3-27-2018

Ultrahigh Speed Direct PCR: A Method for Obtaining Y-STR and STR Based Genotypes in Under 20 Minutes

Georgiana C. Gibson-Daw
Ggibs005@fiu.edu

FLORIDA INTERNATIONAL UNIVERSITY

Miami, Florida

ULTRAHIGH SPEED DIRECT PCR:

A METHOD FOR OBTAINING Y-STR AND STR BASED GENOTYPES IN UNDER

20 MINUTES

A dissertation submitted in partial fulfilment of

the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

CHEMISTRY

by

Georgiana Charlotte Gibson-Daw

2018

To: Dean Michael Heithaus
    College of Arts, Sciences and Education

This dissertation, written by Georgiana Charlotte Gibson-Daw, and entitled Ultrahigh Speed Direct PCR: A Method for Obtaining Y-STR and STR Based Genotypes in Under 20 Minutes, having been approved in respect to style and intellectual content, is referred to you for judgment.

We have read this dissertation and recommend that it be approved.

_____
George Duncan

_____
Watson Lees

_____
Fenfei Leng

_____
Yuan Liu

_____
DeEtta Mills

_____
Bruce McCord, Major Professor

Date of Defence: March 27, 2018

The dissertation of Georgiana Charlotte Gibson-Daw is approved.

_____
Dean Michael R. Heithaus
College of Arts, Sciences and Education

_____
Andrés G. Gil
Vice President for Research and Economic Development
and Dean of the University Graduate School

Florida International University, 2018

DEDICATION

I would like to dedicate this dissertation to Angelique Patricia Gibson-Daw and Philip

Arthur Gibson-Daw, my mother and father, who gave me the strength and courage to

persevere, and to my circle of friends that have always been there with patience, love and

unending support. Without these amazing people and the lessons, they have taught me,

the completion of this work would not have been possible.

ACKNOWLEDGMENTS

In the rush of joy and fulfilment we feel in the completion of a difficult task it is often the case that we forget the trials and tribulations we went through to get there and the people that stood by us and helped us get through the tough times. Therefore, I would like to take a moment to thank a very special group of people, without who this dissertation and the research and degree it represents would not have been possible. Thank you all for your hard work, your words of encouragement, your support and your patience while I completed this research and dissertation, which I can honestly say has been one of the most difficult challenges I have faced so far. Thank you for helping me to grow in knowledge, skill and confidence as the months and years have gone by.

First, I would like to thank Dr. Bruce McCord, my advisor. I can't begin to detail all the many times over the years that you have been of vital assistance to me, whether it was positive encouragement when my experiments weren't working, fixing an instrument in lab or even helping me out with my car on some of the many times it broke down. It has always been obvious to us that you really care about the overall well-being of your students and not just how much research they put out. You always find time to help us with any problems we have, no matter how busy you already are and we greatly appreciate you for this. I feel blessed and thankful that I was accepted into your lab. I would also like to thank Dr. McCord wife Margery Sue Phipps for not only putting up with all of taking up so much of Dr. McCord's time but also for helping us with our writing. Thank you Margery for proof reading all our papers and especially this dissertation, your hard work does not go unappreciated.

To members of my committee Dr. George Duncan, Dr. Watson Lees, Dr. Fenfei Leng, Dr. Yuan Liu, Dr. DeEtta Mills and Dr. Martin Tracey many thanks for your presence on my committee and your insightful questions and comments on my research as it progressed to completion. Thank you especially to Dr. Mills who joined my committee at the last moment and under such short notice. Members of my research lab, thank you for all the great ideas and input you have given me over the years and the moral and emotional support you have provided, I wish you all the best and hope we will remain colleagues for many years to come. To all the members of the Department of Chemistry and at FIU, thanks for everything you have done and to assist me with paperwork, travel reimbursement and other administrative tasks. A special thank you to Maggie and Jackie who have always been there to answer paperwork questions and have gone out of their way on numerous occasions to help with any administrative problems that occurred. I would also like to make special acknowledgement to Fabiana Taglia and Karin Crenshaw who have helped me complete experiments when I needed the help and have supported me throughout my job search.

Thank you to my father Philip Arthur Gibson-Daw who has spent many hours proof reading my work and helping me with imaging and formatting this dissertation. I'm sure he deserves half a PhD himself for all this work. To all my friends outside of FIU I also owe you a debt of appreciation, you have been my family and my support over these last few years and without all of you I would not be writing this today, thank you for all the time you spent listening to me talk about my research, all the positive energy you have sent me and all the candles you have lit for me.

ABSTRACT OF THE DISSERTATION

ULTRAHIGH SPEED DIRECT PCR:

A METHOD FOR OBTAINING Y-STR AND STR BASED GENOTYPES IN UNDER

20 MINUTES

by

Georgiana Charlotte Gibson-Daw

Florida International University, 2018

Miami, Florida

Professor Bruce McCord, Major Professor

There are many situations in forensic DNA typing where the time it takes to identify an individual is critical. Examples include suspects under arrest and individuals detained at ports of entry. It is also important to identify victims quickly when dealing with mass disasters such as terrorist attacks, airplane crashes and natural disasters in order to notify families. The goal of this project was to develop an ultrafast method for screening saliva samples through the use of rapid direct PCR coupled to microfluidic separation and detection.

Fast amplification was achieved through the use of high speed thermocyclers and the experimental optimization of PCR reactants and polymerases. High processivity mutant polymerases were tested at increased speeds along with PCR enhancers. Experimental optimizations and designs were used to ultimately achieve amplification of a 4 loci Y-STR multiplex in 12.5 minutes and a 7 loci STR multiplex in 6.5 minutes.

The use of direct PCR not only decreases the overall time of sample processing through the elimination of the extraction, but also, by the removal of the extraction step cuts down on the possibility of contamination to the sample. By using inhibitor resistant fast polymerases such as Omnitaq, BSA and the PEC-1 enhancer, and optimizing cycling conditions, a direct PCR method was developed where amplification from diluted saliva was successfully achieved in 13.5 minutes.

The products from these ultrafast amplifications can be coupled to a microfluidic chip for analysis. The electrophoretic system separates the PCR products and provides complete genotypes for both Y and autosomal STR multiplexes in 80 seconds. This permits complete sample processing in under 15 minutes for the rapid direct PCR protocol. This screening method can exclude individuals who do not match evidentiary material and test evidence to see if it is be viable for full analysis. The result of this work was a reliable and robust method for the rapid genotyping of forensic samples which can be used in situations where rapid single sample DNA analysis is necessary.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACRONYMS AND ABBREVIATIONS

ºC	Degree Celsius
A	Adenine
ABI	Applied Biosystem by Life Technologies
Bp	Base Pair
BSA	Bovine Serum Albumin
C	Cytosine
CE	Capillary Electrophoresis
CODIS	Combined DNA Index System
Ct	Critical Threshold
DNA	Deoxyribonucleic Acid
dNTP	Deoxyribonucleotide Triphosphate
dsDNA	Double Stranded DNA
EDTA	Ethylenediaminetetraacetic Acid
G	Guanine
HEC	Hydroxyethyl Cellulose
ILS	Internal Lane Standard
LR	Likelihood Ratio

NRC	National Research Council
PCIA	Phenol-Chloroform Isoamyl
PCR	Polymerase Chain Reaction
PHR	Peak Height Ratio
POP	Performance Optimized Polymer
PVP	Polyvinyl Pyrrolidinone
QC	Quality Check
qPCR	Quantitative PCR
RFLP	Restriction fragment length polymorphism
RMP	Random Match Probability
ssDNA	Single Stranded DNA
STRs	Short Tandem Repeats
SSR	Single Sequence Repeats
SWGDAM	Scientific Working Group for DNA Analysis Methods
T	Thymine
T_a	Primer Annealing Temperature
TAPS	3-{[1,3-Dihydroxy-2-(hydroxymethyl)-2-propanyl]amino}-1-propane-1-sulfonic Acid
Taq	Thermus Aquaticus Polymerase

T_m	Primer Melting Temperature
UV	Ultraviolet Radiation
VNTR	Variable Number Tandem Repeats
Y-STRs	Y-Chromosome STRs

CHAPTER I. FORENSIC DNA ANALYSIS

A. History & Technology

Forensic science is the application of science to the law. Among its most important tenets is that of Edmond Locard which states that "every contact leaves a trace". This tenet is true of both biological and non-biological evidence. Using scientific principles to help determine the details of a crime is not a new idea, as there are many examples of this happening in ancient history. Among the most famous examples is that of Julius Caesar's murder where it was determined that although he was stabbed a total of 23 times, only one of the stab wounds was fatal [Suetonius c.82, Greg 2006]. Another early example of forensic science is the Chinese book titled "His Duan Yu" translated as the "Washing away of Wrongs." This book, written by Song Chi in 1248 AD, teaches of how to differentiate a death by drowning or strangulation [Bell, 2008]. These early accounts show that the application of logical reasoning to criminalistics was established many years ago. However, recent advances in science and technology have truly aided juries in solving crimes. These advances started to appear around the 1800s when Mathieu Orfila the "father of toxicology" developed a method for determining the presence of blood. Later that century Sir Francis Galton and Sir Edward Henry created a method for classifying and identifying individuals by their fingerprints and in the early 1900s Leone Lattes developed blood typing from dried bloodstains. This was the first real step on the road towards DNA typing. However, it was not until the end of the century that the field of DNA profiling really evolved. In the early 1950s Erwin Chargraff published a series of papers detailing the different percentages of each of the 4 bases that make up DNA. In this work, he showed these ratios to be ~30% each of Adenine (A) and Thymine (T) and ~20% each of

Cytosine (C) and Guanine (G) [Chargaff, 1950-51]. Chargaff's work along with the efforts of Rosalind Franklin greatly aided Watson and Crick in their discovery of the double helix structure of DNA in 1953. They discovered that the A/T and CG interactions kept the two strands of the helix from dissociating and thus gave the structure its necessary stability [Watson and Crick 1953].

On March 19th, 1953, weeks before it was announced to the public, scientist Francis Crick excitedly wrote a letter to his son and told him of one of the most important scientific developments of modern times: his co-discovery of the "beautiful" structure of DNA, the molecule responsible for carrying the genetic instructions of living organisms; or, as Crick explained it to 12-year-old Michael, "the basic copying mechanism by which life comes from life". Although DNA was isolated back in the 1860s by Friedrich Miescher, its now famous double helix structure wasn't correctly modelled until the early 1950s by Crick and his colleague, James Watson, thanks in no small part to work already done by Maurice Wilkins, Rosalind Franklin and Raymond Gosling. In 1962, Crick, Watson and Wilkins were awarded the Nobel Prize for their efforts.

In April 2013, this letter became the most expensive in history after being sold at auction for $5.3million. A copy follows:

FRANCIS CRICK to MICHAEL CRICK – A MOST IMPORTANT DISCOVERY[*]

My Dear Michael,   *March 19th, 1953*

Jim Watson and I have probably made a most important discovery. We built a model for the structure of dex-oxi-ribose-nucleic-acid (read it carefully)

---

called D.N.A. for short. You may remember that the genes of the chromosomes —which carry the hereditary factors — are made up of protein and D.N.A.

Our structure is very beautiful. D.N.A. can be thought of roughly as a very long chain with flat bits sticking out. The flat bits are called the "bases". The formula is rather like this.

```
        |
sugar ——— base
        |
phosphorus
        |
sugar ——— base
        |
phosphorus
        |
sugar ——— base
        |
phosphorus
        :
and so on.
```
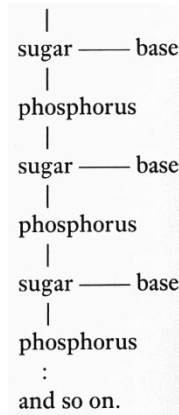
Now we have two of these chains winding round each other — each one is a helix — and the chain, made up sugar and phosphorus, is on the <u>outside</u>, and the bases are all on the <u>inside</u>. I can't draw it very well, but it looks like this.

[Diagram of the double helix as taken from page 3 of the original letter]
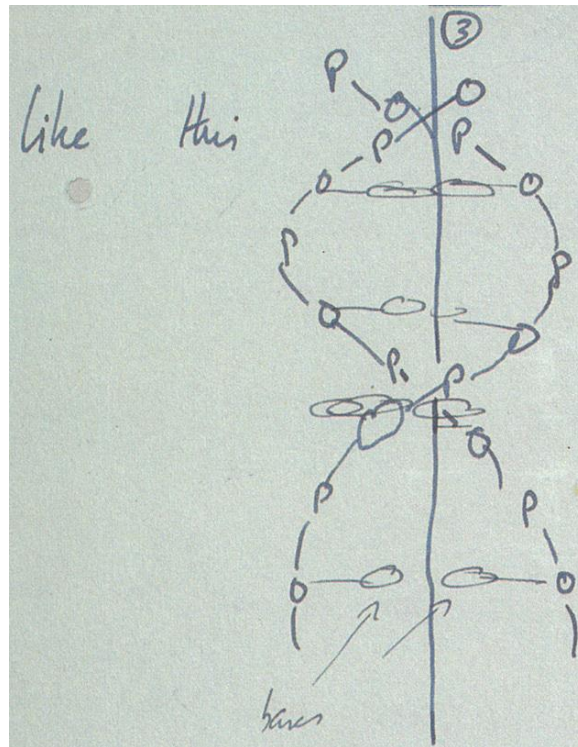
The model looks <u>much</u> nicer than this.

Now the exciting thing is that while there are 4 <u>different</u> bases, we find we can only put certain pairs of them together. The bases have names. They are Adenine, Guanine, Thymine & Cytosine. I will call them A, G, T and C. Now we find that the pairs we can make — which have one base from one chain joined to one base from another — are only

> *A with T*
> *and G with C.*

Now on one chain, as far as we can see, one can have the bases in any order, but if their order is <u>fixed</u>, then the order on the other chain is also fixed. For example, suppose the first chain goes ↓, then the second <u>must</u> go on

```
A---------------------------- T
T---------------------------- A
C---------------------------- G
A---------------------------- T
G---------------------------- A
T---------------------------- A
T---------------------------- A
```

It is like a code. If you are given one set of letters you can write down the orders.

Now we believe that the D.N. A. is a code. That is, the order of the bases (the letters) makes one gene different from another gene (just as one page of print is different from another). You can now see how Nature makes copies of the genes. Because if the two chains unwind into two separate chains, and if each chain then makes another chain come together on it, then because A always goes with T, and G with C, we shall get two copies where we had one before.

For example

```
                    A — T
                    T — A
                    C — G
                    A — T
                    G — C
                    T — A
               /    T — A    \
              /                \
            ↙       chains       ↘
                   separate

        A                          T
        T                          A
        C                          G
        A                          T
        G                          C
        T                          A
        T                          A

                      ↓
              new chains form

    A — T                  T — A
    T — A                  A — T
    C — G                  G — C
    A — T                  T — A
    G — C                  C — G
    T — A                  A — T
    T — A                  A — T
```
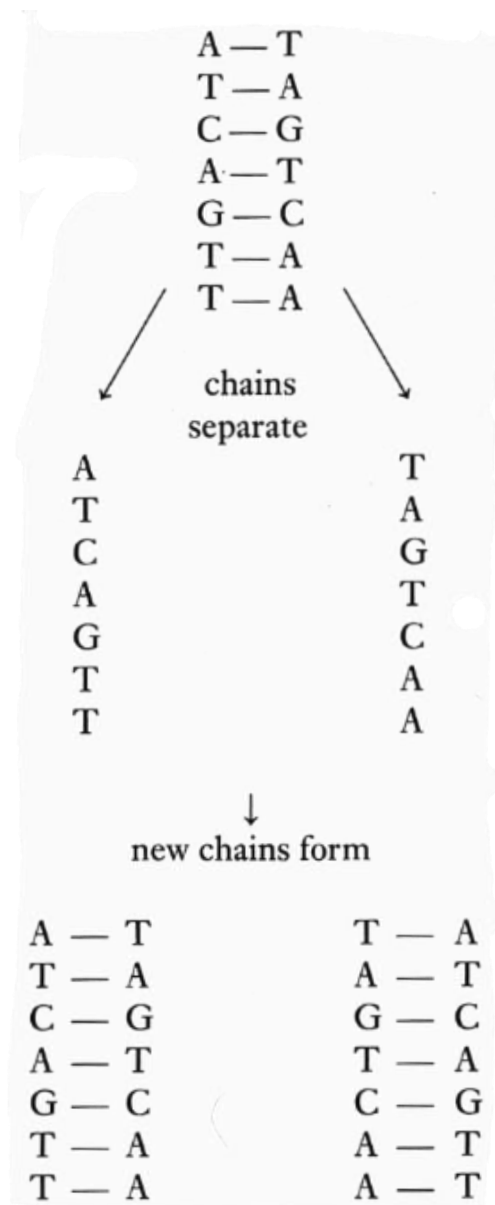
In others words we think we have found the basic copying mechanism by which life comes from life. The beauty of our model is that the shape of it is such that <u>only</u> these pairs can go together, though they could pair up in other ways if they were floating about freely. You can understand that we are very excited. We have to have a letter off to Nature in a day or so.

Read this carefully so that you understand it. When you come home we will show you the model.

Lots of love, Daddy

In 1984 Sir Alec Jeffreys made a discovery that would be the beginning of DNA typing methodology. Jeffreys discovered that the human genome contained a number of variable regions that contained repeat sequences that differ from one individual to the next. These regions were named Variable Number of Tandem Repeats (VNTRs). The differences in the number of length variations permitted discrimination of individuals, a process initially termed DNA fingerprinting [Jeffreys, 1985-1993]. The repetitive sequences were analysed using a method developed in 1980 by Arlene Wyman and Ray White [Wyman 1980], known as restriction fragment length polymorphism (RFLP). The RFLP process was originally designed to detect single base variations in the human genome [Botstein 1980]. Wyman discovered that DNA could be cut up into segments using restriction enzymes (endonucleases) harvested from bacteria, which cut the DNA at specific recognition sites. These fragments could then be separated by their sizes using agarose gel electrophoresis, in which smaller fragments move faster through the gel than the larger ones (Figure 1.1). Southern blotting was used to detect fragments in the gel by transferring them onto a nylon membrane and hybridizing them with radioactive (P_{32}) probes containing sequences complementary to the desired fragments. The radioactive phosphates probes were detected using radiographic film, resulting in a barcode like image (Figure 1.2) in which different patterns of bands in the resulting gel can be statistically analysed to permit individualization of human genotypes. [Butler 2009].

2) DNA is extracted from blood cells

3) DNA is cut into fragments by a restriction enzyme

4) The DNA fragments are separated into bands during electrophoresis in an agarose gel

1) Blood sample

5) The DNA band pattern in the gel is transfered to a nylon membrane by a technique known as Southern Blotting

11) The X-ray film is developed to make visible the pattern of bands which is known as a DNA fingerprint

6) The radioactive DNA probe is prepared

7) The DNA probe binds to specific DNA sequences on the membrane

10) X-ray film is placed next to the membrane to detect the radioactive pattern

8) Excess DNA probe is washed off

9) At this stage the radioactive probe is bound to the DNA pattern on the membrane

Fig.1.1 Schematic of the VNTR/RFPL/Southern Blot process. DNA is extracted and fragmented using restriction enzymes, these fragments are then separated by gel electrophoresis. After electrophoresis, the fragments are transferred to a nylon membrane and a radioactive probe binds to its complimentary position. The membrane is then exposed to X-rays and fragments show up as dark bands on the film which can then be analysed and compared with the bands obtained from different evidentiary or reference samples. From https://liamdnafingerprinting.weebly.com/techniques-and-process.html
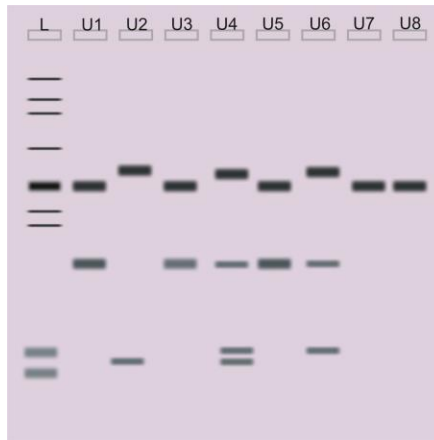
Fig.1.2 Example of what the barcode view of DNA fragments looks like. L indicates the ladder while U1-8 indicate unknown samples (crime scene samples from an unknown source). A match occurs at that marker when two bands align with each other. All bands would have to match at each marker for two samples to be considered as coming from the same source. Image created by author.

Dr. Jeffreys used the RFLP/Southern Blot process to analyse his VNTRs [Jeffreys 1985].

Originally, he used a restriction enzyme known as Hinf I, along with multi-locus probes to

digest the DNA. These probes were made of sequences that were common for several

different VNTR loci. This made interpretation complex and confusing, especially when

there could have more than one contributor. To simplify analysis Dr. Jeffreys began to use

single locus probes instead and he termed his method DNA fingerprinting [Jeffreys, 1985-

1993]. Jeffreys's method was first used in a paternity case in 1985 to reunite an immigrant

child with his family in England [Jeffreys 1993] and then in a most famous homicide case

in 1986 [Wambaugh, 1995]. The case involved the rape and murder of two teenage girls in

a small village in England. A suspect had confessed to one of the murders but not the other.

His mental stability was dubious. There was a lack of evidence other than the DNA

retrieved from the victims, so the police decided to use Jeffreys DNA fingerprinting

method to test the DNA of everyone in the village and to compare it to the evidence found

at the crime scene. After all the samples had been tested and compared, including that of

8

the suspect who had confessed, no matches were found. The police continued to search unsuccessfully for answers, until a man was heard boasting that he had submitted a DNA sample for a friend. When the police had the friend's DNA tested, it was a positive match to the DNA from the crime scene and he was arrested and later convicted for both of the rapes and murder [Wambaugh, 1995]. This case proved the usefulness of the VNTR/RFLP method. However, although it was accurate and reproducible, it was also lengthy, labour intensive and needed harmful reagents and a relatively large amount of DNA [Butler 2009].

Around the same time that Dr. Jeffreys was implementing his DNA fingerprinting method, Kary Mullis theorized that if he could make more copies of the DNA it would be easier to analyse and extract information [Mullis 1990]. To do this he would use a set of two primers per segment (loci) of DNA he wanted copied. By using polymerases, nucleotides and the correct heating and cooling cycles he theorized these segments could be copied indefinitely. The developed method was called the Polymerase Chain Reaction (PCR) [Mullis 1987]. The Polymerase chain reaction (PCR) proved to be a sound method and was able to make millions of copies of the same DNA sequence leading to exponential increases to the amount of DNA that biologists had access to. The amount of DNA required for PCR was 0.5ng or less, rather than the 0.5µg or more previously needed for the VNTR/RFLP method [Butler 2005]. The PCR method when applied to the forensic field meant that scientists could get profiles from crime scene samples that previously would have been too degraded or not contained enough DNA to work by VNTR/RFLP.

Initially new polymerase needed to be added after each denaturation step as the heat of this denaturing degraded the enzyme beyond functionality. This made the process laborious

[Mullis 1990]. However, the discovery of new polymerases that were thermally stable at high temperatures (ex. *Thermus aquaticus*, a.k.a *Taq*. Bacteria found in geothermal springs) rendered this step unnecessary [Innis 1999]. Further developments in polymerase technology led to what we now call "hot-start" polymerases. These are polymerases that are bound to a repressor molecule inhibiting their action. When the polymerase is heated to a certain temperature, this repressor separates from its substrate allowing the polymerase to activate and begin its function. The hot start mechanism helps to improve the specificity and efficiency of the reaction as it ensures that the polymerase does not start until the desired temperature has been reached and the primers have been annealed [Innis 1999, Butler 2009].

The first PCR assay used to detect sequence variations in humans was the leukocyte antigen (HLA) locus, DQα1 marker. This area of the genome codes for the α-subunit of the DQ protein of the major histo-complex on chromosome 6 and it has a 242 bp region which has at least eight different known alleles [Sajantila 1991, Blake 1992]. As this PCR assay was being developed a new detection method was also being used, known as the "dot blot", in which the DNA is bound to a substrate and a probe complementary to the desired sequence is hybridized to the target. Reagents were then added which caused a colour change at the location of the successfully hybridized probe [Sajantila 1991]. There were two problems with this initial assay. The first was that the results were difficult to interpret if the sample was not obtained from a single source. Secondly while the procedure was capable of analysing smaller amounts of DNA, its power of discrimination was low. To improve the discrimination power, more loci were added to the assay. The new assay was known as the AmpliType Polymarker kit. It included five extra polymorphic loci. These five extra

markers had two to four known alleles, and included: D7S8, haemoglobin gamma globin (HBGG), Gycophorin A (GYPA), low density lipoprotein receptor (LDLR) and a group specific component (GC) [Koblinsky 2004].

To improve discrimination power, an alternative PCR method was developed that detected length variations known as variable length tandem repeats (VNTRs). These were also known as amplified fragment length polymorphisms (AmpFLPs). One of the favoured AmpFLPs for forensic purposes was known as D1S80. It amplified a 432 bp region containing a 16 bp sequence that contained 23 alleles [Duncan 1997]. The PCR products were separated by size in an agarose gel and visualised using the intercalating agent ethidium bromide and UV light. The alleles observed were sized in comparison to an allelic ladder run in parallel with the samples and contained all of the known alleles for the studied locus.

The discovery of Short Tandem Repeats (STRs) rendered previous methods obsolete, at least in forensics. The STRs or microsatellites were similar to VNTRs in that they included a region containing differing number of repeats in different individuals. However, while VNTRs have repeats of 8-100 bp [Butler 2009] in STRs, these STR repeat units are much shorter 2-7 bp [Butler 2007, Lins 1998]. A number of considerations are given for choosing STR markers to probe for forensic purposes and for protecting personal information of the individuals tested. First, only those from non-coding regions are used. Secondly, it is desirable that the STRs chosen, be short in sequence while having as many different alleles as possible. Short sequences more readily amplify DNA, maximizing the number of alleles and enhancing the power of discrimination. Another benefit of STRs is that they can easily

be combined in multiplex assays. Commercial kits now contain up to 24 loci, saving time, labour and quantity of sample used, while significantly improving both sensitivity and power of discrimination (>1 in 10^6). For these reasons the combination of STR, PCR and CE have replaced previous DNA fingerprinting methods and technologies in current crime laboratories.

To use STR methodologies in conventional forensic laboratories it has become necessary for standardized guidelines to be produced and implemented. The National Research Council (NRC) was tasked with producing the procedural guidelines and developing statistical methods to be used when interpreting the resulting profiles, such as the random match probability (RMP) and the likelihood ratio (LR) [NRC 1996]. The FBI created a database of 13 core loci called the Combined DNA Index System (CODIS) in order to enhance cohesion and standardisation between the labs across the nation, as well as to aid in the sharing of information. The CODIS database was designed to include the profiles of criminals convicted of certain crimes as well as crime scene samples. This enabled searches between crime scene samples and individuals in the database (missing persons can be added to this database as well). Originally the 13 loci that for this core set were: CSF1PO, FGA, TH01, TPOX, VWA, D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539, D18S51, D21S11 and a sex marker, Amelogenin (Figure 1.3) [Butler 2005, Butler 2006]. These 13 loci gave a good power of discrimination (one in a trillion, except for identical twins [Hill 2012]) and they were validated extensively for the purpose of creating the CODIS database [Budowle 1999]. The sex marker Amelogenin presents a 6 bp difference in size between the X and Y chromosomes, which allows scientists to determine if a subject is a male or a female. Amelogenin can also be used to detect the presence of male/female

mixtures [Sullivan 1993, Mannucci 1994]. As genotyping technological advanced, additional loci were added to commercial kits (although the original 13 were usually preserved) [Collins 2004]. Eventually CODIS increased its core set to include: D1S1656, D2S441, D2S1338, D3S1358, D5S818, D7S820, D8S1179, D10S1248, D12S391, D13S317, D16S539, D18S51, D19S433, D21S11, D22S1045, CSF1PO, FGA, Penta D, Penta E, SE33, TH01, TPOX and vWA [Hares 2012-2015, Moretti 2016]. The addition of extra loci gave a higher power of discrimination, and assisted in the discrimination of close relatives.
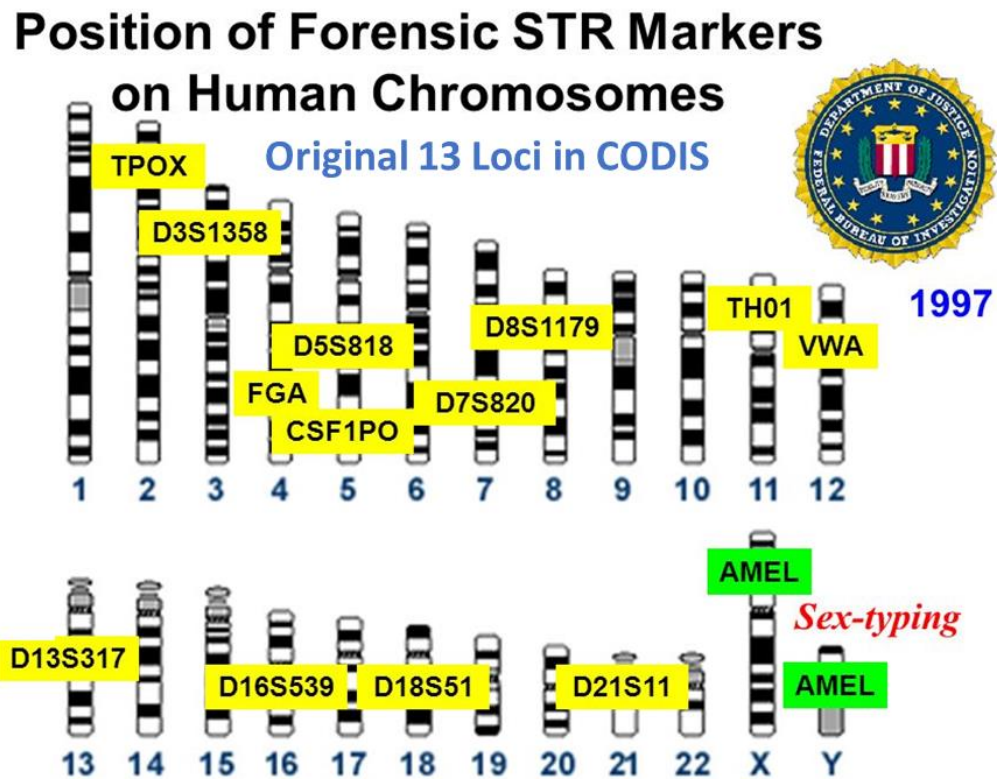
Fig.1.3 Original 13 loci included in CODIS in 1997. Adapted from http://strbase.nist.gov/fbicore.htm

The current method for separating and detecting the STRs copied by PCR is through capillary electrophoresis (CE). Capillary electrophoresis (CE) is similar to agarose gel electrophoresis but is much faster and has better resolution and sensitivity. STRs are separated by size and then different coloured dyes (which are attached to the forward primers at the 5' end of the DNA strands) are used to allow loci of interest to be separated in the resulting profile without overlap. The CE instrument detects the dyes by exciting the molecules with a laser and capturing the emission wavelengths. The most common instruments used today in forensic laboratories today for STR DNA profiling are the Applied Biosystems (ABI), now Life Technologies, Genetic Analyzers. These are CE instruments that can be used with one capillary or with 16-capillary arrays. Recently, forensic laboratories have begun a transition from 5 dye systems 6 dye systems, allowing for the analysis of more multiplexed loci which leads to an increase in the power of discrimination. These new instruments also accommodate up to 24 capillaries, allowing for a further increase in throughput.

Current work continues in an effort to add more loci and to develop faster methods of amplifying DNA, while maintaining the level current level of sensitivity and power of discrimination. Massively parallel sequencing (MPS) methods can also detect STRs and may take the place of CE based analysis in the future. However, these techniques are slow and far more complex than the PCR/CE based methods. In particular, MPS procedures are slow and expensive when rapid single sample analysis is required. It is the goal of this thesis to demonstrate and extend the capability of high speed analysis of singles samples through the use of rapid and rapid direct PCR methods coupled to microfluidic electrophoresis separation and detection instruments.

B. DNA

B.1. Introduction

Although the cell was first discovered in 1665 by the English microscopyst Robert Hooke, it took almost a century before scientists began to theorize that this tiny molecule might be the basic unit of all life [Mazzarello, 1999]. The cell theory was validated as cells were found to be responsible for the activity and energy storage and use of the organism. Furthermore, they were found to contain, preserve and transmit hereditary information, which was passed down through independent replication and division of cells through the process of mitosis and meiosis. Deoxyribonucleic acid (DNA) stores all of the genetic material and information that make up living organisms. This genetic information is stored in each and every cell of the host organism (except in mature red blood cells) [Alberts 2014]. In simpler prokaryotic organisms, the DNA is present in the cytoplasm because these organisms lack a nucleus, nor do they have histones associated with their DNA. This makes the processes of replication, transcription and translation much simpler in prokaryotes than in eukaryotes. In complex organisms, known as eukaryotes, the DNA is complexed with histone proteins and found in the nucleus in the form chromosomes (Figure 1.4), which act as a blueprint for the manufacture of proteins that make up the organism [Pierce, 2013]. In humans, there are two copies of each chromosome. One set of chromosomes is inherited from each parent. In this way the genetic information is passed down through the generations [Nussbaum 2007]. The DNA sequence in the cell is 99.7% identical among humans (97% identical with apes and 60% identical with chickens and fruit flies). The remaining 0.3% of the human genome is used for forensic purposes. For

example, polymorphic regions of the genome can be used to produce profiles unique to an individual, and so can be used to identify and distinguish among individuals. DNA evidence can also be used for investigation of crime scenes, in paternity testing, for identifying human remains from missing persons and to aid in diagnosis and treatment of genetic diseases [Butler, 2006, 2011].
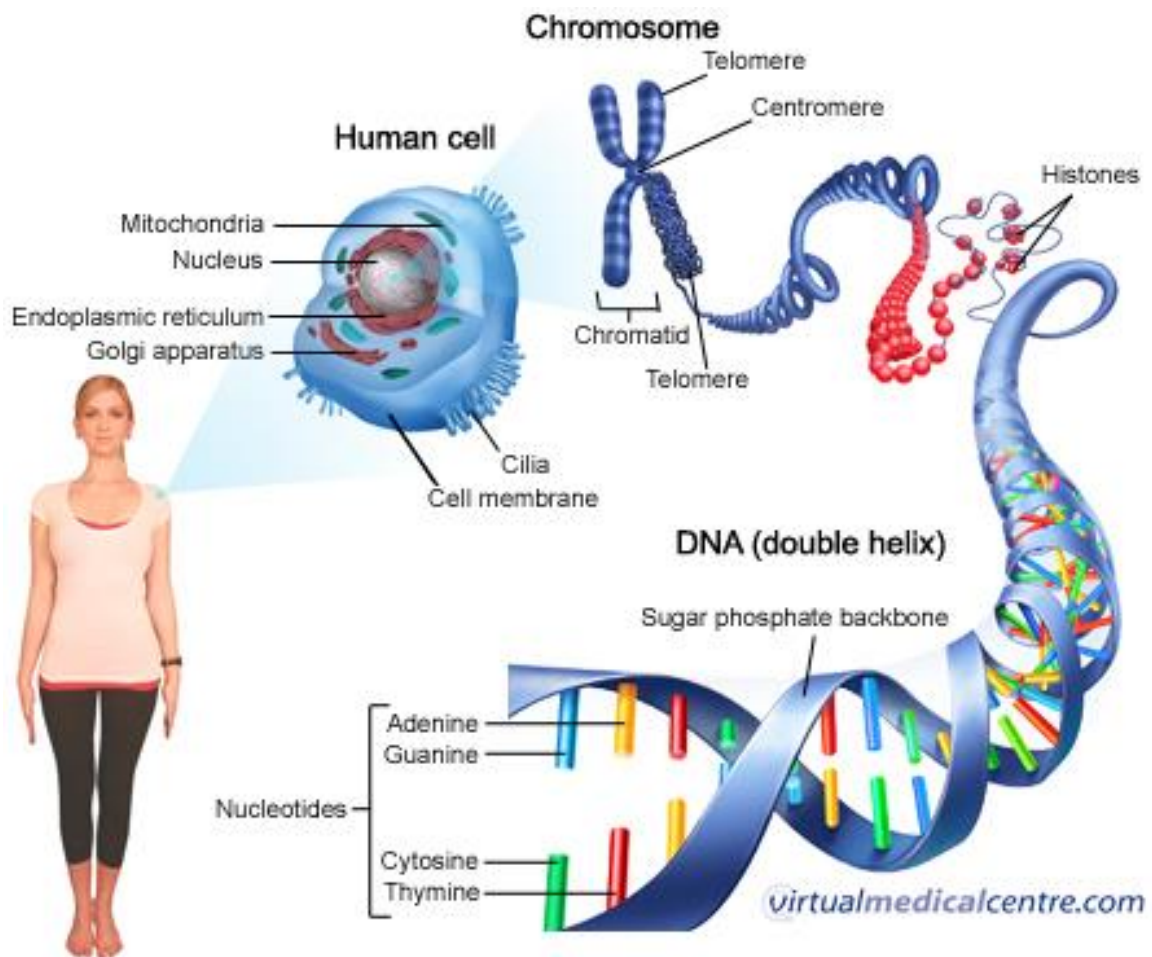
Fig.1.4 Structure of a cell. The nucleus is the home of nuclear DNA, which complexes with histones to form nucleosomes. The nucleosomes then coil to form chromatin which during division supercoils into the more compact chromosomes. Mitochondria inherited from the mother is present outside of the nucleus and has its own DNA which is different from the nuclear DNA. Image from http://slideplayer.com/slide/10841754/

B.2. DNA Structure

DNA, a double helix of two complementary strands (Fig. 1.5), is a large molecule which is built of repeating units called nucleotides. Nucleotides are composed of a 5-carbon sugar known as 2'deoxyribose, a phosphoric acid and then, depending on the nucleotide, one of the four nitrogen containing bases Adenine, Guanine, Cytosine and Thymine denoted A, G, C and T. These bases, which account for the structural variation of the DNA, are divided into 2 types: Purines which have a double ring structure (A and G) and Pyrimidines which have a single ring structure (C and T). The structures of these bases are shown in Figure 1.6 with the N that is joined to the deoxyribose and in turn to the phosphoric acid circled in red.
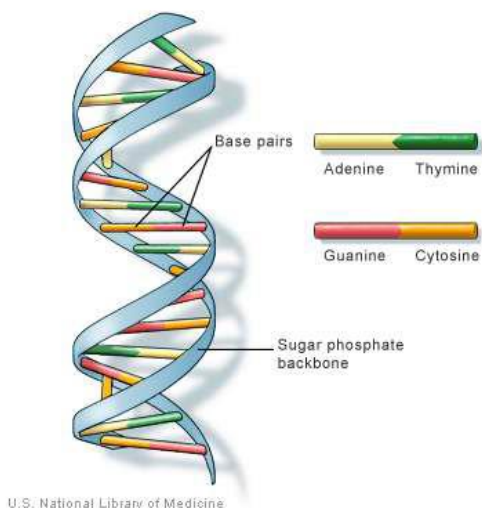
Fig.1.5 DNA Double Helix Structure. From https://www.sciencebuddies.org/blog/celebrating-dna-and-the-history-of-the-double-helix
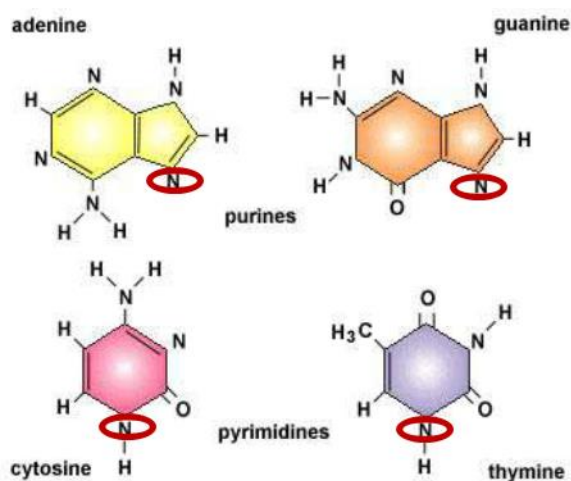
Fig. 1.6 Nucleotide bases. Circled in red is the N that joins the nucleotides to the deoxyribose. Adapted from https://chelseaharripersad.wordpress.com/tag/structure-of-dna/

The nucleotide bases attach to the sugar phosphate backbone, which is formed by linkage between the 5' hydroxyl of one pentose sugar to the 3' hydroxyl of the next sugar. The

DNA molecule is a double stranded helix containing two complementary polynucleotide chains (Fig.1.7). These two complementary chains are associated with one another by base pairing in which the nitrogenous base of the nucleotide on one chain is paired to its complementary base (A-T and C-G) on the other chain. This association of the two chains is achieved through hydrogen bonding (Fig. 1.7), where a hydrogen atom is shared between two other atoms. Within the DNA molecule the H-bonds act as structural elements, (2 H-bonds per A/T pair and 3 H-bonds per G/C pair). In this way, the strands are held together and twist around each other to form the stable double helix structure of DNA [Hartl 2008] (Fig.1.5).
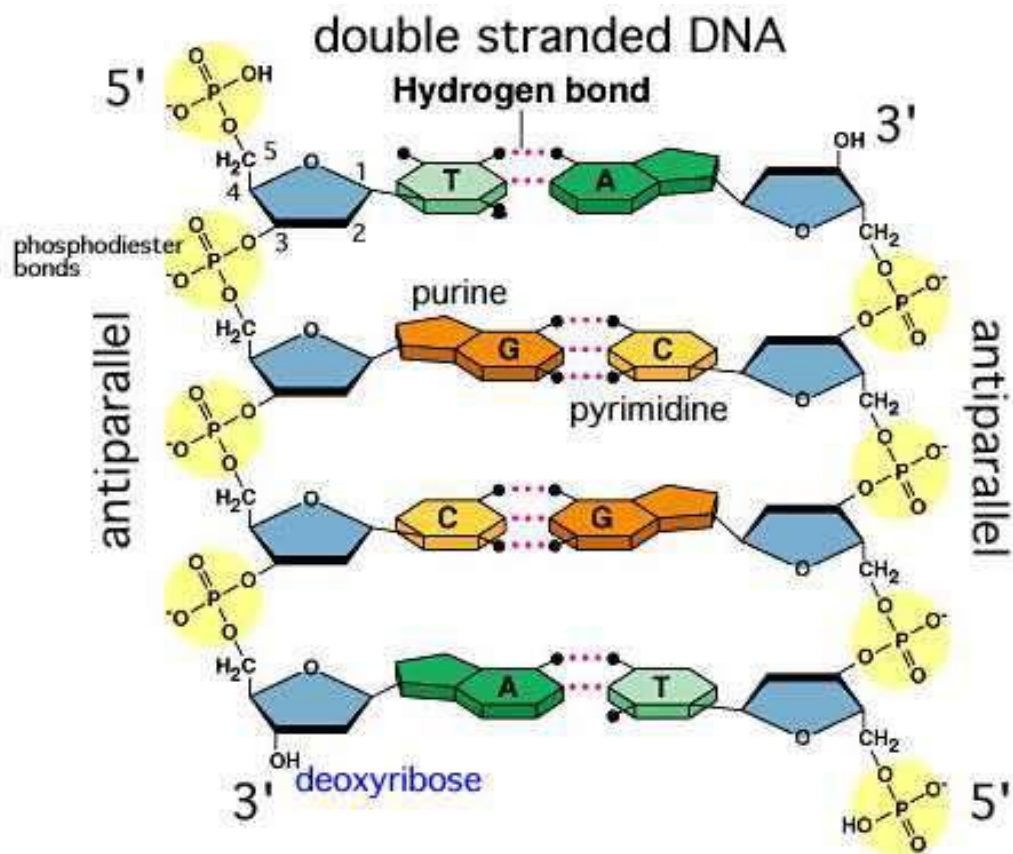
Fig.1.7 Base Pairing. Adapted from https://www.slideshare.net/sofipaz1/dna-double-helix-2-class

It is the different sequence variations and combinations of the nucleotide bases that lead to different messages coded in DNA, thus allowing for the many genetically different and complex organisms.

B.3. DNA Arrangement: Chromosomes, Genes and Alleles

Once the DNA double helix has formed it is then supercoiled with the help of proteins known as histones into what is known as chromatin. Chromatin condenses during cell division into structures known as chromosomes Figure1.8. Different species have different numbers and sizes of chromosomes. In *Homo Sapiens*, there are 46 chromosomes forming 23 pairs of chromosomes, with one set inherited from the maternal parent and one from the paternal parent. Two of these chromosomes (the X and Y chromosomes) determine the sex of the individual with females having two X chromosomes (XX genotype) and males having both an X and a Y (XY genotype).
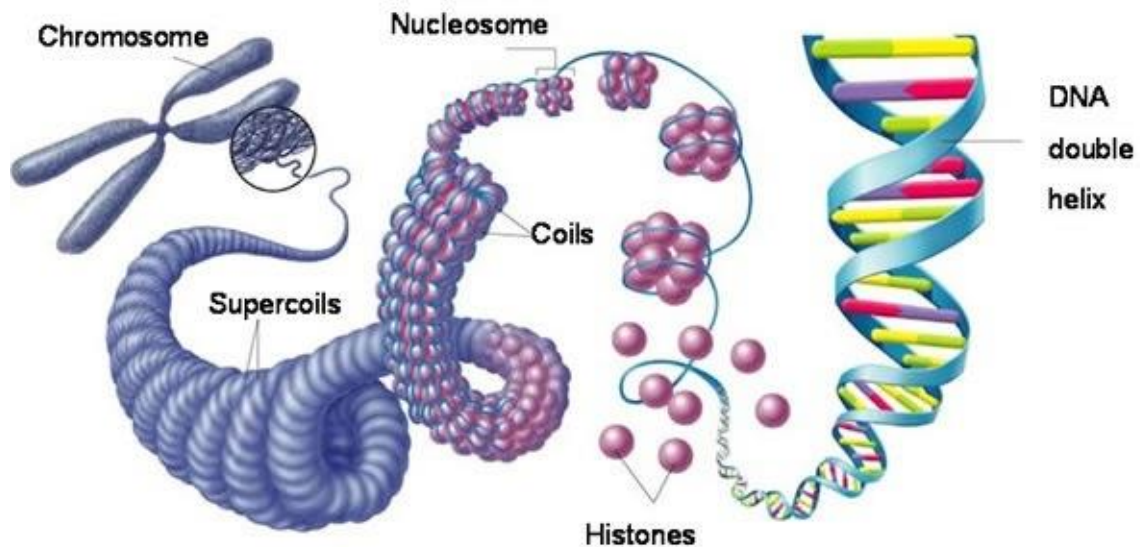
Fig.1.8 DNA structures from basic double helix, associated and coiled around histones to form nucleosomes and chromatin then supercoiling to form chromosomes. Image from http://slideplayer.com/slide/6324087/

The DNA in each chromosome can be divided into coding regions (exons) and non-coding regions (introns). The exons are sections of a gene that encode for a protein or peptide sequence while the introns (formerly known as "junk DNA") do not code for anything and seemed to merely interrupt the gene sequence [Alberts 2014]. Recently these non-coding regions have been found to have a role in gene activity and expression [Barret 2012]. Genes are considered the working subunits of DNA. Each gene consists of a number of nucleotide bases which then act as a set of instructions that code for the production of a messenger RNA and subsequently a particular polypeptide chain with a particular function in the organism [Pearson 2006]. Genes are present in different forms (alleles) and the number of variant alleles present in a population depends on the gene under consideration.

