A method of direct PCR without DNA extraction for rapid detection of begomoviruses infecting jute and mesta.* Letters in Applied Microbiology, 2014. **58**(4): p. 350-355.
4. Satya, P., et al., *Rapid and inexpensive NaOH based direct PCR for amplification of nuclear and organelle DNA from ramie (Boehmeria nivea), a bast fibre crop containing complex polysaccharides.* Industrial Crops and Products, 2013. **50**: p. 532-536.
5. Ohta, S., K. Yano, and Y. Kurita, *A Sample Preparation Method for Direct and Non-direct PCR in Woody Plants.* Journal of the Japanese Society for Horticultural Science, 2013. **82**(1): p. 14-21.
6. Kranzfelder, P., T. Ekrem, and E. Stur, *Trace DNA from insect skins: a comparison of five extraction protocols and direct PCR on chironomid pupal exuviae.* Molecular Ecology Resources, 2016. **16**(1): p. 353-363.
7. Aboud, M., H.H. Oh, and B. McCord, *Rapid direct PCR for forensic genotyping in under 25 min.* Electrophoresis, 2013. **34**(11): p. 1539-1547.
8. Hassab., et al., *Rapid and simple forensic technique for detection of DNA profiles from saliva samples obtained from different individuals and loaded directly into the PCR.* The Egyptian Journal of Experimental Biology (Zoology), 2014. **10**(2): p. 145-160.
9. Gray, K., D. Crowle, and P. Scott, *Direct amplification of casework bloodstains using the Promega PowerPlex® 21 PCR Amplification System.* Forensic Science International: Genetics, 2014. **12**: p. 86-92.
10. Sorensen, A., et al., *Direct-to-PCR tissue preservation for DNA profiling.* International Journal of Legal Medicine, 2016. **130**(3): p. 607-613.

11. Kitpipit, T., et al., *Low-cost direct PCR for aged and processed wildlife sample analysis.* Forensic Science International: Genetics Supplement Series, 2013. **4**(1): p. e71-e72.
12. Kitpipit, T., et al., *Forensic animal DNA analysis using economical two-step direct PCR.* Forensic Science, Medicine, and Pathology, 2014. **10**(1): p. 29-38.
13. Foran, D.R., M.E. Gehring, and S.E. Stallworth, *The recovery and analysis of mitochondrial DNA from exploded pipe bombs.* Journal of Forensic Sciences, 2009. **54**(1): p. 90-94.
14. Linacre, A., and Shanan Tobe, *Wildlife DNA analysis: Applications in Forensic Science*. 1st Edition ed. 2013: John Wiley & Sons.
15. Linacre, A. and J. Templeton, *Forensic DNA profiling: state of the art.* Research Reports Forensic Medical Science, 2014. **4**: p. 25-36.
16. Altshuler, H. and R. Roy, *Evaluation of direct PCR amplification using various swabs and washing Reagents.* Journal of Forensic Sciences, 2015. **60**(6): p. 1542-1552.
17. Thomasma, S.M. and D.R. Foran, *The influence of swabbing solutions on DNA recovery from touch samples.* Journal of Forensic Sciences, 2013. **58**(2): p. 465-469.
18. McNevin, D., et al., *Reduced reaction volumes and increased Taq DNA polymerase concentration improve STR profiling outcomes from a real-world low template DNA source: telogen hairs.* Forensic Science, Medicine, and Pathology, 2015. **11**(3): p. 326-338.
19. Meakin, G. and A. Jamieson, *DNA transfer: review and implications for casework.* Forensic Science International: Genetics, 2013. **7**(4): p. 434-443.
20. Gill-King, H., *Chemical and ultrastructural aspects of decomposition.* Forensic Taphonomy: The Postmortem Fate of Human Remains. CRC, Boca Raton, FL, 1997: p. 93-108.
21. Alaeddini, R., *Forensic implications of PCR inhibition - A review.* Forensic Science International Genetics, 2012. **6**(3): p. 297-305.
22. van Oorschot, R.A., G. Glavich, and R.J. Mitchell, *Persistence of DNA deposited by the original user on objects after subsequent use by a second person.* Forensic Science International: Genetics, 2014. **8**(1): p. 219-225.

23. van Oorschot, R.A., et al., *DNA transfer: the role of temperature and drying time.* Legal Medicine, 2014. **16**(3): p. 161-163.
24. Verdon, T.J., R.J. Mitchell, and R.A. van Oorschot, *Evaluation of tapelifting as a collection method for touch DNA.* Forensic Science International: Genetics, 2014. **8**(1): p. 179-186.
25. Cale, C.M., et al., *Could Secondary DNA Transfer Falsely Place Someone at the Scene of a Crime?* Journal of Forensic Sciences, 2016. **61**(1): p. 196-203.
26. Goray, M., et al., *Secondary DNA transfer of biological substances under varying test conditions.* Forensic Science International: Genetics, 2010. **4**(2): p. 62-67.
27. Goray, M. and R.A. van Oorschot, *The complexities of DNA transfer during a social setting.* Legal Medicine, 2015. **17**(2): p. 82-91.
28. Goray, M., et al., *Shedder status - An analysis of self and non-self DNA in multiple handprints deposited by the same individuals over time.* Forensic Science International: Genetics, 2016. **23**: p. 190-196.
29. Farmen, R.K., et al., *Assessment of individual shedder status and implication for secondary DNA transfer.* Forensic Science International: Genetics Supplement Series, 2008. **1**(1): p. 415-417.
30. Helmus, J., T. Bajanowski, and M. Poetsch, *DNA transfer—a never ending story. A study on scenarios involving a second person as carrier.* International Journal of Legal Medicine, 2016. **130**(1): p. 121-125.
31. Lacerenza, D., et al., *A molecular exploration of human DNA/RNA co-extracted from the palmar surface of the hands and fingers.* Forensic Science International: Genetics, 2016. **22**: p. 44-53.
32. Ladd, C., et al., *A systematic analysis of secondary DNA transfer.* Journal of Forensic Sciences, 1999. **44**: p. 1270-1272.
33. Poy, A.L. and R.A. van Oorschot, *Trace DNA presence, origin, and transfer within a forensic biology laboratory and its potential effect on casework.* Journal of Forensic Identification, 2006. **56**(4): p. 558.
34. Fox, A., et al., *Exploring the recovery and detection of messenger RNA and DNA from enhanced fingermarks in blood.* Science & Justice, 2014. **54**(3): p. 192-198.

35. Haines, A.M., et al., *Detection of DNA within fingermarks.* Forensic Science International: Genetics Supplement Series, 2013. **4**(1): p. e290-e291.
36. Haines, A.M., et al., *Properties of nucleic acid staining dyes used in gel electrophoresis.* Electrophoresis, 2015. **36**(6): p. 941-944.
37. Haines, A.M., et al., *Finding DNA: Using fluorescent in situ detection.* Forensic Science International: Genetics Supplement Series, 2015. **5**: p. e501-e502.
38. Haines, A.M., et al., *Successful direct STR amplification of hair follicles after nuclear staining.* Forensic Science International: Genetics Supplement Series, 2015. **5**: p. e65-e66.
39. Ottens, R., et al., *Successful direct amplification of nuclear markers from a single hair follicle.* Forensic Science, Medicine, and Pathology, 2013. **9**(2): p. 238-243.
40. Ottens, R., et al., *Optimising direct PCR from anagen hair samples.* Forensic Science International: Genetics Supplement Series, 2013. **4**(1): p. e109-e110.
41. Blackie, R., D. Taylor, and A. Linacre, *Successful direct amplification of nuclear markers from single dog hairs using Dog Filer multiplex.* Electrophoresis, 2015. **36**(17): p. 2082-2085.
42. Blackie, R., D. Taylor, and A. Linacre, *DNA profiles from clothing fibers using direct PCR.* Forensic Science, Medicine, and Pathology, 2016. **12**(3): p. 331-335.
43. Ottens, R., D. Taylor, and A. Linacre, *DNA profiles from fingernails using direct PCR.* Forensic Science, Medicine, and Pathology, 2015. **11**(1): p. 99-103.
44. *Reed & Anor, R v EWCA Crim 2698*. 2009; Case 2007/04708/B. Available from: http://www.bailii.org/ew/cases/EWCA/Crim/2009/2698.html.
45. *Nucleic acid collection device and method - AU2016902269*. 2016; Patent application. Available from: http://www.ipaustralia.com.au/applicant/flinders-university-of-south-australia/patents/AU2016902269/.
46. Berti, A., et al., *The recovery of DNA profiles from saliva and touch evidences after postal bomb explosion.* Forensic Science International: Genetics Supplement Series, 2011. **3**(1): p. e471-e472.

47. Horsman-Hall, K.M., et al., *Development of STR profiles from firearms and fired cartridge cases.* Forensic Science International: Genetics, 2009. **3**(4): p. 242-250.
48. Montpetit, S. and P. O'Donnell, *An optimized procedure for obtaining DNA from fired and unfired ammunition.* Forensic Science International: Genetics, 2015. **17**: p. 70-74.
49. Polley, D., et al., *An investigation of DNA recovery from firearms and cartridge cases.* Canadian Society of Forensic Science Journal, 2006. **39**(4): p. 217-228.
50. Gashi, B., et al., *Measurement of 9mm cartridge case external temperatures and its forensic application.* Forensic Science International, 2010. **200**(1): p. 21-27.
51. Xu, Y., et al., *Design and manufacture of surface textures on gun cartridge cases to trap DNA material.* Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture, 2010. **224**(8): p. 1229-1238.
52. Sanchez, J.J., et al., *A multiplex assay with 52 single nucleotide polymorphisms for human identification.* Electrophoresis, 2006. **27**(9): p. 1713-1724.
53. Musgrave-Brown, E., et al., *Forensic validation of the SNPforID 52-plex assay.* Forensic Science International: Genetics, 2007. **1**(2): p. 186-190.
54. Sanchez, J.J., et al., *Forensic typing of autosomal SNPs with a 29 SNP-multiplex—results of a collaborative EDNAP exercise.* Forensic Science International: Genetics, 2008. **2**(3): p. 176-183.
55. Walsh, S., et al., *Developmental validation of the IrisPlex system: determination of blue and brown iris colour for forensic intelligence.* Forensic Science International: Genetics, 2011. **5**(5): p. 464-471.
56. Walsh, S., et al., *Developmental validation of the HIrisPlex system: DNA-based eye and hair colour prediction for forensic and anthropological usage.* Forensic Science International: Genetics, 2014. **9**: p. 150-161.
57. Alimat, S., S. Hadi, and W. Goodwin, *SNP genotyping of forensic casework samples using the 52 SNPforID markers.* Forensic Science International: Genetics Supplement Series, 2013. **4**(1): p. e178-e179.

58. Børsting, C., et al., *Performance of the SNPforID 52 SNP-plex assay in paternity testing.* Forensic Science International: Genetics, 2008. **2**(4): p. 292-300.
59. Bouslimani, A., et al., *Lifestyle chemistries from phones for individual profiling.* Proceedings of the National Academy of Sciences, 2016. **113**(48): p. E7645-E7654.
60. Brotherton, P., et al., *Neolithic mitochondrial haplogroup H genomes and the genetic origins of Europeans.* Nature communications, 2013. **4**: p. 1764.
61. Templeton, J.E., et al., *DNA capture and next-generation sequencing can recover whole mitochondrial genomes from highly degraded samples for human identification.* Investigative Genetics, 2013. **4**(1): p. 1-13.
62. Fordyce, S.L., et al., *Second-generation sequencing of forensic STRs using the Ion Torrent™ HID STR 10-plex and the Ion PGM™.* Forensic Science International: Genetics, 2015. **14**: p. 132-140.
63. Daniel, R., et al., *A SNaPshot of next generation sequencing for forensic SNP analysis.* Forensic Science International: Genetics, 2015. **14**: p. 50-60.
64. Xavier, C. and W. Parson, *Evaluation of the Illumina ForenSeq™ DNA Signature Prep Kit–MPS forensic application for the MiSeq FGx™ benchtop sequencer.* Forensic Science International: Genetics, 2017. **28**: p. 188-194.

Appendices - Chapter VI

Additional manuscripts

Below are a list of manuscripts that were published within the course of this PhD candidature that focus on the application of massive parallel sequencing technology using DNA recovered from human and animal remains. DNA was considered chemically damaged, degraded, and present in low quantities, and these techniques could be applied to other forensic applications.

Manuscript (2013):

1) **Templeton, J.E**., Brotherton, P.M., Llamas, B., Soubrier, J., Haak, W., Cooper, A. and Austin, J.J., 2013. DNA capture and next-generation sequencing can recover whole mitochondrial genomes from highly degraded samples for human identification. ***Investigative genetics***, *4*(1), p.1.

Summary of manuscript:

Mitochondrial DNA (mtDNA) typing can be a useful tool to aid human identification when nuclear DNA is too damaged, degraded or low-template. Standard mtDNA typing (i.e. sequencing HVSI and HVSII control regions) disregards 70% of the variation that exists in the mtDNA genome. The method described in this manuscript first 'immortalises' the valuable DNA extract by creating DNA libraries, followed by in-solution based DNA hydridisation to target and enrich mtDNA where a large fraction of the DNA is damaged and degraded (i.e. < 100 bp in length). Enriched DNA was then characterised by Next Generation Sequencing (NGS) and high-resolution data was used to aid the identification of a World War II soldier and other ancient human post-mortem remains.

Manuscript (2013):

2) Brotherton, P., Haak, W., **Templeton, J**., Brandt, G., Soubrier, J., Adler, C.J., Richards, S.M., Der Sarkissian, C., Ganslmeier, R., Friederich, S. and Dresely, V., 2013. Neolithic mitochondrial haplogroup H genomes and the genetic origins of Europeans. ***Nature communications***, *4*, p.1764.

Summary of manuscript:

To throw light on the genetic origins of present day Europeans, ancient DNA was recovered from 39 ancient skeletal remains across a number of archaeological cultures (~ 7,500-4,000 years old) at a single site in Germany. Whole mitochondrial genome sequence data from haplogroup H individuals (a haplogroup virtually absent in Mesolithic hunter-gatherers, but now carried by up to 45% of Europeans) was used to reconstruct human evolutionary history and reveal dramatic migratory events; including the genetic origins of Germany's earliest farmers in Neolithic Turkey and the Near East. This was the first ancient population study published where a large sample size had been used for next generation sequencing (i.e. 39 ancient bones and teeth).

Manuscript (2014):

3) Der Sarkissian, C., Brotherton, P., Balanovsky, O., **Templeton, J.E**., Llamas, B., Soubrier, J., Moiseyev, V., Khartanovich, V., Cooper, A., Haak, W. and Genographic Consortium, 2014. Mitochondrial genome sequencing in Mesolithic North East Europe Unearths a new sub-clade within the broadly distributed human haplogroup C1. ***PLoS One***, *9*(2), p.e87612.

Summary of manuscript:

Whole mitochondrial genomes were retrieved from ancient remains and sequenced by next generation DNA to establish phylogenetic relationships that could not be defined using control region sequence data. High-resolution data characterized a new distinct clade of the Mesolithic C1 haplogroup; coined "C1f".

Manuscript (2015):

4) Llamas, B., Brotherton, P., Mitchell, K.J., **Templeton, J.E**., Thomson, V.A., Metcalf, J.L., Armstrong, K.N., Kasper, M., Richards, S.M., Camens, A.B. and Lee, M.S., 2015. Late Pleistocene Australian marsupial DNA clarifies the affinities of extinct megafaunal kangaroos and wallabies. ***Molecular biology and evolution***, *32*(3), pp.574-584.

Summary of manuscript:

Hybridization DNA-capture methodology - developed at the Australian Centre for Ancient DNA by Paul Brotherton and Jennifer Templeton - was used to recover

mtDNA sequence data and reconstruct the evolutionary history of the Sthenurine kangaroos. Partial mitochondrial genomes were retrieved from two extinct macropods (46-50 ka and 40-45 ka); the oldest known Australian fossil to generate DNA data.

Manuscript (2015):

5) Santos, C., Fondevila, M., Ballard, D., Banemann, R., Bento, AM., Børsting, C., Branicki, W., Brisighelli, F., Burrington, M., Capal, T., Chaitanya, L., Daniel, R., Decroyer, V., England, R., Gettings, KB., Gross, TE., Haas, C., Harteveld, J., Hoff-Olsen, P., Hoffmann, A., Kayser, M., Kohler, P., Linacre, A., Mayr-Eduardoff, M., McGovern, C., Morling, N., O'Donnell, G., Parson, W., Pascali, VL., Porto, MJ., Roseth, A., Schneider, PM., Sijen, T., Stenzl, V., Court, DS., **Templeton, JE.**, Turanska, M., Vallone, PM., van Oorschot, RA., Zatkalikova, L., Carracedo, Á., Phillips, C.; EUROFORGEN-NoE Consortium. 2015. Forensic ancestry analysis with two capillary electrophoresis ancestry informative marker (AIM) panels: Results of a collaborative EDNAP exercise. ***Forensic Science International: Genetics***, *19*, pp.56-67.

Summary of manuscript:

A collaborative inter-laboratory exercise took place where participants from 19 laboratories assessed a panel of 34 ancestry markers and 46 INDELs using PCR-based SNP methodology and CE detection, with the capability of detecting mixtures.

Manuscript 1: DNA capture and next-generation sequencing can recover whole mitochondrial genomes from highly degraded samples for human identification.

Templeton *et al. Investigative Genetics* 2013, **4**:26
http://www.investigativegenetics.com/content/4/1/26

Investigative Genetics

METHODOLOGY Open Access

DNA capture and next-generation sequencing can recover whole mitochondrial genomes from highly degraded samples for human identification

Jennifer E L Templeton[1,2], Paul M Brotherton[1,3], Bastien Llamas[1], Julien Soubrier[1], Wolfgang Haak[1], Alan Cooper[1] and Jeremy J Austin[1,4*]

Abstract

Background: Mitochondrial DNA (mtDNA) typing can be a useful aid for identifying people from compromised samples when nuclear DNA is too damaged, degraded or below detection thresholds for routine short tandem repeat (STR)-based analysis. Standard mtDNA typing, focused on PCR amplicon sequencing of the control region (HVS I and HVS II), is limited by the resolving power of this short sequence, which misses up to 70% of the variation present in the mtDNA genome.

Methods: We used in-solution hybridisation-based DNA capture (using DNA capture probes prepared from modern human mtDNA) to recover mtDNA from post-mortem human remains in which the majority of DNA is both highly fragmented (<100 base pairs in length) and chemically damaged. The method 'immortalises' the finite quantities of DNA in valuable extracts as DNA libraries, which is followed by the targeted enrichment of endogenous mtDNA sequences and characterisation by next-generation sequencing (NGS).

Results: We sequenced whole mitochondrial genomes for human identification from samples where standard nuclear STR typing produced only partial profiles or demonstrably failed and/or where standard mtDNA hypervariable region sequences lacked resolving power. Multiple rounds of enrichment can substantially improve coverage and sequencing depth of mtDNA genomes from highly degraded samples. The application of this method has led to the reliable mitochondrial sequencing of human skeletal remains from unidentified World War Two (WWII) casualties approximately 70 years old and from archaeological remains (up to 2,500 years old).

Conclusions: This approach has potential applications in forensic science, historical human identification cases, archived medical samples, kinship analysis and population studies. In particular the methodology can be applied to any case, involving human or non-human species, where whole mitochondrial genome sequences are required to provide the highest level of maternal lineage discrimination. Multiple rounds of in-solution hybridisation-based DNA capture can retrieve whole mitochondrial genome sequences from even the most challenging samples.

Keywords: Mitochondrial DNA, Degraded DNA, Ancient DNA, DNA hybridisation, DNA enrichment, Forensic science, Next-generation sequencing

* Correspondence: jeremy.austin@adelaide.edu.au
[1]Australian Centre for Ancient DNA, School of Earth and Environmental Sciences, The University of Adelaide, Adelaide, South Australia 5005, Australia
[4]Sciences Department, Museum Victoria, Carlton Gardens, Melbourne, Vic 3001, Australia
Full list of author information is available at the end of the article

Background

Nuclear DNA short tandem repeat (STR) profiling is currently the preferred method for human identification in forensic practice [1]. However, analysis of low copy number (LCN) and highly damaged or degraded DNA from trace sources or poorly preserved human remains is challenging due to stochastic effects and can often fail completely [2]. Typical complications observed in the analysis of trace amounts of DNA include issues with contamination, amplification failure (allele and locus dropout) [3], preferential amplification of shorter amplicons [4] and artefacts (enzymatic stutter, allele drop-in and off-ladder peaks). Complete amplification failure can be due to PCR inhibition or the fragmentation of all DNA templates below target amplicon sizes, which generally range from 100 to 400 base pairs (bp) [5]. Another complication – 'jumping PCR' – can generate non-authentic chimeric amplicons from discrete DNA template molecules, particularly when DNA fragmentation levels are high [6,7]. Additionally, chemical DNA modification due to miscoding lesions can terminate amplification reactions by halting DNA polymerase extension [7]. A combination of all these factors can lead to a poor or misleading DNA profile, or no profile at all in extreme cases.

The development and optimisation of nuclear SNP (single nucleotide polymorphism) typing protocols, shorter amplicon commercial STR kits (mini-STRs) [8], optimisation of PCR conditions, capillary electrophoresis and statistical interpretation techniques [9] have improved standard profiling methods [10-12]. However, in spite of these developments, the STR profiling of degraded, low-template DNA often has limited success. Furthermore, a large number of nuclear SNPs are required (50 to 80 loci) to obtain a similar level of discrimination as a full nuclear 16-loci STR profile [13]. In these cases, genetic identification from degraded samples may succeed through the analysis of mitochondrial DNA (mtDNA).

Mitochondrial DNA has several features that can make it a useful marker for human identification. As there can be thousands of copies of the mitochondrial genome in many cells (compared to only two copies for autosomal nuclear DNA), mtDNA typing is well suited to biological specimens where DNA fragmentation has occurred or the total DNA copy numbers are naturally low or have been severely reduced through post-mortem damage and degradation [14]. Suitable materials include bones, teeth, hair shafts, faeces and other biological materials. The lack of recombination events in the mtDNA genome and strict uniparental inheritance, in contrast to the nuclear genome, can allow maternal relatives separated by several generations to serve as reference samples. This latter feature is particularly beneficial in missing-person identification, where suitable ante-mortem or family reference samples may be unavailable.

Standard PCR-based sequencing approaches for mitochondrial hypervariable regions I and II (HVS I and II) typically amplify 2 to 12 overlapping fragments of approximately 150 bp to 600 bp in length [15-18] but are labour intensive, consume significant amounts of valuable DNA extract and can be template-length dependent and costly. Repeated singleplex PCR amplifications also bring an increased risk of contamination with exogenous human DNA due to the multiple lab steps required. Multiplex PCR amplification could in theory provide a solution for medium-sized PCR target fragments but still require hundreds of overlapping amplicons [1] in cases where whole mitochondrial genome sequences are needed for high-resolution identification.

Another disadvantage of typing just the mitochondrial HVS I/II is that short sequences from this single locus are far less powerful for identification purposes than a full multi-locus STR profile [19]. This can become a significant problem when many individuals in a population share a common haplogroup, such as the >40% of Western Europeans who belong to mitochondrial haplogroup H, or when distantly related individuals share a maternal ancestry that may not be known [20]. Recent studies sequencing whole mitochondrial genomes have shown that >70% of the mtDNA variation can be located outside HVS I/II for some haplogroups [21], so that full mitochondrial genome sequences provide far greater resolving power for human identification [22,23].

Ancient DNA studies of human archaeological samples routinely generate complete mitochondrial genomes via DNA hybridisation-based enrichment of mtDNA target sequences [21,24-26], and the creation of barcoded/indexed DNA libraries, followed by next-generation sequencing (NGS). Multiple samples can be processed in parallel in a high-throughput fashion [25], greatly reducing processing contamination risks, labour and costs compared to traditional Sanger sequencing approaches. These kinds of DNA capture strategies generally rely on the hybridisation of target DNA sequences to probes that are either immobilised on a surface (such as a microarray) or in solution [27,28]. Despite the significant potential of these new approaches, they have not been applied or examined in a forensic context for human identification.

The aim of this study was to sequence whole mitochondrial genomes from a range of human skeletal samples (in this case ranging from 10 to 2,500 years old) at an affordable cost using standard laboratory equipment and home-made DNA-capture probes for use in hybridisation-based target enrichment (Figure 1). Our previous application of this method [21] used three rounds of in-solution capture-based enrichment so we also aimed to explore the efficacy of using one or two rounds of enrichment to reduce costs and improve workflows. We deliberately used samples that had previously failed or had the potential to

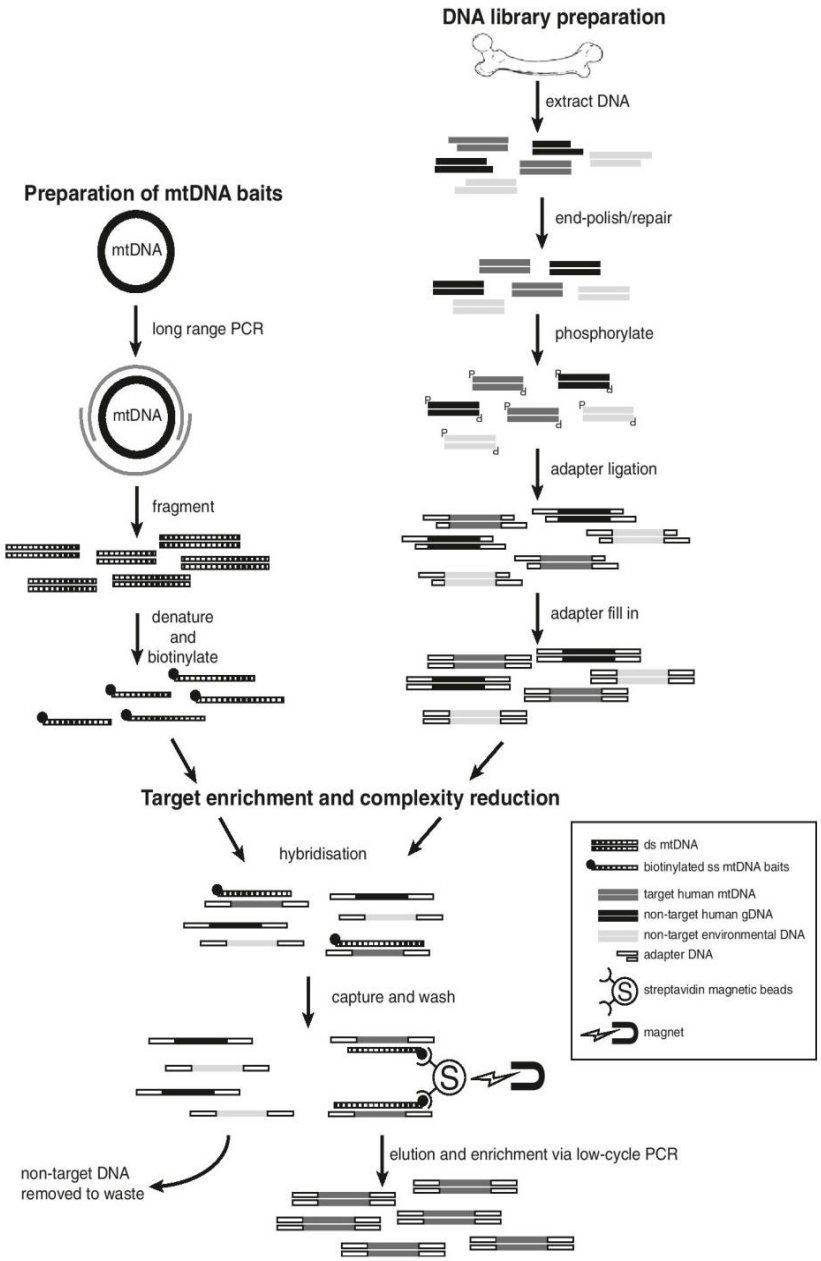


Figure 1 mtDNA hybridisation enrichment protocol. mtDNA baits were prepared using 2 × 8 kb mtDNA long-range PCR products (spanning 16,569 base pairs), generated using a DNA extract from a present-day sample (of a known haplotype) and the Roche LR Expand PCR kit. The products were fragmented by physical shearing to create 200 bp to 600 bp fragments prior to end labelling with biotin. The DNA library was prepared as follows. Damaged DNA leaves 5′ and 3′ overhangs. T4 DNA polymerase was used to polish the DNA by creating blunt ends and T4 PNK phosphorylated 5′ ends, which is required for adaptor ligation. T4 ligase attached universal hybridisation adaptors (Uni-hyb A and Uni-hyb B) to the phosphorylated ends. Klenow polymerase filled in the short-arm adaptor ligation to create double-stranded adaptors (through the use of deoxyribonucleotide triphosphates - dNTPs). Adaptor complementary primers and Taq polymerase amplified the entire library to immortalise the sample. Single-stranded probe DNA was mixed with single-stranded library DNA and left to hybridise overnight (in the presence of blocking oligos). Biotinylated probe and bound library DNA were fixed to streptavidin beads on a magnetic rack, and non-specific or weakly bound library DNA was washed away through a series of three stringency washes (by increasing temperature and decreasing salt concentration) from the library–probe–streptavidin interaction. The single-stranded library DNA was converted to double stranded DNA and eluted from the probe–streptavidin interaction using the Bst strand-displacing enzyme (in the presence of dNTPs). Bst recognises nicks in the template and displaces library DNA into solution. Probe DNA remained bound to the magnet. Eluted library DNA was enriched through low cycle PCR, using adaptor complementary primers. Library DNA was then prepared for next-generation sequencing. mtDNA, mitochondrial DNA; PCR, polymerase chain reaction.

fail nuclear STR typing (Table 1). STR profiling was predominantly performed to assess the likelihood of obtaining full STR profiles from degraded samples and not to identify the samples. To identify the samples would require reference profiles for comparison and replicate testing of the samples by LCN analyses. The capture-probe method is designed to focus on the recovery of human mtDNA fragments <100 bp in length (with the vast majority in the 20 to 70 bp range) (Figure 2), from samples that yield highly damaged and fragmented DNA templates available only with low copy number. We anticipate the method will be useful for samples that cannot be typed successfully using standard STR kits and for detecting key or private SNPs within whole mitochondrial genome sequences (that would otherwise remain undetected with traditional mtDNA HVS I/II sequencing) for human identification.

Methods

Samples

Bone and tooth samples were selected representing a range of ages, preservation conditions and contexts (Table 1). Three samples were from missing-person cases, two of which were from Australian servicemen killed in World War II. Two samples were recovered from archaeological contexts.

Degraded DNA work

To avoid the potential for contamination of samples with contemporary human DNA or previously amplified PCR products, all steps preceding DNA library amplification were carried out in a dedicated ancient DNA laboratory geographically separated (by approximately 1.5 km) from post-PCR and other molecular biology laboratories at the University of Adelaide.

Strict decontamination procedures were followed [29]. There were ultraviolet lights in every room. There was positive air pressure and the one-way airflow was filtered using high efficiency particulate air filters. There were separate workrooms each containing dead-air glove boxes. Equipment and work areas were cleaned with sodium hypochlorite and isopropanol before and after use. Personnel protective clothing included a full-body suit, face mask, face shield, boots and triple-gloves. There was a strict one-way movement of personnel (from shower to freshly laundered clothes to ancient DNA laboratory to post-PCR laboratory).

Non-template controls and extraction blanks were included in each experiment to monitor potential contamination from exogenous human DNA sources and cross-contamination from other samples. The complete mitochondrial genomes of all staff involved in the handling of the samples (JT, WH, BL and PB) were sequenced to monitor potential contamination (Additional file 1: Table S1). The mitochondrial genome of an anonymised present-day sample used to generate mtDNA capture probes was also fully sequenced (haplotype: J1c8a) to monitor contamination (Additional file 1: Table S1).

DNA extraction, quantification and STR profiling

DNA was extracted from five samples using an optimised method as previously described [21]; see Additional file 1.

For sample 9210A, a small quantity of DNA extract was available so only mitochondrial testing was performed. The four additional samples (4464B, 10730A, 8727C and 11995A) were subjected to both nuclear STR and mitochondrial sequence analyses. Nuclear DNA and mtDNA were quantified in all extracts using quantitative PCR (qPCR) with SYBR® green chemistry and previously published 67 bp nuclear [30] and 77 bp mitochondrial [31] PCR targets (Additional file 1: Table S5). The total 10 μL qPCR reaction mix consisted of 1× Brilliant II SYBR® green master mix (Agilent Technologies, USA), 0.15 μM forward primer (Additional file 1: Table S5), 0.15 μM reverse primer (Additional file 1: Table S5), 400 ng/μL rabbit serum albumin and 1 μL DNA extract. Samples were run in triplicate, and negative (no template) and positive controls (male genomic control DNA, Applied

Table 1 Samples used in this study

ACAD ID	Description	Age	Excavated	Locality	Environmental conditions of site	Preservation status on collection
8727C	Cranium fragment	Approximately 10 years	Unknown	South-east Queensland, Australia	Surface deposit	Well preserved
11995A	Long bone fragment	Approximately 70 years	2011	Papua New Guinea	Tropical lowland battlefield, wet	Poor
9210A	Tooth	Approximately 70 years	2006	Christmas Island, Australia	Tropical lowland, burial, wet	Well preserved
4464B	Long bone fragment	Approximately 2500 years (Iron Age)	2007	Latsch, South Tyrol, Italy	Temperate, burial	Well preserved
10730A	Long bone fragment	Approximately 600 to 1,000 years (Ychsma culture)	2002	Huaca Pucllana, Lima, Peru	Temperate, high altitude, burial	Well preserved

ACAD, Australian Centre for Ancient DNA.

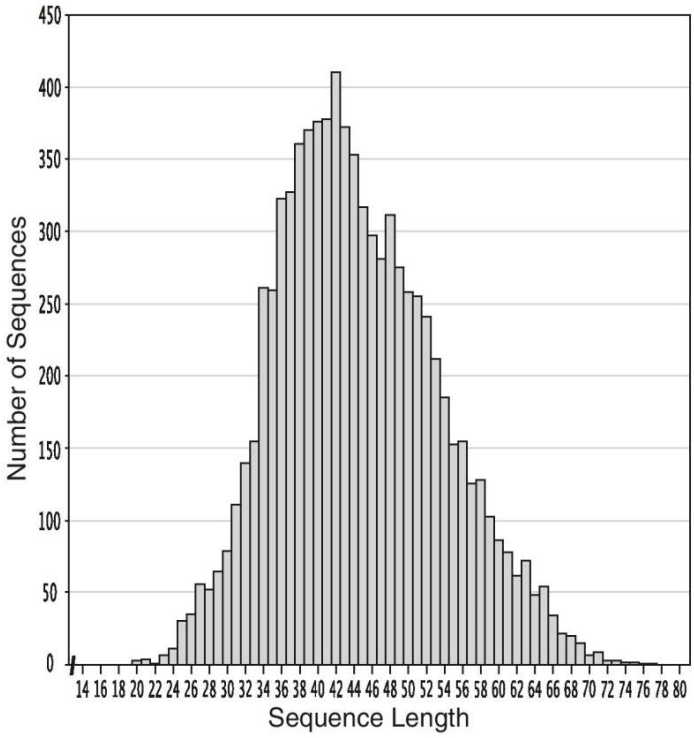


Figure 2 Distribution of the length of human mtDNA reads from sample ID 8727C. In total, 277,470 reads were trimmed to remove primers and the resulting 218,102 mapped (to the Reconstructed Sapiens Reference Sequence) reads were filtered to remove duplicates as a result of clonal reads. The resulting 7,756 unique mapped reads (2.8% of the total reads) gave 99.5% mitochondrial genome sequence (16,487 base pairs) coverage. The mean fold coverage was 20x (SD ± 9.7), mean unique read length was 43.2 bp (SD ± 9.3) and all fragments captured were <80 bp.

Biosystems, USA) were included in all runs. Extraction blank controls were also quantified. Cycling was performed using a Corbett 6000 Rotor-Gene real-time PCR thermocycler and consisted of an initial 5 min denaturation step at 95°C, followed by 45 cycles of 95°C for 10 s, 59°C for 20 s and 72°C for 15 s. Results were analysed using the Rotor-Gene 6000 Series Software 1.7. The DNA concentration was determined using the comparative cycle threshold method where unknown samples are compared to a standard curve. The standard curve for the nuclear target was created using male genomic control DNA (Applied Biosystems, USA). The standard curve for mitochondrial DNA was created using a PCR product (Additional file 1: Table S5).

STR typing was performed using AmpFLSTR ProfilerPlus™ (Applied Biosystems, USA). The final 12.5 μL reaction volume consisted of 4.6 μL ProfilerPlus™ reaction mix, 2.5 μL of ProfilerPlus™ primer mix, 0.4 μL AmpliTaq Gold™ and 5 μL of DNA extract. Cycling was performed on a 9700 GeneAmp thermal cycler and consisted of an initial denaturation at 95°C for 10 min followed by 34 cycles of 94°C for 1 min, 59°C for 1 min and 72°C for 1 min, then a final extension at 60°C for 45 min. PCR products were analysed on a 3130xl Genetic Analyser in a 17.3 μL final volume that consisted of 2 μL of PCR product, 15 μL HiDi™ formamide and 0.3 μL ROX-500 size standard (Applied Biosystems, USA). Results were analysed using Genemapper ID software (v3.2.1). Alleles were interpreted based on peak heights reaching a set threshold value of 25 relative fluorescence units (RFU) above a clean baseline. A wild-card designation was used, with peak heights <150 RFU, to account for potential allele dropout (for example, '11, F' instead of '11,11'). A profile was considered full when all alleles were detected above the threshold RFU. A profile was defined as partial when peaks were detected above the threshold RFU and when at least one locus was successfully called. A profile in forensic terms can be described as partial when at least one locus has been called (even if this is not an informative profile).

Mitochondrial DNA capture and enrichment

Biotinylated DNA-capture probes of a known haplotype and DNA libraries were generated as described in Additional file 1 and as previously described [21]. Whilst Brotherton *et al.* [21] used three rounds of enrichment for all samples, we explored the effects of using one or two rounds of enrichment on the number of unique reads, coverage and sequencing depth. Hybridisation was carried out in a final

volume of 30 μL consisting of 100 ng of probe and 400 ng of library DNA. The thermal profile used was: denature DNA for 5 min at 95°C, followed by 14 to 18 hours incubation at 50°C to allow the DNA-capture probes to hybridise to fragments of DNA with closely matched sequences from complementary human DNA regions. The two library primers (Additional file 1: Table S4) were included as part of the hybridisation mix at 0.67 μM to 1.0 μM, to act as 'blocking' oligonucleotides. The blocking oligonucleotides are complementary to the library adaptors and have a dual role during the hybridisation reaction: (i) to minimise unwanted hybridisation between the adaptor-tagged flanking regions of otherwise unrelated single-stranded library DNA molecules and (ii) to enable strand displacement of probe DNA from library DNA as explained below.

Following overnight hybridisation at 50°C, 50 μL of magnetic streptavidin beads in solution (Invitrogen) were added to 30 μL of hybrid DNA and the beads were immobilised on a magnetic rack. The clear supernatant was discarded. The bead complex (DNA–capture probe/library DNA) immobilised to the magnet was subjected to successively increased-stringency washes, to remove progressively non- or weakly-hybridised single-stranded library DNA molecules, using decreased salt and increased temperature: 2× saline sodium citrate (SSC)/0.1% sodium dodecyl sulphate (SDS) at 37°C for 1 min; 2× SSC/0.1% SDS at 42°C for 10 min; 1× SSC/0.1% SDS at 43°C for 10 min; 0.5× SSC/0.1% SDS at 44°C for 10 min; 0.5× SSC/0.1% SDS at 45°C for 10 min.

The strand-displacing Bst DNA polymerase enzyme (large fragment, New England Biolabs) was used to release library DNA from the DNA-capture probe (immobilised to beads on the magnet). Reactions were performed at 35 μL final volume comprising 1× Thermopol buffer (New England Biolabs), 200 μM of each dNTP (to convert single-stranded library DNA to dsDNA), and 100 μg/mL of Bovine Serum Albumin (New England Biolabs). The reaction was pre-heated to 60°C and 2 μL of Bst enzyme was added last to each reaction. Tubes were incubated at 60°C for 5 min with regular agitation. The reaction tube was then applied to the magnetic rack at 60°C and 35 μL of supernatant was transferred to a fresh PCR tube. This tube was immediately incubated at 80°C for 20 min to inactivate the enzyme.

The heat-inactivated supernatant was split between eight PCR re-amplification reactions (total combined volume 140 μL), designed so that upon the addition of the subportion of Bst buffer, the final composition of the reactions would be 1× AmpliTaq Gold buffer II, 2.5 mM $MgCl_2$, 250 μM of each dNTP, 1.0 U AmpliTaq Gold (Applied Biosystems), and 0.5 μM of PCR primers UniHyb-PCR-A and UniHyb-PCR-B (Additional file 1: Table S4). Thermocycling was at 94°C for 11 min, followed by 12 cycles of 30 s at 95°C, 30 s at 60°C and 1 min (+2 seconds per cycle) at 72°C, followed by a final 10 min at 72°C. Amplification reactions were pooled and library amplicons purified using MinElute spin columns (Qiagen) and eluted into 15 μL as per the manufacturer's instructions. These comprised the 'first enrichment' DNA libraries and amplification products were sized and quantified via gel electrophoresis against size markers (HyperLadder™ V, Bioline) and a Nanodrop 2000 (Thermo Scientific).

For cases where a second round of enrichment took place, the overnight hybridisation and wash steps were repeated to produce 'second enrichment' DNA libraries highly enriched for mtDNA sequences.

Ion Torrent PGM sequencing

Enriched library DNA was prepared for Ion Torrent sequencing by re-amplification using Ion Torrent barcoded primers (Additional file 1: Table S6). Eight 24 μL reaction volumes per sample were re-amplified using 1 μL of purified library DNA as the template. Final reactions conditions comprised of 1× AmpliTaq Gold buffer II, 2.5 mM $MgCl_2$, 2.5 U AmpliTaq Gold (Applied Biosystems), 250 μM of each dNTP (Invitrogen), and 0.5 μM of each PCR primer. The thermocycling profile was 94°C for 12 min, followed by 12 cycles of 30 s at 95°C, 30 s at 60°C and 45 s at 72°C, followed by a final 10 min at 72°C. The eight amplified samples per reaction were pooled and purified using MinElute spin columns (Qiagen) and eluted into 15 μL as per the manufacturer's instructions. The DNA was sized and quantified via gel electrophoresis against size markers (HyperLadder™ V, Bioline) and a Nanodrop 2000 (Thermo Scientific). Library DNA was size-selected above 120 bp and further purified to remove adaptor dimer, using Qiagen's gel extraction purification kit following the manufacturer's instructions.

Prior to sequencing, the fragment size distribution and DNA concentration of individual libraries were measured using a Bioanalyzer 2100 (Agilent Technologies) following the manufacturer's instructions. The quantified indexed library DNA was pooled to an equimolar concentration prior to the One Touch. The pooled library DNA was adjusted to a final concentration of 10 to 15 pM prior to amplification (by emulsion PCR) and enriched for positive ion sphere particles (ISPs) using the Ion Torrent One Touch System II (Life Technologies) and the Ion One Touch 200 template kit v2 DL (Life Technologies), following the manufacturer's instructions.

Templated ISPs were sequenced on a 316 micro-chip (up to 100 Mb of data expected) using the Ion Torrent Personal Genome Machine (PGM; Life Technologies) and the Ion PGM 200 sequencing kit v2 chemistry (Life Technologies) for 130 cycles (520 flows). After sequencing, the individual sequence reads were filtered within the PGM software to remove low-quality and polyclonal sequences.

