Characterisation of X Chromosomal Short Tandem Repeat Markers for Forensic Use

Flinders University School of Biological Sciences

Toni Marie Diegoli, B.A., M.F.S.

Declaration Page

I certify that this thesis does not incorporate without acknowledgment any material previously submitted for a degree or diploma in 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.

The opinions or assertions presented hereafter are the private views of the author and should not be construed as official or as reflecting the views of the Department of Defense, its branches, the US Army Medical Research and Materiel Command or the Armed Forces Medical Examiner System.

Portions of this project were supported by Award No. 2011-DN-BX-K401, awarded by the National Institute of Justice, Office of Justice Programs, U.S. Department of Justice. The opinions, findings, and conclusions or recommendations expressed in this thesis are those of the author and do not necessarily reflect those of the Department of Justice.

Toni Marie Diegoli 1/15/14

Toni Marie Diegoli Date

Abstract

The use of X chromosomal short tandem repeat (STR) markers has been greatly increasing in the forensic setting. The marker system offers the potential to provide information in addition to that which is obtained from autosomal STR systems currently employed at forensic and paternity laboratories and in the courtroom. In certain scenarios, markers on the X chromosome may be the only means of obtaining the required information. Any investigated relationship situation where at least one female is involved will benefit from the use of X STRs, which can be applied to cases of missing persons, criminal incest, immigration, deficiency paternity or other questioned relationships. In-depth characterisation of the marker system is the first step in maximizing the power of this additional tool in the forensic arsenal.

Using guidelines set forth within the 1991 report of the International Society for Forensic Genetics (ISFG) relating to the use of DNA polymorphisms, all aspects of the feasibility of routine X STR use were evaluated. Two mini-X chromosomal STR multiplexes capable of amplifying 15 total markers (DXS6795, DXS9902, DXS8378, DXS7132, DXS6803, DXS6789, DXS7424, DXS101, GATA172D05, DXS7130, GATA165B12, HPRTB, GATA31E08, DXS10147, and DXS7423) were developed and optimised for use in the investigation and characterisation of allele nomenclature, allele/genotype frequencies, mutation rates, and linkage between markers. To simplify the transition into routine use, the assays utilised techniques and instrumentation already present in most forensic laboratories as well as analyses familiar to forensic scientists. Several large sample sets from four U.S. populations (African Americans, U.S. Asians, U.S. Caucasians, and U.S. Hispanics) were studied to address the lack of relevant data available for the United States, where the use of X STRs remains limited. Much of the knowledge gained, however, is universally applicable.

Though commercial kits are available that probe a wide variety of genetic markers on both the Y chromosome and the autosomes, there is only one commercial kit currently manufactured assaying markers on the X chromosome, the Investigator Argus X-12 kit (QIAGEN, Hilden, Germany). Due to patent and intellectual property issues between the United States and European STR kit manufacturers, the

kit cannot be sold or marketed in the U.S. at this time. Nonetheless, four markers overlapping with those present in the Argus X-12 kit were included in the optimised multiplexes and were the subject of an extensive concordance study. With greater than 99% concordance across 975 alleles, comparability of the data was established for these four markers while the utility of the kit with U.S. populations was evaluated for the first time.

This extensive developmental validation study investigated each aspect of the X STR marker system that would require consideration before implementation by a forensic laboratory. The optimised assays were found to be robust and the markers discriminating, while the mutation rate was estimated with high accuracy and the extent of linkage between the 15 markers was thoroughly evaluated. The combination of the results obtained as part of this study form the foundation upon which the introduction and routine use of X STRs may be built.

Acknowledgements

Though the following dissertation will be viewed as an individual accomplishment, the work involved could never have been completed without the encouragement, support, and help of a number of individuals that has meant so much over the past three years (and really longer if the months of indecision about heading back to school are considered). First and foremost I must thank my academic advisor Professor Adrian Linacre who has, from the start, been more than willing to work with and champion an atypical student who resides on the other side of the globe. You welcomed me into your professional life without hesitation, and I am grateful to have gotten to know you on a personal level as well. It has been an honour to study under the guidance of such an accomplished scientist, even more so because you are also an inspiring individual. For all the learning you have supported and constant encouragement you have provided, my gratitude is most sincere.