B.4. DNA Marker nomenclature

There are roughly 3 billion base pairs making up the 23 pairs of chromosomes observed in the human genome. Each of these chromosomes can contain several hundred to several thousand genes. These genes account for the coding sequences of the DNA (1.3%), while non-coding sequences make up the rest (98.7%). Many of the non-coding regions are highly polymorphic between individuals (have a high variability in bp sequence and bp sequence repeats) and so these regions are of forensic value for use in identification of individuals. The variances in sequence arise as a result of random mutations. The resultant alleles are passed down through the generations. The DNA markers used in forensic science [Pierce 2016] are DNA sequences which have a known location on the chromosome. The term DNA markers can refer to indels (insertions/deletions), single nucleotide polymorphisms (SNPs) sequence variants, short tandem repeats (STRs), and

VNTRs. The naming of these loci gives information about individual sequences and their location on the chromosome. For example, an STR locus having the following designation D5S818, includes a D for DNA, a 5 for the chromosome number, an S indicates a single copy sequence and a number, in this case 818, indicating the order in which it was discovered. Thus, this location was the 818th marker described on chromosome 5. These locus names are often abbreviated to include the first letter and number for ease of use. For example, locus discussed above D5S818 is often known as D5 among specialists in the field. Some markers have a different nomenclature, due to the locus falling within the coding region of a gene. For example, the TPOX locus falls within the coding region of the thyroid peroxidase gene.

CHAPTER II. COLLECTION AND ISOLATION OF BIOLOGICAL EVIDENCE

A. Collection of Biological Evidence

A.1. Introduction

Physical evidence found at a crime scene plays a vital role in both the conviction of a guilty suspect or the acquittal of an innocent suspect [Lee 2001]. Forensic evidence can be compromised if the samples are not correctly collected, stored and transported. This is especially true in the case of biological evidence, which if handled improperly can become contaminated or degraded over time. To avoid these problems, protocols for collection, storage and transportation, as well as chain of custody, must be carefully followed and documented. These protocols and the treatment of the evidence are contingent on the type of sample being collected and on what substrate it is on. The most common biological samples collected from crimes scenes are blood, semen, saliva, hair, teeth, tissue, bone and urine. Aside from the samples collected from a substrate, it is also important to collect a substrate blank which is taken from a section of the substrate uncorrupted by biological evidence.

In addition, reference samples from the victim, the victim's belongings, family members and any possible suspects should also be collected so that the resulting profiles can be compared with those obtained from the crime scene evidence [Tilstone 2006]. These reference samples can be collected in a variety of ways, but the least invasive sample is a buccal swab.

General collections of biological samples for DNA testing can include both wet and dry samples, which are collected by employing a double swab technique. This method utilizes first a wet swab to collect the sample and then a dry swab to remove any remaining DNA that may have been loosened by the first wet swab but not collected. All wet samples and stains must be air-dried before further packaging in paper or card containers, as any moisture in the sample will promote bacterial and fungal growth leading to DNA degradation. Once collected, samples can be stored for short periods of time in a refrigerator at 4 ºC. They can also be kept for longer periods of time in a freezer at -20 ºC [Butler 2005] or transferred to an FTA card which permits storage at room temperature. Reference samples should also be collected from individuals of interest to the case.

If the composition of the sample to be collected is unknown, then presumptive serological testing can be performed to give an indication of what type of sample is being analysed. It is also important to check for trace evidence and such as touch DNA and to make sure these pieces of evidences are handled accordingly so the DNA evidence is not lost [Raymond 2009, Butler 2011].

A.2. Blood Evidence

Blood is the most commonly found biological evidence at crime scenes. It is used as a source of DNA and to provide information on the events of crime with its spatter pattern [Lee 2014]. The majority of the DNA obtained from blood comes from the nucleus of the white blood cells [Baird1978, Barni 2007]. The more numerous, red blood cells do not have a nucleus and so do not contain DNA but they contain haemoglobin which can be

used to indicate the presence of blood. Haemoglobin can also behave as a PCR inhibitor and therefore careful DNA extraction procedures must be employed to reduce its effect.

When bloodstains are present on small, moveable substrates, the whole sample is collected and dried. It is then packaged in a paper or card container for transportation and storage [Lee 2014]. Wet bloodstains on unmovable substrates are collected using a sterile swab. Dried bloodstains are collected either by scraping with a sterile blade, by using fingerprint tape to lift it off the substrate or using the double swab technique [Lee 2001]. All samples are stored in paper or card to prevent growth of mould and bacteria that's results in DNA degradation.

Presumptive tests for blood focus on detecting the haemoglobin (HB). Its function is to transport oxygen from the lungs to the rest of the tissues and organs of the body, while removing CO_2. One presumptive test for blood uses luminol (5-amino-2, 3-dihydro-1, 4 phthalazinedione) (C8H7N3O2) and hydrogen peroxide (H_2O_2) to produce a chemiluminescent signal. In the presence of hydrogen peroxide the luminol is oxidised to form 3-aminophthalate, which is in an excited state. Once it returns to ground state, it emits blue light [Barni 2007] which can be detected photographically (Figure 2.1). The luminol method is effective in detecting even trace amounts of blood (diluted X 10^6) while not compromising downstream DNA analysis of the sample. However, the procedure can produce false positives if the reagents come into contact with other substances that contain iron and similar metals (copper, nickel, chromium). Bleach may also produce a false positive. Also, when using luminol the blood spatter patterns may be destroyed and thus it

is important to photograph and document the scene before the luminol spray is used [Barni 2007].
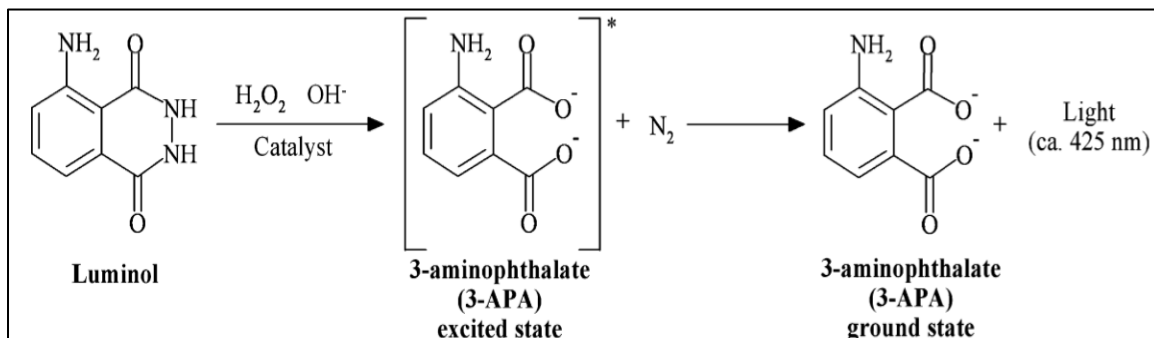
Fig.2.1 Mechanism by which luminol chemiluminesces to produce blue light when coming into contact with the iron in blood. In the presence of hydrogen peroxide the luminol is oxidised to form 3-aminophthalate, which is in an excited state it returns to ground state by emitting blue light at 425 nm [Barni 2007].

Another presumptive test for blood is the ABAcard HemaTrace kit (Abacus diagnostics Inc., West Hills, CA). This kit uses strips containing monoclonal antihuman haemoglobin antibodies which bind with the antigens of human haemoglobin as they migrate through the strip. In the presence of haemoglobin, the stationary antibodies, labelled with a pink dye, will form antibody-antigen-antibody complexes which can accumulate and produce a pink band on the strip. The appearance of this pink band indicates the presence of blood. This presumptive test is both sensitive and specific to human blood but false positives can occur in the presence of ferret or higher primate blood [Lee 2001, Lee 2014].

A.3. Semen Evidence

Seminal fluid is another sample commonly found at crime scenes, especially in rape cases, often in the form of a dry white stain. Semen samples are most often found on the victim's clothes, body or bedsheets. Where possible, the entire item that the stain is found on is collected and sent to laboratory, as additional trace evidence may be present that is not seen

by the naked eye. Semen stains are collected and dried in the same way as blood evidence, and are commonly stored in paper or cardboard to prevent bacterial growth.

Seminal fluid contains both enzymes that assist in its identification and sperm cells that contain the nuclear DNA. An alternative light source (ALS) at a wavelength of 450 nm can be used to visualize fluorescence of the sample which can then be marked on the substrate before sending it to a lab for analysis [Lincoln 2006].

Presumptive tests for semen are obtained by the detection of the acid phosphatase (AP) enzyme. Prostate Specific Antigen (PSA), also known as p30 [Denison 2004] can also be used. A spot test can be performed using Brentamine, in which the acid phosphatase from seminal fluid reacts with sodium alpha naphthyl phosphate and diazo blue dye solution to produce in a colour change to purple [An 2012]. Tests that use the detection of acid phosphatase are only considered presumptive as AP is also present in other body fluids, such as vaginal secretions and breast milk. However, the levels of acid phosphatase in these fluids are present in lower concentrations than those in seminal fluid [Virkler 2009]. Confirmatory tests can also be made through the direct visualization of sperm. This visualization can be achieved using a microscope and a method known as Christmas tree staining, in which nuclear red dye stains the nuclei present in the sperm head while picroindigocarmine targets the tails, and stains them a blue/green colour [McClintock 2014]. This confirmatory test only works if there is sperm present in the seminal fluid, which is not the case in vasectomized or azospermic males. In these cases, confirmatory tests detect PSA (which is a glycoprotein produced by the prostate gland). Commercially available test kits are available from Abacus diagnostics Inc. (West Hills, CA) and they

work by the same principle as the HemaTrace kit described above for blood detection except they use anti-PSA antibodies instead of anti-haemoglobin antibodies [Hochmeister 1999].

A.4. Saliva Evidence

Saliva evidence can be collected either as straight saliva or through the use of buccal swabs (most common form of reference sample collected). Saliva can also be collected from items such as cigarette butts, chewing gum, drinking containers, toothbrush, envelopes, stamps and bite marks. Saliva stains are collected the in the same way as blood using a sterile swab and the double swab technique [Pang 2007, Lee 2014]. The swab must then be left to air dry and then packaged in a paper envelope or cardboard swab box to prevent moisture collection, condensation and bacterial growth during storage and transportation. Saliva can contain epithelial cells from inside the cheek which are a rich source of DNA. Saliva stains can be difficult to identify with the naked eye and thus require other means of visualisation such as alternate light sources (ALS) at a wavelength of 450 nm that causes them to fluoresce [McClintock 2014].

Presumptive tests for saliva target amylase, an enzyme present in high concentrations in saliva. The enzyme is responsible for the digestion of starch [Kipps 1975]. Amylase is also present in other body fluids so its presence is only considered presumptive for saliva. It is detected using the Phadebas test, which involves a special filter paper containing a starch-dye complex that breaks down when the amylase in the sample digests the starch releasing a blue dye and producing a colour change [Hedman 2011].

A.4.1 Saliva Reference Samples

Reference samples are most commonly collected as cheek swabs. This technique involves rubbing a cotton swab against the inside of an individual's cheek to collect buccal epithelia. The swab is then air dried and packaged in a paper bag or cardboard box for transportation.

A.5. Hair, Tissue, Bone, Urine and Teeth Evidence

Hair samples can be found everywhere and are easily transferred at a crime scene and from victim to suspect or vice versa. These samples are considered as both trace evidence and as a source of DNA. When considering hair evidence for DNA analysis, the root of the hair is generally used as it contains skin cells known as keratinocytes which contain nuclear DNA. If hair is pulled out violently, then it is probable that this root bulb is still attached to the hair and so can be used to create a DNA profile [McNevin 2005]. Hair evidence is collected with sterile tweezers and transferred onto a paper. The paper is then folded with a druggist fold and then packaged in a paper container. The double packaging acts as a precaution against potential loss or damage to the hair. Each hair found at the crime scene should be packaged separately to avoid cross contamination between hairs from different individuals. If hair evidence is very small or difficult to visualize or collect, then a tape lifting method may be employed. This involves applying a piece of clear adhesive tape to the area where the hairs are suspected to be present. The tape is then carefully lifted off and affixed to a clear backed card which is packaged in a paper container. Another method for collecting hair and fibre evidence from larger areas is by vacuuming where the vacuum has a special evidence collecting filter attached to it. The problem with the vacuum method

is the increased risk of contamination and damage to any hair evidence collected especially if the filter is not properly cleaned before use [Saferstein 2011].

Biological samples such as tissue, bone, urine and teeth can also be used as sources of DNA. However, it is important to consider other forensic applications of this evidence first. For example, dental x-rays, teeth, teeth impressions and bite marks [Atsü 1998] for identifications of victims [Olaisen 1997]. Urine can also yield useful information apart from DNA such as drugs and toxins that may have been taken by the victim or suspect [Drummer 2004]. Forensic anthropologists utilize skeletal remains to gain information on the life and death of a victim by studying the bones and any trauma they have suffered [Blau 2011].

Because biological evidence can be used for a variety of forensic applications it is important that the isolation of DNA from such samples be as unobtrusive as possible while still providing enough material for the analysis process. In addition, the type of sampling employed and the reagents used in the collection and isolation processes need to be compatible with the samples downstream applications. Special consideration must be used to avoid methods that could cause PCR inhibition as this would decrease the chances of a successful amplification.

B. Isolation (Extraction) of Biological Evidence

B.1. Introduction

Once it has been determined that a sample is indeed of biological origin and can thus yield DNA evidence, there are a variety of ways the sample can be handled, and the DNA

extracted. This is a threefold issue, as the cells must be separated from the surface they are deposited on, lysed open to release the DNA, and finally the DNA purified by removal of proteins and other inhibitors that could negatively affect downstream processes [Demeke 2010]. Only a small amount of the sample collected may be extracted, while the rest should be prepared for long term storage in case re-analysis becomes necessary later on.

When performing extractions, it is important to include a negative reagent control which is treated the same as the sample but with no biological material included. Adding a negative control ensures that reagents and consumables are free from contamination. If allele peaks are detected in the negative control then the contaminating source must be found and the samples re-extracted [Giglio 2003].

B.2. Phenol-Chloroform Isoamyl Alcohol (PCIA)

Phenol-Chloroform Isoamyl Alchohol (PCIA) extraction, also known as organic extraction, is one of the earliest and most common extraction methods used to isolate DNA from biological matrices. It is still a popular technique even though newer and simpler procedures exist because of its high recovery of relatively clean DNA from a variety of different biological samples. Its downside is that in addition to requiring toxic phenol, which must be carefully handled in a fume hood with proper ventilation, it also has many steps and is difficult to automate. PCIA isolates DNA using liquid/ liquid extraction with organic reagents. First the cells are lysed and the proteins digested through the addition of an enzyme known as Proteinase K along with sodium dodecyl sulphate (SDS) a detergent and Dithiothreitol (DTT) which acts as a disulphide bond breaking agent. The SDS and proteinase K lyse the cell membrane and dissolve the proteins while the DTT reduces

disulfide bonds in the nuclear protein which further helps release the DNA [Gill 1985, Marmur 1961]. This process is often performed in a water bath heated to 56 °C for 2 - 4 hours to aid the cell lysis and protein degradation.

In a second step, the proteins are separated from the DNA by liquid-liquid extraction involving the addition of phenol, chloroform and isoamyl alcohol to the aqueous solution containing the DNA sample (Figure 2.2). The isoamyl alcohol acts as an anti-foaming agent while the phenol-chloroform partitions the hydrolyzed proteins into the organic phase [Marmur 1961]. Following centrifugation, the DNA molecules (which are more polar) are captured in the aqueous phase, while the protein and cell fragments remain in the organic phase. The aqueous phase is then pipetted into a new tube leaving the remnants of the cell in the organic sample matrix (Figure 2.2).
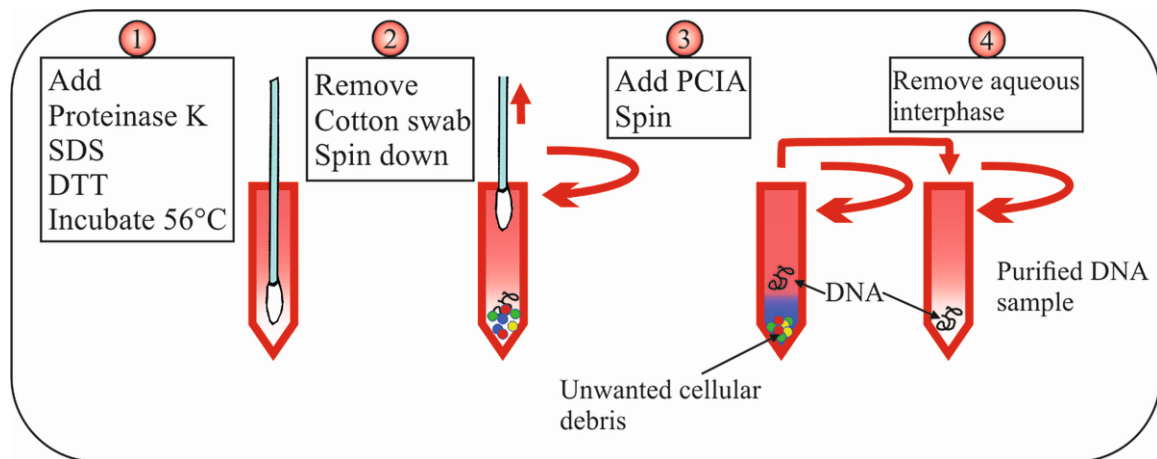
Fig.2.2 Schematic of the 4 steps of PCIA extraction process. Step 1: shows the addition of proteinase K, SDS, DTT incubated at 56°C for 2-4hrs. Step 2: shows the removal of the cotton swab and the DNA and cell fragments spun down to the bottom of the tube. Step 3: shows the addition of PCIA and the spin step that aids in the DNA migration to the aqueous phase while the cell remnants are spun to the bottom of the tube trapped in the organic phase. Step 4 shows the transfer of the DNA in the aqueous phase to a new tube ready for ethanol precipitation to be performed. Image created by Author.

B.2.1 Ethanol Precipitation

Following PCIA extraction the DNA is usually recovered and cleaned up using an ethanol precipitation step. This also helps to remove any remaining PCR inhibitors [Comey 1994]. The DNA has a highly charged phosphate backbone, making it a polar molecule that is soluble in H_2O. The polarity of water has the effect of diminishing the electrical force between any two charges present in the solution when compared to their normal force when not in solution. The result of this is that the water molecules form a "hydration shell" around the charged DNA fragments (Figure 2.3), which prevents them from precipitating out of the aqueous solution.
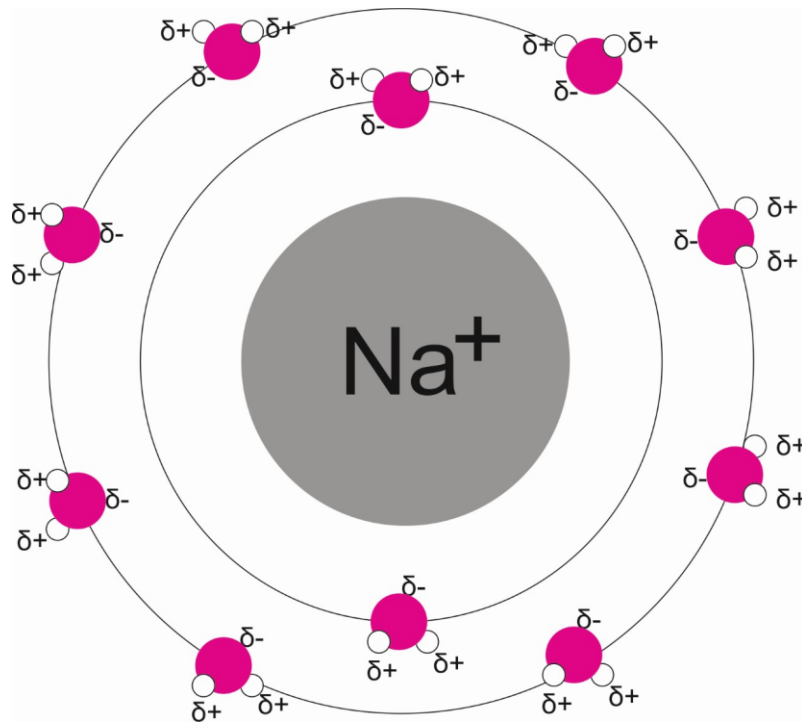
Fig.2.3 Schematic of the hydrogen shell of positive Na ion that forms around the charged DNA fragments preventing them from forming stable bonds and precipitating out of the aqueous solution. Image created by Author.

Ethanol is much less polar than water and when present in high concentrations, the shielding effect of the water is counteracted and the DNA fragments form H-bonds with salts in the solution and precipitate [Butler 2009]. The ethanol used needs to be of high concentration for this to occur, with most protocols using requiring first a dilution and centrifugation with 100% ethanol, followed by a second and third wash step using 70% ethanol. Cooling of the solution can also aid in precipitating the DNA and residual ethanol can then be left to evaporate from the sample.

B.3. Chelex Extraction

Chelex 100 (Bio-Rad Laboratories, Hercules, CA) is a highly selective cation exchange resin, composed of styrene divinylbenzene copolymer containing iminodiacetic acid groups. The iminodiacetic ions bind to metal ions such as the Magnesium (Mg^{2+}) ion and in doing so minimize the effects of PCR inhibitors and DNA nucleases which can degrade DNA and affect downstream processes [Walsh 1991, Hoff-Olsen 1999]. When used in DNA extraction, the sample is boiled in a 5% solution of deionized water and Chelex resin for several minutes. The boiling causes the cell membrane to lyse and the DNA to be released into the supernatant [Willard 1998]. The resin has a selectivity of 5000:1 for divalent versus monovalent ions, even in high salt solutions. The Chelex extraction method requires less pipetting and tube transfer steps than the PCIA method. While also being much faster, the downside of this method is that the high temperatures used cause the resulting DNA to be in single stranded form which may not be as stable as in its double stranded form. In addition, Chelex extraction should not be used on samples suspected of containing low levels or degraded DNA as it could further decrease the quality of these samples. Finally, it is important that the washing steps after extraction be carefully carried

out as any carryover of the resin into the extracted DNA can cause PCR inhibition [Walsh 1991, Hoff-Olsen 1999].

B.4. FTA® Paper

The Fast Technology for Analysis of nucleic acids (FTA) method was developed at Flinders University in Australia in the late 1980's as a method to store and preserve biological samples [Burgoyne 1994]. FTA paper is a special cellulose based paper containing a weak base, chelating reagents, an anionic surfactant and uric acid which lyses the cells and denatures the protein. The released DNA is then immobilised within the fibres of the paper where it is protected from degradation by nucleases, bacteria, UV radiation and other environmental elements, thus preserving it for long term storage [Thacker 2000]. Biological samples spotted on FTA paper can be stored at room temperature for many years with little to no DNA degradation. Because of their small size FTA papers are easy to package, transport and store, which has greatly increased their popularity in forensic use.

A biological sample is simply spotted on the FTA paper where the reagents present lyse and immobilize the DNA. Next a small punch is washed, and the DNA is eluted with a buffer that is added to PCR reactions. Similar amounts of DNA can be added to PCR reactions simply by using the same size of punch which allows the quantitation step, usually required for forensic samples, to be skipped and thus saves even more time. Direct PCR methods even permit the elimination of the wash and elution steps. The direct PCR procedure utilizes the PCR cocktail to perform elution, further reducing run times.

Modern FTA papers include colour indicators showing where the DNA is present on the paper to ensure that areas punched are those containing the DNA. The main limiting factor of this extraction method, is that during loading, the sample punches may float out of the

sample well due to static electricity, which could lead to sample loss or contamination [Smith 2004].

B.5. Solid Phase Extractions

Solid phase extraction (SPE) methods, selectively bound DNA to a solid substrate (usually silica particles or coated magnetic beads). Following cell lysis, while cellular debris, proteins and other artefacts are removed from solution by multiple wash steps and centrifugation. The DNA, now purified, is released back into solution with a final elution step.

A number of commercial SPE kits have become available in recent years such as the QIAamp® Kit (QIAGEN, Valencia, CA) [Greenspoon 1998] and the DNA IQ™ system (Promega Corporation, Madison, WI) [Frégeau 2012]. Different kits are available for tissue specific DNA extraction. These kits mainly differ in their use of buffers and bead chemistry. The underlying mechanisms of these kits involve the lysis of cell followed by the DNA selectively binding of nucleic acids to a bead based support. Up to 95% coverage of the beads is typically possible. DNA binding is performed under acidic conditions below a pH of 7.5 and in the presence of high concentrations of a chaotropic salt solution which disrupts the intra-molecular forces such as hydrogen bonding, thus aiding in cell lysis. The process also stabilizes the denatured DNA. The remaining unbound cellular debris and proteins are then washed away with water and a second wash step is performed with ethanol to remove the chaotropic salts followed by a third wash with water to remove any remaining ethanol. The DNA can then be eluted back into solution by raising the pH, making the silica beads more negatively charged [Greenspoon 1998, Nagy 2005]. The beads are then spun down by centrifugation and the DNA collected in the aqueous layer

[Nagy 2005]. These kits have been adapted by several robotic platforms such as the BioRobot EZ1 workstation (Qiagen, Inc., Valencia, CA) and QIAcube system (Qiagen, Inc., Valencia, CA) to allow for automation of the extraction process [Budowle 2005].

Another commercial kit, that employs similar chemistry but uses magnetic beads instead of silica coated glass beads, is the DNA Isolation Quantitation kit (DNA IQ) (Promega Corporation, Madison, WI). The DNA bonds to the magnetic beads and this complex then drawn to the side of the tube and immobilized using a magnet or a magnetized tube rack. The DNA IQ design allows for less pipetting steps and tube transfers as all washes can be performed in the same tube without disturbing the DNA-magnetic bead pellet. After a series of wash steps to remove cellular debris and PCR inhibitors the DNA is eluted back into solution through the use of an elution buffer and heated to 65°C [Frégeau 2012]. The main problem with this method is that if any compound (for example is chemicals from a presumptive test for blood [Poon 2009]) interferes with the DNA binding to the magnetic beads then the DNA will be lost during the wash steps. Other SPE chemistries include ion exchange, charge switch such as the ChargeSwitch Technology® (CST®, Life Technologies, Grand Island, New York) [Witt 2012], sucrose based methods and microchip based methods [Wolfe 2002, Dauphin 2010].

B.6. Differential extraction

Differential extraction was first used by Peter Gill in 1985. It is the preferred method for extracting DNA from a mixture of male and female cells, and is commonly used with sexual assault evidence [Gill 1985]. Differential extraction separates the male from the female fractions resulting suspect profiles that are less convoluted and easier to interpret [Schoell 1999, Schoell 1999]. Differential extraction exploits the different DNA packaging

in sperm and other tissues. Sperm contains protamines which are proteins that contain high concentrations of arginine, cysteine and charged protein residues. These proteins form disulphide bonds which causes the resulting DNA to be packed much more densely. This process makes it harder to denature and release the DNA [Oliva 2006]. In order for the DNA in sperm to be released, these disulphide bonds must be broken, requiring much harsher lysis conditions then those employed with other cell types.

Organic differential extraction is a two-step process. First the vaginal epithelial cells are lysed with protinease K and SDS in the solution. This step does not affect the sperm cells and these are collected and removed from the solution by centrifugation (Figure 2.4). The second step consists of lysing the remaining sperm cells using a buffer containing proteinase K, SDS and DTT reducing agent. The DTT targets the cross-linked thiol-rich nuclear proteins of the sperm, weakening them and making the cells more susceptible to organic extraction by the Proteinase K and SDS [Gill 1985, Getts 1999]. Both the female and male fractions are purified through PCIA and ethanol precipitation which can then each produce cleaner profiles of their individual contributors.
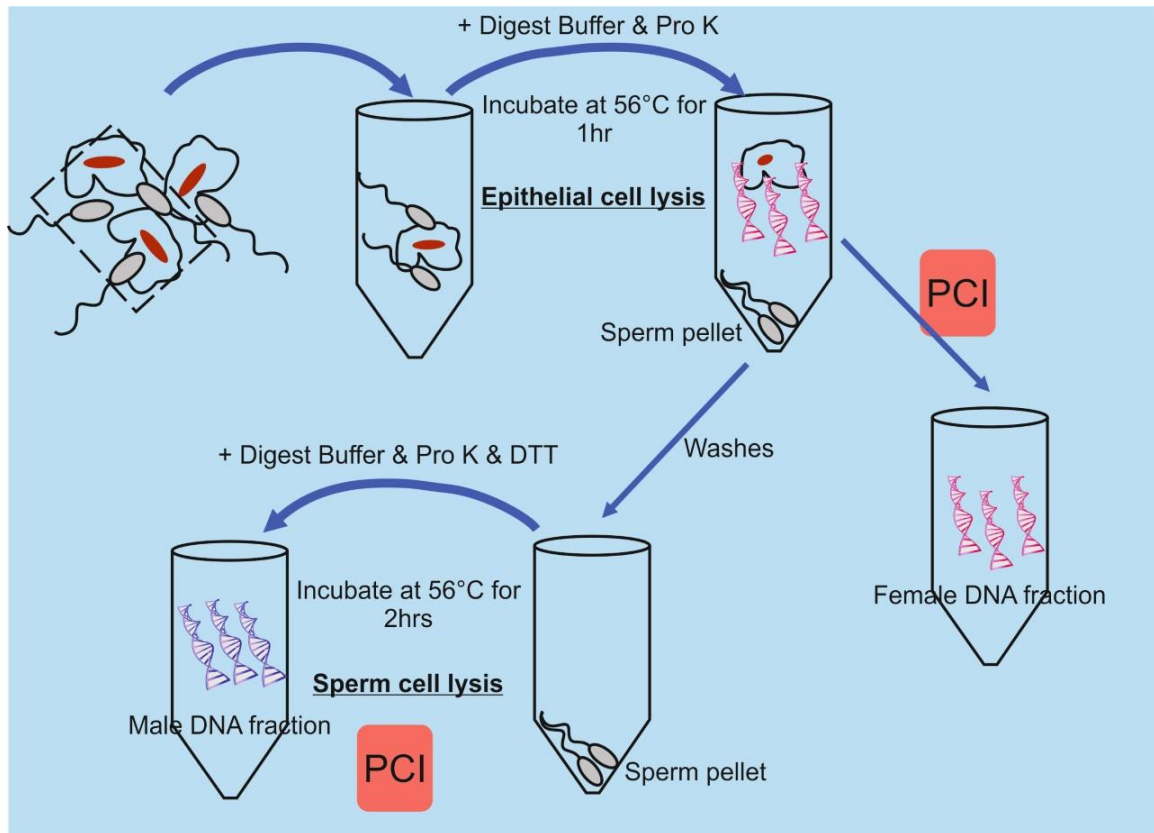
Fig.2.4 Schematic of Organic Differential Extraction. First the biological sample containing male and female cells is incubated in a cell lysis buffer of SDS and Proteinase K, resulting in the lysis of the female epithelia. Centrifugation allows female fraction to be removed with top aqueous layer. DTT, SDS and Proteinase K buffer then used to lyse the sperm pellet and release the male DNA fraction, which is isolated and purified through normal PCIA and ethanol precipitation methods. Image created by Author.

While this method is still in favour for separating male to female mixtures it is time consuming, with multiple transfer steps and is difficult to automate. If proper control of the process is not employed, then inefficient separation may occur and the resulting profile becomes a confusing mixture of male and female DNA [Vhen 1998, Hudlow 2012]. An additional problem arises if the male contributor was vasectomized or azospermic, as no sperm cells would be present and thus this method would not be able to separate the two fractions [Butler 2011].

B.7. Laser microdissection

Laser microdissection (LM) or laser-capture microdissection (LCM) is a method of selectively recovering cells visualised by use of a microscope. This method can be used to separate and isolate sperm cells from others in evidence from sexual assaults [Elliot 2003, Sanders 2006]. Laser microdissection is performed is by spreading some of the sample on a microscope slide and covering it with a plastic film. Next, the cell of interest is located and then a laser is utilized to cut the cells from the plastic film and capture them. This process is repeated for all desired cells and their DNA can be extracted by any standard method. There are two concerns with this procedure, the first involving the possibility of contamination due to exogenous cells. The second concern is that the heat produced by the laser to melt the film could cause degradation of the cell and the DNA it contains [Murray 2007]. To solve the contamination issue, a UV system can be used in which the sample is transferred to a polymer slide that is then cut out along with the cell of interest and collected in a tube without ever contacting other substrates that may contaminate the sample [Vandewoestyne 2010]. Laser microdissection has proved to work more effectively than differential extraction in singling out the male sperm cells from male and female mixtures [Elliot 2003] but it is not as widely used as it is less cost effective and also yields fewer cells for later extraction.

CHAPTER III. DNA AMPLIFICATION: POLYMERASE CHAIN REACTION

A. PCR: Basics

In 1985 Kary Mullis, discovered the Polymerase Chain Reaction (PCR) [Schochetman1988]. The PCR can produce millions of exact copies of a single DNA sequence. In the field of forensics the invention of PCR permitted successful sample analysis of samples which previously would not have contained enough DNA or where too degraded for downstream analysis. This discovery led to Kary Mullis being awarded the Nobel Prize in 1993, further illustrating the significance of his discovery and its contribution to science [Mullis 2000].

The Polymerase Chain Reaction (PCR) is a selective in vitro amplification of a defined target DNA sequence. PCR includes 25-45 repeating cycles in which not only is the DNA present doubled each cycle (if 100% efficiency is maintained) but the new DNA created in one cycle can serve as the template for the next, giving an exponential rise in the amount of DNA [Butler 2011]. The PCR takes advantage of an enzyme that uses a defined sequence in a strand of DNA as a template for synthesising a complementary strand. The reaction occurs as a result of thermal cycling (repeated heating and cooling of a DNA sample), using primers containing short single stranded DNA fragments complementary to the target region, along with a thermally stable DNA polymerase. The primers used are generally labelled with fluorescent dyes which can be excited to permit detection of the different fragments. Over the years there has been much research conducted on improvements in amplification by PCR with special interest being given to the reduction of cost [Schuelke 2000] and automation.

PCR is completed in three basic steps (denaturing, annealing and elongation), all occurring at different temperatures. Additional steps may be included, depending on the protocol, such as an initialisation step at the beginning of the process in which the sample is heated to 98°C for 5-15 minutes. The initialisation step is needed to activate certain stabilized (hot start) enzymes. The denaturing step consists of separating the double stranded DNA into two single strands (>90 °C) by increasing the temperature to a point where the hydrogen bonds holding the 2 strands together weaken and break. During annealing, the temperature is lowered to 40-65°C to allow primers to bind to their complementary region on the single stranded template DNA. The temperature at which this occurs depends on the melting temperatures of the primers used. If using multiple sets of primers, a median temperature will be selected. Primers must be sufficiently long to uniquely bind the selected regions that will be amplified. A forward and reverse primer is used to bracket the location of interest. Finally, an elongation step is used in which the sample is heated to ~72°C. At this temperature the polymerase is optimally active and binds to the primers adding complementary nucleotides to the template strand [Delidow 1993, Demeke 2010]. Temperatures and times at each step vary depending on protocol and sample. This 3-step process can be seen in Figure 3.1 and 3.2.
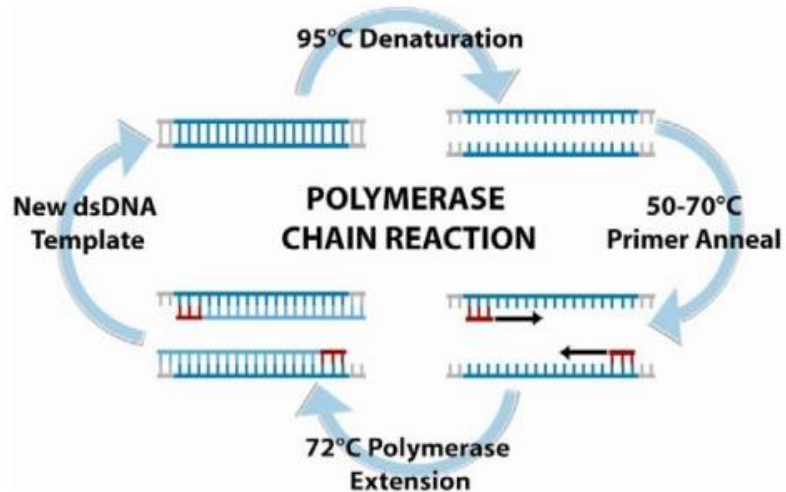
Fig.3.1 Schematic overview of the PCR process. First denaturation takes place once the sample is heated 95°C and the 2 strands separate from each other forming single stranded DNA. Next primers anneal to their target regions takes at a lower temperature (50-70°C). Polymerase then binds to primers and adds nucleotides to the new growing strand of DNA (72°C) resulting in a copy of the original DNA template. This process is repeated multiple times depending on cycle number until enough DNA has been produced to be visualized at a later step by electrophoresis. Image from Lui 2009.

In a forensic setting, most PCR reactions amplify STRs of ~60-500 base pairs (bp) in length. The final amount of amplified product is determined by the available reagents in the reaction, and the feed-back and inhibitory effects produced by the reaction products [Schochetman1988]. Theoretically, the number of copies of each STR at the end of a PCR is roughly 2^n (n= number of amplification cycles) permitting an estimation of the number of theoretical copies of the DNA produced depending on the number of cycles used. For example, 35 cycles can produce $2^{35} = 3.43597 \times 10^9$ copies [Butler 2011]. Of course in actuality fewer copies are produced due to effects of polymerase efficiency, crowding, and limiting reagents [Butler 2011].

To check whether the PCR generated the anticipated DNA fragment (amplicon), the PCR products are separated by either agarose gel or capillary electrophoresis in which separation of the fragments occurs by size. These sizes can then be compared to an internal lane

standard that contains fragments of known sizes to allow size identification of the PCR amplicon.

B. PCR Component

The polymerase chain reaction requires a mixture of reagents in optimal concentrations in order to successfully proceed to produce a desired product, also known as an amplicon. The components in the reaction mix include the polymerase, $MgCl_2$, forward and reverse primers, dNTPs, buffer, PCR enhancers and template DNA.

The polymerase used is an enzyme that binds to the primers and extends the new DNA strand by incorporating complementary nucleotides to those on the template strandInitially a Klenow fragment (derived from *E. Coli.*) was used in PCR [Mullis 1990]. It was found that the enzyme was rendered inactive after each cycle of the PCR by denaturation at the high temperatures needed to separate the DNA strands. This meant that it had to be replenished manually after every denaturation step which proved labour intensive and undesirable. A solution to the problem was the discovery of the bacteria *Thermus Aquaticus* and its derived polymerase Taq. This bacterium was found living in hot springs and its proteins (including its polymerase Taq) were thermostable at temperatures of 95-98 ºC allowing it to remain active through the various high temperature cycles of PCR without the need of replenishing. This allowed for the automation of the PCR process in instruments now known as thermocyclers. For this reason, the polymerase used in most modern PCR reactions is some form of Taq (AmpliTaqGold, Z-Taq, OmniTaq, Cesium Taq, Etc.) [Innis 1999].

The speed at which the polymerase incorporates new nucleotides is known as processivity, although what processivity actually represents is the average number of nucleotides added by a polymerase before it dissociates from the template strand. The percentage of correct nucleotides added is known as fidelity. Persistence represents how stable the enzyme is at high temperatures. All three are important factors when considering and comparing polymerases for different uses, but for STR amplification the effect of low fidelity polymerases or those lacking proofreading activity is negligible as the regions amplified are relatively short.

In recent years mutant Taqs have been created in laboratories to adhere to ever more specific needs but these will be discussed in detail in the chapter: Fast Thermocyclers and Rapid Polymerases.

Magnesium Chloride ($MgCl_2$) acts as a co-factor that binds and activates the polymerase [Innis 1999], while also maintaining its stability. The concentration of $MgCl_2$ is important, as too much can cause unspecific amplification while a deficiency will lead to polymerase activation failure and no reaction.

Primers are small single strands of DNA that are complementary to a target site on the DNA template [Frégeau 2012]. Two primers are needed to successfully amplify a region and they are known as the forward and reverse primers due to their 5'-3' directionality. Primers are designed to bind to a target region on either side of a designated amplification region in a bookend like fashion. Optimal primer length is between 15-35 bp and primers should be short enough to not add too many nucleotides to the amplicon sequence while remaining long enough to ensure specific binding to the target site. Primers should also

have a similar GC percentage for optimal binding to the DNA template [Wu 1991]. The size and location of primer binding sites can be changed to alter the size of the amplicon as desired. This design feature is employed when preparing multiplex amplifications in which the different amplicons must not produce amplicon size overlap. It is also important when designing primers that they do not have sequences complementary to themselves, their partner or any other pairs of primers. If using a multiplex, this complementarity can cause the formation of secondary structures such as primer dimers and hairpins reducing the primers ability to bind to the template [Vallone 2004]. For the purposes of rapid PCR elevated concentrations of primer need to be used as this improves amplification rates due to diffusion effects.

Deoxynucleotide Triphosphates (dNTPs) are the building blocks of DNA and they are used by the polymerase to create each new strand of DNA. Each of the four dNTPs are added to the PCR in a mixture of equal amounts.

The buffer solution is an important component to the PCR as it maintains an environment that is stable at the optimum pH and salt levels for the polymerase activity. This buffer usually contains $MgCl_2$ but it is not uncommon for it to also contain PCR enhancers (also known as inhibitor facilitators) such as bovine serum albumin (BSA) and Dimethyl sulfoxide (DMSO). These compounds are known to increase reaction efficiency and reduce the effects of inhibitors [Mamedov 2008]. Bovine serum albumin (BSA) aids in the reaction by binding to inhibitors, preventing them from binding to the polymerase and halting its activity [Satoh 1998]. DMSO is thought to facilitate the unwinding of the DNA from its associated histone proteins, making it more available for polymerase binding. This

increases its processivity while also decreasing the amount of non-specific binding of the enzyme [Delidow 1993]. Other compounds added to the PCR mixture include commercial PCR enhancers cocktails, such as PEC-1, PEC-2 (DNA Polymerase Technology) which are designed to counteract inhibitors or improve the recognition of sequences that are GC rich. These additives are often used in combination with inhibitor resistant polymerases such as OmniTaq (DNA Polymerase Technology).

The final component in PCR is the template strand which is the genomic DNA sample that is copied. Template DNA is usually extracted from a biological sample. In the case of Direct PCR, no initial extraction is performed and the biological sample is added in its untreated form to the PCR. Inhibitor facilitators and PCR enhancers are also used to assist in this process. It is also important that the DNA template be as clean and uncontaminated as possible as due to the sensitivity of the PCR, as even small amounts of contamination can produce copies that will later confuse data interpretation. This is particularly true when small amounts of template DNA are used. Finally, the amount of input template must also be addressed as too high a concentration of template can cause non-specific amplification or overloading of sample. If the input DNA concentration is too low allele dropout or amplification failure can occur [Bell 2008, Butler 2005].

C. PCR Controls

One commonly employed quality assurance methods for PCR is the use of control samples that are run alongside normal samples, these controls allow the analyst to check that the system is working correctly and aid in troubleshooting. There are three types of controls used in forensic laboratories: a positive control, a negative control and a blank. The positive

control is a high-quality DNA standard for which the origin and the concentration is established and the profile is known. The positive control is used to ensure all reagents are working properly, and that there are no problems in the amplification process. The negative control contains all the reagents of the PCR master mix but no template DNA, thus the resultant amplification should produce no peaks in the resulting profile. The negative control can also be used to detect primer dimers, dye blobs and contamination events. Finally, a blank is used which contains just the reagents for running samples on the electrophoresis instrument. This typically consists of a mixture of formamide and internal lane standard (ILS). The blank helps detect the presence of post PCR contamination as well as making sure there are no issues with reagents or capillary degradation. If using extracted DNA then another type of control, known as an extraction control may also be included. The extraction control is a sample that has no DNA template but goes through all the processes from extraction through to detection and can help detect additional contamination that may have occurred during the extraction process. These controls are important especially for quality assurance and accreditation purposes in forensic laboratories and they should be run with all samples [Innis 1999].

D. Thermal Cycling

As mentioned above, the PCR follows the principle of thermocycling in which the analyte mixture is treated with a series of 20-40 repeated temperature changes (cycles), with each cycle normally having 3 temperature steps or changes. The temperatures used and the length of time they are applied in each cycle depend on a variety of parameters, such as the

enzyme used, the concentration of divalent ions and dNTPs in the reaction, and the melting temperature of the primers [Rychlik 1990].

The steps involved are shown in Figure 3.2 and are as follows:

1. Initialization step: heating the reaction to a temperature of 94–98 °C which is held for 1-15 minutes to activate hot start polymerases and aid in denaturing of dsDNA.

2. Denaturation: First part of a cycle, heats the reaction to 94–98 °C for ~30 s, breaking hydrogen bonds between the dsDNA to give ssDNA.