Sequences matching the PGM 3′ adaptor were also automatically trimmed prior to bioinformatics analysis.

Bioinformatic sequence analysis

Ion Torrent PGM data from the mtDNA capture was processed using a customisable analytical pipeline. The scripts fastx_barcode_splitter.pl and fastx_trimmer (from the FASTX toolkit [32]) were used to demultiplex the reads by barcode, using a strict zero mismatch threshold. Cutadapt v1.1 [33] was then used to trim adapters using a maximum error rate of 0.33 (–e 0.3333), and to remove short (–m 25), long (–M 110) and low-quality sequences (–q 20), for a total of five passes (–n 5). The filtered reads were checked with FastQC [34] before being mapped against the Reconstructed Sapiens Reference Sequence (RSRS) [35] using TMAP v3.2.1 [36] with the following options: –g 3 -M 3 -n 7 -v stage1 –stage-keep-all map1 –seed-length 12 –seed-max-diff 4 stage2 map2 –z-best 5 map3 –max-seed-hits 10. The program TMAP has been optimised to align Ion Torrent PGM reads against a reference genome [37]. Mapped reads with mapping quality below Phred 30 and read duplicates were removed using Samtools v0.1.18 [38] and the MarkDuplicates tool of Picard Tools v1.79 [39]. The GC content of mapped reads was analysed using the CollectGcBiasMetrics tool of Picard Tools v1.79. Misincorporation patterns were assessed using mapDamage v0.3.6 [40]. The resulting sequence assembly was visualised using Biomatters Geneious Pro v5.6.2 software [41] and mitochondrial haplotypes were defined for each individual according to phylotree.org [4].

Confirming SNP calls by hypervariable region I sequencing

HVS I was amplified using a minimum of four short overlapping primer pairs, as previously described [42,43]. Minisequencing of 22 coding region SNPs (GenoCoRe22) using a multiplex and SNaPshot based approach was carried out, as previously described [42,43].

Results

Quantitative PCR on four of the five samples with sufficient DNA extract volume indicate a 14,000 to 300,000-fold difference in the amount of recovered nuclear DNA:mtDNA, highlighting the greater potential for mtDNA typing in degraded remains. The total nuclear DNA in all four samples was very low (<2 pg/μL). Subsequently, all four samples only produced partial STR profiles using low copy number techniques (34 cycles of PCR and reduced reaction volumes with higher concentrations of *Taq* DNA polymerase) (Table 2). Locus dropout was observed in each degraded sample analysed for nuclear STR typing. Only the positive control DNA produced a full STR-DNA profile, which matched the reference profile at all ten loci examined (Table 2). All negative controls were blank.

In contrast, a higher concentration of mitochondrial DNA was detected in all four samples using qPCR (Table 2). DNA library preparation, mtDNA enrichment and NGS were completed for all five samples. After one round of hybridisation and enrichment we obtained 96% to 97% of the mitochondrial genome at an average 15 to 18-fold coverage from two well-preserved samples but only 62% of the mitochondrial genome at an average 1-fold coverage on a poorly preserved sample (Table 3, Figure 3). However, after two rounds of hybridisation and enrichment we obtained 98% to 100% of the mitochondrial genome at an average 1646-fold coverage from all five samples, irrespective of morphological preservation of the sample (Table 3, Figure 3). Complete or near complete mitogenomes were recovered from samples with as few as 350 copies/μL of the 77 bp mtDNA fragment.

Two rounds of enrichment substantially improved the number of unique reads that mapped to the mtDNA genome (from 2- to 11-fold) (Figure 4, Additional file 1: Figure S1) and the average redundancy per site of the genome (from 1 to 18× to 16 to 46×) but did not alter the mean fragment length of mtDNA recovered (42 bp after one round and 45 bp after two rounds) (Table 3). The second round of enrichment proved to be particularly important for the less well-preserved sample 11995A, for which it provided an 11-fold increase in the total number of unique reads, which also substantially improved the coverage from 62% to a near complete mitochondrial genome (98%) and therefore allowed an unambiguous haplotype designation (Table 3).

Coverage was uneven across all five samples after one and two rounds of enrichment (Additional file 1: Figure S1). This variation in coverage has been reported previously for modern and ancient human and Neanderthal mtDNA genomes [5,24,44-46] and is positively correlated with GC content. This may be due to loss (denaturation) of short AT-rich sequences before or during the library preparation [5,46].

Damage patterns in all samples followed expectations for degraded DNA, with a larger than usual amount of deaminated cytosine residues accumulated towards the ends of the molecules. In addition, we could observe a high frequency of indels, a well-identified homopolymer sequencing error characteristic of PGM technology [47-50]. However, indels were randomly distributed and did not affect the final consensus sequences, as each called position was covered with enough depth to prevent false-positive base calls.

Stringent quality filtering during analysis removed a large proportion of the total reads for each sample. Post-filtering provided on average, across all samples, a very small proportion of unique mapped reads vs total reads (0.04% to 2.4%) (Figure 5). However, the pattern of mapped reads had an adequate level of coverage for each sample to allow

Table 2 Mitochondrial and nuclear quantitative PCR (qPCR) and STR typing result

Sample ID	Mitochondrial DNA	Nuclear DNA				
	qPCR	qPCR	qPCR	STR typing (Profiler Plus kit, 10 loci) 34 cycles		
	77 bp target (copies/μL)	67 bp target (copies/μL)	67 bp target (ng/μL)	Number of loci	Percentage genotyping success	None, partial or full profile obtained
4464B	15,727	<1 copy (0.54)	0.001794	8	80%	Partial
10730A	62,592	0	0	Amel only	10%	Partial
8727C	350	<1 copy (0.025)	0.000082	3	30%	Partial
11995A	2,715	<1 copy (0.0091)	0.00003	2	20%	Partial
9210A	n/a	n/a	n/a	n/a	n/a	n/a
Negative control (H20)	0	<1 copy (0.034)	0.000111	0	0%	None
Female positive control	7,896,895	543	1.79213	10	100%	Full

n/a, non-applicable (The sample was excluded from testing due to the low quantity of DNA extract. Quantitative PCR data is presented as a geometric average using triplicate values.).

detection of variants in the mitochondrial genome. Traditional HVS I sequencing and coding region SNP mini-sequencing of all five samples produced identical results to those obtained by whole mtDNA genome sequencing.

Discussion

Low amounts of DNA combined with high levels of damage and fragmentation make STR typing of degraded samples challenging. DNA capture coupled with next-generation sequencing can retrieve whole mitochondrial genome sequences from degraded samples when nuclear DNA is below detection levels. Despite high levels of DNA decay in skeletal remains, whole mtDNA genome sequencing is possible due to the copy-number advantage and reduced rate of fragmentation of mtDNA (compared with nuclear DNA) combined with the ability to capture and sequence DNA fragments in the 20 bp to 70 bp range. Quantitative PCR can be used to determine the amount of DNA available from extracted materials and will indicate the likelihood of obtaining a nuclear STR-DNA profile from a degraded sample. This is of particular importance in cases where total nuclear DNA quantity is <100 pg, which reduces the likelihood of obtaining a full nuclear STR DNA profile even when applying LCN techniques. In contrast, near complete mitochondrial genome sequences can be obtained with a single round of enrichment from samples with >10,000 77 bp mtDNA copies/μL and with two rounds of enrichment from samples with <3,000 77 bp mtDNA copies/μL. Our work builds on previous in-solution capture-based enrichment methodologies [21,24-26] and demonstrates the importance of using multiple rounds of enrichment to improve mtDNA genome recovery from samples with low amounts of endogenous DNA. Repeating the enrichment process on samples with very low amounts of mtDNA can more than double the number of unique reads and average coverage, and substantially improve the overall coverage of the mtDNA genome (Additional file 1: Figure S1). The methodology has the ability to capture DNA templates that are damaged and fragmented (<100 bp in length) (Figure 2) and that are generally difficult to recover using traditional methods

Table 3 Ion Torrent PGM whole mitochondrial DNA sequencing data, after one and two rounds of enrichment

ACAD ID	Rounds capture	Percentage coverage	Average length (± SD)	Number of bases covered	Unique reads	Mean X coverage (± SD)	Haplotype
4464B	1	97.4%	42 (± 11.0)	16,136	7,069	18 (± 12.3)	HV0e
4464B (ancient)	2	99.60%	45 (± 11.7)	16,507	14,837	40 (± 19.8)	HV0e
10730A	1	96.6%	43 (± 11.3)	16,004	5,760	15 (± 10.7)	B2b
10730A (ancient)	2	99.80%	47 (± 12.2)	16,540	16,138	46 (± 20.2)	B2b
11995A	1	62.3%	40 (± 10.2)	10,321	553	1 (± 1.6)	Low coverage
11995A (forensic)	2	97.70%	45 (± 12.9)	16,188	6,012	16 (± 11.2)	H1a
Extraction blank	1	1.5%	41 (± 18.4)	248	6	0	n/a
Extraction blank (EBC11049)	2	0.20%	0	28	1	0	n/a
8727C (forensic)	2	99.50%	43 (± 9.3)	16,487	7,756	20 (± 9.7)	U5a2a1fNEW
9210A (forensic)	2	100%	65 (± 20.4)	16,569	7,865	31 (± 14.1)	J1c12

ACAD, Australian Centre for Ancient DNA; n/a, non-applicable.

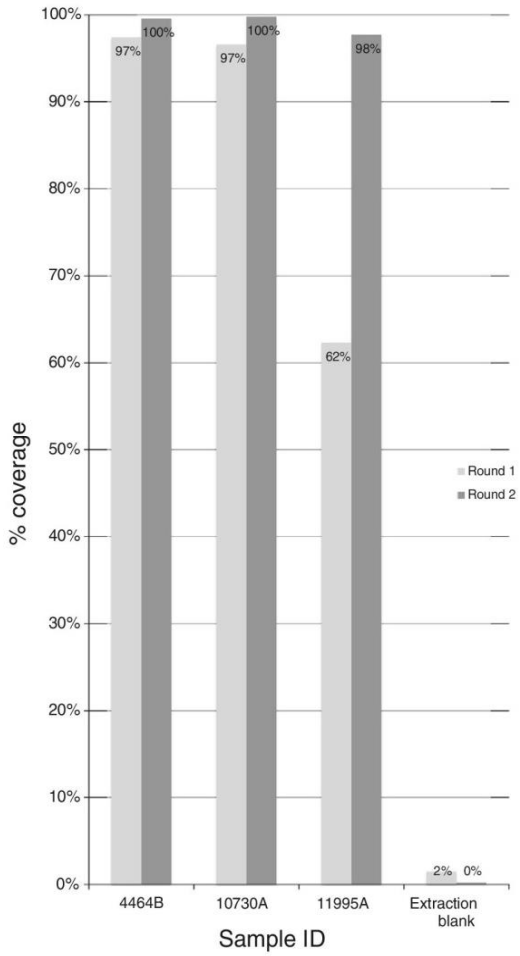


Figure 3 Percent coverage of human mitochondrial genomes sequenced after one (light grey bars) and two (dark grey bars) rounds of enrichment for three degraded samples and an extraction blank.

of PCR-based amplification and sequencing [51]. This is of particular importance in cases where DNA has been exposed to prolonged heat, moisture, ultraviolet light and microbial attack, which generally results in template fragmentation (in extreme cases there can be no surviving endogenous DNA templates >100 bp) [52].

Two common concerns with mitochondrial DNA testing can be eliminated or reduced using this whole mtDNA genome sequencing approach. Traditional HVS I/II sequencing requires 2 to 12 separate PCR amplifications and up to 24 separate DNA sequencing reactions. This multi-tube, multi-step approach introduces the potential for sample mix-up during laboratory processing and increases the risk of introducing contaminating DNA. Our whole mtDNA genome approach eliminates this risk, massively reducing opportunities for sample mix-up, while the barcoded adapters ligated to the DNA provide an additional means to eliminate (or identify and screen out) contamination introduced in later steps. In addition, barcoding allows many samples to be pooled for high-throughput screening efforts and can reduce the cost of sequencing.

DNA capture and related approaches have been shown to give preferential enrichment of short endogenous DNA templates over longer exogenous contaminant DNA in a sample [7]. This is particularly important where small quantities of endogenous DNA in a sample have become saturated by larger quantities of exogenous contamination (human and microbial), consequently leading to poor PCR amplification, mistyping of target loci via artefacts or even complete PCR amplification failure [1].

Traditional forensic and archaeogenetic studies using mtDNA have relied on HVS I/II sequencing. However, this relatively short sequence has limited resolving power and can fail to discriminate between distinct maternal lineages [20]. Outside the control region, coding region SNPs provide additional resolution and discriminatory power [20,53]. To date, this additional information has been obtained via case [20,54], region [55,56], continental [57] or haplogroup [58] specific SNP multiplex assays. In contrast, our whole mtDNA genome sequencing approach is a universal solution for obtaining high-resolution mtDNA data, which can discriminate between closely related maternal lineages. However, although our methodology provides a mechanism to generate whole mtDNA genome sequences from difficult and degraded samples, there is a clear need for the parallel development of high-quality mitochondrial genome databases [20,53].

Complete mitochondrial genomes sequences can aid human identification efforts by placing an individual into specific haplotypes based on private SNPs. This high-resolution discrimination can be used to include or exclude closely related maternal lineages [21], especially in populations with high frequencies of particular haplotypes. By resorting to whole mtDNA sequencing, we were able to gain additional haplogroup and haplotype resolution relative to traditional HVS I/II sequencing. This information has already proved critical in a comparison with maternal relatives in a case where the HVS I/II sequence alone could not exclude a maternal relationship. Our approach could assist large-scale identification efforts when more comprehensive mtDNA reference databases become available to the forensic community.

Validation studies have confirmed that mtDNA typing is a reliable means of forensic identification [59]. However, a worldwidewide effort will be required with labs collaborating and producing large databases, estimating the frequency of particular mtDNA haplotypes and improving the statistical basis of the databases. In the meantime, techniques used to sequence whole mtDNA in archaeological and population studies will continue to advance at a rapid pace.

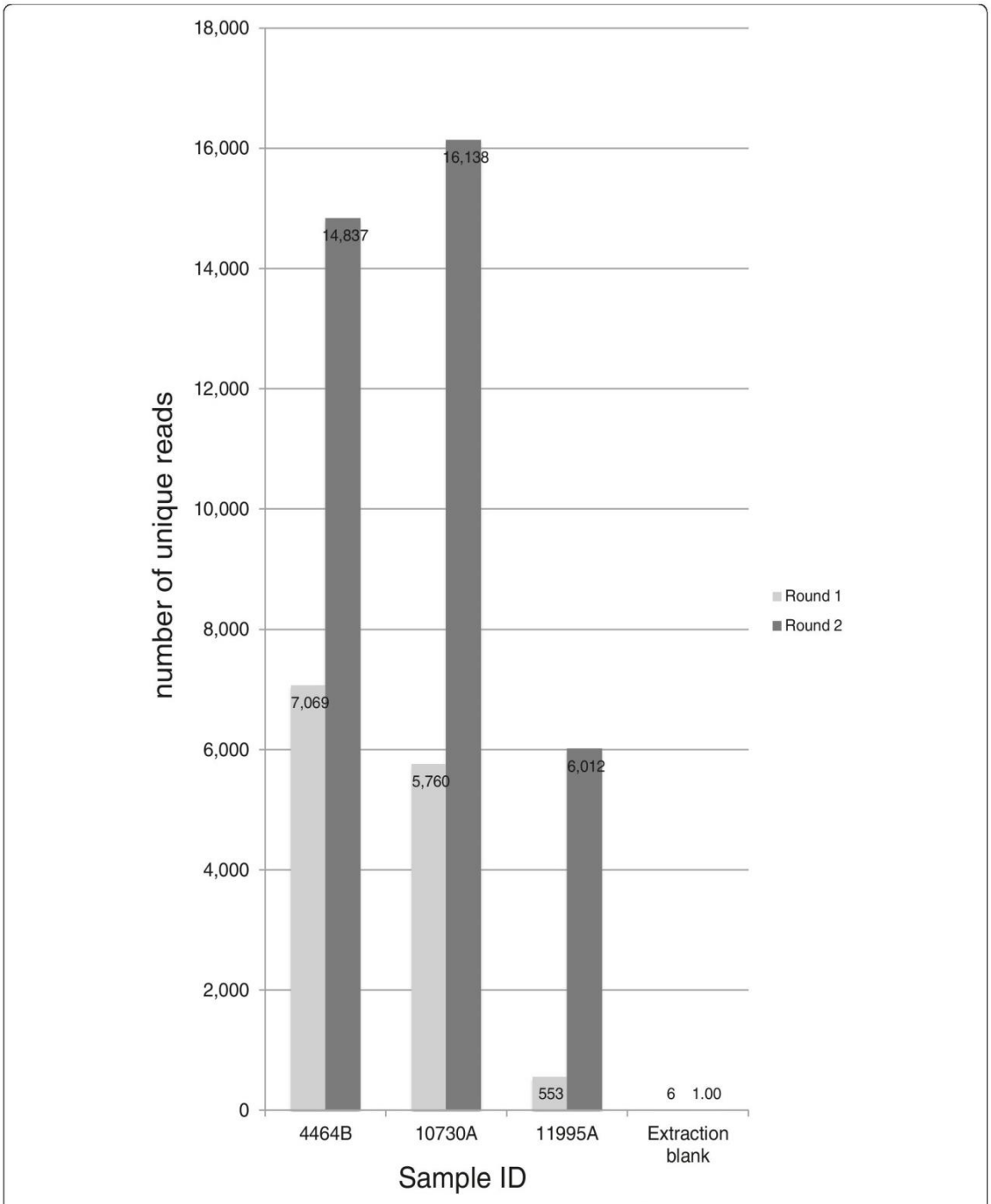


Figure 4 Number of unique sequencing reads mapped to the Reconstructed Sapiens Reference Sequence mitochondrial genome after one (light grey bars) and two (dark grey bars) rounds of enrichment for three degraded samples and an extraction blank.

Conclusions

In-solution capture-based whole mitochondrial genome sequencing immortalises the limited and important contents of the DNA extract in the form of a DNA library, and is followed by targeted enrichment of mtDNA sequences. The application of these methods using hybridisation enrichment and NGS has led to the reliable genotyping of human remains for which standard nuclear PCR

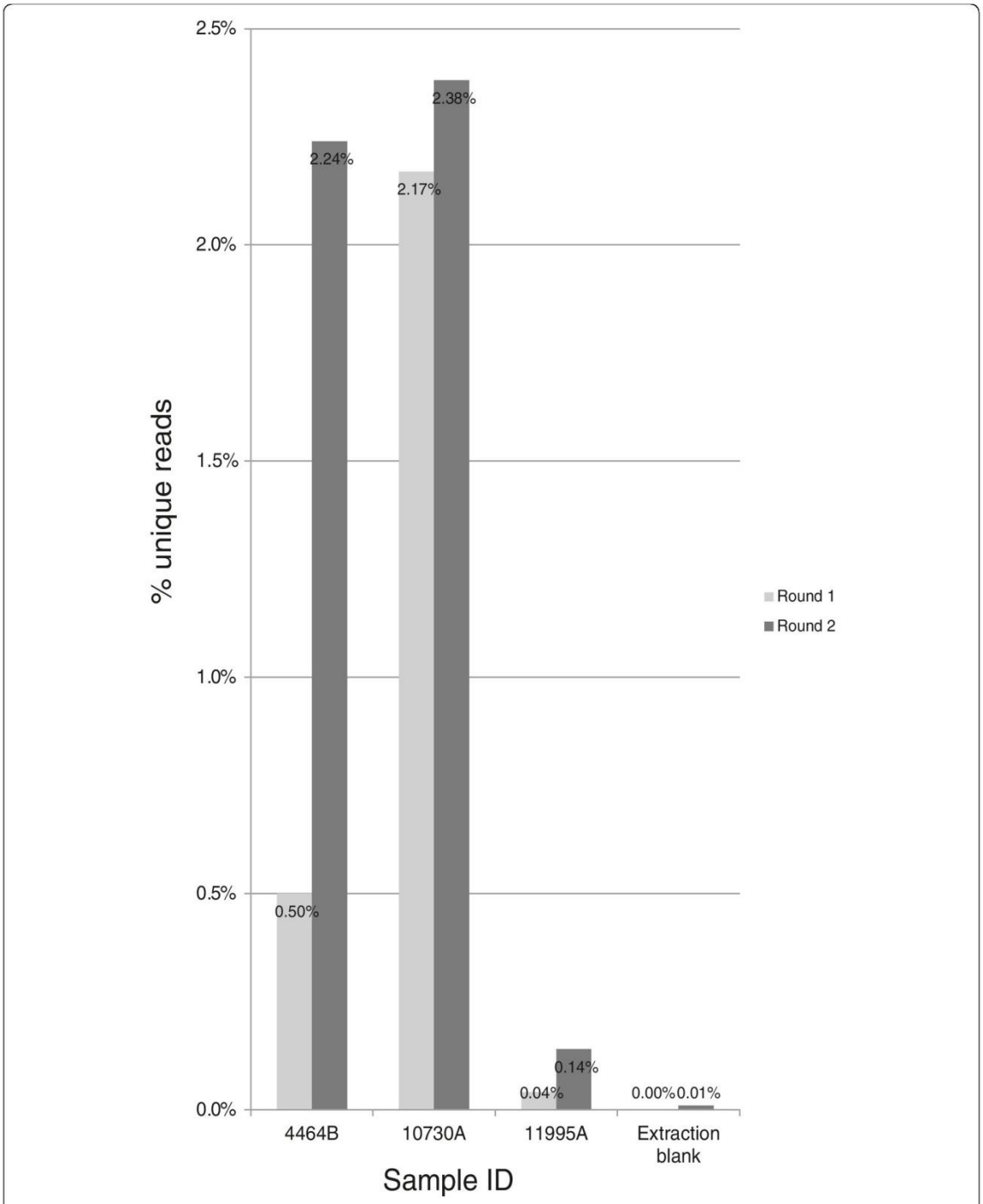


Figure 5 Percentage of unique sequencing reads that map to the Reconstructed Sapiens Reference Sequence mitochondrial genome after one (light grey bars) and two (dark grey bars) rounds of enrichment for three degraded samples and an extraction blank.

protocols had been unsuccessful. This result indicates that the technique can be applied to obtain whole mitochondrial genomes even from particularly challenging samples. Additionally, as NGS platforms become more affordable and widely available and with the advent of DNA library, barcoding (to monitor contamination and allow multiple samples to be processed), new methods for mtDNA analysis should be considered.

Additional file

Additional file 1: Detailed description of the methods used to extract DNA from bone samples, prepare capture-bait library and prepare libraries from degraded DNA. Primer sequences are shown in Tables S2, S3, S4, S5, and S6. Mitochondrial genome haplotypes for laboratory staff and degraded bone samples are shown in Tables S1 and S8, respectively. Details of Ion Torrent sample barcoding and sequencing runs are shown in Table S7. Mapping of individual sequence reads to the reference mitochondrial genome for all five degraded samples are shown in Figure S1.

Abbreviations

ACAD: Australian Centre for Ancient DNA; bp: base pair; HVS: Hypervariable region; ISP: Ion sphere particle; kb: kilobase; LCN: Low copy number; mtDNA: mitochondrial DNA; NGS: Next-generation sequencing; PCR: Polymerase chain reaction; PGM: Personal Genome Machine; qPCR: quantitative PCR; RFU: Relative fluorescence units; SSC: Saline sodium citrate; SDS: Sodium dodecyl sulphate; SNP: Single nucleotide polymorphism; STR: Short tandem repeat.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

JELT co-developed the protocol, processed samples, performed next-generation sequencing, co-analysed data and wrote the manuscript. PB designed and developed the DNA extraction, library preparation and targeted enrichment protocol, co-developed the underlying research concept and assisted manuscript preparation. BL processed sample extraction and co-developed the analytical pipeline for data analysis with JS. BL and JS performed data analysis and assisted with manuscript preparation. WH contributed to the experimental design, provided archaeological samples, processed sample extractions and library preparations and assisted with manuscript preparation. AC co-developed the underlying concept, contributed to the experimental design and assisted with manuscript preparation. JA provided forensic samples, processed sample extractions, contributed to the experimental design and assisted with manuscript preparation. All authors read and approved the final manuscript.

Acknowledgments

We thank the Queensland Police, Australian Navy and the Unrecovered War Casualties – Army (Australian Defence Force) for permission to use samples. Members of the Australian Centre for Ancient DNA, in particular Dr Denice Higgins, provided helpful and constructive comments on the research and manuscript. We thank Dr Hubert Steiner, Dr Isabel Flores Espinoza and Mrs Maria Ines Barreto for providing archaeological samples. Additional thanks to Professor Jeremy Timmis, Rosalie Kenyon, Dr Mark Van der Hoek, Dr Paul Gooding, Dr Steve Richards and Dr Hugh Cross for technical support. This research was funded by an Australian Research Council Linkage project to AC (LP0822622), an ARC Future Fellowship to JA (FT100100108) and an ARC Discovery Project to WH and BL (DP1095782).

Author details

[1]Australian Centre for Ancient DNA, School of Earth and Environmental Sciences, The University of Adelaide, Adelaide, South Australia 5005, Australia. [2]School of Biological Sciences, Flinders University, Bedford Park, Adelaide, South Australia 5001, Australia. [3]Archaeogenetics Research Group, School of Applied Sciences, University of Huddersfield, Queensgate, Huddersfield HD1 3DH, UK. [4]Sciences Department, Museum Victoria, Carlton Gardens, Melbourne, Vic 3001, Australia.

Received: 19 July 2013 Accepted: 18 October 2013
Published: 2 December 2013

References

1. Butler JM: *Advanced Topics in Forensic DNA Typing: Methodology.* London: Academic; 2011.
2. Nandineni MR, Vedanayagam JP: **Selective enrichment of human DNA from non-human DNAs for DNA typing of decomposed skeletal remains.** *Forensic Sci Int-Gen Supplement Series* 2009, **2:**520–521.
3. Alonso A, Martín P, Albarrán C, García P, Primorac D, García O, de Fernández Simón L, García-Hirschfeld J, Sancho M, Fernández-Piqueras J: **Specific quantification of human genomes from low copy number DNA samples in forensic and ancient DNA studies.** *Croat Med J* 2003, **44:**273–280.
4. van Oven M, Kayser M: **Updated comprehensive phylogenetic tree of global human mitochondrial DNA variation.** *Hum. Mutat.* 2009, **30:**E386–E394.
5. Briggs AW, Good JM, Green RE, Krause J, Maricic T, Stenzel U, Lalueza-Fox C, Rudan P, Brajković D, Kućan Z, Gušić I, Schmitz R, Doronichev VB, Golovanova LV, Rasilla M, Fortea J, Rosas A, Pääbo S: **Targeted retrieval and analysis of five Neandertal mtDNA genomes.** *Science* 2009, **325:**318–321.
6. Anderung C, Persson P, Bouwman A, Elburg R, Götherström A: **Fishing for ancient DNA.** *Forensic Sci Int-Gen* 2008, **2:**104–107.
7. Brotherton P, Sanchez JJ, Cooper A, Endicott P: **Preferential access to genetic information from endogenous hominin ancient DNA and accurate quantitative SNP-typing via SPEX.** *Nucleic Acids Res.* 2010, **38:**e7.
8. Parsons TJ, Huel R, Davoren J, Katzmarzyk C, Miloš A, Selmanović A, Smajlović L, Coble MD, Rizvić A: **Application of novel 'mini-amplicon' STR multiplexes to high volume casework on degraded skeletal remains.** *Forensic Sci Int-Gen* 2007, **1:**175–179.
9. Bright J-A, Taylor D, Curran JM, Buckleton JS: **Developing allelic and stutter peak height models for a continuous method of DNA interpretation.** *Forensic Sci Int-Gen* 2013, **7:**296–304.
10. Hill CR, Kline MC, Coble MD, Butler JM: **Characterization of 26 miniSTR loci for improved analysis of degraded DNA samples.** *J Forensic Sci* 2008, **53:**73–80.
11. Lee HY, Park MJ, Kim NY, Sim JE, Yang WI, Shin K-J: **Simple and highly effective DNA extraction methods from old skeletal remains using silica columns.** *Forensic Sci Int-Gen* 2010, **4:**275–280.
12. Fregeau CJ, De Moors A: **Competition for DNA binding sites using Promega DNA IQ (TM) paramagnetic beads.** *Forensic Sci Int-Gen* 2012, **6:**511–522.
13. Gill P: **An assessment of the utility of single nucleotide polymorphisms (SNPs) for forensic purposes.** *Int J Legal Med* 2001, **114:**204–210.
14. Budowle B, Wilson MR, DiZinno JA, Stauffer C, Fasano MA, Holland MM, Monson KL: **Mitochondrial DNA regions HVI and HVII population data.** *Forensic Sci Int* 1999, **103:**23–35.
15. Allen M, Engström AS, Meyers S, Handt O, Saldeen T, Von Haeseler A, Pääbo S, Gyllensten U: **Mitochondrial DNA sequencing of shed hairs and saliva on robbery caps: sensitivity and matching probabilities.** *J Forensic Sci* 1998, **43:**453–464.
16. Boles TC, Snow CC, Stover E: **Forensic DNA testing on skeletal remains from mass graves: a pilot project in Guatemala.** *J Forensic Sci* 1995, **40:**349–355.
17. Gill P, Ivanov PL, Kimpton C, Piercy R, Benson N, Tully G, Evett I, Hagelberg E, Sullivan K: **Identification of the remains of the Romanov family by DNA analysis.** *Nat Genet* 1994, **6:**130–135.
18. Ginther C, Issel-Tarver L, King M-C: **Identifying individuals by sequencing mitochondrial DNA from teeth.** *Nat Genet* 1992, **2:**135–138.
19. Butler JM, Levin BC: **Forensic applications of mitochondrial DNA.** *Trends Biotechnol* 1998, **16:**158–162.
20. Just RS, Loreille OM, Molto JE, Merriwether AD, Woodward SR, Matheson C, Creed J, McGrath SE, Sturk-Andreaggi K, Coble MD, Irwin JA, Ruffman A, Parr RL: **Titanic's unknown child: the critical role of the mitochondrial DNA coding region in a re-identification effort.** *Forensic Sci Int-Gen* 2011, **5:**231–235.
21. Brotherton P, Haak W, Templeton J, Brandt G, Soubrier J, Adler CJ, Richards SM, Der Sarkissian C, Ganslmeier R, Friederich S, Dresely V, van Oven M, Kenyon R, Van der Hoek MB, Korlach J, Luong K, Ho SYW, Quintana-Murci L, Behar D, Meller H, Alt KW, Cooper A, the Genographic Consortium: **Neolithic mitochondrial haplogroup H genomes and the genetic origins of Europeans.** *Nat Commun* 2013, **4:**1764.
22. Melton T, Clifford S, Kayser M, Nasidze I, Batzer M, Stoneking M: **Diversity and heterogeneity in mitochondrial DNA of North American populations.** *J Forensic Sci* 2001, **46:**46–52.
23. Grubwieser P, Mühlmann R, Parson W: **New sensitive amplification primers for the STR locus D2S1338 for degraded casework DNA.** *Int J Legal Med* 2003, **117:**185–188.
24. Maricic T, Whitten M, Pääbo S: **Multiplexed DNA sequence capture of mitochondrial genomes using PCR products.** *PLoS One* 2010, **5:**e14004.
25. Knapp M, Stiller M, Meyer M: **Generating barcoded libraries for multiplex high-throughput sequencing.** In *Edited by Shapiro B, Hofreiter M.* Edited by Ancient DNA. Humana Press: New York; 2012:155–170.

26. Fu Q, Mittnik A, Johnson PLF, Bos K, Lari M, Bollongino R, Sun C, Giemsch L, Schmitz R, Burger J, Ronchitelli AM, Martini F, Cremonesi RG, Svoboda J, Bauer P, Caramelli D, Castellano S, Reich D, Pääbo S, Krause J: **A revised timescale for human evolution based on ancient mitochondrial genomes.** *Curr Biol* 2013, **23:**553–559.
27. Hodges E, Xuan Z, Balija V, Kramer M, Molla MN, Smith SW, Middle CM, Rodesch MJ, Albert TJ, Hannon GJ, McCombie WR: **Genome-wide *in situ* exon capture for selective resequencing.** *Nat Genet* 2007, **39:**1522–1527.
28. Gnirke A, Melnikov A, Maguire J, Rogov P, LeProust EM, Brockman W, Fennell T, Giannoukos G, Fisher S, Russ C, Gabriel S, Jaffe DB, Lander ES, Nussbaum C: **Solution hybrid selection with ultra-long oligonucleotides for massively parallel targeted sequencing.** *Nat Biotechnol* 2009, **27:**182–189.
29. Cooper A, Poinar HN: **Ancient DNA: do it right or not at all.** *Science* 2000, **289:**1139–1139.
30. Swango KL, Timken MD, Chong MD, Buoncristiani MR: **A quantitative PCR assay for the assessment of DNA degradation in forensic samples.** *Forensic Sci Int* 2006, **158:**14–26.
31. Adler CJ, Haak W, Donlon D, Cooper A, the Genographic Consortium: **Survival and recovery of DNA from ancient teeth and bones.** *J Archaeol Sci* 2011, **38:**956–964.
32. **FASTX-Toolkit.** http://hannonlab.cshl.edu/fastx_toolkit/index.html.
33. Martin M: **Cutadapt removes adapter sequences from high-throughput sequencing reads.** *EMBnet Journal* 2011, **1:**1–10.
34. Andrews S: **FastQC. A quality control tool for high throughput sequence data.** http://www.bioinformatics.bbsrc.ac.uk/projects/fastqc.
35. Behar DM, van Oven M, Rosset S, Metspalu M, Loogväli E-L, Silva NM, Kivisild T, Torroni A, Villems R: **A 'Copernican' reassessment of the human mitochondrial DNA tree from its root.** *Am J Hum Gen* 2012, **90:**675–684.
36. **TMAP v3.2.1.** https://github.com/nh13/TMAP.
37. Merriman B, Torrent I, Rothberg JM: **Progress in Ion Torrent semiconductor chip based sequencing.** *Electrophoresis* 2012, **33:**3397–3417.
38. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R: **The sequence alignment/map format and SAMtools.** *Bioinformatics* 2009, **25:**2078–2079.
39. **Picard Tools.** http://picard.sourceforge.net.
40. Ginolhac A, Rasmussen M, Gilbert MTP, Willerslev E, Orlando L: **mapDamage: testing for damage patterns in ancient DNA sequences.** *Bioinformatics* 2011, **27:**2153–2155.
41. **Geneious v6.0 created by Biomatters.** http://www.geneious.com/.
42. Haak W, Forster P, Bramanti B, Matsumura S, Brandt G, Tanzer M, Villems R, Renfrew C, Gronenborn D, Alt KW, Burger J: **Ancient DNA from the first European farmers in 7500-year-old Neolithic sites.** *Science* 2005, **310:**1016–1018.
43. Haak W, Balanovsky O, Sanchez JJ, Koshel S, Zaporozhchenko V, Adler CJ, der Sarkissian C, Brandt G, Schwarz C, Nicklisch N, Dresely V, Fritsch B, Balanovska E, Villems R, Meller H, Alt KW, Cooper A, the Genographic Consortium: **Ancient DNA from European early Neolithic farmers reveals their near eastern affinities.** *PLoS Biol.* 2010, **8:**e1000536.
44. Sánchez-Quinto F, Schroeder H, Ramirez O, Ávila-Arcos MC, Pybus M, Olalde I, Velazquez A, Marcos MEP, Encinas JMV, Bertranpetit J, Orlando L, Gilbert MTP, Lalueza-Fox C: **Genomic affinities of two 7,000-year-old Iberian hunter-gatherers.** *Curr Biol* 2012, **22:**1494–1499.
45. Krause J, Briggs AW, Kircher M, Maricic T, Zwyns N, Derevianko A, Pääbo S: **A complete mtDNA genome of an early modern human from Kostenki, Russia.** *Curr Biol* 2010, **20:**231–236.
46. Green RE, Malaspinas A-S, Krause J, Briggs AW, Johnson PL, Uhler C, Meyer M, Good JM, Maricic T, Stenzel U, Prüfer K, Siebauer M, Burbano H, Ronan M, Rothberg JM, Egholm M, Rudan P, Brajković, Kućan Z, Gušić I, Wikström M, Laakkonen L, Kelso J, Slatkin M, Pääbo S: **A complete Neandertal mitochondrial genome sequence determined by high-throughput sequencing.** *Cell* 2008, **134:**416–426.
47. Yeo ZX, Chan M, Yap YS, Ang P, Rozen S, Lee ASG: **Improving indel detection specificity of the Ion Torrent PGM benchtop sequencer.** *PLoS One* 2012, **7:**e45798.
48. Quail MA, Smith M, Coupland P, Otto TD, Harris SR, Connor TR, Bertoni A, Swerdlow HP, Gu Y: **A tale of three next generation sequencing platforms: comparison of Ion Torrent, Pacific Biosciences and Illumina MiSeq sequencers.** *BMC genomics* 2012, **13:**341.
49. Golan D, Medvedev P: **Using state machines to model the Ion Torrent sequencing process and to improve read error rates.** *Bioinformatics* 2013, **29:**i344–i351.
50. Bragg LM, Stone G, Butler MK, Hugenholtz P, Tyson GW: **Shining a light on dark sequencing: characterising errors in Ion Torrent PGM data.** *PLoS Comput Biol* 2013, **9:**e1003031.
51. Koehnemann S, Pfeiffer H: **Application of mtDNA SNP analysis in forensic casework.** *Forensic Sci Int-Gen* 2011, **5:**216–221.
52. Allentoft ME, Collins M, Harker D, Haile J, Oskam CL, Hale ML, Campos PF, Samaniego JA, Gilbert MTP, Willerslev E: **The half-life of DNA in bone: measuring decay kinetics in 158 dated fossils.** *Proc Roy Soc B: Biol Sci* 2012, **279:**4724–4733.
53. Irwin JA, Parson W, Coble MD, Just RS: **mtGenome reference population databases and the future of forensic mtDNA analysis.** *Forensic Sci Int-Gen* 2011, **5:**222–225.
54. Just RS, Leney MD, Barritt SM, Los CW, Smith BC, Holland TD, Parsons TJ: **The use of mitochondrial DNA single nucleotide polymorphisms to assist in the resolution of three challenging forensic cases.** *J Forensic Sci* 2009, **54:**887–891.
55. Coble MD, Just RS, O'Callaghan JE, Letmanyi IH, Peterson CT, Irwin JA, Parsons TJ: **Single nucleotide polymorphisms over the entire mtDNA genome that increase the power of forensic testing in Caucasians.** *Int J Legal Med* 2004, **118:**137–146.
56. Ballantyne KN, van Oven M, Ralf A, Stoneking M, Mitchell RJ, van Oorschot RAH, Kayser M: **MtDNA SNP multiplexes for efficient inference of matrilineal genetic ancestry within Oceania.** *Forensic Sci Int-Gen* 2012, **6:**425–436.
57. van Oven M, Vermeulen M, Kayser M: **Multiplex genotyping system for efficient inference of matrilineal genetic ancestry with continental resolution.** *Invest Genet* 2011, **2:**6–6.
58. Quintans B, Alvarez-Iglesias V, Salas A, Phillips C, Lareu MV, Carracedo A: **Typing of mitochondrial DNA coding region SNPs of forensic and anthropological interest using SNaPshot minisequencing.** *Forensic Sci Int* 2004, **140:**251–257.
59. Wilson MR, DiZinno JA, Polanskey D, Replogle J, Budowle B: **Validation of mitochondrial DNA sequencing for forensic casework analysis.** *Int J Legal Med* 1995, **108:**68–74.

doi:10.1186/2041-2223-4-26
Cite this article as: Templeton *et al.*: DNA capture and next-generation sequencing can recover whole mitochondrial genomes from highly degraded samples for human identification. *Investigative Genetics* 2013 **4**:26.

Manuscript 2: Neolithic mitochondrial haplogroup H genomes and the genetic origins of Europeans.

nature COMMUNICATIONS

ARTICLE

Received 20 Sep 2012 | Accepted 27 Feb 2013 | Published 23 Apr 2013

DOI: 10.1038/ncomms2656

Neolithic mitochondrial haplogroup H genomes and the genetic origins of Europeans

Paul Brotherton[1,2,*], Wolfgang Haak[1,*], Jennifer Templeton[1], Guido Brandt[3], Julien Soubrier[1], Christina Jane Adler[1,†], Stephen M. Richards[1], Clio Der Sarkissian[1,†], Robert Ganslmeier[4], Susanne Friederich[4], Veit Dresely[4], Mannis van Oven[5], Rosalie Kenyon[6], Mark B. Van der Hoek[6], Jonas Korlach[7], Khai Luong[7], Simon Y.W. Ho[8], Lluis Quintana-Murci[9], Doron M. Behar[10], Harald Meller[4], Kurt W. Alt[3], Alan Cooper[1] & The Genographic Consortium[‡]

Haplogroup H dominates present-day Western European mitochondrial DNA variability (>40%), yet was less common (~19%) among Early Neolithic farmers (~5450 BC) and virtually absent in Mesolithic hunter-gatherers. Here we investigate this major component of the maternal population history of modern Europeans and sequence 39 complete haplogroup H mitochondrial genomes from ancient human remains. We then compare this 'real-time' genetic data with cultural changes taking place between the Early Neolithic (~5450 BC) and Bronze Age (~2200 BC) in Central Europe. Our results reveal that the current diversity and distribution of haplogroup H were largely established by the Mid Neolithic (~4000 BC), but with substantial genetic contributions from subsequent pan-European cultures such as the Bell Beakers expanding out of Iberia in the Late Neolithic (~2800 BC). Dated haplogroup H genomes allow us to reconstruct the recent evolutionary history of haplogroup H and reveal a mutation rate 45% higher than current estimates for human mitochondria.

[1] The Australian Centre for Ancient DNA, School of Earth and Environmental Sciences, University of Adelaide, Adelaide, South Australia 5005, Australia. [2] Archaeogenetics Research Group, School of Applied Sciences, University of Huddersfield, Huddersfield HD1 3DH, UK. [3] Institute of Anthropology, Colonel-Kleinmann Weg 2, Johannes Gutenberg University Mainz, D-55128 Mainz, Germany. [4] State Office for Heritage Management and Archaeology Saxony-Anhalt/State Museum for Prehistory Halle, Richard-Wagner-Straße 9, D-06114 Halle/Saale, Germany. [5] Department of Forensic Molecular Biology, Erasmus MC, University Medical Centre Rotterdam, 3000 CA Rotterdam, The Netherlands. [6] SA Pathology, SA Health, Adelaide, South Australia 5000, Australia. [7] Pacific Biosciences, Menlo Park, California 94025, USA. [8] School of Biological Sciences, The University of Sydney, Sydney, New South Wales 2006, Australia. [9] Institut Pasteur, Unit of Evolutionary Genetics, 75015 Paris, France. [10] Rambam Medical Centre, 31096 Haifa, Israel. * These authors contributed equally to this work. † Present address: Institute of Dental Research, Westmead Centre for Oral Health, The University of Sydney, Sydney, New South Wales 2145, Australia (C.J.A.); Centre for Geogenetics, Natural History Museum of Denmark, 1350 Copenhagen, Denmark (C.D.S.). Correspondence and requests for materials should be addressed to P.B. (email: p.m.brotherton@hud.ac.uk) or to W.H. (email: wolfgang.haak@adelaide.edu.au).
‡A full list of authors for the Genographic Consortium and their affiliations appears at the end of the paper.

