Design of a low-cost real-time PCR machine

By

Lingchen Wang
Master of Engineering (biomedical)

Thesis
Submitted to Flinders University
for the degree of Master of Engineering (biomedical)

Master of Engineering (biomedical)
College of Science and Engineering
15/11/2021

I have provided feedback on this document and the student has implemented it fully
Russell Brinkworth

Declaration

I hereby acknowledge that in accordance with the University's policy on plagiarism, and unless otherwise referenced, all material presented in this thesis is my own.

_____

Lingchen wang
November 2021

Acknowledgement

I would like to acknowledge the following for their assistance in producing this report:

- Russell Brinkworth for supervising the project.
- Adrian Linacre for providing laboratory access and guidance.
- Duncan Taylor for providing experiment material and guidance.
- Ula Alexander for optical technology support.
- Youhong Tang for chemistry technology support.

Executive summary

Introduction: Polymerase chain reaction (PCR) is an important tool for DNA profiling in forensic science. The PCR machine can generate millions/billions of targeted DNA copies with only several copies to start. However, the process of PCR is almost fixed since it was discovered in the 1980s. We believe that we can add a sensor to the PCR to monitor the process, and test whether we can apply machine learning to the PCR machine to improve the PCR behaviour. The optimised PCR process might reduce time duration and increase specificity which leads to the increased efficiency of PCR and save money.

Aim: The first half of the aim of this thesis is to design a low-cost optical system that can quantify the DNA concentration of the PCR sample in real-time; the second half is to design a real-time PCR machine that uses the optical system.

Methodology and result: We have designed an optical system that can measure the concentration of DNA within a PCR product that contains fluorescence dye. The experiment uses a blue LED as the light resource and a spectrometer to detect the percentage ratio between the green and blue light from the sample. The result shows that both Green/Blue (%) and the green light intensity increase as the DNA concentration increases.

Discussion and conclusion: Our design is able is distinguish PCR products with different concentrations; however, a minimum loading volume is required to conduct a valid measurement. The future work includes experiments with more samples to complete the conversion between the fraction and DNA concentration and upgrading an open-source PCR machine to a qPCR machine (also called real-time PCR) which measures the quantification of the PCR process in real-time. Future experiments with the PCR parameters can be conducted to sketch the machine learning algorithm for PCR process optimisation.

Table of content

List of Figures and Tables

Introduction

Before the 1880s when PCR has not been discovered by Kary Mullis, biology scientists have been struggled to acquire enough copies of DNA for genetic analysis. The PCR is a simple assay that allows the amplification of DNA with specific content from a complex DNA pool, the process only requires a few copies of the DNA and will yield millions or billions of copies of it. (Garibyan and Avashia, 2013) In the forensic science field, the steps for generating DNA profiling includes extraction to separate DNA from proteins, membranes and other cellular material; quantification with UV absorbance to identify the quantity of original DNA sample copies; PCR to amplify the DNA copies, and electrophoresis to separate DNA with different sizes and electrical charge. (Morling, 2009) It used to take 6 to 8 weeks to do a DNA analysis back to the time when PCR has not been discovered, and large samples such as bloodstains in the size of a $1 coin were needed to get enough DNA for profiling. With the benefit from PCR, now the DNA found in a few cells is enough to yield a profile and the time for DNA analysing is reduced to less than a day by loading sufficient DNA samples to a modern genetic analyser. (Alaeddini, 2012)

The process of PCR involves multiple cycles of thermal changes of the sample, each cycle has 3 main phases which ideally double the copies of the DNA at the end of each cycle. The cycle starts with denaturing which heat the sample to 94°C-98°C to break the double-strand DNA to 2 single-strand DNA; after that, annealing/priming happens when the temperature drop to 55°C -70°C to allow the primer to bind to the single-strand DNA; in the end, the extension happens when the temperature is back to 68°C to 72°C so the polymerase can bind to the primer and extend the single-strand to a new double-strand DNA. (Mullis, 1990) The temperature, duration and cycling number only vary slightly depending on the targeted DNA, since certain phases only happen at a certain range of the temperature and there is limited nucleotide in the buffer for the polymerase to construct new copies. However, the protocol of PCR has not been changed much since it was discovered and there are reasons to believe that the process can still be optimised. There are researches been testing different cycling parameters that are most predictive for the success rate of the PCR, (Cordaro et al., 2021) many smart PCR systems use pre-setting to calculate

the parameters for optimizing the process rather than look at the PCR performance itself to analyse the optimized cycling parameters that specific to the sample. The goal of this project is to add a sensor to monitor the process and to use the recorded data to determine whether we can apply machine learning to the machine for optimizing the PCR process. Many studies have been developing machine learning algorithms to predict the result of PCR, and we only found one study that tried to interpret real-time PCR data independent of threshold cycle value to increase the accuracy of qPCR. (Alouani et al., 2021) There is not any study that applied machine learning to optimise the PCR behaviour, there is a necessity to upgrade the PCR technology by increasing specificity and reducing the time consumption to potentially save money and time from optimised PCR experiments. A 55-63% PCR success rate implies that roughly $2.6-3.2 million dollars in materials cost are wasted on failed PCRs and 1 million hours that could be reallocated to each advanced science institute each year. (Cordaro et al., 2021) A real-time PCR, which is also called quantitative PCR (qPCR), can provide the quantification information of the DNA in real-time. The qPCR works by using a fluorescence dye bond to the double-strand DNA, the sample with more double-strand DNA in it will excite more fluorescence; then, an optical detection system is used to detect the fluorescence intensity in each cycle to achieve real-time quantification of the DNA. (Higuchi et al., 1993) The first half of the aim of this thesis is to design a low-cost optical system that can quantify the DNA concentration of the PCR sample in real-time; the second half is to design a real-time PCR machine that uses the optical system. The principle for our optical system is using the percentage fraction between the excitation light intensity and the emission light intensity to quantify the DNA concentration. Since each fluorescence dye will be excited by a certain wavelength of the light then emit light with a longer wavelength, the percentage fraction between the excitation light intensity will vary at different concentrations of DNA. In this thesis, a blue LED is used to excite the PCR product with 6-FAM fluorescence dye (requires 480nm blue dominant light for excitation and emits 520nm green dominant light) and a spectrometer is used to measure the intensity for the excitation light (blue) and the emission light (green). The hypothesis is that the PCR product with a higher DNA

concentration will have a higher Green/Blue(%) fraction than the PCR product with a lower DNA concentration.

This thesis firstly annotates research in related areas such as PCR, PCR parameters, qPCR, optical system and machine learning to address the importance of our project and how can we approach the aim. Then, the methodology and the result are described to determine whether our approach is working. The design part shows a detailed refitting plan that applies the optical system to an open-source PCR along with the physical block diagram and functional block diagram for the overall qPCR design. The comparison between other research and ours for the methodology, result and design are discussed in the discussion and a list of future work is described for the further research of the project.

Literature review

The polymer chain reaction (PCR) was discovered by Mullis KB in 1987, this method of DNA amplification requires only small numbers of the specified DNA to start, then the DNA can be amplified exponentially in cycles of temperature changes. With PCR, scientists can get a large amount of the DNA that they are interested in. (Mullis, 1990) The discovery of PCR has boosted many amazing biological approaches that require DNA profiling, such as Human Genome Project which aims to map all the gene from the human genome. (Garibyan and Avashia, 2013) The PCR assays have been widely used in medical fields such as virology and bacteriology to detect pathogens since the technique can easily amplify naturally obtained DNA to a quantity that is seen as significant for detection and analysis in the lab. The genome for viruses and bacteria is much shorter than humans; therefore, it is convenient to map the virus' and bacteria's genome using PCR with less time and materials. (Garibyan and Avashia, 2013)

In the forensic science field, PCR largely improves the procedure of criminal identification by speeding up the process of DNA profiling. A large sample such as a bloodstain with a size of a $1 coin was needed for sufficient DNA profiling. (Morling, 2009) The invention of PCR has allowed the improvement of a variety of forensic genetic techniques, such as genetic fingerprinting which uses genomic regions that contain short tandem repeats (STRs) to identify individuals. (Morling, 2009) In forensic casework, human remains have usually been identified by PCR application; (Alaeddini, 2012) however, there are sometimes other materials that can contaminate DNA such as soil and recent sediments. Those materials usually inhibit the PCR process from normal functioning when there is adequate DNA present in the sample. The inhibitory activity can affect every element of a PCR reaction such as DNA template, nucleotide and primer; combined impurities could also affect the cell lysis during DNA extraction that the cell might break down abnormally and the DNA extracted will be contaminated. (Alaeddini, 2012) The thought occurred to us that we can program our PCR machine into an adaptive PCR. By measuring the concertration of the DNA in each cycle, if the DNA amplification reachs the factor of two, the machine will automatically move to the next step instead of following the

setup time; if the DNA amplification have been ceased or slowed, the machine can stop imediated and report failure of the PCR process. For example, if the presence of inhibition is detected, our PCR machine can stop the procedure and report inhibition to prevent time and material wasting.

15 parameters that are required for optimizing the PCR during experiments are listed by Cordaro et al. in decreasing order by their contribution level to the PCR optimizing: 'Max primer Tm which is the highest melting temperature for the primer, Amplification Cycles, Total Primer, Extension Temperature, Product Band Size, Minimum GC Clamp Strength which is the lowest strength of the interaction of the 5,3' base pair, Annealing Cycle Temp, Template Amount, Final Extension Time, Melting Cycle Time, Final Extension Cycle Temp, Max Homodimer Tm, Max GC Clamp Strength, Melting Cycle Temp, and Minimum Primer Tm'. (Cordaro et al., 2021) Thermal fisher scientific also listed 6 key considerations relate to cycling parameters for the success and efficient DNA amplification, including template DNA denaturation assessment, DNA denaturation considerations, primer annealing optimization, primer extension considerations, PCR cycle number determination, final extension evaluation. (ThermoFisherScientific, 2017) Those parameters and considerations are closely related to each other, such as the melting temperature Tm for primer annealing optimization is influenced by the GC content of the DNA sample. Those factors are essential which have been proved to have a significant correlation to the specificity and success rate of PCR. It is necessary to test how will the parameters influence the PCR process and what setting of those parameters could optimise the process in certain condition (sample DNA content, size and buffering), and how can we use those data for constructing our future machine learning algorithm.