3. Annealing step: Temperature is lowered to 50–70 °C (primer Tm dependent) for ~30 s so that the primers anneal to the new single stranded template molecule.

Fig.3.2 Schematic of the PCR thermocycling steps. Deneaturation at ~95°C, Annealing ~50°C and Elongation ~72°C. Image from https://www.floridamuseum.ufl.edu/cowries/amp

4. Extension/elongation step: The temperature at this step depends on the DNA polymerase used. If using Taq polymerase then a temperature of 70-72 °C is set. The addition of nucleotides occurs in the 5'-3' direction and continues until the polymerase reaches the end of the strand and disassociates from the DNA. The time it takes to synthesis a new

strand depends on the polymerase and the size of the desired product, 1-3 minutes is the typical.

5. Final elongation or incubation: After all the cycles are complete, a step is added in which a temperature of 60–74 °C is used for ~ 10 minutes after the last PCR cycle to ensure that any remaining single stranded DNA is fully extended. This also helps to standardize adenylation, a situation in which the polymerase adds an additional base at the end of a fragment.

The cycling process is completed in an instrument known as a thermal cycler, which rapidly heats and cools a metal block in cycles set to the PCR parameters. These are much advanced from the first thermocyclers which required the addition of mineral oil to the top of each sample in order to prevent condensation and evaporation from occurring (the heated lid solved this problem) [Innis 1999]. There are many instrument designs differing in sample numbers, tube sizes and ramp rates. One example (Figure 3.3) is the GeneAmp 9700 from Life Technologies (previously Applied Biosystems).

Fig.3.3 GeneAmp® PCR system 9700 thermal cycler from Life Technologies. Temperature ramp rate 4°C/s, fits 96 samples at once and has heated lid to prevent condensation. From Thermo Fisher Scientific website.

All thermal cyclers include computer control to adjust the cycling parameters that vary by reaction protocol. Parameters that can be adjusted include the temperature and time for each step of the PCR, the number of cycles of denaturing, annealing and elongation and the addition of extra steps before or after these cycles.

E. PCR Inhibition

E.1 PCR Inhibition

Perhaps the major problem affecting DNA analysis is PCR inhibition. Certain compounds present in the sample matrix or used to preserve it inhibit PCR by making cell lysis more difficult or through binding to the DNA polymerase or template [Funes-Huacca 2011]. The process can lead to poor or no amplification [Demeke 2010, Alaeddini 2012]. Poor amplification is a particularly problematic when analysing blood samples, as not only do such samples contain PCR inhibitors such as haemoglobin, heparin, lactoferrin and immunoglobulin G, they may also contain anticoagulants such as EDTA and sodium citrate. Other inhibitors include haematin, calcium, humic acid (from soil) and indigo (dye used in denim) [Wilson 1997, Larkin 1999, Young 1993, Opel 2010, Thompson 2012]. Haematin and calcium from blood and bone inhibit the PCR by competing with the magnesium to bind to the polymerase [Akane 1994, Bickley 1996]. Some inhibitors can also be introduced into the sample during collection or preamplification steps [Willard 1998, Butler 2011], such as traces of phenol from improperly washing or pipetting during a PCIA extraction [Demeke 2010]. For this reason, a number of studies have been performed to determine what compounds are inhibitory and what polymerases work with each type of inhibitor (Table 3.1) [Al-Soud 2000- 2001, Funes-Huacca 2011]. It has been

found that many of the less inhibiting factors could be removed simply by reducing their concentration. The removal of inhibiting factors is a relatively easy process for forensic purposes, as sample quantities used are already extremely small and often dilution is possible [Byrnes 1975]. Other methods to counteract inhibition include the addition of higher concentrations of DNA polymerase as even if some polymerase was bound to inhibitors and rendered non-functional the rest would still be active and able to complete the PCR [Butler 2011].

| DNA polymerase | PCR result with the indicated concn of[a]: | | | | | | | | | | | | | |
| | Hemoglobin (μg/reaction tube) | | | | | | | | Lactoferrin (ng/reaction tube) | | | | | |
	0	1	1.3	0.4	2	4	10	100	0	0.5	1	2.5	5	25
AmpliTaq Gold	+, +	±, ±	−	−, −	−	−, −	−	−, −	+, +	+, +	+, +	+, +	+, +	−, −
DyNAzyme II	+, +	+, +	+	+, +	+	+, +	+	±, −	+, +	+, +	+, +	+, +	+, +	−, −
DyNAzyme EXL	+, +	+, +	+	+, +	+	+, +	+	+, ±	+, +	+, +	+, +	+, +	+, ±	−, −
HotTub	+, +	+, +	+	+, +	+	+, +	+	+, −	+, +	+, +	+, +	+, +	+, +	−, −
Pwo	+, ±[b]	−, −	−	−, −	−	−, −	−	−, −	±, ±	±, ±	±, ±	±, ±	−, −	−, −
rTth	+, +	+, +	+	+, +	+	+, +	+	+, +	+, +	+, +	+, +	+, +	+, +	+, −
Taq	+, +	+, +	+	+, +	+	+, +	+	±, −	+, +	+, +	+, +	+, +	+, +	−, −
Tfl	+, +	+, +	+	+, +	+	+, +	+	+, ±	+, +	+, +	+, +	+, +	+, +	+, −
Tli	+, +	+, +	+	+, +	+	+, +	+	+, +	+, +	+, +	+, +	+, +	+, +	−, −
Ultma	+, ±	−, −	−	−, −	−	−, −	−	−, −	+, +	+, +	+, −	±, −	−, −	−, −

[a] Results of two independent PCRs. +, PCR product of high yield; ±, PCR product of low yield; −, no PCR product. Purified hemoglobin conc. and human milk lactoferrin. (Sigma Chemical) were added to the PCR mixtures of the 10 DNA polymerases containing 1 ng of *L. monocytogenes* DNA.
[b] *Pwo* had many nonspecific products.

Table 3.1 Inhibitory effects of different concentrations of haemoglobin and lactoferrin on the amplification capacities of 10 thermostable DNA polymerases. From Al-Soud 2001.

Overall though, it was considered that avoidance of co-extracting the known inhibitors along with the DNA when possible was the best approach [Alaeddini 2012]. But how can we tell if a sample is degraded and thus not producing a useable profile, or if the PCR is simply inhibited? This is important to know so that the sample can be treated accordingly and the quality of the results improved. Real time PCR and internal positive controls (IPCs) can be used to determine if the sample is degraded or suffering from inhibitory effects [Opel 2010, Thompson 2012].

E.2 Inhibition Facilitators

Al-Soud tested ten facilitators in combination with different amounts of haemoglobin and lactoferrin using AmpliTaq Gold [Al-Soud 2001]. The results of the effect of facilitators to improve inhibition resistance can be seen in Table 3.2. [Al-Soud 2001]. It was found that the addition of small amounts of bovine serum albumin (BSA), approximately ~0.4%, was the most efficient way of counteracting inhibition by both haemoglobin and lactoferrin. Betaine and gp32 (11.7% and 0.01% respectively) also showed some relief of inhibition for haemoglobin but only gp32 had an effect on lactoferrin inhibition.

| Facilitator (%, wt/vol) | PCR result[b] with the indicated concn of: | | | | | | | | | | | | | | |
| | Hemoglobin (µg/reaction tube) | | | | | | | | Lactoferrin (ng/reaction tube) | | | | | | |
	0	1	1.3	2	4	10	20	100	0	5	25	50	125	250	500
Water (0)	+, +	±, ±	−, −	−, −	−, −	−, −	−, −	−, −	+, +	+, +	−, −	−, −	−, −	−, −	−, −
Betaine (11.7)	+, +	+, +	+, +	±, ±	±, ±	±, −	−, −	−, −	+, +	+, +	−, −	−, −	−, −	−, −	−, −
BSA (0.4)	+, +	+, +	+, +	+, +	+, +	+, −	−, −	−, −	+, +	+, +	+, +	+, +	+, +	+, +	+, +
Dextran 50 (1)	+, +	±, −	−, −	−, −	−, −	−, −	−, −	−, −	+, +	+, +	−, −	−, −	−, −	−, −	−, −
Dimethyl sulfoxide (1)	+, +	±, −	−, −	−, −	−, −	−, −	−, −	−, −	+, +	+, +	+, −	−, −	−, −	−, −	−, −
Glycerol (5)	+, +	±, −	−, −	−, −	−, −	−, −	−, −	−, −	+, +	+, +	−, −	−, −	−, −	−, −	−, −
gp32 (0.01)	+, +	+, +	+, +	+, +	+, +	±, ±	−, −	−, −	+, +	+, +	+, +	+, ±	−, −	−, −	−, −
Nonidet NP-40 (0.1)	+, +	±, −	−, −	−, −	−, −	−, −	−, −	−, −	+, +	+, +	+, −	−, −	−, −	−, −	−, −
Polyethylene glycol 400 (1)	+, +	±, −	−, −	−, −	−, −	−, −	−, −	−, −	+, +	+, +	−, −	−, −	−, −	−, −	−, −
Tween 80 (0.1)	+, +	±, ±	±, −	−, −	−, −	−, −	−, −	−, −	+, +	+, +	−, −	−, −	−, −	−, −	−, −

[a] Amplification facilitators were added to reaction mixtures of Ampli*Taq* Gold containing 1 ng of *L. monocytogenes* DNA and different concentrations of purified hemoglobin and human milk lactoferrin.
[b] Results of two independent PCRs. +, PCR product of high yield; ±, PCR product of low yield; −, no PCR product.

Table 3.2 Effects of nine amplification facilitators on the inhibitory effects of haemoglobin and lactoferrin on the amplification capacity of AmpliTaq Gold. From Al-Soud 2001.

These results indicate the importance of adding optimal concentrations of BSA to PCR master mixes when amplifying whole blood or saliva samples on FTA paper.

Another method alters normal extraction procedures by lysing the cell using alkali conditions. An aliquot of the released DNA in the solution can be transferred to the PCR reaction mixture. Usually this type of lysis requires additional steps such as neutralisation of the lysates prior to PCR. However, it was found that polyethylene glycol (PEG) at a concentration of over 50% at alkaline pH does not require these additional steps

[Chomczynski 2006]. Chomczynski added 60% PEG 200 (tetraethylene glycol) at a pH of 13.5 with KOH. This works as the PEG 200 stabilises the solution as it reduces the solubility of CO_2, which if left destabilised would neutralize the reagent over time [Chomczynski 2006].

Samples taken directly from cheek swabs show much less inhibition as they are cleaner samples than others such as blood and do not contain many important inhibitory factors such as haematin and lactoferrin.

F. Real Time PCR

Real time or quantitative PCR(QPCR) is a method used to quantify the amount of DNA present in a biological sample. QPCR can also be used to predict how well a sample will amplify during subsequent PCR analysis conditions and detect the presence of inhibitors by analysing amplification efficiency [Higuchi 1993, Park 2007]. The amount of DNA present in the sample is important for downstream DNA analysis as an optimal amount of template is needed for good quality data to be produced at the end of the analysis. By first determining how much template is in a sample by QPCR the sample can then either be diluted or concentrated as necessary ensuring the optimal amount of DNA is used for STR amplification.

Quantitative PCR proceeds as a normal PCR, however in addition the amplification process can be monitored through the addition of a fluorescent dye, which can either be attached to the primers/probe or using a dsDNA intercalating dye such as SYBR Green. The increase in product DNA is proportional to the fluorescence produced. There are two ways this can occur: fluorescent accumulation or quenching. In the first method, the fluorescent signal

53

produced by the dyes increases with the increase in DNA product, while in the latter, less common method, the fluorescent signal decreases as the fluorescent probe is displaced or quenched following DNA product formation.

A typical QPCR will have four distinct phases (Figure 3.4): the initial phase where little to no product is visible, the exponential phase where the DNA amplifies exponentially with each cycle, the linear phase where the reaction begins to slow down due to reagent depletion [Delidow 1993] and the plateau phase where the reaction produces no more PCR product as all the reagents have been depleted or the fluorescence has exceeded the limits of detection [Butler 2011]. The cycle threshold (Ct) is what the DNA quantitation is derived from, this is the cycle at which the fluorescence intensity exceeds the background noise.
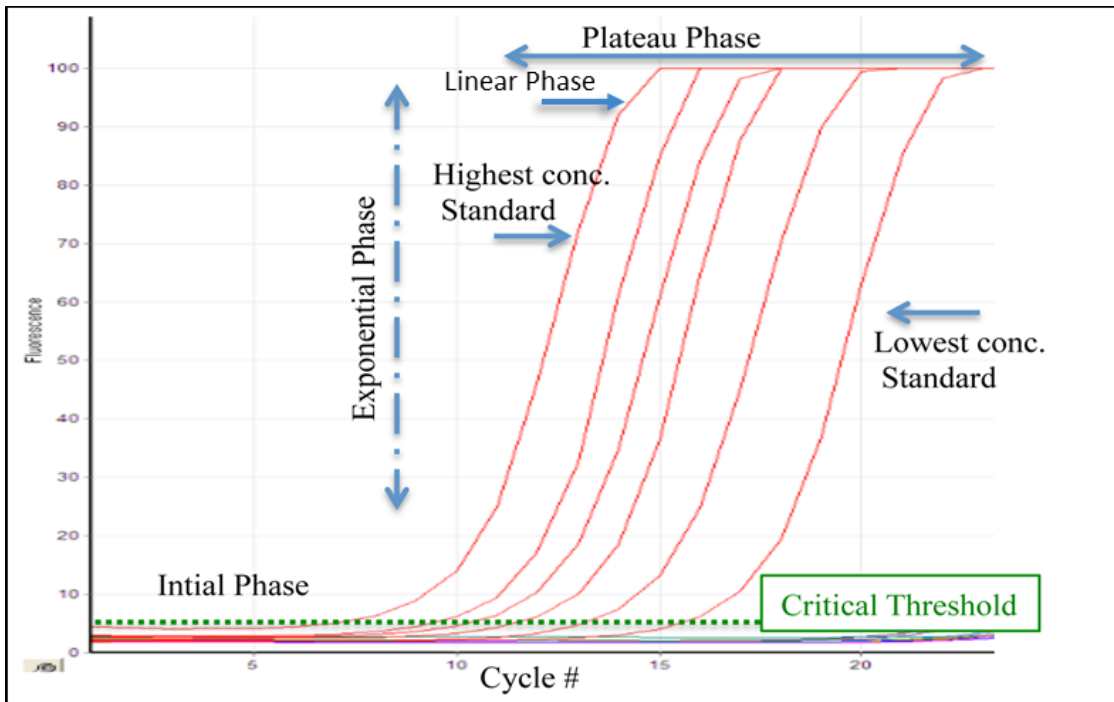
Fig 3.4 An overview of the QPCR process. The initial phase where there is little product, the exponential phase where the amount of product is produced exponentially, the linear phase where the reaction starts to slow down and the plateau phase where reagents are consumed and the reaction completed. Image created by Author.

54

There are numerous methods for fluorescence-based quantitation but the most common methods use intercalating dyes (such as SYBR® Green), hydrolysis probes (Taqman) or Plexor® dye chemistry (commercial kit from Promega). Intercalating dyes such as SYBR® Green [Giglio 2003], can be used to monitor the concentration of double stranded DNA by intercalating between the strands and binding to the minor groove of the DNA double helix. When bound to DNA in solution, the fluorescence of the dyes is enhanced as their internal pi-bonds (of their aromatic structure) can interact with each other [Kubista 2006]. For forensic purposes, only the quantity of human DNA present in a sample needs to be determined, so it is important to use real time PCR methods that only target human DNA [Williams 1994].

Hydrolysis or Taqman probes consist of sequences of DNA that are complementary to a sequence on the template strand. On the 5' end of the probe is a fluorescent labelled oligonucleotide, acting as a reporter dye, while on the 3' end is a quencher [Bell 2008], which prevents fluorescence when in close proximity to the reporter dye through fluorescence resonance energy transfer (FRET). As the probe binds to the DNA template and PCR begins, the Taq polymerase will make its way to the 3' end of the probe and as a result of its exonuclease activity, it will degrade the probe, removing the quencher molecule and allowing the reporter dye to fluoresce [Heid 1996]. In forensic analysis, both intercalating dyes and Taqman probes have shown to give similarly accurate results [Srinivasan 1993], but the Taqman probes are sequence specific permitting multiplex analysis.

The last type of chemistry used for forensic quantification is found in the Plexor quantification kit by Promega. This system utilizes specific interactions between the modified nucleotides isoguanine (Iso-G) and 5'- methylisocytosine (Iso-C). One primer is labelled with a fluorescent dye and an Iso-C at the 5' end, while the other primer remains unlabelled. dNTPs modified with iso-dGTP with Dabcyl, which is a quencher are used in the PCR mix and as amplification takes place the specific interactions between the Iso-C and the Dabcyl Iso-G are preferential at the complimentary position resulting in the quenching of the fluorescence signal. As a result, the fluorescent signal decreases as the reaction progresses and more dsDNA is produced (Figure. 3.5)
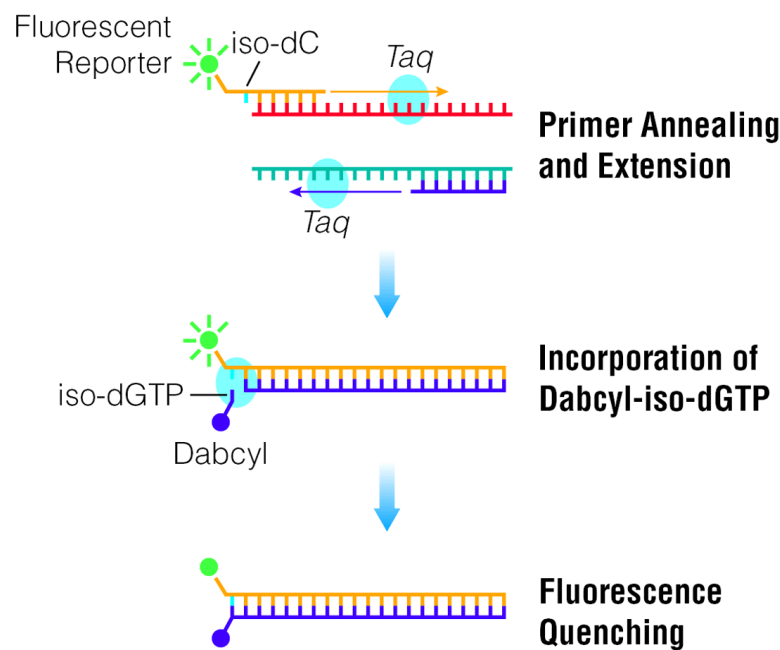
Fig.3.5 Schematic representation of the Plexor® QPCR chemistry. Fluorescence signal is originally high and then decreases. This quenching occurs when the nucleotides modified with quenching molecules get incorporated opposite the complementary nucleotide tagged with the fluorescent reporter. Adapted from Promega Corporation website.

Some of the modern quantification kits offer the ability to determine the total amount of human DNA and male-specific DNA in a sample at the same time (Plexor® HY and

Quantifiler® DUO). It is also possible to include mitochondrial DNA quantification [Walker 2005]. Multiplex analysis is important when the sample may need to be analyzed for Y-STRs or when there it is a mixture with a high female-to-male ratio.

The type of QPCR employed can affect the sensitivity of the system and its ability to correctly determine the concentration of low copies of DNA. Multi-locus probes can allow for multiplexing and detection of autosomal, mitochondrial and Y-chromosome DNA and are capable of detecting low amounts of DNA [Walker 2005]. For example, the Plexor HY can detect levels of input DNA as low as 6.4 pg/µ.

G. Rapid PCR

Research efforts to improve the speed of the PCR date back to the work of Widmer in the early 90s, in which his group achieved high speed DNA amplification using rapid heating in glass capillaries. Demonstration of rapid PCR for forensic analysis has also been discussed by Vallone et. al. [Vallone 2008] among others [Laurin 2012, Verheij 2012, Aboud 2013, French 2002].

In his early work Wittwer discovered through a series of experiments [Wittwer 1991] that speeding up the PCR reaction gave better results than normal speed PCR as it reduced the amount of undesired side products that can form in these reactions. Figure 3.6 shows the results of a PCR done using different speeds for the different cycle steps.
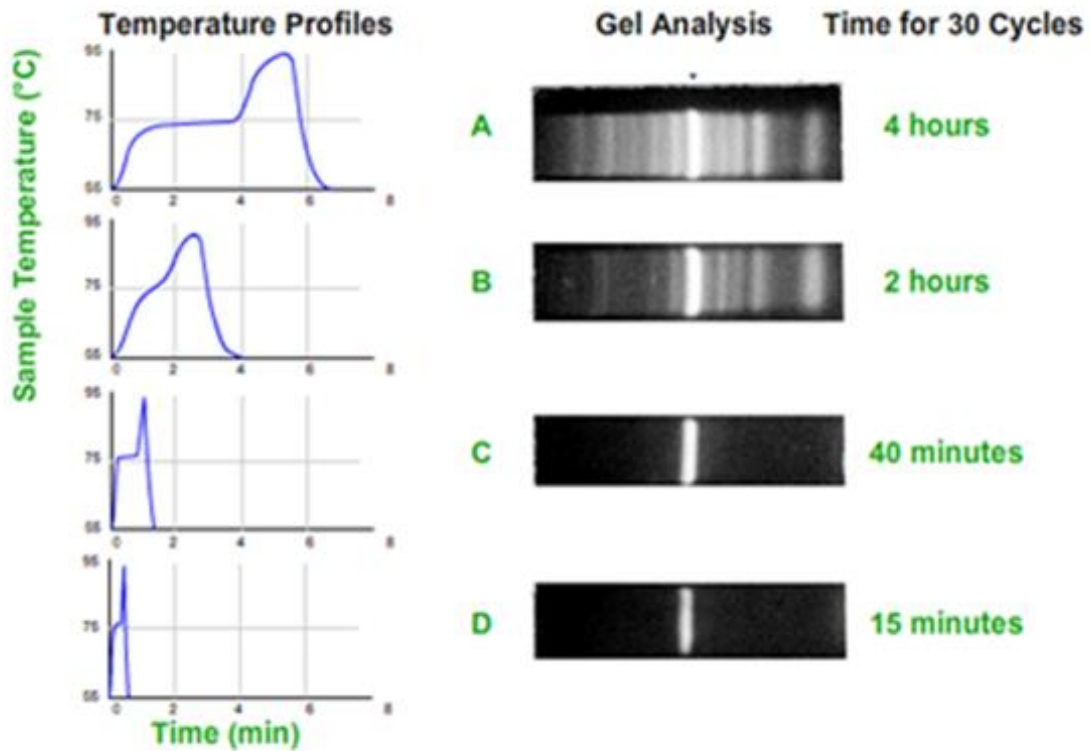
Fig.3.6 Amplification of a 536bp ß-Globin fragment of genomic DNA using different cycling conditions to give decreasing overall time for the same 30 cycles. Frome Wittwer 1991.

As can be seen from the Fig.3.6 both the 4 and 2 hour PCRs show large amounts of nonspecific product formation on either side of the desired product, while the 40 minute and even the 15 minute runs show none of these extra bands. This result indicates that faster cycling conditions do indeed increase specificity of the reaction. Wittwer does say however, that this increase in speed leads to a slight decrease in amount of desired product produced, but that this mainly affects larger amplicons, with products of <200bp showing no sign of this decrease in product formation as speed is increased. This is good news as the alleles used for forensic purposes are small, and rarely over the 536bp amplicons used in the above example.
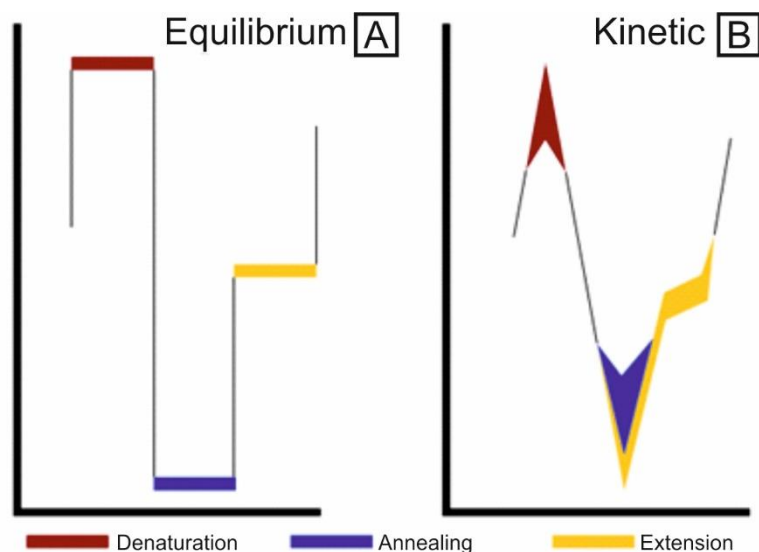
Fig.3.7 Wittwers' two described models of PCR. From Wittwer 1991.

It is generally thought that the PCR reaction occurs according to model A (Figure 3.7.), that is, that each step occurs at a different temperature and separately from each of the other steps. While this makes it easier to understand and graph, Wittwer described that this is not entirely the case and what really happens is closer to what is illustrated in model B (Figure 3.7.). Model B shows that annealing and extension happens concurrently, with the extension beginning as soon as the primers anneal and the polymerase attaches. The kinetic model also shows that there is a range of temperature on both sides of a programmed temperature where activity is taking place. For this reason, the time spent at these steps can be reduced, as activity is occurring both as the reaction is heating up to the desired temperature and as it is cooling down or moving on to the next step.

Wittwer ran a series of tests [Wittwer 1991] to discover which steps could be sped up and by how much. The results of his study can be seen in Figure 3.8.

Denaturation Time at 93°C (sec)	Ramp Time 92° to 55°C(sec)	Annealing Time at 55°C	Elongation Time at 77°C (sec)
<1	9	<1	160
2	25	<1	80
4	25	5	60
8	25	10	40
16	25	20	20
32	25	40	10
64	25	80	5
PhiX174 RF Hae III Digest	25	PhiX174 RF Hae III Digest	2.5
			PhiX174 RF Hae III Digest
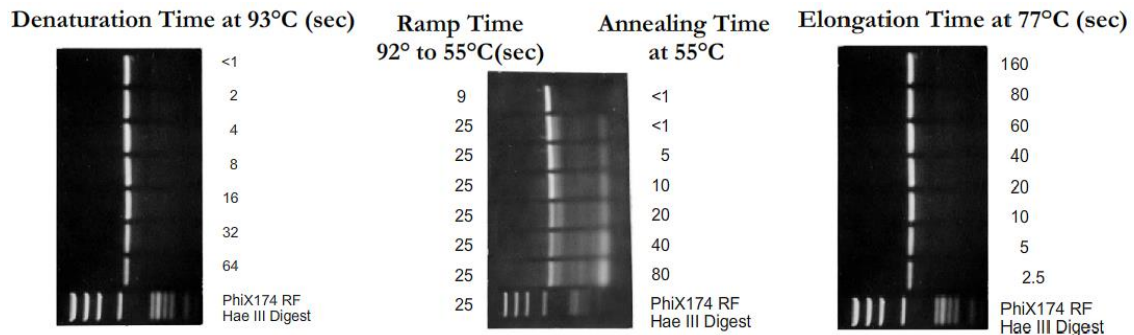
Fig.3.8 Gel results of different times spent at A) denaturation step, B) annealing step, C) Extension step. From Wittwer 1991.

The above results shown in Figure 3.8 A show that there is no effect in speeding up the denaturation time, even to the point of <1s. However, at 32s and especially at 64s slightly less product was produced than at 16s, probably because of lower specificity and more undesired products being formed. Figure 3.8 B shows that better results were obtained when using more rapid annealing times (1s-5s), as well as when faster ramp rates were employed. For example, when the sample was cooled from 92°C to 55°C in 9s better results were obtained. Wittwer explains that this effect results from the fact that less time was needed for PCR products formed in previous cycles to reanneal to the template strands. Thus, better product yields and lower concentrations of undesired side products were produced. Finally, in C we see how extension/elongation times affect results. Here we do see a slight decrease in product produced as the time at this temperature was decreased. However, the effect was small and good results were still observed using both 2.5s and 5s extension times, with no decrease in product produced if amplicons were <200bp.

The use of Rapid PCR does have some drawbacks [Wittwer 1991], these include the need for a higher concentration of polymerase, so that it is more readily available for the reaction to use. The improvement in yield of already short amplicons, is usually not a problem for forensic purposes, as most of the desired alleles are small. High melting temperatures of primers are also needed so that their activity can start at higher temperatures, and thus lower the time needed for annealing. Temperature homogeneity and transfer are also very important, and for this reason the kind of receptacle used for the reaction can be optimal. For example, normal plastic PCR tubes show poor thermal transfer. Wittwer suggested that it would be better to use glass capillaries as they are thin walled, have a high surface area to volume and low heat capacity [Wittwer 1991]. This and their small volume of sample will allow internal temperature of the sample to change rapidly. However, they are difficult to handle and so a plastic receptacle was formed to fit on the end of the capillary for easier loading and use (these were then called plastic cuvettes and can be seen in Figure 3.9).

Fig.3.9 Example plastic cuvettes designed by Wittwer. From Wittwer 1991.

Later studies involved the use of rapid PCR technique to speed up the genotyping of ABO blood groups from whole blood [O'Keefe 1996, Huh 2011]. These studies, while showing faster amplification times then their contemporary methods were ~1 hour long and so

61

further research was needed to increase speed. More recent methods use either liquid samples or FTA paper punches and faster enzymes. These permit much shorter cycling conditions resulting in successful amplifications in ~70 mins [Lee 2011]. Integrated systems involving PCR coupled to microfluidic devices have also been developed [Lagally, 2001; Mitnik, 2008; Shi, 2003] but these focus more on the speed of separation and detection rather than increasing the speed of the PCR. More recently, rapid DNA analysis techniques have been developed; however, these techniques continue to include an extraction step within these processes to complete the analysis, which can take up to an extra 90 minutes.

Rapid PCR analysis has been performed using commercial kits such as the AmpFlSTR Profiler Plus, COfiler, Identifiler and PowerPlex 16. The AmpFlSTR Profiler Plus kit was used with each of the following polymerases: SpeedSTAR HS, KOD hot start and PyroStart fast PCR master mix [Giese 2009]. Of these it was found that SpeedSTAR gave the best results in the fastest time (19.1 mins), however it was noted that some incomplete non-template addition (NTA) was observed above the usual level, and stutter was reported as being slightly higher than in normal PCR [Giese 2009].

A second study tested a different set of enzymes and compared them to the results produced using Amplitaq Gold. The enzymes tested were Phusion, AB77, AB95, AB-1 and AB-3. The results showed that Phusion, AB77 and AB95 all gave non-specific amplification (probably due to these enzymes lacking hot start functionality), AB-3 and AB-1 gave cleaner profiles than the other enzymes, but produced n+1 and n-2 stutter peaks and double the amount of cross reactivity as what [Wang 2009]. Another study looked at using the

AmpFlSTR Profiler Plus in combination with SpeedSTAR HS and Kappa2G fast hot start [Laurin 2012]. Laurin improved amplification by combining the use SpeedSTAR with leaving the PCR product at room temperature for ~15 minutes, along with optimization of other PCR reagents [Laurin 2012].

In 2012 Verheij used a newly engineered enzyme (Phusion flash) to test a rapid/direct process with a variety of different sample types of forensic interest including saliva, blood, bone, semen, hairs and touch samples. The resultant amplification produced a ~30-40 mins PCR [Verheij 2012]. This process worked well for saliva and blood with full profiles being obtained after amplification directly from whole blood. Results for touch DNA were not as reliable with Verheij noting a 46% decrease in success rate for these types of samples when compared with results from blood or saliva [Verheij 2012].

H. Direct PCR

Originally the direct PCR process involved the deposition of blood or saliva samples on to FTA paper. The FTA paper stabilizes the DNA and keeps it free of pathogens while halting the DNAs degradation.   A small punch from this material can be amplified directly off the paper without the necessity of cell lysis or DNA isolation. The chemical mechanism within the paper permits lysis and purification to take place but does not alter the buffering capacity of the PCR cocktail. [Chomczynski 2006].   This type of direct amplification eliminates the extraction process as well as the need for quantification as the size of the punch can be the same for all reactions. This simple procedure also reduces the possibility of contamination by minimizing sample processing steps.

Paper based stabilization of nucleic acids was first developed in the late 1980's at Flinders University in Australia. Burgoyne and Fowler were able to design a special paper that was capable of protecting nucleic acids from depredating agents, such as nucleases and bacterial activity, and preserving the samples for long-term storage [Fowler 1988]. They applied a weak base, a chelating reagent, an anionic surfactant and uric acid into a cellulose based paper. The chemicals embedded on the paper lysed the cells, releasing the DNA and denaturing the proteins. DNA samples on Watman FTA® (Flinders Technology Associates) paper are stabilized and can be stored at room temperature for years without loss of integrity and this method of long-term storage requires less space than traditional methods.

The main problem with direct amplification is that it requires relatively clean samples. Otherwise, inhibitors can produce allele dropout especially when the analysis speed is increased in combination with direct amplification [Aboud 2013, Verheij 2012]. Careful selection of the enzyme and optimization of buffer system, magnesium content, enzyme concentration, salts, BSA and other additives have been found to reduce problems with allele dropout. Recently developed enzymes have included: OmniTaq, a polymerase series capable of performing amplification of blood and soil DNA; Rapid Klentaq, a mutant form of Klentaq that was designed to provide robust amplification with very short extension times, permitting faster amplifications. Other such examples are Phusion Polymerase (Thermo Scientific) which has been designed for direct amplification of whole blood and resistant to the inhibitors such as haematin that are present in this type of sample [Fuehrer 2011, Verheij 2012], with a variation of this, Phusion Flash, being used in rapid amplification for different crime scene samples [Verheih 2012].

Polymerases developed for speed and sensitivity include Pyrostar (Thermo scientific) and SpeedStar (Takara) which were used by Vallone et.al. to speed up the PCR of commercially available STR analysis kits [Vallone 2009]. Other kits include KAPA2G (Kapa Biosystems), AptaTaq (Roche), and inhibitor resistant polymerases such as Omnitaq (DNA polymerase technology).

The direct amplification of blood and saliva samples has been shown using both FTA paper based methods [Mercier et al 1990] and non-FTA based methods. The non-FTA based methods include procedures involving FoLT PCR (formamide low temperature PCR), nested PCR, buffer Ph adjustment and microwave-based techniques [Panaccio 1993, Qu 2007, Chomczynski 2006, Bu 2008, Fuehrer 2011, McCusker 1992, Goodwin 1993]. Other sample types have been used including tissue [Panaccio 1993], bone, semen [Verheij 2012] and touch DNA [Templeton 2013]. However most of these studies have not focused on analyses in which sample sizes are limited, and generally focus on eliminating inhibition and not on high speed PCR.

Specially designed buffers (known as direct buffers) are becoming commercially available for improved amplification of FTA punches. One method involves a pre-treatment of DNA transferred from a saliva swab (although it also can be done with whole blood) to an FTA paper punch (2mm), by soaking it in 40 µL of Direct-N-Elute elution buffer (BioQuest, Seoul, Korea) at room temperature for 15 minutes. 1µL of this solution is included in a PCR mixture containing a zwitterionic buffer and non-reducing carbohydrates which together help to overcome any inhibiting factors present in the sample [Park 2008]. A Zwitterionic buffer is a buffer that consists of a molecular compound with both acidic and

basic functional groups. This means that at a neutral pH the zwitterionic buffer solution will have both anions and cations at the same time. Components of the Zwitterionic buffer can be HEPES (4- (2-Hydroyethyl) -1- piperazineethanesulfonic acid), Bicine (NN-bis (2-hydroxyethyl) glycine) or preferably Tricine (N- Tris (hydroxymethyl)methylglycine). The buffer works in a wide variety of pH conditions (6.8-9.5) but is preferable kept between 8-8.7 [Yang 2006]. The "non-reducing carbohydrate" that also form part of the Direct-N-Elute are hydrocarbons that have more than two hydroxyl groups while also not having reducing capabilities. These include polyols (polyhydroxy compounds) and their derivatives and non-reducing hydrates of carbon, called non-reducing sugars or saccharides, and their derivatives. Examples of these non-reducing carbohydrates include sorbitol, mannitol, trehalose, raffinose and sucrose to be included in a 2-30 % (w/v) of the reaction mix [Yang 2006]. These compounds along with BSA help counteract the effect of inhibitors in direct PCR reactions where biological matrices are used directly with no cleanup or purification step being performed prior to their incorporation into the PCR mixture [Park 2008].

Further work was performed to evaluate the effect of size and location of the saliva FTA paper punch to see if these had any effect on results [Wang 2009]. It was found that the optimal size of punch for both blood and buccal samples is 1.2mm, as larger sizes delivered more sample and increased the quantity of inhibitors to the point at which the buffer could no longer counteract inhibitory effects. Another buffer of interest is the AmpFLSTR® Identifiler® Direct buffer developed by Applied Biosystems. This buffer contains salts, detergents, carrier protein and 0.05% sodium azide to aid the direct PCR process. This direct buffer has been tested in the form of a master mix, and with the use of AmpliTaq

Gold polymerase to give successful amplification from single source blood and buccal swabs on FTA paper without purification steps [Wang 2011]. The AmpFLSTR® Identifiler® Direct buffer previously has also been tested with non-FTA samples [Brito 2011]. This procedure worked well with blood samples, but not as well with buccal swabs. The authors suggested that the amount of DNA collected on a swab varies much more than the amount present in a blood drop deposited on FTA paper.

The literature does not suggest that there is a difference between blood and saliva, when both are transferred to FTA paper for DNA extraction and storage samples [Brito 2011, Wang 2011]. It is however possible that the reagents in the paper itself can cause inhibition if wash steps are not employed or specific buffers used. In general, direct buffers are used to counteract all these types of inhibition as we have seen.

It may be possible to amplify directly from saliva as this is a relatively clean sample and would not require the use of specialized buffers, as is necessary with FTA paper punches. Different methods are being investigated including microwaving the saliva for different periods of time [Goodwin 1993], diluting the saliva and addition of different facilitators (such as BSA).

I. Rapid Direct PCR

Early methods to increase the speed of the PCR, led to the idea of simultaneously performing rapid and direct PCR [Lee 2011]. Lee reports the amplification of whole blood for ABO blood typing in 70 minutes with no prior extraction step. As can be seen from the previous discussion on rapid PCR it is only the instrumentation and not the biochemistry itself that acts as the limiting factor for the speed at which the PCR reaction can take place.

With new thermocycling technology such as the Streck Philisa rapid thermocycler, these reactions can be sped up considerably, thus increasing throughput and saving time and money.

Faster processing, using rapid direct PCR with high speed enzymes and microfluidics, can produce a genotype in under 25 minutes [Aboud 2013], as no extraction step is necessary. One of the newer areas of research is to test these new enzymes with off the shelf high speed thermal cyclers to provide crime laboratories with a method for screening and producing genotypes for crime scene samples of all types (such as saliva, blood, tissue, semen, bone and touch/fingerprint samples). This has already been achieved with saliva samples and FTA paper [Aboud 2013], but additional optimization is needed to further reduce analysis time. The key to this process is direct PCR, which utilizes special buffers and sets of enzymes to amplify DNA directly. Currently direct PCR utilizes FTA paper as the substrate for amplification. The buccal swab is wiped on the FTA paper to transfer the saliva sample. This can be processed in the same way as blood, with no extraction or wash steps necessary. This method can be altered to utilize saliva directly (without prior transferal to FTA paper) as this would remove another step in the process as well as not having the problem of inhibitors from the FTA paper present in the PCR reaction. Touch or fingerprint DNA will also be an important part of the project as many crime scenes contain evidence that has been handled by a suspect. While a fair amount of research has focused on methods for better collection and analysis of touch samples [Gibson-Daw 2013, Nunn 2013, Sangeeta 2011, Daly 2012] there has been little work performed on the use of such sample for direct [Templeton 2013] or rapid PCR processes.

For rapid direct PCR to work with samples other than saliva, new enzymes must be identified and optimized that increase the speed of the process despite the presence of inhibitors. Companies such as Takara, Promega and DNA Polymerase Technology have developed such enzymes with improved processing speeds and better capability to handle PCR inhibition. The concentration of dNTPs is also increased in rapid direct PCR to avoid problems caused by diffusion. High speed thermal cyclers will also be required. The combination of fast enzymes and quick temperature cycling should permit amplifications to be completed in under 15 minutes. Optimization of buffer conditions (BSA), and reagents such as dNTPs will still be necessary to achieve high fidelity amplifications at these speeds.

When coupled to a standard STR typing instrument, this optimized process can take an analyst from 25 min to an hour [Aboud 2010] from sample preparation to getting a complete result. Many researchers are looking into ways to speed up the genotyping step, such as using microfluidic chip based electrophoresis. With the purchase of a fast-thermal cycler and appropriate enzymes, any laboratory could easily perform a fast DNA amplification [Jensen 1990, Kupfer 2006]. Specialized laboratories at remote sites and at mass disasters could perform the task in under 25 minutes using microfluidic separations [Aboud 2013]. Figure 3.10 demonstrates the result of a 16-minute amplification and a 2-minute electrophoretic separation. These results were obtained using a Streck Philisa thermal cycler and a modified Agilent Bioanalyzer. The genotype could also be obtained using a standard ABI 310 or 3100, however it would take an addition 30 mins.
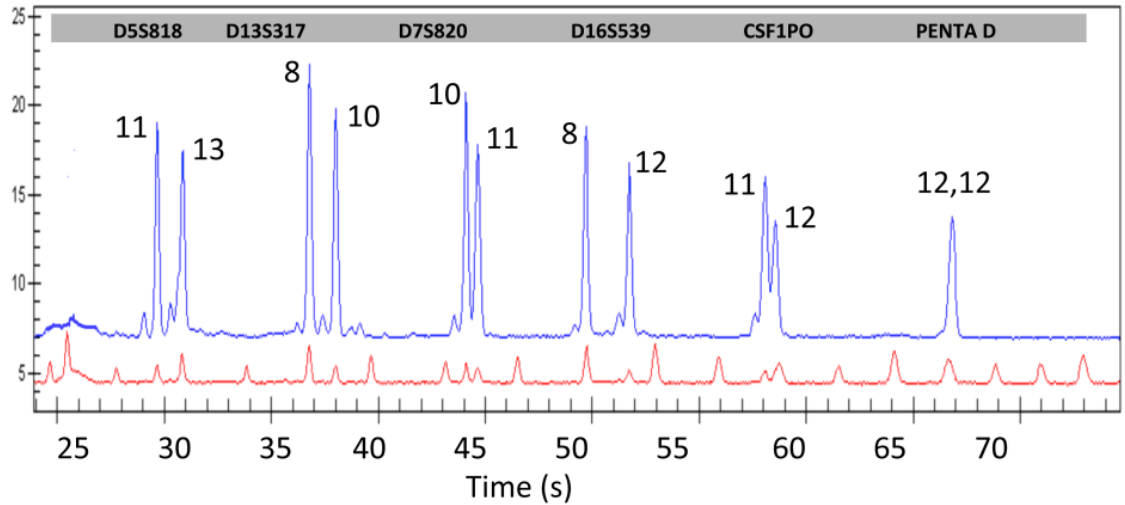
Fig.3.10 The production of a genotype in 70 seconds using a modified Agilent 2100 Bioanalyzer.

Because this process only uses a small amount of the total sample, the rest can be stored for further testing. Contamination is also minimized as there is minimal sample transfer between steps as extraction and amplification take place in the same tube.

In this way, large numbers of samples can be analysed quickly and efficiently. Furthermore, this procedure is capable of being used as a portable technique that can be taken on site to crime scenes or mass disaster sites. Rapid PCR methods are of a particular interest as they can be used for on-site screening of samples being collected [Paradela 2006]. While rapid direct PCR shows a much-increased speed of analysis, it may be possible to improve on it further. With the development of new rapid polymerases and instruments further decreases in analysis time are possible by reducing amplification time [Gibson-Daw 2018] and removing the FTA paper step through the use pf direct PCR.

CHAPTER IV. SHORT TANDEM REPEATS (STRs)

A. STRs: Basics

Only 0.3% of the human genome is made of regions that differ from individual to individual. Typically, these regions are located between genes in what are known as introns. These regions contain repetitive sequences of nucleotides [Butler 2009] which vary in size from person to person, usually with no detriment or benefit to the different individuals. There are three types of variable regions located within the human genome depending on respective length. The long regions are known as satellites, the medium length ones are called minisatellites (or VNTRs) and have repeats of 10-60 base pairs, and the shortest ones (those which are normally used in forensic analysis) are known as microsatellites, simple sequence repeats (SSR) or more commonly short tandem repeats (STR). Their repeat units are 2-6 base pairs.

Sort tandem repeats are the most reliable DNA markers for forensic use not only due to their highly polymorphic nature. Furthermore, as a result of their small size, they can be easily amplified by PCR even when samples are significantly degraded. This means that very small amounts of DNA from a crime scene (100 pg) can be amplified, giving the scientist much more material to test. Forensic STRs are usually made up of 4 repetitive sequences of nucleotides but di-, tri-, tetra-, penta- and hexa-nucleotides, tri-nucleotides, tetra-nucleotides or hexanucleotides are also possible (Figure 4.1). Sequences of 3-5 nucleotides repeats are commonly used in commercial kits.
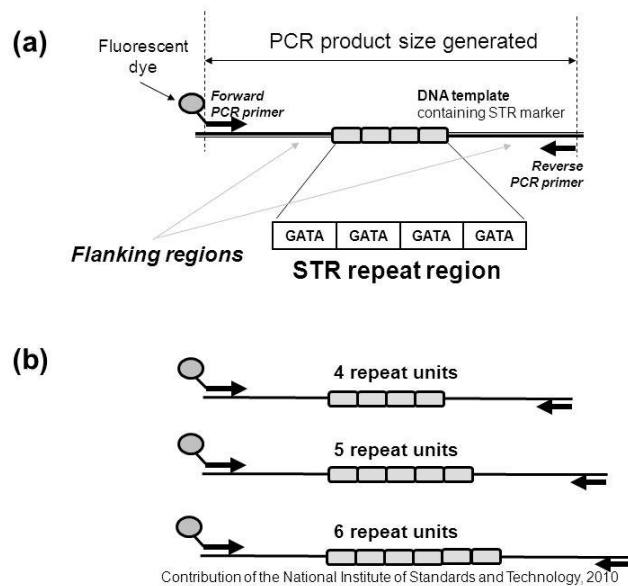
Fig.4.1 Schematic of Short tandem repeats. A) shows primer binding sites, flanking regions and repeat motif. B) shows differing number of repeat units that differentiate alleles. Adapted from Butler 2009.