A key unanswered issue in human prehistory is the extent to which cultural change identifiable in the archaeological record can be ascribed to the movements of people, as opposed to the movements of just their ideas and artefacts. The Central European archaeological record identifies a succession of profound cultural and economic changes between the last hunter-gatherers of the Mesolithic and the first farmers of the Early Neolithic (ENE), through to the socially stratified chiefdoms of the Early Bronze Age[1–3]. The exact nature and genetic context of the transformative changes that took place over these four millennia remain unclear[4,5], although current genetic patterns of mitochondrial DNA (mtDNA) haplogroup (hg) distribution suggest a complex series of events in European prehistory[4–9] and hint at multiple inputs from outside Central Europe[4,10,11].

Phylogeographic studies suggest that mt hg H arrived in Europe from the Near East before the Last Glacial Maximum (22,000 BP), and survived in glacial refugia in Southwest Europe before undergoing a post-glacial re-expansion[4,12]. Haplogroup H now accounts for over 40% of mtDNA variation in anatomically modern humans across much of Western Eurasia, with declining frequencies south and east to ~10–30% in the Near East and Caucasus[10]. However, it remains uncertain when and how H became the dominant European hg. Traditional approaches (including ancient DNA studies) have been unable to resolve either the phylogeny or phylogeographic distribution of H sub-haplogroups (sub-hgs)[6], however, they have generally relied on sequencing only 300–400 bp of the mt D-loop or control region[10,13]. A number of studies based on complete 16.6 kb human mt genomes have revealed a complex evolutionary history for hg H (for example, refs 12,14–18, with phylogenetic analyses recognizing 87 H sub-hgs[19]). These complete mt genomes revealed that 71% of hg H polymorphic diversity is located outside the D-loop, in the coding region[20] and, as a result, this diversity has not yet been exploited at the population genetics level.

To investigate the relationship between the European genetic and archaeological records, we sequenced whole hg H mt genomes from skeletal remains directly assigned to distinct Central European archaeological cultures. Owing to its excellently preserved human skeletal remains, forming a continuous record across a series of archaeological cultures since Palaeolithic times, the Mittelelbe-Saale region of Saxony-Anhalt (Germany) provided a unique opportunity to address this issue. We analysed a time transect spanning the >3,500 years of the Central European Neolithic period (Table 1, Supplementary Table S1), from the first farmers of the ENE linear pottery culture (LBK, 5450–4775 BC), through the subsequent Rössen (4625–4250 BC), Schöningen (4100–3950 BC), Baalberge (3950–3400 BC) and Salzmünde (3400–3025 BC) cultures. These were followed by two of the first pan-European Late Neolithic (LNE) cultural complexes, the Corded Ware (CWC, 2800–2050 BC) and Bell Beaker (BBC, 2500–2050 BC) cultures, before the emergence of the Early Bronze Age with the Unetice culture (2200–1575 BC). We chose to focus on hg H because of its recent dramatic rise in frequency to become the dominant hg in Europe, because of its presence in all Neolithic cultures in the Mittelelbe-Saale region, and the potential it provided to explore detailed genetic structure on a sub-hg level. Overall, our results suggest that the broad foundations of the Central European mtDNA pool, here approximated via hg H, were formed during the Neolithic rather than the post-glacial period.

Table 1 | Summary of genotyping data against the Reconstructed Sapiens Reference Sequence (RSRS).

Culture/age	Individual	Hg*	Hg H sequence variants compared with RSRS
LBK (5450–4775 BC)	HAL36	H23	**C10211T**
	HAL11	H	T16093C, G16129A!
	HAL32	H26	**T11152C**
	HAL39	H1e	**G3010A**, **G5460A**
	DEB9	H88	**A8596G**
	DEB21	H1j	**G3010A**, **T4733C**
	KAR6a	H1bz	G1719A, **G3010A**, C14380T
	KAR11b	H	T152C!
	KAR16a	H46b	**C2772T**, A11893G
Rössen (4625–4475/4250 BC)	OSH2	H89	A6932G, C8068T, T12696C
	OSH3	H1	**G3010A**
	OSH1	H16	**T152C!**, **C10394T**
	OSH7	H5b	**C456T, G5471A, T16304C**, C16519T
Schöningen (4100–3950 BC)	SALZ18a	H10i	C13503T, **T14470a**, T**16093C**
	SALZ21b	H1e7	T1766C, **G3010A**, **G5460A**
Baalberge (3950–3400 BC)	ESP30	H1e1a5	**G3010A**, **G5460A**, (C5960T), **A8512G**, G8865A, **C14902T**
	HQU4	H7d5	**A4793G**, **C15409T**, G16388A
Salzmünde (3400–3100/3025 BC)	SALZ57a	H3	T152C!, **T6776C**
	SALZ77a	H3	**T6776C**
Corded Ware (2800–2200/2050 BC)	ESP15	H6a1a	**T239C**, **G3915A**, **A4727G**, **G9380A**, **T11253C**, **T16362C**, **A16482G**, C16519T
	BZH6	H1_TBD	**G3010A**, A8149G, A9377G, T9467C, A13671G, T14319C, **T16189C!**
Bell Beaker	BZH4	H1e7	**G3010A**, **G5460A**, A15220G, A15401G, A16293G
(2500–2200/2050 BC)	ROT6	H5a3	**C456T**, **G513A**, **T4336C**, **G15884A**, **T16304C**, C16519T
	ALB1	H3b	**A2581G**, **T6776C**
	ROT1	H3ao2	C4577T, **T6776C**, **C16256T**
	ROT2	H5a3	**C456T**, **G513A**, **T4336C**, **G15884A**, **T16304C**, C16519T
	QUEXII1	H4a1	**C3992T**, **A4024G**, **T5004C**, **G9123A**, **C14365T**, **A14582G**, C16519T
	QUEXII2	H4a1	**C3992T**, **A4024G**, **T5004C**, **G9123A**, **C14365T**, **A14582G**, C16519T
	QLB26a	H1	**G3010A**
	QUEXII3	H13a1a2c	**C2259T**, **A4745G**, G9025A, **A13542G**, **C13680T**, **C14872T**, C16519T
	QLB28b	H1	**G3010A**
Unetice (2200–1575 BC)	BZH1	H11a	**T195C!**, **T961g**, **T8448C**, **(G13759A)**, **A16293G**, **T16311C!**, C16519T
	BZH8	H2a1a3	**G951A**, **G1438A**, **G4769A**, **C6173T**, **T13095C**, A16240t, **C16354T**, C16519T
	BZH14	H82a	T195C!, **A16220G**
	EUL41a	H4a1a1a5	**A73G!**, **C3992T**, **A4024G**, **T5004C**, **G8269A**, **G9123A**, **A10044G**, C13545T, **C14365T**, **A14582G**, C16519T
	EUL57B	H3	T152C!, **T6776C**
	QUEVIII4	H7h	**A4793G**, **G16213A**
Nuragic Bronze Age (1624 BC)	—	H1aw1	**G3010A**, **A8701G!**, C15912T
Iron Age (500 BC)	—	H90	C5435T, T8911C, T10237C, T15109C

Abbreviation: SNP, single-nucletide polymorphism.
Sub-haplogroup defining diagnostic SNPs are shown in bold and 'private'/as-yet-unknown sequence variants in regular print.
*Haplogroup H designations based on the http://www.phylotree.org mtDNA tree Build 14 (5 April 2012)[19,20].

Results

Sequence and network analyses. From a collection of over 400 European prehistoric human archaeological remains we selected 37 Mittelelbe-Saale individuals, as well as two samples from Italy (Supplementary Table S1), previously assigned to hg H by simplex and multiplex PCR[7]. Work was independently replicated for two samples per individual (Supplementary Methods). We

designed and optimised a hybridisation-based DNA-capture system to sequence complete mt genomes on the Affymetrix MitoChip v2.0 (ref. 21) (Supplementary Methods, Supplementary Fig. S1,S2, and Supplementary Tables S2-S3) via immortalised libraries prepared from the highly damaged and degraded endogenous DNA recovered from archaeological remains[22,23]. Six of the 39 target-enriched libraries were also analysed via a single-molecule, real-time (SMRT[24]) Pacific Biosciences *RS* sequencing platform (Supplementary Table S4, Supplementary Dataset). In addition, 35/391 (9%) of all SNPs identified via the MitoChip were independently confirmed by direct PCR and Sanger sequencing (Supplementary Methods, Supplementary Tables S5 and S6). Mt genomes from all 39 individuals were unambiguously assignable to individual sub-hgs of hg H[20], confirming that a single human was typed in each case (Table 1). The mt hypervariable region I sequences matched those previously determined for each individual. The ancient hg H mt genomes were highly diverse, with 34 distinct haplotypes attributed to 20 major sub-hgs (gene diversity $H = 0.997 \pm 0.0071$; nucleotide diversity 0.000421 ± 0.000225), including three novel lineages (provisionally named H88–H90).

Phylogenetic network analysis of these ancient mt genomes reveals evidence of dynamic changes in the composition of H sub-hgs over the ~3,500-year time transect (Fig. 1). Importantly, sequences from older samples (and cultures) tend to represent basal lineages, only one to three mutations away from the ancestral root of hg H, while younger samples (after ~4000 BC) largely comprise more derived haplotypes appearing on longer

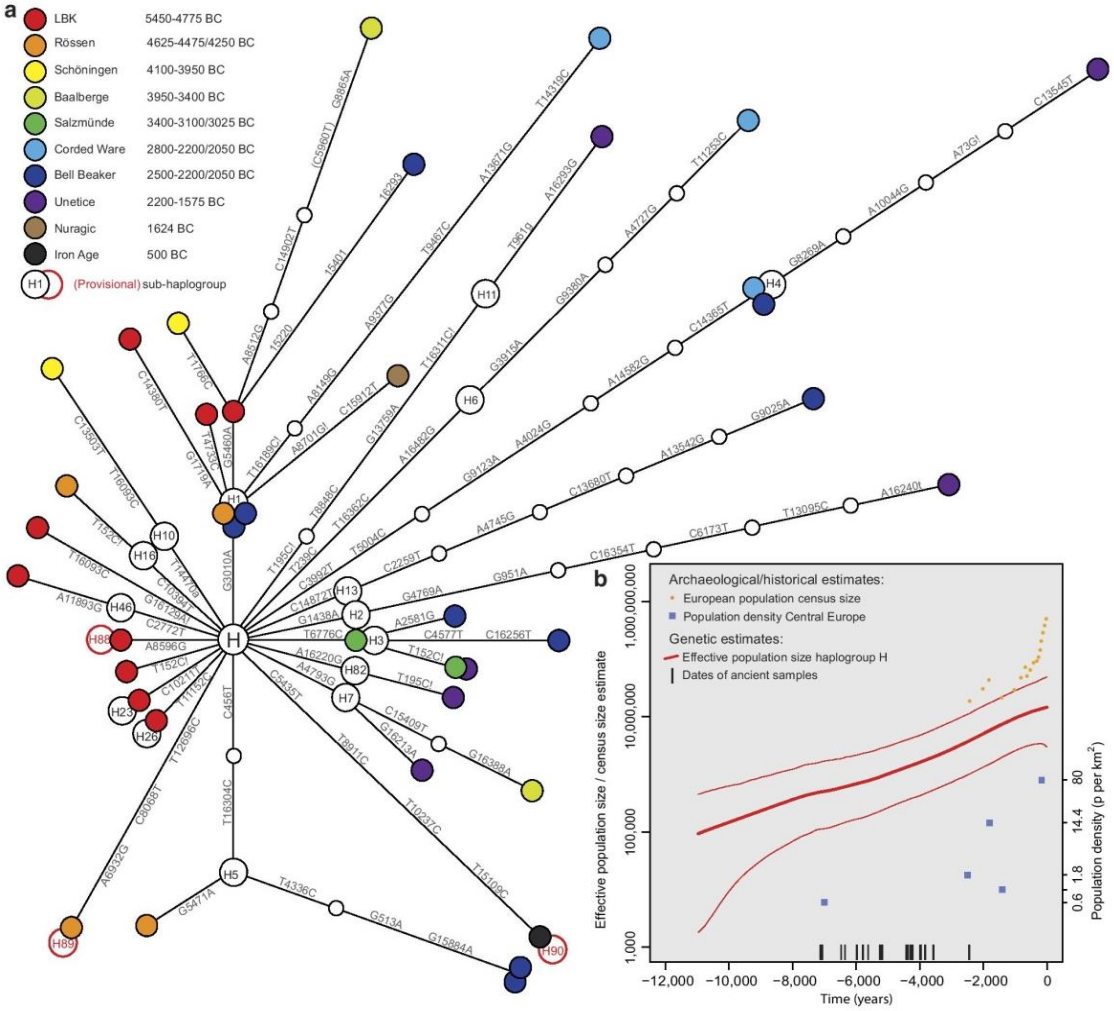


Figure 1 | Mitochondrial haplogroup H sequence evolution. (**a**) Phylogenetic network of 39 prehistoric mitochondrial genomes sorted into two temporal groupings: Early Neolithic (left) and Mid-to-Late Neolithic (right). Node colours represent archaeological cultures. (**b**) A Bayesian skyride plot of 200 representative present-day and 39 ancient hg H mt genomes (the thick red line denotes the posterior median, thinner flanking lines denote the 95% credibility interval; note the logarithmic scale of the *y* axis). Prehistoric samples (18 radiocarbon and 21 mean archaeological dates) served as internal calibration points (black bars). For comparison, census size estimates for the European population are shown as orange dots. Population density estimates from the archaeological record for key periods in Central Europe are plotted as blue squares in chronological order: LBK, Iron Age, Roman period, Merovingian and Pre-industrial modern times (*y* axis on the right)[28].

branches. This temporal relationship provides further support for the authenticity of the ancient mt genomes.

Network analysis (Fig. 1) reveals pronounced differences in the composition of sub-hgs between the ENE cultures (LBK, Rössen, Schöningen), and those of the Mid Neolithic (MNE)/LNE to Early Bronze Age (Baalberge, Salzmünde, Corded Ware, Bell Beaker, Unetice). ENE (and in particular LBK) mt genomes are either rare today (H16, H23 and H26), extinct or have not yet been observed in present-day populations (H46b, H88 and H89). In sharp contrast, most of the later H sub-hgs are more common in present-day European populations (for example, hg H3, H4, H6, H7, H11 and H13)[12,14–16]. Of the 39 haplotypes detected, only three (within the common, basal, sub-hg H1) were shared between ENE and MNE/LNE cultures. As the observed gene diversity is high, we might expect the number of shared haplotypes within and between cultures to be low. However, as the MNE/LNE haplotypes are on different sub-hg branches from the ENE haplotypes, these patterns combined show minimal local genetic continuity over this time period (Table 1).

Genetic distances. To further examine these apparent temporal shifts in sub-hg distribution, we tested whether hg H individuals represent different meta-populations by pooling them into different cultural and/or temporal groups of ENE versus LNE (Table 2, Supplementary Table S7). When pooled in four groups (ENE, MNE, LNE and Bronze Age), pairwise population comparisons via F_{ST} values based on sequence data showed that genetic distances increased with time over the duration of the Neolithic, reaching a significant value ($F_{ST} = 0.08722$; $P = 0.00386 +/- 0.0006$) between the ENE and the early Bronze Age (Table 2). This suggests a transformation of hg H diversity during the Neolithic period. This effect was less apparent (non-significant F_{ST} values) when samples from various sites were pooled in larger temporal groups (Table 2). However, non-parametric multivariate analysis of variance (NP-MANOVA, $P = 0.0072$) also confirmed a significant difference between pooled groups of ENE and LNE individuals when comparisons were based on the presence or absence of sub-hgs (Table 2).

Table 2 | Population pairwise and linearised Slatkin's F_{ST}s and NP-MANOVA tests.

(a) NP-MANOVA four time periods ($P = 0.0696$)

	Early Neolithic	Middle Neolithic	Late Neolithic	Bronze Age
Early Neolithic (13)	0	0.1262	0.024	0.0574
Middle Neolithic (6)	**0.7572**	0	0.8575	0.7782
Late Neolithic (9)	**0.144**	**1**	0	0.742
Bronze Age (6)	**0.3444**	**1**	**1**	0

F_{ST} four time periods F_{ST}

	Early Neolithic	Middle Neolithic	Late Neolithic	Bronze Age
Early Neolithic (13)	0	*0*	*0.03179*	*0.09555*
Middle Neolithic (6)	0.01135	0	*0*	*0.02299*
Late Neolithic (9)	0.02247	−0.01165	0	*0.01148*
Bronze Age (6)	**0.08722**	0.03081	−0.02250	0

(b) NP-MANOVA LBK, BBC and pooled intermediate Neolithic (MNE) as used in PCA ($P = 0.2355$)

	LBK	MNE	BBC
LBK (9)	0	0.2084	0.0916
MNE (10)	**0.6252**	0	0.8025
BBC (7)	**0.2748**	**1**	0

F_{ST} LBK, BBC, and pooled intermediate Neolithic (MNE) as used in PCA

	LBK	MNE	BBC
LBK (9)	0	*0*	*0.03369*
MNE (10)	−0.02587	0	*0*
BBC (7)	0.03260	−0.00704	0

(c) NP-MANOVA two time periods ($P = 0.0072$)

	Early Neolithic	Late Neolithic
Early Neolithic (13)	0	0.0109
Late Neolithic (16)	**0.0109**	0

F_{ST} two time periods

	Early Neolithic	Late Neolithic
Early Neolithic (13)	0	*0.01459*
Late Neolithic (16)	0.01438	0

(d) NP-MANOVA Cultures grouped with geographic regions as in Fig. 2 *($P < 0.0001$)*

	Iberia	Near East	Mainland Europe
Iberia	0	0	0.0001
Near East	**0**	0	0.0004
Mainland Europe	**0.0003**	**0.0012**	0

Abbreviations: BBC, Bell Beaker culture; MNE, Mid Neolithic; NP-MANOVA, non-parametric multivariate analysis of variance; PCA, principal component analysis.
Neolithic samples pooled in different time periods: (a) four time periods; (b) three time periods; (c) two time periods, and (d) from cultures grouped with geographic regions. For NP-MANOVA, Bonferroni corrected values are given in bold print and areas shaded grey indicate significant values ($P < 0.05$). Slatkin F_{ST}s are italicised (upper diagonal) and significant pairwise distances are given in bold print (lower diagonal).

Genetic affinities. To examine potential geographic origins for Neolithic cultures (Supplementary Table S1) and to assess their contribution to present-day Central European mtDNA diversity, we used principal component analysis (PCA) to investigate genetic affinities between three ancient culturally/temporally pooled groups (LBK, MNE and BBC) and 37 present-day Western Eurasian populations (Supplementary Table S8). PCA of the frequencies of the 15 most common H sub-hgs showed that the present-day populations form three significantly supported geographic clusters (a grouping which was also supported using NP-MANOVA, $P < 0.0001$; Table 2: (i) Iberia in the west; (ii) the Caucasus, the Near East and Anatolia; and (iii) Central and Eastern Europe from the Urals to France (Fig. 2a). This particular number of clusters was also the best supported in a model-based test on sub-hg H frequencies followed by Ward clustering (Fig. 2c,d). We also used Procrustes analysis to quantify the relationship between hg H substructure and the geographic locations of both the present-day Western Eurasian and the Mittelelbe-Saale ancient populations. For this analysis, we superimposed the PCA coordinates on the geographic map of the present-day and ancient sampling locations. We found a striking resemblance between the genetic and geographic maps, with a highly significant Procrustes similarity score ($t_0 = 0.733$) obtained for the comparison ($P < 10^{-6}$; 100,000 permutations). The analysis supported a clustering of the transformed genetic data from present-day populations into the three major groups described above (Fig. 2b). In contrast, Procrustes analysis clearly showed that the genetic data for LBK and BBC samples were not related to their geographic location. Although all three ancient groups were sampled from the same Central European location only the MNE group genetically resembles present-day populations from this region.

The combined set of analyses (PCA, Procrustes and Ward clustering) revealed that Mittelelbe-Saale's earliest farmers (LBK; $n = 9$) cluster with present-day Caucasus, Near Eastern and Anatolian populations, as previously noted[7]. In contrast, individuals from the successor series of regional post-LBK (and MNE) Rössen, Schöningen, Baalberge and Salzmünde cultures (ca. 4625–3025 BC, MNE; $n = 10$) cluster with present-day Central European populations (Fig. 2). Mitochondrial genomes from BBC individuals in Mittelelbe-Saale (BBC; $n = 7$) display close genetic affinities to present-day Iberian populations (Fig. 2). The component loadings of the PCA biplot indicate that this is largely based on high frequencies of sub-hgs H1 and H3, which are thought to have spread from a glacial Iberian refugium[13] and which have also been reported from ancient Neolithic sites from France and Spain[8,25]. Other LNE samples add further to the genetic complexity. Individuals from the CWC (2800–2200 BC), which has archaeological associations towards North-Eastern Europe, produced two distinct mt genomes (H1_TBD and H6a1a), which have not been found in their contemporaneous Bell Beaker neighbours, nor in preceding Central European cultures. Similarly, data from the subsequent Early Bronze Age

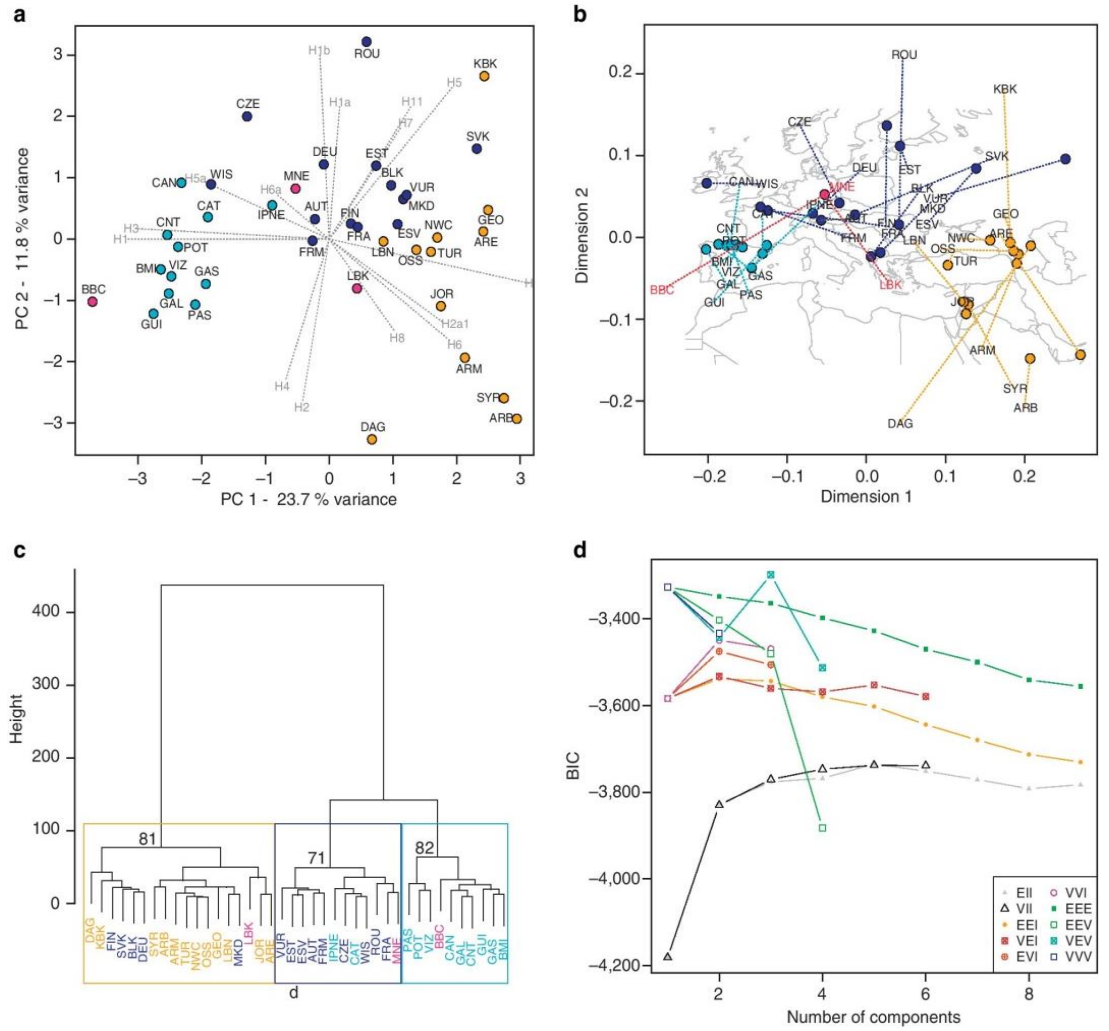


Figure 2 | Population affinities of select Neolithic cultures. (**a**) PCA biplot based on the frequencies of 15 hg H sub-haplogroups (component loadings) from 37 present-day Western Eurasian and three ancient populations (light blue: Western Europe; dark blue: Central and Eastern Europe; orange; Near East, Caucasus and Anatolia; and pink: ancient samples). Populations are abbreviated as follows: GAL, Galicia; CNT, Cantabria; CAT, Catalonia; GAS, Galicia/Asturia; CAN, Cantabria2; POT, Potes; PAS, Pasiegos; VIZ, Vizcaya; GUI, Guipuzcoa; BMI, Basques; IPNE, Iberian Peninsula Northeast; TUR, Turkey; ARM, Armenia; GEO, Georgia; NWC, Northwest Caucasus; DAG, Dagestan; OSS, Ossetia; SYR, Syria; LBN, Lebanon; JOR, Jordan; ARB, Arabian Peninsula; ARE, Arabian Peninsula2; KBK, Karachay-Balkaria; MKD, Macedonia; VUR, Volga-Ural region; FIN, Finland; EST, Estonia; ESV, Eastern Slavs; SVK, Slovakia; FRA, France; BLK, Balkans; DEU, Germany; AUT, Austria, ROU, Romania; FRM, France Normandy; WIS, Western Isles; CZE, Czech Republic; LBK, Linear pottery culture; BBC, Bell Beaker culture; MNE, Middle Neolithic. (**b**) Procrustes analyses of geographic coordinates and PCA scores of the same data set (similarity score $t_0 = 0.733$, $P < 10^{-6}$, 100,000 permutations). (**c**) Ward clustering dendrogram of the three ancient groups and present-day populations (colour code as above and p values in % of approximately unbiased bootstrapping for the following three main clusters). (**d**) Results of the model-based test to identify the number of clusters by the model with the highest support (highest Bayes Information criterion (BIC); VEV = multivariate mixture model (ellipsoidal, equal shape)).

Unetice culture revealed haplotypes with genetic affinities to both the East (sub-hg H2a, H7 and H11) and the West (sub-hg H3 and H4), based on frequency distributions of these sub-hgs in present-day populations[13]. We also included two individuals from outside Central Europe (Sardinia and South Tyrol) and from different time periods (Nuragic Bronze Age and Iron Age, respectively) to further investigate genetic diversity within hg H and to test the power of resolution of complete mt genomes. Both individuals had mt genomes that are not found in samples from the Mittelelbe-Saale region. The Iron Age sample from South Tyrol produced another new sub-hg (provisional H90) and the Bronze Age individual from Sardinia a new H1 haplotype (H1aw1).

Reconstructing the demographic history of mtDNA hg H. It has previously proved difficult to use present-day data alone to determine when hg H became the predominant hg in Europe, as archaeogenetic and palaeodemographic reconstructions have very large uncertainties[4,26]. However, as our 39-dated ancient mt genome sequences provide precise temporal calibration points,

we performed a Bayesian skyride analysis with 200 random present-day mt genome sequences to reconstruct the lineage history of hg H through time (with the caveat of assuming a continuous and panmictic population). The resulting skyride plot (Fig. 1b) is the first real-time estimation of the European hg H population size (and consequently its contribution to Europe's effective population size and demographic history) with a broad temporal coverage over ~3,500 years of the Neolithic period in Central Europe (5500–2000 BC). Hg H shows a consistent and strong exponential growth over the entire course of the Neolithic. The estimated population size tracks the European census size[27] and population density estimates from archaeological sites[28] in the Late Holocene, but also provides detailed estimates for prehistoric times for which data points remain very scarce (Fig. 1b).

Another major advantage of the temporal calibration points provided by ancient hg H mt genomes is that the data allow a relatively precise estimate of the evolutionary substitution rate for human mtDNA. The temporal dependency of evolutionary rates predicts that rate estimates measured over short timespans will be considerably higher than those using deep fossil calibrations, such as the human/chimpanzee split at ~6 million years[29]. The rate calibrated by the Neolithic and Bronze Age sequences is 2.4×10^{-8} substitutions per site per year ($1.7–3.2 \times 10^{-8}$; 95% high posterior density) for the entire mt genome, which is 1.45× (44.5%) higher than current estimates based on the traditional human/chimp split (for example, 1.66×10^{-8} for the entire mt genome[30] and 1.26×10^{-8} for the coding region[31]). Consequently, the calibrated 'Neolithic' rate infers a considerably younger coalescence date for hg H (10.9–19.1 kya) than those previously reported (19.2–21.4 kya for HVSI[10], 15.7–22.5 kya for the mt coding region[31] or 14.7–22.6 kya when corrected for purifying selection[30]).

Discussion

Despite recent successes in sequencing portions of nuclear genomes from Meso- and Neolithic samples[11,32,33], mtDNA remains the most widely studied and best described marker in population genetics. Although its interpretation is limited to the matrilineal genetic history[4,13,19], this can be an important socio-cultural and demographic signal additional to that gained from autosomal loci[34,35]. Our results clearly demonstrate that high-resolution full mt genome-typing, combined with the ability to analyse large numbers of individuals from multiple cultural layers, can provide highly resolved temporal views that are not yet practical with nuclear DNA studies.

The phylogenetic network analysis of our chronological hg H mt genome data set (Fig. 1a) provides the first detailed real-time view of mutations in human mtDNA. It has enabled the direct observation of the mutation rate over thousands of years and revealed a distinct temporal distribution pattern of hg H diversity. Although a temporal pattern could be expected in an expanding population with stable/increasing hg H frequencies (Fig. 1b, Supplementary Fig. S3), ENE and MNE/LNE/Bronze Age samples clearly show a mutually exclusive sub-hg distribution with the exception of sub-hg H1, which is the most common and basal sub-hg within H[14–16]. Under an assumption of genetic continuity, we would expect MNE/LNE and Bronze Age individuals to be on the same sub-hg branches as ENE individuals. Instead, ENE mt genomes are generally either rare today[19] or have not yet been observed in present-day populations, possibly owing to subsequent extinction of these lineages. In contrast, most MNE/LNE and Bronze Age sub-hgs are still common today. This suggests that individuals from the ENE made a marginal contribution to LNE and present-day hg H diversity. Although the relatively small sample numbers from each time period limit detailed analyses of the causes of the distribution shifts, we interpret this phylogenetic pattern as a genetic discontinuity between Early and subsequent Neolithic cultures in Europe, potentially mirroring genetic structure in Neolithic European populations. Genetic drift could also have played a role in generating discrepant hg distributions over time and space. However, if drift was the sole cause we would expect a random distribution across all sub-hgs rather than a clear distinction between ENE and MNE/LNE/Bronze Age mt genomes.

Our genetic distance data also indicate minimal local genetic continuity between the ENE and the MNE/LNE in Central Europe (Fig. 1; Table 1), again suggesting that ENE lineages were largely superseded during the MNE/LNE (~4100–2200 BC) in a previously unrecognised major genetic transition. This pronounced genetic changeover between ENE and MNE/LNE cultures is comparable to other known major genetic transition, thus far revealed by ancient DNA and coalescent simulations (between indigenous European hunter-gatherers and incoming early farmers from the Near East during the initial Meso-Neolithic transition from ~7500 BC in Central Europe)[6,7]. When compared with hg H diversity in present-day Central Europe[14,15,18,36], the network in Fig. 1 suggests that much of the present-day diversity can be attributed to the incorporation of new lineages in the MNE/LNE and emerging Bronze Age (from 2200 BC). The LNE in particular is known to have been a period of profound cultural and economic change[37], with newly emerging pan-European cultures such as the Bell Beaker phenomenon in Western Europe and the Corded Ware culture in north-eastern Europe. It therefore seems likely that these pan-European cultures were associated with the introduction of lineages from outside Central Europe. Fortunately, the ranges of both these groups overlapped in the Mittelelbe-Saale sample area (Supplementary Methods), allowing this possibility to be further investigated.

Our data on genetic affinities (PCA, Procrustes and Ward clustering) revealed that Mittelelbe-Saale's earliest farmers (LBK; $n = 9$) cluster with present-day Caucasus, Near Eastern, and Anatolian populations. These findings are consistent with a highly detailed archaeological record tracing the temporal and spatial spread of agriculture into Central Europe; beginning initially in Anatolia and the Near East, where farming originated ~12,000 years ago[7].

Our observation that individuals from the successor series of regional post-LBK and MNE cultures (Rössen, Schöningen, Baalberge and Salzmünde) cluster with present-day Central European populations could be explained by a loss of lineages from the ENE LBK period during a short phase of population decline in the centuries after 5000 BC (as proposed in some archaeological models)[38]. However, our results suggest that mtDNA H sub-hg diversity established during the MNE is still present in Central European populations today. This is consistent with independent archaeological evidence of a phase of more localised cultural development during the MNE period, potentially involving influences from contemporaneous MNE cultures outside Mittelelbe-Saale, which (perhaps in concert with LBK population decline) could have resulted in a replacement of most ENE H sub-hgs. Together, the genetic and archaeological evidence highlight the complexities of both the formative and consolidation phases in Central Europe.

From around 2800 BC, the LNE Bell Beaker culture emerged from the Iberian Peninsula to form one of the first pan-European archaeological complexes. This cultural phenomenon is recognised by a distinctive package of rich grave goods including the eponymous bell-shaped ceramic beakers. The genetic affinities between Central Europe's Bell Beakers and present-day Iberian

populations (Fig. 2) is striking and throws fresh light on long-disputed archaeological models[3]. We suggest these data indicate a considerable genetic influx from the West during the LNE. These far-Western genetic affinities of Mittelelbe-Saale's Bell Beaker folk may also have intriguing linguistic implications, as the archaeologically-identified eastward movement of the Bell Beaker culture has recently been linked to the initial spread of the Celtic language family across Western Europe[39]. This hypothesis suggests that early members of the Celtic language family (for example, Tartessian)[40] initially developed from Indo-European precursors in Iberia and subsequently spread throughout the Atlantic Zone; before a period of rapid mobility, reflected by the Beaker phenomenon, carried Celtic languages across much of Western Europe. This idea not only challenges traditional views of a linguistic spread of Celtic westwards from Central Europe during the Iron Age, but also implies that Indo-European languages arrived in Western Europe substantially earlier, presumably with the arrival of farming from the Near East[41].

Other LNE population movements appear to have added further genetic complexity, as exemplified by the CWC (2800–2200 BC), which preceded the Bell Beaker culture in Mittelelbe-Saale and has archaeological associations with North-Eastern Europe. A genetic affinity to eastern populations is consistent with two distinct CWC mt genomes (H1_TBD and H6a1a) not identified in either their contemporaneous Bell Beaker neighbours or in preceding Central European cultures. The subsequent Early Bronze Age Unetice culture, associated with emerging metallurgy and increasingly stratified societies[37,42], marks a consolidation of social and cultural systems in Mittelelbe-Saale that were established during the LNE by the two pan-European Bell Beaker and CWCs. The Unetice culture appears contemporaneously with the last Neolithic horizon (~2200 BC) in areas where elements of both the Bell Beaker and CWCs are present, sometimes overlapping at the same sites. It is therefore not surprising that individuals ascribed to the newly emerging Unetice culture carry mt genomes with both Western (sub-hgs H3 and H4) and Eastern (sub-hgs H2a, H7 and H11) associations.

The demographic reconstruction, which is based on direct calibration points, has major implications for understanding post-glacial human history in Europe. Our new estimate is incompatible with traditional views that the majority of present-day hg H lineages were carried into Central, Northern and Eastern Europe via a post-glacial human population expansion before the Holocene (12 kya)[13]. Our data complement a recent study, based on present-day mt genomes, which describes a pronounced population increase at ~7000 BC (interpreted as a Neolithic expansion into Europe), but followed by a slow population growth until the present day[26]. By including ancient DNA data from across the critical time points in question, our skyride plot corrects for missing temporal data and suggests substantial growth of hg H from the beginning of the Neolithic and continuing throughout the entire Neolithic period. This emphasizes the role of farming practices and cultural developments in the demographic expansions inferred in subsequent time periods, which have not yet been explored genetically.

Although an expansion of hg H could in principle be compatible with a post-glacial resettling of Northern and Central Europe from southwestern refugia[12,16] (as indicated by our population skyride and PCA plots), we instead propose that the rise of hg H to become the predominant mtDNA branch in Europe was mediated by subsequent demographic events during the Neolithic, as shown by a general increase in hg H frequency and strong population growth during this period (Fig. 1b). Support for this position comes from data suggesting that hg H was virtually absent among Central and Northern European hunter-gatherers[6,43] and formed only 19% in LBK individuals, most likely introduced from Southeast Europe and/or the Near East[7]. In our updated data set from Mittelelbe-Saale, hg H appears to have been established by the LBK period and increased in frequency after 4000 BC (Supplementary Fig. S3). Interestingly, MNE/LNE cultures with cultural associations to the North and Northeast, such as the Bernburg and CWCs, show reduced hg H frequencies and hg H only moved northwards into southern Scandinavia during the Neolithisation of Northern Europe around the Middle Neolithic, as exemplified by individuals from the Funnel Beaker Culture[11,43]. However, hg H appears to have been generally more frequent in prehistoric Western Europe: at 20% from a Middle Neolithic (3030–2890 calBC) site in France[9]; at ~25% from Iberian (Epi-)Cardial Neolithic samples[8,25]; at 36% from a Neolithic site in Catalonia[44]; and at 44% from Neolithic sites from the Basque Country and Navarre[45]. Importantly, a recent study on Iberian hunter-gatherers revealed the presence of hg H there in Mesolithic times[45]. In Mittelelbe-Saale, the Bell Beaker samples signpost a significant increase in hg H frequency (the 95% confidence intervals do not overlap with earlier LBK and Schöningen Neolithic cultures; Supplementary Fig. S3). In conclusion, the Western European Neolithic and the widespread pan-European Bell Beaker phenomenon appear to be important factors in driving the spread of H sub-hgs throughout large parts of Western Europe. In particular, high proportions of sub-hgs H1 and H3 seem to have made substantial contributions to the hg H diversity that exists in Western and Central Europe today[16]. Having reached significant levels, and assuming a generally higher rate of population growth in southern and western Europe in post-Neolithic times[27], these Neolithic processes appear to have been the major factor in hg H becoming the predominant European mtDNA hg.

Overall, our results suggest that the broad foundations of the Central European mtDNA pool, here approximated via hg H, were formed during the Neolithic rather than the post-glacial period. ENE hg H mt lineages brought in from the Near East by Central Europe's first farmers do not appear to have contributed significantly to present-day Central Europe's hg H diversity, instead being largely superseded during the MNE and LNE (with the process starting around 4000 BC), after which there appears to have been substantial genetic continuity to the present-day in Central Europe. These developments have been revealed by comparative full mt genome sequencing and would have remained obscure using standard HVS I data.

In conclusion, demographic changes across the MNE, followed by the widespread Bell Beaker cultural phenomenon, are likely to have been the key factors in the expansion of hg H across Western Europe and the eventual rise of hg H to become the predominant mtDNA hg. However, LNE Corded Ware and Early Bronze Age data suggest a complex series of additional genetic contributions, which require further investigation.

Methods

Ancient DNA analyses. DNA was extracted from two independent samples for each individual (Supplementary Methods). HVS I was amplified using a minimum of four short overlapping primer pairs, following established protocols and authentication criteria as described previously[7,46]. Multiplex SNP typing of 22 hg informative SNPs (GenoCoRe22) was carried out using a SNaPshot-based protocol as described previously[7].

Ancient DNA Library preparation. Ancient DNA extract polishing, phosphorylation, adaptor ligation and polymerase 'fill-in' reactions were used sequentially to create fully double-stranded adaptor-tagged aDNA libraries (Fig. 3). Following every step, DNA was purified using MinElute spin columns (Qiagen) as per the

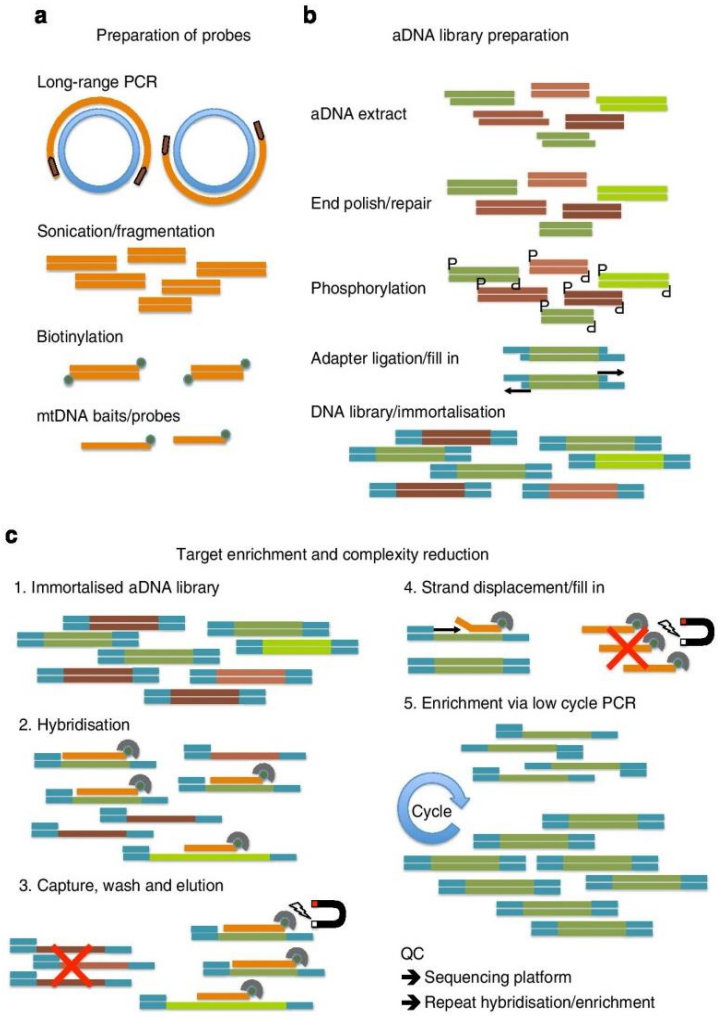


Figure 3 | Schematic representation of experimental steps. (**a**) Probe DNA was prepared by amplifying a complete mitochondrial genome in two overlapping fragments by long-range PCR, followed by DNA fragmentation and biotinylation to form mtDNA 'baits' for targeted hybridisation. (**b**) Ancient DNA was enzymatically blunt-ended and phosphorylated, ligated to custom library adaptors, followed by polymerase 'fill-in' to create 'immortalised' double-stranded DNA libraries. (**c**) Hybridisation-based DNA-capture using biotinylated probe bound to Streptavidin magnetic beads; following stringency washes, captured library constructs enriched in mtDNA sequences are eluted from the beads/probe via a novel polymerase strand-displacement reaction followed by PCR library reamplification. These steps can be carried out iteratively to maximise mtDNA content in enriched libraries (see Supplementary Methods for full details).

manufacturer's instructions. PCR amplification reactions were then performed to create 'primary' DNA libraries, ready for DNA-capture hybridisation steps, and amplification products were sized and quantified (Supplementary Methods).