Russell Higuchi has been researching to monitor the PCR process using fluorescence probes. The research has facilitated quantitative PCR (qPCR) which allow real-time quantification of the DNA during the PCR. (Higuchi et al., 1993) The discovery further enhances the accuracy and specificity of DNA amplification by using fluorescence reporter probes that bind to a specific sequence of DNA. (Deepak et al., 2007) Russell has described a quantitative assay that use a video camera to monitor

multiple PCR reactions during the thermocycling through fluorescence change. The design for the assay is shown in Figure 1 with a camera on top of the thermocycler; the data from the camera is sent to the computer with a Frame grabber for offsite data analysis. Russell's design is the fundamental reference for this thesis since part of the aim for this thesis is upgrading the open-source PCR into a real-time PCR. The linear relationship between the fluorescence and the DNA concentration (shown in Figure 2) is the gold standard for DNA quantification.

Figure 1. Block diagram of video camera system used to monitor amplification in thermocycler. (Higuchi et al., 1993)

Figure 2. The fluorescence of DNA solutions at different concentrations measured using the video-camera set-up. (Higuchi et al., 1993)

A real-time PCR machine usually combines thermal cycler (PCR), optical detection system and instrument software. The optical detection system design differs with different qPCR machines. Figure 3-5 shows schematic graphs for 3 different qPCR optical system designs. (BioRadLaboratories, 2021) All of those designs start with a light source that has a broad wavelength that is filtered by filters to provide the excitation light with a certain wavelength. CFX96 and MiniOpticon use similar designs which they both use LED as the light source and photodiodes for the detector.

Figure removed due to copyright restriction

Figure 3. CFX96 Touch optics shuttle. Six filtered LEDs and six filtered photodiodes are used for excitation and detection. Light is filtered in a narrow bandwidth, ensuring that only data from the desired fluorophores are collected. (BioRadLaboratories, 2021)

Figure removed due to copyright restriction

Figure 4. MiniOpticon™ real-time PCR detection system optics. Forty-eight LEDs fire in rapid succession, illuminating a single sample at a time, while a Fresnel lens focuses each beam directly down into the centre of the corresponding well, minimizing light loss. The emitted fluorescence is split into two beams that pass through separate filters to two sensitive photodiodes. (BioRadLaboratories, 2021)

Figure removed due to copyright restriction

Figure 5. iQ™5 real-time PCR detection system optics. All 96 wells are excited by a combination of narrow-bandpass filters and a tungsten-halogen lamp. Filtered light from the lamp is reflected off mirrors, passes through a condensing lens, and is focused into the centre of each well. Fluorescent light emitted from the wells reflects off the main fold mirror, passes through an emission filter, and is detected by a 12-bit charge-coupled device (CCD). (BioRadLaboratories, 2021)

Figure 6 shows the schematic graph of a qPCR machine using laser and optical fibre and Figure 7 shows the fluorescence intensity at different cycle numbers when running the qPCR machine. The results demonstrate that this system can generate the correctly amplified PCR product and perform complete real-time PCR detection. (Xiang et al., 2007) Figure 8 shows the schematic graph of a low-power and portable qPCR machine that employs LED and multi-channel fluorescence detection. (Hatch, 2014) The machine is a continuous flow real-time PCR that has a unique sample loading and flow system. This design also uses optical fibre as the light-transmitting system which allows transmitting the excitation light into the thermal cycler and transmitting the emission light out of the thermal cycler for detection. Same with the qPCR in *Xiang et al., 2007*. The optical fibre saves a lot of space inside the machine and prevent the light detecting device from heat damage during thermal cycling. And the qPCR result (shown in figure 9) shows a similar trend with *Xiang et al., 2007, in* which the fluorescence intensity increase as the cycle number increase.

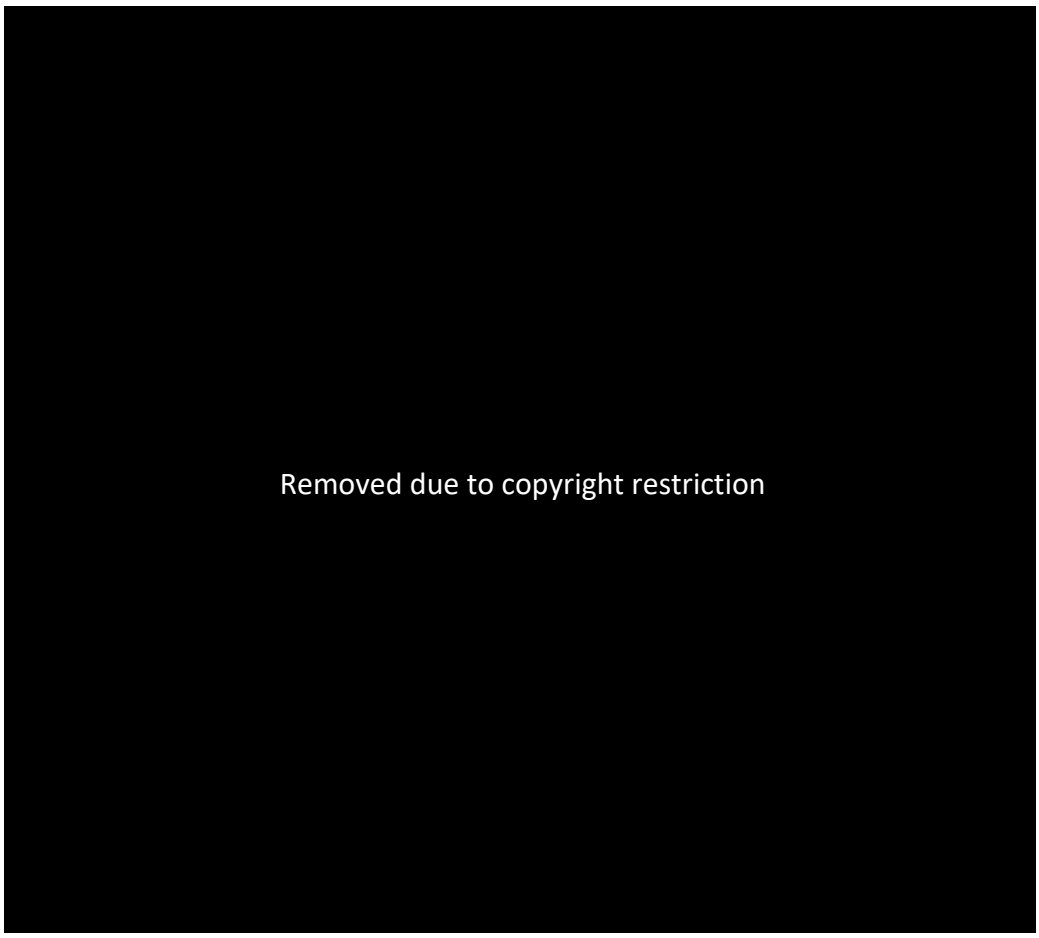

Figure 6. Schematic diagram of the overall qPCR. (a) A schematic diagram of fibre optical fluorescence excitation and detection system. A fibre optical switch was used to realize the multiple-well detection. (b) PCR chip with a fibre embedded in the reaction well. (c) Fibre-lens-coupling devise. (Xiang et al., 2007)

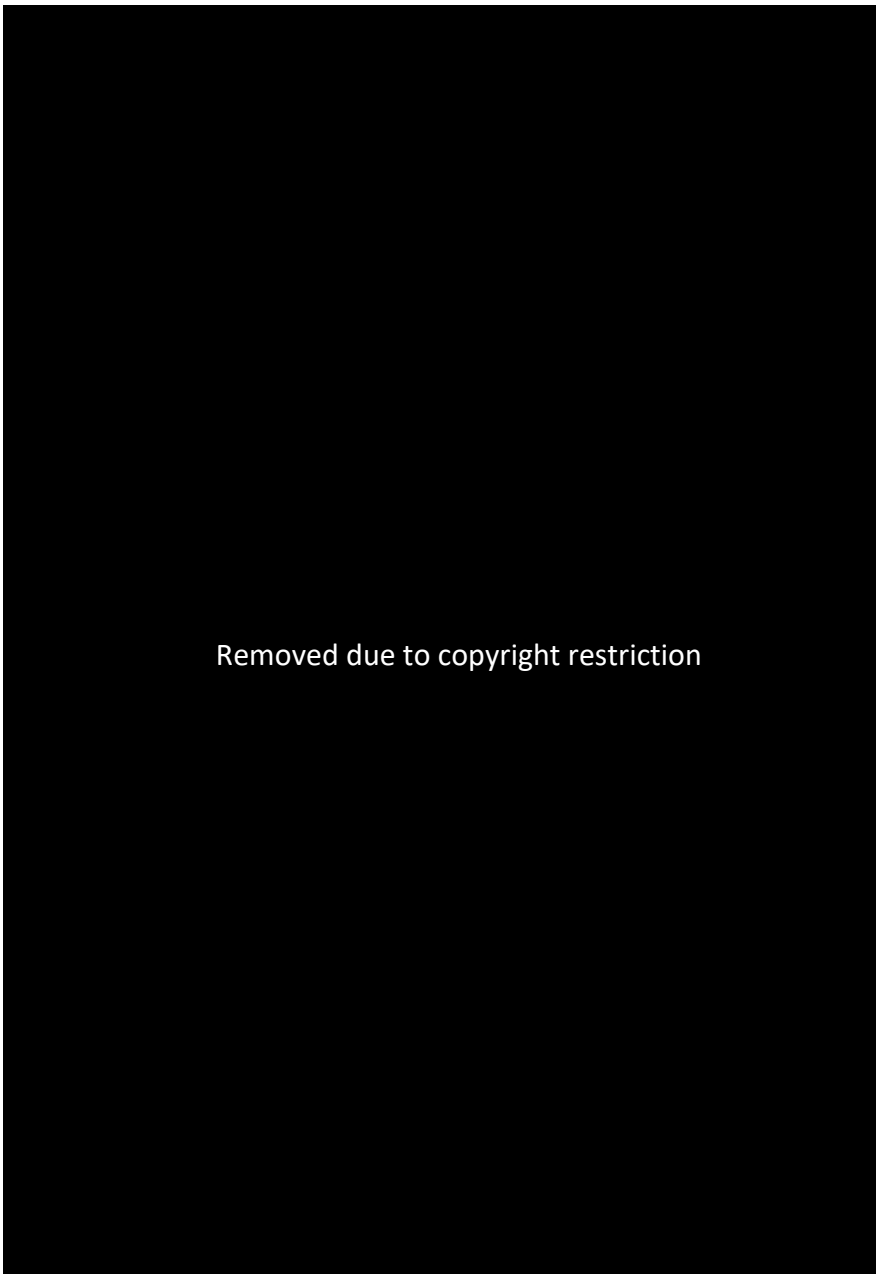

Figure 7. Fluorescence intensity of two runs of real-time PCR with different concentration of the initial DNA template. (a) Fiber-in-well detection method was used during the test. The intensity started to increase earlier for higher initial DNA concentration than lower initial DNA concentration. (b) Fiber-lens-coupling detection method was used during the tests. The results are similar to that using fiber-in-well detection method.