The longer lengths of the repeat motif produced decreased stutter (see section 4.D.1), while also offering improved resolution by electrophoresis than the smaller repeat motifs (Figure 4.1). There are different types of STRs: Simple, Compound and Complex. Simple STRs, are simple repeats of a 2-7 bp motif. Compound STRs are made of two or more simple repeats and are a complex mixture of these simple repeats. Complex STRs are made of several repeat blocks of variable unit length with the presence of intervening sequences between them [Urquhart 1994, Butler 2009]. Many STRs also have microvariants which involve incomplete alleles, containing variations in the standard 4 base repeat motifs. An example is the human TH01 locus which contains a 3 base variant along with 9 repeat unit giving rise to what is known as the 9.3 allele [Butler 2011]. Typically, in a forensic investigation a set of 15-24 STR are examined. Statistical analysis of these STRs is based upon fundamental concepts in population genetics including the Hardy-Weinberg

equilibrium (HWE) and the independent assortment of alleles (product rule). The Hardy-Weinberg equilibrium is a principle that states that allele frequencies observed in a population remain stable across many generations. This principle holds true so long as random mating occurs and there is no evolutionary influences favouring or disfavouring some over the others. Using the HWE it is possible to calculate genotype frequencies. For a locus with two alleles per individual (A and a), the genotype frequencies for AA, Aa and aa will be given by p^2, 2pq and q^2 respectively [Butler 2009].

The product rule expresses the probability of two (or more) separate events occurring at once and it is calculated by multiplying together the probabilities of each individual event. This can only be used if the events are completely independent from one another. The product rule is employed in forensics to calculate the frequency of a profile expected to appear in the population. It is calculated by multiplying the frequencies of the observed alleles at each locus together including both alleles of a heterozygote. The product rule is also sometimes known as the random match probability [Butler 2009].

B. Allelic Ladders

Allelic ladders are an assembled mix of DNA fragments that are used as reference standards. They consist of co-amplified different alleles from many individuals to produce fragments representing all alleles found and sequenced [Smit 1995] in the population for the loci under analysis. The ladder is created by taking all alleles possibly present at a locus and amplifying them. These are then separated in polyacrylamide gels and individual alleles are cut out and mixed together at similar concentrations. This process is performed for all known alleles for the different loci in a multiplex. The fragments are combined to

create the ladder [Kopka 2011]. Allelic ladders can be remade or resized by diluting the original ladder and re-amplifying it with the original primers [Butler 2003]. This method can be utilized in situations in which different primers are used, by diluting the original ladder and then amplifying it with the new primer sets selected for the loci. Modern forensic kits supply their own ladders (Figure 4.2). The allelic ladder is run with every batch of samples and the unknown peaks from these samples are compared with the known peaks of the ladder for use in allele assessment [Schumm 1997]. If an allele has the same fluorescent label and is within 0.5 bp of the known allele from the allelic ladder, then the software will identify it as that allele [Butler 2006]. With modern capillary electrophoresis instruments, this is achieved automatically using an internal size standard. An internal size standard or internal lane standard (ILS) is a standard control that is added to every sample run by electrophoresis. It consists of a number of fragments of DNA of known sizes ranging from small >100 bp to large >1000 (the range varies by the ILS used). These fragments run along with the sample fragments and are used to size them by the data analysis software. The sizing occurs by comparing the ILS fragments retention time to that of the sample fragments and using a sizing algorithm that assigns the corresponding size to the sample. Allelic ladders are important for the automatic calling of alleles by the instrument software, while also helping correct for minor changes in size of alleles that may arise from temperature and mobility shifts between samples, instruments and labs.
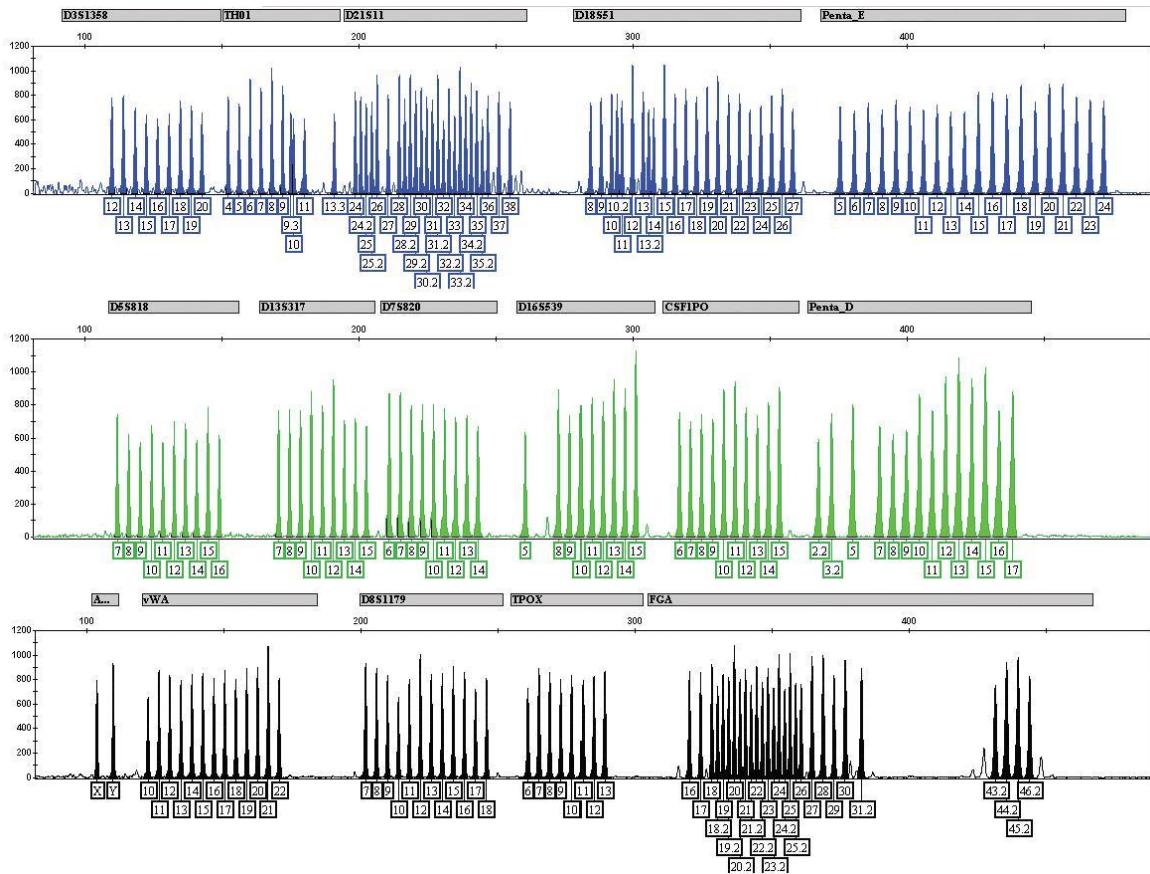
Fig.4.2 An electropherogram of the PowerPlex® 16 HS Allelic Ladder, showing the different alleles for each locus in the PowerPlex® 16 HS kit. Adapted from Promega website.

C. Commercial STR kits

Since STRs are the current method of choice in forensic DNA typing, several commercial STR kits have been created [Krenke 2002, Greenspoon 2004, Wang 2011, Laurin 2012]. The modern kits permit the user to multiplex PCR amplification of STR markers and tag them with up to 6 fluorescent dyes to allow analysis in instruments such as the Thermo Fisher Scientific 310 Genetic Analyser or the multi capillary based 3130 or 3500 CEs. These kits not only supply the primer multiplex, but also come with all the PCR regents needed for successful amplification [Promega technical manual 2016]. Matrix standards (used to calibrate the instrument), internal size standard (added to all samples in order to

later size them) and allelic ladders (used for allele calling) are generally either included in the kits or are available for purchase. Commercial kits originally used relatively few loci, were then expanded to include the 13 original core CODIS loci. Some of the new kits contain a mixture of normal STRs and Y-STRs, such as the PowerPlex® Fusion 6C (containing 23 autosomal STRs, 3 Y-STRs and Amelogenin), which can help in the separation of male from female contributors in mixtures. Before a commercial kit can be sold and used by laboratories it must go through both developmental validation (by the company producing the product) and internal validation (by the crime lab implementing the product), these validations must comply with the up to date SWGDAM guidelines [SWGDAM 2012]. The use of commercial kits allows for standardized procedures across laboratories and makes proficiency testing easier as well as saving the laboratories the time and money it would take to develop their own kits.

D. STRs: Chemical and Biological Artefacts

Current methods for STR typing sometimes include artefacts, than can appear in the electropherograms produced. Some of these are artefacts are the result of PCR process (stutter, microvariants, null alleles, drop-in and drop-out and non-template addition) while others are inherent to the capillary electrophoresis method (pullup spikes). If improperly identified these artefacts can cause misinterpretation of the data. It is important to understand them and have protocols in place for mitigation of these effects. Common artefacts are mentioned in more detail in the following sections.

D.1. Stutter

One artefact that is often seen in STR profiles is a stutter peak. Stutter peaks are extra peaks in the profile that can be either one to two units smaller or larger than the actual allele amplicon. Stutter peaks arise from strand slippage of the polymerase during the primer annealing step of PCR, where the primers mispair to their target regions on the template DNA. Depending on which strand is affected this can cause either deletion or insertions into the product strand. If slippage occurs to the bottom strand, i.e., the 3'-5' one, then a deletion (n-4) occurs causing a small peak of around 5-15% of the height of the true allele [Hauge 1993]. If slippage occurs to the top strand, i.e. the 5'-3' one, then an insertion (n+4) is caused which appears as a small peak 4 bp after that of the allele peak. The n+4 peaks are rarely visualized as they are usually ~2% of the allele height and so get lost in the noise [Leclair 2004] (Figure 4.3). The percentage of stutter is calculated using Equation 4.1:

Equation 4.1: Stutter %

$$\frac{(stutter\_product\_peak\_height)}{(main\_product\_peak\_height)} * 100$$
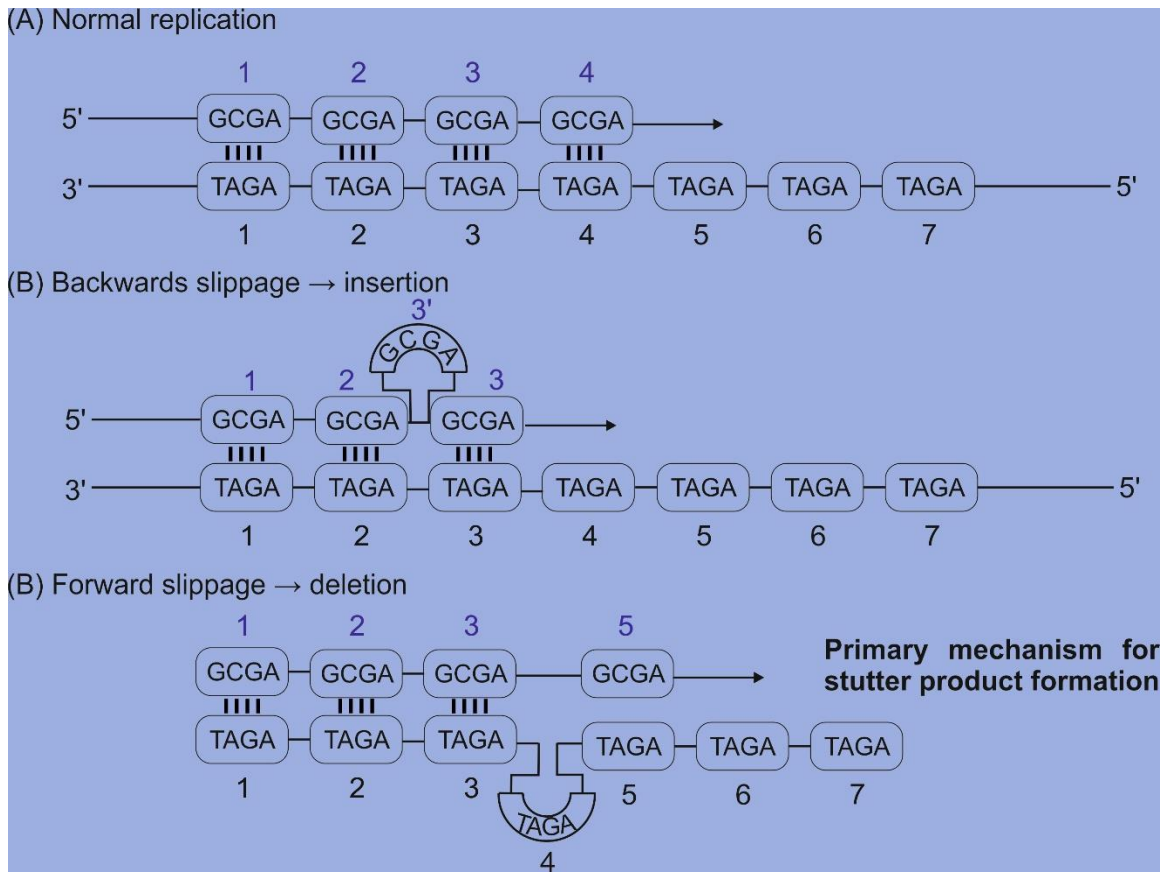
Fig.4.3 Schematic of the two mechanisms of stutter product formation during PCR. A) normal PCR process with no stutter formation. B) Stutter product formed by insertion by backward slippage, where resulting product is one repeat longer. C) Stutter product formation by deletion caused by forward slippage, causes -4 bp stutter which is the most commonly observed stutter product where the final product is one repeat short. Image recreated by author, original by Butler 2009.

Appearance of stutter decreases both as the length of the repeat motif increases (pentameric repeats show less stutter) and as number of repeats decreases (alleles with less repeats show less stutter) [Walsh 1996].

Stutter is a common occurrence and can hamper analysis of complex mixtures. Most laboratories have a special stutter threshold set which differentiate between true alleles and stutter peaks that may be present in mixtures. It is important that stutter thresholds are set properly as true stutter can be indistinguishable from low level peaks [Moretti 2001].

D.2. Microvariants

Many STR loci repeats contain alleles that include incomplete repeat units, which are known as a microvariant repeat units. These can be formed by deletions, insertions, etc. One example is the FGA locus which contains a 2 base microvariant in the repeat motif which is the result of a –TT– dinucleotide partial repeat after the fifth full repeat. Alleles containing this 2 base variant are labelled using the repeat number along with a number indicating the length of the variant allele in bp. The D21S11 locus also has a 2 bp microvariant caused by a–TA– dinucleotide partial repeat which is present before the last full TCTA repeat. This creates alleles such as 28.2 and 33.2 [McBeth 2006]. There are many loci that have differing numbers of these microvariencies. However, most common microvariant alleles have been identified and documented using large scale population studies [Ruitberg 2001]. Minor alleles that are not used in allelic ladders are collected in a database maintained by NIST (STRbase).

D.3. Null Alleles

Null alleles occur when there are point mutations or indels that have occurred at the primer binding sites [Dakin 2004]. These mutations can cause the primers to improperly anneal to their binding sites and prevent replication of one or both alleles of the loci in question. The closer the variation is to the 3' end of the primer, the more likely drop out will occur [Budowle 2001]. To test if this has occurred, primers with two different annealing sites can be used, in which one of these should bind correctly and produce a PCR product. Sequencing of the DNA can also indicate the presence of point mutations in the primer binding sites [Kline 2011]. Many of these mutations have now been documented, and so commercial multiplex kits have designed their primers to avoid these regions. In general,

the primer sequences should be as close as possible to the STRs as this decreases the size of the amplicon and therefore minimizes the potential for insertions and deletions in the region between the primer binding sites and the STR region. The presence of null alleles or unusually sized alleles normally does not affect interpretation of the resulting data as long as a single STR kit is used. However, when different kits containing different primer binding sites are used, lack of concordance in the resulting genotypes can occur. Database searches are designed to correct for these issues.

D.4. Non-template addition

Non-template addition is the result of the Taq polymerase adding an extra nucleotide to the 3' end of a PCR product. This is often an adenine, so the process is known as adenylation. This phenomenon can result in the production of split peaks in which the peak that is missing an adenine is labelled -A, and the peak with the added A is known as the +A allele (Figure 4.4) [Butler 2009]. These split peaks (A-/A+) are undesirable, as they can appear to be two different alleles. The effect can also cause confusion during the profile analysis, as the allele peak height is reduced by an amount equal to the area of the -A peak.
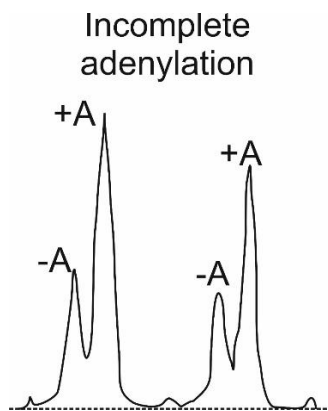
Fig.4.4 An example of an electropherogram showing split peaks (-A/+A) which result from incomplete adenylation. Image recreated by author, original by Butler 2009.

80

The cause of this non-template addition is usually excessive concentrations of input DNA template and inadequate time for the polymerase to finish replicating the alleles. For this reason, non-template addition is seen often in rapid PCR as limited time is available for the polymerase finish replication. Even under conditions of excess DNA template, it is possible to diminish the probability of adenylation. For instance, primers can be designed to help promote complete replication. This is done by adding a guanine to the 5' end of said primer [Butler 2009], addition of ATT to the 5' end of the unlabelled reverse primer can also help complete adenylation [Krenke 2002]. Another method to reduce adenylation is to include a final soak step of 15-30 minutes at 72°C to give the polymerase time to correctly complete the adenylation of all PCR products. This is not usually an option for rapid PCR as it would add unnecessary time to the procedure.

D.5. Stochastic Effects

Stochastic amplification can sometimes occur in samples during the PCR process resulting in uneven allele peak heights between heterozygous alleles, allele dropout (one allele missing), or locus dropout (both alleles missing). These effects occur usually during low template amplification and are caused by there not being enough of one or more template strands available to be copied [Butler 2005]. Stochastic issues generally occur in samples containing less than 100 pg of DNA, which is known as low template DNA or low copy number (LCN) DNA [Butler 2005]. When dealing with low template samples, different methods for analysis must be used [Caragine 2009]. These methods must be validated for these sample types, in order that correct interpretation of the data is maintained. It is also important that individuals involved in the analysis of LCN sample have appropriate methodology training and proficiency testing [Caragine 2009, Caragine 2011, Budowle

2009]. However, there are things the analyst can do to circumvent the negative effects of these samples, for example adding sufficient sample to reach the minimum necessary input concentration of DNA by first quantifying the samples then adjusting input volume as necessary. Running samples in replicates of 3 or more and then pooling the results can also reduce allele or locus dropout problems as it is unlikely that the same allele or locus would dropout in the same way in each replicate [Budowle 2009, Gill 2001, Budowle 2011].

E. Y-STRs

Y-STRs are short tandem repeats present on the Y chromosome of males. They are important as they can be used for situations such as sexual assaults, paternity cases or mass disasters in which the DNA from close relatives is utilised for comparison purposes to determine the identification of unknown victims [Walker 2005, Park 2007]. Y-STRs can also be of use in sexual assault cases to separate out the male from the female DNA in circumstances in which the female victim's contribution overwhelms that of the male perpetrator, or in cases with multiple male contributors. In forensic situations evidence can be collected from a wide variety of sources such as fingernail scrapings, car or home door handles, etc. In situations in which close relatives are involved, a special class of Y-STRs known as rapidly mutating Y-STRs can be used which show high mutation rates [Albani 2013]. Rapidly mutating Y-STRs e have an improved discrimination power between males of related paternal lineages.

A selection of 11 standard Y-STR core loci were originally chosen in 1997 by the Scientific Working Group for DNA Analysis Methods (SWGDAM) using 9 primer sets (Figure 4.5) [Sinha 2004].

Fig.4.5 The chromosome location and nomenclature of the 11 core Y-STR loci as suggested by SWGDAM, including repeat motif, allele range and mutation rate. Adapted from Butler 2006.

Since then additional loci have been added to various commercial kits to increase the power of discrimination. Examples of these expanded kits are the PowerPlex® Y23 System (Promega) which includes 23 loci across 4 dye channels (DYS576, DYS389I/II, DYS448, DYS19, DYS391, DYS481, DYS549, DYS533, DYS438 (penta), DYS437, DYS570, DYS635, DYS390, DYS439, DYS392, DYS643 (penta), DYS393, DYS458, DYS385a/b, DYS456 and Y-GATA-H4) [Promega website] and the Yfiler® Plus Kit (Thermo Fisher Scientific) which includes 27 loci across 6 dye channels (DYS19, DYS385 a/b, DYS387S1 a/b, DYS389 I/II, DYS390, DYS391, DYS392, DYS393, DYS437, DYS438, DYS439, DYS448, DYS449, DYS456, DYS458, DYS460, DYS481, DYS518, DYS533, DYS570, DYS576, DYS627, DYS635 (Y GATA C4), and Y GATA H4.) [Thermo Fisher website].

CHAPTER V. FAST THERMOCYCLERS AND RAPID POLYMERASES

A. Fast Thermocyclers

In recent years much research and development has taken place to create new styles of thermocyclers capable of reducing amplification time. This has been done by increasing the heating and cooling ramp rates through the use of special coating of metal blocks, changes to tube design, and instrument configuration. Peltier based thermal cyclers increase the ramp speed by using specially coated heating blocks. In addition, faster cooling fans and specialized ultra-thin tubes (Figure 5.1) that permit faster, more efficient heat transfer between the sample and heating block are also used. An example of a rapid Peltier based thermocycler is the Philisa (Streck) seen in Figure 5.1.

Systems where the PCR mix is transferred between areas maintained at different temperatures are also being developed as this would decrease amplification time by removing the need to heat and cool a metal block in order to change the temperature of the PCR mix. However, these systems have yet to be commercialized.

Fig.5.1 Philisa ultra-fast thermal cycler. Instrument shown top left. The silver heat block with PCR tube shown bottom left. Specially designed thin, flat PCR tube shown on right. The Philisa cycler is has heating rates of 15°C/s and 12°C/s cooling.

B. DNA Polymerases

DNA polymerases are a sub type of enzymes that are vital components in DNA replication, repair and recombination. Polymerases bind to the template strand of DNA by interacting via electrostatic interactions with the sugar phosphate backbone and the minor groove of the DNA. In many cases, it is this binding that determines processivity of a given enzyme. Processivity is "an enzymes ability to catalyse consecutive reactions without releasing its substrate" [Stryer 2002]. This amounts to the average number of nucleotide bases added to the new DNA strand per association/disassociation event of the polymerase with the

85

template DNA strand (Amplitaq Gold has a processivity of 50-60 nt/s according to Life Technologies).

Many new enzymes have become available commercially that have been modified or engineered to improve this rate of processivity. The precise modifications performed on commercial enzymes are unavailable due to patent issues, however we can hypothesize that these are based on research performed to improve and develop polymerases of higher processivity and resistant to inhibitors. One such study by Wang, details how both Taq polymerase and Pfu can be fused to a sequence non-specific double-stranded dsDNA binding protein (in this case Sso9d). This design results in an increase of processivity while showing no negative effects on either stability or catalytic activity of the polymerases [Wang 2004]. The original Taq polymerase was made up of two main sections, a polymerase domain and a 5'-3' exonuclease domain. Studies have shown that this exonuclease domain takes part in interactions with the template DNA of a PCR reaction, since mutants without this domain show less processivity [Wang 2004, Murali 1998]. The fusion of both the original Taq and the mutant Taq to Sso9d showed increased processivity. The author suggests that the increase in processivity is due to the Trp24 residue of the SSo7d protein and the fact that this Trp24 residue plays multiple roles in the binding to ds DNA [Wang 2004, Gao 1998].

The above polymerase design is an example of what is called domain swapping or domain tagging [Motz 2002, Davidson 2003, Bedford 1997], in which domains from different enzymes are combined to express desirable traits from each component. Some of the best results were produced by the flexible attachment of nonspecific DNA binding domains of

other enzymes to the catalytic domains of the polymerase [Pavlov 2002, Belova 2001, Wang 2004]. This creates hybrid enzymes that have a high processivity even in the presence of high concentrations of salts and other PCR inhibitors [Pavlov 2004]. Another way to improve processivity of Taq polymerase is by site-directed mutagenesis, which is the intentional change in a DNA sequence. In one study, a change made by replacing Phe667 with Tyr, led to an increase in processivity [Pavlov 2004].

In addition to increasing processivity, researchers have also been interested in creating inhibitor resistant polymerases. One study showed that an N-terminal 278 amino acid deletion of Taq (Kletaq1) along with mutations at codons 706-708 of both Kletaq 1 and wild type Taq increased the resistance of these polymerases to various inhibitors commonly found in crime scene samples such as haemoglobin, lactoferrin IgG soil extracts and humic acid. This polymerase was among several inhibitor resistant polymerases tested showing a resistance to an overloading of intercalating dyes [Kermekchiev 2009].

A proven method for cultivating polymerases for enhanced inhibitor resistance is compartmentalized self-replication (CSR). CSR works like a feedback loop where the polymerase replicates only its own encoding genes (self-replication). This process occurs in spatially separate, noncommunicating compartments formed by a heat-stable water-in-oil emulsion. This compartmentalization permits the genotype to be traced or linked to its corresponding phenotype, as each polymerase is only replicating its own encoding genes and not those of other polymerases or components that are contained in the other compartments. This means that any mutations that occur in the polymerases in order to adapt to their environment will be replicated for the next generation and in this way the

polymerases will become less affected by inhibitors with each generational mutation that occurs [Ghadessy 2007]. Using this method in a multicomponent inhibitor cocktail a number of different chimeric polymerases were tested for inhibitor resistance by Baar [Baar 2011]. In his study Baar found that polymerases labelled P4F12 (Hu1) and P6F3 showed the best resistance to peat extract (an 8-fold increase), both polymerases shared a similar design where their N-Terminal and C-terminal regions came from Tth polymerase while the protein core was originally from Taq polymerase. They both also contain in their main polymerase region a section derived from Tfl and a number of point mutations. When looking at polymerases resistant to bone powder Baar found that the chimeric polymerase 2D9 performed the best, this polymerase was unique in that in contained an array of sections from a number of different polymerases including: T. thermophiles, T. oshimai, T. brockianus and Taq which granted it resistance to the inhibitors present in bone samples.

The polymerases used in the research presented in this manuscript were Z-Taq polymerase (Takara Bio, Clontech Laboratories, Inc., Mountain View, CA) and Omnitaq (DNA Polymerase Technology, Inc., St. Louis, MO). Z-Taq, supplied with 10 X Z-Taq buffer and 2.5mM dNTP mixture, had previously been shown in studies by Aboud [Aboud 2013] to provide optimal peak balance at high amplification rates when compared to other DNA polymerases [Gibson-Daw 2017] Z-Taq is a mutated Thermococcus kodakarensis polymerase with alterations in its EXO I and 3'-5' exonuclease regions, leading to a more rapid amplification without proofreading capabilities or the need for activation steps [Imanaka 2001]. This enzyme was also engineered to have a high processivity (over 5X speed of regular Taq which has a processivity of 60 nt/s) while retaining high fidelity [Imanaka 2001, Wang 2004]. Omnitaq, supplied with 10 X Taq buffer and PEC-1 buffer

is a double mutant (meaning it has two mutations in genes of interest) of Taq polymerase that makes the enzyme resistant to the inhibitory effects of blood, soil and more, while having a high processivity of 120 nt/s. The use of this polymerase may permit direct amplification from biological sample with no prior extraction step [DNA Polymerase Technology Website].

CHAPTER VI. SAMPLE ANALYSIS BY ELECTROPHORESIS

A. Gel Electrophoresis

Gel electrophoresis consists of a slab of agarose or acrylamide polymer gel containing a number of small indentations known as wells, in which the samples will be placed. The gel is placed in an electrophoresis tank, filled with buffer solution, and is attached to an anode at one end and a cathode at the other. Voltage is next applied to create an electric field. Once the electric field is generated, the negatively charged DNA fragments (STRs) move through the entangled gel matrix away from their beginning position near the cathode and toward the anode, attracted by its positive charge. The smaller DNA fragments will migrate faster than their larger counterparts and the rate of fragment migration depends on the composition of the gel and on the strength of the electric field. Each fragment will show up as a dark band when observed under UV light or as a coloured band if a dye has been added to the product before running the gel. The resulting bands and their position can be used to determine the size (in bp) of each fragment and create a barcode like profile that can be compared to others of known sources [Butler 2009]. An example can be seen in Figure 1.2.

B. Capillary Electrophoresis (CE)

Early analysis of STRs was performed using slab gel electrophoresis but this was later replaced by capillary gel electrophoresis (CE), in which the DNA fragment separation occurs in small diameter (50-100 μm) silica glass tubes [Jorgenso 1981, Mikkers 1979] (36-50 cm long) filled with a viscous polymer matrix (generally polyacrylamide). This polymer fills the capillary and coats its walls reducing electroosmotic flow (EOF) while

also separating the DNA fragments by size as they migrate through the capillary drawn, drawn though the capillary by the application of an electric field and detected through a window etched in the capillary [Butler 2009]. Early systems used UV detection methods but this was quickly replaced by laser induced fluorescence [Swerdlow1990, McCord 1992]. In 1995 the ABI Prism 310 Genetic Analyzer was introduced to the commercial market and was quickly installed in many crime laboratories. This system was with simultaneous multi-wavelength laser induced fluorescence detection which permitted its use with STR multiplexes where multiple loci could be detected at once if labelled with different fluorescent dyes resulting in faster run times, increased throughput and higher power of discrimination. The fluorescence emitted by the dyes attached to the DNA fragments is plotted as relative fluorescence intensity (as RFUs) as they pass the detector.
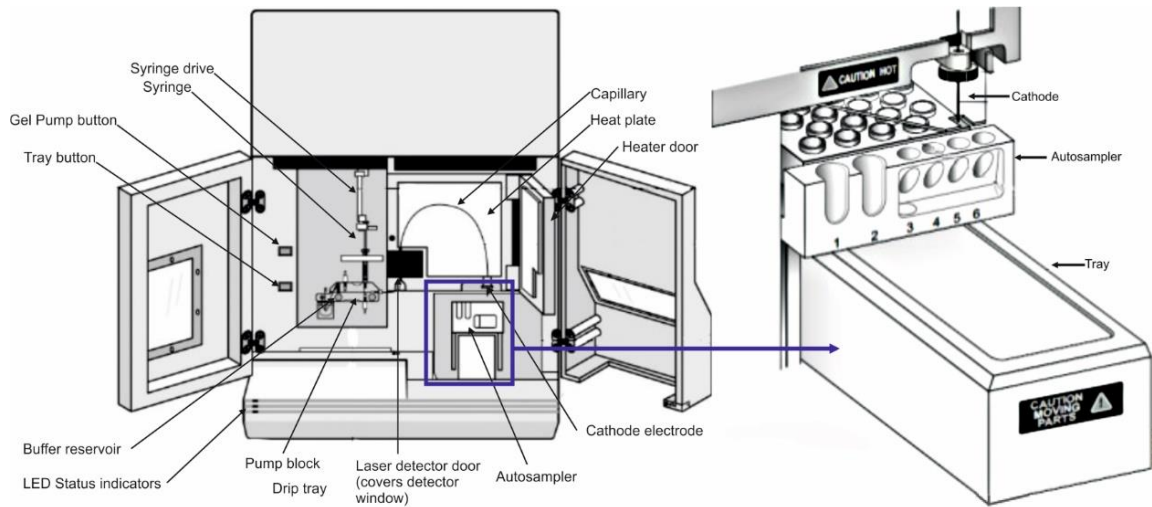
Fig.6.1 On the left schematic of the ABI Prism®310 Genetic Analyzer. On the right schematic of the ABI Prism®310 Genetic Analyzer Autosampler, a mobile sample tray that presents sample, buffer, wash water and waste to the capillary as needed. Injection occurs by the electrode attracting the DNA molecules to it where they then get taken up into the capillary and move towards the anode separating by sizes as they navigate the polymer matrix before being detected by laser fluorescence detection. This data gets turned into peaks on the electropherogram which gives the resulting profiles commonly used in forensic DNA analysis. Image recreated by author, original by Watts 2001.

The small internal diameter of the capillaries requires very small amounts of sample (~1μl), and permits faster run times. Another advantage to using CE is the ease with which the process can be automated at both the injection and detection stages.

CE results were first deemed admissible in court in 1996 [Kuffner 1996] and have been used in the analysis of gunshot and explosive residues and drugs as well as for DNA typing [Northrop 1994] and the sequencing of the human genome [Carson 1993].

B.1. Sample Preparation

Once amplified the DNA fragments present in the PCR product need to be separated and visualized using CE. To do this the PCR product is added to a mixture of deionized formamide (Hi-Di formamide), which reduces the ionic strength of the sample relative to the buffer while also denaturing the dsDNA to ssDNA. An internal lane standard (ILS) is also added to permit fragment size detection. In the past water was used instead of formamide and the sample rapidly heated and then snap cooled on ice to reduce the amount of dsDNA in the sample [Williams 1994] but with the change from water to deionized (Hi-Di) formamide this rapid and cooling step was no longer necessary. It is important that the conductivity of the formamide used be routinely checked as formamide degradation produces negatively charged by-products such as formic acid. This can compete with the DNA and be preferentially injected into the capillary of the CE instrument, negatively impacting the resulting resolution of the sample peaks [Buel 1998]. A formamide conductivity of <100 μS/cm has shown to give optimal performance while reducing amount of freeze and thaw cycles to minimize the possibility of degradation [Butler 2004, Butler 2011].

Once the sample is prepared, data collection software is utilized to automate analysis and recording of results.

B.2. Injection

The most common method of introducing the sample into the capillary is through electrokinetic injection. This works by placing the capillary in close proximity to the electrode and immersing them both into the sample tube while a voltage is applied for a set period of time. Under these conditions the negatively charged DNA fragments in the sample flow towards the positive charge of the electrode where they move into the capillary. The amount of DNA injected into the capillary in this manner can be calculated by Equation 6.1 as it is dependent on the electric field that is applied (E), the injection time (t), electrophoretic mobility (μ_{EP}), electroosmotic flow (μ_{EOF}) the concentration of DNA in the sample, the area of the capillary opening (πr^2), the ionic strength of the sample (λ_{sample}) and the ionic strength of the buffer (λ_{buffer}) [Sinha 2002, Liu 2009, Mitnik 2002, Kan 2004].

$$\left[DNA_{injected}\right] = \frac{Et(\pi r^2)(\mu_{EP} + \mu_{EOF})[DNA_{sample}](\lambda_{buffer})}{\lambda_{sample}}$$

Equation 6.1 Estimation of amount of DNA injected into a capillary by electrokinetic injection.

As indicated by equation 6.1, the amount of sample injected can be controlled by modifying the voltage or the injection time. Poor injection will result from the presence of competing ions such as chloride ions (CL^-), as the amount of DNA injected is inversely proportional to the ionic strength of the sample.

PCR products diluted in low conductive formamide have reduced ionic strength when compared to the CE buffer used. As the sample is introduced into the highly conductive environment within the capillary, a phenomenon known as field amplified injection or sample stacking occurs. In sample stacking a narrow band of the analyte is concentrated at the interface between the low conductive sample (high electric field) and the high conductive buffer (low electric field). This phenomenon occurs due to mobility differences between the two zones and leads to improved sensitivity of injection [Sinha 2002].

When a voltage is applied the speed at which the charged DNA fragments and other ionic molecules migrate towards the electrode is directly proportional to the combination of the electrophoretic mobility (μ_{EP}) and electroosmotic mobility field (E). Electrophoretic mobility (described in equation 2) increases with the decrease in frictional force (f) in the buffer, which is related to the viscosity of the medium and the size of the analyte. Equation 2 also indicates that mobility of the ions is affected by charge to size ration where the net charge (q) is inversely proportional to the mobility and directly related to the radius of the ion (r).

$$\mu_{EP} = \frac{q}{f} = \frac{q}{6\pi\eta r}$$

Equation 6.2 Electrophoretic mobility. Shows factors affecting the rate at which the DNA fragments migrate towards electrode.

B.3. Separation

DNA fragments are sieved through the capillary once a voltage is applied, with the smaller fragments moving faster than the larger fragment through the capillary due to the larger fragments being delayed more by the sieving matrix. It is important to use a method that

yields the best separation and resolution in the shortest time. This is achieved by optimization of the capillary, the sieving matrix and the buffer used for CE.

B.3.1 Capillary

Current methods for separation of DNA fragments occur in a hollow fused silica capillary. This capillary is usually between 25-75 cm in length (most commonly 47 cm), with a 50-100 µm internal diameter (ID), and is coated on the outside with polyimide to increase strength and durability (Figure 6.2). Fused silica capillaries dissipate heat very efficiently [Weinberger 2000] which combined with a small ID (usually 50 µm) permit the usage of high voltages (up to 15-25 kV) with minimal joule heating and little to no loss of resolution. Joule heating refers to the heat that is generated when an electric current flow through another resistive medium, in this case that would be the polymer and low joule heating allows for increased speed of separation.

Electroosmotic flow (EOF) will greatly affect the separation of DNA fragments in the case of uncoated silica capillaries due to the residual silanol groups on the silica surface. At a pH >5 the hydroxyl groups on the silanol molecules undergo ionization to SiO^- and attract the cations from the buffer solution, interacting with them electrostatically [Camilleri 1998]. An electrical double layer is created (Figure 6.2) with the outer layer being mobile and moving slowly towards the negative electrode (when electric field is applied) while the inner layer remains stationary [Butler 2011]. This is problematic for reproducibility as it causes differences in the separation velocities of the DNA fragments from run to run. This EOF can be reduced by coating the interior of the capillary walls with the linear polymers of the sieving matrix.
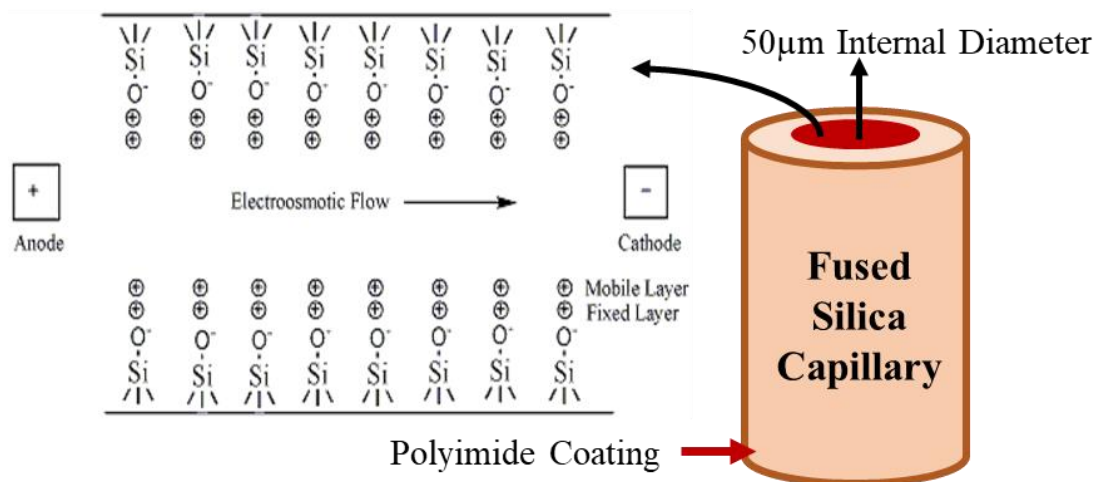
Fig.6.2 On the left a schematic of a fused silica capillary with polyimide coating on the outside and a 50µm internal diameter. On the right, a schematic of the interior of the silica capillary with the silanol groups along the wall surface. Coating of the wall surface can reduce the electroosmotic flow. Image created by author.

In each capillary, a small transparent optical window is burned through the polyimide coating allowing the laser to penetrate the capillary and excite the dye labels of the DNA fragments while also permitting the resulting fluorescence to be captured by the detection device [Yeung 1991, Butler 2011].

B.3.2 Sieving Matrix

The sieving matrix is the most important factor in achieving fast, high resolution separation of DNA fragments, due to the fact that DNA mobility is independent of [Olivera 1964]. The sieving matrix can be prepared from a number of different polymers or mixtures of polymers. Early CE methods utilized agarose or cross-linked polyacrylamide gels. These however were difficult to prepare and could not be reused [Barron 1995 a,b] and were later replaced by the development of aqueous polymer solutions. The polymer acts as a sieve for the DNA fragments as they migrate through the capillary. The method by which these interactions and sieving occurs is not fully understood however there are two simplistic

96

models that effectively describe this mobility of the DNA through the polymer matrix, known as Ogston sieving and reptation. Ogston sieving, first described in 1958 describes the movement of spherical particles through a mas of fibres [Ogston 1958]. When applied to the separation of DNA fragments this would assume that these fragments were in a spherical form with a set radius of gyration as they migrate through the polymer matrix [Rodbard 1970]. The smaller fragments of DNA will have smaller radii of gyration allowing them to move more freely through the pores created by the polymer entanglement, while the larger fragments with their larger radii of gyration will be more hindered as they progress through the polymer [Grossman 1991 a,b]. Movement of the DNA fragments through the polymer is thus determined by both the size of the fragments and the size of the pore created by the polymer strands, which is determined by the concentration of the polymer and its molecular weight in solution. The concentration of polymer in solution must be high enough that the strands interact with each other to form an entangled matrix but not so high that they form a conglomerate that the DNA fragments would not be able to pass through

Another theory of how the DNA migrates through the polymer matrix is the reputation model. This model is used to describe situations when the DNA fragments have a radius of gyration larger than the pore size of the entangled polymers. These fragments under the Ogston model would not fit through the pores in the matrix and so they would not separate. The reputation model suggests that in these cases the DNA would unwind and thus move through the pores in a snake-like manner [Butler 2005], which leads to the mobility of the fragments being related to their length once unwound (Equation 6.3).

$$m = \frac{1}{L}$$

Equation. 6.3 Reptation model equation showing the relationship between the length of the DNA fragment and its mobility. μ is the mobility of the DNA molecule and L is the length of the fragment.

Equation 6.3 can be adapted to incorporate the effect of field strength on the DNA molecule, which increased field strength causing the DNA to elongate to a rod like shape (Equation 6.4) [Grossman 1991]. As can be seen from equation 6.4, with this increase in field strength the length of the DNA fragments becomes less important in their overall mobility.

$$m = \left(\frac{1}{L}\right) + bE^2$$

Equation 6.4 Adapted reputation model equation to account for increase electric field strength. Where b is a function of the pore size and E is the electric field strength.

Both models are able to predict the separation of the DNA fragments through the polymer matrix but each on its own does not explain how this separation occurs in dilute polymer solution. In order to explain this a combination of these models must be used.

B.3.3 Buffer

The buffer composition and concertation used in conjunction with the sieving matrix plays a critical role in DNA fragment separation, as its contents are responsible for the solubility of the DNA, the ionic strength of the solution as well as its denaturing capability and pH. These in turn affect the quality of injection and the current flow. Electroosmotic flow (EOF) is highly sensitive to changes in pH of the buffer solution [Rose 1998].

$$\mu_{EOF} = \frac{\in \zeta}{4\pi\eta} E$$

Equation 6.5. Electroosmotic flow (EOF). Rate at which sample moves through capillary

In Equation 6.5, $\in$ is the buffer's dielectric constant, η is the viscosity of the buffer, E is the field strength and ζ is the zeta potential at the interface of the capillary and the buffer solution.

The pH also determines the degree of ionization of the silanol groups present on the capillary walls, so at high pH these become fully ionized resulting in increased electroosmotic mobility. Th pH can also change the state of the DNA and affect its mobility [Hartzell 2005].

Increasing buffer concentrations can help contain EOF [Rose 1998] due to the inverse relationship between the rate of EOF and the ionic strength of the buffer, but if the concentration is too high then the current to can increase too much leading to excessive heating and poor resolution of DNA fragments. If the buffer concentration is too low the EOF is also low which results in long retention times leading to broad peaks and poor restoration of the capillary due to diffusion.

Capillary electrophoresis relies on the creation of an electrical current which requires buffer reservoirs at both the cathode and the anode. These reservoirs should be periodically refreshed with new buffer to avoid changes in the concentration between the two, which can arise from degradation of ions during the runs and can lead to current problems. The buffer used is normally created with similar components as the sieving matrix employed. The purpose of the buffer is to stabilize the pH while also adding EDTA to chelate metals, urea and pyrrolidine to denature the DNA and salts to help regulate EOF. reduce chemical

interaction between compounds. Most commonly the buffer composition used consists of 100mM 3-[[1,3-dihydroxy-2-(hydroxymethyl)propan-2-yl]amino] propane-1-sulfonic acid (TAPS) (Figure.6.3), with the pyrrolidinone and urea acting as denaturants. The addition of 1mM ethylenediaminetetraacetic acid (EDTA) (Figure.6.3) to act as a chelating agent removing metals from the solution is also commonly observed. The buffer is then adjusted to a pH of 8 with sodium hydroxide [Butler 2006, Butler 2011]. It is important to note that the compounds and any additives used must be stable at the high denaturing temperatures employed during CE (60°C).
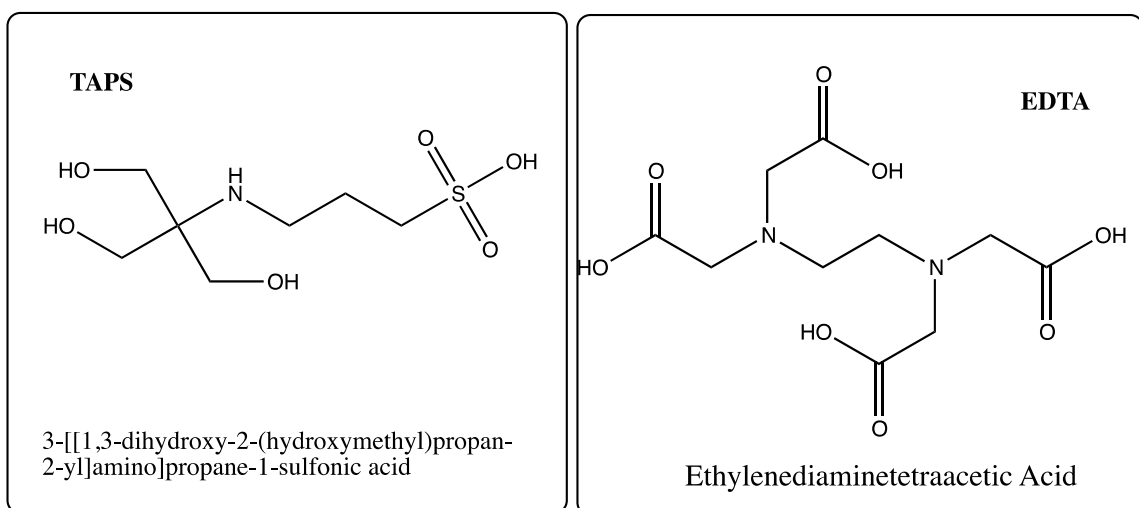
Fig.6.3 Chemical structures of TAPS and EDTA, the two main components of the buffers used in capillary electrophoreses for the separation of DNA fragments.