Hybridisation-based enrichment of human mtDNA. The basic conceptual design for the hybridisation of tracer DNA (aDNA library) to biotinylated driver DNA sequences (human mt probe) was previously described[47] and the overall scheme is outlined in Fig. 3. The two library-specific PCR primers were included as part of the hybridisation mix as blocking oligonucleotides to minimise unwanted hybridisation between the adaptor-tagged flanking regions of otherwise unrelated single-stranded library DNA molecules[48]. A key innovation of this methodology was the use of a DNA polymerase with strand-displacing activity after post-hybridisation stringency washes. This allowed primer extension from the bound library (blocking) primers to disrupt the double-stranded region of stable hybridisation between human mt probe DNA sequences and single-stranded library DNA molecules that had inserts with complementary sequences. These mtDNA-enriched library DNA molecules captured in the hybridisation step could thereby be cleanly separated from biotinylated probe molecules, which remained bound to magnetic Streptavidin beads. PCR reamplification reactions from the mt-enriched library DNA molecules comprised the 'first enrichment' DNA libraries. In general, we used three cycles of hybridisation/enrichment/reamplification to produce DNA libraries highly enriched for short endogenous mtDNA sequence fragments ready for genotyping (Supplementary Methods).

Affymetrix Mitochip v2.0 array typing and Pacific Biosciences SMRT sequencing. MtDNA-enriched libraries underwent biotin labelling using terminal deoxynucleotidyl transferase (TdT) as per the Affymetrix GeneChip Whole-Transcript Sense Target Labelling Assay Manual (P/N 701880, rev. 4). Biotin-labelled DNA libraries were hybridised to Affymetrix GeneChip Human mt Resequencing 2.0 Arrays for 17 h at 49 °C. Arrays were washed, stained and scanned as per the GeneChip CustomSeq Resequencing Array Protocol (P/N 701231, rev. 5). Affymetrix GeneChip Command Console software (v3.2) was used to generate CEL files, which were then analysed using GeneChip Sequence Analysis Software (GSEQ v4.1, Affymetrix) and validated using the software Geneious[49] (Supplementary Fig. S1,S2, Supplementary Tables S2,S3). Six of the mt-enriched

DNA libraries were also converted to SMRTbell template libraries for sequencing on a Pacific Biosciences *RS* platform (Supplementary Methods).

Network analyses. A median joining network of all ancient hg H mt genomes (Fig. 1a) was constructed manually using the most up-to-date version of the mt phylogenetic tree (PhyloTree.org, mtDNA tree Build 14) as a scaffold on which to place the observed hg H lineages[19,20]. This version included a revised version of the hg H sub-tree comprising 1203 sequences in total. As per convention, insertions at np 309.1C(C), 315.1C, 523-524d (aka 522-523d), 16182C, 16183C, 16193.1C(C) and mutation 16519 were not considered for phylogenetic reconstruction[20].

Procrustes-based PCA and Ward Clustering. PCA was used to describe and visualise the maternal genetic relationships among the Neolithic cultures investigated, as well as to 37 present-day European and Near Eastern populations (Fig. 2a). PCA was performed on the frequency of H sub-hgs taken from the literature (Supplementary Table S8). To minimise statistical noise caused by rare sub-hgs and to allow for data compatibility across published studies, we considered only the following 15 most common H sub-hgs in Europe and the Near East: H*, H1, H1a, H1b, H2, H2a1, H3, H4, H5, H5a, H6, H6a, H7, H8 and H11. PCAs were performed and visualised in R version 2.11.1 (ref. 50) using a customised script based on the function prcomp.

Ancient hg H individuals were pooled into three different groups based on the numbers of samples available: two for 'pan-European' archaeological phenomena/cultures alongside hypothesised geographic origins (LBK, $n=9$ and BBC, $n=7$); and a temporally transitional group pooling regional (mostly MNE) cultures (MNE, $n=10$). Small sample sets such as the Corded Ware ($n=2$) and later Bronze Age Unetice ($n=5$) were excluded. To test whether the clustering pattern observed in the PCA was significantly supported, we performed a number of statistical tests including Ward clustering, Procrustes analysis and NP-MANOVA (as described below). First, we performed model-based cluster tests to identify the number of clusters via the model with the best support (highest Bayes Information criterion) followed by Ward hierarchical clustering of sub-hg H frequencies using the packages mclust, pvclust (for bootstrap values) and hclust in R, respectively. Procrustes analysis was also performed in R using the package vegan based on PCA scores and geographic coordinates (Supplementary Table S8) and the function protest to calculate the similarity score (100,000 permutations).

Summary statistics. Population pairwise F_{ST}, Slatkin's linearised F_{ST} and haplotype diversity were calculated in Arlequin version 3.5 (ref. 51). We used jMODELTEST 0.1.1[52] in order to find the best fitting evolutionary model and, if required, to estimate a discrete γ shape parameter for our 39 non-partitioned mt genomes. Based on the resulting scores for each model (AIC and Bayes Information criterion), we subsequently used the Tamura and Nei model and a γ value of 0.049 for our calculations of population distances in Arlequin. The ancient hg H individuals were pooled into different temporal/cultural groups in order to calculate genetic diversity indices and to test for genetic differentiation (Table 2, Supplementary Table S7).

Multivariate analysis of variance. We performed a NP-MANOVA to test whether the temporal grouping of ancient individuals according to archaeological time periods are statistically supported. The NP-MANOVA was performed on a Raup-Crick distance matrix, which was produced from the presence/absence of the 15 hg H sub-hgs used in the PCA. Calculations were performed in PAST version 2.09 with 10,000 permutations per test and *post hoc* Bonferroni correction to account for multiple comparisons and small sample sizes[53]. We also tested whether the clustering pattern between the ancient and present-day populations observed in the PCA was significantly supported.

Bayesian skyride analyses and mutation rate calculation. The data set comprised 37 newly sequenced, non-related, ancient mt genomes, five sets of randomly chosen, distinct, present-day hg H mt genomes from Phylotree (http://www.phylotree.org, mtDNA tree Build 12 (20th July 2011)) and 420 newly available hg H sequences[17]. The Sequences were manually aligned to the revised Cambridge Reference Sequence (rCRS: AC_000021)[54] using the program SeaView[55]. The alignment was partitioned into four subsets, representing the D-loop, the protein-coding regions (1st + 2nd codon positions and 3rd codon position) and a concatenation of tRNA and RNA genes. Insertions at nps 309.1C(C), 315.1C, 523-524d (aka 522-523d), 16182C, 16183C, 16193.1C(C) were not considered for phylogenetic reconstruction and position 16519 was removed from the D-loop subset[20]. The best substitution models were selected using ModelGenerator 0.85 (ref. 56). by comparison of Bayesian Information Criterion scores: HKY + G for D-loop, TN + G for protein-coding regions and HKY for RNA genes. Considering the short evolutionary timescale being studied (intra-hg), models including a proportion of invariant sites were excluded. A Bayesian skyride analysis[57] was performed using the phylogenetic software BEAST 1.6.1 (ref. 58). and calibrated using radiocarbon dates from 18 of the ancient individuals and mean archaeological dates for the remaining individuals. This allowed us to achieve a broad temporal coverage for ~3500 years of the Neolithic period in Central Europe (5500–2000 BC) and to generate the most precise demographic reconstruction of hg H. Results were replicated using independent sets of 100 (1 ×), 200 (3 ×), and 300 (1 ×) mt genomes. A strict molecular clock was used, allowing for a distinct rate in each subset of the alignment. Additional analysis using an uncorrelated log normal relaxed clock to account for potential rate variations could not reject the strict clock assumption. Convergence was checked by sampling from two independent Markov chains. Each MCMC analysis was run for 100,000,000 steps and samples from the two chains were combined, after discarding the first 10% of samples as burn-in. All parameters showed sufficient sampling, indicated by effective sample sizes above 200. Tracer 1.5 was used to produce the skyride plot (Fig. 1b)[59].

We carried out a 'date randomisation test', to test whether the signal from the radiocarbon dates associated with the ancient sequences was sufficient to calibrate the hg H phylogeny[60]. This test randomises all dates associated with the sequences (including present-day ones) and replicates of the phylogenetic analysis as described above. If the structure and spread of the ancient sequences in the tree were sufficient to calibrate the analysis, the inferred mean rate of the randomised analysis should be significantly different from the rate calculated using the correct association date/sequence. In other words, the 95% HPD of the randomised analysis should not overlap with the mean rate estimated without randomisation. The comparison of estimated rates from the main analysis and from 10 replicates with randomised dates presented in Supplementary Fig. S4 confirms the presence of sufficient signal to calibrate the tree provided by dates from the 37 ancient samples.

References

1. Whittle, A. W. R. & Cummings, V. *Going over: The Mesolithic-Neolithic Transition in North-West Europe* 632 (Oxford University Press, Oxford, 2007).
2. Sherratt, A. Plough and pastoralism: aspects of the secondary products revolution. *Patterns of the Past: Studies in honour of David Clarke.* In: Hodder, I., Isaac, G. & Hammond, N. (eds). 261–305 (Cambridge University Press, Cambridge, 1981).
3. Bogucki, P. I. & Crabtree, P. J. *Ancient Europe 8000 B.C.-A.D. 1000: Encyclopedia of the Barbarian World,* 1221p (Charles Scribner's Sons, 2004).
4. Soares, P. *et al.* The archaeogenetics of Europe. *Curr. Biol.* **20,** R174–R183 (2010).
5. Pinhasi, R., Thomas, M. G., Hofreiter, M., Currat, M. & Burger, J. The genetic history of Europeans. *Trends genet.* **28,** 496–505 (2012).
6. Bramanti, B. *et al.* Genetic discontinuity between local hunter-gatherers and central Europe's first farmers. *Science* **326,** 137–140 (2009).
7. Haak, W. *et al.* Ancient DNA from European Early Neolithic Farmers Reveals Their Near Eastern Affinities. *PLoS Biol.* **8,** e1000536 (2010).
8. Gamba, C. *et al.* Ancient DNA from an Early Neolithic Iberian population supports a pioneer colonization by first farmers. *Mol. Ecol.* **21,** 45–56 (2012).
9. Lacan, M. *et al.* Ancient DNA reveals male diffusion through the Neolithic Mediterranean route. *Proc. Natl Acad. Sci. USA* **108,** 9788–9791 (2011).
10. Richards, M. *et al.* Tracing European founder lineages in the Near Eastern mtDNA pool. *Am. J. Hum. Genet.* **67,** 1251–1276 (2000).
11. Skoglund, P. *et al.* Origins and genetic legacy of Neolithic farmers and hunter-gatherers in Europe. *Science* **336,** 466–469 (2012).
12. Pereira, L. *et al.* High-resolution mtDNA evidence for the late-glacial resettlement of Europe from an Iberian refugium. *Genome Res.* **15,** 19–24 (2005).
13. Torroni, A., Achilli, A., Macaulay, V., Richards, M. & Bandelt, H. J. Harvesting the fruit of the human mtDNA tree. *Trends Genet.* **22,** 339–345 (2006).
14. Roostalu, U. *et al.* Origin and expansion of haplogroup H, the dominant human mitochondrial DNA lineage in West Eurasia: The near eastern and Caucasian perspective. *Mol. Biol. Evol.* **24,** 436–448 (2007).
15. Loogväli, E. L. *et al.* Disuniting uniformity: a pied cladistic canvas of mtDNA haplogroup H in Eurasia. *Mol. Biol. Evol.* **21,** 2012–2021 (2004).
16. Achilli, A. *et al.* The molecular dissection of mtDNA haplogroup H confirs that the Franco-Cantabrian glacial refuge was a major source for the European gene pool. *Am. J. Hum. Genet.* **75,** 910–918 (2004).
17. Behar, D. M. *et al.* The Basque paradigm: genetic evidence of a maternal continuity in the Franco-Cantabrian region since Pre-Neolithic times. *Am. J. Hum. Genet.* **90,** 486–493 (2012).
18. Alvarez-Iglesias, V. *et al.* New population and phylogenetic features of the internal variation within mitochondrial DNA macro-haplogroup R0. *PLoS ONE* **4,** e5112 (2009).
19. Behar, D. M. *et al.* A "Copernican" reassessment of the human mitochondrial DNA tree from its root. *Am. J. Hum. Genet.* **90,** 675–684 (2012).
20. van Oven, M. & Kayser, M. Updated comprehensive phylogenetic tree of global human mitochondrial DNA variation. *Hum. Mutat.* **30,** E386–E394 (2009).
21. Hartmann, A. *et al.* Validation of microarray-based resequencing of 93 worldwide mitochondrial genomes. *Hum. Mutat.* **30,** 115–122 (2009).
22. Maricic, T., Whitten, M. & Paabo, S. Multiplexed DNA sequence capture of mitochondrial genomes using PCR products. *PLoS ONE* **5,** e14004 (2010).

23. Margulies, M. *et al.* Genome sequencing in microfabricated high-density picolitre reactors. *Nature* **437,** 376–380 (2005).
24. Korlach, J. *et al.* Real-time DNA sequencing from single polymerase molecules. *Methods Enzymol.* **472,** 431–455 (2010).
25. Lacan, M. *et al.* Ancient DNA suggests the leading role played by men in the Neolithic dissemination. *Proc. Natl Acad. Sci. USA* **108,** 18255–18259 (2011).
26. Fu, Q., Rudan, P., Pääbo, S. & Krause, J. Complete mitochondrial genomes reveal Neolithic Expansion into Europe. *PLoS ONE* **7,** e32473 (2012).
27. Livi-Bacci, M. *A Consice History of World Population* 279 (Blackwell Publishing, Malden, Oxford, Carlton, 2007).
28. Zimmermann, A., Hilpert, J. & Wendt, K. P. Estimations of population density for selected periods between the Neolithic and AD 1800. *Hum. Biol.* **81,** 357–380 (2009).
29. Ho, S. Y., Shapiro, B., Phillips, M. J., Cooper, A. & Drummond, A. J. Evidence for time dependency of molecular rate estimates. *Syst. Biol.* **56,** 515–522 (2007).
30. Soares, P. *et al.* Correcting for purifying selection: an improved human mitochondrial molecular clock. *Am. J. Hum. Genet.* **84,** 740–759 (2009).
31. Mishmar, D. *et al.* Natural selection shaped regional mtDNA variation in humans. *Proc. Natl Acad. Sci. USA* **100,** 171–176 (2003).
32. Keller, A. *et al.* New insights into the Tyrolean Iceman's origin and phenotype as inferred by whole-genome sequencing. *Nat. Commun.* **3,** 698 (2012).
33. Sanchez-Quinto, F. *et al.* Genomic affinities of two 7,000-year-old iberian hunter-gatherers. *Curr. Biol.* **22,** 1494–1499 (2012).
34. Behar, D. M. *et al.* The Basque paradigm: genetic evidence of a maternal continuity in the Franco-Cantabrian region since pre-Neolithic times. *Am. J. Hum. Genet.* **90,** 486–493 (2012).
35. Gunnarsdottir, E. D. *et al.* Larger mitochondrial DNA than Y-chromosome differences between matrilocal and patrilocal groups from Sumatra. *Nat. Commun.* **2,** 228 (2011).
36. Brandstätter, A. *et al.* Dissection of mitochondrial superhaplogroup H using coding region SNPs. *Electrophoresis* **27,** 2541–2550 (2006).
37. Heyd, V. Families, prestige goods, warriors & complex societies: Beaker groups of the 3rd millennium cal BC along the upper & middle Danube. *Proc. Prehist. Soc.* **73,** 327–379 (2007).
38. Shennan, S. & Edinborough, K. Prehistoric population history: from the late glacial to the late neolithic in central and northern Europe. *J. Archaeol. Sci.* **34,** 1339–1345 (2007).
39. Cunliffe, B. & Koch, J. T. (eds) *Celtic from the West: Alternative Perspectives from Archaeology, Genetics, Language and Literature,* 384 (Oxbow Books, Oxford, 2010).
40. Koch, J. T. *Tartessian. Celtic in the South-west at the Dawn of History* (Aberystwyth, 2009).
41. Renfrew, C. *Archaeology and Language: The Puzzle of Indo-European Origins.* XIV, 346 S. Ill., Kt(Cape, London, 1988).
42. Nowak, M. Transformations in East-Central Europe from 6000 to 3000 BC: local vs. foreign patterns. *Documenta Praehistorica* **XXXIII,** 143–158 (2006).
43. Malmström, H. *et al.* Ancient DNA reveals lack of continuity between neolithic hunter-gatherers and contemporary Scandinavians. *Curr. Biol.* **19,** 1758–1762 (2009).
44. Sampietro, M. L. *et al.* Palaeogenetic evidence supports a dual model of Neolithic spreading into Europe. *Proc. Biol. Sci./Royal Soc.* **274,** 2161–2167 (2007).
45. Hervella, M. *et al.* Ancient DNA from hunter-gatherer and farmer groups from Northern Spain supports a random dispersion model for the Neolithic expansion into Europe. *PLoS ONE* **7,** e34417 (2012).
46. Haak, W. *et al.* Ancient DNA from the first European farmers in 7500-year-old Neolithic sites. *Science* **310,** 1016–1018 (2005).
47. Patel, M. & Sive, H. PCR-based subtractive cDNA cloning. *Curr. Protocols Mol. Biol.*. Chapter 25, Unit **25B,** 2 (2001).
48. Tao, S. C., Gao, H. F., Cao, F., Ma, X. M. & Cheng, J. Blocking oligo--a novel approach for improving chip-based DNA hybridization efficiency. *Mol. Cell. Probes* **17,** 197–202 (2003).
49. Drummond, A. J. *et al.* Geneious v5.4, Available from http://www.geneious.com/ (2011).
50. R Development Core Team. *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, Vienna, Austria, 2010).
51. Excoffier, L. & Lischer, H. E. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Mol. Ecol. Resources* **10,** 564–567 (2010).
52. Posada, D. jModelTest: phylogenetic model averaging. *Mol. Biol. Evol.* **25,** 1253–1256 (2008).
53. Hammer, O., Harper, D. A. T. & Ryan, P. D. PAST: paleontological statistics software package for education and data analysis. *Palaeontologia Electronica* **4** (2001).
54. Andrews, R. M. *et al.* Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. *Nat. Genet* **23,** 147 (1999).
55. Gouy, M., Guindon, S. & Gascuel, O. SeaView version 4: a multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Mol. Biol. Evol.* **27,** 221–224 (2010).
56. Keane, T. M., Creevey, C. J., Pentony, M. M., Naughton, T. J. & McLnerney, J. O. Assessment of methods for amino acid matrix selection and their use on empirical data shows that ad hoc assumptions for choice of matrix are not justified. *BMC Evol. Biol.* **6,** 29 (2006).
57. Minin, V. N., Bloomquist, E. W. & Suchard, M. A. Smooth skyride through a rough skyline: Bayesian coalescent-based inference of population dynamics. *Mol. Biol. Evol.* **25,** 1459–1471 (2008).
58. Drummond, A. J. & Rambaut, A. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol. Biol.* **7,** 214 (2007).
59. Rambaut, A. & Drummond, A. Tracer v1.4, Available from http://beast.bio.ed.ac.uk/Tracer (2007).
60. Ho, S. Y. *et al.* Bayesian estimation of substitution rates from ancient DNA sequences with low information content. *Syst. Biol.* **60,** 366–375 (2011).

Acknowledgements

We are indebted to Matt Kaplan and Ryan Spriggs at Arizona Research Laboratories, Division of Biotechnology, University of Arizona Genetics Core Facility, http://uagc.arl.arizona.edu/, Tyson Clark, Michael Brown, Kristi Spittle and Matthew Boitano (Pacific Biosciences) for sequencing work, Jeremy Timmis for help with DNA sonication protocols, and Robin Skeates, and Hubert Steiner for additional samples and contextual information. We thank the Australian Research Council (grant LP0882622), the Deutsche Forschungsgemeinschaft (Al 287/7-1 and Me 3245/1-1) and National Geographic's Genographic Project for funding. M.v.O. was supported in part by the Netherlands Forensic Institute (NFI) and a grant from the Netherlands Genomics Initiative (NGI)/Netherlands Organization for Scientific Research (NWO) within the framework of the Forensic Genomics Consortium Netherlands (FGCN).

Author contributions

P.B., W.H. and A.C. conceived and designed the project. P.B. designed and developed the DNA extraction, DNA library construction and hybridisation-based DNA-capture protocols (with assistance from J.T.). P.B., J.T. and W.H. generated and analysed the data. S.M.R., C.D., R.K. and M.B.v.d.H. contributed experimental steps and C.J.A., J.S., S.Y.W.H., J.K. and K.L. contributed analytical steps. G.B., R.G., S.F., V.D., M.v.O., L.Q., D.M.B., H.M. and K.W.A. provided ancient samples, contextual information, radiocarbon dating and access to critical population data. P.B., W.H. and A.C. wrote the manuscript with input from C.J.A., J.S., S.Y.W.H., S.M.R., J.K. and members of the Genographic Consortium. All authors discussed the paper and gave comments.

Additional information

Accession codes: The complete consensus mt genome sequences have been deposited to NCBI GenBank under accession numbers KC553980 to KC554018.

Supplementary Information accompanies this paper at http://www.nature.com/naturecommunications

Competing financial interests: The authors claim no competing financial interests associated with this paper.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

How to cite this article: Brotherton, P. *et al.* Neolithic mitochondrial haplogroup H genomes and the genetic origins of Europeans. *Nat. Commun.* 4:1764 doi: 10.1038/ncomms2656 (2013).

Syama Adhikarla[11], Arun Kumar Ganesh Prasad[11], Ramasamy Pitchappan[11], Arun Varatharajan Santhakumari[11], Elena Balanovska[12], Oleg Balanovsky[12], Jaume Bertranpetit[13], David Comas[13], Begoña Martínez-Cruz[13], Marta Melé[13], Andrew C. Clarke[14], Elizabeth A. Matisoo-Smith[14], Matthew C. Dulik[15], Jill B. Gaieski[15],

Amanda C. Owings[15], Theodore G. Schurr[15], Miguel G. Vilar[15], Angela Hobbs[16], Himla Soodyall[16], Asif Javed[17], Laxmi Parida[17], Daniel E. Platt[17], Ajay K. Royyuru[17], Li Jin[18], Shilin Li[18], Matthew E. Kaplan[19], Nirav C. Merchant[19], R. John Mitchell[20], Colin Renfrew[21], Daniela R. Lacerda[22], Fabrício R. Santos[22], David F. Soria Hernanz[23], R. Spencer Wells[23], Pandikumar Swamikrishnan[24], Chris Tyler-Smith[25], Pedro Paulo Vieira[26] & Janet S. Ziegle[27]

[11] The Genographic Laboratory, School of Biological Sciences, Madurai Kamaraj University, Madurai 625 021, Tamil Nadu, India. [12] Research Centre for Medical Genetics, Russian Academy of Medical Sciences, 115478 Moscow, Russia. [13] Institut de Biologia Evolutiva (CSIC-UPF), Departament de Ciences de la Salut i de la Vida, Universitat Pompeu Fabra, 08003 Barcelona, Spain. [14] Department of Anatomy, University of Otago, Dunedin 9054, New Zealand. [15] Department of Anthropology, University of Pennsylvania, Philadelphia, Pennsylvania, 19104-6398, USA. [16] National Health Laboratory Service, Sandringham 2131, Johannesburg, South Africa. [17] IBM, Yorktown Heights, New York 10598, USA. [18] School of Life Sciences and Institutes of Biomedical Sciences, Fudan University, Shanghai 200433, People's Republic of China. [19] Arizona Research Laboratories, University of Arizona, Tucson, Arizona 85721, USA. [20] Department of Genetics, School of Molecular Sciences, La Trobe University, Melbourne, Victoria 3086, Australia. [21] McDonald Institute for Archaeological Research, University of Cambridge, Cambridge, CB2 3ER, UK. [22] Departamento de Biologia Geral, Universidade Federal de Minas Gerais, Belo Horizonte, Minas Gerais, CEP 31270-901, Brazil. [23] National Geographic Society, Washington, District of Columbia 20036-4688, USA. [24] IBM, Somers, New York 10589, USA. [25] Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge, CB10 1SA, UK. [26] Universidade Federal do Rio de Janeiro, Rio de Janeiro, CEP 21941-901, Brazil. [27] Applied Biosystems, Foster City, California 94494, USA

Manuscript 3: Mitochondrial genome sequencing in Mesolithic North East Europe unearths a new sub-clade within the broadly distributed human haplogroup C1.

OPEN ACCESS Freely available online

PLOS ONE

Mitochondrial Genome Sequencing in Mesolithic North East Europe Unearths a New Sub-Clade within the Broadly Distributed Human Haplogroup C1

Clio Der Sarkissian[1*¤a], Paul Brotherton[1¤b], Oleg Balanovsky[2,3], Jennifer E. L. Templeton[1¤c], Bastien Llamas[1], Julien Soubrier[1], Vyacheslav Moiseyev[4], Valery Khartanovich[4], Alan Cooper[1], Wolfgang Haak[1*], The Genographic Consortium[¶]

1 Australian Centre for Ancient DNA, School of Earth and Environmental Sciences, University of Adelaide, Adelaide, South Australia, Australia, **2** Vavilov Institute for General Genetics, Russian Academy of Sciences, Moscow, Russia, **3** Research Centre for Medical Genetics, Russian Academy of Medical Sciences, Moscow, Russia, **4** Peter the Great Museum of Anthropology and Ethnography (Kunstkamera) RAS, St Petersburg, Russia

Abstract

The human mitochondrial haplogroup C1 has a broad global distribution but is extremely rare in Europe today. Recent ancient DNA evidence has demonstrated its presence in European Mesolithic individuals. Three individuals from the 7,500 year old Mesolithic site of Yuzhnyy Oleni Ostrov, Western Russia, could be assigned to haplogroup C1 based on mitochondrial hypervariable region I sequences. However, hypervariable region I data alone could not provide enough resolution to establish the phylogenetic relationship of these Mesolithic haplotypes with haplogroup C1 mitochondrial DNA sequences found today in populations of Europe, Asia and the Americas. In order to obtain high-resolution data and shed light on the origin of this European Mesolithic C1 haplotype, we target-enriched and sequenced the complete mitochondrial genome of one Yuzhnyy Oleni Ostrov C1 individual. The updated phylogeny of C1 haplogroups indicated that the Yuzhnyy Oleni Ostrov haplotype represents a new distinct clade, provisionally coined "C1f". We show that all three C1 carriers of Yuzhnyy Oleni Ostrov belong to this clade. No haplotype closely related to the C1f sequence could be found in the large current database of ancient and present-day mitochondrial genomes. Hence, we have discovered past human mitochondrial diversity that has not been observed in modern-day populations so far. The lack of positive matches in modern populations may be explained by under-sampling of rare modern C1 carriers or by demographic processes, population extinction or replacement, that may have impacted on populations of Northeast Europe since prehistoric times.

Citation: Der Sarkissian C, Brotherton P, Balanovsky O, Templeton JEL, Llamas B, et al. (2014) Mitochondrial Genome Sequencing in Mesolithic North East Europe Unearths a New Sub-Clade within the Broadly Distributed Human Haplogroup C1. PLoS ONE 9(2): e87612. doi:10.1371/journal.pone.0087612

Editor: Alessandro Achilli, University of Perugia, Italy

Received November 4, 2013; **Accepted** December 23, 2013; **Published** February 4, 2014

Funding: This research was supported by The Genographic Project, which is supported by funding from the National Geographic Society, IBM, and the Waitt Family Foundation. PB was funded by the Australian Research Council Linkage Project LP0882622 awarded to AC, and the Discovery Projects DP1095782 and DP130102158 awarded to WH and BL. OB was funded by the RAS Programmes "Molecular and cell biology", "Gene pool dynamics" and RFBR grants 13-04-01711, Cover Letter 13-06-00670. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors declare having received funding from a commercial partner (IBM). This does not alter the authors' adherence to all the PLOS ONE policies on sharing data and materials.

* E-mail: clio.dersarkissian@gmail.com (CDS); wolfgang.haak@adelaide.edu.au (WH)

¤a Current address: Centre for Geogenetics, Natural History Museum of Denmark, University of Copenhagen, Copenhagen, Denmark
¤b Current address: Archaeogenetics Research Group, School of Applied Sciences, University of Huddersfield, Huddersfield, United Kingdom
¤c Current address: School of Biological Sciences, Flinders University, Adelaide, South Australia, Australia

¶Membership of The Genographic Consortium is provided in the Acknowledgments.

Introduction

Human mitochondrial haplogroup (hg) C is part of the non-African macro-haplogroup M. Most of the diversity of hg C is found today in indigenous populations of Asia and the Americas [1–2]. In northern Asia, hg C represents, together with hg D, more than half of the present-day mitochondrial (mtDNA) diversity [3]. Haplogroup Z, the sister-clade of hg C, has a broad distribution ranging from northern Scandinavia (in Saami) to central Asia, Siberia, northern China and Korea.

Phylogenetic analyses of complete mtDNA genomes revealed four major sub-clades of hg C, termed C1, C4, C5 and C7 (e.g., [3–7]). Of these, haplogroup C1 has one of the broadest distributions of all human mtDNA hgs in the world, ranging from Iceland to East Asia and the Americas. The C1 basal haplotype is defined by the hypervariable region I and II (HVR-I and HVR-II) motif: A16129G, T16187C, C16189T, G16230A, T16278C, T16298C, C16311T, T16325C, C16327T (HVR-I; numbering according to the Reconstructed Sapiens Reference Sequence RSRS; [8]) and C146T, C152T, C195T, A247G, A249d, 290–291d and T489C (HVR-II).

The phylogeny of hg C1 is structured into five distinct monophyletic sub-clades, C1a, C1b, C1c, C1d and C1e, which exhibit a clear geographical distribution pattern ([4], [7], [9–10]; Figure 1). Three of the C1 sub-clades (C1b, C1c and C1d) are restricted to Native American populations, although spread widely across the American continent [11–12]. It was proposed that these three Native American C1 sub-clades were among the ancestral

founder lineages, along with hg A2, B2 and D1, which reached the Americas during the initial human colonisation of the continent [4–5], [7], [9]. The source population of this migration was assumed to be in eastern Asia, where most of the diversity of hg C is observed today, and where C1a, a sister clade of the American C1 clades, is found at low frequencies in diverse indigenous populations [9]. The peopling of the Americas was made possible by the Beringian ice-free land bridge that connected north-east Siberia and Alaska before (~30,000 years Before Present, BP) and after (~13,000 yBP) the last Ice Age [9], [13]. The place of origin of ancestral hg C1 was approximated in the Amur River region just south of Beringia (eastern Asia) on the basis of the current frequency distribution of hg C1 in Asia [9]. The last hg C1 clade to have been described, C1e, was only found recently in a few individuals in Iceland and was shown to be distinct from any of the previously defined Asian and American clades on the basis of seven coding region and three control region mutations [10].

In Europe, the dense and extensive sampling of the HVR-I diversity has revealed extremely low frequencies of hg C1, with very few haplotypes found in Germans [14], Canarians [15], Icelanders [16–17] and Bashkirs [18] (Figure 2). These sequences lack HVR-I Single Nucleotide Polymorphisms (SNPs) diagnostic of the sub-clades C1a (T16356C) and C1d (A16051G). However, a more detailed assignment of the European haplotypes into sub-haplogroups is limited by the low resolution provided by HVR-I and the lack of information from the coding region thus far. These limitations therefore impede the reconstruction of their precise phylogenetic placement, origin and relation to the Asian, American and Icelandic sister-clades.

Three hypotheses for the origins of the C1 lineages in Europe can be put forward [10], [16–17]. Hypothesis 1 proposes a recent genetic input from Asia into Europe during historical times. Historically, Central and East Europe experienced repeated influences from invading groups from the neighbouring Asian steppes, which could have introduced C1 into Europe. Well-documented examples include the Huns from Mongolia in the 4th–5th centuries Anno Domini (A.D.) and the Mongols in the 13th century A.D. [19]. However, the common Asian C1a clade is characterised by the HVR-I transition T16356C, which has not been found in any European C1 haplotype. In the case of a recent Asian ancestry, a reversal of the mutation at nucleotide position

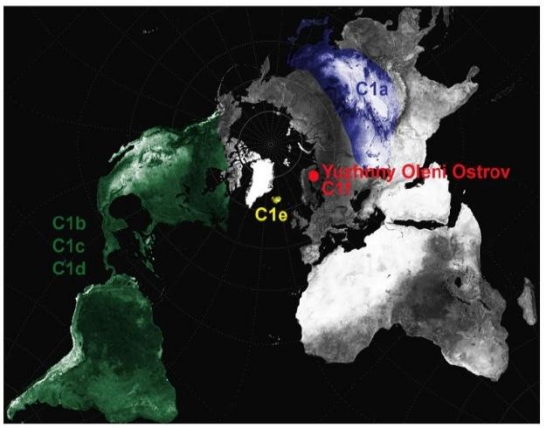


Figure 1. Approximate geographical distribution of the C1 sub-clades in modern and Mesolithic Yuzhnyy Oleni Ostrov populations.
doi:10.1371/journal.pone.0087612.g001

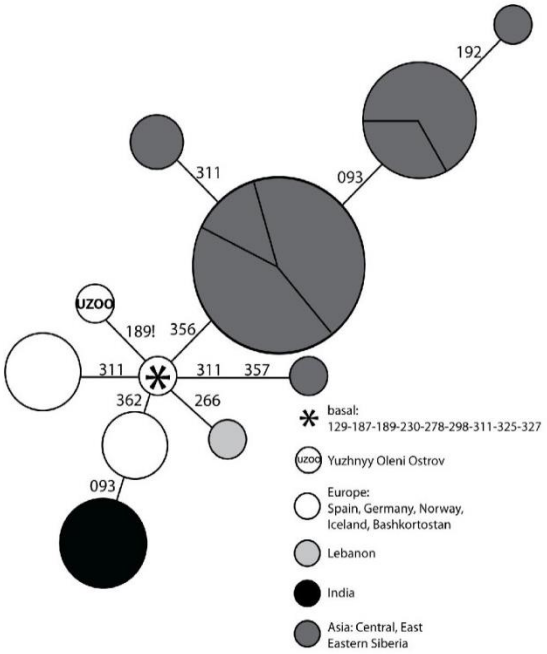


Figure 2. Network representation of C1 HVR-I sequences in Mesolithic Yuzhnyy Oleni Ostrov and modern Eurasian populations. Each haplotype is represented by a circle, the area of which is proportional to the number of individuals that were found to carry this haplotype in the literature. The haplotypes are colour-coded according to their geographical location: India (black), Asia (dark grey), Lebanon (light grey), and Europe (white). Each section of the circles represents individuals sampled from a same population. Mutations are all substitutions and are reported according to the Reconstructed Sapiens Reference Sequence minus 16000. The star represents the hypervariable region-I haplotype that characterizes the root of the C1 clade. The haplotype labeled 'UZOO' is the hypervariable region-I haplotype sequenced from individuals of the archaeological site of Yuzhnyy Oleni Ostrov. All the other haplotypes were found in modern populations.
doi:10.1371/journal.pone.0087612.g002

(np) 16356 in all the European sequences would be required. Hypothesis 2 assumes an American origin, where hg C1 would have reached Europe through admixture between Native Americans and Europeans. This gene flow may have occurred during and after the colonization of the New World by Europeans in the 15th century A.D., i.e. in post-Columbian times [16–17]. Alternatively, one explanation for the presence of the sub-clade C1e in Iceland was a pre-Columbian admixture between Native Americans and Icelandic Vikings, which are widely acknowledged to have built temporary pioneer settlements in the north-western coast of the Americas in the 10th century A.D. [10]. In accordance with the hypothesis of the American origin, few European C1 HVR-I sequences could belong to either the C1b or C1c American clades, as diagnostic SNPs for these two clades are located outside HVR-I [16–17]. Hypothesis 3, proposes that hg C1 has been present in Europe since prehistoric times in the light of the recent finding of hg C1 HVR-I haplotypes in three individuals of the 7,500-year-old Mesolithic site of Yuzhnyy Oleni Ostrov (individuals UZOO-7, UZOO-8, and UZOO-74), North West Russia (Figure 1; [20]). The classification of the corresponding mtDNA haplotype within hg C1 was previously determined by HVR-I sequencing (hg C1 defining mutation T16325C) as well as by typing informative SNPs in the coding region (hg C defining

mutation A13263G [20]), but the lack of resolution of the HVR-I sequence prevented establishing clear phylogenetic relationships with currently known hg C1 clades.

In this study, we sequenced the complete mtDNA genome of one of the three Mesolithic hg C1 carriers from the Yuzhnyy Oleni Ostrov archaeological site (individual UZOO-74) in order to shed further light on the population history of the Yuzhnyy Oleni Ostrov hunter-gatherers and to contribute to the characterisation of the mtDNA diversity, evolutionary history and phylogeography within hg C1.

To this day, complete mtDNA genome sequences from ancient specimens have successfully been determined for a Palaeo-Eskimo from Greenland [21], the 5,000-year-old Tyrolean Iceman [22], a 700-year-old individual from New Zealand [23], a Palaeolithic individual from Tianyan, China [24], as well as several Palaeolithic, Mesolithic and Neolithic individuals from Europe [25–28]. Most of these mtDNA genomes were obtained on high-throughput 'next-generation' sequencing platforms (e.g., [29]). In accordance with these recent studies, we first created a genomic library, which was subsequently enriched for mtDNA in two iterative rounds of hybridisation to in-house designed biotinylated DNA probes, following the protocol by [28]. The enriched DNA library was sequenced on an Ion Torrent PGM platform. We analysed the resulting mtDNA genome from the Yuzhnyy Oleni Ostrov specimen in the light of an updated phylogeny of all currently available hg C1 lineages. The resulting mtDNA genome sequence allowed us to identify a novel C1 sub-clade, coined "C1f", which fills a gap in the knowledge of the hg C1 distribution in West Eurasia.

Results

Ancient Mitochondrial Genome Sequencing

Our ancient DNA (aDNA) enrichment, followed by sequencing on an Ion Torrent PGM platform, allowed the unambiguous determination of 99.8% (16537 out of 16569 base pairs, bp) of the UZOO-74 mtDNA genome with 20,579 unique reads assembled to the RSRS at an average coverage of 68X (Table 1) and average read length of 55±14.5 bp. Indels, a well-defined homopolymer sequencing error, were observed in the resulting data set. However, adequate depth and coverage of the mtDNA genome sequence data prevented false-positive base calls.

Missing data comprised 32 consecutive bp, spanning nps 7525–7556 (Table 1). Other mtDNA genome sequences that we have generated following the same protocol have also exhibited a low coverage or dropout in exactly the same region [28]. Figure S1 shows that low GC content regions are characterised by a poor coverage; in particular in the region 7525–7556, GC content is only 25.0% compared to 44.4% GC for the whole mtDNA. It is therefore suspected that this region of the mtDNA genome is energetically sub-optimal (AT-rich) for two rounds of hybridisation and stringency washes in the ionic and temperature conditions used here [28], which may produce secondary DNA structures that adversely affect hybridisation-based DNA capture [30]. In addition, re-amplification of the enriched libraries was done with AmpliTaq Gold (Applied Biosystems; see [28]), a *Taq* polymerase known to be biased towards high GC content [31].

The ancient mtDNA haplotype of individual UZOO-74 differed from the RSRS at 58 nucleotide positions (Table 1). Among these, 51 substitutions define the sub-hg C1 in accordance with the current phylogeny (www.PhyloTree.org), including back mutations at T182C! and G11914A!, which are identical by state to the RSRS. In addition, UZOO-74 showed five private substitutions (G247A!, A8577G, A11605t, A12217G and T16189C!, including two additional back mutations in the hypervariable region). These five additional nucleotide differences were directly amplified and sequenced from two different extracts in order to verify whether they represented true private mutations defining a novel C1 sub-clade (Table S1). All five mutations have been confirmed by direct sequencing and were taken into account in further phylogenetic analysis of hg C1 sequences. Importantly, we did not observe any SNPs characteristic of other hgs, nor any mixed signals that could indicate systematic DNA degradation or DNA contamination from exogenous sources.

We analysed the pattern of nucleotide misincorporation at the 3′ and 5′-ends of the DNA fragments in order to assess whether the estimated age of the molecules reflects the age of the sample [32–33]. We observed a C-to-T substitution frequency of 22.4% at the 5′-end (Figure S2), which sits well with previous findings that suggested a correlation between this frequency and the age of the samples [34], whereby samples older than 500 years had a C-to-T substitution frequency >10%. The Bayesian statistical framework implemented in mapDamage v2.0.1 [33] also provided simulated posterior distribution of three parameters of the damage model: λ, probability of terminating in overhang; ∂D, probability of cytosine deamination in double strands; and ∂S, probability of cytosine deamination in single strands. The posterior distribution of these parameters all departed from 0 (λ: mean, 0.582; standard deviation, 0.021; ∂D: mean, 0.036; standard deviation, 0.001; ∂S: mean, 0.651; standard deviation, 0.041), in accordance with

Table 1. Positions and nucleotide changes in the Yuzhnyy Oleni Ostrov C1f haplotype when compared to the Reconstructed Sapiens Reference Sequence.

Coverage	99.8%
Missing bases	np 7525–7556
Average redundancy	68.5 (23.2 stdv)
Minimum redundancy	0
Maximum redundancy	126
Mutations	C146T, C152T, T182C!, C195T, **G247A!**, A249d, 290–291d, T489C, *522.AC*, A769G, A825t, A1018G, A2758G, C2885T, T3552a, T3594C, G4104A, T4312C, A4715G, G7146A, C7196a, T7256C, A7521G, T8468C, **A8577G**, G8584A, T8655C, A9545G, C10400T, T10664C, A10688G, C10810T, C10915T, **A11605t**, G11914A!, **A12217G**, G13105A, A13263G, G13276A, T13506C, T13650C, T14318C, T14783C, G15043A, G15301A, A15487t, A16129G, *A16183c*, T16187C, **T16189C!**, G16230A, T16278C, T16298C, C16311T, T16325C, C16327T, *C16519T*

Nucleotide changes in bold represent mutations in the Yuzhnyy Oleni Ostrov haplotype that are new within the C1 clade. Transitions are reported with upper case letters and transversions with lower case letters. "!" indicates a back mutation. np, nucleotide position. stdv, standard deviation.
doi:10.1371/journal.pone.0087612.t001

the DNA sequences generated from the UZOO-74 individual arising from aDNA molecules, and not from contamination by more recent DNA molecules during post-excavation handling. The results of the DNA damage analyses support the authenticity of the aDNA data presented here.

The Sequence from Mesolithic Yuzhnyy Oleni Ostrov Defines a Novel Lineage within the C1 Phylogeny

Upon confirmation of the five novel mutations in the C1 mtDNA genome of individual UZOO-74, we genotyped the same SNPs in the two other C1 individuals UZOO-7 and UZOO-8 from Yuzhnyy Oleni Ostrov. Direct sequencing confirmed the presence of all five novel SNPs, suggesting that the three C1 individuals from Yuzhnyy Oleni Ostrov were maternally related. Their mtDNA genomes may be strictly identical, or they may display differences in the form of additional private SNPs at coding region positions that have not been sequenced in these remaining individuals. The precise nature of the genetic relationships between these individuals cannot be inferred from the archaeological and genetic data currently available.