I am also deeply indebted to Dr. Michael Coble, for whom I hold the utmost respect and admiration. From the very first interactions with you as my new boss early in my career at the Armed Forces DNA Identification Laboratory, you made me hungry for the answers only scientific research could provide. It was exhilarating to set goals working under your direction, and then surprise myself by actually reaching them. Though I had been considering both a career in research as well as a return to school, interactions with you solidified both decisions for me. And, of course, I never would have started down the path that led me to the most inspiring of all the chromosomes without you first recognizing its potential utility in forensic investigations, and then allowing me to pursue all aspects of its implementation. Though you are no longer my boss – and have not been through much of the dissertation process – you continue to offer advice when I need it and support me in every way. Over the years I have learned so much from you both personally and professionally, and I feel privileged to be able to call you a friend.

To Dr. John Butler, I am thankful for your drive and insight that has pushed me to become a better student and scientist. Your breadth of knowledge and continued pursuit of it has provided a model towards which I could strive in each of my scientific endeavours. Though I remain far from grasping at the bar you have set so

high, I am thankful to have had the opportunity to interact with you throughout this process.

To Dr. Tom Parsons, Dr. Jodi Irwin, and Rebecca Just, thank you for setting an example of what it takes to become a successful researcher, and for recognizing and supporting those qualities in your employees. I have learned so much from each of you, and your dedication to your work has always been inspiring to me. Jodi, you encouraged me to think seriously about the dream of pursuing a doctoral degree, and gave me the courage to believe that I could actually make it happen while working full-time. To my co-worker and friend Kim Andreaggi, you make research a truly fun experience; experiments-gone-wrong never feel as bad when you have someone with which to commiserate. I am so thankful that our professional relationship turned into such a close friendship, and your support along the way has been invaluable. To the rest of the research team at the AFDIL – Dr. Odile Loreille, Melissa Scheible, Jennifer Higginbotham, Spence Fast, Jocelyn Bush, Michelle Peck, and Joseph Ring – thank you for all the hard work you do that keeps me on my toes, and thank you for being patient with me when I locked myself in my office with the door closed. Last but never least, I must thank Col. Louis Finelli, Lt Col Laura Regan, James Canik, and Dr. Timothy McMahon, the entire Armed Forces Medical Examiner System, and the American Registry of Pathology for providing such a stimulating work environment. You have challenged me to become a better scientist and leader while providing me with the tools and support with which to grow.

Of course, heaps of gratitude belong with my husband-to-be, Dr. Peter Vallone. I am most appreciative of your patience with me as I worked through the process of considering and then pursuing this degree. Your personal sacrifices that have allowed me to chase this professional goal have not gone unnoticed, and the unwavering support you have given to me has kept me focused. Ours is a particularly unique relationship, being that we are both forensic researchers, but it also allows me a unique opportunity to understand what personal and professional qualities have led to your intelligence and success, which I have admired from the start. You have always been more than willing to help me professionally – from demonstrations of your so-called lab hands to finding typos in my poster presentations – but you have also made me a better person. This process has proven

that we both want the same things out of life, and I look forward to spending the rest of mine (free from schoolwork) with you.

Though (through no fault of hers) we have recently been out of touch, I want to thank my college roommate Jenn Hoang for her confidence in me. I will always remember the conversation we had when I first told you of this crazy idea I had about pursuing a doctoral degree, and how you encouraged me to go for it. Not once did you doubt what I could accomplish. Thank you. My dearest friend, Kerry O'Brien, also deserves much credit for lending a sympathetic ear over long runs during the early years of this process and never failing to be both enthusiastic and sincere about every little bit of progress. Thank you for your friendship, which has been vital and motivating.