B.4. Sample Detection

The primer pairs used in forensic STR analysis are labelled with fluorescent dyes which are covalently bonded to the 5' end of the forward primer. These labels allow the DNA fragments produced during PCR to be detected by laser-induced fluorescence (LIF) when they flow through a window burned into the wall of the capillary. The laser used to excite the dyes is a single argon-ion laser at 488 nm, and the dye labelled amplicons emit at

multiple wavelengths ranging from 520-680 nm. These emissions of the excited fluorophores are recorded by a charge coupled device (CCD) camera allowing for the discrimination between dyes [Butler 2004] that fluoresce in different regions of the spectrum. A deconvolution of the signal is utilized to isolate the spectral contribution of each dye. This calculation relies on the linear relationship between fluorescence and the concentration of the DNA fragments and permits the relative concentration of each dye to be determined at a given wavelength. Excess fluorescent signal creates a non-linear condition producing cross over of the dye signal into neighbouring lanes. This produces artefactual peaks which show up in neighbouring dye lanes as small peaks. The artefacts produced by this process are known as "pull up" [Butler 2009].

B.5. Interpretation

The spectral data collected is transformed into a multi-coloured electropherograms which are then analysed using the GeneMapper® Analysis software (Applied Biosystems, Valencia, CA) [Capt 2005].

This software takes the raw data and determines the size of each STR allele by plotting it on a standard curve made from the data received from an internal size standard (ILS). Determining allele size by comparison to the ILS using the local southern or the global southern methods curve fitting [Butler 2011]. These methods differ in that the local method uses an interpolation method by using two size peaks before and two size peaks after the unknown allele to calculate its size, while the global method uses all of the ILS peaks to calculate the unknown alleles size by regression analysis [Hartzell 2003].

Once the alleles are sized STR genotyping can then be performed by comparing the allele sizes in each sample to those present in the allelic ladder [Schumm 1997] and inserting

them into virtual labelling bins, created by the user or commercial kit developer, for those alleles. These bins represent the possible size range for each allele. They must be of sufficient width to encompass statistical variations in allele size. An allelic ladder composed of known alleles is used to adjust bin positions, which can vary slightly between runs due to changes in mobility. The result of the software analysis is a genotype of the DNA contributor which can be compared to other known profiles or uploaded into a national or international database for comparison.

C. Microfluidic Separation

An alternative to capillary based analysis is electrophoresis by microchip. This technique is performed using glass on plastic chips in which small channels are etched or embedded on the substrate. The electrophoresis mechanism is similar to CE but has an improved injection process which decreases peak width and reduces run time. The microchips are typically only used once as they are difficult to refill. For these chips to work efficiently it is important that the design of the channels and the viscosity of the sieving matrix be optimal. If the channels are too small or the polymer to viscous then the chip cannot be properly loaded. If channels are too large this can lead to excess diffusion resulting in poor separation. Commonly used substrates for microchips include glass, silica, polydimethylsiloxane (PDMS) and polymethyl methacrylate (PMMA) [Horseman 2007]. The channels etched into the substrate are usually 2-20 cm long, 4-100 µm wide and 10-30 µm deep [Heller 2002]. These dimensions can vary depending on the needs of the manufacturer and the desired resolution.

Injection of sample is performed using a crossed T or a double T intersection (Figure 6.4) [Thomas 2004]. The design of the chip is chosen based on the amount of sample desired to be injected. Injections can occur by either a gated or a pinched mechanism. The gated mechanism is used when sample injection needs to occur continuously. This is accomplished by the chip having a crossed T design where continuous flow perpendicular to the injection channel pushes the sample directly to the waste with none entering the separation channel. This flow is stopped and sample then enters the separation channel and the flow started again (in this way the amount of sample injected can be controlled by however long the perpendicular flow is stopped [Landers 2008].

A pinched injection, is more common in the double T design. In this type of injection, the sample is held between the sample and waste wells by a "push back" voltage pulling the sample backwards and preventing it from entering the separation channel. This is important because if some of sample enters the separation channel too early (while the rest of the sample is still being injected) it could lead to peak broadening and decrease of resolution [Landers 2008].
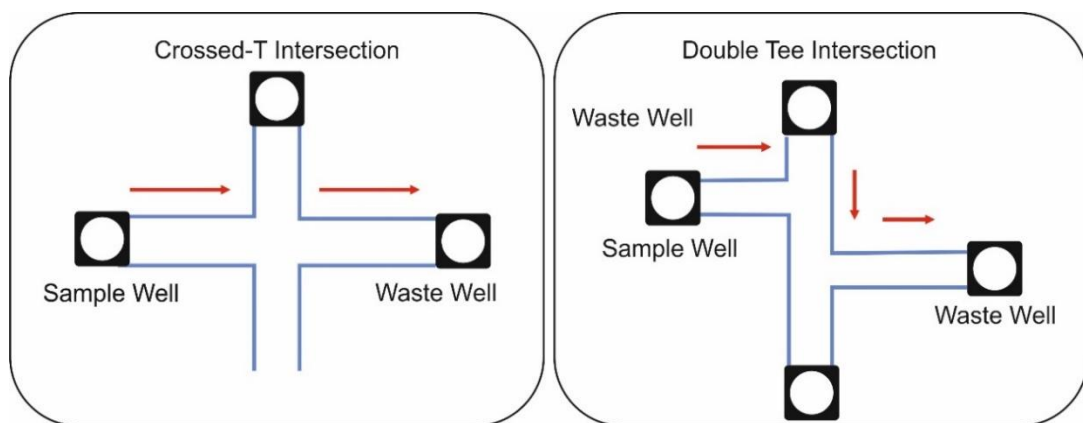
Fig.6.4 Schematic of two types of microchip design. On the Left, is shown the most commonly used Crossed-T Intersection, while on the right is shown the Double-T Intersection which allows for multiple rapid injections. Image recreated by author, original by Aboud 2012.

The filling of the microchip channels is accomplished with the aid of either a vacuum or a syringe to ensure no bubbles are present. Care must be taken to avoid bubble formation as this can block the channels necessary for injection and separation and interrupt the conductance.

Applied voltages are used in ME to control sample injection and migration of the DNA fragments while detection is achieved through laser induced fluorescence. A schematic of a pinched, cross-tee injection can be seen in Figure 6.5.
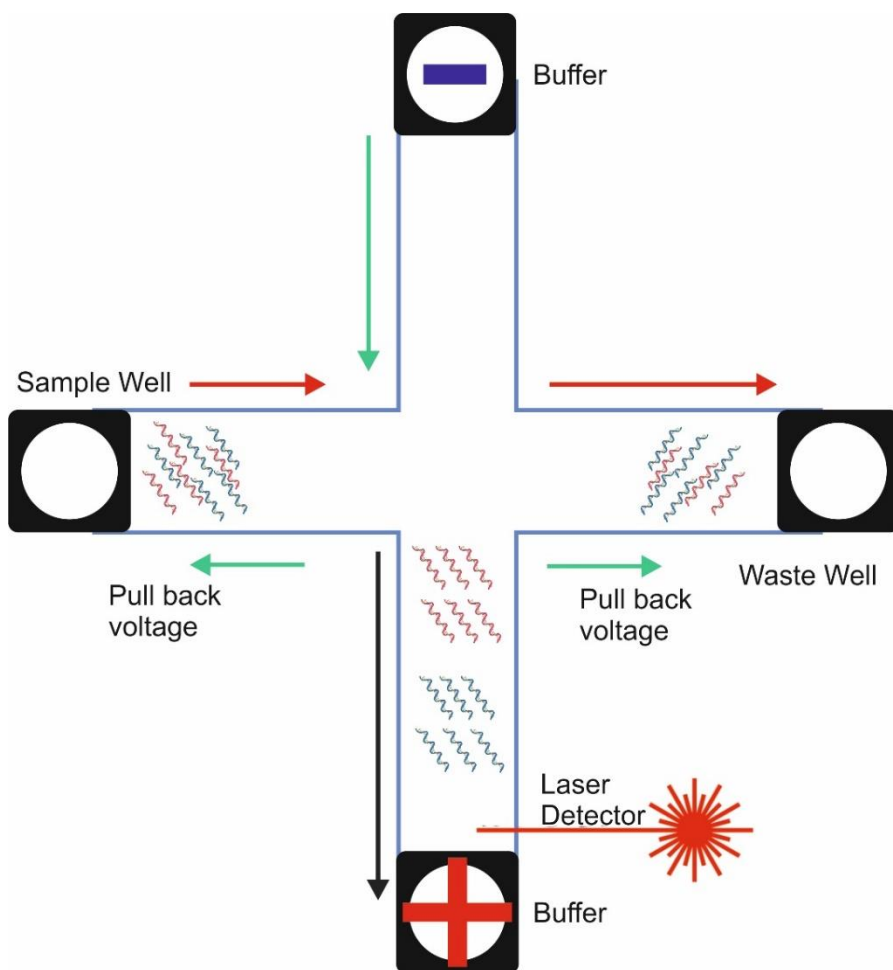
Fig.6.5 Schematic of a crossed-T microchip using a 'pinched injection' with pull back voltages. Separation occurs once the sample migrates through the separation channel. Same as in CE the DNA fragments separate by size due to their interaction with the entangled polymer matrix. Detection occurs using laser induced fluoresce of fluorescently labelled primers of the fragment [Aboud 2012]. Image recreated by author, original by Aboud 2012.

An example of a microfluidic electrophoresis system is shown in Figure 6.6. The main components of this method are microfluidic chips, which are glass plates containing etched-in 1.8 cm long separation channels of µm scale [Aboud 2013]. These glass plates are glued to a plastic caddy providing wells for the application of the samples. The system utilises a voltage of 1500 V for sample separation and a single excitation wavelength of 532 nm with emission wavelengths of 575 and 670 nm for sample detection.
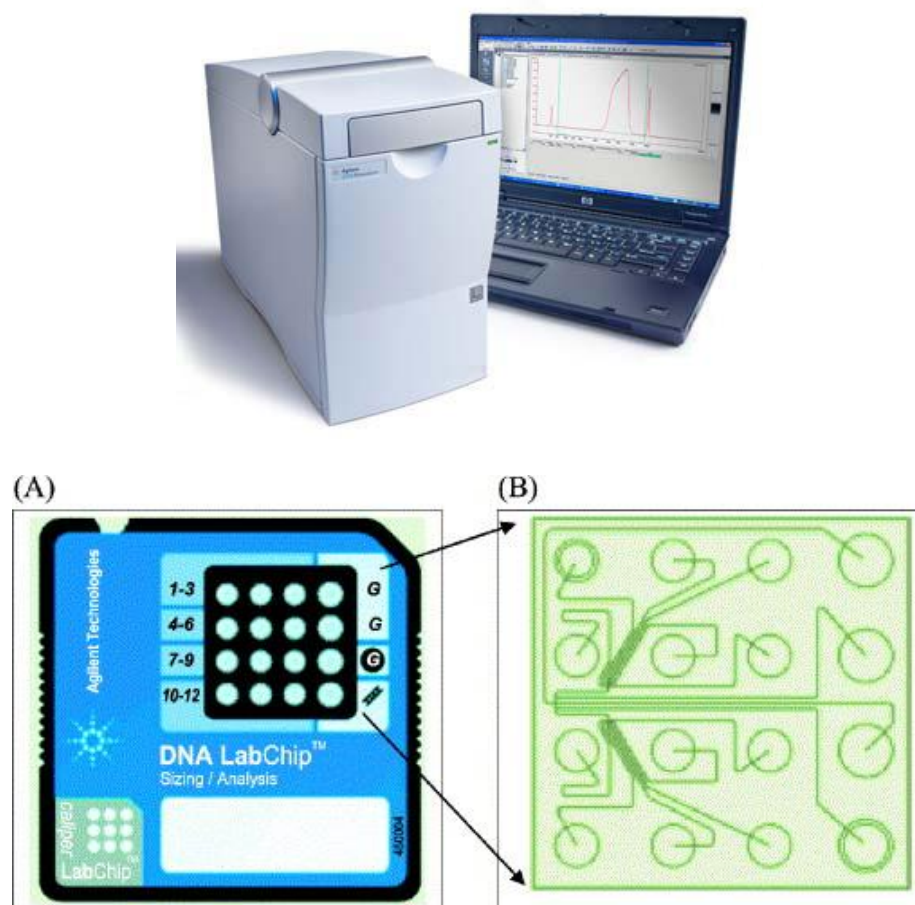
Fig.6.6 On top is the Agilent 2100 Bioanalyzer. the bottom images are schematics of the microchip used with the 2100 Bioanalyzer for microfluidic electrophoresis. (A) Image of the chip as seen by user (actual size), showing wells marked 1-12 along with 3 G wells and a final well for the ladder. (B) view of the inside the chip with the dark green lines indicating the microchannels etched into the glass pane. Adapted from archimede.bibl.ulaval.ca.

The low volume of reagents and sample used allows for run times of 80s per sample. While the small size of the instrument facilitates its portability, leading to its possible use for on-site sample analysis [Aboud 2015].

CHAPTER VII. DEVELOPMENT OF RAPID MICRFLUIDIC ANALYSIS OF A Y-STR MULTIPLEX FOR SCREENING OF FORENSIC SAMPLES. [*]

A. Introduction

It is often important to rapidly screen suspects who may have been involved in a crime, especially when time is of the essence or when large numbers of samples may need to be quickly processed.   These situations create a need for rapid screening of DNA samples. While current DNA typing methods provide the best biometric information yielding identity, kinship and geographical origin, they are not sufficiently fast to permit identification of a suspect's DNA in real time. Rapid polymerase chain reaction (PCR) procedures [Vallone 2008, Vallone 2009, Giese 2009] can greatly speed the processing time of forensic samples [Verheij 2012] because of the high processivity of newly designed fast polymerases, particularly when combined with newly developed rapid thermocyclers and microfluidics. The decrease in processing time and minimal use of reagents can lead to improved processing of forensic samples [Vallone 2008, Vallone 2009, Aboud 2013]. However, the current generation of rapid PCR systems are expensive to utilize, and most require an extraction step prior to analysis [Mapes 2016]. We have found that the speed of these analyses can be greatly increased by utilizing a smaller set of PCR amplicons along with new fast polymerases and fast thermocyclers [Aboud 2013]. When used for sample screening, such a system could rapidly triage different suspects, excluding those not involved in criminal activities. Furthermore, such a system could be useful in situations

---

[*] This chapter has been published: Gibson-Daw G, Albani P, Gassmann M, McCord B. Rapid microfluidic analysis of a Y-STR multiplex for screening of forensic samples. Anal. Bioanal. Chem. 2017;409(4):939–47

such as mass disasters where it is necessary to search for close relatives or to analyse evidence that is comingled. Thus, having the capability to perform Y chromosomal short tandem repeat (Y-STR) typing rapidly could be very useful. It was the goal of this project was to develop a rapid method for screening male DNA samples. The use of microfluidics in place of regular capillary electrophoresis (CE) methods permits analysis to be performed in under 80 seconds.

Y-STRs were chosen for this project as they are present only on the male chromosome, which means that in mixtures of male and female DNA separation of the two would be unnecessary. Also in situations in which there is a large excess of female DNA it is often difficult to obtain a clear autosomal male profile [Redd 2002]. Y-STRs show a high level of diversity when compared to other genetic polymorphisms, because of their high mutability [Ravid-Amir 2010] which has been documented at a range of 10-6 to 10-2 per meiotic division [Schlötterer 2009, Leclerq 2010]. These mutations are usually caused by slippage of the DNA polymerase during replication [Schlötterer 2009, Payseur 2011], leading to insertions or deletions in the Y-STRs repeat motif [Kelkar 2011]. The markers selected for this project show a higher polymorphism rate than other Y-STRs and so are known as rapidly mutating Y-STRs (RM Y-STRs) [Ballantyne 2012]. These loci should prove particularly useful in sample screening where related individuals are present [Redd 2002].   Recently published work has demonstrated that fast amplification of Y-STRs is possible.   For example, Abuidrees et al. have developed a 28-minute amplification of rapidly mutating Y-STR markers using Phusion Flash high fidelity polymerase [Abuidrees 2016], while Kim et al. have successfully demonstrated an integrated system for standard Y-STR markers using a slidable chip format that performs a 26-minute amplification [Kim

2016]. Our system can amplify a small Y-STR multiplex, suitable for screening evidence and suspect samples in 16 minutes. Subsequent analysis by microchip takes 80 seconds. While not integrated, the technology described here is inexpensive to implement and can easily be adopted by current forensic laboratories for sample screening using standard capillary electrophoresis methods or microfluidics.

Specially engineered enzymes [Wang 2004, Giese 2009, Kermekchiev 2009], high speed thermal cyclers (capable of running 28 cycles in 16 minutes) and microfluidic chip electrophoresis [Woolley 1998, Shi 2003, Goedecke 2004, Shi 2006, Hopwood 2010] were implemented to process a specifically designed Y-STR multiplex. The designed multiplex included four rapidly mutating (RM) Y-STRs DYS526a/b, DYS576, DYS626, and DYS570 between 137 and 402 bp in size, with mutation rates of 10-2 per meiosis or greater. By using off the shelf instruments and commercially available enzymes it was possible to create a procedure that acts as a quick, highly informative sample screening process that also retains sufficient DNA for later manual processing using standard STR or Y-STR kits.

In the first phase of this study we designed a 4 loci Y-STR multiplex and utilized both a conventional 310 CE as well as a beta version of a denaturing microfluidic electrophoresis (ME) system provided by Agilent Technologies. The equipment was tested on control DNA standards 2800 M DNA and 9948 DNA (Promega) as well as with donated saliva samples from 20 adult males. The multiplex was next analysed using a rapid PCR protocol, using rapid polymerase Z-Taq (Takara) and the Philisa rapid thermocycler (Streck) in an effort to optimize the speed and balance of the rapid amplification. We then optimized the

different aspects of this procedure such as concentrations of different reagents, number of cycles, annealing temperatures, and time spent at denaturing, annealing and elongation steps. This was then coupled with microfluidic separations to further reduce sample analysis times to under 25 minutes. A sensitivity study to determine the lowest concentration of input DNA needed for successful amplification was also completed. The results of this study demonstrate a new application of rapid PCR and microfluidic separation for the analysis of Y-STRs for suspect screening.

B. Methods and Materials

B.1. DNA Samples

Male DNA standards 2800 and 9948, were obtained from Promega (Madison, WI, USA) and used at a concentration of 1ng/μL. These were then used as positive controls for the genotyping process. A test population was obtained from 20 individuals using Sterile Cotton Tipped Applicator swabs from Puritan Medical Products (Guildford, ME, USA) to swab the inside of their cheeks and extracted by PCIA [Budowle 1990, Comey 1994]. This method of extraction usually yields from 3-50 ng/ μL of DNA but for reasons of developing a rapid method the samples in this study were not quantified after extraction. The collection and analysis of these samples was approved by the Institutional Review Board (IRB) at FIU, under the reference number #101831.

B.2. RM4 multiplex and Primer sequences

A multiplex containing rapidly mutating Y-STRs (RM4) was used for these experiments (Figure 7.1.). This multiplex was assembled from 5 rapidly mutating (rates of 10^{-2} per

meiosis or greater) Y-STRs. The selected loci showed a mixture of simple 4 base repeat motifs and complex repeat motifs, these included DYS526 a/b, DYS576, DYS626, and DYS570. While 5 loci were amplified only 4 primer sets were used, due to the fact that the primers designed for locus DYS526 bind to two separate locations on the chromosome (therefore these loci are known as DYS526 a and DYS526 b respectively). The loci combined in the RM4 multiplex were selected on the basis of size so that when multiplexed their allele ranges would not overlap with one another (Figure. 7.1.) while also allowing the mixture to be run on both the Applied Biosystems ABI 310 and the Agilent Bioanalyzer. Additionally, when combined these loci had sufficient discriminatory power to screen samples. The primers associated with these loci were labelled with the TAMRA fluorescent dye and can be seen in Table 7.1. along with details of the loci including repeat motif, size range, alleles size range and the genetic diversity for each locus.
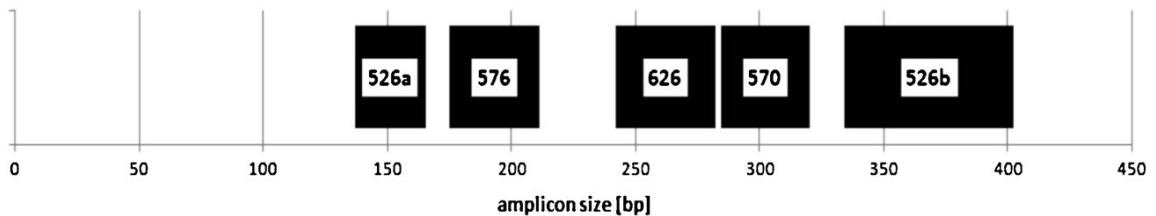
Fig.7.1 Size range of each loci included in the RM4 multiplex

Table 7.1: Loci and Primer Information

Loci	Repeat motif	Primer Sequence 5'-3'	Product sizes	Allele Range	Genetic Diversity
DYS526a	(CCTT)10-17	56-TAMN-TCTGGTGAACTGATCCAAACC GGGTTACTTCGCCAGAAGGT	137-165	10-17	0.8267(23)
DYS576	(AAAG)13-22	56-TAMN- TTGGGCTGAGGAGTTCAATC GGCAGTCTCATTTCCTGGAG	175-211	13-22	0.7836(24)
DYS626	(GAAA)14-23N24(GAAA)3N6(GAAA)5(AAA)1(GAAA)2-3(GAAG)1(GAAA)3	56-TAMN-TGCAAGACCCCATAGCAAAAG AAGAAGAATTTTGGGACATGTTT	242-282	16-26	0.8365(25)
DYS570	(TTTC)14-24	56-TAMN-GAACTGTCTACAATGGCTCACG AGAAGGGCTTCTAAGGGATG	284-320	14-24	0.7972(24)
DYS526b	(CCCT)3N20(CTTT)11-17(CCTT)6-10N113(CCTT)10-17	56-TAMN -TCTGGTGAACTGATCCAAACC GGGTTACTTCGCCAGAAGGT	334-402	27-44	0.8821(23)

Table 7.1 Loci used in multiplex, the repeat motif of each loci, forward and reverse primer sequences used to amplify each locus, product size, allele ranged and genetic diversity of each locus [Butler 2006, Zhang 2014, Alghafri 2015].

B.3. Thermal Cycling and Rapid Polymerase

The thermal cycler used in this research was the Philisa from Streck (Nebraska, US), see Figure 5.1. The Philisa has a cooling rate of 12ºC/s and a heating rate of 15°C/s. This instrument is comprised of an 8-well low thermal mass silver sample block and has dimensions of 285 mm x 202 mm x 215 mm, with a weight of 7.5 lbs making it easy to transport and field friendly as it can also be controlled by either a desktop or laptop computer. By using a dual Peltier module for heating and cooling the Philisa® achieves heating rates of 15ºC/s and cooling rates of 12ºC/s. The instrument produces high temperature ramp rates using flat thin walled 50 µL polypropylene tubes containing a relatively high surface to volume ratio (Figure 5.1). This aids in heat transfer from the silver block to the PCR mix.

For this study Z-Taq polymerase (Takara Bio, Clontech Laboratories, Inc., Mountain View, CA), supplied with 10 X Z-Taq buffer and 2.5mM dNTP mixture, was used which had previously been shown in studies by Aboud [Aboud 2013] to work well for achieving ultrafast PCR amplifications due to its high processivity (over 5X speed of regular Taq) and high fidelity [Wang 2004, Kermekchiev 2009, de Vega 2010].

B.4. PCR Optimization

B.4.1. Testing of multiplex

Preliminary testing of the RM4 multiplex was amplified in a 50 µL total volume master mix. The master mix was made up of 1 X Ramp-Taq buffer (Promega), 2 mM MgCl$_2$, 0.6 mM each dNTPs (Promega), TAMRA fluorescently labelled primers at concentrations of

0.72 µM DYS526, 0.19 µM DYS570, 0.48 µM DYS576 and 0.60 µM DYS626 (IDT), 1.25

Units/50 µL of Ramp-Taq polymerase (Promega). 1 ng of DNA was added and the mix

amplified with the following cycle conditions using a GeneAmp system 9700 with a ramp

rate of 5ºC/s (Applied Biosystems): 95°C for 10 mins, 25 cycles of 94°C for 30 s, 60°C for

30 s, 75°C for 25 s, then a final extension step of 72°C for 7 mins.

B.4.2. Optimization of Master Mix

Once the RM4 multiplex was successfully tested the master mix was optimized for faster

amplification speeds and lower total volume (Table 2). Amplification was performed on a

Philisa thermocycler (Streck) using the following cycling conditions: 25 cycles of 94°C for

30 s, 60°C for 30 s, 75°C for 25 s. PCR products were then run on the ABI 310 Following

directions seen in section B.8.1.

Table 7.2. Master Mix Optimization. Total volume 10 µL							
Master Mix	2.5 U/ µL Z-Taq (µL)	Primer Mix (µL)	10X Z-Taq Buffer (µL)	2.5 mM dNTPs (µL)	20 mg/mL BSA (µL)	25 ng DNA (µL)	Water (µL)
A	0.3	2	2	0.8	0	1	3.90
B	0.2	2	2	0.8	0	1	4.00
C	0.2	1.5	2	0.8	0	1	4.50
D	0.3	1.5	2	0.8	0	1	4.40
E	0.25	2	2	0.8	0	1	3.95
F	0.25	2	2	0.8	0	1	3.45
G	0.25	2	2	0.8	0.5	1	3.40

Table 7.2 Shows the Master mix reagent combinations and concentrations. Concentrations of Z-Taq Polymerase and primer mix were varied to determine the optimum concentrations of these reagents. G also has the addition of BSA to help counteract any inhibition from the sample.

B.4.3. Optimization of Cycling Conditions

Amplification was performed on the Philisa thermocycler (Streck) using a variety of different annealing temperatures. The annealing temperatures tested with Master Mix A were 70°C, 65°C, 62°C, 60°C, 58°C. In addition, the initiation step and the final elongation step were removed to speed up the amplification time, the denaturing temperature was increased to 95°C and the elongation temperature lowered to 72°C. Time at each cycle was decreased 1 s at a time until optimal conditions were determined and cycle number was increased. These optimal conditions included heat denaturation at 95 °C for 5 s, followed by 29 cycles of 98 °C for 5 s, 62 °C for 9 s and 72 °C for 12 s, yielding amplification in 16 min and 11 s. PCR products were then run on the ABI 310 Following directions seen in section B.8.1.

B.5 Rapid PCR Amplification

The final conditions for the optimized process used in all subsequent studies in this chapter were: 9 µL of Master Mix A with 1 µL of control DNA or DNA extracted from the volunteer's cheek swabs to give a total volume of 10 µL. This master mix included 2 µL 10 X Z-Taq buffer (Takara Bio), 0.8 µL 2.5 mM each dNTPs (Takara Bio), 2 µL Primer mix (IDT), 3.9 µL of H_2O and 0.3 µL of 2.5 U/µL Z-Taq polymerase (Takara Bio). Final concentrations per 10 µL reaction were 2 X Z-Taq buffer, 0.2 mM each dNTPs, primers: 0.72 µM DYS526, 0.19 µM DYS570, 0.48 µM DYS576 and 0.60 µM DYS626 and 0.075 U/10µL Z-Taq polymerase. Amplifications were performed on a Philisa thermocycler (Streck) in just over 16 min. Optimal conditions were 29 cycles of heat denaturation at 95°C for 5 s, followed by an annealing at 62°C for 9 s and elongation at 72°C for 12 s. PCR products were then run on the ABI 310 Following directions seen in section B.8.1.

B.6. Sensitivity Study of Rapid Amplification

To test the sensitivity of the system a range of input DNA concentrations (9948 standard DNA) from 2 ng down to 0.25 ng was amplified using the above optimized method and analysed using the ABI 310 for separation and detection. Three replicates of each experiment were run on the ABI 310.

B.7 Mixture of Male and Female DNA

Mixtures of varying ratios of female to male DNA (9947/9948 DNA standards) were created in order to test the system's capability of generating useable male profiles in the presence of an overabundance of female DNA. These females to male ratios ranged from

0.25:1 to 40:1, with a consistent input of 1ng of male DNA. Concomitantly, the amount of input of female DNA increased according to the input ratio of F:M used. Mixtures were amplified and analysed using the same method as seen in section B.5. All experiments were run in 3 replicates on the ABI 310 as in previous sections.

B.8. Separation, Detection and Analysis

B.8.1 Applied Biosystems ABI 310 Capillary Electrophoresis Genotyper

All PCR products were analysed using an ABI 310 (Applied Biosystems) by adding 1 µL of PCR product to a mixture of 12.5 µL HiDiTM Formamide (Applied Biosystems) and 0.5 µL internal lane standard (ILS). The ILS used throughout this paper was LIZ 500 (Applied Biosystems) which is labelled with the orange LIZ dye. The mix was then run on the ABI 310 using the following conditions: Polymer used was POP4, Capillary length to detector: 36 cm, module GS STR POP4 (1 mL), injection voltage 15 kV, injection time 5 s, run time 28 min, run temperature 60°C, run voltage 15 kV. Software used for Analysis was Genemapper.

B.8.2. Agilent Technologies Microfluidic Chip Bioanalyzer 2100[*]

The modified Agilent Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA) equipped with a heat plate and denaturing polymer was used in these experiments and compared with the standard ABI 310 (Applied Biosystems, Life Technologies Corporation, Carlsbad, CA, USA). In the device, electrophoresis takes place on a glass microchip with a 1.5-cm separation channel. Up to 12 samples may be analysed in series, with microfluidic

---

[*] Work completed on Bioanalyzer 2100 and data presented was a collaboration between the author and Patricia Albani a former Masters student of the McCord DNA lab.

separations of all four loci completed in less than 80 s. Full details of the system are provided in previous work by Aboud et al. [Aboud 2013]. The system utilizes a voltage of 1500 V for sample separation and a single excitation wavelength of 532 nm with emission wavelengths of 575 and 670 nm for sample detection. The two fluorescent dyes utilized in these experiments were TAMRA and LIZ as their emission wavelengths were compatible with the detection capabilities of both the ABI 310 and the Agilent Bioanalyzer. DNA amplicons were labelled with TAMRA dye, and the GeneScan LIZ 500 Size Standard was used as an internal lane standard in the red channel for sizing precision [Aboud 2015]. The chip was prepared by pipetting 15 µL of the denaturing polymer from the Agilent high-resolution DNA prototype reagent kit, and then pressure was applied with a syringe for 4 min to allow the polymer to flow through all areas of the chip channels. Fifteen microlitres of the polymer was then added to the G and ladder wells. All samples were mixed with formamide prior to analysis. A master mix of 150 µL HiDi™ formamide (Applied Biosystems), 150 µL green Agilent HR marker buffer and 7.5 µL LIZ 500 Size Standard was prepared. Two microliters of the sample were added to 12 µL of this mix, and 6 µL of this was then added to the sample wells of the chip with the 12th well holding just the ladder sample. Sizes were determined using the peak retention times in an Excel software table and comparison to the simultaneously run sizing standard LIZ 500 (Applied Biosystems). The results from this system were compared to the conventional CE method using an ABI 310 (Applied Biosystems). Samples were mixed with formamide prior to analysis. This mix consisted of 12.5 µL HiDi™ formamide (Applied Biosystems) and 0.5 µL LIZ 500 Size Standard per 1 µL sample. Run conditions of ABI 310 are as follows: POP4 used as the polymer, capillary length to detector 36 cm, module GS STR POP4 (1 mL), injection

voltage 15 kV, injection time 5 s, run time 28 min, run temperature 60 °C and run voltage 15 kV. The samples used in this experiment were extracted male buccal sample, run in duplicate on both this system and the ABI 310 as seen above and the results compared for concordance.

C. Results and Discussion

Optimisation tests were performed using two DNA standards (2800 male DNA and 9948 male DNA; Promega) as well as the donated saliva samples from 20 adult males, and mixture tests were performed using the 9947 female DNA and 9948 male DNA (Promega). The multiplex was next analysed using a rapid PCR protocol, rapid polymerase Z-Taq (Takara) and the Philisa rapid thermocycler (Streck) in an effort to optimize the speed and balance of the rapid amplification. We then optimised the different aspects of this procedure such as concentrations of different reagents, number of cycles, annealing temperatures and time spent at denaturing, annealing and elongation steps. A number of samples were run on both the ABI 310 and the Bioanalyzer 2100 to test for concordance between the two systems.

C.1 Y-STR multiplex

The first step in developing this procedure was to create a Y-STR multiplex using specific loci that had previously been shown to have high mutability rates. Higher mutation rates can lead to improved power of discrimination when comparing samples from multiple males, especially if some of these may be related. To do this, we chose four loci that did not overlap, and once this was completed, we tested the multiplex with both control DNA and extracted saliva samples. The ratio of each primer added to the mixture was optimized

in order to obtain the best overall peak balance. Final values of 3:2:0.8:2.5 were used for DYS526a/b, DYS576, DYS570 and DYS626, respectively. Figure 2 illustrates an electropherogram of the RM4 multiplex. The result demonstrates the successful amplification of four separate Y-STR loci. Of note here is that the locus DYS526 shows two peaks, as this primer pair anneals to two different loci. This result was expected based on the observed literature [Ballantyne 2012, Alghafri 2015], and these loci will be known henceforth as DYS526a and DYS526b, respectively.
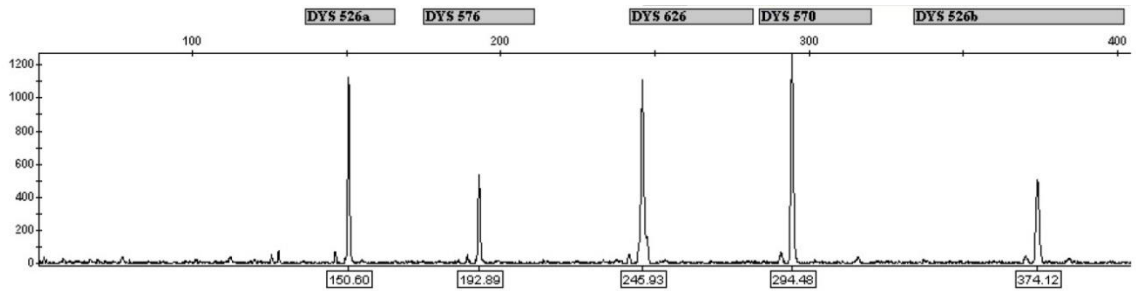
Fig.7.2 Multiplex run with extracted DNA from male Vat an input concentration of 1 ng in ~90 min. Run on ABI 310

C.2. Rapid Y-STR PCR amplification

In order to obtain profiles with no allele dropout and good peak heights (above 150 relative fluorescence units (RFU)) and balance, the concentration of Z-Taq, primers and BSA was optimised along with the annealing temperature. The optimisation was conducted by testing a series of lower and higher concentrations and annealing temperatures compared to those used with the non-rapid PCR seen in the previous section. It was expected that due to the rapid nature of the process the possible addition of BSA may have been needed in order to help counteract any inhibition.

C.2.1. Optimization of Master Mix

The results of the optimisation of master mix reagents by their observed peak heights can be seen in Table 7. 3. The Master Mix A (MM A) combination of Z-Taq (0.075 U/ 10 µL) and primer mix (0.72 µM DYS526, 0.19 µM DYS570, 0.48 µM DYS576 and 0.60 µM DYS626) gave the best results with no allele dropout and more intense peak heights noted. This was also tested with added 0.5 µL BSA to give a final concentration of 1µg/µL per 10 µL reaction mix (master mix G), however no further increase in peak height was observed. This was possibly due to the fact that the samples tested in this series of experiments were extremely clean and thus less subject to inhibition, which BSA would help counteract in less well-purified samples.

Master Mix	Results for each loci									
	DYS526a		DYS576		DYS626		DYS570		DYS526b	
	Allele size	Peak height	Allele size	Peak height	Allele size	Peak height	Allele size	Peak height	Allele size	Peak height
A	152.53 ± 0.12	281 ±18	185.79 ± 0.06	1250 ± 63	242.40 ± 0.11	1295 ± 83	301.18 ± 0.13	1048 ± 84	372.40 ± 0.15	890 ± 110
B	152.70 ± 0.11	189 ±18	185.86 ±0.06	610 ± 63	242.56 ± 0.10	652 ± 83	301.45 ±0.13	514 ± 84	372.66 ± 0.15	760 ± 110
C	152.68* ± 0.22	184.5 ± 28	185.88 ± 0.13	447 ± 59	242.61 ± 0.21	418 ± 65	301.46 ± 0.21	187 ± 34	372.70 ±0.23	420 ± 57
D	152.63 ±0.15	202 ± 28	185.87 ± 0.09	582 ± 76	242.55 ± 0.16	551 ± 90	301.29 ± 0.16	322 ± 36	372.49 ± 0.19	787 ± 56
E	152.69* ± 0.01	192 ± 49	185.98 ± 0.16	440 ± 140	242.73 ± 0.36	460 ± 180	301.50 ± 0.33	350 ±150	372.72 ± 0.33	430 ± 160
F	152.73* ± 0.02	170 ± 20	185.86 ± 0.05	510 ± 200	242.57 ± 0.05	530 ± 200	301.32 ± 0.10	460 ± 200	372.55 ± 0.10	550 ±270
G	152.88 ± 0.25	195 ± 17	185.99 ± 0.17	870 ± 91	242.76 ± 0.28	970 ± 110	301.55 ± 0.31	727 ± 65	372.80 ± 0.34	1050 ± 110

Table 7.3: Optimization of Master Mix

Table 7.3 Shows the allele size and peak height for each of the loci when amplified with different master mix compositions. The size and height variation between replicates demarked by the ± symbol. On the basis of these results it was determined that Master Mix A: 0.075 U/ 10 µL Z-Taq and primer mix (0.72 µM DYS526, 0.19 µM DYS570, 0.48 µM DYS576 and 0.60 µM DYS626) contained the optimum reagent concentrations to amplify the Y-STR multiplex.

C.2.2 Optimization of Cycling Conditions

The annealing temperature was also optimised to obtain optimum sensitivity and peak balance. Since the melting temperature (Tm) varied for the different primers used in the multiplex, annealing temperatures ranging from 58 to 70 °C were employed. The results of the optimisation of the annealing temperature in terms of peak height are listed in Table

7.4. An annealing temperature of 62 °C was found optimal as the resulting profiles displayed maximum peak heights (most of these over 1000 RFU with the exception of DYS526a as this is a secondary binding location of the primers). Furthermore, no allele dropout was observed, and peak heights were balanced across all loci. The results of the optimisation experiments demonstrated that the most important parameter for peak height and balance was the annealing temperature (Table 7.4) as this produced the most significant changes in peak height. Changing reagent concentrations (Table 7.3) resulted in only slight variations in overall peak height. Therefore, optimum conditions were set at an annealing temperature of 62°C using Master Mix A: 0.075 U/10 µL Z-Taq and primer mix (0.72 µM DYS526, 0.19 µM DYS570, 0.48 µM DYS576 and 0.60 µM DYS626). These parameters gave profiles with good reproducibility for all alleles in terms of size and peak height.

Table 7.4: Optimization of Annealing temperature

Annealing Temp.	DYS526a Allele size	DYS526a Peak height	DYS576 Allele size	DYS576 Peak height	DYS626 Allele size	DYS626 Peak height	DYS570 Allele size	DYS570 Peak height	DYS526b Allele size	DYS526b Peak height
70°C	-	-	-	-	-	-	-	-	-	-
65°C	-	-	185.37 ±0.09	681 ± 83	242.66 ± 0.10	299 ± 68	300.18 ±0.17	328 ± 74	372.61 ± 0.20	910 ± 150
62°C	152.53 ± 0.10	282 ± 52	185.79 ± 0.04	1260 ± 260	242.40 ± 0.05	1295.25 ± 98	301.18 ± 0.11	1050 ± 150	372.40 ±0.06	894 ±99
60°C	152.63 ±0.13	410 ± 120	185.85 ± 0.09	660 ± 180	242.44 ± 0.15	1010 ± 170	301.26 ± 0.10	990 ± 100	372.41 ± 0.18	1030 ± 130
58°C	152.52 ± 0.11	327 ± 43	185.68 ± 0.01	347 ± 33	242.33 ± 0.10	950 ± 140	301.25 ± 0.09	703 ± 86	372.45 ± 0.06	820 ± 180

Results for each loci

Table 7.4 Shows the allele size and peak height for each of the loci at different annealing times along with the size and height variation between replicates demarked by the ± symbol. Based on these results it was determined that 62ºC was the optimum annealing temperature for the primers of the Y-STR multiplex.