Figure 8. Overview of the system architecture of the real-time PCR instrument. Sample injection takes place through the sample loading device, the plug/droplet selector determines if the sample is run as a single plug (path a) or broken up into multiple droplets (path b), the sample then feeds through the thermocycler while fluorescence detection is measured by the optical components. (Hatch, 2014)

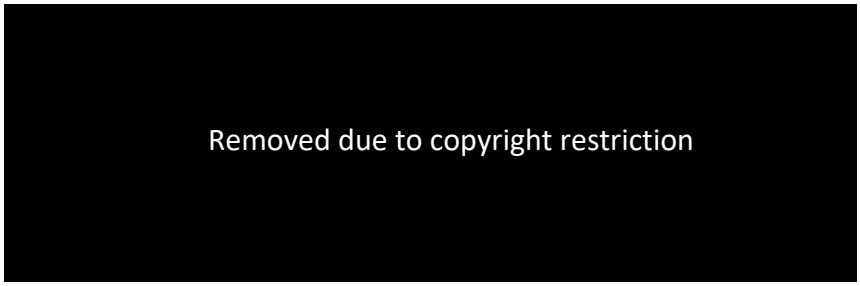

Figure 9. Real-time PCR amplification curves with fluorescence normalisation.

There are various PCR and qPCR machines available on the market. The ProFlex PCR System provides the operator with flexible configuration and control features including assessing the system remotely with a mobile app, interchangeable block formats which supports multiple experiments at once and a simple-to-use touchscreen interface. (FisherScientific, 2021) The VeritiPro Thermal Cycler is another PCR machine that has been widely used in the laboratory due to its temperature optimization through VeriFlex technology, which allows six zones of independent temperature control to optimize PCR temperature with precision. (ThermalFisherScientific, 2021b) Rotor-Gene Q MDx 5plex HRM is a 6-channel qPCR machine that spins the qPCR tubes in a chamber of moving air to keep all samples at the same temperature precisely, the rotary design provides 20 times better thermal accuracy compared with classic block base instrument. (QIAGEN, 2021) Chai Open qPCR is an open-sourced qPCR machine, it only weighs 4 kilograms and pricing at 4,999 American dollars (One of the cheapest qPCR machines) with a single channel and 6,999 American dollars with dual channel. (Chai, 2021)

A machine learning model for optimizing PCR was designed to give a success or failure result assumption of a PCR reaction from all-encompassing input PCR parameter data. (Cordaro et al., 2021) The study claimed that the success rate of PCR can be improved by 18-26% with a standard machine learning approach that transforms the recorded data into relevant PCR parameters. Another trained artificial intelligence algorithm was developed to predict the correct result in real-time PCR for SARS-COV-2. (Langer et al., 2020) The research used several machine learning algorithms available on WEKA data mining software and on Semeion Research Centre depository (Massimo Buscema, Deep Supervised ANNs, Semeion SW #12, version 33.0, 1999–2019) were trained to predict RT-PCR results using the Training and Testing validation protocol, and the best machine learning system has achieved 91.4 % in accuracy, 94.1% in sensitivity and 88.7% in specificity. (Langer et al., 2020) The 2 studies mentioned above are mainly focusing on predicting the success rate and specificity of the PCR process, however, there is not much research on using machine learning to optimize the PCR performance since most study uses machine learning algorithms as a prediction tool for the result. The final goal of the project is to optimise the PCR process during cycles by adding a real-time DNA quantification feature to the PCR machine. The quantification data can be used as feedback to complete the closed-loop system for optimising the PCR process rather than one setting fits for all. A recent study trained a deep learning model to analyse the fluorescence data as a global image during the real-time PCR independent of the threshold cycle (C_T), which C_T is traditionally used as the interpretation of the qPCR data. (Alouani et al., 2021) Noise early in the PCR process may lead to incorrect threshold establishment that causes errors in the threshold cycle. The author believes that the novel approach is potentially more accurate than the interpretation of C_T. This study is using a new approach to analyse qPCR data which can potentially improve the qPCR data analysis by avoiding the error from threshold quantitative interpretation. (Alouani et al., 2021)

Methodology

Step 1: Choose the right PCR machine

First of all, this project requires a PCR machine or a qPCR machine to perform the PCR. Many commercialised PCR/qPCR machines have great accuracy and control; however, those machines like ProFlex PCR System and VeritiPro Thermal Cycler are closed systems that cannot be accessed or modified by the user. Thus, we need an open-sourced machine to acquire all the information during the PCR progress and program the machine to make modifications. As shown from the left part of figure 10, the machine is called Ninja PCR which was created by a Japanese couple. We selected this device because it is cheap and The machine is running as a normal thermal cycler with a 4x4 well PCR holder. The machine is open source which means the design details and code are all available online. The software for the machine is Arduino based which can be modified for personal research. The basic PCR protocol was set as 95° C for denaturing temperature, 58° C for annealing temperature, 72° C for extending temperature, the step duration for each procedure is 30 seconds per cycle, the ramp duration (time taken for getting to the next temperature) is 10 seconds, 30 as the number of cycles. The temperature changing curve can be visualised from the right. The heater on the top lid is maintained at around 110 ° C (shown in blue line) to prevent condensation on the lid of the PCR tube. There are 2 temperature sensors in the well which the red line represents the temperature for the sample and the green line represent the temperature for the well.

Step 2: Design the optical detection system

As mentioned before, there are many choices for the optical detection system, such as using laser or LED with filters as the light source and photodiodes as the light sensor. We did consider using those designs; however, the laser requires a power output that excesses the Australian law and Flinders University standards, and the filters are too big to be put within an existing PCR machine and the professional filters with a narrow wavelength range are quite expensive. Thus, we chose to use blue LEDs with 470nm dominant wavelength as the light source so we do not need the filter for the light source, and spectrometers with build-in filters are chosen to detect the emission light intensity from different light channels. (so no filter need for

the emission light) The spectrometer is Adafruit AS7341 10-Channel Light / Colour Sensor Breakout which contains 8 visible light channels from 415nm to 680nm, 1 clear and 1 IR channel. We selected this sensor since it has a narrow wavelength range for each channel and it covers all the visible light wavelengths, and it is much cheaper than the professional VIS-UV spectrometer.

Step 3: Design the testing tool

A 3D printed experiment tool is made to mimic the designed DNA quantification process. (Shown in figure 11) The reason we created this tool is that we cannot have a stable positioning with the spectrometer and the PCR tube to do the measurement without refitting the PCR machine, the 3D printed tool is used for the proof of concept whether our optical system can real-time quantify the DNA outside the PCR machine. The top left is the bottom block which provides stability to the whole structure and allows the wire of LED goes through the bottom. The top right picture is the middle block which is on top of the bottom block which holds the blue LED to allow it to illuminate upwards. The bottom left is a top block that holds the PCR tube and allows the light from the bottom to go through the tube. The bottom right picture is the full 3D print with the spectrometer on the side to analyse the light refracts and emit from the PCR tube at a 90-degree angle. Since the 3D print material is semi-transparent, the surfaces of the blocks are painted black to reduce the light refraction within the 3D print.

Step 4: Prepare PCR product samples

To prove whether our optical system can quantify the DNA concentration, PCR samples with different DNA concentrations are needed to be tested. A pack of PCR products with different concentrations of DNA and different volumes was provided by Forensic science at Flinders University. (see Figure 12) The first 4 PCR tubes on the first row of the tube holder contain 5µl of neat human DNA PCR product sample. The 5th-9th PCR tubes on the first row each contains 20µl of neat human DNA PCR product sample. The first 4 PCR tubes on the 5th row each contains 100µl sample 1 human DNA PCR product diluted with distilled water, the concentration of DNA decrease from left to right with neat, 1:2, 1:4 and 1:8. The concentration gradient is diluted by a factor of two to mimic the concentration changes through cycles of the

PCR process. (The concentration of the DNA will be ideally doubled in each cycle) The 5th PCR tubes on the 5th raw are controlled with 100 µl distilled water. The 4 PCR tubes on the 7th row each contains 100µl sample 2 human DNA PCR product diluted with distilled water, the concentration of DNA decrease from left to right with neat, 1:2, 1:4 and 1:8.

Step 5: Conduct the experiment

Each of the PCR tubes was loaded to the 3D printed experiment tool for the spectrometer reading with 10 repeats, each repeat is performed by removing the tube from the tool and loading again to determine the level of consistency in the measurement. The data analysis is done by excel to find the percentage fraction between green and blue light, calculated by $\frac{I_{515nm}+I_{555nm}}{I_{445nm}+I_{480nm}} \times 100$ with I stand for light intensity. The average percentage fraction is used for graph display and the data analysis is run by a data analysis tool in excel. Another method by measuring the emission light (green light) intensity only ($I_{515nm} + I_{555nm}$) to compare with the Green/Blue(%). This experiment was designed to determine whether it is possible to detect different amplified DNA amounts using real-time feedback. Therefore, as a proof-of-concept, we had DNA PCR products generated with a 6-FAM fluorescent label at different concentrations, and a control with no DNA. If the optical system is going to work, a positive trend of fluorescence and DNA concentration should be able to be observed.

Figure removed due to copyright restriction

Figure 10. Ninja PCR machine(left) and performance during PCR (right).