Finally, I must thank my family, whose unconditional support throughout this process – and my life – has meant the world. My parents, Carole and Dennis Diegoli, provided me with every opportunity that has led me to this point, and have helped me pick myself up and make the best of difficult situations. My sister, Kara Yifru – and her wonderful family, Danny, Corey, Sam, and George – is always there for me, motivating me to "get writing" so I could come home to visit my adorable nephews. My grandmothers, Gram and Nani, my Aunt Patty, and my cousin Brett have all been behind me every step of the way, with words and handwritten notes of encouragement. Gram never failed to warm my heart by reminding me how proud my late grandfather, Papi, would have been. Though I will miss the phone calls, emails, letters, and conversations asking when I will be done, how many pages I have written today, and what percentage I have completed so far, I am grateful that you all kept me true to my goals and cared enough to know how important this degree was to me.

Abbreviations

µL	Microliter
µM	Micromolar
A	Adenine
AFDIL	Armed Forces DNA Identification Laboratory
BLAST	Basic Local Alignment Search Tool
BLAT	BLAST-like Alignment Tool
bp	Base pair(s)
BSA	Bovine serum albumin
C	Cytosine
CEPH	Centre d'Etudes du Polymorphisme Humaine
CI	Confidence interval
cM	Centimorgan
cm	Centimeter(s)
CODIS	Combined DNA Index System
Ct	Cycle threshold
DAB	DNA Advisory Board
DF	Dilution factor
dGTP	Deoxyguanosine triphosphate
dH_2O	Distilled water
DNA	Deoxyribonucleic acid
dNTP	Deoxyribonucleotide triphosphate
EDNAP	European DNA Profiling Group
EDTA	Ethylenediaminetetraacetic acid
EMPOP	EDNAP mtDNA Population Database
FamID	Family identification number
FID	Father's identification number
F_{ST}	Fixation index
G	Guanine
H(exp)	Expected heterozygosity
H(obs)	Observed heterozygosity
INDEL	Insertion/deletion marker
IPC	Internal PCR control
IRBO	Institutional Review Board Office
ISFG	International Society for Forensic Genetics
ISFH	International Society for Forensic Hameogenetics
kb	Kilobase(s)
kV	Kilovolts
LD	Linkage disequilibrium
LG	Linkage group
LISA	Laboratory Information Systems Applications
LOD	Logarithm of the odds
LR	Likelihood ratio
Mb	Megabase(s)
MEC	Mean exclusion chance
MECI	Mean exclusion chance in trios involving daughters
MECII	Mean exclusion chance in father/daughter duos
$MgCl_2$	Magnesium chloride
MID	Mother's identification number
MIXplex	Mixture multiplex

mM	Millimolar
MRMC	U.S. Army Medical Research and Materiel Command
mtDNA	Mitochondrial DNA
NCBI	National Center for Biotechnology Information
ng	Nanograms
NIGMS	National Institute of General Medical Sciences
NIST	National Institute of Standards and Technology
^{o}C	Degrees Celcius
p (HWE)	P value of the exact test for Hardy-Weinberg equilibrium
PCR	Polymerase chain reaction
PDf	Power of discrimination in females
PDm	Power of discrimination in males
pg	Picograms
PIC	Polymorphism information content
PID	Patient identification number
Ref.	Reference
RF or Θ	Recombination fraction (or recombination rate)
SNP	Single nucleotide polymorphism
SRY	Sex-determining region of the Y chromosome
STR	Short tandem repeat
STRBase	Short Tandem Repeat Database
SWGDAM	Scientific Working Group on DNA Analysis Methods
T	Thymine
TLE	Tris-low-EDTA buffer
T_m	Melting temperature
U.S.	United States
UCSC	University of California Santa Cruz
YHRD	Y Chromosome Haplotype Reference Database
Z	LOD score

Table of Contents