C.2.3 Rapid PCR Amplification

Figure 7.3 shows the results for the optimized procedure for high-speed PCR using the rapid enzyme Z-Taq. This polymerase has a high processivity [Aboud 2013] rate that permits the fast amplification time observed. All 20 extracted individual samples produced positive results with no allele dropout. Electropherograms A, B and C provide examples of the resultant profiles for control DNA and human samples. Examining the 372.57-bp allele in panel A demonstrates <4 bp resolution between the main allele and a four-base stutter peak one repeat unit upstream of the main allele.
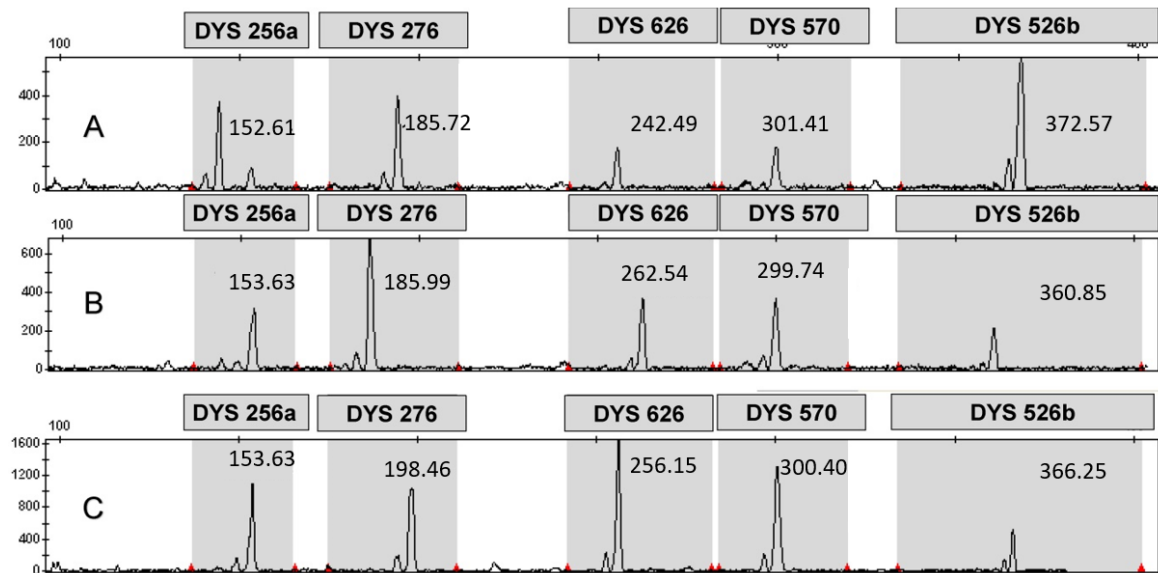
Fig.7.3 Rapid PCR of (A) DNA standard 9948, (B) DNA standard 2800 and (C) extracted DNA from subject K. Run on ABI 310

C.3. Sensitivity study of rapid amplification

Figure 7.4 shows a series of dilutions using the 9948 control DNA. Input levels ranged from 2 to 0.5 ng resulted in useable profiles; however, the 0.25 ng sample showed low threshold peaks (<150 RFU). Nevertheless, all expected peaks were still present even at

150 pg of input DNA. Based on these results, the recommended input level for this procedure is greater than 250 pg male DNA. The results show an average standard deviation of 0.15 bp and 102 RFU for peak size and height, respectively.
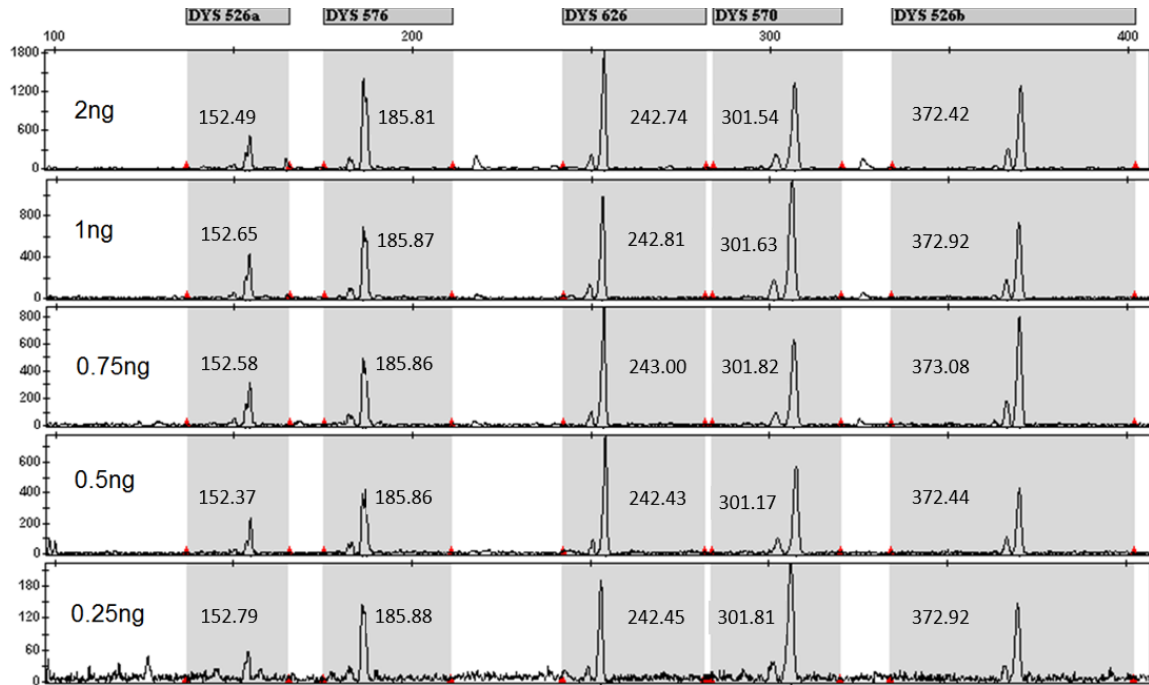
Fig.7.4 Rapid PCR of DNA standard 9948. Input concentrations decreasing from 2 to 0.25 ng. Run on ABI 310

C.4. Mixture of male and female DNA

Profiles of the 9948 male DNA were successfully recovered for all ratios of female-to-male DNA from 0.25:1 up to 40:1. The peak balance remained consistent throughout all ratios tested; however, peak heights decreased slightly when the ratio of female-to-male DNA exceeded 20:1, although mixtures of up to 40:1 female-to-male DNA still produced Y-STR profiles. Figure 7.5 shows the profile generated for the 40:1 ratio of female-to-male DNA. The results show a standard deviation of 0.14 bp and 12 RFU for peak size and height, respectively.
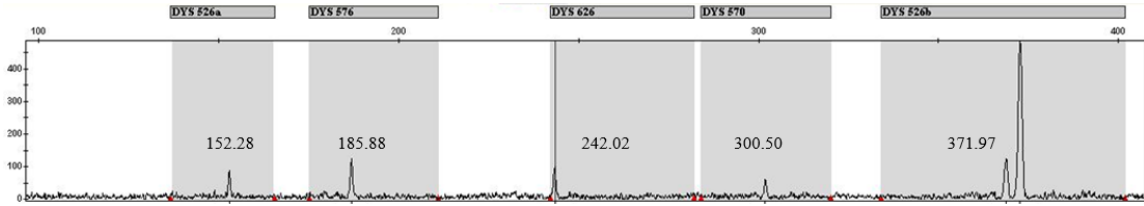
Fig.7.5 A 40:1 ratio of female 9947 DNA-to-male 9948 DNA. Run on ABI 310

C.5. Microfluidic separation

Because each chip may suffer a slight mobility shift as a consequence of temperature fluctuations and buffer composition, sample signals collected at 570 nm were compared to those of a size standard run simultaneously at 670 nm on the same chip. A regression line was created using the peaks from the LIZ dye-labelled size standard. Overall, precision was below 0.21 bp, and resolution varied from 2.4 bp to 4 bp over a peak size range from 100 to 450 bp [Aboud 2013]. Figure 7.6 shows the resulting profile from an extracted buccal sample taken from individual P. This was done in replicate, and these separate runs showed consistent peak patterns, with each peak appearing sharp, although some of the larger peaks showed stutter peaks. The peaks were slightly unbalanced; however, this is thought to be due to the fact that the DYS526 primers bind to two separate chromosomal locations, creating DYS526a and DYS526b amplicons, with DYS526b being consistently higher than DYS526a. The results showed a standard deviation of 0.57 bp and 65 RFU for peak size and height, respectively. Furthermore, the results were comparable in location and quality to those of the CE method, thus showing that the microfluidic chip separation can be used successfully in place of CE. When comparing these results to those created by the ABI

310, we noted some variation in size. Performing a comparison of the average peak sizes between the Agilent Bioanalyzer and the ABI 310 produced a slope of 0.9294 with an intercept of 20.1. The correlation coefficient was 0.97. An average difference of 1.6 bp appeared between the two instrument methods; however, each method in itself was reproducible [Aboud 2013]. This difference was expected, as variations in sizing are a natural consequence of the differences in separation polymers and temperature profiles between the two electrophoresis techniques [Hartzell 2003].
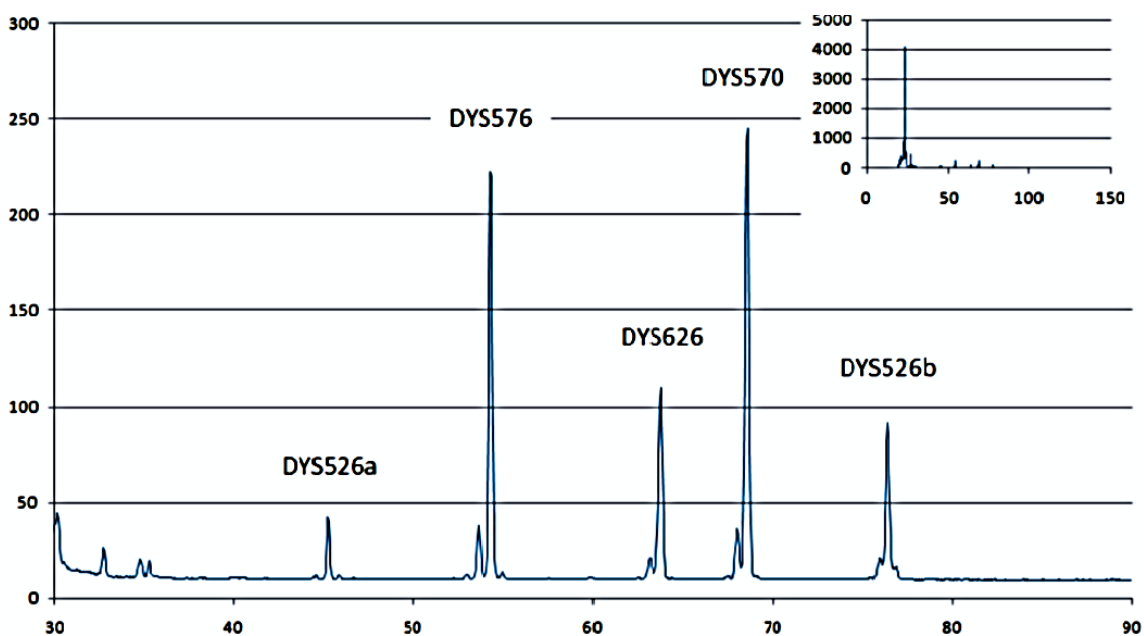
Fig.7.6 Multiplex run with extracted DNA from subject P at an input concentration of 1 ng and run on the Agilent 2100. Top right shows full electropherogram, while full picture shows zoomed-in version to better view peaks. Chip work was completed by Patricia Albani [Albani 2013, Gibson-Daw 2017].

D. Conclusion

In this chapter, we have shown the successful creation of a four locus Y-STR multiplex that can be used to screen crime scene samples for male DNA. We have also demonstrated that the combination of a high-speed thermal cycler (Streck) and a rapid polymerase (Takara) permits the decrease of amplification time for the Y-STR multiplex to less than

17 min. This can be coupled with a microfluidic separation device to further reduce analysis time to less than 25 min per profile after sample extraction. This procedure should prove useful for screening sexual assault casework samples and exclusion/ presumptive identification of male suspects.

CHAPTER VIII. DEVELOPMENT OF ULTRAHIGH SPEED PCR METHOD FOR OBTAINING STR BASED GENOTYPES IN 8 MINUTES OR LESS[*]

A. Introduction

In this chapter process is described for the optimization of multiplex PCR. The goal was to optimize a high speed amplification of forensic short tandem repeats using commercially available instrumentation. While there has been much work on improved extraction methods [Aboud 2013, Qiagen 2014, President's DNA Initiative, Qiagen 2016, ZyGEM 2017] and microfluidic analysis [Horseman 2007], work on improving the speed of multiplex amplifications has been more limited. In 1989 Wittwer described a rapid thermocycler that used glass capillaries that were heated and cooled through the addition and removal of hot air from a chamber [Wittwer 1989]. Since then instrumentation has continued to develop and improve. Rapid hot air and Peltier based thermal cyclers are now commercially available, and faster DNA polymerases have also been produced. Small volume microfluidic systems have also been produced which can be rapidly heated and cooled [Belgrader 1999, Giordano 2001]. This technology includes lab on a chip based systems in which amplification is performed along with separation and detection in a single integrated device [Easley 2006, Liu 2007, Liu 2011] with total times ranging from 20 to 90 minutes [Khandurina 2000, Hopwood 2010, Estes 2012, Lounsbury 2013]. While much research has been done on improving the instrumentation used in PCR amplification, cycling conditions and the use of newly developed faster polymerases [Butss 2014, Romsos 2015], little work has focused on the actual optimization of the amplification chemistry.

---

[*] This chapter has been published: Gibson-Daw G, Crenshaw K, McCord B. Optimization of ultrahigh speed multiplex PCR for forensic analysis. Anal. Bioanal. Chem. 2018;410(1):235-45

130

Many studies involving rapid PCR amplification, especially in the clinical field, have focused on reactions involving single locus amplifications rather than multiplex amplification [Belgrader 1998, Farrer 2015]. For example, Wittwer recently discussed a method for fast amplification where two water baths at two different temperatures were used with a stepper motor to move the sample between the two in <0.1 s. This lead to the successful amplification of a 60 bp fragment with 35 cycles in 15 s [Farrar 2015]. However, this fragment was quite small and not multiplexed.

For forensic purposes, multiplex STR amplifications are needed which can involve up to 23 simultaneous reactions with amplicons that range in size from 60 to 500bp. The complexity involved in balancing peak heights and optimizing a multiplicity of primer binding and extension naturally results in a need to reduce the overall amplification speed in order to optimize sensitivity and peak balance. Some recent work has been performed in an effort to increase amplification speed of commercially available kits [Vallone 2008, Vallone 2009, Foster 2012, Laurin 2012, Romsos 2015]. For example, Vallone et al. examined the effect of polymerase, buffer conditions and cycle time in an effort to increase the speed of commercially available forensic kits [Vallone 2008, Vallone 2009, Romsos 2015]. And while many of these studies have successfully reduced amplification time to ~60 minutes [Budowle 1990, Gibson-Daw 2017], only a few manage to reduce this below 30 minutes [Foster 2012, Laurin 2012, Aboud 2013, Romsos 2015] and only one below 15 minutes [Butts 2014, Romsos 2015]. A recent paper by Butts et al. shows the successful amplification of the Identifiler STR kit using the SpeedSTAR HS Polymerase and the Philisa thermocycler in 14 minutes [Butts 2014]. Previous research in this laboratory involved the development of a rapid direct method which utilized 2mm punches of DNA

on FTA paper to amplify a 7 loci multiplex in 16 minutes using Z-taq and a Philisa thermocycler [Aboud 2013]. While this amplification procedure and the method outlined in the Butts 2014 paper are very fast, they are twice as long as the method we propose in this paper.

In this laboratory, we have been working on the development of rapid amplification of small STR multiplexes for use in rapid screening of individuals. To do this we have examined specially engineered DNA polymerases [Giese 2009, Kermekchiev 2009, Verheij 2012] and buffer systems to rapidly amplify small autosomal and Y STR multiplexes [Aboud 2013, Gibson-Daw 2017]. The autosomal multiplex we have developed includes the STR loci D5S181, D13S317, D7S820, CSF1PO, D16S539, Penta D and Amelogenin with amplicon sizes between 106 and 454 bp. This multiplex set is specifically designed to be run on microfluidic chips as well as larger fluorescent genotyping systems [Aboud 2013]. Here our aim was to utilize experimental optimization procedures to produce a rapid and robust multiplex PCR that would permit high speed DNA amplification in 6.5 minutes or less in a relatively low reaction volume. This result should prove useful for screening of individuals at ports of entry or during mass disasters while still leaving sufficient sample for full amplification at a later time.

B. Methods and Materials

B.1. DNA Samples

DNA standard K562 was obtained from Promega Corporation (Madison, WI, USA) and was used at various dilutions ranging from 100-2 ng/µL in the optimization process as a positive control for the genotyping process. A small population of DNA samples was

obtained from 32 cheek swabs of volunteers using Sterile Cotton Tipped Applicator swabs from Puritan Medical Products (Guildford, ME, USA). These samples were lysed, digested and the DNA extracted and purified using standard phenol chloroform isoamyl alcohol (PCIA) methods [Budowle 1990, Comey 1994] This method of extraction usually yields from 3-50 ng/ μL of DNA but for reasons of developing a rapid method the samples in this study were not quantified after extraction. The collection and analysis of these samples was approved by the Institutional Review Board (IRB) at FIU, under the reference number #101831.

B.2. MP7 Multiplex and Primer Sequences

A custom multiplex (MP7, Table 8.1.) was assembled from 7 Loci chosen from a subset of standard forensic STR loci containing 4 and 5 base repeat motifs. The selected loci included D5S818 (D5), D13S317 (D13), D7S820 (D7), D16S539 (D16), CSF1PO (CSF), and Penta D, as well as the Amelogenin locus. The Amelogenin (AM) locus was labelled with Liz fluorescent dye, while the other six of the loci were labelled with the TAMRA fluorescent dye (IDT). These loci were chosen based on size, primer melting temperature, and low levels sequence variants (improved electrophoretic resolution). The multiplex was designed to be compatible with a two-channel microfluidic chip based system (modified Bioanalyzer by Agilent) as well as the Life Technologies capillary based DNA sequencers. Table 8.1 shows the details of the loci including repeat motif, Primer sequences, alleles and size range, location on the chromosome and Genebank Accession number.

Table 8.1: Locus Information					
Loci	**Repeat Motif**	**Primer Sequence 5'-3'**	**Allele Range**	**Chromosome Position**	**Genebank Acc No.**
D5S818	AGAT	F GGTGATTTTCCTCTTTGGTATCC R AGCCACAGTTTACAACATTTGTATCT	6 (115bp) 18 (163bp)	5q23.2	G08446
D13S317	TATC	F ATTACAGAAGTCTGGGATGTGGAGGA R GGCAGCCCAAAAAGACAGA	5 (157bp) 17 (205bp)	13q31.1	G09017
D7S820	GATA	F ATGTTGGTCAGGCTGACTATG R GATTCCACATTTATCCTCATTGAC	5 (211bp) 15 (251bp)	7q21.11	G08616
D16S539	GATA	F GGGGGTCTAAGAGCTTGTAAAAAG R GTTTGTGTGTGCATCTGTAAGCATGTATC	4 (260bp) 16 (308bp)	16q24.1	G07925
CSF1PO	AGAT	F CCGGAGGTAAAGGTGTCTTAAAGT R ATTTCCTGTGTCAGACCCTGTT	5 (317bp) 16 (361bp)	5q33.1	X14720
Penta D	AAAGA	F GAAGGTCGAAGCTGAAGTG R ATTAGAATTCTTTAATCTGGACACAAG	1.1 (370bp) 18 (454bp)	21q22.3	AP001752
Amel	NA	F CCCTGGGCTCTGTAAAGAATAGTG R ATCAGAGCTTAAACTGGGAAGCTG	X – 106bp Y – 112bp	Xp22.1-Xp22.3 Yp 11.2	M55418 / M55419

Table 8.1 Information on the loci that make up the MP7 multiplex, including locus name, repeat motif, primer sequence, product size range, allele range, chromosome position and gene bank Acc number.

B.3. Thermal Cycling and Rapid Polymerase

The thermal cycler used for all experimental amplifications was the Philisa from Streck (Nebraska, US). Which with its thin walled tubes achieved heating rates of 15ºC/s and cooling rates of 12ºC/s allows for fast PCR to occur.

For this study Z-Taq polymerase (Takara Bio, Clontech Laboratories, Inc., Mountain View, CA), supplied with 10 X Z-Taq buffer and 2.5mM dNTP mixture, was used which had previously been shown in studies by Aboud [Aboud 2013] to work well for achieving

ultrafast PCR amplifications due to its high processivity (over 5X speed of regular Taq which has a processivity of 60 nt/s) and high fidelity [Wang 2004, Kermekchiev 2009, de Vega 2010].

B. 4. PCR Optimization

B.4.1 Optimization of Cycling Conditions

Initial amplification conditions were selected based on the 16 min amplification of a Y STR multiplex discussed in chapter 7, in which the amplification occurred in a final PCR master mix volume of 9 µL, to which 1 ng of 2800 male control DNA (Promega) was added. The master mix was made up of 2 X Z-Taq buffer (Takara Bio), 0.2 mM each dNTPs (Takara Bio), TAMRA fluorescently labelled primers at concentrations of 0.20 µM D5S818, 0.20 µM D13S317, 0.36 µM D7S820, 0.24 µM D16S539, 0.21 µM CSF1PO, 0.55 µM Penta D and 0.24 µM Amelogenin. (IDT), 0.0625 Units/10µL of Z-Taq polymerase (Takara Bio). Cycling conditions were denaturation at 95°C for 5 s, followed by 29 cycles of 98°C for 5 s, 62°C for 9 s, 72°C for 12 s. Allele calling threshold was set at 100 RFUs, so any peak visible above this threshold was considered to be a true peak separate from the noise and thus called as an allele for the loci it presented at. All experiments were tested in replicates in order to determine reproducibility and render results valid.

These conditions were used in this study as a starting point for the optimization of master mix reagent concentrations and cycling conditions for increased speed and robustness. 100ng of input DNA was used for this series of experiments in order to increase the chance of success with lower cycle numbers and shorter annealing/elongation times. Once

optimization was complete, the initial amount of input DNA was reduced. Initial work was performed to speed up the reaction by optimizing the cycling conditions, Table 8.2. One of the first steps was the elimination of the initial activation step changing the 4-step PCR to a 3- step PCR. Next, a two-step thermal cycling process was initiated by combining annealing and elongation into a single step at $62^{O}C$. Additionally, the overall time of both this combined step and the denaturation step were reduced, as was as the total number of cycles (A and B). Longer times and higher cycle numbers, more similar to those seen in previously, were also used (C and D) to act as a control for use in comparing the effect of allele size and peak height when amplified at faster conditions. Throughout the optimization process, the denaturing and annealing times and cycle numbers were decreased both individually and in a variety of combinations to achieve the fastest and most robust amplifications. Cycling conditions were optimized based on 3 parameters, total reaction time, peak height and peak balance.

Also tested were certain methods in which two different types of thermal cycling reactions were used in the PCR reaction as a form of touchdown PCR, in which the initial 2-3 cycles were longer than the remaining ones. This was expected to aid in achieving more intense peak heights and decreased non-specific amplification but in the end, was deemed unnecessary. The conditions shown as Y were selected as optimal for use in subsequent experiments.

Table 8.2 Cycling Conditions Optimization					
Cycling Condition	Denaturing (s)	Annealing (s)	Cycles	Total Time	
Set	(98ºC)	(62ºC)		Minutes	Seconds
A	3	10	25	8	19
B	3	5	29	7	0
C	5	9	29	10	5
D	5	12	29	11	21
E	3	9	25	7	49
F	3	8	25	7	15
G	2	10	25	7	49
H	3	10	24	7	45
I	3	7	25	7	0
J	3	7	29	7	56
K	2	9	23	6	29
L	3	7	24	6	37
M	2	8	25	7	1
N	1	10	25	7	15
O	2	6	24	5	54
P	2	5	25	5	35
Q	2	10	22	6	27
R	2	6	25	6	0
S	2	8	21	5	55
T*	3	10	3		

	2	6	21	6	30
U*	3	10	3		
	2	8	21	6	53
V*	3	10	2		
	2	6	25	6	39
W*	3	10	3		
	2	8	23	7	26
X	2	8	24	6	33
Y	2	7	24	6	21
Z	2	8	23	6	19
* two separate cycling conditions					

Table 8.2 Cycling conditions, showing time at both annealing and denaturing, cycle number and total time the amplification took to complete. Conditions marked with * were those which changed conditions after a certain number of cycles as detailed in the table.

B.4.2 Optimization of Master Mix

After optimization of the cycling conditions to get the fastest amplification time possible it was necessary to optimize the concentrations of the different reagents making up the master mix in order to improve peak height, quality and balance. The concentrations tested and the experimental order designed by the software can be seen in (Table 8.3). These values were generated by a software program called Design Expert v.10.0.4.0 and the order is randomized by the software for statistical purposes. To set up these parameters, a maximum and minimum concentration was inputted for Z-Taq, primer mix, buffer and dNTPs (2 µL of DNA was used for every reaction and water was adjusted accordingly to

reach a total volume of 10 µL). The basic study type selected was for a reagent mixture, and the module chosen was characterization.

Master Mix	2.5 U/ µL Z-Taq (µL)	Primer Mix (µL)	10X Z-Taq Buffer (µL)	2.5 mM dNTPs (µL)	25 mM MgCl$_2$ (µL)	25 ng DNA (µL)	Water (µL)
A	1.24	2.00	1.06	2.69	1.00	2.00	0.00
B	0.30	2.18	1.00	3.00	1.28	2.00	0.25
C	0.81	2.59	1.70	1.82	1.08	2.00	0.00
D	0.30	2.81	1.00	0.89	3.00	2.00	0.00
E	0.44	3.55	1.00	2.01	1.00	2.00	0.00
F	0.94	2.00	1.26	0.80	3.00	2.00	0.00
G	0.32	2.79	1.00	2.09	1.00	2.00	0.80
H	0.30	2.00	1.00	0.80	1.00	2.00	2.90
I	0.30	2.00	2.58	2.12	1.00	2.00	0.00
J	0.56	2.00	1.00	2.01	2.43	2.00	0.00
K	0.30	2.00	1.00	0.80	1.00	2.00	2.90
L	1.50	2.03	1.74	0.83	1.00	2.00	0.90
M	0.30	2.04	2.61	0.95	2.11	2.00	0.00
N	0.81	3.51	1.00	0.80	1.88	2.00	0.00
O	0.30	2.00	3.90	0.80	1.00	2.00	0.00
P	0.30	2.04	2.61	0.95	2.11	2.00	0.00
Q	0.41	2.00	1.00	0.80	2.31	2.00	1.48
R	0.56	2.00	1.00	2.01	2.43	2.00	0.00
S	1.50	2.00	1.00	1.22	1.97	2.00	0.31
T	0.30	2.69	3.07	0.80	1.14	2.00	0.00

Table 8.3 Master Mix Optimization. Total volume 10 µL

U	0.81	3.51	1.00	0.80	1.88	2.00	0.00
V	0.30	3.48	1.00	0.80	1.00	2.00	1.42
W	0.49	2.00	1.00	1.95	1.00	2.00	1.55
X	0.30	4.90	1.00	0.80	1.00	2.00	0.00
Y	0.42	3.45	2.33	0.80	1.00	2.00	0.00
Z	1.50	3.70	1.00	0.80	1.00	2.00	0.00
AA	0.39	2.70	1.69	0.86	1.61	2.00	0.74
AB	1.11	2.07	1.00	0.80	1.00	2.00	2.02
AC	0.49	2.00	1.00	1.95	1.00	2.00	1.55
AD	1.22	2.00	2.78	0.92	1.09	2.00	0.00
AE	0.36	2.00	2.44	0.80	1.00	2.00	1.40

Table 8.3 Master mix optimization, the output of the experimental software Design Expert 10. The order is randomized for statistical purposes. Table 8.3 Shows amounts (µL) of each reagent used and their starting concentration, along with how much DNA and water was added to each reaction to keep the total volume at 10 µL and the amount of input DNA constant.

B.4.3 Optimization Total Volume

The next step was to observe the effects of reducing the overall volume of the PCR mix in an attempt to increase observable peak heights as well as reducing the quantity of input DNA. 1 µL of 100 ng, 75 ng (Figure 8.4), 50 ng and 25 ng DNA were each tested within a total reaction volume of 8 µL, 6 µL and 4 µL. All samples were tested in triplicate.

B.5 Ultra High Speed PCR Amplification

The final conditions for the optimized process used in all subsequent studies in this chapter were: 3 µL of Master Mix X with1 µL of 50 ng DNA (K562, Promega) to give a total volume of 4 µL. This master mix included 1 µL of 10 X Z-Taq buffer (Takara Bio), 0.8 µL

of 2.5 mM each dNTPs (Takara Bio), 4.9 µL Primer mix made up of 10 µM TAMRA fluorescently labelled primers: 3.25 µL D5S818, 3.25 µL D13S317, 6 µL D7S820, 4 µL D16S539, 3.5 µL CSF1PO, 9 µL Penta D and 4 µL Amelogenin. (IDT), 1 µL of 25 mM $MgCl_2$ and 0.48 µL of 2.5 U/µL of Z-Taq polymerase (Takara Bio). Total concentrations for a 10 µL reaction were: 2.5 X Z-Taq buffer, 0.5 mM of each dNTP primer mix (0.24 µM D5S818, 0.24 µM D13S317, 0.44 µM D7S820, 0.30 µM D16S539, 0.26 µM CSF1PO, 0.66 µM Penta D and 0.30 µM Amelogenin), 6.25mM $MgCl_2$ 0.3 Units/4 µL of Z-Taq polymerase. Amplifications were performed on a Philisa thermocycler (Streck) in just under 6.5 min. Optimal conditions were 24 cycles of heat denaturation at 98°C for 2 s, followed by a combined annealing elongation step of 62°C for 7 s (Figure 8.5).

B.6. Sensitivity Study

To test the sensitivity of the system a range of input DNA concentrations (K562 standard DNA) from 100 ng to 0.5 ng were amplified and analysed using the ABI 310 for separation and detection (Figure 8.6). 3 replicates of each experiment were run in the same manner as in the previous section.

B.7. Cheek Swab Samples

Cheek swabs from 32 individuals were extracted using PCIA and of these samples 1 µL was used for PCR with the concentration unchanged from that seen after extraction. These were then amplified using the optimized method detailed in section B.5. and run on the ABI 310 as well as the Agilent Bioanalyzer as outlined in section B.8. (Figure 8.7-8.8)

B.8. Separation, Detection and Analysis

B.8.1 Applied Biosystems ABI 310 Capillary Electrophoresis Genotyper

All PCR products were analysed using an ABI 310 (Applied Biosystems) by adding 1 µL of PCR product to a mixture of 12.5 µL HiDiTM Formamide (Applied Biosystems) and 0.5 µL internal lane standard (ILS). The ILS used throughout this paper was LIZ 500 (Applied Biosystems) which was labelled with the orange LIZ dye. This mix was then run on the ABI 310 using the following conditions: POP4 polymer, Capillary length to detector: 36 cm, module GS STR POP4 (1 mL), injection voltage 15 kV, injection time 5 s, run time 28 min, run temperature $60^{O}C$, run voltage 15 kV. Genemapper software was used for sample analysis.

B.8.2 Agilent Technologies Microfluidic Chip Bioanalyzer 2100[*]

The samples from section B.7. were also run on a modified Agilent 2100 Bioanalyzer that contained a heat plate and dual detector for high resolution separations. DNA Lab Chips (Agilent Technologies) were prepared for microfluidic electrophoresis of DNA samples using the High-Resolution DNA Gel Matrix (Agilent Technologies) as the polymer. The loading of the polymer into the chip capillaries was accomplished using a Chip Priming Station (Agilent Technologies). Sample wells were loaded with 5.5 µL of a mixture of formamide: ILS at a 9:1 ratio. Approximately 1 to 1.5 µL of sample was added to each sample well. Every chip also contained one well (Well 12) designated for ladder, which

---

[*] Work completed on Bioanalyser 2100 and data presented was a collaboration between the author and Karin Crenshaw a PhD student of the McCord DNA lab.

included only CC5 ILS 500 (Promega). Chip run times of approximately 2 minutes per sample were obtained. The data was generated in real time and each sample was run sequentially. The ILS in each well is then used to analyse the data and generate allele sizing using the Agilent 2100 Expert software (see Figure 8.8.A).

C. Results and Discussion

The 7 locus multiplex MP7 was used in the development of the ultrafast PCR protocol, along with a rapid polymerase Z-Taq (Takara) and the Philisa rapid thermocycler (Streck). The goal of the optimization was to increase the amplification speed, while maintaining high peak intensities and balance between alleles and loci. Optimization of the cycling conditions, master mix reagent concentrations and total volume were all carried out independently. These tests were performed using the K562 DNA standard (Promega), using 100 ng, 50 ng and 75 ng of input DNA. This amount of DNA was used initially for optimization to better visualize any variances in parameters tested, once these were optimized the concentration of input DNA was reduced to 2ng as can be seen in section 3.2. DNA Once these steps were completed all optimized conditions were combined and a sensitivity study was performed. The final method was then tested with samples collected from the cheek swabs of 30 adult male and female volunteers. Good profiles were those showing all alleles with peak intensities of >100 RFUs and peak balances of >60%. Results were achieved in replicates and determined to be reproducible when peak size showed no more than a 0.3bp change. This procedure resulted in a full profile being achieved in 6 mins 21 s. with 100% allele recovery and an average peak balance of ~70%. The

calculations for peak to height ratio (PHR) were based on recommendations in the SWGDAM guidelines [SWGDAM 2010].

C.1 PCR Optimization

C.1.1 Optimization of Cycling Conditions

Optimization of the cycling parameters can be seen in Figure 8.1. In the figure, each set of cycling parameters are shown with the resulting peak height for each allele for the loci under study. The results are organized based on the total time of the amplification and overall allele peak height. Total run times range from 11.5-5.5 mins. As can be observed from the data, as the total cycling time decreases, there is a concomitant reduction in peak height. Cycling times below 6 minutes produced noticeable allele dropout. The upper right-hand side of the figure provides information on the peak heights and balance produced for cycling times from 6-6.5 minutes. These conditions were selected for further optimization.
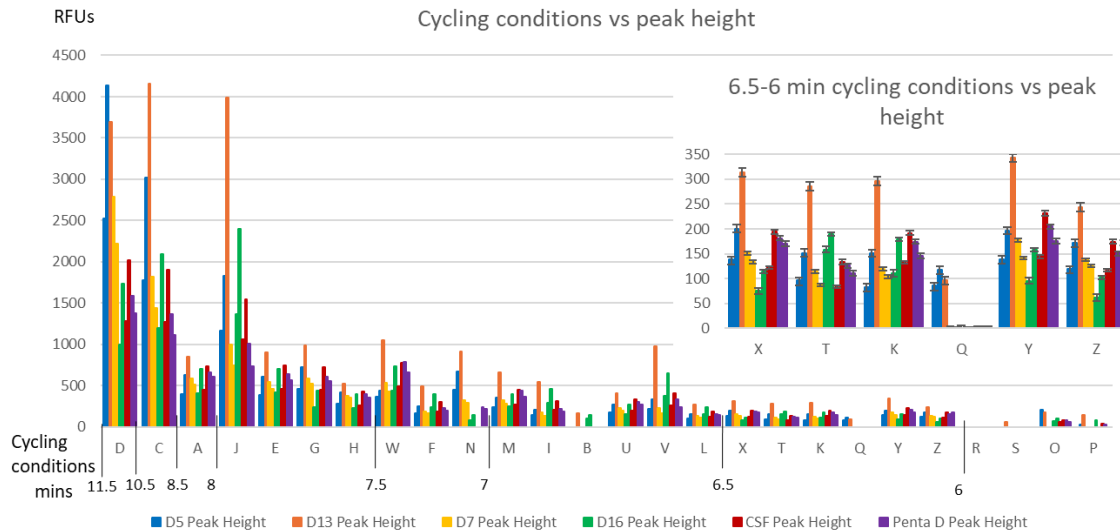
Fig.8.1 Cycling conditions vs. peak height. This figure shows each set of cycling conditions plotted against the peak height of each allele in RFUs and along the bottom time brackets have been added to show the relative times of the conditions. In the top right of the figure is an insert containing an expanded view of the conditions which permitted a 6- 6.5 min amplification. These conditions also include error bars representing the standard deviation of the process (n=3). It is clear from this image that amplification conditions below 6 mins result in an increase in allelic dropout. Samples analysed using ABI 310.

The goal for optimization was to obtain the largest peak height and best peak balance between loci in the fastest time. As can be seen in the enlarged section in the top right of Figure 8.1, with the exception of condition Q, all of these cycling conditions gave similar results. However, on close examination conditions X, Y and Z consistently gave higher, more reproducible peak heights, and condition Y was also approximately 10 s faster than X with peaks that were consistently higher than those from conditions Z. For these reasons the cycling conditions represented by condition Y were chosen as the optimal cycling conditions. Other cycling conditions (R, S, O and P) were a minute or so faster but at these speeds allele dropout or high variations in peak heights were observed. The cycling conditions for Y included a denaturing step at 98ºC for 2 seconds, and an annealing step at 62ºC for 7 seconds. These steps were repeated for a total of 24 cycles. This gave the total

amplification time of 6 minutes and 21 seconds. This cycling protocol was used for all subsequent experiments. The resulting electrophoretic profile can be seen in Figure 8.2.
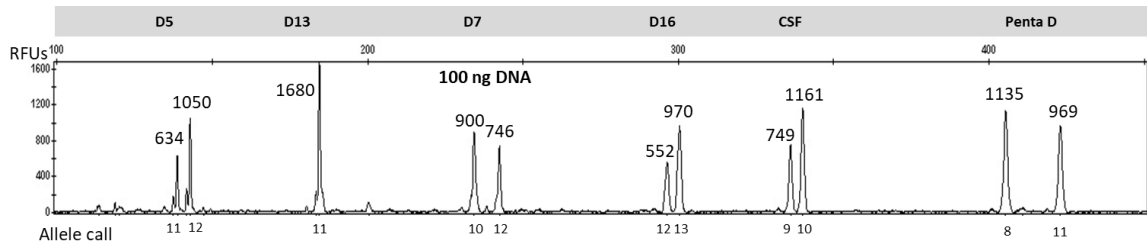
Fig.8.2 Profile resulting from the determined optimal conditions labelled Y in table 2: denaturing at 98°C for 2 seconds, annealing at 62°C for 7s, 24 cycles. This gave the total amplification time of 6 minutes and 21 seconds. Input DNA was 100ng of K562 control DNA standard. Overall peak balance was 70.0%.

C.1.2 Optimization of Master Mix

The next step in the optimization process was to examine the composition of the master mix. A variety of different concentrations were examined with a focus on the primer mixture, dNTP concentration and $MgCl_2$ concentrations, as previous research demonstrated that these conditions were key factors in improving the quality of amplification [de Vega 2010, Aboud 2013, Gibson-Daw 2017].

The resulting peak heights for the alleles of each loci were plotted against their respective master mix conditions and can be seen in Figure 8.3. These results show that the master mixes with the highest peak heights and best peak ratios were obtained using conditions D and X. Figure 8.3, top right inset, shows a closer view of how master mixes D and X compare to each other. From this comparison, it is clear that master mix X produces higher overall peak intensities and a better peak to peak balance. For this reason, the master mix concentrations represented by X were determined to be optimal and chosen for subsequent experiments. In these experiments, the factors having the most pronounced effect were the

volume of the primer mixture and the concentration of the dNTPs. A higher concentration of primers was necessary to successfully amplify genomic DNA at these high speeds, Table 8.3. The fact that relatively high concentrations of dNTPs improve the quality of the results may be the result of local depletion due to diffusive effects at lower dNTP concentration. In general, when using these high amounts of both primer and dNTPs, it is also important to increase the $MgCl_2$ to make sure the reaction stays activated.
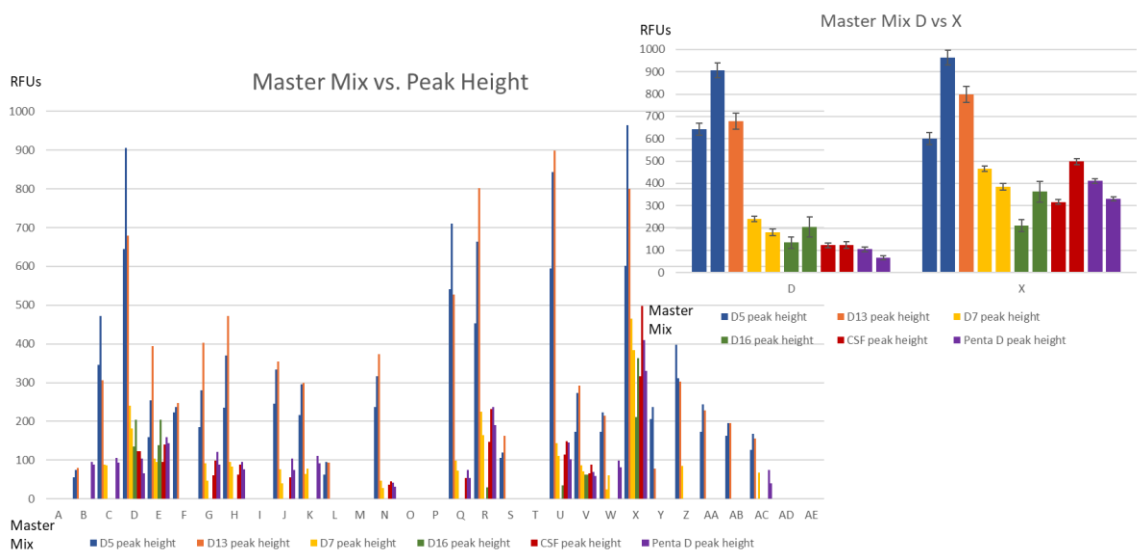
Fig.8.3 Master mix reagent concentrations for 10 µL final volume vs. peak height in RFUs. Showing which sets of conditions show allele dropout and which ones do not as well as which conditions give higher peak intensities and more equal peak balance between alleles. Top right: Close up comparison of allele peak heights of the two best performing master mix combinations D and X. Error bars show the standard deviation for each allele. Samples analysed using ABI 310.

C.1.3 Optimization of Total Volume

Once the cycling conditions and master mix were optimized, further experiments were performed to reduce the total volume of the PCR master mix. The goal of this study was to reduce the total amount of input DNA while maintaining peak height and sensitivity. Figure 4 shows the results of these tests for 75 ng of input DNA. Amelogenin while not shown here was present in every sample. The reduction of total volume from 8 to 6 µL produces

147

an average 157 % increase in peak height across the loci. When reduced from 6 to 4 µL little change in peak height was observed (an average 106 % increase in peak height across the loci). Each profile had an average peak balance of 73.0%, 66% and 72% respectively. The 6 µL samples showed peak balances from 37% - 93%, while both the 8 µL and 4 µL samples showed peak balances between 57% and 89% and the standard deviation was ±5, ±8, and ±12 respectively for the 8, 6 and 4 µL volumes. The 6 µL protocol was at this point dismissed from further consideration as some of the peak balances were below 55% which was undesirable. 4 µL was then chosen as the optimal total volume as it used less reagents and gave an increased peak height.
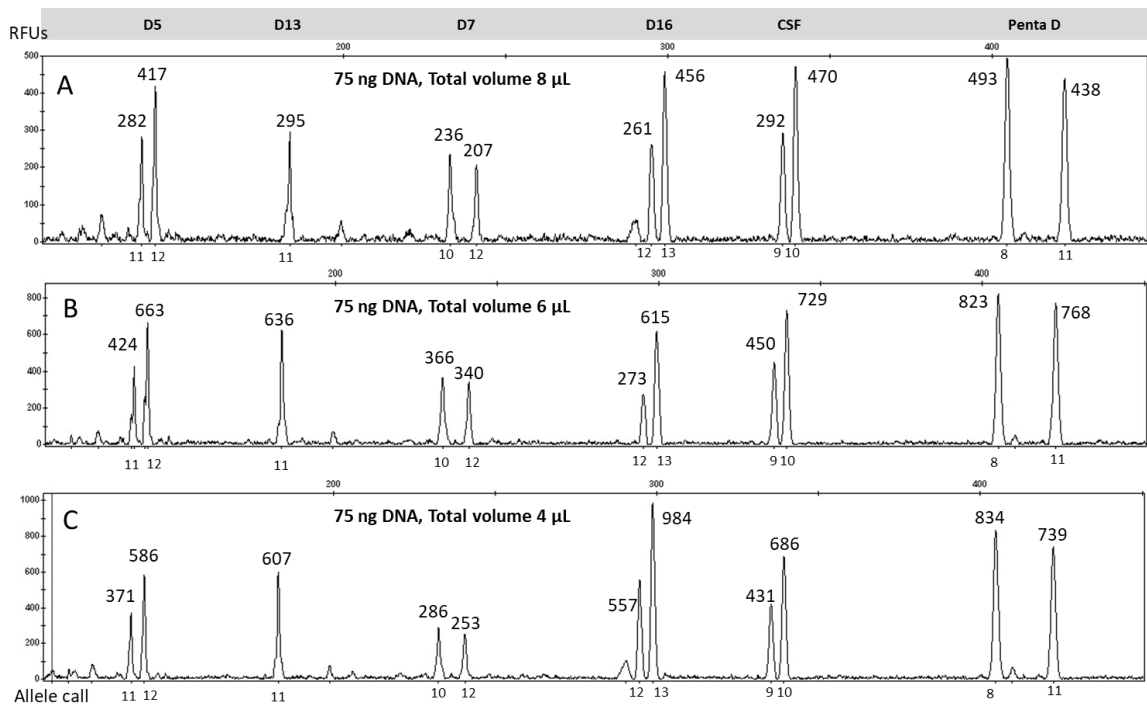
Fig.8.4 75ngs of K562 control DNA standard at differing total volumes: 8 µL, 6 µL and 4 µL. Each profile had an average peak balance of 73.0%, 66% and 72% respectively. The 6 µL samples showed peak balances from 37% - 93%, while both the 8 µL and 4 µL samples showed peak balances between 57% and 89%.

C.2 Ultra High Speed PCR Amplification

In the last part of the optimization process we combined all of the previously optimized parameters into the final method and used with 50 µL of DNA, Figure 8.5 shows the resulting profile. The top lane shows the 6 STR loci labelled with TAMRA dye, while the bottom lane shows the orange lane with the ILS Liz 500 and the Amelogenin locus labelled with Liz.
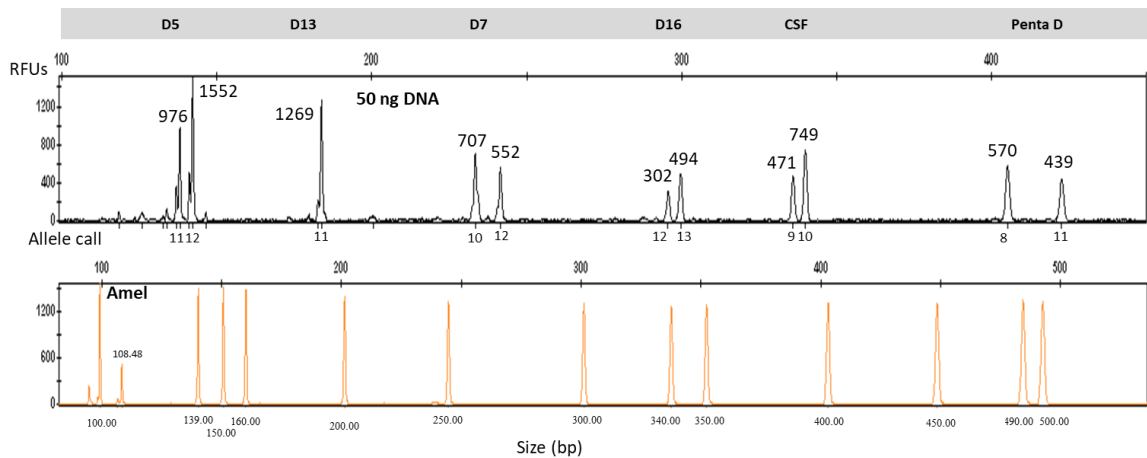
Fig.8.5 Profile generated from 50 ng of the K562 control DNA standard using the master mix concentrations X from Table 3. The top lane shows the 6 STR loci labelled with TAMRA dye, while the bottom lane shows the orange lane with the ILS Liz 500 and the Amelogenin locus labelled with Liz. Overall peak balance was 68.4%. Samples analysed using ABI 310.

C.3 Sensitivity Study

Once the method was developed and optimized we tested the sensitivity of the system using a series of dilutions using the K562 control DNA, Figure 8.6. Input levels ranging from 100 ng-2 ng gave useable profiles however samples below 2 ng of input DNA showed low threshold peaks indistinguishable from the noise, for this as with all previous experiments the calling threshold was set at 100 RFUs, Amelogenin while not shown here was successfully amplified in every sample. As expected peak height was observed to decrease

with decrease of input DNA and thus, the recommended minimum input level for this procedure is >2 ng of DNA.
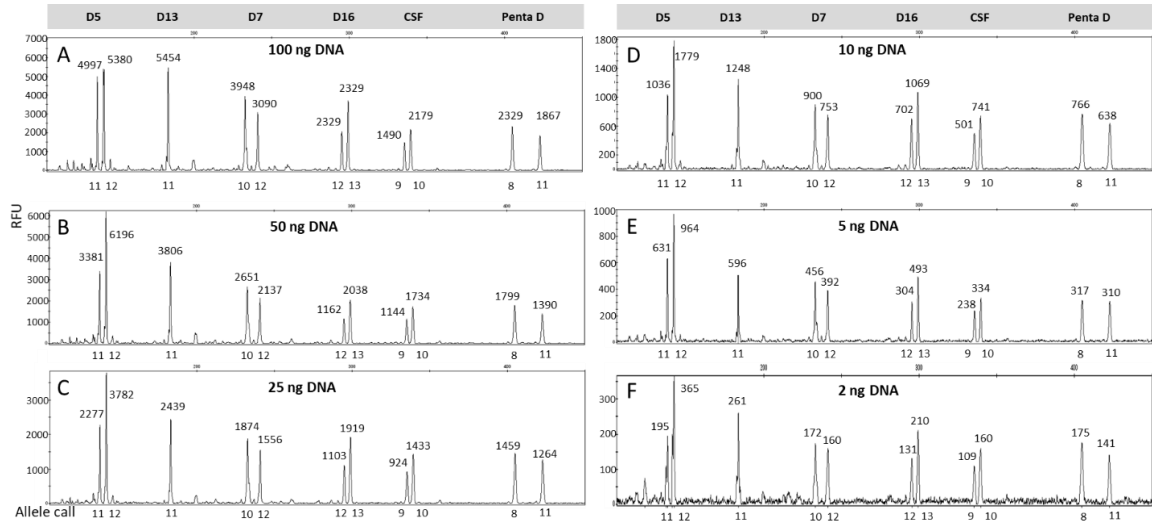
Fig.8.6 Profiles obtained from optimized method as the amount of input DNA (K562 control standard) was reduced from 100 ng to 2ng. Overall peak balances for the respective profiles are as follow: A) 76.0%, B) 67.1%, C) 70.4%, D) 71.7%, E) 76.8%, F) 71.5%. Scale in RFUs can be seen on the left of each electropherogram and is set to the largest peak height in order to improve visual determination of peak height and threshold. Samples analysed using ABI 310.