A search against the public Phylotree database yielded no match for the newly sequenced C1 mtDNA haplotype in 16810 modern complete mtDNA genomes (entries, mtDNA tree build 15 (30 Sep 2012) on PhyloTree.org [35]). The five SNPs identified in these individuals, and among these, the three coding region mutations A8577G, A11605t and A12217G in particular, represent novel sub-clade defining mutations that have not been reported together within a single hg C1 haplotype before. We therefore assigned them to a distinct new clade, which we tentatively named "C1f" following the conventional nomenclature (Figure 3). The resulting phylogenetic reconstruction shows that clade C1 is now characterised by six monophyletic sub-clades: C1a, C1b, C1c, C1d, C1e and C1f. The tree topology suggests that the Eurasian C1 sub-clades, the East Asian C1a, the rare C1f branch from Yuzhnyy Oleni Ostrov and the Icelandic C1e split early from the most recent common ancestor of the C1 clades and evolved independently (Figure 3).

Discussion

Under-sampling of the Mitochondrial Genome Diversity

In the present study, we established that the hg C1 mtDNA genome sequence carried by the Mesolithic individuals of the Yuzhnyy Oleni Ostrov site in north-western Russia defines a new clade, C1f, within the hg C1 phylogeny. Because of the polytomous topology of the hg C1 tree, no direct phylogenetic relationship could be established between C1f and the other well geographically defined C1 clades. As a result, clear inferences regarding the origin and evolutionary history of the C1f clade will remain difficult to draw, unless future sequencing of complete mtDNA genomes uncovers sequences closely related to the C1f genome sequenced here.

The absence of a direct match with sequences in databases of complete mtDNA genomes could be explained by under-sampling of mtDNA genomes in modern human populations. The number of published modern-day *Homo sapiens* complete mtDNA genome sequences is still small compared to that of HVR-I sequences. As such it is not too surprising that studies regularly report the discovery of novel clades and lineages (e.g., hg C1e [10]; within hg C1d [7]). Furthermore, the geographical coverage of modern-day populations for complete mtDNA genome sequencing is still unequally distributed, and the sampling so far has focused either on few specific populations or on particular hgs (e.g., [36–38]). As a consequence, mtDNA genomes available from the literature can still only provide an incomplete yet biased picture of the full, extant mtDNA diversity.

Absence of Match for C1f in Asia

Asia, and more precisely Siberia, could be considered as potential places of origin for the C1f clade identified in the Mesolithic site of Yuzhnyy Oleni Ostrov. This hunter-gatherer group was indeed shown to exhibit mtDNA affinities with modern-day populations of western and southern Siberia, the Altai region, or Mongolia [20]. The hypothesis of an Asian origin for the C1f sub-clade is also supported by the fact that most of the diversity of hg C is found in present-day populations of East Eurasia [3]. Sequences closely related to hg C1f may persist in modern-day populations of East Eurasia but remain undetected to date, as mtDNA genomes for these populations have not been as densely and extensively sampled as, for example, European populations.

Absence of Match for C1f in Europe

Despite the dense sampling of mtDNA in modern-day populations of Europe, only a few hg C1 HVR-I and no hg C1f mtDNA genome sequences were detected. The close matches for the HVR-I sequence of C1f did not display the back mutation T16189C! (Figure 2) and hence, none matched the C1f HVR-I haplotype exactly. However, np 16189 has been described as one of the top five transitional hotspots in the human control region [39], and hence provides little phylogenetic discrimination power. It is possible that these European haplotypes belong to the C1f clade without harbouring the mutation at np 16189. Therefore, additional SNPs in the coding region are required to definitely rule out these Eurasian C1 haplotypes as potential members of the C1f clade, and potential persistence of hg C1f in Europe since the Mesolithic.

Extinction or Near-extinction of C1f due to Post-Mesolithic Population Dynamics

Low frequencies and a restricted distribution seem to have been characteristic of hg C1 already in Mesolithic times, as hg C1 could not be detected in any of the other European Mesolithic populations sampled for ancient mtDNA in Eurasia further west [20]: in central/eastern Europe [25], [27], [40], and in Scandinavia [41]. This suggests an under-sampling of Mesolithic populations for aDNA, mating isolation of the Yuzhnyy Oleni Ostrov population, and/or influences from Siberian populations that had not reached Central Europe. Because of its low frequency, the distribution of hg C1 is prone to be affected by demographic processes, such as genetic drift or population replacements that may have occurred since Mesolithic times. Eventually, hg C1 may have reached extremely low frequencies or have gone extinct, thus preventing it from being detected in present-day European populations. The effects of these population processes can be observed at the population level, as the Yuzhnyy Oleni Ostrov group, similarly to the other described Mesolithic populations of Europe, was indeed shown to exhibit little genetic continuity with present-day Europeans [20]. Significant dissimilarities have been shown between the mtDNA gene pool of European Mesolithic populations characterised by a low diversity and high frequencies of hg U sub-clades (U2, U4, U5 and U8 in particular), and the rather homogeneous mtDNA makeup of present-day Europeans, which arrived during the Neolithic transition and subsequent periods [20], [40–42].

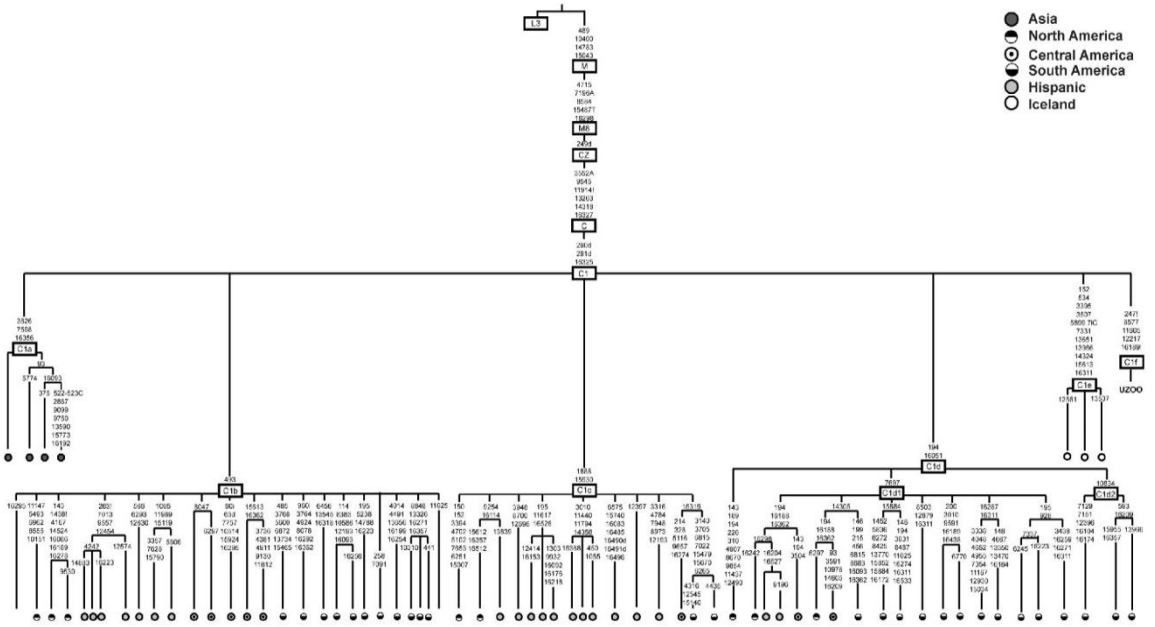


Figure 3. Median joining phylogenetic tree of haplogroup C1 complete mitochondrial genomes. A haplogroup sequence L3 sequence was chosen as the root of the tree. Mutations are reported according to the Reconstructed Sapiens Reference Sequence. "d" represents deletions. "i" represents insertions.
doi:10.1371/journal.pone.0087612.g003

Absence of Match for C1f in the Americas

The Americas also remain under-sampled for complete mtDNA genomes and could be suggested as a potential geographical origin for the C1f lineage, as it has been for the Iceland-restricted C1e sub-clade [10]. For C1e, an American origin through mating of Viking explorers with Native American women sometime earlier than 300 years ago was proposed by [10]. Among other hypotheses including that of a European origin, an American origin was favoured on the basis that most of the hg C1 diversity is found on the American continent, despite the fact that no sequence belonging to hg C1e could be detected in the Americas (or anywhere else). This lack of match was explained by under-sampling of the American mtDNA genome diversity [10]. In any case, if admixture between Native Americans and Vikings did occur, it must have been limited, as no other American-specific lineage (e.g. hg A2, B2, D1, C1b, C1c, C1d) was detected in Iceland.

As for Mesolithic Europe, the possibility of a direct prehistoric genetic influence from the Americas is highly unlikely. However, in the eventuality that further sampling of complete mtDNA genomes in the Americas reveals the presence of additional haplotypes belonging to C1f, it would suggest an evolutionary history similar to that of mtDNA hg X2. Like hg C1, hg X2 displays relatively low frequencies albeit with a global distribution in the Northern hemisphere. For example, clade X2a was observed in Europe in the West, in the Near East, Europe, Central Asia, Siberia as well as North America [43]. One model for the present-day distribution of hg X2 suggests that clade X2a split early from the rest of the X2 lineages in the Near East, and reached east Siberia before participating in the second wave of migration into the Americas through admixture with Beringian populations [44]. A similar scenario involving an early split of the different C1 clades in Asia followed by their spread and subsequently isolated evolution could be considered as an explanation for the wide geographical distribution of hg C1 in general. However, this scenario currently lacks substantial support.

Similar Genetic Pre-history for the Icelandic-specific C1e and the Mesolithic C1f European Sub-clades

While the updated phylogeography of hg C1 does not allow defining the precise origins and divergence times of the C1f and C1e clades, the observation of C1f in Mesolithic Yuzhnyy Oleni Ostrov brings us to reconsider the hypotheses concerning the origins of C1e. Building on a hypothesis proposed by [10], we suggest that the Icelandic-specific C1e sub-clade could have had a recent origin in northern Europe rather than an American origin. This hypothesis is relevant with regard to the origins of the Icelandic population, as Iceland was discovered and first settled by Scandinavian Vikings around 1,130 years ago. Vikings raids extended as far from their homeland in Scandinavia as France, Spain and Sicily, but their main expansion range comprised western Russia, the Baltic region, Scandinavia, and the British Isles [16]. The study of the mtDNA diversity of present-day Icelanders identified that most of the Icelandic mtDNA lineages had Norse (from Scandinavia) or Gaelic origins (from the British Isles) and that the Icelandic gene pool had strongly been impacted by genetic drift [16–17,45].

Considering the Scandinavian origins of Icelanders and the identification of the sister clade C1f in Mesolithic North East Europe, it can be proposed that the Icelandic-specific C1e and C1f sub-clades might have both split from the common ancestors of the C1 lineages somewhere in Eurasia and later reached northern Europe during independent or similar migrations (before the Mesolithic for C1f). Therefore, the rare occurrence of the C1e and C1f sub-clades in Europe could be the result of their dilution within the pre-existing European mtDNA diversity when these

lineages reached Europe. Of note, a contrasting pattern of elevated frequency and diversity was observed for the American C1 sister-clades (C1b, C1c and C1d): all three American sub-clades signal important population expansion during the initial peopling of the continent, which was void of human occupation and thus competing lineages. The distribution of the C1e sub-clade restricted to Iceland, associated with the presence of the novel sub-clade C1f in a region neighbouring the homeland of Vikings and clearly predating the Viking expansion, lends support to the hypothesis that hg C1e might have been brought in by the Vikings who first colonised Iceland. The presence of a novel sub-clade (C1f) closely related to the Icelandic-specific C1e sub-clade in a region neighbouring the homeland of Vikings and clearly predating the Viking expansion lends support to the hypothesis that hg C1e might have been brought in by the Vikings who first colonised Iceland. While the C1e sub-clade might have been preserved at detectable frequencies in the Icelandic population due to the effects of founder event, it most likely has gone extinct in the source population in northern Europe as a consequence of its low frequency. In contrast, due to the small size of the population through time, Icelandic mtDNA diversity has been greatly affected by genetic drift and increased rates of mtDNA haplotype extinctions [45]. As such, the C1e clade would be more likely to survive in the potential North European source population than in Iceland [45], but the extensive sampling of the Icelandic population makes it more likely to be detected there than anywhere else in North Europe. The potential long-term survival of C1 lineages in prehistoric Europe is highly relevant to the discussion about the prehistoric interactions between the ancestral populations of Europeans, Siberians and Native Americans. It is consistent with recently published genomic data from a 24,000 year-old Upper Paleolithic individual from Mal'ta, South Siberia [46]. Interestingly, this individual was shown to belong to the western Eurasian hg U, which was also the most frequent hg found in Yuzhnyy Oleni Ostrov Mesolithic individuals (64%) [20]. Genome-wide data from Upper Palaeolithic Mal'ta revealed affinities with both present-day western Eurasian and Native Americans, and further supports gene-flow between the ancestral populations of Europeans and Native Americans prior to the colonisation of the Americas [46]. The new C1f lineage thus bridges the geographic gap between the Icelandic, the Siberian and the Native American C1 lineages and argues for the presence of C1 lineages, albeit at low frequency, in prehistoric West Eurasia.

Materials and Methods

Ethics Statement

No specific permits were required for the described field studies.

Archaeological Samples

The three tooth samples analysed in this study were collected at the Mesolithic site of Yuzhnyy Oleni Ostrov, Onega Lake, Karelia, Russian Federation (61°30′N 35°45′E). These samples are under the custody of V.K. at the Peter the Great Museum of Anthropology and Ethnography (Kunstkamera) RAS, St Petersburg, Russian Federation, and were previously subjected to aDNA analyses in [20]. The three samples are identified as follows: UZOO-7 (MAE RAS collection number 5773–7, grave number 56), UZOO-8 (MAE RAS collection number 5773–8, grave number 57), and UZOO-74 (MAE RAS collection number 5773–74, grave number 114). Individuals UZOO-7 and UZOO-8 were found in adjacent graves (grave number 56 and grave number 57), whereas UZOO-74 was found in a grave located at the other end of the graveyard (grave number 114; see map in [47]).

Ancient DNA Extraction

Among the three Mesolithic individuals from Yuzhnyy Oleni Ostrov shown to carry hg C1 (UZOO-7, UZOO-8, UZOO-74) in [20], individual UZOO-74 was selected for mtDNA genome sequencing on the basis of its subjectively good preservation and robust performance in previous Polymerase Chain Reaction (PCR) amplification experiments [20]. DNA extractions followed established protocols as described previously [20].

Enrichment of Ancient Human Mitochondrial DNA

Ancient DNA extracts from specimens preserved in soil are expected to contain DNA molecules of various origins. In addition to the highly degraded DNA of the specimen under study, environmental, microbial DNA (bacteria and fungi), as well as from unidentified sources, has been shown to constitute a major proportion in the pool of DNA molecules present in aDNA extracts (e.g., [48]). The presence of a mixed population of DNA molecules from various organisms hampers the reliable sequencing of the DNA fragments of interest. Here, targeted enrichment of ancient human mtDNA of the extract for individual UZOO-74 was a crucial step prior to sequencing on the Ion Torrent PGM. Our aim was to increase the concentration of mtDNA fragments above the concentration threshold required for obtaining unambiguous sequencing at sufficient mtDNA genome coverage on an Ion Torrent PGM 316 chip. Ancient DNA libraries were enriched for human mtDNA using a hybridisation-based method described in [28]. All enrichment steps were performed twice for sample UZOO-74 to produce a 'second round enrichment' DNA library.

Sequencing on the Ion Torrent PGM

The enriched library DNA was prepared for Ion Torrent sequencing by re-amplification using Ion Torrent barcoded adapters (Adapter A: 5′-CCATCTCATCCCTGCGTGTCT-CCGACTCAGAAAAAGGTGTTGTTAGGAATGCGAGA-3′; Adapter B: 5′-CCTCTCTATGGGCAGTCGGTGATAGGA-TAGGTCGTTGCTGTGTA-3′). Eight reactions with a total volume of 25 µL were re-amplified using 1 µL of purified library DNA as template. Final reaction conditions comprised of 1x AmpliTaq Gold buffer II, 2.5 mM $MgCl_2$, 2.5 U AmpliTaq Gold (Applied Biosystems), 250 µM of each dNTP (Invitrogen), and 0.5 µM of each PCR primer. The thermocycling profile consisted of 94°C for 12 min, followed by 12 cycles of 30 s at 95°C, 30 s at 60°C and 45 s at 72°C, followed by a final 10 min at 72°C. The eight PCRs were pooled and purified using MiniElute spin columns (Qiagen), then eluted into 15 µL as per the manufacturer's instructions. The DNA was sized and quantified via gel electrophoresis against size markers (HyperLadder V, Bioline) and a Nanodrop 2000 (Thermo Scientific). Library DNA was size-selected above 120 bp and further purified, to remove adaptor dimers, using Qiagen's gel extraction purification kit following manufacturer's instructions.

Prior to sequencing, individual libraries were assessed for fragment size distribution and DNA concentration using a Bioanalyzer 2100 (Agilent Technologies) following manufacturer's instructions. The quantified indexed library DNA was pooled to an equimolar concentration alongside other samples. The pooled library DNA was adjusted to a final concentration of 10–15 pM prior to amplification (by emulsion PCR) and enriched for positive Ion Sphere Particles (ISPs) using the Ion Torrent One Touch System II (Life Technologies) and the Ion One Touch 200 template kit v2 DL (Life Technologies), following manufacturer's

instructions. Templated ISPs were sequenced on a 316 micro-chip (up to 100 Mb of data) using the Ion Torrent Personal Genome Machine (PGM; Life Technologies) and the Ion PGM 200 sequencing kit v2 chemistry (Life Technologies) for 130 cycles (520 flows). After sequencing, the individual sequence reads were filtered within the PGM software to remove low quality and polyclonal sequences. Sequences matching the PGM 3′ adaptor were also automatically trimmed prior to bioinformatics analysis.

Bioinformatics and Sequence Analysis

Next generation sequencing data (Ion Torrent PGM platform) from the mtDNA capture was processed using an in-house customizable analytical pipeline based on available scripts. Reads were de-multiplexed by barcode, not allowing for any mismatch, using FASTX-Toolkit (http://hannonlab.cshl.edu/fastx_toolkit/index.html). Cutadapt v1.1 [49] was then used to trim adapters and filter the reads for quality and length (five successive rounds of trimming with 33.3% error rate allowed; minimum quality value of Phred 20; minimum and maximum length of 25 and 110 bp after trimming). The filtered reads were checked with FastQC (http://www.bioinformatics.bbsrc.ac.uk/projects/fastqc), before mapping against the RSRS [8] using TMAP v3.2.1 (https://github.com/nh13/TMAP) with the following options: -g 3 -M 3 -n 7 -v stage1–stage-keep-all map1–seed-length 12–seed-max-diff 4 stage2 map2–z-best 5 map3–max-seed-hits 10. Mapped reads with mapping quality below Phred 30 and all duplicates were removed using both Samtools v0.1.18 [50] and the MarkDuplicates option of Picard Tools v1.79 (http://picard.sourceforge.net). GC content of mapped reads was analysed using the CollectGcBiasMetrics tool of Picard Tools v1.79. Misincorporation patterns were assessed using MapDamage v2.0.1 [10–11] and default parameters. The resulting sequence assembly was analyzed using Geneious Pro (v5.6.2) software in order to build the consensus sequence and assign a haplotype following the latest nomenclature and phylogeny on PhyloTree.org (Build 15 (30 Sep 2012)).

Genomic coverage by unique reads was compared to the distribution of GC content across the mtDNA using a custom R script (see Figure S1). Coverage depth was directly estimated from the unique read mapping file by counting the number of reads covering each mtDNA position. We also calculated the average GC percentage in a sliding window of 55 nt centered around each mtDNA position, taking in account the circular nature of the molecule. The size of the window corresponds to the average size of the unique mapped reads (see the Results section and Figure S2).

SNP Confirmation by Direct Sequencing and Minisequencing

Selected regions of the mtDNA genome were sequenced independently via single and multiplex PCR amplification followed by direct sequencing (Table S1), and SNaPshot minisequencing following established PCR conditions [30] in order to:

1) verify the HVR-I sequence between np 15997 and 16410;
2) verify the 22 coding region hg diagnostic SNPs targeted by the GenoCoRe22 reaction;
3) interrogate the deletions at np 249, 290, 291, which are diagnostic of the CZ and C1 sub-haplogroups (Figure 3);
4) confirm the new SNPs identified here in the mtDNA genome sequence of specimen UZOO-74, but also to type these in the two other C1 carriers from Yuzhnyy Oleni Ostrov, UZOO-7 and UZOO-8.

Typing of coding-region SNPs using the GenoCore22 reaction was performed using the protocol described in [20].

Authentication of the Ancient mtDNA Sequence

The aDNA sequence data was validated using three lines of evidence: 1) monitoring of contamination, 2) reproducibility, 3) phylogenetic consistency.

1) Pre-PCR DNA work was carried out at the Australian Centre for Ancient DNA (ACAD), University of Adelaide, a purpose-built positive air pressure laboratory dedicated to aDNA studies, which is physically isolated from any molecular biology laboratory amplifying DNA. Routine decontamination of the laboratory surfaces and instruments involves exposure to UV radiation and thorough cleaning using DNA oxidants such as bleach, Decon (Decon labs) and Isopropanol. In order to protect the laboratory environment from human DNA, researchers wear protective clothes including body suit, a facemask, a face shield, gumboots, and three layers of surgical gloves that are changed on a regular basis. Obvious large-scale contamination within the laboratory or in the reagents were monitored and controlled by blank controls (one extraction blank for every five ancient samples and two PCR/GenoCoRe22 blank controls for every six reactions). In addition, no haplotype similar to any of those possessed by laboratory members was consistently amplified from aDNA extracts.
2) For all three Yuzhnyy Oleni Ostrov individuals, the HVR-I sequences and GenoCoRe22 profiles could be replicated from two samples extracted independently [20]. For UZOO-74, the same HVR-I sequence and coding region SNPs were obtained by direct sequencing/GenoCoRe22 minisequencing and sequencing the complete mtDNA genome. Private SNPs identified by complete mtDNA genome sequencing were also confirmed by direct sequencing.
3) The phylogenetic consistency of the combination of variable positions in the mtDNA genome was an additional indicator of the authenticity of the sequence. The fifty-one nucleotide differences with RSRS were 100% consistent with the phylogenetic position of UZOO-74 on the hg C1 branch. No deviation from this position could be detected. The combination of five additional mutations was found to be unique to the haplotype sequenced here. None of these additional mutations was found to define branches within the human mtDNA tree, thus providing little support for them arising from contamination. In addition, considering the multiple replications performed to type the SNPs of interest, jumping PCR is thought to have had little or no impact on the hg C1 haplotype presented here.

Phylogenetic Analysis of Haplogroup C1

In order to construct a phylogenetic network of the hg C1 HVR-I sequences displaying the C1 mutational pattern A16129G, T16187C, C16189T, G16230A, T16278C, T16298C, C16311T, T16325C, C16327T in Eurasia were gathered from the literature (Table S2). Sequences of whole mtDNA genomes belonging to hg C1 were also compiled from the online GenBank database on the basis of the list published in [10] (Table S2) and relevant new publications since. Sequences were corrected for known sequencing errors and ambiguities and, in particular, sequence length polymorphisms following recommendations in [39] and [51]. Mutations 309.1C, 315.1C, A16182c, A16183c, 16193.1C and C16519T were systematically ignored, as these positions are

known to represent mutational hotspots and/or recurrent sequencing artefacts, which create reticulations in the phylogenetic analysis [35]. A tree was then constructed manually for complete mtDNA genome sequences on the basis of the tree constructed in [10] taking the latest mtDNA phylogeny into account (www.phylotree.org; [35]).

Divergence Time Estimate

The split of the six clades from the hg C1 root could theoretically be dated using mtDNA genome sequences. However, the use of molecular data for dating divergence times can be problematic. The most commonly used method in modern mtDNA studies is based on the calculation of the ρ statistic [52]. However, this method has been shown to produce inaccurate dates, especially when the genetic data was collected from populations with complex demographic histories (e.g., [53]). The tree topology of hg C1 suggests that this is likely to be the case for the six C1 clades. The tree obtained is indeed very imbalanced: the three American clades show strong signals of a recent expansion, contrary to the three clades in Asia, Iceland and Mesolithic Yuzhnyy Oleni Ostrov. Moreover, a limitation of molecular dating methods is the inaccuracy of the estimation of the human mtDNA substitution rate and the fact that they rely on the use of a constant mutation rate through time (e.g., [54]). A recent study could largely improve the accuracy of such substitution rate estimates by correcting for purifying selection but still rely on a fossil calibration point, i.e., the human/chimp split ~6.5 million years ago [55]. Based on this method, divergence time estimate were calculated for the modern subclades C1a, C1b and C1c at around 17,100 yBP (95% Confidence Interval: 12,000–22,500 yBP). The use of dated ancient mtDNA genome sequences to calibrate the molecular clock has so far not been considered by this method. However, recent studies have utilised ancient mtDNA genomes as tip calibrations in a Bayesian Markov Chain Monte Carlo framework using the program BEAST [56], which allows estimation of a substitution rate that is allowed to vary in time and can also take temporally heterogeneous mtDNA genome sequences into account [27–28]. However, the case of the polytomous hg C1 tree revealed problematic [7] it was recently demonstrated that a biased or incomplete representation of C1d lineages could lead to a much younger divergence time estimate (7,000–9,000 yBP; [7]). The addition of new data resulted in a revised divergence time to 18,700+/−1,400 yBP, in accordance with estimates for all other Pan-American haplogroups (15,000–18,000 yBP; [4], [7], [42], [57]). With only one C1f and very few Icelandic C1e mtDNA genomes, it is apparent that the sampling of the genetic diversity for these clades is not yet extensive enough to reliably calculate a divergence time estimates. Similarly, it seems also unreasonable to reconstruct the timing of the arrival of hg C1 lineages in Europe via coalescence age dating and/or determine whether all European C1 lineages reached Europe as part of the same migration as the Yuzhnyy Oleni Ostrov C1f branch or as part of other movements from the East.

Accession Numbers

The deep-sequencing data were deposited on the NCBI Sequence Read Archive, accession number: SRP033724.

Supporting Information

Figure S1 Coverage depth and GC content for the haplogroup C1f mitochondrial genome (individual UZOO-74). Mapping coverage of unique reads (in blue) is given per base. Local GC content (red) is shown for 55-bp intervals. The arrow indicate missing data between nucleotide positions 7277 and 7556, a region also characterized by low GC content.
(PDF)

Figure S2 Analysis of DNA damage patterns. A. Four upper plots: Frequencies of the A, C, G and T bases according to the nucleotide positions within the read (within the grey box) and outside the read (outside the grey box). Two lower plots: Frequency distribution of specific substitutions from the 5′-end (left) to the 3′-end (right) of the read sequence. B. Two upper plots: Read length distribution. Two lower plots: Observed cumulative frequency of C to T and G to A misincorporations. C. Observed frequencies of nucleotide misincorporation and simulated Bayesian posterior predictive intervals obtain from model fitting. D Simulated posterior distribution of model parameters: Lambda, probability of terminating in overhang (λ); DeltaD, probability of cytosine deamination in double strands (∂D); and DeltaS, probability of cytosine deamination in single strands (∂S).
(PDF)

Table S1 Primer sequences.
(PDF)

Table S2 Haplogroup C1 mitochondrial sequences used to construct control region and complete mitochondrial genome C1 phylogenies.
(PDF)

Acknowledgments

We thank Marta Kasper for help with project logistics; Jeremy Austin, and Jessica Metcalf for helpful comments; Thomas Duffey for support with match searches. Members of The Genographic Project Consortium: Syama Adhikarla (Madurai Kamaraj University, Madurai, Tamil Nadu, India), Christina J. Adler (University of Adelaide, South Australia, Australia), Jaume Bertranpetit (Universitat Pompeu Fabra, Barcelona, Spain), Andrew C. Clarke (University of Otago, Dunedin, New Zealand), David Comas (Universitat Pompeu Fabra, Barcelona, Spain), Matthew C. Dulik (University of Pennsylvania, Philadelphia, Pennsylvania, United States), Jill B. Gaieski (University of Pennsylvania, Philadelphia, Pennsylvania, United States), Arun Kumar Ganesh Prasad (Madurai Kamaraj University, Madurai, Tamil Nadu, India), Marc Haber (Universitat Pompeu Fabra, Barcelona, Spain; Lebanese American University, Chouran, Beirut, Lebanon), Li Jin (Fudan University, Shanghai, China), Matthew E. Kaplan (University of Arizona, Tucson, Arizona, United States), Shilin Li (Fudan University, Shanghai, China), Begoña Martínez-Cruz (Universitat Pompeu Fabra, Barcelona, Spain), Elizabeth A. Matisoo-Smith (University of Otago, Dunedin, New Zealand), Nirav C. Merchant (University of Arizona, Tucson, Arizona, United States), R. John Mitchell (La Trobe University, Melbourne, Victoria, Australia), Amanda C. Owings (University of Pennsylvania, Philadelphia, Pennsylvania, United States), Laxmi Parida (IBM, Yorktown Heights, New York, United States), Ramasamy Pitchappan (Madurai Kamaraj University, Madurai, Tamil Nadu, India), Daniel E. Platt (IBM, Yorktown Heights, New York, United States), Lluis Quintana-Murci (Institut Pasteur, Paris, France), Colin Renfrew (University of Cambridge, Cambridge, United Kingdom), Daniela R. Lacerda (Universidade Federal de Minas Gerais, Belo Horizonte, Minas Gerais, Brazil), Ajay K. Royyuru (IBM, Yorktow Heights, New York, United States), Fabrício R. Santos (Universidade Federal de Minas Gerais, Belo Horizonte, Minas Gerais, Brazil), Theodore G. Schurr (University of Pennsylvania, Philadelphia, Pennsylvania, United States), Himla Soodyall (National Health Laboratory Service, Johannesburg, South Africa), David F. Soria Hernanz (National Geographic Society, Washington, District of Columbia, United States), Pandikumar Swamikrishnan (IBM, Somers, New York, United States), Chris Tyler-Smith (The Wellcome Trust Sanger Institute, Hinxton, United Kingdom), Arun Varatharajan Santhakumari (Madurai Kamaraj University, Madurai, Tamil Nadu, India), Pedro Paulo Vieira (Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil), Miguel G. Vilar (University of Pennsylvania, Philadelphia, Pennsylvania, United States), R. Spencer Wells

(National Geographic Society, Washington, District of Columbia, United States), Pierre A. Zalloua (Lebanese American University, Chouran, Beirut, Lebanon), Janet S. Ziegle (Applied Biosystems, Foster City, California, United States).

Author Contributions

Conceived and designed the experiments: CDS WH PB BL AC. Performed the experiments: CDS WH PB BL JT. Analyzed the data: CDS WH PB BL JS. Contributed reagents/materials/analysis tools: CDS WH PB BL JS JT OB VM VK AC TGC. Wrote the paper: CDS WH AC.

References

1. Kong QP, Yao YG, Sun C, Bandelt HJ, Zhu CL, et al. (2003) Phylogeny of East Asian mitochondrial DNA lineages inferred from complete sequences. Am J Hum Genet 73: 671–676.
2. Metspalu M, Kivisild T, Bandelt HJ, Richards M, Villems R (2006) The pioneer settlement of modern humans in Asia. In: Bandelt HJ, Macaulay V, Richards M, editors. Human mitochondrial DNA and the evolution of Homo sapiens. Berlin: Springer-Verlag. 181–199.
3. Derenko M, Malyarchuk B, Grzybowski T, Denisova G, Rogalla U, et al. (2010) Origin and post-glacial dispersal of mitochondrial DNA haplogroups C and D in Northern Asia. PLoS One 5: e15214.
4. Tamm E, Kivisild T, Reidla M, Metspalu M, Smith DG, et al. (2007) Beringian standstill and spread of Native American founders. PLoS One 2(9): e829.
5. Achilli A, Perego UA, Bravi CM, Coble MD, Kong QP, et al. (2008) The phylogeny of the four pan-American MtDNA haplogroups: implications for evolutionary and disease studies. PLoS One 3(3): e1764.
6. Volodko N, Starikovskaya E, Mazunin I, Eltsov N, Naidenko P, et al. (2008) Mitochondrial genome diversity in arctic Siberians, with particular reference to the evolutionary history of Beringia and Pleistocenic peopling of the Americas. Am J Hum Genet 82: 1084–1100.
7. Perego UA, Angerhofer N, Pala M, Olivieri A, Lancioni H, et al. (2010) The initial peopling of the Americas: a growing number of founding mitochondrial genomes from Beringia. Genome Res 20(9): 1174–1179.
8. Behar DM, van Oven M, Rosset S, Metspalu M, Loogväli EL, et al. (2012) A "Copernican" reassessment of the human mitochondrial DNA tree from its root. Am J Hum Genet 90(4): 675–684.
9. Starikovskaya EB, Sukernik RI, Derbeneva OA, Volodko NV, Ruiz-Pesini E, et al. (2005) Mitochondrial DNA diversity in indigenous populations of the southern extent of Siberia, and the origins of Native American haplogroups. Ann Hum Genet 69: 67–89.
10. Ebenesersdóttir SS, Sigurðsson A, Sánchez-Quinto F, Lalueza-Fox C, Stefánsson K, et al. (2011) A new subclade of mtDNA haplogroup C1 found in Icelanders: evidence of pre-Columbian contact? Am J Phys Anthropol 144(1): 92–99.
11. Forster P, Harding R, Torroni A, Bandelt H (1996) Origin and evolution of Native American mtDNA variation: a reappraisal. Am J Hum Genet 59: 935–945.
12. Bandelt HJ, Herrnstadt C, Yao YG, Kong QP, Kivisild T, et al. (2003) Identification of Native American founder mtDNAs through the analysis of complete mtDNA sequences: some caveats. Ann Hum Genet 67(6): 512–524.
13. Shields G, Schmiechen A, Frazier B, Redd A, Voevoda M, et al. (1993) mtDNA sequences suggest a recent evolutionary divergence for Beringian and northern North American populations. Am J Hum Genet 53: 549–562.
14. Pfeiffer H, Brinkmann B, Hühne J, Rolf B, Morris AA, et al. (1999) Expanding the forensic German mitochondrial DNA control region database: genetic diversity as a function of sample size and microgeography. Int J Legal Med 112: 291–298.
15. Rando JC, Cabrera VM, Larruga JM, Hernandez M, Gonzalez AM, et al. (1999) Phylogeographic patterns of mtDNA reflecting the colonization of the Canary Islands. Ann Hum Genet 63: 413–428.
16. Helgason A, Sigurðardóttir S, Gulcher JR, Ward R, Stefánsson K (2000) mtDNA and the origin of the Icelanders: deciphering signals of recent population history. Am J Hum Genet 66: 999–1016.
17. Helgason A, Nicholson G, Stefánsson K, Donnelly P (2003) A reassessment of genetic diversity in Icelanders: strong evidence from multiple loci for relative homogeneity caused by genetic drift. Ann Hum Genet 67: 281–297.
18. Bermisheva MA, Kutuev IA, Spitsyn VA, Villems R, Batyrova AZ, et al. (2005) Analysis of mitochondrial DNA variation in the population of Oroks. Genetika 41: 78–84.
19. Grousset R (1970) The Empire of the Steppes: History of Central Asia. Rutgers University Press.
20. Der Sarkissian C, Balanovsky O, Brandt G, Khartanovich V, Buzhilova A, et al. (2013) Ancient DNA reveals prehistoric gene-flow from siberia in the complex human population history of North East Europe. PLoS Genet 9(2): e1003296.
21. Gilbert MT, Kivisild T, Grønnow B, Andersen PK, Metspalu E, et al. (2008) Paleo-Eskimo mtDNA genome reveals matrilineal discontinuity in Greenland. Science 320: 1787–1789.
22. Ermini L, Olivieri C, Rizzi E, Corti G, Bonnal R, et al. (2008) Complete mitochondrial genome sequence of the Tyrolean Iceman. Curr Biol 18: 1687–1693.
23. Knapp M, Horsburgh KA, Prost S, Stanton JA, Buckley HR, et al. (2012) Complete mitochondrial DNA genome sequences from the first New Zealanders. Proc Natl Acad Sci U S A 109(45): 18350–18354.
24. Fu Q, Rudan P, Pääbo S, Krause J (2012) Complete mitochondrial genomes reveal neolithic expansion into Europe. PLoS One (3): e32473.
25. Krause J, Briggs A, Kircher M, Maricic T, Zwyns N, et al. (2010) A complete mtDNA genome of an early modern human from Kostenki, Russia. Curr Biol 20: 231–236.
26. Sánchez-Quinto F, Schroeder H, Ramirez O, Avila-Arcos MC, Pybus M, et al. (2012) Genomic affinities of two 7,000-year-old Iberian hunter-gatherers. Curr Biol 22(16): 1494–1499.
27. Fu Q, Mittnik A, Johnson PL, Bos K, Lari M, et al. (2013) A revised timescale for human evolution based on ancient mitochondrial genomes. Curr Biol 23(7): 553–559.
28. Brotherton P, Haak W, Templeton J, Brandt G, Soubrier J, et al. (2013) Neolithic mitochondrial haplogroup H genomes and the genetic origins of Europeans. Nat Commun 4: 1764.
29. Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, et al. (2005) Genome sequencing in microfabricated high-density picolitre reactors. Nature 437: 376–380.
30. Burrow AA, Marullo A, Holder LR, Wang Y-H (2010) Secondary structure formation and DNA instability at fragile site FRA16B. Nuc Acids Res 38: 2865–2877.
31. Dabney J, Meyer M (2012) Length and GC-biases during sequencing library amplification: a comparison of various polymerase-buffer systems with ancient and modern DNA sequencing libraries. Biotechniques 52: 87–94.
32. Ginolhac A, Rasmussen M, Gilbert MT, Willerslev E, Orlando L (2011) mapDamage: testing for damage patterns in ancient DNA sequences. Bioinformatics 27(15): 2153–2155.
33. Jónsson H, Ginolhac A, Schubert M, Johnson PL, Orlando L (2013) mapDamage2.0: fast approximate Bayesian estimates of ancient DNA damage parameters. Bioinformatics 29(13): 1682–1684.
34. Sawyer S, Krause J, Guschanski K, Savolainen V, Pääbo S (2012) Temporal patterns of nucleotide misincorporations and DNA fragmentation in ancient DNA. PLoS One 7(3): e34131.
35. van Oven M, Kayser M (2009) Updated comprehensive phylogenetic tree of global human mitochondrial DNA variation. Hum Mutat 30: E386–394. www.phylotree.org/.
36. Achilli A, Rengo C, Magri C, Battaglia V, Olivieri A, et al. (2004) The molecular dissection of mtDNA haplogroup H confirms that the Franco-Cantabrian glacial refuge was a major source for the European gene pool. Am J Hum Genet 75: 910–918.
37. Pereira L, Richards M, Goios A, Alonso A, Albarrán C, et al. (2005) High-resolution mtDNA evidence for the late-glacial resettlement of Europe from an Iberian refugium. Genome Res 15: 19–24.
38. Pala M, Olivieri A, Achilli A, Accetturo M, Metspalu E, et al. (2012) Mitochondrial DNA signals of late glacial recolonization of Europe from near eastern refugia. Am J Hum Genet 90: 915–24.
39. Bandelt H, Parson W (2008) Consistent treatment of length variants in the human mtDNA control region: a reappraisal. Int J Legal Med 122: 11–21.
40. Bramanti B, Thomas M, Haak W, Unterlaender M, Jores P, et al. (2009) Genetic discontinuity between local hunter-gatherers and central Europe's first farmers. Science 326: 137–140.
41. Malmström H, Gilbert M, Thomas M, Brandström M, Storå J, et al. (2009) Ancient DNA reveals lack of continuity between neolithic hunter-gatherers and contemporary Scandinavians. Curr Biol 19: 1758–1762.
42. Haak W, Balanovsky O, Sanchez JJ, Koshel S, Zaporozhchenko V, et al. (2010) Ancient DNA from European early Neolithic farmers reveals their near eastern affinities. PLoS Biol 8: e1000536.
43. Reidla M, Kivisild T, Metspalu E, Kaldma K, Tambets K, et al. (2003) Origin and diffusion of mtDNA haplogroup X. Am J Hum Genet 73: 1178–1190.
44. Perego UA, Achilli A, Angerhofer N, Accetturo M, Pala M, et al. (2009) Distinctive Paleo-Indian migration routes from Beringia marked by two rare mtDNA haplogroups. Curr Biol 19: 1–8.
45. Helgason A, Lalueza-Fox C, Ghosh S, Sigurethardóttir S, Sampietro ML, et al. (2009) Sequences from first settlers reveal rapid evolution in Icelandic mtDNA pool. PLoS Genet 5(1): e1000343.
46. Raghavan M, Skoglund P, Graf KE, Metspalu M, Albrechtsen A, et al. (2013) Upper Palaeolithic Siberian genome reveals dual ancestry of Native Americans. Nature. 2013 Nov 20.
47. O'Shea J, Zvelebil M (1984) Oleneostrovskii Mogilnik: Reconstructing Social and Economic Organisation of Prehistoric Hunter-Fishers in Northern Russia. J Anthropol Archaeol 1: 1–40.
48. Green RE, Krause J, Ptak SE, Briggs AW, Ronan MT, et al. (2006) Analysis of one million base pairs of Neanderthal DNA. Nature 444: 330–336.
49. Martin M (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet journal 17(1): 10.
50. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, et al. (2009) The Sequence Alignment/Map format and SAMtools. Bioinformatics 25(16): 2078–2079.

51. Bandelt HJ (2006) Estimation of Mutation Rates and Coalescence Times: Some Caveats. In: Bandelt HJ, Macaulay V, Richards M. Mitochondrial DNA and the evolution of Homo sapiens. Berlin: Springer-Verlag. 64.
52. Saillard J, Forster P, Lynnerup N, Bandelt HJ, Nørby S (2000) mtDNA variation among Greenland Eskimos: the edge of the Beringian expansion. Am J Hum Genet 67(3): 718–726.
53. Cox MP (2008) Accuracy of molecular dating with the rho statistic: deviations from coalescent expectations under a range of demographic models. Hum Biol 80(4): 335–357.
54. Ho SY, Lanfear R, Bromham L, Phillips MJ, Soubrier J, et al. (2011) Time-dependent rates of molecular evolution. Mol Ecol 20(15): 3087–3101.
55. Soares P, Ermini L, Thomson N, Mormina M, Rito T, et al. (2009) Correcting for Purifying Selection: An Improved Human Mitochondrial Molecular Clock. Am J Hum Genet 84(6): 740–59.
56. Drummond A, Rambaut A (2007) BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol Biol 7: 214.
57. Fagundes N, Kanitz R, Bonatto S (2008) A reevaluation of the Native American mtDNA genome diversity and its bearing on the models of early colonization of Beringia. PLoS One 3: e3157.

Manuscript 4: Late Pleistocene Australian Marsupial DNA clarifies the affinities of extinct megafaunal kangaroos and wallabies.

Late Pleistocene Australian Marsupial DNA Clarifies the Affinities of Extinct Megafaunal Kangaroos and Wallabies

Bastien Llamas,*[,†,1] Paul Brotherton,[†,‡,1] Kieren J. Mitchell,[1] Jennifer E.L. Templeton,[1,2] Vicki A. Thomson,[1] Jessica L. Metcalf,[§,1] Kyle N. Armstrong,[1] Marta Kasper,[¶,1] Stephen M. Richards,[1] Aaron B. Camens,[2,3] Michael S.Y. Lee,[3,4] and Alan Cooper*[,1]

[1]Australian Centre for Ancient DNA, School of Earth and Environmental Sciences, University of Adelaide, SA, Australia

[2]School of Biological Sciences, Flinders University, SA, Australia

[3]School of Earth and Environmental Sciences, University of Adelaide, SA, Australia

[4]Earth Sciences Section, The South Australian Museum, SA, Australia

[†]These authors contributed equally to this work.