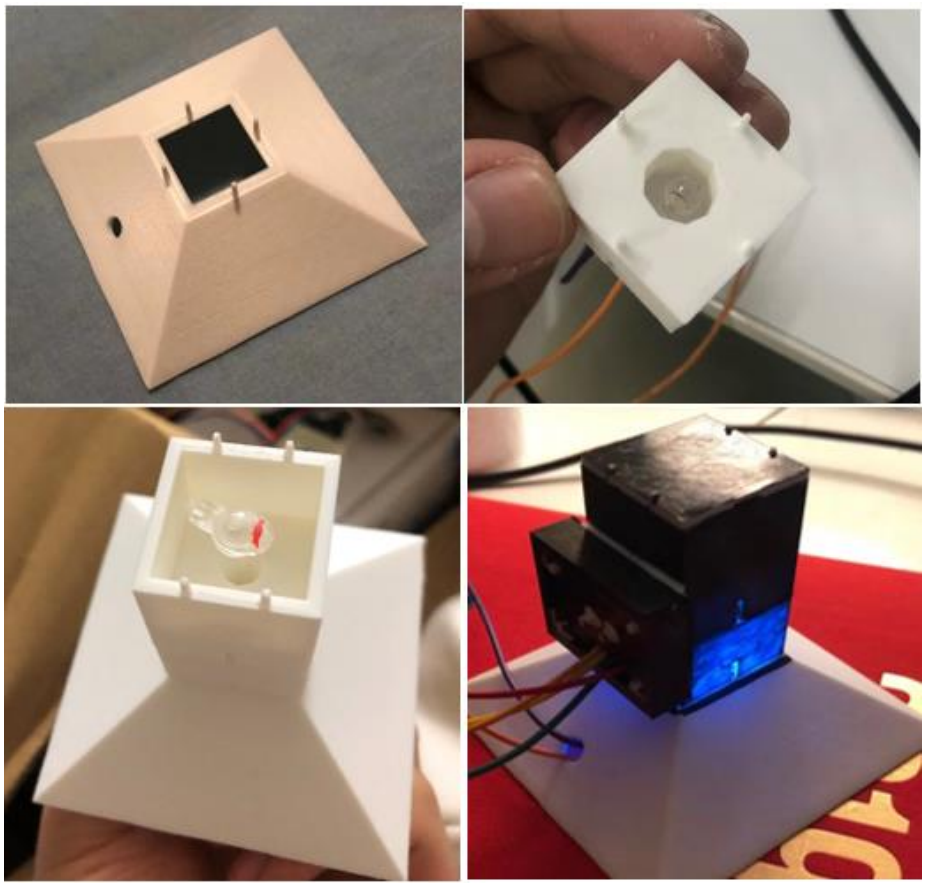

Figure 11. 3D printed experiment tool.

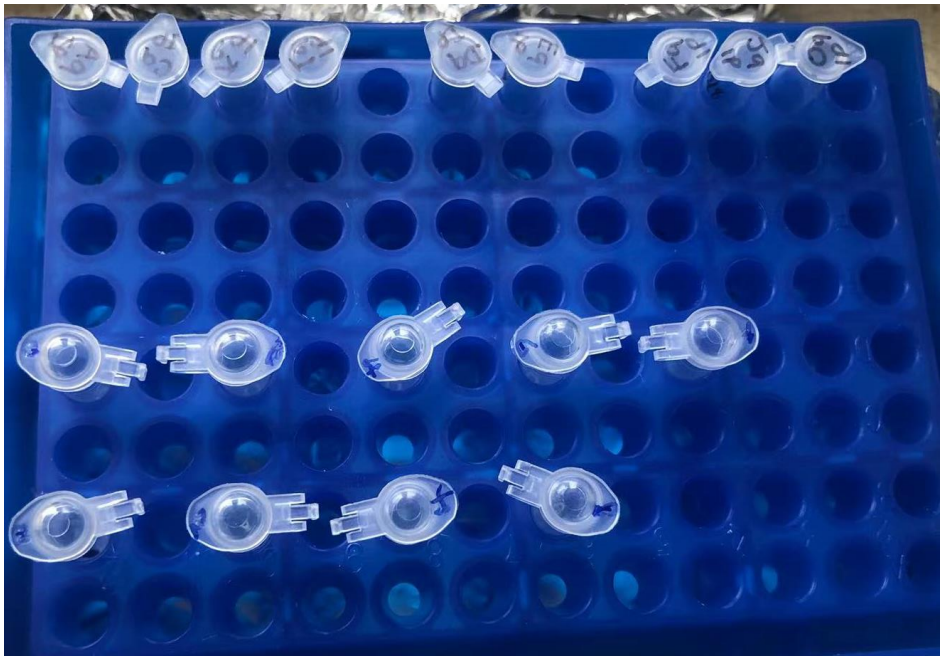

Figure 12. PCR products with different concentrations of DNA and different volume.

Result

High accuracy of the result

The average Green/Blue(%) value was graphed with box and whisker plot for the 8 samples with neat PCR product with 5µl and 20 µl (to determine the influence of loading volume). (Shown in Figure 13) The bottom error bar for each sample represents the first quartile and the top error bar represent the third quartile, the middle line is the median and the cross is the mean. There is no sample with a big interquartile range that have been found since the error bar represents the interquartile range of each sample, which means the measurement method was providing consistent results and the error is small. The low 95% confidence level (with maximum 0.087 and minimum 0.051) and low standard error (with maximum 0.039) for each sample in Table 1 is small also indicate the high accuracy of the result.

Finding minimum sample volume

The result indicates that the first four samples with 5µl of neat human DNA have a 1.4 faction percentage mean difference between the four 5µl neat samples while there is only 0.34 faction percentage mean difference between the 5 20µl neat samples. The P-value between any of the four 5µl neat samples is much less than 5% which indicates that there is a significant difference between the result. The P-value for sample Dq&Eq, Fq&Jq is higher than 5% which means there is no significant difference between the Dq&Eq group and Fq&Jq group. The 20µl group produce gives a more constant and accurate reading compared to the 5µl group. This suggests that the method we use to measure the sample DNA quantification might not be able to provide a valid representation for the DNA concentration when the volume of the sample is too small. The minimum volume of the sample would be less or equal to 20µl and more than 5µl, future experiments with more samples with volumes between 5µl and 20µl will identify the minimum sample volume for the device.

The Green/Blue(%) drops as the decreased concentration of DNA

Figure 14 shows the bar chart comparison of Green/Blue(%) among the 2 samples of 100 µl PCR product with DNA concentration in a downgradient (neat, 1:2, 1:4 and 1:8) and control (with no DNA). Blue is sample 1, orange is sample 2 and green with a dashed line is control. The error bar for each column represents the 95% CI. The Green/Blue(%)faction for control is less than 4% while the neat samples have an average of 6%. There is a clear trend that the Green/Blue(%) drops with the decreasing concentration of DNA for both sample 1 and sample 2, and the control have the lowest Green/Blue(%) compared to any other 100 µl samples. There is a significant Green/Blue(%) difference in samples with different concentrations of DNA from neat to control (P-value much less than 5%).

The green light intensity increase as the increased concentration of DNA

Figure 15 shows the scatter chart of the green light intensity with increasing DNA concentration. The x-axis is the standard concentration which the neat is 1 and the control is 0. The emission light (green light) intensity is calculated by the sum of I_{515nm} and I_{555nm}. The error bar stands for the 95 confidence interval of each sample. There is a clear trend that the green light intensity increases as the DNA concentration increases. There is a significant difference between the green light intensity in different concentrations of DNA from 0 to 1 (P-value much less than 5%).

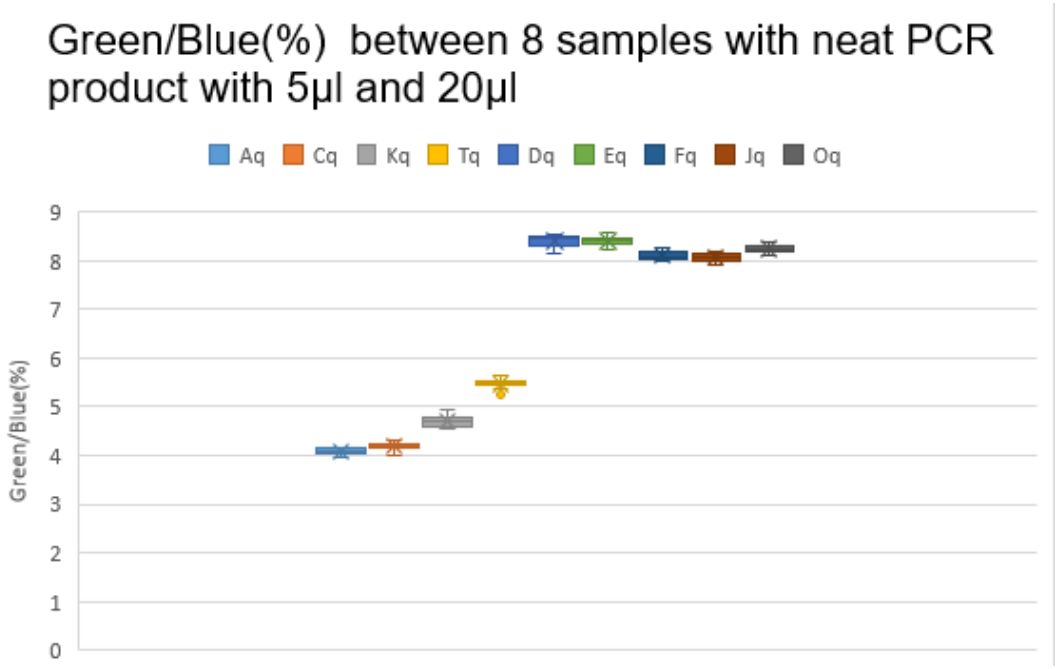

Figure 13. Green/Blue (%) between 8 samples with neat PCR product. Tube Aq, Cq, Kq and Tq with 5 µl neat PCR product; tube Dq, Eq, Fq Jq and Oq with 20 µl neat PCR product. E each sample was assigned a letter with a 'q' for labelling.

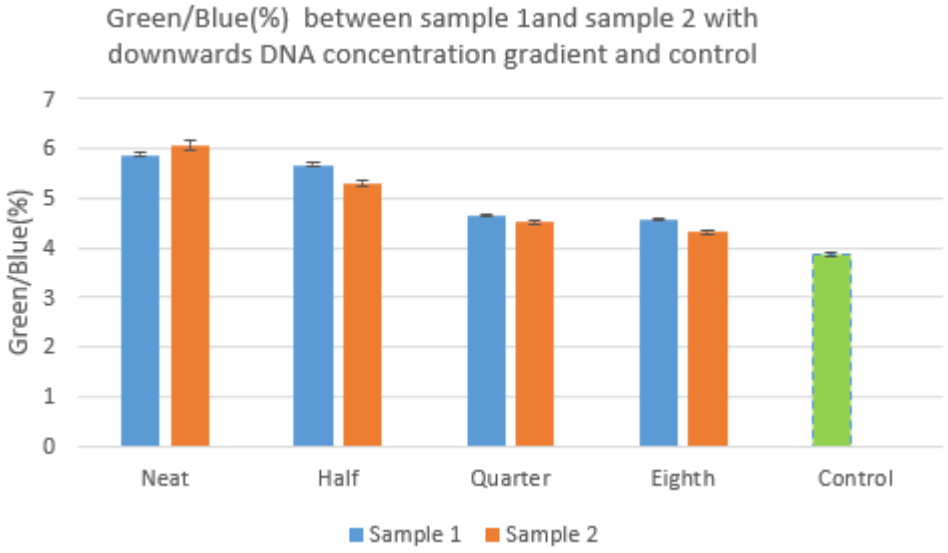

Figure 14. Green/Blue(%) between sample 1 and sample 2 with downwards DNA concentration gradient and control.