C.4 Cheek Swab Samples

A study population of 30 volunteers was used to test the applicability of the developed method for the use with cheek swab samples. Figure 8.7 Shows 3 of the profiles generated and as can be observed from these profiles all loci and alleles are clearly visible and discernible from the noise. The profiles on the left shows the 6 STR loci labelled with TAMRA dye, while those on the right shows the respective orange lanes with the ILS Liz 500 and the Amelogenin locus labelled with Liz. The peak heights differed from sample to sample as each cheek swab yielded slightly different amounts of DNA (using PCIA in our lab the yield range achieved was 20-50 ng/µL) and these samples were not quantified as the amount needed for successful amplification (2 ng) is less than a 10th of the total amount

150

recovered, leaving a plethora of additional DNA for downstream genotyping and quantification. These results demonstrate that the method here developed can be used successfully with both control DNA and extracted DNA from cheek swabs.
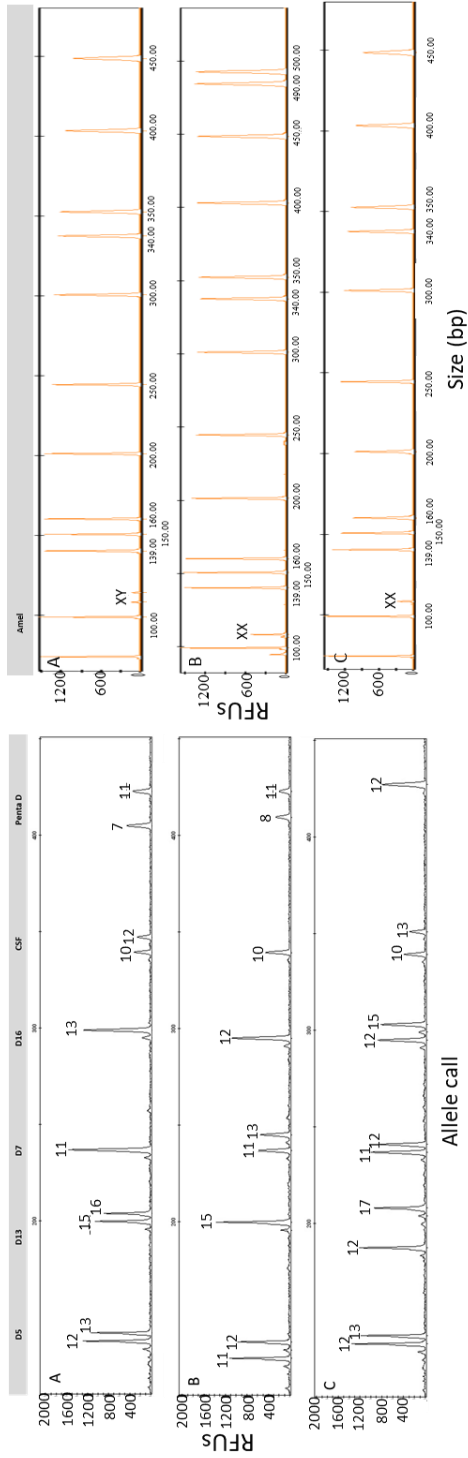
Fig.8.7 Profiles from 3 extracted DNA samples using the optimized method. The profiles on the left shows the 6 STR loci labelled with TAMRA dye, while those on the right shows the respective orange lanes with the ILS Liz 500 and the Amelogenin locus labelled with Liz. Overall peak balance for these profiles was 83.6%. Samples analysed using ABI 310.

C.5 Separation detection and Analysis (Bioanalyzer 2100 vs ABI 310)

Nineteen of the thirty cheek swabs were also examined using microfluidic electrophoresis on an Agilent 2100 Bioanalyzer, which had been modified with a heat plate to permit multichannel denaturing electrophoresis and detection with an excitation wavelength of 532nm, and detection at 575 and 670 nm. [Aboud 2015]. The resulting microfluidic profiles were next compared to the profiles obtained from the ABI 310 CE. Overall peak height ratios and balance for heterozygous loci were compared on each system as shown in Figure 8.8, resulting in an overall profile balance ratio of 82.4% +/- 6.6% (standard deviation) for the ABI310 while the chip system had an overall profile balance ratio of 84.6 +/- 7.4. In addition, all allele calls were concordant between the two systems. This data demonstrates shows that either system may be used to obtain accurate allele calls from the PCR products.
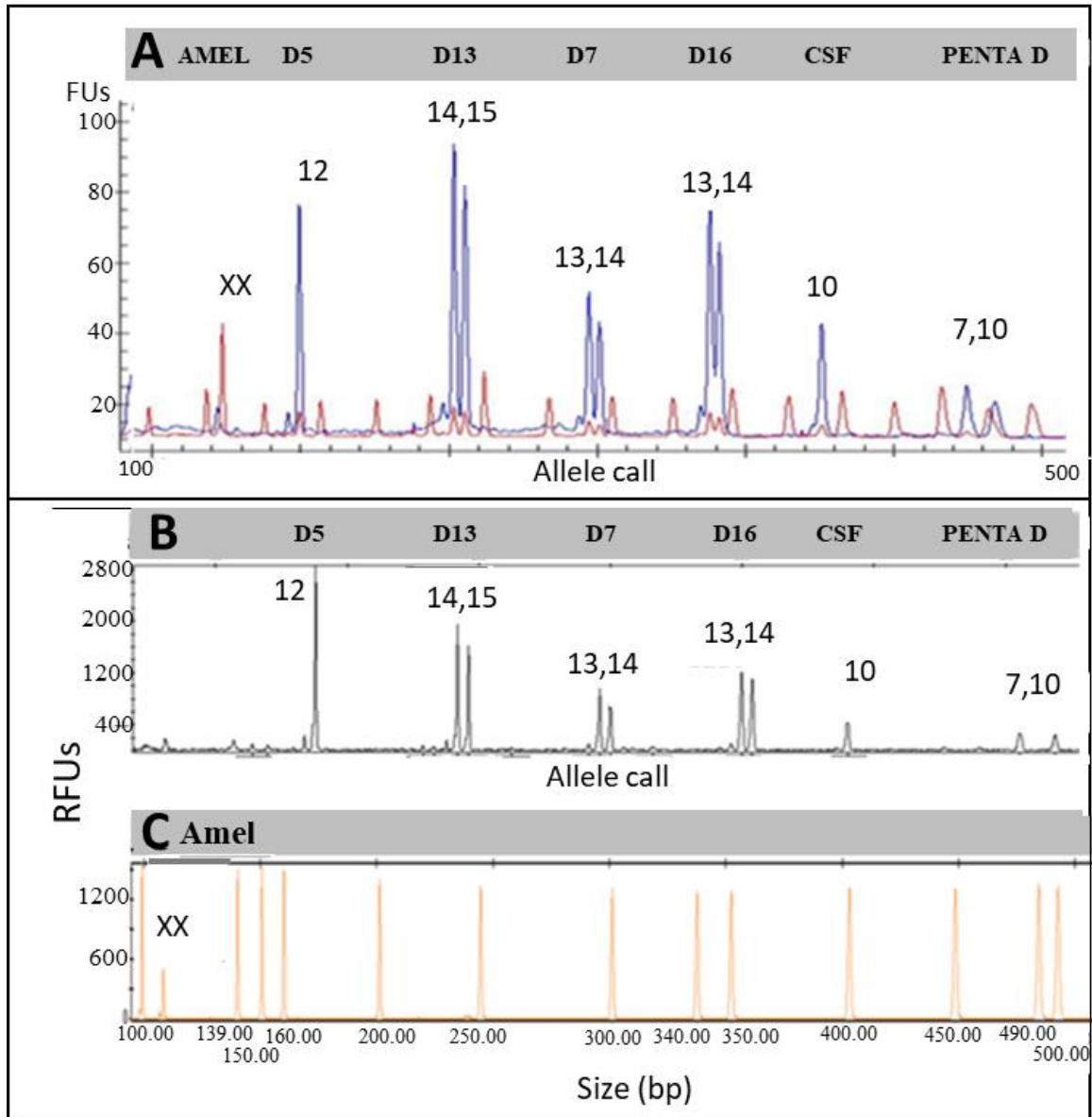
Fig.8.8 A. Electropherogram results from a microfluidic chip analysis run on a modified 2100 Bioanalyzer using denaturing electrophoresis and a two channel detector The 6 STR loci peaks (labelled with TAMRA) can be seen in the blue channel, while the ILS (CC5 ILS 500) and Amelogenin locus (labelled with Liz) can be seen in the red channel. B. Electropherogram results from the ABI 310 (capillary electrophoresis) analysis of the same sample. C. Liz channel with Amelogenin locus from the 310 analyses. Both systems successfully called the same alleles from the cheek swab D.

This 7 locus multiplex has a power of discrimination of 1 X 10^6 or greater, which is

sufficient to identify a potential suspect, once identified a full laboratory analysis would be

required using standard genotyping methods to confirm a match.

D. Conclusions

In this chapter, has been demonstrated the ultrafast amplification of a 7 Loci multiplex in 6.5 minutes. This was achieved through the use of a high-speed thermocycler and rapid polymerase and involved the optimization of cycling conditions, reagent concentrations and sample volume. Once developed, this method was tested with DNA samples extracted from cheek swabs of 30 volunteers. The sensitivity of the method was also tested and showed that complete profiles could be generated using DNA quantities as low as 2ng. Allele recovery for the final optimized protocol outlined at the end of section 2.5 was found to be 100% while peak balance across the profiles was determined to be 70%. The range of PHR observed was 65%-85%, however, most profiles showed a PHR of around 70%. The PHR calculations are based on recommendations in the SWGDAM guidelines. Concordance studies between microfluidic chip and CE separation showed identical allele calls for all samples thus proving the rapid detection and separation using the chip useful for further reducing analysis time. Overall, this method provides an ultrafast amplification that can speed up the processing of forensic cheek swabs while still leaving more than sufficient sample for later analysis and confirmation.

CHAPTER IX. DEVELOPMENT OF A RAPID DIRECT PCR

A. Introduction

In this project, we are working on the development of rapid direct amplification of small STR multiplexes for use in rapid screening of individuals. The term rapid amplification means that no sample extraction is performed prior to the PCR reaction. To create this method, we examined specially engineered DNA polymerases and buffer systems to rapidly amplify a small autosomal STR multiplex. The autosomal multiplex used includes specifically designed primers for amplicon sizes between 106 and 454 bp. This multiplex set was specifically designed [Aboud 2012] to be run on microfluidic chips as well as larger fluorescent genotyping systems. It is the goal of this project to utilize experimental optimization procedures to produce a direct rapid and robust multiplex PCR amplification that would permit high speed sample processing in 15 minutes or less. The thermal cycler used for all experimental amplifications was the Philisa from Streck (Nebraska, US). The polymerases under consideration are Z-Taq and Omnitaq. The use of these along with buffer optimization and addition of enhancers such as the PEC-1 (PCR enhancer cocktail 1) will allow the rapid amplification of DNA straight from saliva with no extraction steps being necessary.

Here our aim was to utilize experimental optimization procedures to produce a rapid direct method that would permit the removal of an extraction step and high speed DNA amplification in 15 minutes or less. This method could be implemented in many situations where screening individuals or samples rapidly may be important, such as: ports of entry,

mass disasters, missing persons, etc. while still leaving sufficient sample for full amplification at a later time.

B. Methods and Materials

B.1. DNA Samples

DNA standard K562 (2-10 ng/µL) was obtained from Promega Corporation (Madison, WI, USA) and used as a positive control for the genotyping process. DNA samples for testing the direct method were obtained by either cheek swabs transferring sample to FTA paper or by directly collecting saliva in a tube. The saliva samples were centrifuged at 13000 RPM for 1 min, the supernatant then removed and 40 µL of amplification grade H_2O added then vortexed to resolubilize the DNA in the liquid. The newly washed saliva was then diluted to 10%. At this dilution the sample still contributes a high enough concentration of template DNA while many of the possible inhibitors have been diluted to reduce their effect on the PCR. The collection and analysis of these samples was approved by the Institutional Review Board (IRB) at FIU, under the reference number #101831.

B.2. MP7 Multiplex

A custom multiplex (MP7, Table 8.1.) was assembled from 7 Loci chosen from a subset of standard forensic STR loci containing 4 and 5 base repeat motifs. The selected loci included 6 STRs: D5S818 (D5), D13S317 (D13), D7S820 (D7), D16S539 (D16), CSF1PO (CSF), and Penta D, as well as the Amelogenin locus. The Amelogenin (AM) locus was labelled with Liz fluorescent dye, while the other six of the loci were labelled with the TAMRA fluorescent dye (IDT). These loci were chosen based on size, primer melting

temperature, and low levels sequence variants (improved electrophoretic resolution) [Aboud 2012]. The multiplex was designed to be compatible with a two-channel microfluidic chip based system (modified Bioanalyzer by Agilent) as well as the Life Technologies capillary based DNA sequencers. The primer mix used for all experiments was made up with the following ratios of each 10µM primer: 1.25X D5S818, 1X D13S317, 1.5X D7S820, 1.25X D16S539, 1X CSF1PO, 3X Penta D and 1X Amelogenin (IDT). Extra Penta D was added separately at a later stage for improved peak height at this locus, bringing the total concentrations to: 0.63 µM D5S818, 0.50 µM D13S317, 0.75 µM D7S820, 0.63 µM D16S539, 0.50 µM CSF1PO, 1.50 µM Penta D and 0.50 µM Amelogenin (IDT)

B.3. Thermal Cycling and Rapid Polymerase

The thermal cycler used for all experimental amplifications was the Philisa from Streck (Nebraska, US). This instrument is comprised of an 8-well low thermal mass silver sample block, and uses Peltier heating and cooling to achieve heating rates of 15ºC/s and cooling rates of 12ºC/s. The instrument produces high temperature ramp rates using thin walled 50 µL polypropylene tubes containing a relatively high surface to volume ratio. This aids in heat transfer from the silver block to the PCR mix.

For this study Z-Taq polymerase (Takara Bio, Clontech Laboratories, Inc., Mountain View, CA), supplied with 10 X Z-Taq buffer and 2.5 mM dNTP mixture, was initially used. Omnitaq (DNA Polymerase Technologies, Inc. St. Louis, MO), supplied with 10 X Taq buffer, was also tested and determined to perform better when used with samples of diluted

saliva. This polymerase was specially designed to be resistant to a number of common inhibitors that can be present in biological samples

B. 4. PCR Optimization

To remove the extraction step from the previously developed rapid PCR protocol [Gibson-Daw 2018] we first experimented with taking cheek swabs and the transferal of the cells and DNA from these to FTA paper. Once the cells were fixed to the FTA paper, the chemicals present on the paper can lyse open the cells and release the DNA. A punch of FTA paper was then added to 1.5 mL of H_2O and heated to $37^{\circ}C$ for 15 mins, the water was then replaced and the heating step repeated a second time before the punch was added to the PCR master mix. This was worked well but added 30 minutes to the PCR process. To save time experiments were initiated involving direct PCR in which saliva was collected in a tube. The saliva samples once collected were spun down on a centrifuge at 13000 RPM for 1 min, the supernatant removed and 40 µL of amplification grade H_2O added to resolubilize the DNA in the liquid. The saliva was then vortexed and diluted to 10% for use in the following experiments.

B.4.1 Optimization of BSA Concentration

When using unextracted samples in the master mix there is also the potential that free DNA and intact cells and other compounds (proteins, salts, lipids, etc), many of which may be inhibitory to the PCR reaction, will also be present. In an attempt to counteract any inhibition these elements may cause, we are also adding BSA and optimizing its concentration in the mix. The amplification occurred in a final PCR master mix volume of 8 µL, to which 2 ng of the diluted to 10 % and washed saliva was added. The master mix

included up of 0.86 µL 10 X Z-Taq buffer (Takara Bio), 0.7 µL 2.5 mM each dNTP mix (Takara Bio), 4.2 µL of primer mix (IDT), 0.86 µL 25mM $MgCL_2$ and 1.1 µL of 2.5 Units/µL of Z-Taq polymerase (Takara Bio). Total concentrations for a 10 µL reaction were: 0.86 X Z-Taq buffer, 0.18 mM each dNTPs, primer mix (0.33 µM D5S818, 0.26 µM D13S317, 0.39 µM D7S820, 0.33 µM D16S539, 0.26 µM CSF1PO, 0.79 µM Penta D and 0.39 µM Amelogenin), 2.15 mM $MgCL_2$ and 0.28 Units/10 µL of Z-Taq polymerase. Amounts of 20 mg/mL BSA tested were 0 µL, 0.75 µL, 1.25 µL, 2 µL, 2.75 µL and 3.5 µL with the total concentrations in the final 10 µL reaction being 0, 1.5, 2.5, 4, 5.5 and 7 mg/mL respectively (Figure 9.1A-F). The total PCR reaction volume per sample was 10 µL (8 µL master mix, 2 µL diluted to 10 % saliva). Cycling conditions were 35 cycles of denaturation at 98°C for 3 s, followed annealing/extension at 62°C for 12 s.

B.4.2 Optimization of dNTPs with Omnitaq

Further experiments were performed using OmniTaq Polymerase (DNA Polymerase Technologies) in an attempt to increase inhibitor resistance. This OmniTaq polymerase were tested in combination with regular Taq buffer and with the addition of PEC-1 (PCR enhancer cocktail) as recommended by manufacturer.

The next parameter to undergo optimization was the concentration of dNTPs. For this experiment the master mix used contained: 0.2 µL of 1 X OmniTaq, 5.8 µL primer mix (IDT), 1.2 µL of 10 X Taq buffer, 1.2 µL of 25 mM $MgCl_2$ and 1 µL of 1 X PEC-1. Final concentrations in the 10 µL reaction were: 0.2 X OmniTaq polymerase, primer mix (0.46 µM D5S818, 0.36 µM D13S317, 0.54 µM D7S820, 0.46 µM D16S539, 0.36 µM CSF1PO, 1.11 µM Penta D and 0.54 µM Amelogenin), 1.2 X Taq buffer, 3 mM $MgCl_2$ and 0.1 X PEC-1. Amounts of 2.5 mM dNTP mix tested were 0.75 µL, 1 µL, 1.25 µL, 1.5 µL and 2

µL with the total concentrations in the final 10 µL reaction being 0.19, 0.25, 0.31,0.37 and 2 0.5 mM respectively (Figure 9.2A-E). 2 µL of diluted saliva was added to 8 µL of this master mix. The total PCR reaction volume per sample was 10 µL (8 µL master mix, 2 µL diluted to 10 % saliva). Cycling conditions were 32 cycles of denaturation at 98°C for 3 s, followed annealing/extension at 62°C for 13 s.

B.4.3 Optimization of MgCl₂

The optimization of the MgCl₂ concentration was performed in combination with the optimization of dNTPs. 3 differing amounts of 2.5 mM dNTPs: 1.25 µL, 1.5 µL, 2 µL (C, D, E from previous experiment) with the total concentrations in the final 10 µL reaction being 0.31, 0.38, 0.5 mM respectively were chosen that in the previous experiment had shown the best results. These were tested in combination with using 1uL, 1.4uL, 1.6 µL, 2uL and 2.5 µL of 25 mg/mL MgCl₂ with the total concentrations in the final 10 µL reaction being 2.5, 3.5, 4.0, 5.0, 6.3 mg/mL respectively. Table 1 shows each of the master mixes and the amounts of reagent used for 1 sample. The total PCR reaction volume per sample was 10 µL (8 µL master mix, 2 µL diluted to 10 % saliva). Cycling conditions were 32 cycles of denaturation at 98°C for 3 s, followed annealing/extension at 62°C for 13 s.

Master mix	Omnitaq (µL) 1 X	Primer mix (µL)	Taq Buffer (µL) 10 X	dNTPs (µL) 2.5mM	MgCl$_2$ (µL) 25mM	PEC-1 (µL) 1 X	DNA (µL)	Number of loci amplified
C1	0.2	5.8	1.2	1.25	1.2	2	2	6/7
C1.4	0.2	5.8	1.2	1.25	1.4	2	2	6/7 (Fig. 9.3A)
C1.6	0.2	5.8	1.2	1.25	1.6	2	2	6/7
C2	0.2	5.8	1.2	1.25	2	2	2	5/7
C2.5	0.2	5.8	1.2	1.25	2.5	2	2	5/7
D1	0.2	5.8	1.2	1.5	1	2	2	6/7 (Fig. 9.3B)
D1.4	0.2	5.8	1.2	1.5	1.4	2	2	5/7
D1.6	0.2	5.8	1.2	1.5	1.6	2	2	5/7
D2	0.2	5.8	1.2	1.5	2	2	2	5/7
D2.5	0.2	5.8	1.2	1.5	2.5	2	2	5/7
E1	0.2	5.8	1.2	2	1	2	2	5/7
E1.4	0.2	5.8	1.2	2	1.4	2	2	6/7 (Fig. 9.3C)
E1.6	0.2	5.8	1.2	2	1.6	2	2	6/7
E2	0.2	5.8	1.2	2	2	2	2	6/7
E2.5	0.2	5.8	1.2	2	2.5	2	2	5/7

Table 9.1 Optimization of MgCl$_2$ in relation to dNTP concentration

Table 9.1 Optimization of MgCl$_2$ in relation to dNTP concentration. Table 9.1 shows the amounts of the different reagents used to optimized MgCl$_2$.

B.4.4 Increase of Penta D Concentration

When performing rapid PCR the smaller loci generally amplify more efficiently than the larger ones. This can result in allele dropout at higher amplicon sizes. To remedy the allelic dropout observed at the Penta D locus, a series of experiments were performed in which the concentration of the Penta D primers was increased to determine if this would lead to improved peak height of the alleles at this locus. The master mix used for these experiments was: 0.2 µL of 1 X OmniTaq, 5.8 µL primer mix (IDT), 1.2 µL of 10 X Taq buffer, 1.2 µL of 25 mM $MgCl_2$, 1 µL of 1 X PEC-1, 1.25 µL of 2.5 mM each dNTP and 0.15 µL of 20 mg/mL BSA. Final concentrations in the 10 µL reaction were: 0.2 X OmniTaq polymerase, primer mix (0.46 µM D5S818, 0.36 µM D13S317, 0.54 µM D7S820, 0.46 µM D16S539, 0.36 µM CSF1PO, 1.11 µM Penta D and 0.54 µM Amelogenin), 1.2 X Taq buffer, 3 mM $MgCl_2$ 0.1 X PEC-1, 0.31 mM each dNTP and 0.3 mg/mL BSA. Amounts of extra 10 µM forward and reverse Penta D primer added were 0 µL, 1 µL and 2 µL which correspond to total concentrations of Penta D in the 10 µL reaction being: 1.11, 2.11 and 3.11 µM respectively (Figure 9.4A-C). The total PCR reaction volume per sample was 10 µL (8 µL master mix, 2 µL diluted to 10 % saliva). Cycling conditions was 35 cycles of denaturation at 98^OC for 3 s, and annealing/extension at 62^OC for 20 s. An extra 2 µL of Penta D was also tested under the following cycling conditions (Figure 9.4D) 32 cycles of denaturation at 98^OC for 5 s, followed annealing/extension at 62^OC for 24 s.

B.4.5 Optimization of Cycling Conditions

The cycling conditions used to optimize the total amplification time are shown in Table 9.2. Table 9.2 contains the temperature and time at each step, the number of cycles, the

total amplification time and the number of 7 loci successfully amplified. Throughout the optimization process, the denaturing times, annealing times and cycle numbers were decreased both individually and in a variety of combinations in an effort to achieve the fastest and most robust amplifications. Cycling conditions were optimized based on 3 parameters, total reaction time, peak height and peak balance. The cycling conditions yielding the best profiles can be seen in Figure 5. The master mix used for these experiments contained: 0.2 µL of 1 X OmniTaq, 5.8 µL primer mix (IDT), 1.2 µL of 10 X Taq buffer, 1.2 µL of 25 mM $MgCl_2$, 1 µL of 1 X PEC-1, 1.25 µL of 2.5 mM each dNTP and 0.15 µL of 20 mg/mL BSA and extra 1 µL of 10 µM Penta D forward and reverse primers. Final concentrations in the 10 µL reaction were: 0.2 X OmniTaq polymerase, primer mix (0.46 µM D5S818, 0.36 µM D13S317, 0.54 µM D7S820, 0.46 µM D16S539, 0.36 µM CSF1PO, 2.11 µM Penta D and 0.54 µM Amelogenin), 1.2 X Taq buffer, 3 mM $MgCl_2$ 0.1 X PEC-1, 0.31 mM each dNTP and 0.3 mg/mL BSA. 2µL of diluted to 10 % saliva was added to 8 µL of this master mix while the cycling conditions varied according to Table 9.2.

Table 9.2 Optimization of cycling conditions					
Cycling condition	Temp (^{0}C)	Time (s)	Cycles	Total time	Number of loci amplified
A	98	3	32	12'16"	4/7
	62	13			
B	98	5	32	18'56"	7/7
	62	24			
C	98	3	35	17'23"	7/7
	62	20			
D	98	5	32	18'55"	7/7
	62	24			
E	98	4	32	18'39"	7/7
	62	24			
F	98	3	32	17'50"	7/7
	62	24			
G	98	2	32	17'29"	7/7
	62	24			
H	98	1	32	16'56"	7/7
	62	24			
I	98	2	35	16'36"	7/7
	62	20			

Cycling condition	Temp (^{0}C)	Time (s)	Cycles	Total time	Number of loci amplified
J	98	1	35	16'14"	7/7
	62	20			
K	98	3	35	15'58"	7/7
	62	18			
L	98	3	35	15'36"	7/7
	62	17			
M	98	1	32	17'21"	6/7
	62	25			
N	98	3	32	14'20"	7/7
	62	17			
O	98	1	33	18'00"	7/7
	62	25			
P	98	4	32	13'32"	7/7
	62	15			
Q	98	2	32	13'34"	7/7
	62	17			
R	98	3	31	13'41"	7/7
	62	17			
S	98	3	32	13'00"	7/7
	62	15			

Cycling condition	Temp (^{0}C)	Time (s)	Cycles	Total time	Number of loci amplified
T	98	2	32	13'16"	7/7
	62	16			
U	98	3	32	13'34"	7/7
	62	16			
V	98	3	32	13'37"	7/7
	61	16			
W	98	3	32	13'29"	7/7
	63	16			
X	98	3	31	12'45"	3/7
	62	15			
Y	98	3	30	12'24"	2/7
	62	15			

Table 9.2 Optimization of cycling conditions. This table shows the different times spent at the denaturing and annealing/elongation steps along with how many cycles these steps are repeated for, total time of amplification and how many of the 7 loci amplified successfully. Master mix: 0.2 X OmniTaq, primer mix (0.46 µM D5S818, 0.36 µM D13S317, 0.54 µM D7S820, 0.46 µM D16S539, 0.36 µM CSF1PO, 2.11 µM Penta D and 0.54 µM Amelogenin), 1.2 X Taq buffer, 3.4 mM MgCl$_2$, 0.31 mM each dNTPs, 0.3 mg/mL BSA and 0.1 X PEC-1. 2µL of diluted to 10 % saliva was added to 8 µL of this master mix to give a total volume of 10 µL per sample.

B.5 Rapid Direct Amplification

The final procedure (Figure 9.6) consisted of a master mix containing: 0.2 µL of 1 X OmniTaq, 5.8 µL primer mix (IDT), 1.2 µL of 10 X Taq buffer, 1.2 µL of 25 mM MgCl$_2$, 1 µL of 1 X PEC-1, 1.25 µL of 2.5 mM each dNTP and 0.15 µL of 20 mg/mL BSA and

extra 1 µL of 10 µM Penta D forward and reverse primers. Final concentrations in the 10 µL reaction were: 0.2 X OmniTaq polymerase, primer mix (0.46 µM D5S818, 0.36 µM D13S317, 0.54 µM D7S820, 0.46 µM D16S539, 0.36 µM CSF1PO, 2.11 µM Penta D and 0.54 µM Amelogenin), 1.2 X Taq buffer, 3 mM $MgCl_2$ 0.1 X PEC-1, 0.31 mM each dNTP and 0.3 mg/mL BSA. 2µL of diluted to 10 % saliva was added to 8 µL of this master mix, giving a total volume of 10 µL. Amplifications were performed on a Philisa thermocycler (Streck) at a total cycling time of 13.5 min. Optimal conditions were 32 cycles of 98ºC for 3 seconds followed by 62 ºC for 16 seconds.

B.6. Separation, Detection and Analysis

B.6.1 Applied Biosystems ABI 310 Capillary Electrophoresis Genotyper

All PCR products were analysed using an ABI 310 (Applied Biosystems) by adding 1 µL of PCR product to a mixture of 12.5 µL HiDiTM Formamide (Applied Biosystems) and 0.5 µL internal lane standard (ILS). The ILS used throughout this paper was LIZ 500 (Applied Biosystems) which was labelled with the orange LIZ dye. This mix was then run on the ABI 310 using the following conditions: POP4 polymer, Capillary length to detector: 36 cm, module GS STR POP4 (1 mL), injection voltage 15 kV, injection time 5 s, run time 28 min, run temperature 60ºC, run voltage 15 kV. Software used for Analysis was Genemapper.

C. Results and Discussion

C.1 PCR Optimization

C.1.1 Optimization of BSA Concentration

The profiles resulting from changes to the BSA concentrations can be seen in Figure 9.1.
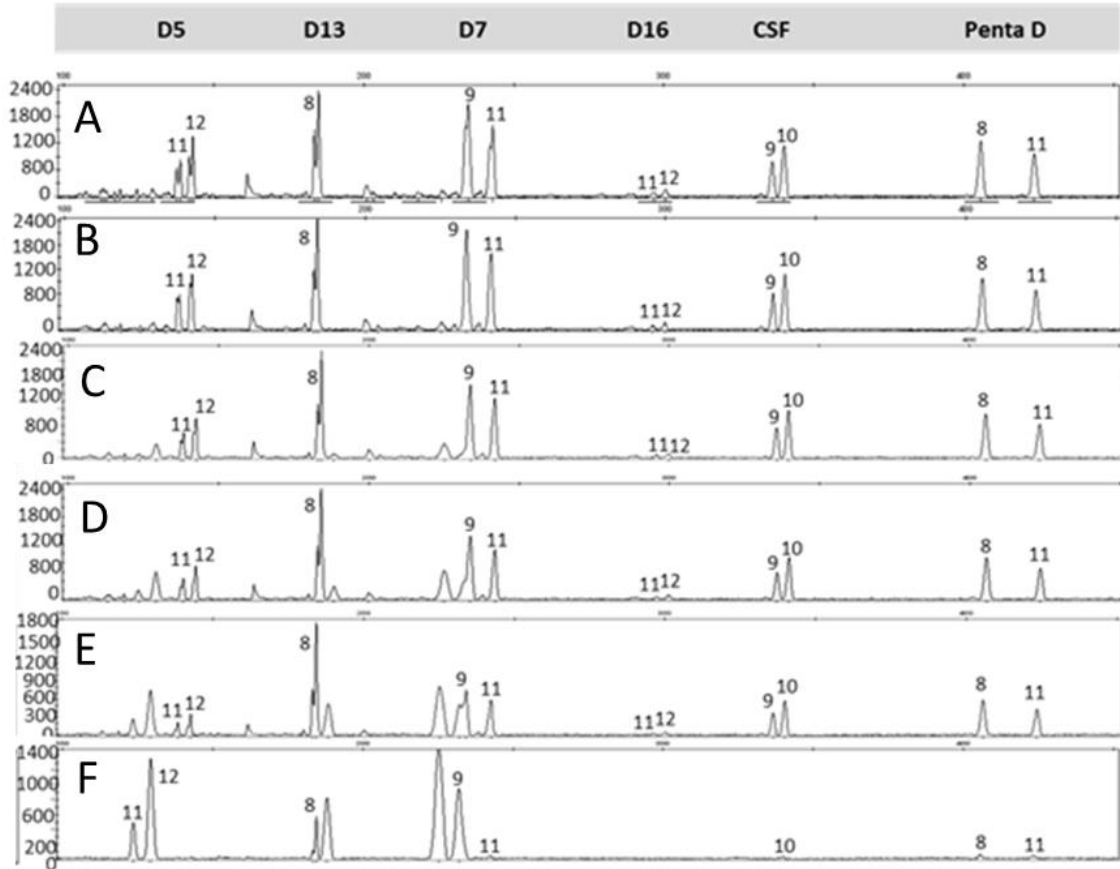
Fig.9.1 Optimization of BSA concentration. BSA concentration was increased for each run starting at (A): 0 µL, (B): 1.5 mg/mL, (C): 2.5 mg/mL, (D):4 mg/mL, (E): 5.5 mg/mL and (F): 7 mg/mL. The master mix was made up of 0.86 X Z-Taq buffer, 0.18 mM each dNTP mix, primer mix (0.33 µM D5S818, 0.26 µM D13S317, 0.39 µM D7S820, 0.33 µM D16S539, 0.26 µM CSF1PO, 0.79 µM Penta D and 0.39 µM Amelogenin), 2.15 mM MgCl$_2$ and 0.28 Units/ 10µL of Z-Taq polymerase. Total volume was 10 µL: 2 µL of diluted to 10 % saliva and 8 µL master mix. Cycling conditions were 35 cycles of denaturation at 98°C for 3 s, followed annealing/extension at 62°C for 12 s.

As can be observed in Figure 9.1 increasing the addition of BSA results in a loss of peak height especially for the larger loci. Profiles D and E also show some non-specific

amplification at the D5 and D7 loci, while profile F shows allele dropout at D16. Profiles A and B from Figure 9.1 show the highest peak heights and best balance, so it was determined that between 0 and 0.75 µL (0-1.5 mg/mL) of BSA should be added per sample.

C.1.2 Optimization of dNTPs with Omnitaq

The switch of polymerase from Z-Taq to Omnitaq resulted in an improvement of peak height for all the loci of the multiplex.

The next step in the optimization process was to determine what concentration of dNTPs would yield the highest peaks and the cleanest profiles. To do this the concentration of dNTPs was increased incrementally, Figure 9.2.
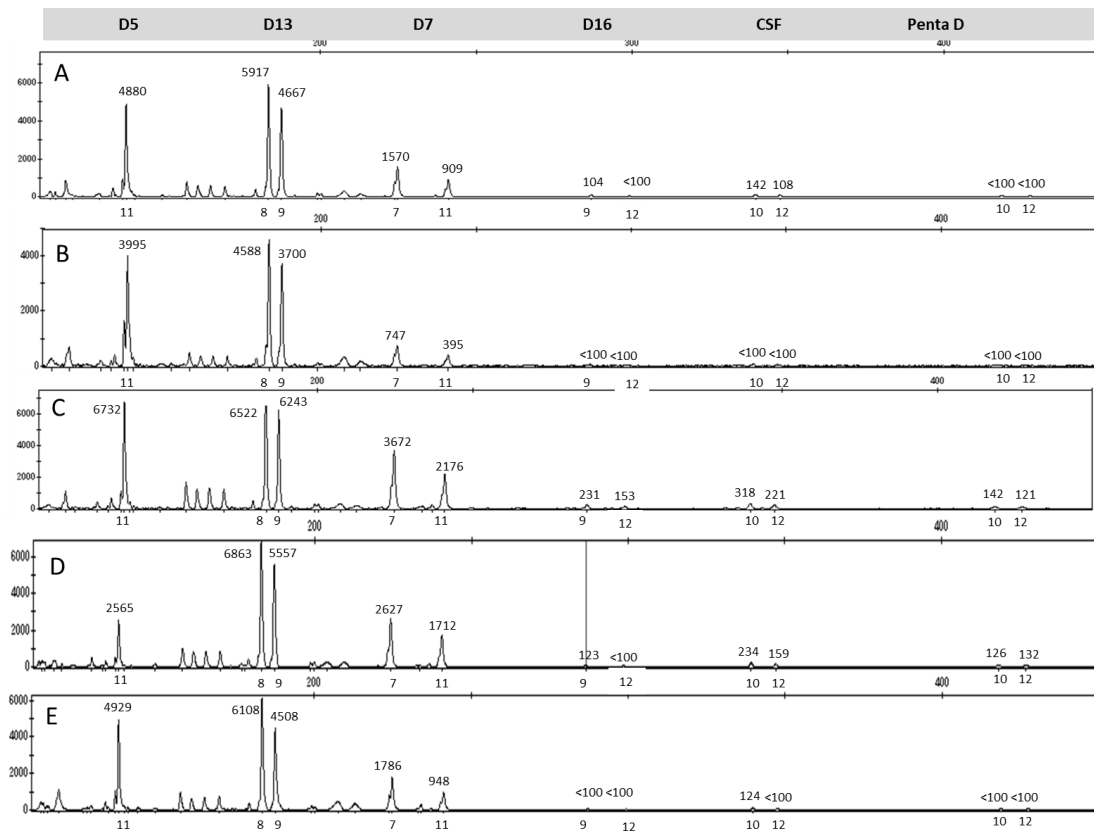
Fig.9.2 Optimization of dNTP concentration. dNTP concentration was increased for each run starting at (A): 0.19 mM, (B): 0.25 mM, (C): 0.31mM, (D): 0.38mM, (E): 0.5mM. Total volume was 10 µL: 2 µL of diluted to 10 % saliva and 8 µL master mix.

170

Figure 9.2. Shows increasing concentrations of dNTPs being added to the samples. Profiles C showed the highest peak heights and best balance so it was determined that 1.25 µL of 2.5 mM each (total concentration in reaction 0.31 mM each) dNTPs should be added per sample.

C.1.3 Optimization of MgCl₂

Due to the fact that MgCl$_2$ interacts directly with dNTPs it is important to optimize its concentration for each of the concentrations of dNTPs determined by the previous experiments (C, D and E). The last column of the table 9.1 shows how many of the 7 loci were successfully amplified for each set of reagent concentrations. The best results were obtained using master mixes C1.4, D1 and E1.4 and can be seen in Figure 9.3A-C.
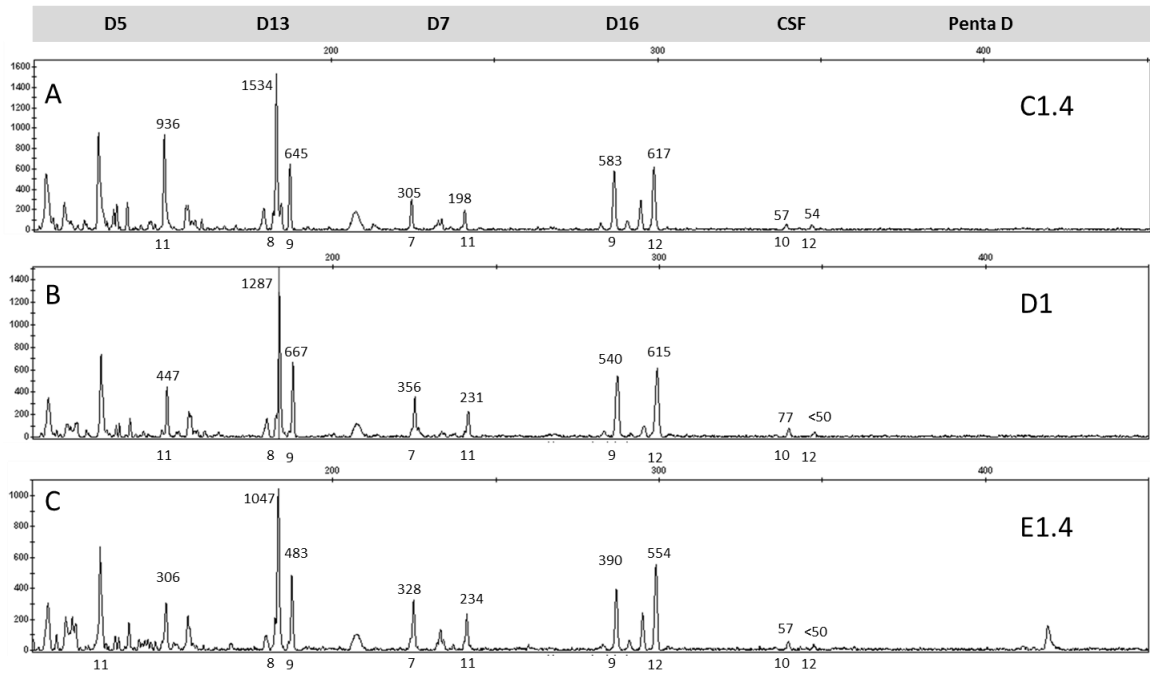
Fig.9.3 Optimization of MgCl$_2$ in relation to differing amounts of dNTPs. Figure 3 shows the best resulting profiles from the differing combinations of dNTPs and MgCl$_2$: (A): C1.4: 0.31 mM dNTPs and 3.5 mM MgCl$_2$, (B): D1: 0.37 mM dNTPs and 2.5 mM MgCl$_2$ and (C): E1.4: 0.5 mM dNTPs and 3.5mM MgCl$_2$.

In all of the previous experiments only 5 of the 6 STR loci. CSF showed low peaks while

Penta D had dropped out. All profiles showed dropout at the Penta D loci so optimization

was deduced using the other 6 loci. By comparing these profiles, it was decided that the

reagent concentrations C1.4: 1.25 µL dNTPs with 1.4 µL of $MgCl_2$ with total

concentrations in the 10 µL reaction being 0.31 mM dNTPs and 3.5 mM of $MgCl_2$ were

optimal.

C.1.4 Increase of Penta D Concentration

In order to remedy the problem observed in the previous experiment with dropout at the

Penta D locus, the addition of extra Penta D primers was tested (Figure 9.4A-C) along with

a combination of extra primer and extra time at the elongation step (Figure 9.4D).
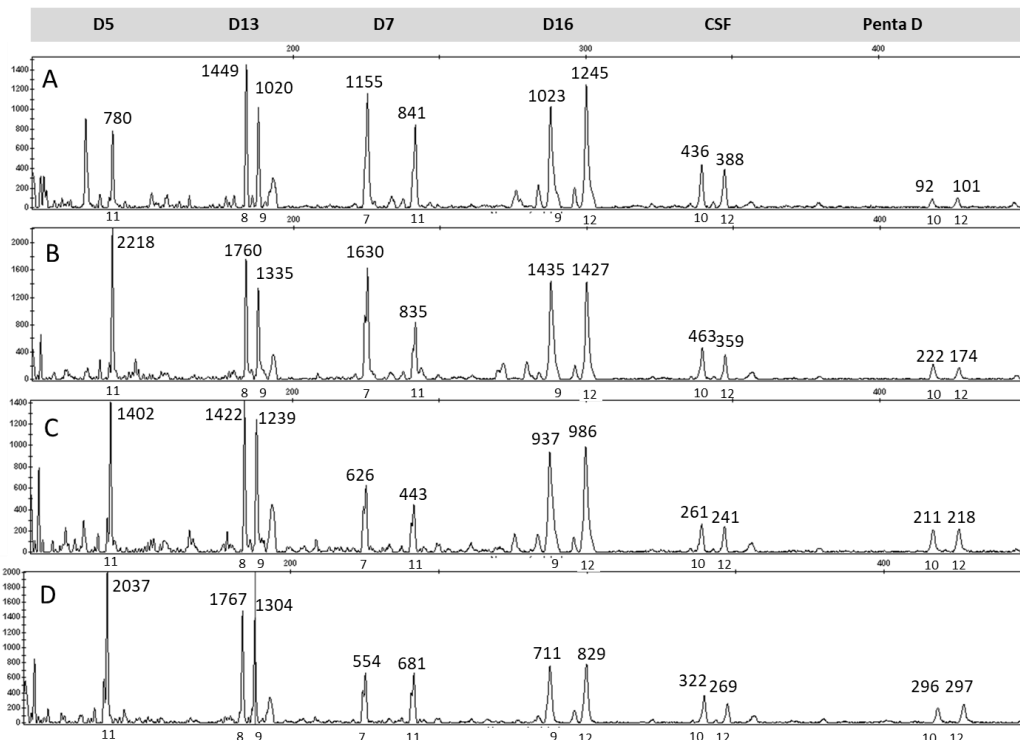
Fig.9.4 Optimization of Penta D concentration. Extra Penta D primers were added to try and increase the peak height of alleles at this locus. (A): 1.11 µM of each forward and reverse primers, (B): 2.11 µM of each forward and reverse primers and (C): 3.11 µM of each forward and reverse primers.

172

Profile D in Figure 9.4 shows the effect of an additional 2 µL (total concentration in reaction 3.11 µM) of each forward and reverse primers with different cycling conditions (32 cycles of denaturation at 98°C for 5 s, followed annealing/extension at 62°C for 24 s). These parameters improved peak height but also increased total amplification time. Based on these profiles it was determined that an extra 1 µL (total concentration in reaction 2.11 µM) per sample of each forward and reverse primers for Penta D should be added to the multiplex.

The best results were obtained with 1 µL of extra Penta D primers added to the master mix which would give the total concentration in a 10 µL reaction to be 2.11 µM of both forward and reverse Penta D primers. Peak heights for the Penta D locus were also higher when an extra 4s was added to the extension time of the PCR cycles.

C.1.5 Optimization of Cycling Conditions

It was also important to optimize the cycling conditions in order to achieve profiles with the highest peak height while maintaining both an inter and intra locus balance and minimizing non-specific amplification and other artefacts. These results had to be carefully balanced with the decrease in total amplification time in order to optimize peak height and balance in the fastest time.