[‡]Present address: BiSN Laboratory Services, Runcorn, United Kingdom

[§]Present address: BioFrontiers Institute, University of Colorado, Boulder, CO.

[¶]Present address: Forensic Science SA, Adelaide, SA, Australia.

***Corresponding author:** E-mail: bastien.llamas@adelaide.edu.au; alan.cooper@adelaide.edu.au.

Associate editor: Michael Rosenberg

Abstract

Understanding the evolution of Australia's extinct marsupial megafauna has been hindered by a relatively incomplete fossil record and convergent or highly specialized morphology, which confound phylogenetic analyses. Further, the harsh Australian climate and early date of most megafaunal extinctions (39–52 ka) means that the vast majority of fossil remains are unsuitable for ancient DNA analyses. Here, we apply cross-species DNA capture to fossils from relatively high latitude, high altitude caves in Tasmania. Using low-stringency hybridization and high-throughput sequencing, we were able to retrieve mitochondrial sequences from two extinct megafaunal macropodid species. The two specimens, *Simosthenurus occidentalis* (giant short-faced kangaroo) and *Protemnodon anak* (giant wallaby), have been radiocarbon dated to 46–50 and 40–45 ka, respectively. This is significantly older than any Australian fossil that has previously yielded DNA sequence information. Processing the raw sequence data from these samples posed a bioinformatic challenge due to the poor preservation of DNA. We explored several approaches in order to maximize the signal-to-noise ratio in retained sequencing reads. Our findings demonstrate the critical importance of adopting stringent processing criteria when distant outgroups are used as references for mapping highly fragmented DNA. Based on the most stringent nucleotide data sets (879 bp for *S. occidentalis* and 2,383 bp for *P. anak*), total-evidence phylogenetic analyses confirm that macropodids consist of three primary lineages: Sthenurines such as *Simosthenurus* (extinct short-faced kangaroos), the macropodines (all other wallabies and kangaroos), and the enigmatic living banded hare-wallaby *Lagostrophus fasciatus* (Lagostrophinae). *Protemnodon* emerges as a close relative of *Macropus* (large living kangaroos), a position not supported by recent morphological phylogenetic analyses.

Key words: **ancient DNA, phylogenetics, Sthenurinae, Lagostrophinae, Macropodinae.**

Introduction

The Late Pleistocene was marked by the extinction of many large terrestrial vertebrates (megafauna) around the world. One of the most remarkable and least understood of these extinction events occurred in Australia, where diverse marsupial, avian, and reptile megafauna dominated a uniquely isolated continent. Further, the relative antiquity of the Australian extinctions (39–52 ka) (Roberts et al. 2001; Price et al. 2011; Gillespie et al. 2012) compared with Late Pleistocene extinctions on other continents (ca. 10-30 ka) (Guthrie 2006) has meant that their drivers remain controversial. Among the many species lost during the Late Pleistocene were enigmatic animals such as the huge "marsupial rhinoceros" *Diprotodon*, the "marsupial lion" *Thylacoleo*, the giant short-faced sthenurine kangaroos, the 5-m-long monitor lizard *Varanus priscus*, and the 200 kg flightless bird *Genyornis*.

Among Australia's extinct megafauna, the short-faced browsing sthenurine kangaroos are a group of particular interest because their phylogenetic relationship to extant macropodines (kangaroos, wallabies, and relatives within the family Macropodidae) is unclear. The Sthenurinae, with 6 genera and 26 species described to date (Prideaux 2004), were an ecologically distinct Miocene radiation that parallel extant macropodines. Some studies of craniodental morphology suggest that the endangered banded hare-wallaby (*Lagostrophus fasciatus*), currently restricted to a relictual

distribution on three small islands off Western Australia, might represent the sole remaining extant sthenurine lineage (Flannery 1983, 1989; Murray 1995). However, more recent morphological studies placed the banded hare-wallaby as an isolated lineage either sister to all modern macropodines, to the exclusion of sthenurines (Prideaux 2004), or sister to a sthenurine/macropodine clade (Prideaux and Warburton 2010; Prideaux and Tedford 2012). Genetic studies have confirmed that *L. fasciatus* is indeed the sole living representative of an ancient kangaroo lineage (Westerman et al. 2002), but cannot determine how it is related to sthenurines due to the lack of molecular data for the latter, or indeed any Late Pleistocene extinct Australian megafaunal taxa.

The problems encountered in determining the phylogenetic relationships of the sthenurine kangaroos are typical of Australia's extinct megafauna. Firstly, the pre-Pleistocene Australian fossil record is quite poor compared with other continents (Archer et al. 1999). Secondly, many extinct Australian marsupial megafauna were morphologically highly divergent, evolving into extreme forms unlike any living species and thereby hindering phylogenetic inference. Finally, unlike recently extinct megafauna from other continents, the antiquity and taphonomy of Australian megafaunal fossil deposits have usually precluded the retrieval of ancient DNA (aDNA). Molecular data have been instrumental in resolving the phylogenetic relationships of many living and extinct species (Bunce et al. 2009; Green et al. 2010; Reich et al. 2010; Rohland et al. 2010; Meyer et al. 2012; Miller et al. 2012, 2013; Mitchell, Llamas, et al. 2014; Mitchell, Wood, et al. 2014), and have the potential to elucidate phylogenetic relationships when morphological data are equivocal or misleading (Springer et al. 2007; Lee and Camens 2009). Although great strides have been made in the molecular analysis of ancient remains from high altitudes/latitudes (where low temperatures permit the survival of aDNA), retrieval of DNA from lower altitudes/latitudes (the usual situation in Australia) remains problematic due to the tight correlation between high temperatures and elevated rates of DNA decay (Smith et al. 2003; Allentoft et al. 2012). As a result, Late Pleistocene Australian megafaunal fossils represent a challenge for current aDNA methodologies and there have been relatively few successful studies: The isolation of emu DNA from eggshells (19 ka) (Oskam et al. 2010); plant and murid DNA from a midden (30.5 ka) (Murray et al. 2012); and bird, reptile, and mammal DNA from bulk extracts of highly fragmented bones (4.3–45.6 uncalibrated ^{14}C ka) (Murray et al. 2013). In the latter study, putative macropodid sequences (<36 uncalibrated ^{14}C ka) were identified that may have been contributed by extinct megafaunal taxa. However, a limitation of bulk DNA extractions is that unequivocal identification of the remains is difficult, as individual sequences are not directly associated with individual fossils.

Recent studies have described methods for selectively capturing and sequencing short, low-concentration endogenous DNA fragments using primer extension (Briggs et al. 2009; Krause et al. 2010), microarrays (Burbano et al. 2010), in-solution molecular baits (Avila-Arcos et al. 2011; Fu, Meyer, et al. 2013; Mitchell, Llamas, et al. 2014; Mitchell, Wood, et al. 2014), and custom polymerase chain reaction (PCR) products (Maricic et al. 2010; Sanchez-Quinto et al. 2012; Brotherton et al. 2013; Fu, Mittnik, et al. 2013; Der Sarkissian et al. 2014). Although these methods appear well suited for Australian megafaunal material, a current limitation for studying extinct species is the need for molecular information from a close phylogenetic relative to design primers or baits for hybridization enrichment. This is a considerable problem in the study of Australian megafauna, as many recently extinct forms are only distantly related to their closest living relatives (e.g., sthenurine kangaroos diverged from kangaroos and wallabies prior to the Middle Miocene approximately 16 Ma [Prideaux and Warburton 2010; Prideaux and Tedford 2012]). However, recent molecular hybridization studies have shown that DNA samples can be successfully enriched using molecular baits designed from quite divergent species (Mason et al. 2011; Li et al. 2013; Mitchell, Llamas, et al. 2014; Mitchell, Wood, et al. 2014). This suggests that hybridization capture might be a viable method for studying the genetics of Australian megafauna.

Here, we report the use of multispecies DNA hybridization capture to characterize partial mitochondrial genomes (879 and 2,383 bp, respectively) from two extinct Australian megafaunal taxa found at Mt Cripps, Tasmania: *Simosthenurus occidentalis* (giant short-faced kangaroo, hereafter referred to as *Simosthenurus*) and *Protemnodon anak* (giant wallaby, hereafter *Protemnodon*). Radiocarbon dating of the *Simosthenurus* and *Protemnodon* material (46–50 and 40–45 ka, respectively) (Gillespie et al. 2012) indicates that these samples are significantly older than any extinct Australian megafaunal remains that have previously yielded DNA. We combine our new DNA sequence data with an existing morphological character matrix to create a total-evidence phylogeny that clarifies the phylogenetic position of these enigmatic megafauna.

Results

aDNA Analysis

Biochemical analyses of the specimens during radiocarbon dating analyses suggested poor preservation of organic material, with very low nitrogen content (0.44–0.45%). As a consequence, only one of three attempts to radiocarbon-date the *Simosthenurus* sample was successful (reference OxA-17143) (Gillespie et al. 2012). Preliminary PCR tests for the presence of marsupial DNA indicated that the DNA was highly fragmented (supplementary note S2, Supplementary Material online), meaning that a phylogenetic study would be unlikely to be successful using a PCR approach. Although molecular hybridization followed by high-throughput Next Generation Sequencing (NGS) seemed appropriate, the lack of a close relative complicated bait design. Osteological analyses suggest that both *Simosthenurus* and *Protemnodon* are part of the same family as modern kangaroos (Macropodidae) (Prideaux and Warburton 2010), so we designed hybridization baits from five divergent macropodid taxa (asterisked in fig. 1A) to enrich the DNA extracts for macropodid DNA prior to

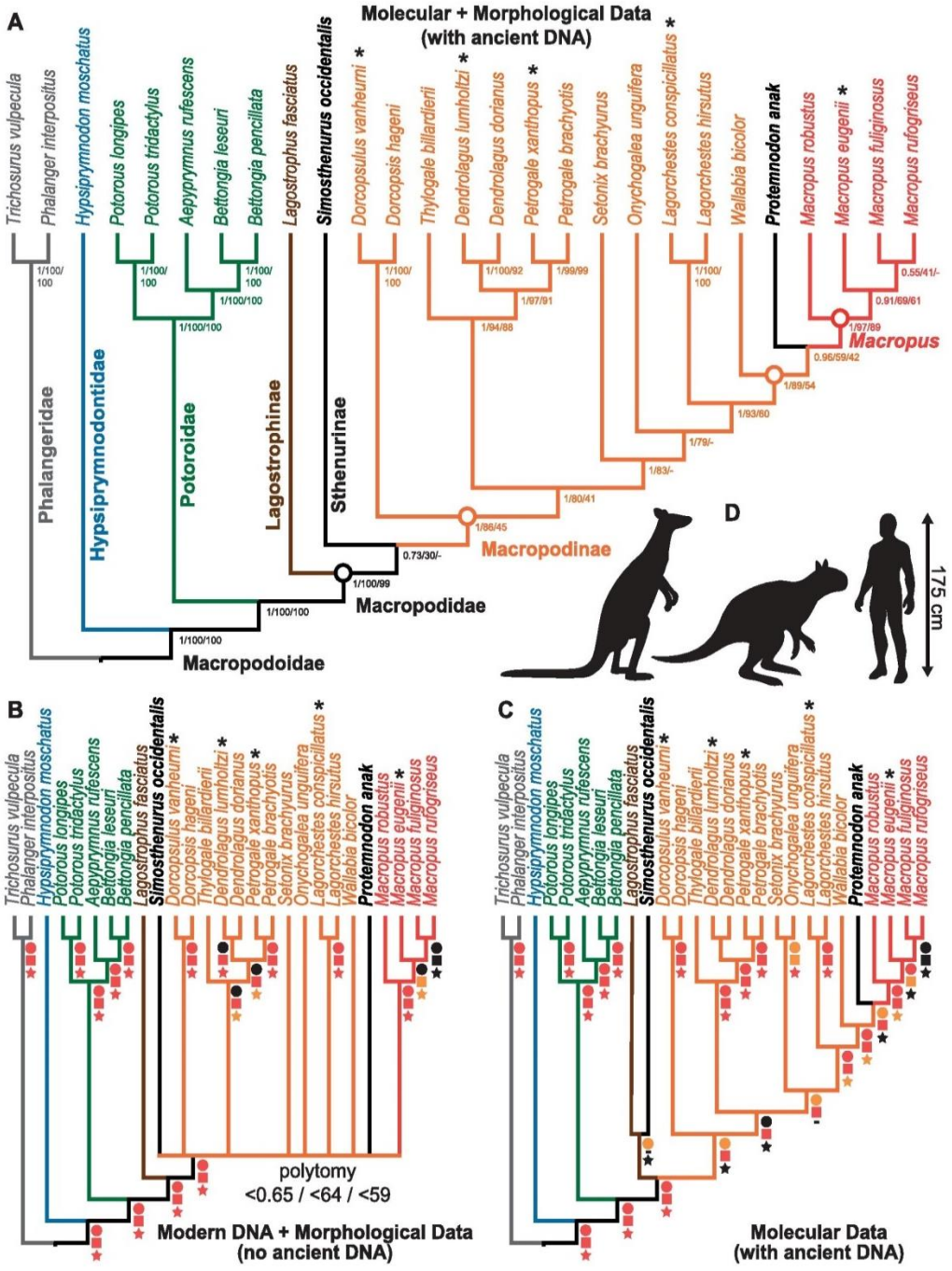


FIG. 1. Phylogenetic relationships of Australian macropodoids (including extinct megafauna) using molecular and morphological data. (*A*) Phylogenetic analyses using all molecular (including aDNA) and morphological data. Taxa used for bait design are marked with an asterisk. Clade support is shown as PP/Maximum Likelihood Bootstrap (ML)/Maximum Parsimony Bootstrap (MP). The Bayesian consensus topology is shown; The "-" sign denotes clade not found in ML and/or MP trees. Open circles denote robustly supported clades that tightly constrain affinities of extinct taxa. (*B*) Phylogenetic analyses using only modern molecular data (excluding aDNA) and morphological data for all taxa, including the extinct megafauna. Support values are given as follows: Filled circles are for PP, filled squares are for ML, and filled stars are for MP; Red shows robust support (>0.9/>70/>70), orange shows weak support (>0.8/>50/>50), and black shows little or no support (<0.8/<50/<50). (*C*) Phylogenetic analyses using only molecular data (including aDNA). Symbols and colors as in (*B*). (*D*) Adult size of *Protemnodon anak* (left silhouette) and *Simosthenurus occidentalis* (middle silhouette) relative to a 175-cm-tall human.

sequencing (see Materials and Methods and supplementary note S3, Supplementary Material online).

We obtained 5,723,580 raw sequencing reads for *Protemnodon* (Ion Torrent PGM only) and 1,039,524 for *Simosthenurus* (Ion Torrent PGM: 678,102; Roche 454 GS-FLX: 286,711; Pacific Biosciences RS: 74,711). Reads were trimmed for residual adapter sequences and low-quality sequences. Reads shorter than 25 nucleotides were filtered out as part of the trimming process (supplementary note S4, Supplementary Material online). The resulting data set had an average read size of 37.1 ± 8.9 nt for *Protemnodon* and 40.8 ± 15.8 nt for *Simosthenurus* (PGM: 47.6 ± 21.9 nt; 454 GS-FLX: 38.6 ± 10.2 nt; RS: 41.6 ± 12.1 nt). The overall short length of the reads confirmed the highly degraded nature of the DNA. Moreover, the taxonomic distribution of the filtered reads showed the large majority of reads were either assigned to prokaryotes or could not be assigned to any known taxonomic group, suggesting a high level of background, exogenous DNA (supplementary fig. S3, Supplementary Material online).

To retrieve the endogenous reads belonging to *Simosthenurus* and *Protemnodon*, we used the short reads mapper BWA (Li and Durbin 2009) following published guidelines for aDNA analysis (Orlando et al. 2011; Kircher 2012; Schubert et al. 2012). As no reference data were available for *Simosthenurus* and *Protemnodon*, we used the five mitochondrial genomes used to design molecular baits as mapping references. We explored the parameter space for mismatches and gap openings to account for the resulting phylogenetic distance between the references and the target organism (supplementary note S4, Supplementary Material online). Analyses using the most stringent parameters resulted in 58 mapped reads for *Simosthenurus* and 161 reads for *Protemnodon* (supplementary table S5, Supplementary Material online). Neither data set contained reads of likely human origin (Kircher 2012). Progressively relaxing the mapping parameters increased the number of mapped reads (up to an order of magnitude), but this was accompanied by the incorporation of a small proportion of potential human DNA contaminants (0.6–3.2%) (supplementary table S5, Supplementary Material online).

Monitoring for likely human reads does not preclude the presence of other contaminating environmental DNA sequences, especially when using Ion Torrent data, which include frequent indel errors (Loman et al. 2012; Quail et al. 2012; Bragg et al. 2013), and when mapping is performed using phylogenetically distant reference genomes. Thus, we further refined the most stringent mapping analysis using BLASTN and discarded all reads that did not readily align to a known marsupial sequence (see supplementary note S4, Supplementary Material online). The final *Simosthenurus* data set included 37 unique reads covering 879 bp of the mitochondrial genome, whereas the final *Protemnodon* data set included 121 unique reads covering 2,383 bp of the mitochondrial genome. Unfortunately, the low number of reads did not allow us to assess the DNA damage pattern characteristic of aDNA reads, that is, accumulation of 5′ C-to-T misincorporations, and presence of purines at the position immediately before the start and after the end of the reads (Briggs et al. 2007; Orlando et al. 2011; Sawyer et al. 2012).

Phylogenetic Analyses

We added our final aDNA sequence data set to a data matrix comprising mitochondrial genomes from 23 extant macropodoids, and two phalangerid possum outgroups (supplementary table S2, Supplementary Material online). We augmented this nucleotide matrix with morphological data (83 characters) obtained from a previous study (Prideaux and Warburton 2010; Prideaux and Tedford 2012). To examine the effect of our new aDNA sequences on phylogenetic resolution among macropodids, we considered three separate data sets: 1) All morphological and all nucleotide data including aDNA (total-evidence), 2) all morphological data plus nucleotide data from extant taxa only (no aDNA), and 3) all nucleotide data without any morphological characters. Bayesian, Maximum Likelihood (ML), and Maximum Parsimony analyses were performed on each data set (see supplementary note S6, Supplementary Material online). The discussion below focuses on the Bayesian and likelihood results, but parsimony also gave consistent trees although sometimes with weaker support.

The total-evidence analyses, which used all morphological and molecular (including aDNA) data, resolved the positions of both *Simosthenurus* and *Protemnodon* (fig. 1A). *Simosthenurus* was robustly placed within the family Macropodidae (Bayesian Posterior Probability [PP] = 1, ML bootstrap = 100%) but outside of the subfamily Macropodinae (PP = 1, ML = 86%). However, the branching order between Lagostrophinae, Sthenurinae, and Macropodinae could not be robustly resolved. We attempted to improve topological support for the branching among Lagostrophinae, Macropodinae, and Sthenurinae by performing a Bayesian clock-based analysis. The resulting phylogeny was largely consistent with the results of undated analyses (supplementary fig. S12, Supplementary Material online). Node age estimates were extremely wide due to difficulties in accurately constraining the age of the root caused by gaps in the Australian stratigraphic record. Consequently, we focused only on undated analyses. In a further attempt to improve phylogenetic resolution, we repeated the dated Bayesian analysis with the addition of morphological characters for four fossil taxa (*Wanburoo*, *Dorcopsoides*, *Hadronomas*, and *Ngamaroo*) from Prideaux and Warburton (2010). The resulting tree was largely consistent with the undated analyses of the core taxa (supplementary fig. S13, Supplementary Material online). The addition of two fossil sthenurine taxa (*Wanburoo* and *Hadronomas*) did not improve topological support for the branching order among the three macropodid subfamilies. Consequently, in all further analyses we excluded taxa for which no molecular data were available.

Protemnodon formed a robust clade with *Wallabia* and *Macropus* (PP = 1, ML = 89%), and within this clade had moderate support as sister to the sampled *Macropus* (PP = 0.96, ML = 59%). The wider phylogenetic relationships of extant taxa were highly consistent with recent trees based on

larger combined mitochondrial and nuclear data sets (e.g., Meredith et al. 2009a; Mitchell, Pratt, et al. 2014). In contrast to some studies (Meredith et al. 2009b), our results support the monophyly of *Macropus* with respect to *Wallabia* (Phillips et al. 2013).

When the aDNA data were excluded from our analyses, resolution was much poorer. Support for both *Simosthenurus* and *Protemnodon* falling within Macropodidae remained (fig. 1*B*), but their positions within Macropodidae were unresolved, with the basal lineages of the family forming a polytomy. Using only nucleotide data (including aDNA) produced a well-resolved tree similar to the total-evidence analysis (fig. 1*C*). *Simosthenurus* was again placed within Macropodidae but outside Macropodinae, and *Protemnodon* was again placed in a clade with *Wallabia* and *Macropus* as a sister lineage to sampled *Macropus* species. However, support for these groupings was generally weaker than in the total-evidence analysis.

Impact of Sequence Data Processing on Phylogenetic Results

The bulk of the sequencing data was generated using an Ion Torrent PGM, a platform known to generate homopolymer-associated indel errors (Loman et al. 2012; Quail et al. 2012; Bragg et al. 2013). The short reads mapper TMAP (https://github.com/nh13/TMAP) has been optimized for Ion Torrent data and previous studies have demonstrated that it can successfully reconstruct the mitochondrial genomes of extinct organisms de novo (Mitchell, Llamas, et al. 2014; Mitchell, Wood, et al. 2014) using an iterative mapping approach (Green et al. 2008; Hahn et al. 2013). In this approach, the information from newly mapped reads is used as a new reference to seed the mapping of reads in more divergent regions during the next mapping iteration. Iterative mapping continues by growing the assembly from existing seeds, until either the mitochondrial genome has been completed or the number of reads added stops increasing.

We used TMAP to perform iterative mapping on the *Protemnodon* and *Simosthenurus* data sets. As expected, we observed an increasing number of reads mapped following each iteration (fig. 2), leading to an increased coverage of the mitochondrial DNA (mtDNA) but still failing to reconstruct complete mitogenomes (supplementary table S6, Supplementary Material online). Misincorporation patterns and purine frequency before the start of the reads were as expected for aDNA, although there was a relatively elevated background noise due to the relatively low number of reads (supplementary figs. S4–S7, Supplementary Material online).

We investigated the taxonomic distribution of the reads incorporated during the iterative mapping by aligning reads from each iteration against the GenBank nucleotide database using BLASTN. No prokaryote reads were detected after the first iteration of mapping. However, we observed an increase in the number of nonmarsupial reads after each subsequent mapping iteration, whereas the number of marsupial reads remained stable after the second iteration (line graphs in fig. 2; supplementary tables S8 and S9, Supplementary Material online). Proportionally, only reads of prokaryote origin increased significantly over successive iterations (bar graphs in fig. 2 and supplementary figs. S9 and S10, Supplementary Material online).

We further explored the impact of the increased incorporation of contaminant reads during the mapping iterations on the phylogenetic analyses. The alignments obtained after the initial and the last iterations were used for ML and Bayesian phylogenetic analyses as described above and in supplementary note S6, Supplementary Material online, using the combined morphological and molecular data sets (supplementary fig. S8, Supplementary Material online). The topologies were strongly supported and remained unchanged from figure 1. However, the terminal branch length for both extinct marsupials increased significantly between the initial and the final mapping iterations (supplementary table S7, Supplementary Material online).

Discussion

Phylogenetic Affinities of Extinct Australian Megafauna

Homoplasy in macropodid morphology means that analyses of morphological data alone have been unable to conclusively resolve the phylogenetic relationships of *Simosthenurus* and *Protemnodon* (Flannery 1989). The addition of ancient mtDNA provides vital independent evidence for the precise affinities of these taxa, although the limited information contained in the short sequences means that the use of both molecular and morphological data resulted in stronger results than molecular data alone. The new aDNA substantially revised the phylogenetic position of *Protemnodon*, and confirmed the position of sthenurines as a distinct lineage, separate from living kangaroos and wallabies, within Macropodidae. *Protemnodon* has long been of contentious and unstable affinities (Prideaux and Warburton 2010). The most recent morphological analysis placed *Protemnodon* as a sister member of the Macropodinae (outside of a clade comprising *Lagorchestes*, *Onychogalea*, *Wallabia*, and *Macropus* [Prideaux and Warburton 2010; Prideaux and Tedford 2012]), whereas the molecular and combined data suggest that *Protemnodon* is more closely related to *Macropus*. Sthenurine kangaroos (as represented by *Simosthenurus*) are identified as the third major lineage of the Macropodidae, along with macropodines and lagostrophines. This is supported by both analyses of the molecular data alone (fig. 1*C*) and recent morphological studies (Prideaux and Warburton 2010; Prideaux and Tedford 2012), strengthening support for the results of our total-evidence analyses (fig. 1*A*). Although relationships between sthenurines, macropodines, and lagostrophines are not robustly resolved in our analyses, the sthenurine–macropodine clade found in the Bayesian and likelihood analyses (fig. 1*A*) is consistent with certain dental and upper appendicular characters, and an inferred increased emphasis on bipedal hopping (Prideaux and Warburton 2010; Prideaux and Tedford 2012)—even if recent evidence argues against hopping as default locomotion for the large Pleistocene sthenurines (Janis et al. 2014).

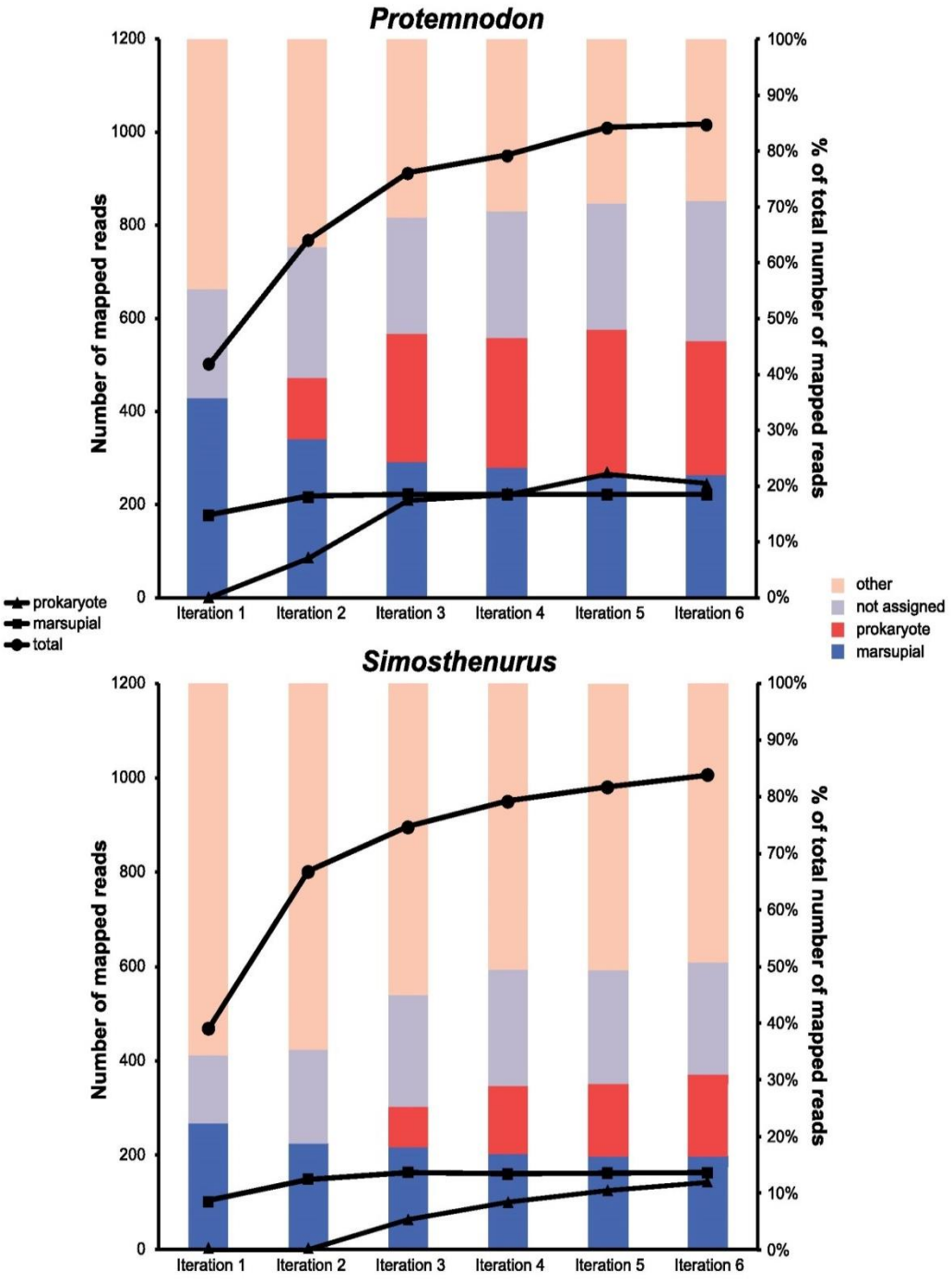


FIG. 2. Taxonomic classification of mapped reads after each mapping iteration. The line graphs (*y* axis and legend on the left) represent the number of mapped reads: Total number (filled circles), reads assigned to marsupials (filled squares), and reads assigned to prokaryotes (filled triangles). The bar graphs (*y* axis and legend on the right) represent the percentage of the total number of mapped reads assigned to the different taxonomic groups. The "not assigned" group is reads that could not be assigned to any organism for which data are available in the GenBank database ("not assigned" and "no hits" reads in supplementary tables S8 and S9, Supplementary Material online).

Limitations of Hybridization Enrichment for Analyzing Upper Pleistocene Australian Remains

Cross-species hybridization capture has previously been applied to specimens for which reference sequences from closely related outgroups are available (Mason et al. 2011; Mitchell, Llamas, et al. 2014; Mitchell, Wood, et al. 2014), but the technique is potentially efficient even across highly phylogenetically divergent taxa (Li et al. 2013). As a result, cross-species hybridization capture holds great potential for the aDNA field, where many extinct species lack close living outgroups. In this study we retrieved only a limited number of unique mtDNA reads, resulting in coverage of a small portion of the mitochondrial genome. However, this is more likely to be due to the poor preservation of DNA than phylogenetic distance: The *Simosthenurus* and *Protemnodon* specimens used in this study were collected as cave floor surface finds, and were consequently exposed to changes in the cave environment without protection from a soil matrix. The overall short length of the DNA reads retrieved for both specimens confirms that the templates have undergone substantial decay (Smith et al. 2003; Allentoft et al. 2012). In summary, Late Pleistocene Australian megafaunal fossils will continue to represent a challenge for aDNA studies due to poor preservation; however, the advent of hybridization capture has made molecular analysis of such specimens possible.

Impact of Contaminant Reads in Phylogenetic Analyses

The low endogenous DNA content in the extinct marsupial DNA extracts seemed to impair our iterative mapping approach. In the first iteration of mapping, only reads from conserved regions were mapped successfully due to the divergence between the extinct taxon and the reference (supplementary fig. S11, Supplementary Material online). A potential limitation is that these newly mapped reads may not lead to a robust consensus sequence due to the presence of artefactual substitutions typical of aDNA data. If the number of reads from authentic endogenous DNA is too low, seeds are more likely to grow through the addition of damaged or contaminant reads and the resulting degenerated consensus seed will then serve as a template for contaminant reads in the next iteration.

Phylogenetic analyses of consensus sequences obtained using multiple rounds of iterative mapping resulted in robustly supported trees with topologies identical to figure 1*A*. However, the incorporation of an increasing number of prokaryote reads during iterative mapping resulted in a substantial increase in branch length for the extinct taxa in the final mapping iteration, whereas branch length remained unchanged for all other taxa in the phylogeny.

Theoretically, contaminant sequencing reads could be mapped to any target genome if the mapping parameters were too relaxed. However, the examination of damage patterns would reveal a distribution of substitutions along the reads instead of an accumulation of 5′ C-to-T transitions. However in this case, iterative mapping resulted in the incorporation of contaminant sequences into the consensus sequence and these regions were also used as a reference during the DNA damage analyses. As a result, the contaminant reads could not be detected through standard examination of damage patterns.

The Importance of Mapping Methods for the Analysis of Phylogenetically Divergent Taxa

The iterative mapping approach was first used in the aDNA field to reconstruct a Neanderthal mitochondrial genome (Green et al. 2008). Most subsequent aDNA studies focused on taxa for which closely related reference genomes were available (i.e., early humans, extinct hominins, or horses). It is only recently that mitochondrial genomes from phylogenetically divergent extinct taxa, such as the elephant bird, have been successfully reconstructed from distantly related reference genomes using iterative mapping (Mitchell, Llamas, et al. 2014; Mitchell, Wood, et al. 2014).

The iterative mapping approach clearly reached operational limits in this study. The first two iterations for *Simosthenurus* and only the first iteration for *Protemnodon* produced data sets apparently free of bacterial DNA contaminants. Then without stringent filtering, the combination of low sequence coverage, DNA damage, and presence of contaminant environmental DNA rapidly led to the reconstruction of a chimeric consensus containing bacterial sequences, although this still produced a similar resulting tree. However, iterative mapping is likely to be efficient with molecular data sets where preservation is reasonably good (Green et al. 2008; Mitchell, Llamas, et al. 2014; Mitchell, Wood, et al. 2014) and should be considered when studying extinct taxa. It is also likely to be relevant in the analysis of ancient pathogen genomes, where variable genomic structure and horizontal gene transfer limit the potential of direct mapping against modern references.

Conclusion

Cross-species DNA capture by hybridization combined with NGS is a promising method to decipher the evolutionary history of the extinct Australian megafauna, and other taxa that are similarly phylogenetically distinct. The analyses of partial mtDNA sequences retrieved from both *Simosthenurus* and *Protemnodon* are broadly consistent with morphological data, but provide clarification of their phylogenetic positions. Sthenurines form one of the three primary lineages of the marsupial family Macropodidae, along with lagostrophines (represented today solely by the banded hare-wallaby *Lagostrophus fasciatus*) and macropodines (all other extant kangaroos and wallabies). These results support the hypothesis that the endangered banded hare wallaby (*L. fasciatus*) is the last surviving member of a distinct macropodid lineage, but is not a sthenurine. In contrast, our results suggest that *Protemnodon* represents a much more recent lineage closely related to *Macropus*, a position not predicted from morphology alone. Further aDNA analysis of enigmatic taxa such as *Thylacoleo* or *Diprotodon* will be required to fully resolve the evolutionary history of the

Australian megafauna, and cross-species DNA capture provides an important new approach for this endeavor.

Materials and Methods

All aDNA work was performed at the Australian Centre for Ancient DNA (ACAD, in Adelaide), a purpose-built laboratory dedicated to aDNA studies using established protocols for aDNA work.

Extinct Marsupial Samples

Skeletal remains of the extinct *S. occidentalis* and *P. anak* specimens were collected at Mt Cripps, Tasmania, in caves CP222 (Calcite Column Chasm) and CP213 (Bone Aven), respectively. (See supplementary note S1, Supplementary Material online, for specimen identification, dating information, sample preparation, and DNA extraction.)

aDNA Libraries Preparation

The *Simosthenurus* DNA extract (ACAD3501B) and *Protemnodon* DNA extract (ACAD9010A) were processed as described previously (Brotherton et al. 2013) to generate aDNA libraries.

Molecular Baits Design Strategy and Preparation

Although no genetic information is available for either extinct marsupial, the latest paleontological studies place *Simosthenurus* and *Protemnodon* within Macropodidae (Prideaux and Warburton 2010). Therefore, we applied a cross-species hybridization strategy to capture mtDNA fragments from the aDNA libraries using complete mitochondrial genomes from five extant macropodids: *Dendrolagus lumholtzi*, *Dorcopsulus vanheurni*, *Lagorchestes conspicillatus*, *Macropus eugenii*, and *Petrogale xanthopus* (supplementary table S2, Supplementary Material online). The preparation of molecular baits was adapted from Brotherton et al. (2013). Details can be found in supplementary note S3, Supplementary Material online.

Hybridization Capture Assay

DNA capture by hybridization was performed following Brotherton et al. (2013), using 100 ng of biotinylated DNA baits (see also supplementary fig. S1, Supplementary Material online). As in Brotherton et al., we selectively enriched for marsupial DNA of significant homology to bait sequence by performing increasingly stringent washes, which involved increasing the temperature and decreasing the salt concentration during washing rounds. The main innovation of the DNA capture technique was the use of strand-displacing polymerases to detach the target DNA fragments from the biotinylated baits to ensure that DNA baits would not contaminate the mtDNA-enriched libraries. We used DNA Polymerase I, Large (Klenow) Fragment (New England Biolabs [NEB]) for *Protemnodon* (Brotherton et al. 2013), and the *Bst* DNA Polymerase, Large Fragment (NEB) for *Simosthenurus*.

Next Generation Sequencing

The short adapters of the mtDNA-enriched *Simosthenurus* DNA library were converted by PCR (see Brotherton et al. 2013) into full-length adapters that included sequencing primers for each of the following sequencing platforms: Roche 454 GS-FLX, Ion Torrent PGM, and Pacific Biosciences RS. Likewise, the mtDNA-enriched *Protemnodon* DNA library was converted into an Ion Torrent PGM sequencing library. Sequencing was performed at Pacific Biosciences (Pacific Biosciences RS), the Australian Genome Research Facility (Roche 454 GS-FLX), and the Australian Cancer Research Foundation Cancer Genomics Facility (Ion Torrent PGM).

Mapping of NGS Reads

After adapter and quality trimming, and filtering of reads shorter than 25 nt, we aligned the filtered data sets against the GenBank nr database (March 2014) to analyze the taxonomic distribution from the filtered reads (supplementary fig. S3, Supplementary Material online). Results were visualized using Megan v5.3.5 (Huson et al. 2011). Reads were then simultaneously mapped against the five marsupial reference genomes used in the baits design using BWA v.0.5.9 (Li and Durbin 2009). As detailed in supplementary note S4, Supplementary Material online, we explored alternative parameter values affecting the mapping stringency, but ultimately retained the most stringent combination at the risk of rejecting some real endogenous sequences. We further refined the mapping analysis by using BLASTN to compare all reads against the NCBI nr database (March 2014) and subsequently discarding all reads that did not align to a known marsupial sequence in order to remove all potential contaminating DNA sequences.

We also performed an iterative mapping using TMAP and the *Dorcopsulus vanheurni* mtDNA sequence as initial reference, following the method described in Mitchell, Llamas, et al. (2014) and Mitchell, Wood, et al. (2014) (see supplementary note S5, Supplementary Material online). Iterative mapping did not incorporate more reads after six iterations for either extinct marsupial. Misincorporation patterns were assessed using MapDamage v0.3.6 (Ginolhac et al. 2011) (supplementary figs. S4–S7, Supplementary Material online). Reads from the initial and final iterations were aligned against the NCBI nr database (March 2014) using BLASTN to evaluate the presence of contaminant reads (fig. 2 and supplementary figs. S9 and S10, Supplementary Material online).

Molecular Data

Consensus sequences were generated for both *Simosthenurus* and *Protemnodon* using the mpileup command, the bcftools utilities, and the vcfutils.pl script from SAMtools v0.1.18 (Li et al. 2009). Nucleotides were called at each position covered by at least one read. Sites that received no coverage or insufficient coverage to confidently call a base were coded with IUPAC ambiguity symbols as appropriate. An alignment of 23 macropodoid and 2 phalangerid mitochondrial genomes (used as outgroup in the phylogenetic analyses; supplementary table S2, Supplementary Material online) was created

using Seaview v.4.2.12 (Gouy et al. 2010). Our consensus sequences were then included manually in the alignment, and we removed three independent single-nucleotide insertions in the ancient *Protemnodon* sequences (most probably PGM sequencing errors) that would disrupt the coding frame of protein-coding genes.

Morphological Data

We augmented our molecular data set with a morphological matrix of 83 skeletal characters taken from several recent studies (Prideaux and Warburton 2010; Prideaux and Tedford 2012). Eight species for which we included genetic data were scored as unknown in the morphological data set (*Trichosurus vulpecula*, *Phalanger interpositus*, *Ptrogale xanthopus*, *Aepyprymnus rufescens*, *Bettongia lesueur*, *Macropus robustus*, *M. rufogriseus*, and *Potorous longipes*), whereas two species were represented by congeners: *Dendrolagus bennettianus* in place of *De. lumholtzi* (McGreevy et al. 2012) and *Dorcopsis veterum* in place of *Do. hageni* (Groves and Flannery 1989). In an additional Bayesian analysis (see below), we added morphological data for four of the most complete kangaroo fossil taxa *Wanburoo*, *Dorcopsoides*, *Hadronomas*, and *Ngamaroo* (Prideaux and Warburton 2010).

Phylogenetic Analyses

The molecular data sets did not include the D-loop and consisted of eight partitions: Codon positions for the protein-coding genes (with ND6 codon positions in separate partitions), RNA stems, and RNA loops. All morphological characters were treated as unordered in all analyses, as per the original study (Prideaux and Warburton 2010). Bayesian Inference (MrBayes 3.2 [Ronquist et al. 2012]), ML (RAxML v.7.2.8 [Stamatakis 2006]), and Maximum Parsimony (PAUP* [Swofford 2002]) analyses were performed on the combined molecular and morphological data (supplementary note S6, Supplementary Material online). For the Bayesian and ML analyses, molecular substitution models and partitioning schemes were selected using the Bayesian information criterion as implemented by PartitionFinder v.1.0.1 (Lanfear et al. 2012), which favored a six-partition scheme with the selected models shown in supplementary table S4, Supplementary Material online. The morphological data formed the seventh partition. To ascertain the impact of the new aDNA data, combined analyses were performed with all morphological and molecular data, and then without aDNA (i.e., the extinct sthenurine and *Protemnodon* scored only for morphological traits, whereas all other taxa scored for molecular and—where available—morphological traits). Analyses were also performed with the molecular data alone (including aDNA).

We attempted to improve topological support by performing two additional Bayesian analyses using the combined molecular and morphological data:

i) A clock-based analysis in MrBayes using the IGR model and all morphological and molecular data (supplementary fig. S12, Supplementary Material online). Root age was modeled as a uniform distribution with a lower bound at approximately 25 Ma (the age of the oldest macropodoids) and an upper bound at 54.6 Ma (the age of the Murgon deposits from which no macropodoids have been described). This necessarily wide prior distribution means that estimated node ages are similarly uninformative.

ii) A dated Bayesian analysis (as above) with the addition of morphological characters from the most complete kangaroo fossil taxa *Wanburoo*, *Dorcopsoides*, *Hadronomas*, and *Ngamaroo* (Prideaux and Warburton 2010) (supplementary fig. S13, Supplementary Material online).

Supplementary Material

Supplementary notes S1–S6, references, figures S1–S13, and tables S1–S9 are available at *Molecular Biology and Evolution* online (http://www.mbe.oxfordjournals.org/).

Acknowledgments

The authors are grateful to R. Beck and an anonymous reviewer for their helpful comments. They thank C. Reid from the Queen Victoria Museum and Art Gallery and M. Binnie from the South Australian Museum for granting access to paleontological collections; P. Darby and L. Gray for guiding us through the Mt Cripps karst; T. Worthy, J. Wood, N. Rawlence, F. Salt, H. Geraldene, and I. Household for assistance in the field; G. Prideaux for assistance in identifying specimens; J. Korlach and T. Clark for their help with sequencing samples at Pacific Biosciences; e-research SA for high-performance computing resources. Funding for field work at Mt Cripps and 14C dating at Oxford was provided by ARC Discovery Project DP0664562 to A.C. B.L. was a postdoctoral fellow from the ARC Discovery grant DP1095782.