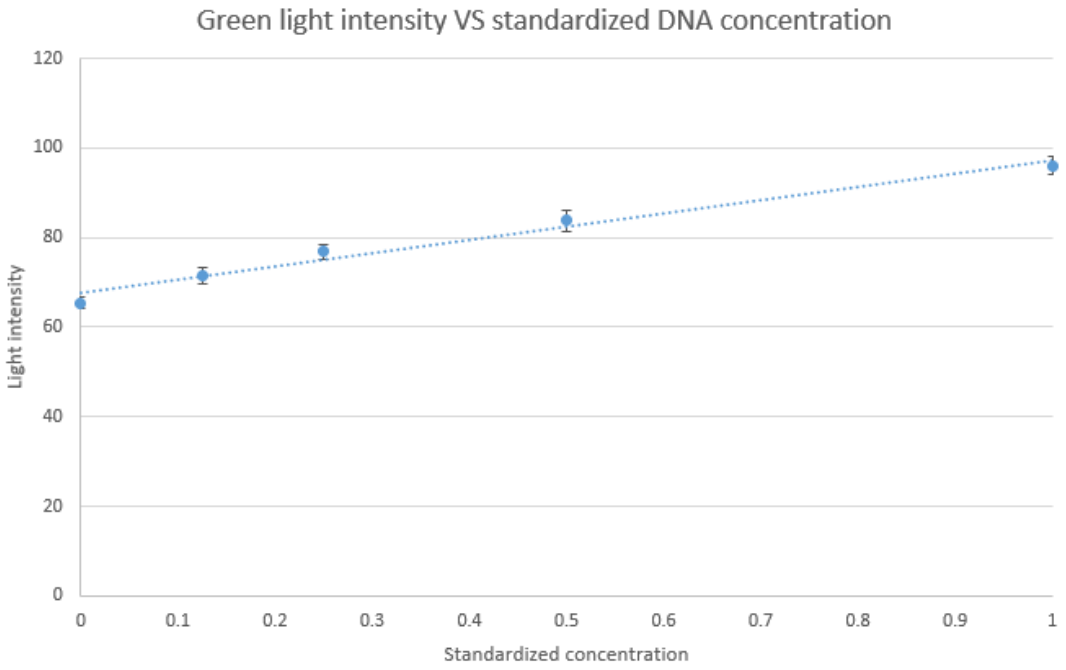

Figure 15. Green light intensity VS standardized DNA concentration.

Table 1. Statistic data analysis for neat samples.

	Aq	*Cq*	*Kq*	*Tq*	*Dq*	*Eq*	*Fq*	*Jq*	*Oq*
Mean	4.078665	4.18938	4.705482	5.472658	8.404925	8.405986	8.107894	8.061152	8.242341
Standard Error	0.022585	0.02485	0.037434	0.031082	0.038535	0.030459	0.026641	0.032485	0.025291
Median	4.088389	4.206782	4.686922	5.484169	8.463384	8.411617	8.085737	8.065261	8.23506
Mode	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A
Standard Deviation	0.071419	0.078581	0.118376	0.098291	0.121858	0.096319	0.084246	0.102728	0.079977
Sample Variance	0.005101	0.006175	0.014013	0.009661	0.014849	0.009277	0.007097	0.010553	0.006396
Kurtosis	-0.603131	2.00636	-0.316311	1.554414	0.832789	0.984973	-0.985734	-1.551063	-0.800282
Skewness	-0.536957	-1.047022	0.617008	-0.783336	-1.131409	-0.181124	0.41193	-0.242739	0.061885
Range	0.217009	0.288439	0.378833	0.36163	0.391388	0.355391	0.252405	0.290966	0.250353
Minimum	3.959013	4.015865	4.552673	5.266075	8.145197	8.222698	7.993668	7.895898	8.114133
Maximum	4.176022	4.304304	4.931507	5.627706	8.536585	8.578089	8.246073	8.186864	8.364486
Sum	40.78665	41.8938	47.05482	54.72658	84.04925	84.05986	81.07894	80.61152	82.42341
Count	10	10	10	10	10	10	10	10	10
Confidence Level(95.0%)	0.05109	0.056214	0.084681	0.070313	0.087172	0.068903	0.060266	0.073487	0.057212

Design

Schematic diagrams interpretation

Figure 16 shows a cartoon schematic of the cross-section view of our design. A hole is drilled downwards from the bottom of each PCR tube slot through the bottom of the well which is attached to the heating pad with thermal glue. A blue SMD LED is placed on the heating pad below each slot so the light illuminate from the LED can go through the drilled hole and reach the PCR tube. The PCR mix within each PCR tube will be excited by the blue light from the LED and emit emission light. A drill horizontal pathway is drilled from the side of the well to each PCR tube slot at the height of the PCR mix liquid level, which allows the emission light from the PCR mix to be transmitted out of the well by optical fibre and be detected by a spectrometer. This system is applied to each PCR tube slot that is assigned for real-time quantification. The assigned slots can be found in Figure 16 with the overall design for the whole well and the circuit design for the LED.

Figure 17 shows a top view schematic diagram of the design with a circuit design of the SMD LEDs. 8 of 16 PCR tube slots are assigned for real-time quantification with the detecting system from figure 16. The 8 SMD LEDs are connected in series with a protection resistor and powered by a 24 V power supply. (The calculation for power supply is shown in power Appendix 2) Digital pin 2 of the Arduino board is used for controlling the on/off of the circuit by using a transistor connected to the circuit.

A 3D graph of our design is shown in figure 18 which was graphed with Inventor 2022. The 3D graph shows an overall 3D design with a real scale. The green fibre is the optical fibre that connects the slots with the spectrometers. The SMD LEDs are below the PCR tube slots which can be better viewed in figure 19 which gives details including the drilled pathway for the LEDs and the optical fibre.

Figure removed due to copyright restriction

Figure 16. Cartoon schematic diagram for the optical system design.

Figure removed due to copyright restriction

24V

Figure 17. Top view schematic diagram with a circuit design of the SMD LEDs.

Figure removed due to copyright restriction

Figure 18. CAD schematic diagram for the design.

Figure removed due to copyright restriction

Figure 19. CAD schematic diagram magnified.

Block diagram interpretations

A physical block diagram shown in figure 20 demonstrates the communication and connection between each part of the design. The Arduino Leonards control the on/off of the SMD LEDs which provides the excitation light for the PCR tube holder (the well). The emission light is transmitted to AS7341 spectrometers for detection through the optical fibre. The spectrometer data is sent to Arduino. The machine

learning algorithm will analyse the data in each cycle and find the best setting for the next cycle; then, the Arduino will send the instruction to the Ninja PCB through the I2C bus (microcontroller of the Ninja PCR) to modify the parameters (such as increase/decrease extension time/temperature) of the next heating cycle. In the end, the heating pad will perform the new setting and the PCR process will be optimized.

The functional block diagram (shown in figure 21) is constructed to show the planning out for the overall system of the designed real-time PCR. Once the machine is turned on, the machine will initialise the setting and start a Wi-fi connection. Once the executing order is received from the user, the LED and Spectrometer will be turned on to initialise the optical system. Then, the baseline for the PCR mix fluorescence will be set and some PCR parameters will be set to the machine. Once the PCR process starts, the quantification of the DNA will be continues monitored and displayed. The quantification data will be analysed in each cycle and the optimised setting for the next cycle will be applied after each cycle. Once the PCR process is finished, the LEDs and the spectrometers will be turned off and the critical experiment data will be recorded to the database for the machine to improve its algorism for optimizing the performance for the next experiment.

Figure removed due to copyright restriction

Figure 20. Physical block diagram for the design.

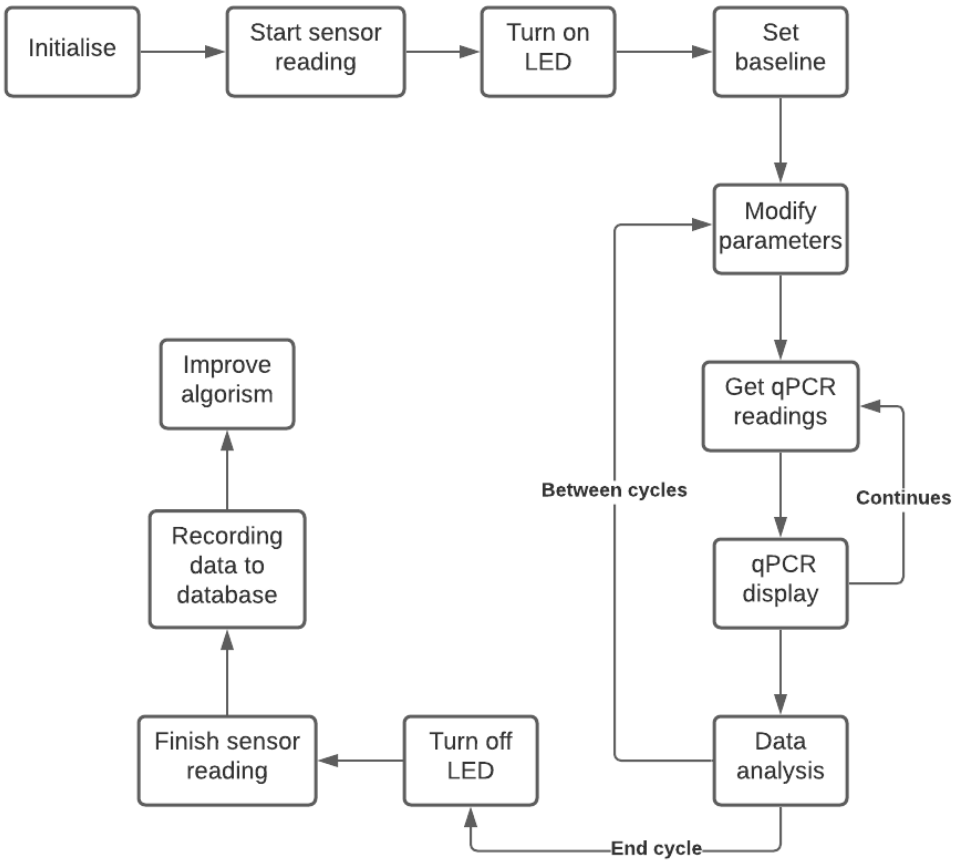

Figure 21. Functional block diagram for the design.