References

Allentoft ME, Collins M, Harker D, Haile J, Oskam CL, Hale ML, Campos PF, Samaniego JA, Gilbert MT, Willerslev E, et al. 2012. The half-life of DNA in bone: measuring decay kinetics in 158 dated fossils. *Proc R Soc Lond B Biol Sci.* 279:4724–4733.

Archer M, Arena R, Bassarova M, Black K, Brammall J, Cooke BN, Crosby K, Godthelp H, Gott M, Hand SJ, et al. 1999. The evolutionary history and diversity of Australia's mammals. *Aust Mammal.* 21:1–45.

Avila-Arcos MC, Cappellini E, Romero-Navarro JA, Wales N, Moreno-Mayar JV, Rasmussen M, Fordyce SL, Montiel R, Vielle-Calzada J-P, Willerslev E, et al. 2011. Application and comparison of large-scale solution-based DNA capture-enrichment methods on ancient DNA. *Sci Rep.* 1:74.

Bragg LM, Stone G, Butler MK, Hugenholtz P, Tyson GW. 2013. Shining a light on dark sequencing: characterising errors in Ion Torrent PGM data. *PLoS Comput Biol.* 9:e1003031.

Briggs AW, Good JM, Green RE, Krause J, Maricic T, Stenzel U, Lalueza-Fox C, Rudan P, Brajkovic D, Kucan Z, et al. 2009. Targeted retrieval and analysis of five Neandertal mtDNA genomes. *Science* 325: 318–321.

Briggs AW, Stenzel U, Johnson PLF, Green RE, Kelso J, Prufer K, Meyer M, Krause J, Ronan MT, Lachmann M, et al. 2007. Patterns of damage in genomic DNA sequences from a Neandertal. *Proc Natl Acad Sci U S A.* 104:14616–14621.

Brotherton P, Haak W, Templeton J, Brandt G, Soubrier J, Jane Adler C, Richards SM, Sarkissian CD, Ganslmeier R, Friederich S, et al. 2013. Neolithic mitochondrial haplogroup H genomes and the genetic origins of Europeans. *Nat Commun.* 4:1764.

Bunce M, Worthy TH, Phillips MJ, Holdaway RN, Willerslev E, Haile J, Shapiro B, Scofield RP, Drummond A, Kamp PJ, et al. 2009. The evolutionary history of the extinct ratite moa and New Zealand Neogene paleogeography. *Proc Natl Acad Sci U S A.* 106: 20646–20651.

Burbano HA, Hodges E, Green RE, Briggs AW, Krause J, Meyer M, Good JM, Maricic T, Johnson PLF, Xuan Z, et al. 2010. Targeted investigation of the Neandertal genome by array-based sequence capture. *Science* 328:723–725.

Der Sarkissian C, Brotherton P, Balanovsky O, Templeton JEL, Llamas B, Soubrier J, Moiseyev V, Khartanovich V, Cooper A, Haak W, et al. 2014. Mitochondrial genome sequencing in Mesolithic North East Europe unearths a new sub-clade within the broadly distributed human haplogroup C1. *PLoS One* 9:e87612.

Flannery TF. 1983. Revision in the macropodid subfamily Sthenurinae (Marsupialia: Macropodoidea) and the relationships of the species of *Troposodon* and *Lagostrophus*. *Aust Mammal.* 6:15–28.

Flannery TF. 1989. Phylogeny of the Macropodidea: a study in convergence. In: Grigg G, Jarman P, Hume I, editors. Kangaroos, wallabies and rat kangaroos. Sydney (NSW): Surrey Beatty and Sons. p. 1–46.

Fu Q, Meyer M, Gao X, Stenzel U, Burbano HA, Kelso J, Paabo S. 2013. DNA analysis of an early modern human from Tianyuan Cave, China. *Proc Natl Acad Sci U S A.* 110:2223–2227.

Fu Q, Mittnik A, Johnson PL, Bos K, Lari M, Bollongino R, Sun C, Giemsch L, Schmitz R, Burger J, et al. 2013. A revised timescale for human evolution based on ancient mitochondrial genomes. *Curr Biol.* 23: 553–559.

Gillespie R, Camens AB, Worthy TH, Rawlence NJ, Reid C, Bertuch F, Levchenko V, Cooper A. 2012. Man and megafauna in Tasmania: closing the gap. *Quat Sci Rev.* 37:38–47.

Ginolhac A, Rasmussen M, Gilbert MT, Willerslev E, Orlando L. 2011. mapDamage: testing for damage patterns in ancient DNA sequences. *Bioinformatics* 27:2153–2155.

Gouy M, Guindon S, Gascuel O. 2010. SeaView version 4: a multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Mol Biol Evol.* 27:221–224.

Green RE, Krause J, Briggs AW, Maricic T, Stenzel U, Kircher M, Patterson N, Li H, Zhai W, Fritz MHY, et al. 2010. A draft sequence of the Neandertal genome. *Science* 328:710–722.

Green RE, Malaspinas AS, Krause J, Briggs AW, Johnson PL, Uhler C, Meyer M, Good JM, Maricic T, Stenzel U, et al. 2008. A complete Neandertal mitochondrial genome sequence determined by high-throughput sequencing. *Cell* 134:416–426.

Groves CP, Flannery TF. 1989. Revision of the genus *Dorcopsis* (Macropodidae: Marsupialia). In: Grigg G, Jarman P, Hume I, editors. Kangaroos, wallabies and rat-kangaroos. Chipping Norton (NSW): Surrey Beatty & Sons Pty Ltd. p. 117–128.

Guthrie RD. 2006. New carbon dates link climatic change with human colonization and Pleistocene extinctions. *Nature* 441: 207–209.

Hahn C, Bachmann L, Chevreux B. 2013. Reconstructing mitochondrial genomes directly from genomic next-generation sequencing reads—a baiting and iterative mapping approach. *Nucleic Acids Res.* 41:e129.

Huson DH, Mitra S, Ruscheweyh H-J, Weber N, Schuster SC. 2011. Integrative analysis of environmental sequences using MEGAN4. *Genome Res.* 21:1552–1560.

Janis CM, Buttrill K, Figueirido B. 2014. Locomotion in extinct giant kangaroos: were sthenurines hop-less monsters? *PLoS One* 9: e109888.

Kircher M. 2012. Analysis of high-throughput ancient DNA sequencing data. *Methods Mol Biol.* 840:197–228.

Krause J, Briggs AW, Kircher M, Maricic T, Zwyns N, Derevianko A, Pääbo S. 2010. A complete mtDNA genome of an early modern human from Kostenki, Russia. *Curr Biol.* 20:231–236.

Lanfear R, Calcott B, Ho SY, Guindon S. 2012. Partitionfinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. *Mol Biol Evol.* 29:1695–1701.

Lee MSY, Camens AB. 2009. Strong morphological support for the molecular evolutionary tree of placental mammals. *J Evol Biol.* 22: 2243–2257.

Li C, Hofreiter M, Straube N, Corrigan S, Naylor GJ. 2013. Capturing protein-coding genes across highly divergent species. *Biotechniques* 54:321–326.

Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25:1754–1760.

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, Proc 1000 Genome Project Data. 2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25: 2078–2079.

Loman NJ, Misra RV, Dallman TJ, Constantinidou C, Gharbia SE, Wain J, Pallen MJ. 2012. Performance comparison of benchtop high-throughput sequencing platforms. *Nat Biotechnol.* 30: 434–439.

Maricic T, Whitten M, Paabo S. 2010. Multiplexed DNA sequence capture of mitochondrial genomes using PCR products. *PLoS One* 5: e14004.

Mason VC, Li G, Helgen KM, Murphy WJ. 2011. Efficient cross-species capture hybridization and next-generation sequencing of mitochondrial genomes from noninvasively sampled museum specimens. *Genome Res.* 21:1695–1704.

McGreevy TJ, Dabek L, Husband TP. 2012. Tree kangaroo molecular systematics based on partial cytochrome b sequences: are Matschie's tree kangaroo (*Dendrolagus matschiei*) and Goodfellow's tree kangaroo (*D. goodfellowi buergersi*) sister taxa? *Aust Mammal.* 34:18.

Meredith RW, Westerman M, Springer MS. 2009a. A phylogeny of Diprotodontia (Marsupialia) based on sequences for five nuclear genes. *Mol Phylogenet Evol.* 51:554–571.

Meredith RW, Westerman M, Springer MS. 2009b. A phylogeny and timescale for the living genera of kangaroos and kin (Macropodiformes:Marsupialia) based on nuclear DNA sequences. *Aust J Zool.* 56:395–410.

Meyer M, Fu Q, Aximu-Petri A, Glocke I, Nickel B, Arsuaga JL, Martinez I, Gracia A, de Castro JM, Carbonell E, et al. 2013. A mitochondrial genome sequence of a hominin from Sima de los Huesos. *Nature* 505:403–406.

Meyer M, Kircher M, Gansauge M-T, Li H, Racimo F, Mallick S, Schraiber JG, Jay F, Pruefer K, de Filippo C, et al. 2012. A high-coverage genome sequence from an archaic Denisovan individual. *Science* 338: 222–226.

Miller W, Schuster SC, Welch AJ, Ratan A, Bedoya-Reina OC, Zhao F, Kim HL, Burhans RC, Drautz DI, Wittekindt NE, et al. 2012. Polar and brown bear genomes reveal ancient admixture and demographic footprints of past climate change. *Proc Natl Acad Sci U S A.* 109: E2382–E2390.

Mitchell KJ, Llamas B, Soubrier J, Rawlence NJ, Worthy TH, Wood J, Lee MSY, Cooper A. 2014. Ancient DNA reveals elephant birds and kiwi are sister taxa and clarifies ratite bird evolution. *Science* 344: 898–900.

Mitchell KJ, Pratt RC, Watson LN, Gibb GC, Llamas B, Kasper M, Edson J, Hopwood B, Male D, Armstrong K, et al. 2014. Molecular phylogeny, biogeography, and habitat preference evolution of marsupials. *Mol Biol Evol.* 31:2322–2330.

Mitchell KJ, Wood JR, Scofield RP, Llamas B, Cooper A. 2014. Ancient mitochondrial genome reveals unsuspected taxonomic affinity of the extinct Chatham duck (*Pachyanas chathamica*) and resolves divergence times for New Zealand and sub-Antarctic brown teals. *Mol Phylogenet Evol.* 70:420–428.

Murray DC, Haile J, Dortch J, White NE, Haouchar D, Bellgard MI, Allcock RJ, Prideaux GJ, Bunce M. 2013. Scrapheap challenge: a novel bulk-bone metabarcoding method to investigate ancient DNA in faunal assemblages. *Sci Rep.* 3:3371.

Murray DC, Pearson SG, Fullagar R, Chase BM, Houston J, Atchison J, White NE, Bellgard MI, Clarke E, Macphail M, et al. 2012. High-throughput sequencing of ancient plant and mammal DNA preserved in herbivore middens. *Quat Sci Rev.* 58:135–145.

Murray PF. 1995. The postcranial skeleton of the Miocene kangaroo, *Hadronomas puckridgi* Woodburne (Marsupialia, Macropodidae). *Alcheringa* 19:119–170.

Orlando L, Ginolhac A, Raghavan M, Vilstrup J, Rasmussen M, Magnussen K, Steinmann K, Kapranov P, Thompson JF, Zazula G, et al. 2011. True single-molecule DNA sequencing of a Pleistocene horse bone. *Genome Res.* 21:1705–1719.

Oskam CL, Haile J, Mclay E, Rigby P, Allentoft ME, Olsen ME, Bengtsson C, Miller GH, Schwenninger JL, Jacomb C, et al. 2010. Fossil avian eggshell preserves ancient DNA. *Proc R Soc Lond B Biol Sci.* 277: 1991–2000.

Phillips MJ, Haouchar D, Pratt RC, Gibb GC, Bunce M. 2013. Inferring kangaroo phylogeny from incongruent nuclear and mitochondrial genes. *PLoS One* 8:e57745.

Price GJ, Webb GE, Zhao J, Feng Y, Murray AS, Cooke BN, Hocknull SA, Sobbe IH. 2011. Dating megafaunal extinction on the Pleistocene Darling Downs, eastern Australia: the promise and pitfalls of dating as a test of extinction hypotheses. *Quat Sci Rev.* 30:899–914.

Prideaux GJ. 2004. Systematics and evolution of the sthenurine kangaroos. Berkeley (CA): University of California Press.

Prideaux GJ, Tedford RH. 2012. *Tjukuru wellsi*, gen. et sp. nov., a lagostrophine kangaroo (Diprotodontia, Macropodidae) from the Pliocene (Tirarian) of northern South Australia. *J Vertebr Paleontol.* 32:717–721.

Prideaux GJ, Warburton NM. 2010. An osteology-based appraisal of the phylogeny and evolution of kangaroos and wallabies (Macropodidae: Marsupialia). *Zool J Linn Soc.* 159:954–987.

Quail MA, Smith M, Coupland P, Otto TD, Harris SR, Connor TR, Bertoni A, Swerdlow HP, Gu Y. 2012. A tale of three next generation sequencing platforms: comparison of Ion Torrent, Pacific Biosciences and Illumina MiSeq sequencers. *BMC Genomics* 13:341.

Reich D, Green RE, Kircher M, Krause J, Patterson N, Durand EY, Viola B, Briggs AW, Stenzel U, Johnson PLF, et al. 2010. Genetic history of an archaic hominin group from Denisova Cave in Siberia. *Nature* 468: 1053–1060.

Roberts RG, Flannery TF, Ayliffe LK, Yoshida H, Olley JM, Prideaux GJ, Laslett GM, Baynes A, Smith MA, Jones R, et al. 2001. New ages for the last Australian megafauna: Continent-wide extinction about 46,000 years ago. *Science* 292:1888–1892.

Rohland N, Reich D, Mallick S, Meyer M, Green RE, Georgiadis NJ, Roca AL, Hofreiter M. 2010. Genomic DNA sequences from mastodon and woolly mammoth reveal deep speciation of forest and savanna elephants. *PLoS Biol.* 8:e1000564.

Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Hohna S, Larget B, Liu L, Suchard MA, Huelsenbeck JP. 2012. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst Biol.* 61:539–542.

Sanchez-Quinto F, Schroeder H, Ramirez O, Avila-Arcos MC, Pybus M, Olalde I, Velazquez AM, Marcos ME, Encinas JM, Bertranpetit J, et al. 2012. Genomic affinities of two 7,000-year-old Iberian hunter-gatherers. *Curr Biol.* 22:1494–1499.

Sawyer S, Krause J, Guschanski K, Savolainen V, Paabo S. 2012. Temporal patterns of nucleotide misincorporations and DNA fragmentation in ancient DNA. *PLoS One* 7:e34131.

Schubert M, Ginolhac A, Lindgreen S, Thompson JF, Al-Rasheid KA, Willerslev E, Krogh A, Orlando L. 2012. Improving ancient DNA read mapping against modern reference genomes. *BMC Genomics* 13:178.

Smith CI, Chamberlain AT, Riley MS, Stringer C, Collins MJ. 2003. The thermal history of human fossils and the likelihood of successful DNA amplification. *J Hum Evol.* 45:203–217.

Springer MS, Burk-Herrick A, Meredith R, Eizirik E, Teeling E, O'Brien SJ, Murphy WJ. 2007. The adequacy of morphology for reconstructing the early history of placental mammals. *Syst Biol.* 56: 673–684.

Stamatakis A. 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* 22:2688–2690.

Swofford DL. 2002. PAUP*: phylogenetic analysis using parsimony (and other methods) 4.0 beta. Sunderland (MA): Sinauer Associates.

Westerman M, Burk A, Amrine-Madsen HM, Prideaux GJ, Case JA, Springer MS. 2002. Molecular evidence for the last survivor of an ancient kangaroo lineage. *J Mamm Evol.* 9:209–223.

Manuscript 5: Forensic ancestry analysis with two capillary electrophoresis ancestry informative marker (AIM) panels: Results of a collaborative EDNAP exercise.

Forensic Science International: Genetics 19 (2015) 56–67

Contents lists available at ScienceDirect

Forensic Science International: Genetics

journal homepage: www.elsevier.com/locate/fsig

Forensic ancestry analysis with two capillary electrophoresis ancestry informative marker (AIM) panels: Results of a collaborative EDNAP exercise

C. Santos[a,1], M. Fondevila[a,1], D. Ballard[b,1], R. Banemann[c], A.M. Bento[d], C. Børsting[e,1], W. Branicki[f,2], F. Brisighelli[g], M. Burrington[h], T. Capal[i], L. Chaitanya[j], R. Daniel[k], V. Decroyer[l], R. England[m], K.B. Gettings[n], T.E. Gross[o,1], C. Haas[p], J. Harteveld[q], P. Hoff-Olsen[r], A. Hoffmann[c], M. Kayser[j], P. Kohler[r,2], A. Linacre[s], M. Mayr-Eduardoff[t,1], C. McGovern[m], N. Morling[e,l,1], G. O'Donnell[h], W. Parson[t,u], V.L. Pascali[g], M.J. Porto[d], A. Roseth[r], P.M. Schneider[o,1], T. Sijen[q], V. Stenzl[i], D. Syndercombe Court[b,1], J.E. Templeton[s], M. Turanska[v], P.M. Vallone[n], R.A.H.van Oorschot[k], L. Zatkalikova[v], The EUROFORGEN-NoE Consortium, Á. Carracedo[a,1], C. Phillips[a,1,*]

[a] Forensic Genetics Unit, Institute of Forensic Sciences, University of Santiago de Compostela, Santiago de Compostela, Spain
[b] Department of Forensic and Analytical Science, Faculty of Life Science, King's College London, UK
[c] Federal Criminal Police Office, Wiesbaden, Germany
[d] Forensic Genetic and Biology Service, Centre Branch, National Institute of Legal Medicine and Forensic Sciences, Coimbra, Portugal
[e] Section of Forensic Genetics, Department of Forensic Medicine, Faculty of Health and Medical Sciences, University of Copenhagen, Frederik V's Vej 11, Copenhagen, Denmark
[f] Section of Forensic Genetics, Institute of Forensic Research, Kraków, Poland
[g] Forensic Genetics Laboratory, Institute of Legal Medicine, Università Cattolica del Sacro Cuore, Rome, Italy
[h] Forensic Science Laboratory, Dublin, Ireland
[i] Department of Forensic Genetics, Institute of Criminalistics, Prague, Czech Republic
[j] Department of Forensic Molecular Biology, Erasmus MC University Medical Centre Rotterdam, Rotterdam, The Netherlands
[k] Office of the Chief Forensic Scientist, Forensic Services Department, Victoria Police, Australia
[l] National Institute of Criminalistics and Criminology, Chaussée de Vilvoorde 100, Brussels, Belgium
[m] ESR, Private Bag 92021, Auckland, New Zealand
[n] Material Measurement Laboratory, National Institute of Standards and Technology, Gaithersburg, MD, USA
[o] Institute of Legal Medicine, Faculty of Medicine, University of Cologne, Cologne, Germany
[p] Zurich Institute of Forensic Medicine, University of Zurich, Zurich, Switzerland
[q] Department of Human Biological Traces, Netherlands Forensic Institute, The Hague, The Netherlands
[r] Department of Forensic Biology, Norwegian Institute of Public Health, Oslo, Norway
[s] School of Biological Sciences, Flinders University, Adelaide, South Australia 5042, Australia
[t] Institute of Legal Medicine, Medical University of Innsbruck, Innsbruck, Austria
[u] Forensic Science Program, The Pennsylvania State University, University Park, PA, USA
[v] Institute of Forensic Science, Ministry of the Interior, Department of Biology and DNA Analysis, Slovenská Lupca, Slovakia

ARTICLE INFO

Article history:
Received 21 January 2015
Received in revised form 6 May 2015
Accepted 9 June 2015

Keywords:
Ancestry
SNPs

ABSTRACT

There is increasing interest in forensic ancestry tests, which are part of a growing number of DNA analyses that can enhance routine profiling by obtaining additional genetic information about unidentified DNA donors. Nearly all ancestry tests use single nucleotide polymorphisms (SNPs), but these currently rely on SNaPshot single base extension chemistry that can fail to detect mixed DNA. Insertion-deletion polymorphism (Indel) tests have been developed using dye-labeled primers that allow direct capillary electrophoresis detection of PCR products (PCR-to-CE). PCR-to-CE maintains the direct relationship between input DNA and signal strength as each marker is detected with a single dye, so mixed DNA is more reliably detected. We report the results of a collaborative inter-laboratory exercise of

* Corresponding author.
E-mail address: c.phillips@mac.com (C. Phillips).
[1] Also participated in the EUROFORGEN-NoE inter-laboratory ancestry analysis pilot exercise.
[2] Participated in the EUROFORGEN-NoE pilot exercise only.

http://dx.doi.org/10.1016/j.fsigen.2015.06.004
1872-4973/

Indels
Aims
Bayes analysis
Principal component analysis (PCA)

19 participants (15 from the EDNAP European DNA Profiling group) that assessed a 34-plex SNP test using SNaPshot and a 46-plex Indel test using PCR-to-CE. Laboratories were asked to type five samples with different ancestries and detect an additional mixed DNA sample. Statistical inference of ancestry was made by participants using the *Snipper* online Bayes analysis portal plus an optional PCA module that analyzes the genotype data alongside calculation of Bayes likelihood ratios. Exercise results indicated consistent genotyping performance from both tests, reaching a particularly high level of reliability for the Indel test. SNP genotyping gave 93.5% concordance (compared to the organizing laboratory's data) that rose to 97.3% excluding one laboratory with a large number of miscalled genotypes. Indel genotyping gave a higher concordance rate of 99.8% and a reduced no-call rate compared to SNP analysis. All participants detected the mixture from their Indel peak height data and successfully assigned the correct ancestry to the other samples using *Snipper*, with the exception of one laboratory with SNP miscalls that incorrectly assigned ancestry of two samples and did not obtain informative likelihood ratios for a third. Therefore, successful ancestry assignments were achieved by participants in 92 of 95 *Snipper* analyses. This exercise demonstrates that ancestry inference tests based on binary marker sets can be readily adopted by laboratories that already have well-established CE regimes in place. The Indel test proved to be easy to use and allowed all exercise participants to detect the DNA mixture as well as achieving complete and concordant profiles in nearly all cases. Lastly, two participants successfully ran parallel next-generation sequencing analyses (each using different systems) and achieved high levels of genotyping concordance using the exercise PCR primer mixes unmodified.

1. Introduction

DNA-based forensic ancestry tests have the capacity to provide key information about unidentified DNA donors, which can be particularly useful when police investigators do not have reliable eyewitness descriptions or if the STR profiling data fails to give a DNA database match [1]. Therefore, tests for the inference of ancestry can be grouped alongside forensic DNA phenotyping (FDP) tests such as *HIrisPlex* [2] in a growing array of new technologies that have the potential to take forensic DNA analysis well beyond simple identification [3,4]. For such tests to be effective in routine forensic use they must be sensitive; easy to run using validated DNA detection instruments; and, being mainly composed of binary loci, they should have a reasonably robust way to detect mixed DNA so that apparent heterozygotes are not mistyped. In addition, the genetic data obtained must be easy to interpret. Ideally, it should be straightforward to use the genotypes to calculate a set of Bayes likelihoods for particular ancestries (or phenotypes) in comparison to reference populations whose patterns of genetic variation are already well defined. Although STRs can provide a degree of ancestry information [5,6] and Y-chromosome/mtDNA variation is highly differentiated geographically, there are widely discussed reasons why stand-alone autosomal SNP tests provide more reliable indications of a person's ancestry [7–9].

For the last ten years, forensic SNP genotyping has relied on the SNaPshot single base extension system to create relatively large-scale PCR and extension multiplexes followed by capillary electrophoresis (CE) of the dye-labeled products using standard run conditions. In this way, FDP and ancestry analysis tests [2,8–12] have been developed using single-tube amplification reactions that are highly sensitive and use validated CE regimes [10,13]. One drawback of SNP genotyping with SNaPshot is the inability to distinguish the highly skewed heterozygote peaks often seen in normal DNA with this technique, from the imbalanced peaks common to mixtures. This is mainly due to the SNaPshot terminator chemistry using dyes with much stronger blue/green fluorescence (G/A) compared to yellow/red (C/T) [14]. Therefore, despite their widespread use and evident sensitivity, forensic SNaPshot tests can be inefficient in detecting mixtures. Indel tests have been developed in recent years for identification [15–17] and ancestry analysis [18–20] detecting dye-labeled PCR products sent directly to CE from the amplification stage (PCR-to-CE). The benefits of short amplicon lengths and high levels of multiplexing that SNPs provide, are kept with Indel genotyping in this way. However, peak height ratios in heterozygotes are more balanced within any one locus than those of SNaPshot so mixed DNA is more easily detected from the resulting imbalanced patterns [17]. Two CE-based forensic ancestry tests have been established that offer complimentary characteristics: a SNaPshot assay of 34 ancestry informative marker (AIM) SNPs containing some of the most population-differentiated loci (herein 34-plex, [11]) plus a PCR-to-CE assay of 46 AIM-Indels [19] that offers comparable population differentiation to AIM-SNPs, but much greater sensitivity to mixed DNA. This report describes the use of these two assays in an inter-laboratory exercise of 15 participants from the European DNA Profiling (EDNAP) group, and 4 overseas participants, organized by the University of Santiago de Compostela (USC). As a preamble to the EDNAP exercise, the EUROFORGEN-NoE Consortium ran a similar small-scale inter-laboratory exercise to establish the test framework and gauge the transportability of the assay primer sets. As part of the Consortium's networking remit, the primer mixes used for the EDNAP exercise were purchased, optimized and packaged by USC along with test DNAs with known ancestries (undisclosed to participants). These test components are freely available in trial quantities for the forensic community to assess for themselves (available from USC upon request).

The exercise had three main goals: (i) for laboratories to assess the relative ease-of-use and reliability of the two assays by genotyping test DNAs, whenever possible, using each participant's own CE regimes; (ii) for laboratories to use the statistical ancestry inference tools developed at USC and part of the Snipper data analysis portal [11]; (iii) to assess the ability of each assay to detect mixtures by including an unmarked mixed-donor sample amongst the test DNAs. This third goal was analyzed further by assessing the Indel heterozygote peak height balance in normal DNA across the range of participant's laboratory setups, in comparison to peaks in the mixed sample. As well as the 15 European laboratories including USC, two participants were from Australia, one from New Zealand and one from the USA. All but three laboratories had participated in the preceding EDNAP *IrisPlex* exercise that applied SNaPshot analysis to the genotyping of six FDP SNPs [21]. Five EDNAP laboratories, were part of the EUROFORGEN-NoE pilot ancestry exercise.

2. Materials and methods

2.1. Primer sets, test DNA samples and assay protocols

Six quantified DNA samples (10 μl volumes at 0.5 ng/μl) plus primer mixes sufficient for 20 reactions were sent to participants who used their own PCR and SNaPshot reaction components. For the Indel assay, PCRs only required the combination of 2× Qiagen Multiplex PCR Master Mix (Qiagen, Hilden, Germany) with the primer mix and DNA. The SNaPshot PCR and extension primer sets plus the Indel PCR primer mix were prepared as previously described [11,19] and were dispatched with the DNA samples at ambient temperature. Some package transit times outside Europe exceeded one week, but the stability of both SNP and Indel primer sets had been previously assessed for the EUROFORGEN-NoE pilot exercise by carefully testing the profile quality obtained from batches of primers originally sent to the US participant and one in Australia, who were also part of the subsequent exercise.

The test DNAs were given anonymized codes and comprised five volunteer donors, each with a different continental origin of: East Asian, European, Oceanian, Native American or African ancestries. With the geographic distribution of these samples, examples of all alleles in 80 markers were observed when genotyped by USC, except SNP: rs1573020 (all A homozygotes) and Indels: rs35451359 and rs33974167 (all short-allele 'A' homozygotes) plus rs2307998 (all long-allele 'C' homozygotes).

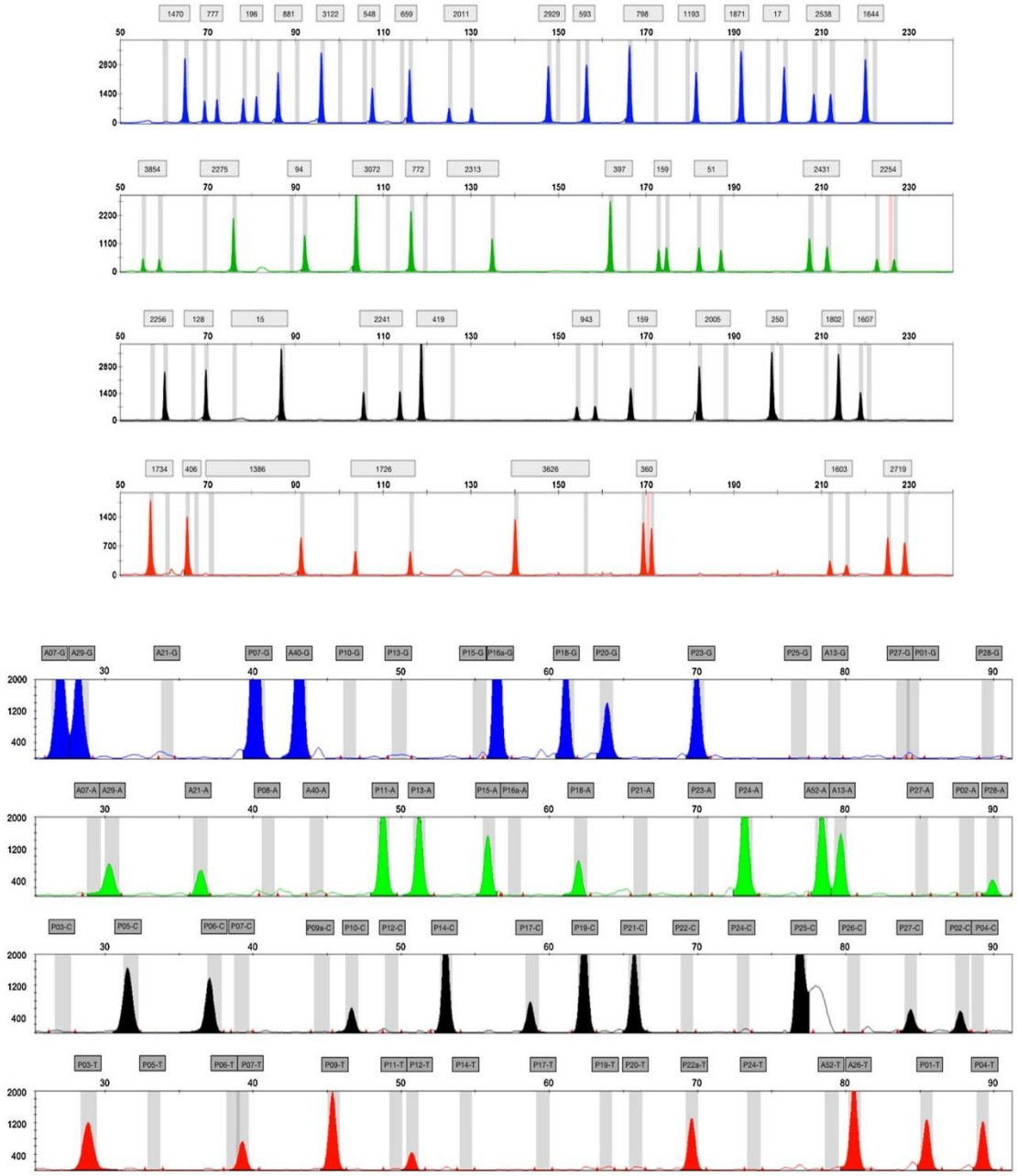


Fig. 1. Electropherograms from the Indel test (upper panel) and the 34-plex SNP test for the 9947A control DNA. Peak positions are labeled with the internal codes used for each marker (internal code-rs-number lists are provided in Fig. 3A; Supplementary Files S3; *Snipper* and in [11]).

In this way, more than 97.6% of component marker alleles could be identifiable in the profiles of the test DNAs. A rare third allele in Indel: rs25584 was found in one test DNA. The sixth test sample was an artificial mixture combining a 1:3 ratio of additional European and East Asian volunteer donors (herein M1 and M3 respectively). Note that 34-plex has two tri-allelic SNPs and one: rs5030240 showed three allele patterns in the mixed DNA sample (other examples in [11]).

Participants were told that one sample was mixed and were asked to identify it, then assign ancestries to the others using Bayes analysis and Principal Component Analysis (PCA) in *Snipper*, as detailed in Section 2.2. The above primer volumes were sufficient to allow participants to begin their analyses with the 9947A positive control DNA used in many STR kits.

Protocols for PCR, SNaPshot extension reactions and CE were sent in the form of an Excel laboratory calculator (Supplementary File S1) plus fragment mobility panels-and-bins files (Supplementary Files S2) that formed templates for participants to adapt to their own CE regimes when necessary. The 9947A DNA acted as a universal point of reference for the peak patterns typical of both assays and example electropherograms were provided to participants, as shown in Fig. 1. Although Indel amplified fragments separate well using all POP polymer types, participants were recommended to use POP-4 for 34-plex genotyping as peak positions are less well separated at the low size range using POP-7. Supplementary Table S1 lists the CE regimes chosen by participants, indicating that most applied a 3130 or 3500 detector with POP-4 (13 and 3, respectively), although two used a 3130 with POP-7 and one successfully typed SNPs with a 3100 and POP-6. Lastly, participants were advised that Indel PCR products could require dilution prior to CE to obtain optimum peak patterns free from excessive signal pull-up.

2.2. Preliminary ancestry checks of test DNAs and use of the Snipper data analysis portal

Although this section reports ancestry analysis results, these analyses were made by USC to evaluate the ancestry of the exercise test samples prior to their dispatch. This process also checked the reference population data supplied and ensured test samples were suitably representative of each of the ancestries the participants were asked to identify.

The *Snipper* portal provides a Bayes classifier accessing population reference data in place in the website, including fixed training sets for three, four or five main continental HGDP-CEPH population groups, for 34 SNPs and/or 46 Indels (these training set genotypes are provided in Supplementary File S3.2). The fixed data options assess one uploaded profile at a time, which is compared to a training set selected by the user. Partial profiles can be uploaded with NN genotypes (or partial genotypes, e.g., 'CN'). Indel data has an identical framework but with 'AC coded' genotypes, where A = short alleles, C = long and G is reserved for novel third alleles. Participants were asked to use the fixed training set option in *Snipper* to make ancestry inferences. However, no guidance was given on choice of training set, which influences calculation of the likelihood ratios (herein LRs). For example, selecting a five-group training set for 34-plex SNP data will lead to lower LRs for East Asian assignments as this marker set lacks AIMs sufficiently

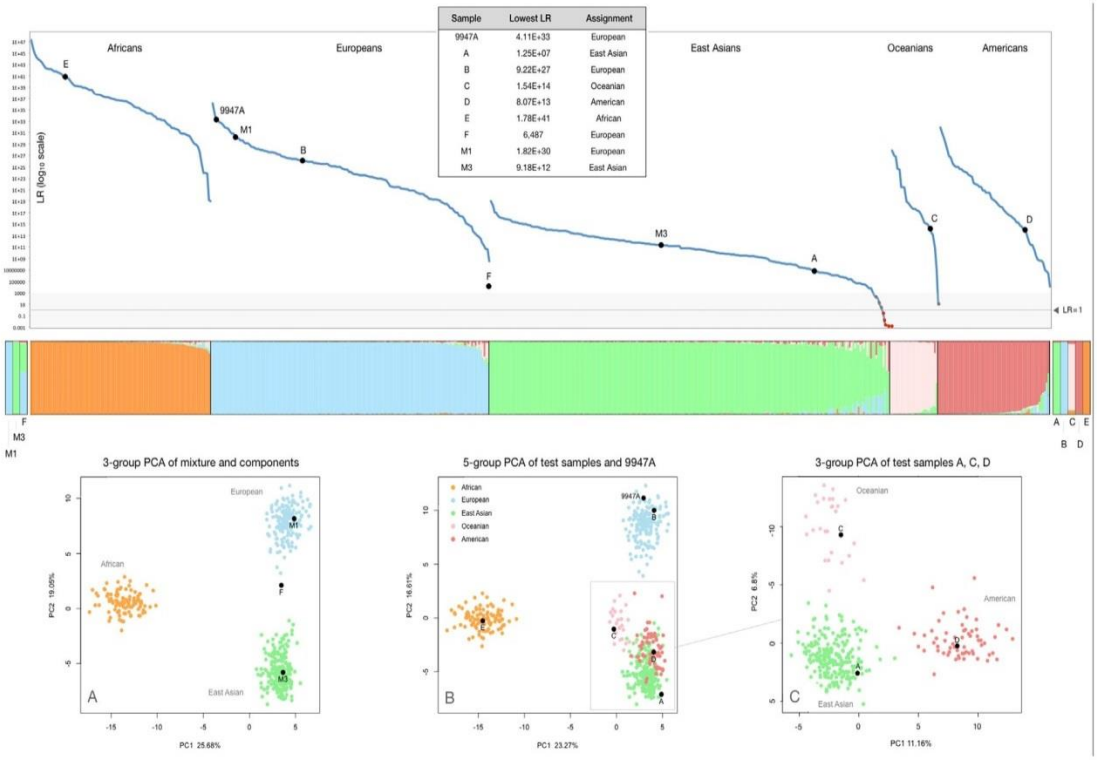


Fig. 2. Ancestry analysis of exercise test samples. 80-marker genotypes were analyzed and HGDP-CEPH training set data was as supplied to participants (Supplementary File S3.2). Top plot shows ranked *Snipper* Bayes analysis LRs from training set cross validation or test profile analysis (black points). Grey points in East Asians/Oceanians indicate LRs below a threshold value of 1000 (the grey shaded log LR range around balanced odds line of LR = 1). Red points indicate East Asian training set LRs that misclassified as Americans. Middle plots show *STRUCTURE* analysis aligned directly to the LR distributions above with separate plots for mixture components, left and test samples, right. Lower plots show 2D PCA analyses of test samples in 3-group or 5-group comparisons. Plot A shows a 3-group comparison of sample F, positioned mid-cluster between contributors M1 and M3. Plot B shows the full 5-group PCA of samples A-E plus 9947A. Plot C shows a restricted comparison of just East Asian, Oceanian and American data to obtain better differentiation of reference population clusters and samples A, C, D; all more closely distributed in plot B.

differentiated to distinguish Oceanians and Native Americans from East Asians. As a rule-of-thumb, 34-plex profiles are optimally analyzed with three-group data (Africa–Europe–East Asia), Indel profiles provide high ancestry assignment LRs for these groups plus Americans, as this differentiation was targeted in their original selection [19]. When combined 80-marker data is used, the differentiation of the fifth Oceanian population group can be accomplished, although Indel data alone can also distinguish Oceanians with minimal error [19].

To check the *Snipper* fixed training sets and test samples used, three ancestry analyses were applied to the genotype data prior to the exercise and results are summarized in Fig. 2. First, the 80-marker reference data was cross-validated with *Snipper* (each training set profile removed and classified by remaining data). Fig. 2 upper plot shows the distribution of probabilities in ranked order of $\log_{10}$ LR values, i.e., the lowest LR from five population comparisons (data in Supplementary File S3.3). The grey line of LR = 1 represents balanced odds, so points below this line show misclassifications. East Asian training set profiles gave five misclassifications, all assigned as American (5/226 = 2.2% error). However, none of their LRs exceeded 750, so applying a threshold value of 1000 led to error-free East Asian assignments, but a non-classification rate of 3.54% (8/226). Fig. 2 indicates the LRs for test samples, mixture donors and 9947A tend to fall in the middle to upper range of training set LRs in nearly all cases.

In addition to obtaining LRs, it can be helpful to compare patterns of variation in reference population data to samples of unknown geographic origin by applying *STRUCTURE* and PCA. Both provide an intuitive way to make such comparisons [3,22,23] and can be useful to alert the analyst that a forensic sample of unknown origin may be from an admixed individual with co-ancestry. Following review of the EUROFORGEN-NoE ancestry exercise results, a two-dimensional PCA module (plotting the first two principal components or PCs) was developed for *Snipper* that allows analysis of multiple profiles plotted directly onto reference data. The *Snipper* output lists the Bayes analysis data for each profile and their positions are labelled on a PC1-PC2 PCA plot (no PC3 estimates are currently made). Participants were provided with the input file of training set genotypes and a link to the *Snipper* PCA module to enable graphical analysis of test DNAs and 9947A.

The middle graphics of Fig. 2 show *STRUCTURE* cluster plots (an optimum $K = 5$ genetic clusters inferred from data) matched to the order of training set LRs charted above. The enlarged cluster plots for samples A-E on the right indicate an absence of co-ancestry, i.e., their cluster plots have almost no membership to multiple genetic clusters. Likewise, cluster plots on the left for mixture components M1 and M3 show no multiple cluster membership, whereas sample F has approximately equal joint membership to the relevant clusters. The lower graphics show three PCAs made with

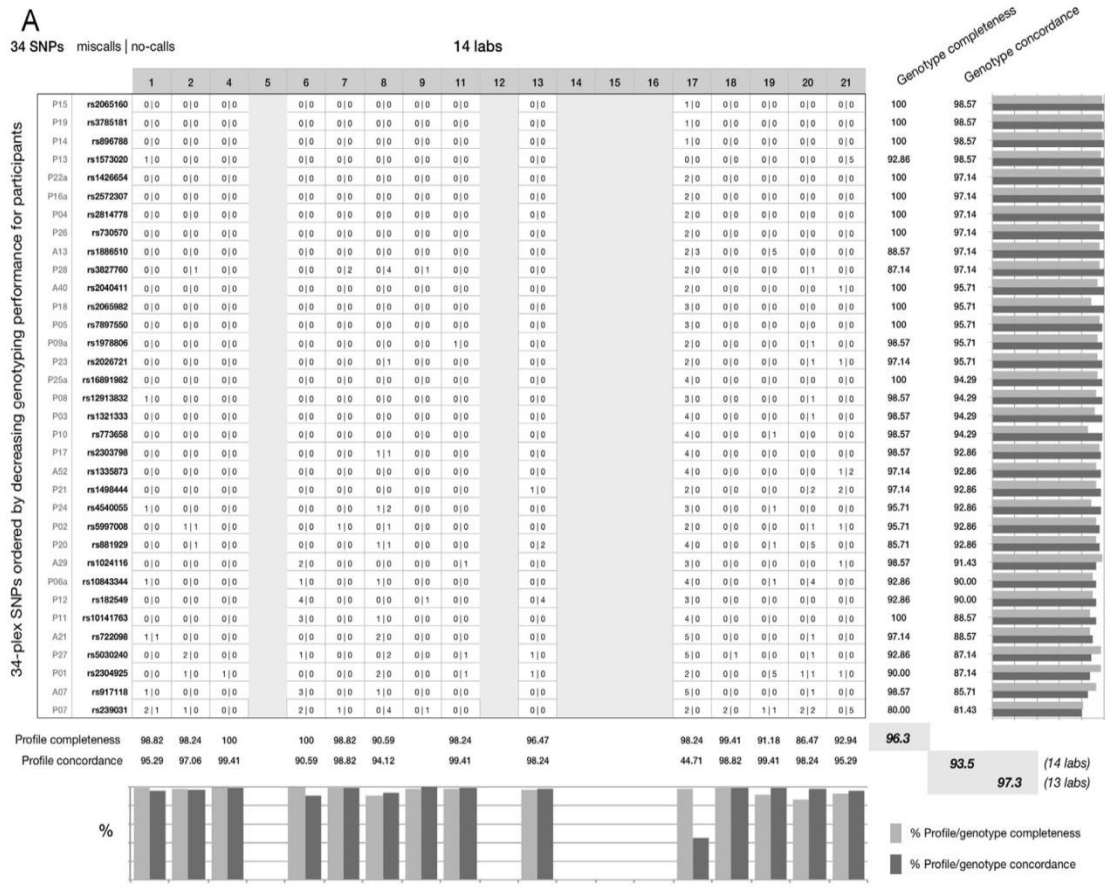


Fig. 3. (A) Genotyping performance of the 34-plex test arranged by SNP (rows) and by 14 participants (columns). Cells record miscalls on the left, and no-calls right. The bar plots on the right summarize total genotype completeness and concordance for each SNP and at the bottom, for each participant. SNPs are ordered by diminishing performance (i.e., decreasing concordance then completeness). Overall genotype concordance is given for 14 and 13 laboratories separately, excluding participant #17 with a very high number of SNP miscalls. (B) Genotyping performance grid for Indel test data from all 19 laboratories. Miscalls are shown as dark grey cells, no-calls light grey.