Discussion

Result comparison

Construction relationship between the two measurements and DNA concentration

Since the study by Russell, 1993 and the green light intensity measurement in this thesis are both measuring the emission light intensity, the results should be matching with each other. Our green light intensity measurement (figure 15) does have a linear trending line that the green light fluorescence is proportional to the DNA concentration, which matches with the fluorescence vs DNA concentration graph by Russell (shown in Figure 2). There is a need for experiment data from more concentrations of DNA to construct a correlation between green light intensity and DNA concentration.

The result with Green/blue(%) (shown in figure 14) shares some similarities to the real-time PCR amplification curves through cycle numbers Xiang et al., 2007 (Figure 7) and Hatch, 2014 (Figure 9), since the DNA concentration is ideally multiplied by a factor of 2 in each cycle and our concentration is diluted by 2 down through the gradient. Even though our result shows a positive relationship between the fluorescence and the DNA concentration (which match with the 2 real-time qPCR curves from the studies), the sample size of our result is too small to give a conclusion. There is a need for more data with different concentrations of DNA to construct a correlation between Green/Blue(%) and DNA concentration.

Comparison between the Green/Blue(%) measurement and green light intensity measurement

The reason we were using Green/Blue(%)is that our spectrometer can detect not only the emission light wavelength range but also the excitation light wavelength range. More excitation light is absorbed and more emission light is emitted when the DNA concentration increases, (The numerator increases and the denominator decreases, will produce a larger result) using the fraction ratio will have a better contrast of the result between different DNA concentrations; and the error created by the minor difference between PCR tubes (transparency/loading positions) can be

minimised (both numerator and the denominator multiplied by a percentage, the result will not change).

However, the error usually is magnified along with the contrast in fraction measurement. The emission light intensity measurement might provide better accuracy compared with the fraction measurement. As we can see from figure 14, the Green/blue(%) hardly have any difference between the quarter and the eighth concentration (can be due to error) while in figure 15 the green light intensity for 0.25 has a significant increase compared with 0.125 (P-value much less than 5%).

Compare to Qubit fluorometer

The optical detection system in qPCR is quite similar to fluorometers since they both measures the fluorescence dye bind to the DNA. One example is the Qubit fluorometer which uses PicoGreen assay. (ThermalFisherScientific, 2021a) Red and Blue LEDs are used for the light source and filters was applied to limit the wavelength each for excitation light and emission light. The system specification for Quibit is shown in Figure 22. The 2 channels of the excitation filters and the 2 channels of the emission light channels are covered by the channels of our spectrometer. This means our device might produce similar accuracy with Qubit if we are using a red and blue LED light source with PicoGreen dye. However, Nanodrop spectrophotometric which measures UV absorption was found more accurate on DNA quantification when compared with Qubit with the reference quantification (qPCR). (Haque et al., 2003) Similar research also shows Qubit 3.0 fluorometer is not always the best method for PCR DNA quantification when the result was compared with NanoDrop UV spectroscopy and a quantitative PCR. (Nakayama et al., 2016) This indicates that we can use a similar method in the future to determine the accuracy of our real-time PCR device with a spectrophotometer like Nanodrop and a professional qPCR. However, the low accuracy of the Qubit fluorometer might suggest that we need to employ more channels of excitation light and emission light to increase the accuracy of our device.

System specifications

Instrument dimensions	5.4 in (w) x 10 in (l) x 2.2 in (h) (13.6 cm x 25 cm x 5.5 cm)
Weight	743 g
Dynamic range	5 orders of magnitude
Processing time	≤5 seconds/sample
Light sources	Blue LED (max ~470 nm) Red LED (max ~635 nm)
Excitation filters	Blue 430–495 nm Red 600–645 nm
Emission filters	Green 510–580 nm Red 665–720 nm
Detectors	Photodiodes: measurement capability from 300–1,000 nm
Warm-up time	<35 seconds
USB drive	4 GB

Figure 22.System specification for Qubit 4.0.

Design comparison

The lining of the sensor to the direction of excitation light

The first attempt of this project was putting a spectrometer at the top of the tube holder rather than a 90 degrees lining, just like the lining from *Rusell, 1993*. (Shown in figure 1) However, the top space is limited for the open-source PCR machine used in this thesis due to the top lid heater (for preventing the condensation at the top of the PCR tube). In addition, since the design for this thesis involves a blue LED illuminate from the bottom of the well, it is better to detect the emission light from a 90-degree angle rather than straight upwards to avoid the saturation of the sensor. The light sensor saturation usually happens when the sensor is exposed directly to the light source, which makes the sensor insensitive to the small change of the input. Especially when the emission light (green) have much less intensity than the excitation light (blue).

Light source and light-transmitting system selection

Similar to the design in CFX96 and Mini (see Figure 3 and 4), this thesis is using an LED as the light source since LEDs are power efficient and the size of an LED can be

small. Blue SMD LED with a dominant wavelength of 470nm is put at the bottom of the well, so minimum space is used within the well and no filter is needed for excitation light. The design with laser with optical fibre which was used by *Xiang et al., 2007* (see Figure 6) could have been another good fit for the project since lasers have very narrow spectra and it transmits well within the optical fibre. However, due to the legal requirement in the faculty and South Australia, the 60mW cyan blue laser is not allowed. (Legal laser with power no larger than 1mW) Nevertheless, the optical fibre transmitting system like *Xiang et al., 2007* and *Hatch, 2014* (see Figure 8) would be a good fit for our machine, since transmitting the emission light out of the thermal cycler for detection will induce minimal change to the pre-existing structure and prevent the detector from thermal damage.

Other design details

The reason why we only plan to include half of the PCR tube hold slots as the qPCR slots are mainly due to the I2C bus limitation in pull-up resistor and limitation to the address. The l2C limitation and solution are described in detail in Appendix 1.

The Arduino board can support all the spectrometers just by its 3.3V power supply, the power budget for the optical system is calculated in Appendix 2.

The choice for optical fibre is Ø200 µm Core Step-Index Multimode Fiber, 0.50 NA, with a total diameter of 925 µm, which can fit inside the drilling hole for optical fibre with a diameter of 1 mm.

The cost estimation is shown in table 2. The cost for each machine will be 1200 Australian dollars, the cost is calculated by the sum for each component from our design and the workmanship that requires to build a qPCR. Since the cheapest qPCR on the market is Chai open qPCR which sale for 4999 American dollars (worth 6700 Australian dollars), our device is only 18% of the price of the cheapest qPCR on the market.

Table 2. Cost estimation for each qPCR machine.

Item	Price/Unit	Number	Total
Ninjar PCR	$670	1	$670
Adafruit AS7341	$31.15	8	$249.20
SMD LED	$0.25	8	$2
Optical fibre	$2/meter	2	4
Arduino Leonardo	$30	1	$30
TCA9548A Multiplexer	$15	1	$15
24V power supply	$30	1	$30
Other accessories			$200
charge for workmanship	$50/hour	4	$200
Total			$1,200

Limitation

This project is adding a sensor to an open-source PCR to achieve real-time PCR, which means the machine is only designed for PCR. We could not put too many things inside the fixed space otherwise we would influence the function of the PCR. The PCR product we got from forensic science Flinders is only from the control human DNA template, and we dilute the PCR mix with distilled water which means the concentration of unbounded fluorescence and other content in the solution is diluted as well. It would be better if we use samples from the same PCR content but with different cycle numbers (with measured concentration) to construct the correlation between Green/Blue(%) to DNA concentration. Also, the same cycle numbers but different amounts of DNA can be put into the PC in the beginning so that we can have different amounts of PCR product at the end.

Conclusion

Future work

First of all, more experiment with more samples (with measured concentration) is needed for constructing the correlation between Green/Blue(%) and green light intensity to DNA concentration. A comparison of the reading with a professional Flame UV-VIS Spectrometer and a professional qPCR would also be useful for validating the result.

Secondly, there is a need to increase the accuracy of our DNA quantification method. The Qubit have demonstrated a method to quantify PCR DNA by using 2 channels of spectra (blue and red), which enlighten a future direction of the project to use more channels to improve the quantification method. The spectrometer that is used in this thesis is Adafruit AS7341 which have 8 visible light channels, 1 clear and 1 IR channel, it does cover all the wavelength channel that the Qubit fluorometer use. Additionally, the current PCR analysis kit used for this thesis is the GlobalFiler 6 dyes STR kit (The Fluorescence Emission Spectra of 6 Dyes Present in the GlobalFiler STR Kit can be seen in figure 23); therefore, all the emission wavelengths have the corresponding channel with our spectrometer. If we can construct the relationship with the Emission/Excitation(%) and emission light intensity for each dye, an algorithm with all 6 dyes may provide more accuracy in DNA quantification. We expect a better result with more channels because the professional PCR machines use by the Forensic science department at Flinders University uses channels for all 6 dyes in GlobalFillller Kit to quantify the DNA sample and the machine is known to have high accuracy.

Furthermore, it is essential to refit the machine as we designed and run some performance tests to make sure the refitting won't break the original function of the PCR. In addition, thermal analysis is needed to run with added components such as the optical fibre and SMD LED. According to the datasheet of the SMD LED, its operating temperature is -40 - 100 °C. Even though the operating temperature for the PCR is 20-95°C, the SMD LED can be heated up to more than 100 °C since it sits

on top of the heating pad. The thermal analysis would help us find out the wearing-off rate for each component so we can determine if we need to find a substitute.

Lastly, we need to develop our own machine learning algorithm. Since most machine learning algorithms in smart PCR machines are only used for predicting the result, we need to do more experiments on the PCR parameters to develop an AI to optimize the PCR process. According to Alouani et al., 2021, the non-threshold reading could also interpret useful information of the PCR process, which occur to us that we can apply this method to our future qPCR machine to focus on the whole process rather than the traditional CT. There might be more information about the cycling condition that can be derived from the interpretation that we can use to construct our machine learning algorithm.