B

46 Indels miscalls | no-calls

19 labs

		1	2	4	5	6	7	8	9	11	12	13	14	15	16	17	18	19	20	21	Genotype completeness	Genotype concordance
MID-1470	rs2307666	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	100	100
MID-777	rs1610863	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	100	100
MID-196	rs16635	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	100	100
MID-881	rs1610965	1|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	100	98.95
MID-3122	rs35451359	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	100	100
MID-548	rs140837	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	100	100
MID-659	rs1160893	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	100	100
MID-2011	rs2308203	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	100	100
MID-2929	rs33974167	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	100	100
MID-593	rs1160852	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	100	100
MID-798	rs1610884	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|1	0|0	0|0	98.95	100
MID-1193	rs2067280	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	100	100
MID-1871	rs2308067	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|1	0|0	0|0	98.95	100
MID-17	rs4183	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	100	100
MID-2538	rs3054057	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	100	100
MID-1644	rs2307840	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	100	100
MID-3854	rs60612424	1|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	100	98.95
MID-2275	rs3033053	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	100	100
MID-94	rs16384	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	100	100
MID-3072	rs34611875	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	100	100
MID-772	rs1610859	0|1	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	98.95	100
MID-2313	rs3045215	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	100	100
MID-397	rs25621	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	100	100
MID-1636	rs2307832	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	100	100
MID-51	rs16343	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	100	100
MID-2431	rs3031979	0|1	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	98.95	100
MID-2264	rs34122827	0|0	0|0	0|0	0|0	0|0	2|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	100	97.89
MID-2256	rs133052	1|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	1|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	100	97.89
MID-128	rs6490	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	100	100
MID-15	rs4181	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	100	100
MID-2241	rs3030826	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	100	100
MID-419	rs140708	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	100	100
MID-943	rs1611026	0|0	0|0	0|0	0|0	0|0	1|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	100	98.95
MID-159	rs16438	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	100	100
MID-2005	rs2308161	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	100	100
MID-250	rs16687	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	100	100
MID-1802	rs2307998	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	100	100
MID-1607	rs2307803	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	100	100
MID-1734	rs2307930	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	100	100
MID-406	rs25630	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	100	100
MID-1386	rs2307582	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	100	100
MID-1726	rs2307922	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|3	0|0	96.84	100
MID-3626	rs11267926	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	100	100
MID-360	rs25584	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	100	100
MID-1603	rs2307799	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	100	100
MID-2719	rs34541393	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	0|0	100	100
Profile completeness		99.1	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	99.1	98.5	100	*99.8*	
Profile concordance		98.7	100	100	100	100	98.7	100	100	100	100	99.6	100	100	100	100	100	100	100	100		*99.8*

Fig. 3. (Continued)

Snipper, with reference cluster colors matched to the *STRUCTURE* data. PCA plot A is a 3-group analysis of sample F and M1-M3 components. The position of F highlights the fact that population admixture and mixed DNA can give indistinguishable PCA patterns, emphasizing the need to efficiently detect mixed DNA in forensic ancestry analysis. PCA plot B is a 5-group analysis showing samples A-E plus 9947A are distributed into their expected clusters, although in these 2-PC plots the Oceanian and American clusters show some overlap with East Asians. To better differentiate these three groups, a PCA can be made of just three possible groups to obtain a more distinct separation, as shown in PCA plot C analyzing the three test samples from less differentiated population groups.

3. Results

3.1. Genotyping performance of the SNP assay

Supplementary Table S1 summarizes the CE regimes used by participants and indicates five did not pursue SNaPshot genotyping of SNPs but elected to just analyze and report Indel genotypes. Given the complexities of reading electropherograms consisting of 32 peak pairs plus two triple-peak positions, this was considered to be a reasonable decision and Indel data alone was collected from these laboratories. The number of SNaPshot no-calls and miscalls recorded for the five test samples A-E, from 14 participants reporting SNP data, are summarized in Fig. 3A. SNPs are listed in order of decreasing genotyping performance for participants, by ranking loci in increasing miscall rate followed by increasing no-calls. Therefore, rs2065160, rs3785181 and rs8986788 are the most robustly genotyped SNPs in 34-plex, with all 14 laboratories identifying peaks in five samples, although laboratory #17 had one genotype miscall in each SNP. At the other extreme, rs239031 was both the most difficult SNP to genotype and the least reliably genotyped, with laboratories #8 and #21 not assigning genotypes to all or most samples, bringing the overall call rate down to 80%, well below those of the other 33 SNPs. Genotyping concordance for rs239031 was also the lowest, with 81.4% of genotypes correctly called. High no-call rates for certain other SNPs tended to cluster with participants: rs1573020 was not genotyped in laboratory #21; rs881929 in #20; rs1886510 and rs2304925 in #19, despite other laboratories genotyping these SNPs without problems. Only 1 of 5 genotypes was called by laboratory #13 for rs182549. Average SNP call and genotype concordance rates shown at the bottom right of Fig. 3A reached 96.3% and 93.5%, respectively. The genotype completeness of ~96% equates to approximately one missing SNP call per 34-plex profile. Laboratory #17 had evident problems recognizing and accurately calling their SNaPshot electropherograms with less than half the successful genotype calls made by the other participants reporting SNPs. Therefore, when considering concordance amongst 13/14 participants, the value rose to 97.3%. Although one other laboratory #6 had slightly

below-average genotyping concordance, no obvious connection could be made between the CE regimes used by participants and miscalls seen in certain SNPs. Nevertheless, there are known issues previously recognized at USC in some 34-plex SNPs and several of these were observed in the electropherograms from participants. Certain mobility or non-specific peak patterns can explain a proportion of the genotype miscalls and these are outlined next.

Examples of three different challenges for SNP genotyping with 34-plex are shown in Supplementary Fig. S1. First, SNPs rs10843344-rs239031 run to positions very close together, with the C peak of rs239031 often having a mobility shift that places it very close to the much higher C peak of rs10843344 (Supplementary Fig. S1.1). The same signal imbalance can be seen in the T peaks but the electrophoretic separation of these peaks remains more distinct. Examination of participant's SNaPshot profiles indicated some laboratories had missed the lower, shifted rs239031-C peak. Second, rs182549, rs881929 and rs3827760 have particularly low signal strengths (Supplementary Fig. S1.2) and the three SNPs show higher than average no-call rates. In the case of rs3827760, there is a very marked disparity in peak heights between the higher East Asian-informative G allele and the A allele (>10:1 peak height ratio in the example shown), so this SNP requires particular care. Third, rs2304925 shows an artifactual G signal in the negative control very close to the G peak of rs5030240 (Supplementary Fig. S1.3). This peak is much higher than the T peak of rs2304925 when it is a true allelic extension product but much lower when artifactual. All participants ran a negative control and most recognized the extra G signal running close to the G peak of rs5030240, although as this is a tri-allelic SNP, when a homozygous A or C allele is present the genotypes can be mistyped as an AG or CG in the absence of the stronger G peak with which to compare the artifact signal.

3.2. Genotyping performance of the Indel assay

All 19 participants successfully completed the genotyping of the samples with Indels. Supplementary Table S1 shows that almost half of the laboratories chose to dilute the PCR products 1:5–1:20 prior to CE detection to control signal pull up. Supplementary Figs. S1.4 and 5 show two examples of minor challenges with genotyping of Indels, consisting of the occurrence of dye blobs (broad non-specific peaks around allele peak positions), identifiable in the negative control, plus signal pull-ups that can occur when the Indel PCR products are not sufficiently diluted. However, there was no evidence that these two profile phenomena interfered with the genotyping performance of the Indel tests in any of the 19 laboratories. In fact, the genotyping completeness and concordance were very high when considering that most

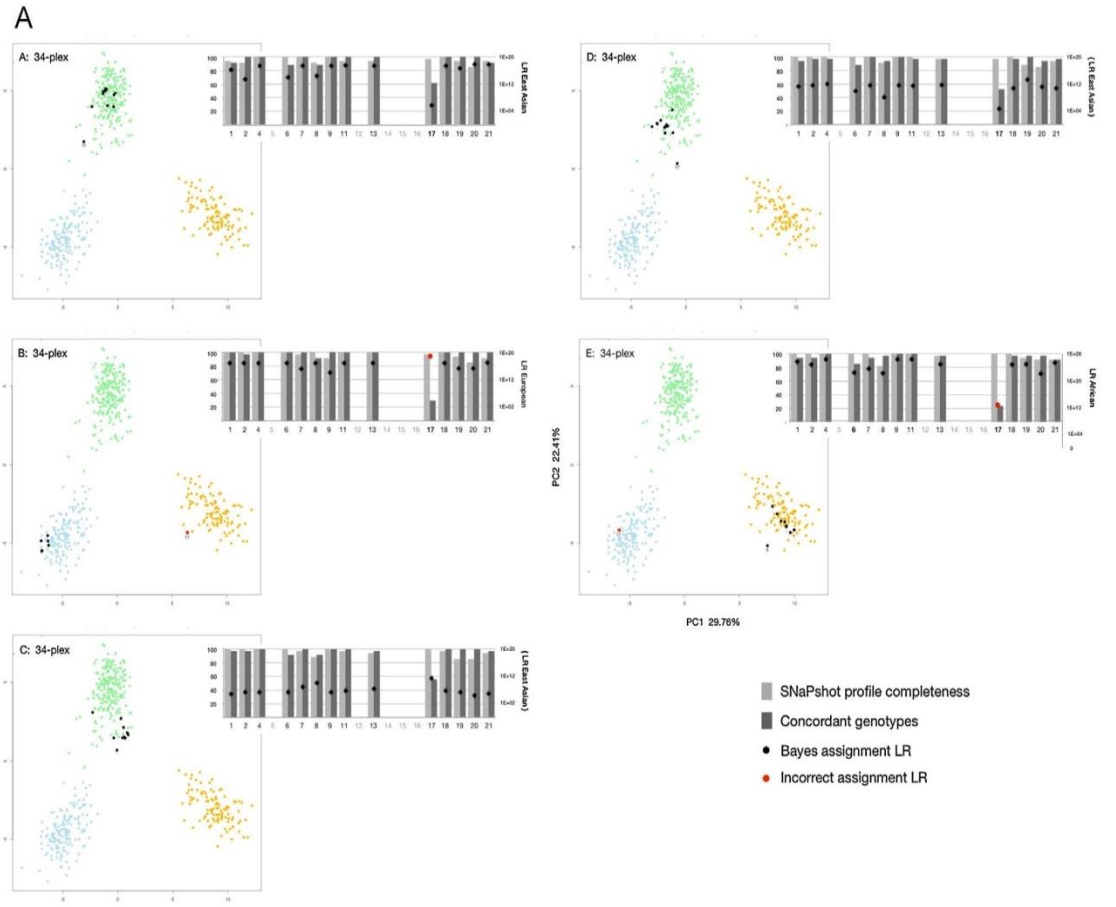


Fig. 4. (A) Participant's SNP-based Bayes LRs and PCA positions for three-group comparisons (AFR-EUR-E ASN) analyzing samples A-E. Genotype completeness and concordance rates are shown as bar charts (left-hand scales) and ancestry assignment LRs (i.e., lowest values) as overlaid points (right-hand scales). Laboratories with some displacement of a sample position from the main PCA cluster are individually labeled and incorrect positions/assignments from miscalled genotypes are shown in red. (B) Participant's Indel-based Bayes LRs and PCAs for 4-group comparisons (including Americans) analyzing A-E. The sixth plot, lower right, shows a 5-group PCA of sample C (adding Oceanian reference data) using 80-marker genotypes. Laboratories only reporting Indel data have Bayes LRs shown in green and one uninformative LR shown in blue. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

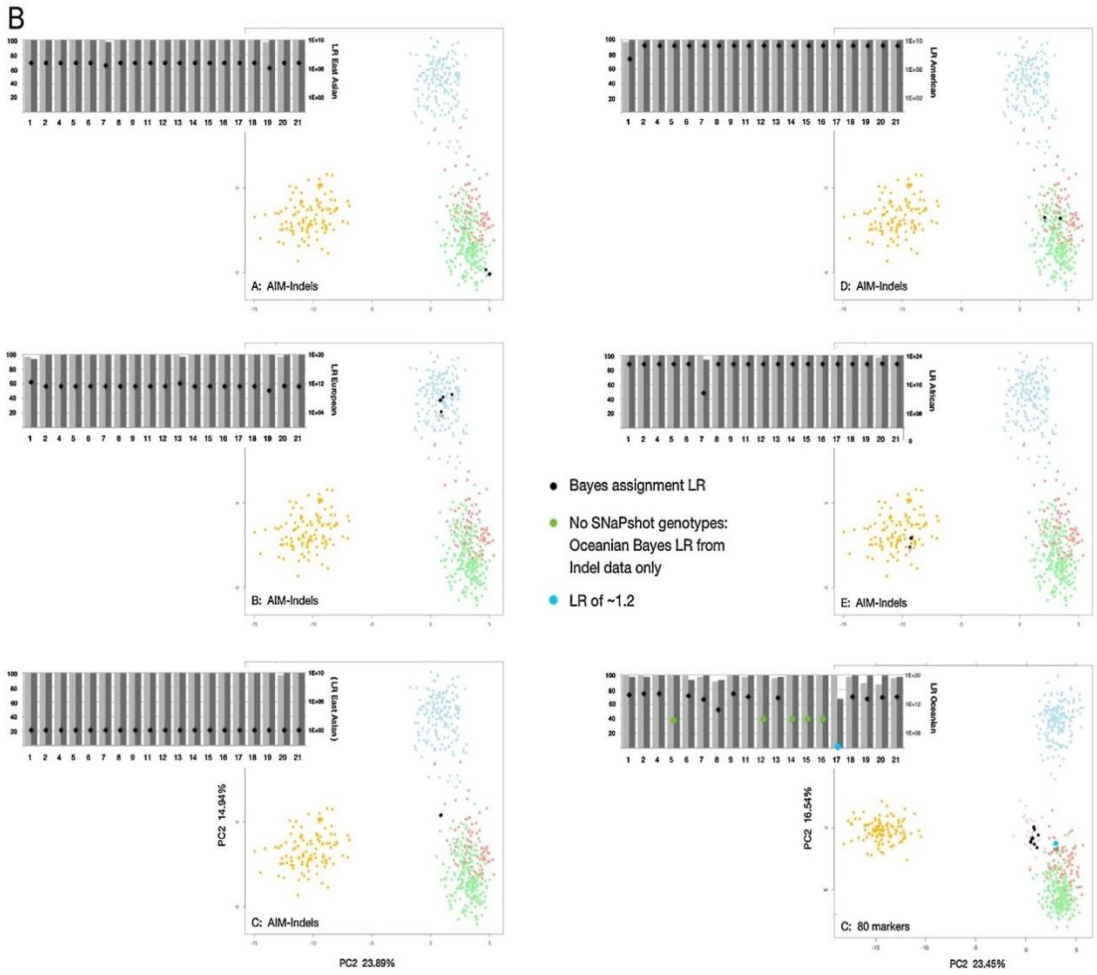


Fig. 4. (Continued)

participants were running the test for the first time and required reading 46 different peak sets in each electropherogram.

Fig. 3B summarizes the Indel genotyping performance and shows participants achieved a very high overall genotyping completeness and concordance rate of 99.8%. Fourteen participants did not have miscalls or no-calls in any test sample profiles. A slight degree of clustering of genotyping miscalls and no-calls is discernible in Fig 3B; for example, laboratory #20 chose not to call 3/5 rs2307922 genotypes, and laboratories #1 and #7 mistyped more than one Indel. It is notable that all 19 participants successfully identified the rare third allele of rs25584 present in test sample C.

3.3. Inference of ancestry

All participants identified F as the mixed DNA sample and made Bayes analysis to infer the ancestry of samples A-E using *Snipper*. The majority, but not all, also made comparisons of the genotypes from A-E with the *Snipper* PCA module using the supplied reference population data. This section summarizes results for all laboratories using both statistical approaches to illustrate that the SNP and Indel data has a degree of ancestry-informativeness redundancy, i.e., the Bayes LRs or PCA positions of samples A-E are very similar despite some genotype miscalls or missing data. Therefore, the ancestry inferences made by participants were correct in all cases apart from those of laboratory #17 that made incorrect ancestry inferences for two samples and had PCA positions markedly displaced from the others in most cases.

Fig. 4A summarizes SNP profile quality (bar-charts, left-hand scale); Bayes LRs (points superimposed on bars, right-hand scale); and PCA positions for the SNaPshot assay data of 14 participants, analyzing samples A-E. Bayes LRs and PCAs from SNP data alone compare African, European and East Asian ancestries; consequently C and D give lower LRs and edge-of-cluster PCA positions that suggest East Asian ancestry despite these being Oceanian and American in origin. For 13/14 laboratories, samples A, B and E give mid-cluster PCA positions and high LRs that varied by four orders of magnitude between 1E+14 to 1E+18 correctly assigning A as East Asian and B as European, and 1E+22 to 1E+26 correctly assigning E as African. The LR values obtained by coordinating laboratory USC for SNP and Indel data are outlined in Table 1 (the 80-marker LRs for all samples are given separately in Fig. 2). Table 1 indicates sample C gave a high LR for Oceanian ancestry with just Indel data used in a 5-group comparison.

Fig. 4B summarizes Indel profile quality, Bayes LRs and PCA positions for a four group comparison using the Indel data of all participants. A sixth PCA plot, bottom right, shows the combined 80-marker analysis for Oceanian sample C. Apart from African sample E, Indel data gives lower LRs than SNPs and the LRs for samples A and E are from different population likelihoods (bold values in Table 1). The improved genotyping consistency of Indels amongst participants is reflected in more uniform sets of Bayes LRs

Table 1
Lowest LR values produced from *Snipper* Bayes analysis of the full SNP and Indel profiles of samples A-E and 9947A with their ancestry inferences. Participant LR values for the same samples are plotted in Fig. 4A/B. Bold values for A and E highlight different population ratios giving the lowest LRs when SNP, 3-group or Indel, 4-group comparisons are made. With Indel, 4-group comparisons the second lowest LRs for samples A and E are based on the same population ratios as the lowest LRs for SNP, 3-group comparisons. Sample C is correctly inferred to be Oceanian with Indel data alone but most participants reported the LR from 80 marker data.

Inference	**34-plex SNPs, 3-group**	
European	9947A is 2118,840,589,047,061,020,672 times more likely EUROPEAN than E ASIAN	
East Asian	**A is 361,148,635,069,545,024 times more likely E ASIAN than EUROPEAN**	
European	B is 64,191,487,284,485,608 times more likely EUROPEAN than E ASIAN	
East Asian	C is 13,115,706 times more likely E ASIAN than AFRICAN	
East Asian	D is 248,539,593,557 times more likely E ASIAN than EUROPEAN	
African	**E is 556,454,701,312,037,054,117,314,560 times more likely AFRICAN than E ASIAN**	
	46-plex Indels, 4-group	**46-plex, 4-group (second lowest LR)**
European	9947A is 1937,432,967,198 times more likely EUROPEAN than E ASIAN	
East Asian	**A is 6993,957 times more likely E ASIAN than AMERICAN**	A is 37,290,377,821,078,192,128 times more likely E ASIAN than EUROPEAN
European	B is 143,659,679,122 times more likely EUROPEAN than E ASIAN	
LR too low	C is 131 times more likely E ASIAN than EUROPEAN	
American	D is 944,698,134 times more likely AMERICAN than E ASIAN	
African	**E is 3229,841,442,838,053,650,432 times more likely AFRICAN than EUROPEAN**	E is 5715,694,248,335,998,122,459,136 times more likely AFRICAN than E ASIAN
	46-plex, 5-group	**80 Markers, 5-group**
Oceanian	C is 24,880,402 times more likely OCEANIAN than E ASIAN	C is 153,747,536,542,653 times more likely OCEANIAN than E ASIAN

and PCA positions that mainly overlay each other (i.e., seen as single points on plots). For the two laboratories with three Indel miscalls, an effect is seen in the Bayes LRs for American sample D and African sample E, with some PCA displacement, indicating that even with just two markers miscalled it can sometimes affect the statistical inference made from other correctly called genotypes (~97% of the data). The Oceanian sample C was correctly identified by 18 participants, with many using both Indel and combined data to make the inference.

Therefore, 18 of 19 laboratories were able to successfully assign ancestry to five samples of undisclosed geographic origin, obtaining unequivocal Bayes LRs and, in most, cases participants constructed PCA plots providing supplementary analyses with good matches to the Bayes results.

3.4. Mixture detection and analysis of participant's Indel peak height data

Although the exercise was not a fully blinded test (i.e., where the presence of a mixed sample is not disclosed), all participants were able to identify sample F as the mixture from the observation of imbalanced signals in the heterozygote peak pairs of the Indel

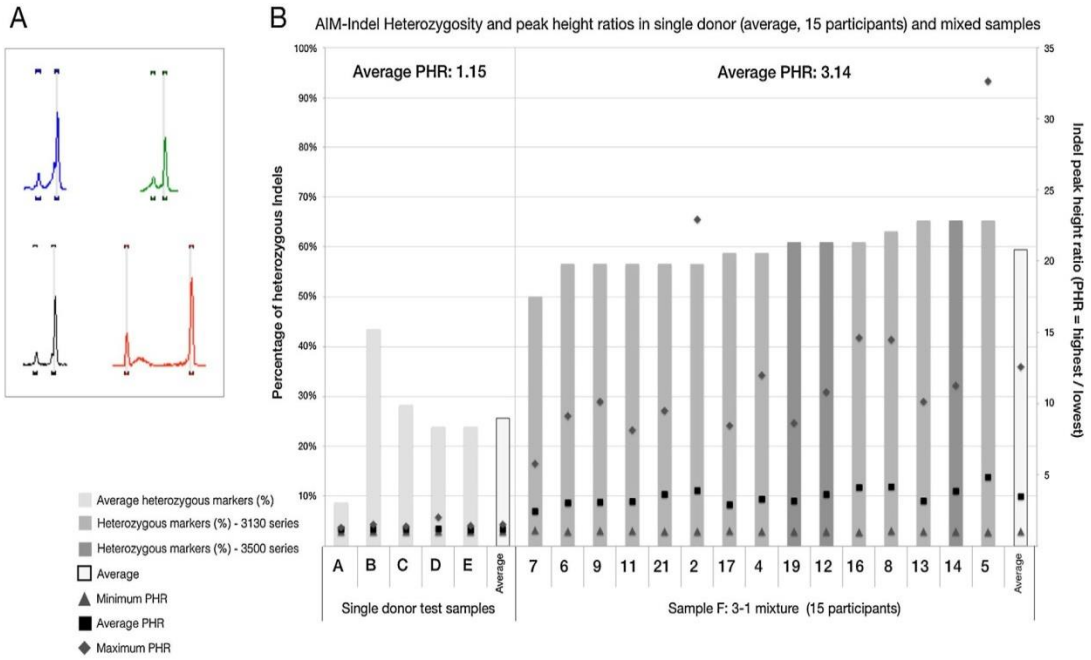


Fig. 5. (A) Example Indel peak pairs for sample F discounted as heterozygotes by one participant. (B) Numbers of Indel heterozygotes (bars) and their peak height ratios (PHR: points) recorded by 15 participants. Unmixed samples A-E are average values from all data and sample F values are shown individually as different numbers of peak pairs were recognized as heterozygotes amongst laboratories.

profile. Therefore, despite a lack of familiarity with Indel peak patterns in most laboratories, there was sufficient contrast between the mixed sample F and the unmixed A-E DNAs for the mixture to be discernible by all participants. In addition, 7 of 14 laboratories reported an ACG triple-peak pattern in the tri-allelic SNP rs5030240, one reported an AC with possible G, one a GG result and the other five gave no-calls. A typical sample F peak pattern for rs5030240 is shown in Supplementary Fig. S1.6.

The detection of peak height imbalances that can indicate mixed DNA has been stated to be an advantage of direct PCR-to-CE Indel genotyping compared to SNaPshot tests [15,17,19], however such patterns have not been properly assessed across a range of CE detectors. For this reason, we decided to ask participants to provide their heterozygote peak height data and then compiled the variation in peak height ratios (PHRs, highest/lowest peaks) recorded in the five unmixed and single mixed DNAs from the range of CE regimes used. Furthermore, when analyzing binary markers the number of heterozygotes observed in mixtures is invariably higher than normal unmixed samples. Although PHR values were distinct between A and E and F, three factors complicated the straightforward statistical comparison of patterns of heterozygosity observed amongst the test samples. First, there was variation in the number of heterozygotes recorded in sample F. Specifically, laboratory #1 identified 18 heterozygotes; #15: 17; #18: 21; and #20: 17, compared to an average number of heterozygotes identified by the other fifteen laboratories of 27. Second, the lower number of identified heterozygotes for F in some participant's data affects the minimum–maximum and average PHR values, particularly when the PHR is extreme and a very low peak is discounted when reading the profile. Four example peak pairs that were recorded as single allele genotypes by one participant but as heterozygotes by the others, are shown in Fig. 5A. Third, due to the contrasting frequencies of most of the 46 Indels between population groups, sample A showed lower numbers of heterozygotes and sample B higher numbers than those seen in C–E.

The numbers of heterozygotes and PHR values are plotted in Fig. 5B. This chart shows data from 15/19 laboratories (excluding #1, #15, #18 and #20). The same chart with all 19 participant's data is shown in Supplementary File S4A. The dark grey bars mark the data from 3500 detectors and indicates that no difference in peak height ratios are discernible in comparison with 31xx CE data.

Statistical assessment of the number of heterozygotes in A-E vs. F was made with a unilateral 2-sample test for equality-of-proportions (with continuity correction). The resulting grid of p-values for pairwise comparisons across all 19 laboratories is shown in Supplementary File S4B, along with the Fig 5B chart re-plotted for full data from all laboratories (Supplementary File S4A). It can be seen from the Supplementary File S4A chart that the numbers of sample F heterozygotes recorded by laboratories #1, #15 and #20 is lower than the average number in unmixed sample B. Inclusion of this data has a direct effect on the distribution of significant p-values obtained from pairwise comparisons. Laboratories #1, #15 and #20 sample F heterozygote numbers are significantly different to those of most of the other laboratories, but not different to heterozygote numbers in unmixed samples B–E, while #18 data for sample F is not significantly different to samples B and C. The high number of heterozygotes in sample B is reflected in significant differences only found for comparisons to those of laboratories #8, #13, #14, and #5, who recorded 29 or more heterozygotes in their sample F profiles. Therefore, we opted to remove #1, #15, #18 and #20 data from the statistical assessment of PHR differences between A-E and F.

The average PHRs shown in Fig. 5B indicate a quite distinct contrast between samples A-E and F, with values of 1.15 compared to 3.14 respectively, which suggests a ratio of 1:2.73 that approximates the actual 1:3 contributor ratio well. Although the PHR values give a clearly discernible difference between mixed and unmixed samples, we completed a formal statistical test of this difference. An ANOVA test is a standard approach for assessing continuous values such as PHR measurements, but a Shapiro–Wilks test indicated that some of the data was not normally distributed (data not shown). Therefore, a Kruskal–Wallis rank sum test was applied and the grid of pairwise p-values comparing the average PHRs of A-E with individual PHRs of F is shown in Supplementary File S4C. The results are completely consistent: the pairwise comparisons of mixed vs. unmixed PHRs give significant p-values in every case and none were detected for comparisons within each sample set.

In summary, despite the need to adjust statistical comparisons by removing 4 of 19 participant's data due to under-reported heterozygote peak pairs, the other laboratories provided a ratio of average peak heights close to 1:3. This ratio is consistent with the mixture that was constructed for the exercise and is statistically significant for all signal strength comparisons made.

3.5. Additional next generation sequencing experiments applied to test DNAs by two laboratories

Two laboratories decided to use their remaining PCR primers to genotype one or both marker sets with different Next Generation Sequencing (NGS) systems, as outlined in Supplementary File S5. One assessed 34-plex SNP typing using an unmodified PCR followed by library preparation and massively parallel sequencing with the Illumina MiSeq system. The other assessed 34-plex SNPs and Indel genotyping in the same way (unmodified PCR in each case) with the Thermo Fisher Scientific-Life Technologies (TFS-LT) Ion PGM™ system.

The 34-plex SNP sequence analyses were successful to a very large degree, as all genotypes were identified and almost fully concordant with each laboratory's SNaPshot data. Sample F was observed to be distinct in a major proportion of its allele-pair sequence ratios (defined as the second allele exceeding 10% of sequence reads), compared to A-E. Supplementary File S5 indicates there were only 5/14 sequence ratios of 1.5 or less (i.e., in the range: 0.4:0.6–0.5:0.5) in the Ion PGM™ data and 3/17 in the MiSeq data. This equates to 64% and 82% of sequence ratios exceeded those of most normal DNA heterozygotes seen in Ion PGM™ and MiSeq respectively, giving unequivocal signals of a mixture in F. Both systems also detected displaced sequence ratios in each of the two tri-allelic SNPs.

The Indel analysis with NGS gave three discordant genotypes in samples B and C, plus an average 8.7% no-calls (coverage too low) and 2.9% missing data (undetected sequence), although not all samples gave the same non-detection rates. Overall, 84% of the NGS genotypes matched the CE calls. However, the alignment of sequences that contain short insertions and deletions is particularly challenging in NGS sequence analysis and it was not possible to be sure how many miscalls or no-calls were due to misalignment issues. Supplementary File S5 shows assessments of Indel sequence ratios for sample F compared to A-E. Given that sequence coverage was low in some loci and this is the first NGS experiment with this type of forensic marker, results need cautious interpretation. However, patterns suggest a degree of displacement in F away from the perfect sequence balance midline (0.5:0.5) compared to many of the heterozygote sequence ratios detected in A-E.

4. Discussion

As forensic NGS analysis gains greater traction, it is the right moment for the forensic community to use inter-laboratory

exercises to assess the binary marker sets that will start to add complementary genetic data to conventional STR polymorphisms. Ancestry inference is seen as a key part of the enhanced characterization of forensic DNA that NGS will allow. Therefore, it is important to evaluate the robustness of existing CE-based ancestry-informative SNP and Indel multiplexes in terms of how easily they can be adopted in laboratories not previously experienced with binary marker genotyping. The statistical analysis of the genotype data obtained from AIMs also needs to be easy to use and interpret by forensic laboratories. The most straightforward approach for inferring ancestry uses Bayesian LR comparisons between the two geographic origin hypotheses with the highest likelihoods. Lastly, binary variation has a reduced capacity to detect mixtures since homozygotes in combination can look like heterozygotes and only a few non-binary SNPs or Indels currently offer the chance to observe more than two alleles. Therefore, the exercise findings for genotyping reliability, ease-of-use of the recommended ancestry inference tools and ability to detect mixed DNA are all relevant to the progress towards adoption of AIMs in forensic analysis.

The principal finding of this exercise was that each of the participants readily established the AIM-Indel 46-plex test in their laboratory. All participants achieved good quality profiles that reached the high level of genotyping concordance of 99.8% and then efficiently detected mixed sample F. In contrast, SNaPshot typing was both more challenging and for many participants less reliable, despite most laboratories having successfully genotyped six SNPs for the preceding *IrisPlex* EDNAP exercise [21]. Miscalled genotypes with SNaPshot produced an overall genotyping concordance rate of 97.3% when a single participant's results were excluded (13/14 laboratories).

We have no explanation for the very high number of SNP miscalls from this one laboratory but it resulted in their statistical analyses producing the only incorrect ancestry inferences for two test DNAs and one uninformative LR of 1.2. All other participants produced correct ancestry predictions from the Bayes LRs calculated in *Snipper* and, for those that created PCA plots, obtained cluster patterns and profile positions that corresponded to these LRs. Therefore, from the review of exercise reports returned from 19 laboratories, we can recommend the use of both of these statistical approaches to ancestry inference, as these proved easy to use and allowed correct ancestry assignments of samples with undisclosed geographic origin in 92 of 95 cases.

Mixture detection achieved from Indel peak patterns was particularly successful, with sample F giving a clear signal of mixed DNA for all participants. Our analysis of peak height ratios made after the exercise finished, gave a good approximation of the actual mixture component ratio, averaging 1.15 and 3.14 for PHRs in unmixed samples and the mixture, respectively. The much higher number of heterozygotes in F could mainly be due to the different ancestries of the mixture contributors. Nevertheless, recording a higher number of heterozygotes than in normal DNA samples and observing PHRs markedly above ~1.2 gives a simple and easily adopted system to detect mixtures with Indels. SNaPshot does not offer the same direct relationship between peak heights and input DNA so there is a risk that simple two-person mixtures mimic the patterns seen in individuals with co-ancestry due to population admixture, as revealed by the PCA plot of sample F in Fig. 2 (plot A). Obviously, single sample experiments are not fully indicative of how well Indels will perform with a range of forensic samples, mixture ratios or component ancestry combinations, but the fact that most participants were running Indels for the first time and all detected the mixture indicates sensitivity to mixed DNA with this assay.

Although the NGS findings from two participants are a set of parallel genotyping experiments using exercise materials that were not part of the study plan, results are included in this report to highlight the enhanced sensitivity to mixtures obtained for SNP analysis with NGS. It is also interesting to note that existing optimized forensic multiplexes work very well in NGS without the need for any modification, confirming the results of a recent study that found the 34-plex PCR primers, amongst four other forensic SNP multiplexes, provide good quality output with the Ion PGM™ system [24]. In addition, the relative success of the initial Indel genotyping experiments with NGS indicate dye-labeled PCR primers do not interfere with library preparation and subsequent sequencing chemistry of the Ion PGM™. This suggests existing forensic CE multiplexes for a range of markers, including STR kits, could be used to prepare target DNA for experimental NGS sequencing runs.

Until NGS systems that incorporate AIMs are widely adopted for forensic use, the results from this EDNAP inter-laboratory exercise indicate the PCR-to-CE Indel test is by far the best current option for forensic ancestry analysis. The Indel multiplex provides a simple, reliable and informative test from a comparatively large marker set that is analyzed using validated CE regimes. Detection of simple two-component mixed DNA from scrutiny of Indel peak patterns was a task accomplished by all exercise participants and gives Indel genotyping a key additional advantage over SNP-based ancestry tests.

Acknowledgements

This work was partially supported through a grant to CS (SFRH/BD/75627/2010) awarded by the Portuguese Foundation for Science and Technology (FCT) and co-financed by the European Social Fund (Human Potential Thematic Operational Programme). The organizing laboratory of the Forensic Genetics Unit, University of Santiago de Compostela, would like to thank Fabio Oldoni for his assistance in piloting the exercise, in partial fulfillment of an M.Sc. in Forensic Science, Department of Forensic Science and Drug Monitoring, King's College, London. The two anonymous donors for the mixture are also thanked for their willingness to provide samples. Zurich Institute of Forensic Medicine thanks Corinne Moser for technical assistance. The pilot studies, Indel primers and test DNAs used in the exercise were funded by the European Forensic Genetics Network of Excellence (EUROFORGEN-NoE, Grant Agreement No. 285487).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.fsigen.2015.06.004.

References

[1] C. Phillips, L. Prieto, M. Fondevila, Á. Salas, Á. Gomez-Tato, J.Á. Alvarez-Dios, Á. Alonso, Á. Blanco-Verea, M. Brio'n, M. Montesino, Á. Carracedo, M.V. Lareu, Ancestry analysis in the 11-M Madrid bomb attack investigation, Plos One 4 (2009) e6583.

[2] S. Walsh, F. Liu, A. Wollstein, L. Kovatsi, A. Ralf, A. Kosiniak-Kamysz, W. Branicki, M. Kayser, The HIrisPlex system for simultaneous prediction of hair and eye colour from DNA, Forensic Sci. Int. Genet. 7 (2013) 98–115.

[3] M. Kayser, P. de Knijff, Improving human forensics through advances in genetics, genomics and molecular biology, Nat. Rev. Genet. 12 (2011) 179–192.

[4] M. Kayser, P.M. Schneider, DNA-based prediction of human externally visible characteristics in forensics: motivations, scientific challenges, and ethical considerations, Forensic Sci. Int. Genet. 3 (2011) 154–161.

[5] C. Phillips, L. Fernandez-Formoso, M. Garciañas, L. Porras, T. Tvedebrink, J. Amigo, M. Fondevila, A. Gomez-Tato, J. Alvarez-Dios, A. Freire-Aradas, A. Gomez-Carballa, A. Mosquera-Miguel, Á. Carracedo, M.V. Lareu, Analysis of global variability in 15 established and 5 new European Standard Set (ESS) STRs using the CEPH human genome diversity panel, Forensic Sci. Int. Genet. 5 (2011) 155–169.

[6] C. Phillips, M. Gelabert-Besada, L. Fernandez-Formoso, M. Garci'a-Magariños, C. Santos, M. Fondevila, D. Ballard, D. Syndercombe Court, Á. Carracedo, M.V. Lareu, New turns from old STaRs: enhancing the capabilities of forensic short tandem repeat analysis, Electrophoresis 35 (2014) 3173–3187.
[7] T.E. King, E.J. Parkin, G. Swinfield, F. Cruciani, R. Scozzari, A. Rosa, S.K. Lim, Y. Xue, C. Tyler-Smith, M.A. Jobling, Africans in Yorkshire? The deepest-rooting clade of the Y phylogeny within an English genealogy, Eur. J. Hum. Genet. 15 (2007) 288–293.
[8] O. Lao, P.M. Vallone, M.D. Coble, T.M. Diegoli, M. van Oven, K.J. van der Gaag, J. Pijpe, P. de Knijff, M. Kayser, Evaluating self-declared ancestry of US Americans with autosomal, Y-chromosomal and mitochondrial DNA, Hum. Mutat. 31 (2010) E1875–E1893.
[9] D. Corach, O. Lao, C. Bobillo, K. van Der Gaag, S. Zuniga, M. Vermeulen, K. van Duijn, M. Goedbloed, P.M. Vallone, W. Parson, P. de Knijff, M. Kayser, Inferring continental ancestry of Argentineans from autosomal, Y-chromosomal and mitochondrial DNA, Ann. Hum. Genet. 74 (2010) 65–76.
[10] C. Bouakaze, C. Keyser, E. Crube'zy, D. Montagnon, B. Ludes, Pigment phenotype and biogeographical ancestry from ancient skeletal remains: inferences from multiplexed autosomal SNP analysis, Int. J. Legal Med. 123 (2009) 315–325.
[11] M. Fondevila, C. Phillips, C. Santos, Á. Freire Aradas, P.M. Vallone, J.M. Butler, M. V. Lareu, Á Carracedo, Revision of the SNPforID 34-plex forensic ancestry test: assay enhancements, standard reference sample genotypes and extended population studies, Forensic Sci. Int. Genet. 7 (2013) 63–74.
[12] C. Phillips, Á Freire Aradas, Á.K. Kriegel, M. Fondevila, O. Bulbul, C. Santos, F. Serrulla Rech, M.D. Perez Carceles, A'. Carracedo, P.M. Schneider, M.V. Lareu, Eurasiaplex: a forensic SNP assay for differentiating European and South Asian ancestries, Forensic Sci. Int. Genet. 7 (2013) 359–366.
[13] M. Fondevila, C. Phillips, N. Naveran, L. Fernandez, M. Cerezo, Á. Salas, Á. Carracedo, M.V. Lareu, Case report: identification of skeletal remains using short-amplicon marker analysis of severely degraded DNA extracted from a decomposed and charred femur, Forensic Sci. Int. Genet. 2 (2008) 212–218.
[14] C. Børsting, E. Rockenbauer, N. Morling, Validation of a single nucleotide polymorphism (SNP) typing assay with 49 SNPs for forensic genetic testing in a laboratory accredited according to the ISO 17025 standard, Forensic Sci. Int. Genet. 4 (2009) 342.
[15] R. Pereira, C. Phillips, C. Alves, Á. Amorim, Á. Carracedo, L. Gusmão, A new multiplex for human identification using insertion/deletion polymorphisms, Electrophoresis 30 (2009) 3682–3690.
[16] S.L. Friis, E. Børsting, L. Poulsen, S.F. Fredslund, C. Tomas, N. Morling, Typing of 30 insertion/deletions in Danes using the first commercial indel kit—Mentype® DIPplex, Forensic Sci. Int. Genet. 6 (2012) e72–e74.
[17] M. Fondevila, C. Phillips, C. Santos, R. Pereira, L. Gusmão, Á. Carracedo, J.M. Butler, M.V. Lareu, P.M. Vallone, Forensic performance of two insertion-deletion marker assays, Int. J. Legal Med. 126 (2012) 725–737.
[18] N.P. Santos, E.M. Ribeiro-Rodrigues, A.K. Ribeiro-dos-Santos, R. Pereira, L. Gusmão, A. Amorim, J.F. Guerreiro, M.A. Zago, C. Matte, M.H. Hutz, S.E. Santos, Assessing individual interethnic admixture and population substructure using a 48-insertion-deletion (INDEL) ancestry- informative marker (AIM) panel, Hum. Mutat. 31 (2010) 184–190.
[19] R. Pereira, C. Phillips, N. Pinto, C. Santos, C.E.B. Santos, A. Amorim, Á. Carracedo, L. Gusmão, Straightforward inference of ancestry and admixture proportions through ancestry-informative insertion deletion multiplexing, Plos One 7 (2012) e29684.
[20] D. Zaumsegel, M.A. Rothschild, P.M. Schneider, A 21 marker insertion deletion polymorphism panel to study bio-geographical ancestry, Forensic Sci. Int. Genet. 7 (2013) 305–312.
[21] L. Chaitanya, S. Walsh, J.D. Andersen, R. Ansell, K. Ballantyne, D. Ballard, R. Banemann, C.M. Bauer, A.M. Bento, F. Brisighelli, et al., Collaborative EDNAP exercise on the IrisPlex system for DNA-based prediction of human eye colour, Forensic Sci. Int. Genet. 11 (2014) 241–251.
[22] J. Novembre, M. Stephens, Interpreting principal component analyses of spatial population genetic variation, Nat. Genet. 40 (2008) 646–649.
[23] D. Reich, A. Price, N. Patterson, Principal component analysis of genetic data, Nat. Genet. 40 (2008) 491–492.
[24] R. Daniel, C. Santos, C. Phillips, M. Fondevila, R.Á. van Oorschot, Á. Carracedo, M.V. Lareu, D. McNevin, A SNaPshot of next generation sequencing for forensic SNP analysis, Forensic Sci. Int. Genet. 14 (2015) 50–60.