Figure removed due to copyright restriction

Figure 23. The Fluorescence Emission Spectra of 6 Dyes Present in the GlobalFiler STR Kit.

Summary

The experiment shows a promising result that our optical system designed for real-time quantification of the DNA is able to distinguish DNA samples with different concentrations. However, a minimum loading volume range between 5 µl and 20µl is required for the system to provide valid measurements.

The low-cost real-time PCR we developed here have a great potential to perform in a similar manner with a professional qPCR machine that measures all 6 dyes. Once the qPCR is successfully been built and programmed, experiment with different PCR parameters can be conducted to sketch the machine learning algorithm for optimising the PCR behaviour.

Reference

ARDUINO. 2021. *Arduino Leonardo* [Online]. Available: https://www.arduino.cc/en/Main/Arduino_BoardLeonardo [Accessed].

BLUEDOT. *How many Devices can you Connect to the I2C Bus?* [Online]. Available: https://www.bluedot.space/tutorials/how-many-devices-can-you-connect-on-i2c-bus/ [Accessed 2021].

DRONEBOTWORKSHOP. 2021. *Multiple I2C Buses with an Arduino* [Online]. Available: https://dronebotworkshop.com/multiple-i2c-bus/ [Accessed 2021].

ALAEDDINI, R. 2012. Forensic implications of PCR inhibition—A review. *Forensic Science International: Genetics,* 6**,** 297-305.

ALOUANI, D. J., RAJAPAKSHA, R. R. P., JANI, M., RHOADS, D. D., SADRI, N. & DEKKER, J. P. 2021. Specificity of SARS-CoV-2 Real-Time PCR Improved by Deep Learning Analysis. *Journal of Clinical Microbiology,* 59**,** e02959-20.

ARDUINO. 2021. *Arduino Leonardo* [Online]. Available: https://www.arduino.cc/en/Main/Arduino_BoardLeonardo [Accessed].

BIORADLABORATORIES. 2021. *Introduction to qPCR Instrumentation* [Online]. Available: https://www.bio-rad.com/en-au/applications-technologies/introduction-qpcr-instrumentation?ID=LUSO5YMNI [Accessed September 2021].

BLUEDOT. *How many Devices can you Connect to the I2C Bus?* [Online]. Available: https://www.bluedot.space/tutorials/how-many-devices-can-you-connect-on-i2c-bus/ [Accessed 2021].

CHAI. 2021. *Chai open qPCR* [Online]. Available: https://www.chaibio.com/openqpcr [Accessed Dec 2021].

CORDARO, N. J., KAVRAN, A. J., SMALLEGAN, M., PALACIO, M., LAMMER, N., BRANT, T. S., DUMONT, V., GARCIA, N. D., MILLER, S., JOURABCHI, T., SAWYER, S. L. & CLAUSET, A. 2021. Optimizing polymerase chain reaction (PCR) using machine learning. *bioRxiv***,** 2021.08.12.455589.

DEEPAK, S., KOTTAPALLI, K., RAKWAL, R., OROS, G., RANGAPPA, K., IWAHASHI, H., MASUO, Y. & AGRAWAL, G. 2007. Real-Time PCR: Revolutionizing Detection and Expression Analysis of Genes. *Current genomics,* 8**,** 234-251.

DRONEBOTWORKSHOP. 2021. *Multiple I2C Buses with an Arduino* [Online]. Available: https://dronebotworkshop.com/multiple-i2c-bus/ [Accessed 2021].

FISHERSCIENTIFIC. 2021. *Applied Biosystems™ ProFlex™ PCR System, 2 x 96-well* [Online]. Available: https://www.fishersci.com/shop/products/proflex-pcr-system-extended-warranty-package-2-x-96-well-5/4484076 [Accessed Dec 2021].

GARIBYAN, L. & AVASHIA, N. 2013. Polymerase chain reaction. *J Invest Dermatol,* 133**,** 1-4.

HAQUE, K. A., PFEIFFER, R. M., BEERMAN, M. B., STRUEWING, J. P., CHANOCK, S. J. & BERGEN, A. W. 2003. Performance of high-throughput DNA quantification methods. *BMC Biotechnology,* 3**,** 20.

HATCH, A. C., RAY, T., LINTECUM, K., YOUNGBULL, C. 2014. Continuous flow real-time PCR device using multi-channel fluorescence excitation and detection. *Lab on a Chip 14,* 562–568.

HIGUCHI, R., FOCKLER, C., DOLLINGER, G. & WATSON, R. 1993. Kinetic PCR Analysis: Real-time Monitoring of DNA Amplification Reactions. *Bio/Technology,* 11**,** 1026-1030.

LANGER, T., FAVARATO, M., GIUDICI, R., BASSI, G., GARBERI, R., VILLA, F., GAY, H., ZEDURI, A., BRAGAGNOLO, S., MOLTENI, A., BERETTA, A., CORRADIN, M., MORENO, M., VISMARA, C., PERNO, C. F., BUSCEMA, M., GROSSI, E. & FUMAGALLI, R. 2020. Development of machine learning models to predict RT-PCR results for severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) in patients with influenza-like

symptoms using only basic clinical data. *Scandinavian Journal of Trauma, Resuscitation and Emergency Medicine,* 28**,** 113.
MORLING, N. 2009. PCR in forensic genetics. *Biochem Soc Trans,* 37**,** 438-40.
MULLIS, K. B. 1990. The unusual origin of the polymerase chain reaction. *Sci Am,* 262**,** 56-61, 64-5.
NAKAYAMA, Y., YAMAGUCHI, H., EINAGA, N. & ESUMI, M. 2016. Pitfalls of DNA Quantification Using DNA-Binding Fluorescent Dyes and Suggested Solutions. *PLOS ONE,* 11**,** e0150528.
QIAGEN. 2021. *Rotor-Gene Q MDx 5plex HRM (CA)* [Online]. Available: https://www.qiagen.com/ca/products/instruments-and-automation/pcr-instruments/rotor-gene-q-mdx-ca/ [Accessed Dec 2021].
THERMALFISHERSCIENTIFIC. 2021a. *Qubit 4 Fluorometer* [Online]. Available: https://www.thermofisher.com/au/en/home/industrial/spectroscopy-elemental-isotope-analysis/molecular-spectroscopy/fluorometers/qubit/qubit-fluorometer.html [Accessed September 2021].
THERMALFISHERSCIENTIFIC. 2021b. *VeritiPro Thermal Cycler* [Online]. Available: https://www.thermofisher.com/au/en/home/life-science/pcr/thermal-cyclers-realtime-instruments/thermal-cyclers/veriti-thermal-cycler.html [Accessed Dec 2021].
THERMOFISHERSCIENTIFIC. 2017. *PCR Cycling Parameters—Six Key Considerations for Success* [Online]. Available: https://www.thermofisher.com/au/en/home/life-science/cloning/cloning-learning-center/invitrogen-school-of-molecular-biology/pcr-education/pcr-reagents-enzymes/pcr-cycling-considerations.html [Accessed September 2021].
XIANG, Q., XU, B. & LI, D. 2007. Miniature real time PCR on chip with multi-channel fiber optical fluorescence detection module. *Biomedical Microdevices,* 9**,** 443-449.

Appendix

Appendix 1: How many spectrometers can we put into the I2C Bus?

The SCL clock line and SDA data line are used for the communication between the spectrometer chips and the Arduino Leonardo board. Therefore, the communication is using the I2C bus protocol which allows multiple I2C devices to communicate together with the 2 lines.

In this project, the Arduino Leonardo board will be the master which is responsible for generating the clock signal on the SCL line and sending the command to the slaves using the SDA line. (Bluedot) The slave here would be the spectrometers responsible for using the SDA line to send the data to the master.

Limitation of the pull-up resistor

One important factor that affects the number of the devices on the same I2C line would be the pull-up resistor, Figure 24 shows a circuit diagram of 2 slaves on the I2C bus connected to the master. Each slave is connected parallelly to the corresponding line of the master (Vcc to Vcc, SDA to SDA, SCL to SCL, GND to GND). The pull-up resistors are between the Vcc line and SDA/SCL line.

Figure removed due to copyright restriction

Figure 24. Circuit diagram of 2 slaves on the I2C bus connection to the master. (Bluedot)

The voltage on the Vcc line is 5V or 3.3V since the spectrometer chip can be supplied by either of the power supply from the Arduino board. The switches connect the GND line and the SCL/SDA line control the signal for SDA and SCL. Figure 25 uses only SCL line as an example since it is the same schematic as the SDA line. When the switch is open, there is no current flow through the chip; therefore, the SCL line is at 5V which the signal of it will be logical high. And when the switch is closed and the current goes through the device, the signal for the SCL line will connect to the ground line and the signal will be logical low.

Figure removed due to copyright restriction

Figure removed due to copyright restriction

Figure 25. Schematic circuit diagram for the SCL signal when the switch is open/closed. (Bluedot)

In the actual circuit, there is no mechanical switch between the GND line and SCL/SDA line. The close/open circuit is controlled by a transistor which works as the switch. (see figure 26) When the transistor is off, the resistance on the transistor between the SDA and the ground line would be very large; there will be almost no current flow through the 2 lines so the SCL line will be really close to 5V and the signal will be logical high. On the other hand, the resistance on the transistor between the SDA and the ground line would be very small when the transistor is on, a small current is now flowing through the transistor; since the voltage drop is still close to 0V, the signal for the SCL line will still identify as logic low. However, when we connect multiple devices into the I2C bus parallelly, the resistance of the pull-up resistor would be reduced a lot, which result in a larger current across the transistor that can lead to overheating within the device. Also, the voltage on the transistor

would be higher than the max low level output voltage (0.4V) which might lead to failure of the I2C bus. (Bluedot)

Figure removed due to copyright restriction

Figure removed due to copyright restriction

Figure removed due to copyright restriction

Figure 26. Schematic circuit diagram for the SCL signal when with the transistor.(Bluedot)

In order to avoid the failure of the I2C bus and overheating within the device, the I2C bus Specification and User Manual have set a maximum sink current which is 3 mA across the transistor. Now with the max low level output voltage (0.4V) and a maximum sink current (3mA), we would be able to calculate the minimum resistance

for the pull-up resistor for 3.3V and 5V power supply respectively by the formula below.

$$R_{Min} = \frac{Vcc - 0.4V}{3mA}$$

For 5V power supply:

$$R_{Min} = \frac{5V - 0.4V}{3mA} = 1.53\ k\Omega$$

For 3.3V power supply:

$$R_{Min} = \frac{3.3V - 0.4V}{3mA} = 967\ \Omega$$

In order to calculate how many spectrometers can be put into the I2C bus, the resistance for the pull-up resistor is needed to be measured. From the datasheet of Adafruit as7341, the pull-up resistor is 10 kΩ. The chip is also connected to a multimeter to measure the resistance between the Vcc line and SCL line, the result is 9.995 kΩ which matches the datasheet. (Shown in Figure 27)

Figure removed due to copyright restriction

Figure 27. Measuring the resistance between the Vcc line and SCL line with a multimeter. (Bluedot)

Let's assume the maximum number of spectrometers for the 5V power supply is n:

$$\frac{1}{R_{total}} = n \times \frac{1}{10\text{k}\Omega}$$

$$R_{total} = \frac{10\text{k}\Omega}{n} \geq 1.53\ k\Omega$$

$$n \leq 6.54$$

Then, let's assume the maximum number of spectrometers for the 5V power supply is m:

$$\frac{1}{R_{total}} = m \times \frac{1}{10\text{k}\Omega}$$

$$R_{total} = \frac{10\text{k}\Omega}{m} \geq 967$$

$$m \leq 10.34$$

Therefore, there can be a maximum of 10 spectrometers connected to the I2C bus when the circuit is powered by 3.3V, and a maximum of 6 spectrometers when they are powered by 5V.

Limitation of the I2C address

The register for the I2C address is 7 bits, which means it can be a maximum of 128 devices on the I2C; however, some ships have fixed address or fixed address bits. Since we cannot have 2 devices with the same address on the I2C bus, the number of the devices on the same line can be reduced to 20-30 or even less.

The I2C address would be the primary limitation for how many devices can be connected to a single I2C bus line. Figure 28 shows the description for the AS7341 slave address. Most likely the device has only fixed the I2C address; since it does not mention how to change the slave address for this device, and there is no solder jumper available for changing the address bit.

I²C Address

Figure 27:
AS7341 I²C Slave Address

Device	I²C Address
AS7341	0x39

Figure 28. I2C address for Adafruit AS7341.

In order to test if the I2C address can be modified using software, a script (see Figure 29) is uploaded to the chip. The code would change the I2C address of the board to 0x31 if the address is changeable. However, the I2C scanner still shows that the I2C device is only detected at address 0x39. (See Figure 29) Thus, the I2C address for

each of the spectrometers cannot be changed digitally with software or physically with jumper/soldering.

```
#include <Adafruit_AS7341.h>
Adafruit_AS7341 adress = Adafruit_AS7341();
void setup() {
  // put your setup code here, to run once:
  adress.begin(0x31);
}

void loop() {
  // put your main code here, to run repeatedly:
}
```

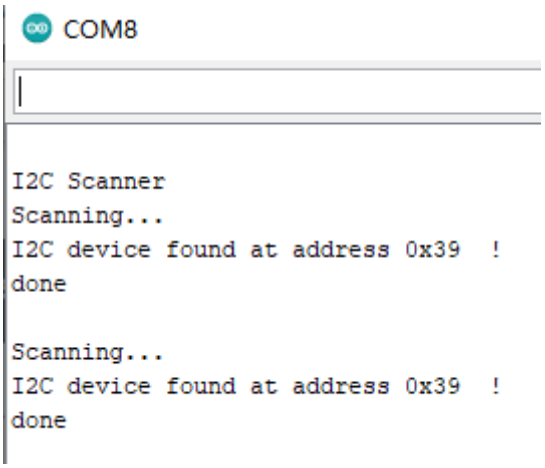

Figure 29. I2C scanner.

Solution

Connecting all spectrometers onto one I2C is not achievable since every spectrometer is having the same address; therefore, we can connect them separately into different I2C buses. Most microcontrollers like Arduino only have a single I2C bus, so how can we actually apply multiple I2C buses onto it? There are 2 ways to achieve it.

The first solution is using software implementation. The standard wire library with Arduino IDE limited the I2C wire to only 2 certain pins on the board. By using alternating wire libraries, the program allows I2C connections on different I/O pins and multiple I2C buses. (DronebotWorkshop, 2021) Some alternative wire libraries will work shown in Figure 30.

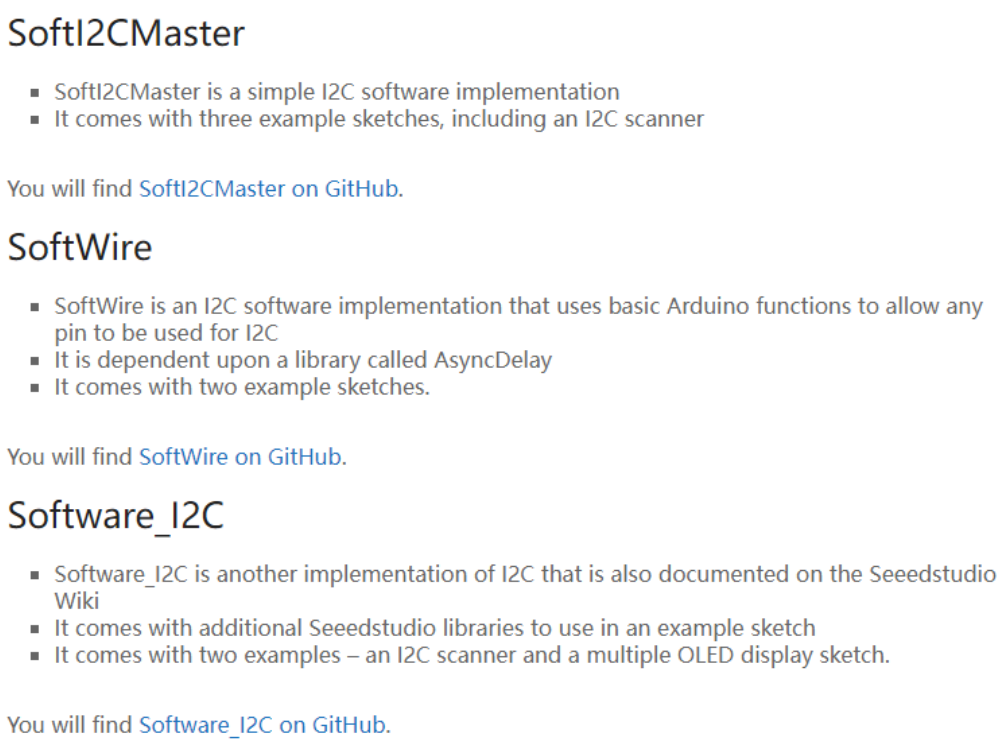

Figure 30. Alternative wire libraries that will work. (DronebotWorkshop, 2021)

The advantage of the software method would be no extra parts are needed for the connection of the I2C except a breadboard for holding the chip and connecting the power supply. All the SCL and SDA lines are connected to the I/O pins on the Arduino Leonardo board directly. The disadvantages include difficulty in programming and the library will not work if the chip is not using the wire library. With this software method, the available number of the spectrometer will be 8 since there are 14 digital I/O lines and a pair of original I2C lines on the Arduino Leonardo board; however, we still need a pair of the I/O line to communicate with the Ninja PCR so the maximum number of the spectrometer would be 7.

The second method is using hardware implementation. A multiplexer is a device that receives signals from multiple inputs and synthesizes a single output signal. In this project, we can use a TCA9548A I2C Multiplexer (see Figure 31) which can get up to 8 same-address devices hooked up to the Arduino Leonardo board.

Figure removed due to copyright restriction

Figure 31. TCA9548A I2C Multiplexer

The advantage of the hardware method would be no need for the wire library alternatives and there is less chance for software failure. The I2C Multiplexer allows 8 spectrometers to connect to it and sends all the data through the I2C line back to the microcontroller. Since the hardware solution provides more spectrometer connection and frees up pins on the Arduino Leonardo board. This method is more suitable for the project.

In conclusion, having multiple same spectrometers with fixed I2C addresses seems a dead-end for the project; however, any I2C address confliction can be solved using multiple I2C buses with an I2C multiplexer or an alternative wire library. The final design is using a TCA9548A I2C Multiplexer to connect 8 spectrometers to the Arduino Leonardo board, the spectrometers are powered by a 3.3 V line of the board.

Appendix 2: Power budget

Arduino

The power supply for the Arduino Leonardo can be 5V from USB or a 12V adaptor. The maximum current draw from a 3.3V pin on an Arduino Leonardo board is 50 mA. (Arduino, 2021) And the max power consumption for the Adafruit AS7341 chip is 300µA according to the datasheet. (See Figure 32) Since the spectrometers are connecting parallelly, the maximum current draw would be the sum of the current

draw for each device. The maximum current draw from the spectrometers will be 8 x 300µA = 2.4mA, which can be easily supported by the 3.3V pin on the board. The overall power consumption for the board still needs to be tested while running.

Power Consumption					
I_{DD}	Supply Current[2]	V_{DD}=1.8V; T_A=25°C Active mode[3]	210	300	µA
		V_{DD}=1.8V; T_A=25°C Idle mode[4]	35	60	µA
		V_{DD}=1.8V; T_A=25°C Sleep mode[5]	0.7	5	µA

Figure 32. Power consumption for Adafruit AS7341.

SMD LED

The power that supplies SMD LED would need to be much higher than 3.3 V since they are connected in series. (the specification for the SMD LED is shown in Figure 33) The forward voltage requirement for each LED is 3V and there are 8 in series, the voltage requirement for output would be 3V x 8 =24 V when a 30mA current is flowing through the circuit. There will be a protective potential meter to limit the forward current to 20mA in order to protect the LED. The current draw for the LED circuit would be 20mA.

SPECIFICATIONS:

LED Diode Colour: Blue
Lens: Rectangle, clear
Luminous Intensity Range: 200-220mcd
Forward Voltage Range: 2.8-3.4V
Forward Current: 30mA
Polarity: Anode (+), Cathode (-) (with marking)
Viewing Angle: 80°
Reverse Voltage: 5V
Operating Temperature: -40 ~ +100°C
Weight: 12.9g
Dimensions: 3.2mm x 1.6mm x 0.7mm

Figure 33. Specification for 1206 BLUE SMD LED DIODE