The best results were obtained using master mixes E, N, P, S, T and U and can be seen in Figure 9.5. These conditions resulted in the optimal peak heights for all loci obtained in the shortest amplification times. As total amplification time decreased, the peak height of the alleles also decreased. However, all loci were present for most of the cycling conditions until total amplification speed dropped below 13 minutes. The cycling conditions that

173

resulted in the profiles with highest peak height and best balance along with the fastest

speeds are illustrated in Figure 9.5.
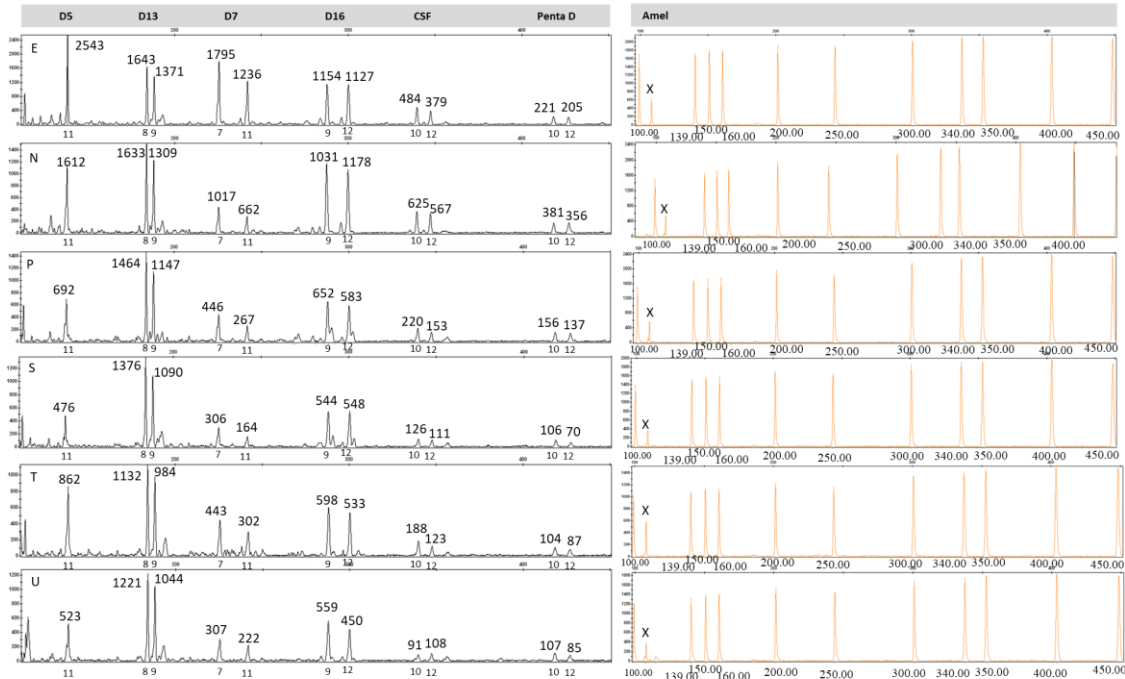
Fig.9.5 Optimization of cycling conditions. Figure 9.5 shows profiles from a selection of cycling conditions that presented both fastest run times along with highest peak heights and balances. U shows overall the best compromise between peak height and quality and total amplification time.

Conditions S (32 cycles of 98°C for 3 s and 62°C 15s) resulted in a total amplification

time of 13 minutes. This was the fastest total amplification time, but these profiles showed

adenylation at many loci and some elevated noise that could prove confusing for later data

interpretation. Cycling conditions P, T and U all had a total amplification times of around

13.5 minutes. Of these conditions P showed very slight increases in peak height but also

in adenylation at the D13 and D16 loci. This may be due to the extra second spent at

denaturation while annealing was one second shorter possibly preventing the polymerase

from completing the adenylation. T also showed some minor adenylation and nonspecific

amplification at these locations as well as at the D7 locus, this is thought to be due to the polymerase not fully completing a chain in the annealing step before the next denaturing step begins. U was overall the cleanest of the three but had slightly lower peek heights than P, since the peaks were still clearly discernible from the noise it was decided that the decrease in height was acceptable in exchange for a cleaner profile overall. It was decided that the cycling conditions seen for conditions U were optimal.

C.2 Rapid Direct Amplification
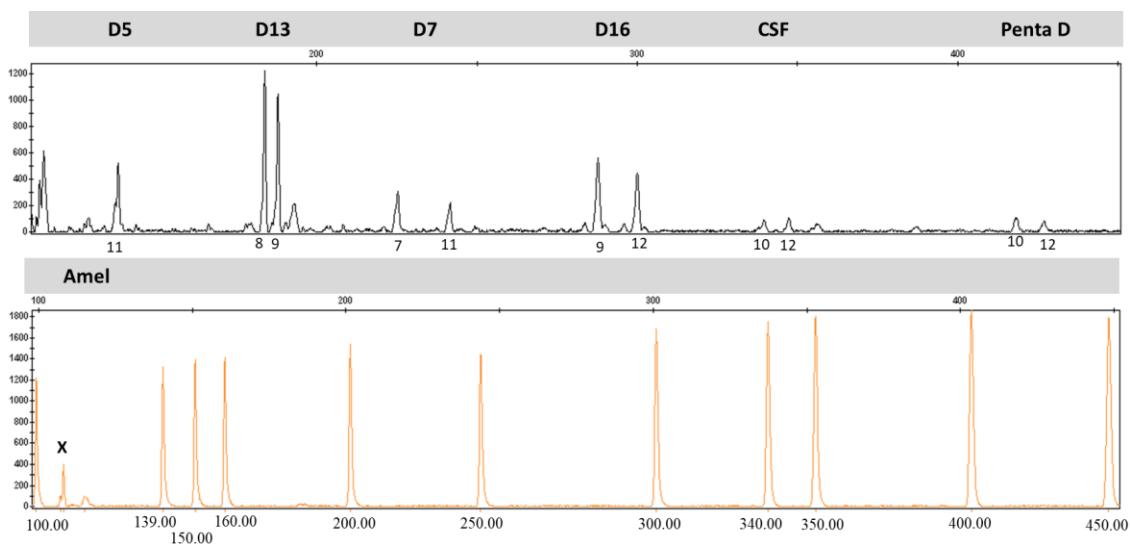
Fig.9.6 Profile generated from dilutes saliva using optimized protocol. The top lane shows the 6 STR loci labelled with TAMRA dye, while the bottom lane shows the orange lane with the ILS Liz 500 and the Amelogenin locus labelled with Liz.

Figure 9.6 shows the successful amplification of a sample of 10% diluted saliva in 13.5 minutes. This 7 locus multiplex has a power of discrimination of 1×10^6 or greater, which is sufficient to identify a potential suspect, once identified a full laboratory analysis would be required using standard genotyping methods to confirm a match.

D. Conclusions

In this chapter, we have demonstrated the rapid direct amplification of a 7 Loci multiplex in 13.5 minutes. This was achieved through the use of a high-speed thermocycler and a rapid and inhibitor resistant polymerase, along with a PCR enhancer cocktail. Optimization of cycling conditions, reagent concentrations and sample volume played a crucial role in the method development. Overall, this method provides a rapid and complete sample analysis with no need for DNA extraction, reducing the need for reagents and money spent on sample processing. This method only uses a small portion of the sample collected, leaving the remainder for conventional analysis if deemed appropriate and necessary at a later date.

CHAPTER X. CONCLUSIONS AND FUTURE WORKS

The overall goal of this project was to develop a rapid PCR method that could amplify both Y and autosomal multiplexes in the quickest possible time. These goals were achieved with the use of the Philisa rapid thermocycler combined with rapid polymerases Z-Taq and Omnitaq and the addition of BSA and PEC-1.

A 4 locus Y-STR multiplex used for screening sexual assault samples for male DNA was amplified in less than 17 minutes by combining the Philisa rapid thermocycler with the use of the rapid Z-Taq polymerase and buffer. These samples could then be run on the Agilent microfluidic chip to produce full sample analysis in under 25 minutes. This method was proved to be an efficient way of detecting and profiling male DNA even in mixtures of female and male DNA to a ratio of 40:1.

The MP7 multiplex designed by Maurice Aboud was successfully amplified in 6.5 minutes also using the Philisa thermocycler and Z-Taq polymerase and buffer. This multiplex which contains a subset of the expanded CODIS loci can act as a screening tool for samples in a variety of situations when analysis needs to occur quickly. The multiplex itself has a power of discrimination of over 1×10^6, while the method developed only requires a small portion of the sample leaving the rest for downstream analysis if required. This method coupled to the microfluidic chip method of separation detection would allow for full sample analysis to occur in under 15 minutes. This system was validated with a small test population of 30 volunteers and was shown to be sensitive down to 2 ng of input DNA.

A rapid direct method was also developed for the MP7 multiplex where saliva was added directly to the PCR reaction thus eliminating an extraction step and further reducing

analysis time. For this rapid direct method amplification was achieved in under 13.5 minutes with the only sample prep being a dilution of the saliva collected to 10% in water.

These methods could easily be implemented in the field as they require minimal equipment, reagents and sample handling, while providing robust and reproducible results with a good power of discrimination. Their use could be invaluable to produce fast results on site at crime scenes, police stations, airport security controls and scenes of mass disaster and terrorist attack.

In the future new thermocyclers will be manufactured that permit even faster amplification times by speeding up the heating and cooling ramp rates or by a completely new design of thermocycling. New and improved polymerases are continuously being created and investigated both in regard to their processivity speeds and their resistance to inhibitors. As these polymerases progress it will be easier and quicker to produce DNA profiles straight from biological samples. Future work will focus on finding new polymerase and optimizing the PCR for their use with a variety of body fluids such as blood. Rapid extraction methods should also be investigated such as short bursts of microwave heat to help lyse the cell and release the DNA. It is the authors opinion that a for further reduction of amplification time and separation time a multiplex could be designed to be run on a six (or more) dye system where one or two loci are observed in each colour channel. These loci would be selected based on their small size and mini primers designed to reduce the size of the DNA fragments produced. Due to the small size of the DNA fragments they will amplify much faster and as there will be only one or two of these loci per dye lane then run times on conventional CE instruments can be reduced considerable. Another area

of focus will be the development and validation of a rapid separation and detection system similar to the microfluidic chip used in this research. With the modern technological advances instrumentation is always being updated and this will hold true for our current methods of electrophoresis. It may be that soon we will see less capillary electrophoresis and completely switch to whole genome sequencing.

LIST OF REFERENCES

Aboud MJ, Gassmann M, McCord B. The development of mini pentameric STR loci for rapid analysis of forensic DNA samples on a microfluidic system. Electrophoresis. 2010;31(15):2672-9

Aboud MJ. The development of direct ultra-fast PCR for forensic genotyping using short channel microfluidic systems with enhanced sieving matrices. PhD Dissertation. Florida International University. 2012

Aboud MJ, Oh HH, McCord B. Rapid direct PCR for forensic genotyping in under 25 min. Electrophoresis. 2013;34(11):1539-47

Aboud MJ, Gassmann M, McCord B. Ultrafast STR separations on short-channel microfluidic systems for forensic screening and genotyping. J Forensic Sci. 2015;60(5):1164–70

Abuidrees AS, Alghafri RH, Hadi S. Rapid amplification of the RM-Yplex assay. Electrophoresis. 2016. doi:10.1002 /elps.201500549.

Akane A, Matsubara K, Nakamura H, Takahashi S, Kimura K. Identification of the heme compound copurified with deoxyribonucleic acid (DNA) from bloodstains, a major inhibitor of polymerase chain reaction (PCR) amplification. J Forensic Sci. 1994;39(2):362-72

Alaeddini R. Forensic implications of PCR inhibition-A review. Forensic Sci. Int. Genet. 2012;6(3):297-305.

Albani P. Rapidly Mutating Y-STR Multiplex Design and Analysis by Microchip Electrophoresis. Masters Thesis. Florida International University.2013.

Alberts B. Molecular Biology of the Cell. Garland Science.2014; 6th Edition

Alghafri R, Goodwin W, Ralf A, Kayser M, Hadi S. A novel multiplex assay for simultaneously analysing 13 rapidly mutating YSTRs. Forensic Sci Int Genet. 2015;17:91–8

Al-Soud WA, Jonsson LJ, Radstrom P. Identification and characterization of Immunoglobulin G in blood as a major inhibitor of diagnostic PCR. J. Clin. Microbiol. 2000;38(1):345-50

Al-Soud WA, Radstrom P. Purification and characterisation of PCR-inhibitory components in blood cells. J. Clin. Microbiol. 2001;39(2):485-93

An JH, Yang WI. Body fluid identification in forensics. BMB Rep. 2012;45(10):545-53

Atsü SS, Gökdemir K, Kedici PS, Ikyaz YY. Bitemarks in forensic odontology. J Forensic Odontostomatol. 1998;16(2):30-4

Baar C, D'abbadie M, Vaisman A, Arana ME, Hofreiter M, Woodgate R, Kunkel TA, Holliger P. Molecular breeding of polymerases for resistance to environmental inhibitors. Nucleic Acids Res. 2011:39(8):e51

Baird JB. The Individuality of Blood and Bloodstains: A Non-Critical Historical Review with Some Reference to the More Recent Canadian Endeavours to Characterize Human Blood. Can. Soc. Forensic Sci. J. 1978;11(2):83-122

Ballantyne KN, Keerl V, Wollstein A, Choi Y, Zuniga SB, Ralf 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):208-18

Barni F, Lewis SW, Berti A, Miskelly GM, Lago G. Forensic application of the luminol reaction as a presumptive test for latent blood detection. Talanta. 2007;72(3):896-913

Barrett LW, Fletcher S,Wilton SD. Regulation of eukaryotic gene expression by the untranslated gene regions and other non-coding elements. Cell Mol Life Sci. 2012;69(21):3613-3634

Barron AE, Sunada WM, Blanch HW. The use of coated and uncoated capillaries for the electrophoretic separation of DNA in dilute polymer solutions. Electrophoresis. 1995;16(1):64-74.a

Barron AE, Blanch HW. DNA Separations by Slab Gel and Capillary Electrophoresis - Theory and Practice. Separ. Purif. Method. 1995;24(1):1-118.b

Bedford E, Tabor S, Richardson CC. The thioredoxin binding domain of bacteriophage T7 DNA polymerase confers processivity on Escherichia coli DNA polymerase I. Proc. Natl. Acad. Sci. U.S.A. 1997;94(1):479-84

Belgrader P, Smith JK, Weedn VW, Northrup MA. Rapid PCR for identity testing using a battery-powered miniature thermal cycler. J Forensic Sci 1998;43(2):315-19

Belgrader P, Benett W, Hadley D, Richards J, Stratton P, Mariella Jr R, Milanovich F. PCR detection of bacteria in seven minutes. Science. 1999;284(5413):449-50

Bell S. Crime and circumstance: investigating the history of forensic science. Praeger. 2008.

Belova GI, Prasad R, Kozyavkin SA, Lake JA, SH, Slesarev AI. A type IB topoisomerase with DNA repair activities. Proc. Natl. Acad. Sci. U.S.A. 2001;98(1):6015-20

Bickley J, Short JK, McDowell DG, Parkes HC. Polymerase chain reaction (PCR) detection of Listeria monocytogenes in diluted milk and reversal of PCR inhibition caused by calcium ions. Lett Appl Microbiol. 1996;22(2):153-8

Blake E, Mihalovich J, Higuchi R, Walsh PS, Erlich H. Polymerase chain reaction (PCR) amplification and human leukocyte antigen (HLA)-DQ alpha oligonucleotide typing on biological evidence samples: casework experience. J Forensic Sci. 1992;37(3):700-26

Blau S, Briggs CA. The role of forensic anthropology in Disaster Victim Identification (DVI). Forensic Sci. Int. 2011;205(1):29-35

Botstein D, White RL, Skolnick M, Davis RW. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. Am J Hum Genet. 1980; 32(3):314-31

Brito P, Lopez V, Bogas V, Balsa F, Andrade L, Serra A, São Brento M, Bento AM, Cunha P, Carvalho M, Corte-real F, AnjosMJ. Amplification of non-FTA samples with AmpFLSTR® Identifiler® Direct PCR Amplification Kit. Forensic Sci. Int. 2011;3(1):e371-2

Bu Y, Huang H, Zhou G. Direct polymerase chain reaction (PCR) from human whole blood and filter-paper-dried blood by using a PCR buffer with a higher PH. Anal. Biochem. 2008;375(1): 370-2

Budowle B, Waye J. Hae III--A suitable restriction endonuclease for restriction fragment length polymorphism Analysis Of Biological Evidence Samples. J Forensic Sci. 1990;35(3):530-36

Budowle B, Moretti TR, Baumstark AL, Defenbaugh DA, Keys KM. Population data on the thirteen CODIS core short tandem repeat loci in African Americans, U.S. Caucasians, Hispanics, Bahamians, Jamaicans, and Trinidadians. J Forensic Sci. 1999;44(6):1277-86

Budowle B, Shea B, Niezgoda S, Chakraborty R. CODIS STR loci data from 41 sample populations. Journal of Forensic Sciences. 2001, 46 (3), 453-489.

Budowle B, Masibay A, Anderson SJ, Barna C, Biega L, Brenneke S, Brown BL, Cramer J, DeGroot GA, Douglas D, Duceman B, Eastman A, Giles R, Hamill J, Haase DJ, Janssen DW, Kupferschmid TD, Lawton T, Lemire C, Llewellyn B, Moretti T, Neves J, Palaski C, Schueler S, Sgueglia J, Sprecher C, Tomsey C, Yet D. STR primer concordance study. Forensic Sci. Int. 2001;124(1):47-54.

Budowle B, Bieber FR, Eisenberg AJ. Forensic aspects of mass disasters: strategic considerations for DNA-based human identification. J. Leg. Med. 2005;7(4):230-43

Budowle B, Eisenberg AJ, van Daal A. Validity of low copy number typing and applications to forensic science. Croat Med J. 2009;50(3):207-17.

Budowle B, van Daal A. Comment on: Gill P, Buckleton J. A universal strategy to interpret DNA profiles that does not require a definition of low copy number. Forensic Sci. Int. Gen. 2010;4: 221-7. Forensic Sci. Int: Gen. 2011;5(1):15

Buel E, Schwartz MB, LaFountain MJ. Capillary electrophoresis STR analysis: comparison to gel-based systems. J Forensic Sci. 1998;43(1):164-70

Burgoyne L, Kijas J, Hallsworth P, Turner J. Safe collection, storage and analysis of DNA from blood. Proceedings of the Fifth International Symposium on Human Identification. Promega Corporation. 1994;163

Butler JM, Shen Y, McCord BR. The development of reduced size STR amplicons as tools for analysis of degraded DNA. J Forensic Sci.2003;48(5):1054-64

Butler JM, Buel E, Crivellente F, McCord BR. Forensic DNA typing by capillary electrophoresis using the ABI Prism 310 and 3100 genetic analyzers for STR analysis. Electrophoresis. 2004;25(10-11):1397-412

Butler JM. Forensic DNA Typing: Biology, Technology, and Genetics of STR Markers. Elsevier Academic Press. 2005; 2nd Edition

Butler JM. Genetics and genomics of core short tandem repeat loci used in human identity testing. J Forensic Sci. 2006;51(2): 253-65.

Butler JM, Decker AE, Vallone PM, Kline MC. Allele frequencies for 27 Y-STR loci with U.S. Caucasian, African American, and Hispanic samples. Forensic Sci Int. 2006;156(2–3):250–60

Butler JM. Short tandem repeat typing technologies used in human identity testing. Biotechniques. 2007; 43(4):2-5

Butler JM. Fundamentals of Forensic DNA Typing. Academic Press. 2009;1st Edition

Butler JM. Advanced Topics in Forensic DNA Typing: Methodology. Elsevier Science & Technology.2011; 1st Edition

Butts ELR, Vallone PM. Rapid PCR protocols for forensic DNA typing on six thermal cycling platforms. Electrophoresis. 2014;35(21-22):3053-61

Byrnes JJ, Downey KM, Esserman L, So LG. Mechanism of hemin inhibition of erythroid cytoplasmic DNA polymerase. Biochemistry. 1975;14(1):796-9

Camilleri P. Capillary Electrophoresis: Theory and Practice. CRC Press: 1998.

Caragine T, Mikulasovich R, Tamariz J, Bajda E, Sebestyen J, Baum H, Prinz M. Validation of testing and interpretation protocols for low template DNA samples using AmpFlSTR Identifiler. Croat Med J. 2009; 50(3):250-67

Caragine T, Prinz M. Comment on: Budowle B, et al. Low copy number typing has yet to achieve "general acceptance". Forensic Sci. Int. Genetics: Supplement Series.2009;2:551-2. Forensic Sci. Int: Gen.2011;5(1):3-4; discussion 5-7

Capt CL. Validation of GeneMapper ID Human Identification Software for Forensic STR DNA Analysis. University of North Texas Health Science Center at Fort Worth. 2005.

Chargaff E, Zamenhof S, Green C. Composition of human desoxypentose nucleic acid. Nature. 1950;165(4202): 756-7;

Chargaff E, Magasanik B, Vischer E, Green C, Doniger R, Elson D. Nucleotide composition of pentose nucleic acids from yeast and mammalian tissues. J. Biol. Chem. 1950;186(1): 51-67

Chargaff E. Chemical specificity of nucleic acids and mechanism of their enzymatic degradation. Experientia. 1950;6(6): 201-9.

Chargaff E. Some recent studies on the composition and structure of nucleic acids. J Cell Physiol Suppl. 1951;38 (Suppl. 1):41-59

Carson S, Cohen AS, Belenkii A, Ruiz-Martinez MC, Berka J, Karger BL. DNA sequencing by capillary electrophoresis: use of a two-laser-two-window intensified diode array detection system. Anal. Chem. 1993;65(22): 3219-26

Chen J, Kobilinsky L, Wolosin D, Shaler R, Baum H. A physical method for separating spermatozoa from epithelial cells in sexual assault evidence. J Forensic Sci. 1998;43(1):114–8

Chomczynski P, Rymaszewski M. Alkaline polyethylene glycol-based method for direct PCR from bacteria, eukaryotic tissue samples, and whole blood. Biotechniques. 2006;40(1):454-8

Collins PJ, Hennessy LK, Leibelt CS, Roby RK, Reeder DJ, Foxall PA. Developmental validation of a single-tube amplification of the 13 CODIS STR loci, D2S1338, D19S433, and amelogenin: the AmpFlSTR Identifiler PCR Amplification Kit. J Forensic Sci. 2004;49(6):1265–77

Comey CT, Koons BW, Presley KW, Smerick JB, Sobieralski CA, Stanley DM, Baechtel ES. DNA extraction strategies for amplified fragment length polymorphism Analysis. J Forensic Sci. 1994;39(5):1254-69

Dakin EE, Avise JC. Microsatellite null alleles in parentage analysis. Heredity. 2004;93(5):504-9

Daly DJ, Murphy C, McDermott SD. The transfer of touch DNA from hands to glass, fabric and wood. Genetics. 2012;6(1):41-6

Dauphin LA, Stephens KW, Eufinger SC, Bowen MD. Comparison of five commercial DNA extraction kits for the recovery of Yersinia pestis DNA from bacterial suspensions and spiked environmental samples. J Appl Microbiol. 2010;108(1):163-72

Davidson JF, Fox R, Harris DD, Lyons-Abbott S, Loeb LA. Insertion of the T3 DNA polymerase thioredoxin binding domain enhances the processivity and fidelity of Taq DNA polymerase. Nucleic Acids Res. 2003;31(16):4702-9

de Vega M, Lazaro J.M, Mencia M, Blanco L, Salas M. DNA polymerase amplification performance by fusion of DNA binding motifs. Proc.Natl.Acad.Sci. 2010;107(38):16506-11

Delidow BC, Lynch JP, Peluso JJ, White BA. Polymerase chain reaction. In PCR Protocols Humana Press. 1993;1-29

Demeke T, Jenkins GR. Influence of DNA extraction methods, PCR inhibitors and quantification methods on real-time PCR assay of biotechnology-derived traits. Anal. Bioanal. Chem. 2010;396(6):1977-90

Denison SJ, Lopes EM, D'Costa L, Newman JC. Positive prostate-specific antigen (PSA) results in semen-free samples. Can. Soc. Forensic Sci. J. 2004, 3 7(4), 197-206

Drummer OH. Postmortem toxicology of drugs of abuse. Forensic Sci. Int. 2004;142(2):101-13

Duncan GT, Balamurugan K, Budowle B, Smerick J, Tracey ML. Microvariation at the human D1S80 locus. Int J Legal Med. 1997;110(3):150-4

Easley CJ, Karlinsey JM, Bienvenue JM, Legendre LA, Roper MG, Feldman SH, Hughes MA, Hewlett EL, Merkel TJ, Ferrance JP, Landers JP. A fully integrated microfluidic genetic analysis system with sample-in-answer-out capability. Proc. Natl. Acad. Sci. 2006;103 (51):19272-77

Elliot K, Hill DS, Lambert C, Burroughes TR, Gill P. Use of laser microdissection greatly improves the recovery of DNA from sperm on microscope slides. Forensic Sci. Int. 2003;137(1):28-36

Estes MD, Yang J, Duane B, Smith S, Brooks C, Nordquist A, Zenhausern F. Optimization of multiplexed PCR on an integrated microfluidic forensic platform for rapid DNA analysis. Analyst. 2012;137:5510-19

Farrar JS and Wittwer CT. Extreme PCR: efficient and specific DNA amplification in 15–60 seconds. Clin. Chem. 2015; 61(1): 145-53

Foster A, Laurin N. Development of a fast PCR protocol enabling rapid generation of AmpFlSTR® Identifiler® profiles for genotyping of human DNA. Investig Genet. 2012;3(1):6

Fowler JCS, Harding HWJ, Burgoyne L. A protocol using an alkali Blotting procedure for the analysis of restriction length fragments of human DNA. Adv. Forensic Haemogenet. 1988;2(1): 339-45

Frégeau CJ, De Moors A. Competition for DNA binding sites using Promega DNA IQ™ paramagnetic beads. Forensic Sci. Int. Genet. 2012;6(5):511-22

French DJ, Archard CL, Anderson MT, McDowell DG. Ultra-rapid DNA analysis using HyBeaconTM probes and direct PCR amplification from saliva. Mol Cell Probes. 2002;16(1):319-26

Fuehrer HP, Fally MA, Habler VE, Starzengruber P, Swoboda P, Noedl H. A novel nested direct PCR technique for malaria diagnosis using filter paper samples. J. Clin. Microbiol. 2011;49(4): 1628-30

Funes-Huacca ME, Opel K, Thompson R, McCord BR. A comparison of the effects of PCR inhibition in quantitative PCR and forensic STR analysis. Electrophoresis. 2011;32(1):1084-9

Gao YG, Su SY, Robinson H, Padmanabham S, Lim L, McCary BS, Edmondson SP, Shriver JW, Wang AH. The crystal structure of the hyperthermophile chromosomal protein Sso7d bound to DNA. Nat. Struct. Mol. Biol. 1998;5(1):782-6

Getz EB, Xiao M, Chakrabarty T, Cooke R, Selvin PR. A comparison between the sulfhydryl reductants tris (2-carboxyethyl) phosphine and dithiothreitol for use in protein biochemistry. Anal Biochem. 1999;273(1):73-80

Ghadessy FJ, Holliger P. Compartmentalized self-replication: a novel method for the directed evolution of polymerases and other enzymes. Methods Mol Biol. 2007;352(1):237-48

Gibson-Daw GC. Forensic Analysis of Loci Obtained From Touch DNA Using the ABI Prism 310 Genetic Analyser. University of East Anglia. 2013.

Gibson-Daw G, Albani P, Gassmann M, McCord B. Rapid microfluidic analysis of a Y-STR multiplex for screening of forensic samples. Anal Bioanal Chem. 2017;409(4): 939-47

Giese H, Lam R, Selden R, Tan E. Fast multiplexed polymerase chain reaction for conventional and microfluidic short tandem repeat analysis. J Forensic Sci. 2009;54(6):1287-96

Giglio S, Monis PT, Saint CP. Demonstration of preferential binding of SYBR Green I to specific DNA fragments in real-time multiplex PCR. Nucleic Acids Res. 2003;31(22):e136-6

Gill P, Jeffreys AJ, Werrett DJ. Forensic Application of DNA "fingerprints". Nature. 1985;318(12):577-9

Gill P. Application of low copy number DNA profiling. Croat Med J. 2001;42(3):229-32

Giordano BC, Ferrance J, Swedberg S, Hühmer AFR, Landers JP. Polymerase chain reaction in polymeric microchips: DNA amplification in less than 240 seconds. Anal Biochem. 2001;291:124-32

Goedecke N, Mckenna B, El-Difrawy S, Carey L, Matsudaira P, Ehrlich D. A high-performance multilane microdevice system designed for the DNA forensics laboratory. Electrophoresis. 2004;25(10–11):1678–86

Goodwin DC, Lee SB. Microwave miniprep of total genomic DNA from fungi, plants, protists and animals for PCR. BioTechniques. 1993;15(3):438, 441-2, 444

Greenspoon SA, Scarpetta MA, Drayton ML, Turek SA. QIAamp spin columns as a method of DNA isolation for forensic casework. J Forensic Sci. 1998;43(5):1024-30

Greenspoon SA, Ban JD, Pablo L, Crouse CA, Validation and implementation of the PowerPlex 16 BIO System STR multiplex for forensic casework. J Forensic Sci. 2004;49(1):71-80

Greg W. Et Tu Brute? – The Murder of Caesar and Political Assassination. Brécourt Academic. 2006

Grossman PD, Soane DS. Capillary electrophoresis of DNA in entangled polymer solutions. J Chromatogr. 1991;559(1-2):257-66. a

Grossman PD, Soane DS. Experimental and theoretical studies of DNA separations by capillary electrophoresis in entangled polymer solutions. Biopolymers. 1991;31(10):1221-8. b

Hares DR. Expanding the CODIS core loci in the United States. Forensic Sci. Int:Gen. 2012;6(1): e52-4

Hares DR. Selection and implementation of expanded CODIS core loci in the United States. Forensic Sci. Int:Gen. 2015;7(1): 33-34

Hartl D, Jones E. Genetics: Analysis of Genes and Genomes. 7th Edition. 2008

Hartzell B, Graham K, McCord B. Response of short tandem repeat systems to temperature and sizing methods. Forensic Sci. Int. 2003;133(3):228-34

Hartzell B, McCord B. Effect of divalent metal ions on DNA studied by capillary electrophoresis. Electrophoresis. 2005; 26(1):1046-56

Hauge XY, Litt M. A study of the origin of 'shadow bands' seen when typing dinucleotide repeat polymorphisms by the PCR. Hum Mol Gen. 1993; 2(4):411-5

Hedman J, Dalin E, Rasmusson B, Ansell R. Evaluation of amylase testing as a tool for saliva screening of crime scene trace swabs. Forensic Sci. Int. Genet. 2011;5(3): 194-8

Heller MJ, Guttman A. Integrated Microfabricated Biodevices: Advanced Technologies for Genomics, Drug Discovery, Bioanalysis, and Clinical Diagnostics. Marcel Dekker Incorporated: 2002

Heid CA, Stevens J, Livak KJ, Williams PM. Real time quantitative PCR. Genome Res. 1996;6(10):986-94

Higuchi R, Fockler C, Dollinger G, Watson R. Kinetic PCR analysis: real-time monitoring of DNA amplification reactions. Biotechnology. 1993;11(9):1026-30

Hill CR. Capillary electrophoresis and 5-channel LIF detection of a 26plex autosomal STR assay for human identification. Methods Mol Biol. 2012;830:17-29

Hochmeister MN, Budowle B, Rudin O, Gehrig C, Borer U, Thali M, Dirnhofer R. Evaluation of prostate-specific antigen (PSA) membrane test assays for the forensic identification of seminal fluid. J Forensic Sci. 1999, 44(5):1057-60

Hoff-Olsen P, Mevåg B, Staalstrøm E, Hovde B, Egeland T, Olaisen B. Extraction of DNA from decomposed human tissue: an evaluation of five extraction methods for short tandem repeat typing. Forensic Sci. Int. 1999;105(3):171-83

Hopwood AJ, Hurth C, YangJ, Cai Z, Moran N, Lee-Edghill JG, Nordquist A, Lenigk R, Estes MD, Haley JP, McAlister CR, Chen X, Brooks C, Smith S, Elliott K, Koumi P, Zenhausern F, Tully F. Integrated microfluidic system for rapid forensic DNA analysis: sample collection to DNA profile. Anal. Chem.2010;82(16):6991-99

Horsman KM, Bienvenue JM, Blasier KR, Landers JP. Forensic DNA Analysis on Microfluidic Devices: A Review. J Forensic Sci.2007;52(4): 784–99

Hudlow WR, Buoncristiani MR. Development of a rapid, 96-well alkaline based differential DNA extraction method for sexual assault evidence. Forensic Sci. Int. Genet. 2012;6(1):1-16

Huh JY, Jang Sj, Moon DS, Park YJ. A rapid long PCR-direct sequencing analysis for ABO genotyping. Ann Clin Lab Sci. 2011;41(4):340-5

Imanaka T, Takagi M. Evolution of PCR enzymes (towards a better PCR system based on a KOD DNA polymerase). Chin. J. Chem. Eng. 2001;32(3):277-88

Innis MA, Sninsky JJ, Gelfand DH. Pcr Applications: Protocols for Functional Genomics. Academic Press. 1999; 33-47

Jeffreys AJ, Wilson V, Thein SL. Individual-specific 'fingerprints' of human DNA. Nature. 1985; 316(6023): 76-9

Jeffreys AJ, Wilson V, Thein SL. Hypervariable 'minisatellite' regions in human DNA. Nature. 1985; 314(6006):67-73

Jeffreys AJ, Monckton DG, Tamaki K, Neil D L, Armour JA, MacLeod A, Collick A, Allen M, Jobling M. Minisatellite variant repeat mapping: application to DNA typing and mutation analysis. Exs. 1993; 67:125-39

Jeffreys AJ, Pena SD. Brief introduction to human DNA fingerprinting. Exs. 1993;67: 1-20.

Jensen M. Use of the Agilent 2100 Bioanalyzer. Applied Biosystems. 1990.

Jorgenson JW, Lukacs KD. Zone Electrophoresis in Open-Tubular Glass Capillaries. Clin Chem. 1981;27(9):1551-3

Kan CW, Fredlake CP, Doherty EAS, Barron AE. DNA sequencing and genotyping in miniaturized electrophoresis systems. Electrophoresis. 2004;25(21-22):3564-88

Kelkar YD, Eckert KA, Chiaromonte F, Makova KD. A matter of life or death: how microsatellites emerge in and vanish from the human genome. Genome Res. 2011;21(12):2038-48

Kermekchiev MB, Kirilova LI, Vail EE, Barnes WM. Mutants of Taq DNA polymerase resistant to PCR inhibitors allow DNA amplification from whole blood and crude soil samples. Nucleic Acids Res. 2009;37(5):e40

Khandurina J, McKnight TE, Jacobson SC, Waters LC, Foote RS, Ramsey JM. Integrated system for rapid PCR-based DNA analysis in microfluidic devices. Anal. Chem. 2000;72:2995-00

Kim YT, Lee D, Heo HY, Sim JE, Woo KM, Kim DH, et al. Total integrated slidable and valveless solid phase extraction-polymerase chain reaction-capillary electrophoresis microdevice for mini Y chromosome short tandem repeat genotyping. Biosens Bioelectron. 2016;78:489-96

Kipps AE, Whitehead PH. The significance of amylase in forensic investigations of body fluids. Forensic Sci. 1975;6(3):137-44

Kist FG, Tomsey CS, Glessner AL, Mihalacki LR, Long TM, Heidebrecht BJ, Braunstein CA, Freeman DA, Soberalski C, Nathan B, Amin AS, Douglas EK, Schumm JW.

Kline MC, Hill CR, Decker AE, Butler JM. STR sequence analysis for characterizing normal, variant, and null alleles. Forensic Sci. Int. Gen. 2011;5(4):329-32

Koblinsky L, Liotti TF, Oeser-Sweat J, Watson J, Witkowski J. DNA: Forensic and Legal Applications. Wiley-Interscience. 2004; 1st Edition

Kopka J, Leder M, Jaureguiberry SM, Brem G, Boselli GO. New Optimized DNA Extraction Protocol for Fingerprints Deposited on a Special Self-Adhesive Security Seal and Other Latent Samples Used for Human Identification. J Forensic Sci. 2011;56:1235-40

Krenke BE, Tereba A, Anderson SJ, Buel E, Culhane S, Finis CJ, Tomsey CS, Zachetti JM, Masibay A, Rabbach DR, Amiott EA, Sprecher CJ. Validation of a 16-locus fluorescent multiplex system. J Forensic Sci. 2002;47(4):773-85

Kubista M, Andrade JM, Bengtsson M, Forootan A, Jonak J, Lind K, Sindelka R, Sjoback R, Sjogreen B, Strombom L, Stahlberg A, Zoric N. The real-time polymerase chain reaction. Mol Aspects Med. 2006;27(2-3):95-125

Kuffner CA Jr, Marchi E, Morgado JM, Rubio CR. Capillary Electrophoresis and Daubert: Time for Admission. J. Anal. Chem. 1996;68:241-6

Kupfer DM, Huggins M, Cassidy B, Vu N, Burian D. Canfield DV. A rapid and Inexpensive PCR-Based STR Genotyping Method for Identifying Forensic Samples. Federal Aviation Administration. 2006; DOT/FAA/AM-06/14, 1-20

Lagally ET, Emrich CA, Mathies RA. Fully integrated PCR-capillary electrophoresis microsystem for DNA analysis. Lab Chip. 2001;1(2):102-7

Landers JP. Handbook of Capillary And Microchip Electrophoresis And Associated Microtechniques. CRC Press: 2008.

Larkin A, Harbison S. An improved method for STR analysis of bloodstained denim. Int J Legal Med. 1999;112(6):388-90

Laurin N, Frégeau C. Optimization and validation of a fast amplification protocol for AmpFℓSTR® Profiler Plus® for rapid forensic human identification. Forensic Sci Int Genet. 2012; 6(1):47-57

Leclair B, Frégeau CJ, Bowen KL, Fourney RM. Systematic analysis of stutter percentages and allele peak height and peak area ratios at heterozygous STR loci for forensic casework and database samples. J Forensic Sci. 2004;49(5):968-80

Leclercq S, Rivals E, Jarne P. DNA slippage occurs at microsatellite loci without minimal threshold length in humans: a comparative genomic approach. Genome Biol Evol. 2010;2:325–35

Lee HC, Ladd C. Preservation and collection of biological evidence. Croat Med J. 2001;42(3):225-8

Lee HY, Park MJ, Kim NY, Yang WI, Shin KJ. Rapid direct PCR for ABO blood typing. J Forensic Sci. 2011;56(S1):S179-82

Lee HC, Palmbach TM, Primorac D, Andelinovic S. Collection and Preservation of Physical Evidence. Forensic DNA Applications: An Interdisciplinary Perspective. 2014:253

Lincoln CA, McBride PM, Turbett GR, Garbin CD, MacDonald EJ. The use of an alternative light source to detect semen in clinical forensic medical practice. J Clin Forensic Med. 2006;13(4):215-8

Lins AM, Micka KA, Sprecher CJ, Taylor JA, Bacher JW, Rabbach DR, Bever RA, Creacy SD, Schumm JW. Development and population study of an eight-locus short tandem repeat (STR) multiplex system. J Forensic Sci. 1998;43(6):1168-80

Liu P, Seo TS, Beyor N, Shin KJ, Scherer JR, Mathies RA, Integrated portable polymerase chain reaction-capillary electrophoresis microsystem for rapid forensic short tandem repeat typing, Anal. Chem. 79 (2007) 1881-89

Liu P, Mathies RA. Integrated microfluidic systems for high-performance genetic analysis. Trends Biotechnol. 2009;27(10):572-81

Liu P, Li X, Greenspoon SA, Scherer J R, Mathies R A. Integrated DNA purification, PCR, sample cleanup, and capillary electrophoresis microchip for forensic human identification. Lab. Chip. 2011; 11(6): 1041-48

Lounsbury JA, Karlsson A, Miranian DC, Cronk SM, Nelson DA, Li J, Haverstick DM, Kinnon P, Saul DJ, Landers JP. From sample to PCR product in under 45 minutes: a polymeric integrated microdevice for clinical and forensic DNA analysis. Lab Chip. 2013;13(7):1384-93

Lui C, Cady N.C, Batt C.A. Nucleic Acid-Based Detection of Bacterial Pathogens Using Integrated Microfluidic Platform Systems. Sensors. 2009;9(1):3713-44

Mamedov TG, Pienaar E, Whitney SE, TerMaat JR, Carvill G, Goliath R, Viljoen HJ. A fundamental study of the PCR amplification of GC-rich DNA templates. Comput. Biol. Chem. 2008;32(6):452-7

Mannucci A, Sullivan KM, Ivanov PL, Gill P. Forensic application of a rapid and quantitative DNA sex test by amplification of the X-Y homologous gene Amelogenin. Int J Legal Med. 1994;106(4):190-3.

Mapes AA, Kloosterman AD, Poot CJ, van Marion V. Objective data on DNA success rates can aid the selection process of crime samples for analysis by rapid mobile DNA technologies. Forensic Sci Int. 2016;264:28-33

Marmur J. A procedure for the isolation of deoxyribonucleic acid from micro-organisms. J. Mol. Biol. 1961;3(2):208-18

Mazzarello P. A unifying concept: the history of cell theory. Nat Cell Biol. 1999; 1(1):E13-5

McBeth M. A Study of the Molecular Basis of Microvariants at the FGA and D21S11 Loci Used in Forensic DNA Testing. Honors scholar thesis. University of Connecticut. 2006

McClintock JT. Forensic Analysis of Biological Evidence: A Laboratory Guide for Serological and DNA Typing. CRC Press. 2014

McCord BR, Jung JM, Holleran EA. High Resolution Capillary Electrophoresis of Forensic DNA Using a Non-Gel Sieving Buffer. J. Liq. Chromatogr. 1993;16(9-10):1963-81

McCusker J, Dawson MT, Noone D, Gannon F, Smith T. Improved method for direct PCR amplification from whole blood. Nucleic Acids Res. 1992;20(24):6747

McNevin D, Wilson-Wilde L, Robertson J, Kyd J, Lennard C. Short tandem repeat (STR) genotyping of keratinised hair: Part 1. Review of current status and knowledge gaps. Forensic Sci. Int. 2005;153(2):237-46

Mercier B, Gaucher C, Feugeas O, Mazurier C. Direct PCR from whole blood, without DNA extraction. Nucleic Acids Res. 1990;18(19); 5908

Mikkers FEP, Everaerts FM, Verheggen TPEM. High-performance zone electrophoresis. J. Chromatogr. 1979;169:11–20

Mitnik L, Carey L, Burger R, Desmarais S, Koutny L, Wernet O, Matsudaira P, Ehrlich D. High-speed analysis of multiplexed short tandem repeats with an electrophoretic microdevice. Electrophoresis. 2002;23(5):719-26

Moretti TR, Baumstark AL, Defenbaugh DA, Keys KM, Smerick JB, Budowle B. Validation of short tandem repeats (STRs) for forensic usage: performance testing of fluorescent multiplex STR systems and analysis of authentic and simulated forensic samples. J Forensic Sci. 2001;46(3):647-60

Moretti TR, Moreno LI, Smerick JB, Pignone ML, Hizon R, Buckleton JS, Bright JA, Onorato AJ. Population data on the expanded CODIS core STR loci for eleven populations of significance for forensic DNA analyses in the United States. J Forensic Sci. Gen. 2016;25(1):175-181

Motz M, Kober I, Girardot C, Loeser E, Bauer U, Moeckel M, Minch E, Voss H, Kilger C, Koegl M. Elucidation of an archael replication protein network to generate enhanced PCR enzymes. J. Biol. Chem. 2002;227(1):16179-88

Mullis KB, Faloona FA. Specific synthesis of DNA in vitro via a polymerase-catalyzed chain reaction. Methods in enzymology. 1987;155: 335-50

Mullis KB. The unusual origin of the polymerase chain reaction. Scientific American. 1990;262 (4):56-61, 64-5

Mullis KB. Target amplification for DNA analysis by the polymerase chain reaction. Annales de biologie Clinique. 1990;48(8):579-82

Mullis K. Dancing Naked in the Mind Field. Vintage. 2000; Reprint Edition

Murali R, Sharkey DJ, Daiss JL, Murthy HM. Crystal structure of Taq DNA polymerase in complex inhibitory Fab: The Fab is directed against an intermediate in the helix-coil dynamics of the enzyme. Proc. Natl. Acad. Sci. U.S.A. 1998;95(1):12562-7

Murray GI. An overview of laser microdissection technologies. Acta Histochem 2007;109(3):171-6

Nagy M, Otremba P, Kruger C, Bergner-Greiner S, Anders P, Henske B, Prinz M, Roewer L. Optimization and validation of a fully automated silica-coated magnetic beads purification technology in forensics. Forensic Sci. Int. 2005;152(1):13-22

National Research Council (US) Committee on DNA Forensic Science: An Update. The Evaluation of Forensic DNA Evidence. National Academies Press (US).1996.

Northrop DM, McCord BR, Butler JM. Forensic Applications of Capillary Electrophoresis. J. Capil. Electrophor. 1994;1:158-68

Nunn S. Touch DNA Collection Versus Firearm Fingerprinting: Comparing Evidence Production and Identification Outcomes. J Forensic Sci. 2013;58(3):601-8

Nussbaum R, McInnes RR, Willard HF. Thompson & Thompson Genetics in Medicine. Elsevier. 2007; 8th Edition

Olaisen B, Stenersen M, Mevåg B. Identification by DNA analysis of the victims of the August 1996 Spitsbergen civil aircraft disaster. Nature Genet. 1997;15(4):402-5

Ogston AG. The spaces in a uniform random suspension of fibres. Trans Farady Soc. 1958;54:1754-57

O'Keefe DS, Dobrovic A. A rapid and reliable PCR method for genotyping the ABO blood group. II:A2 and O2 Alleles. Hum Mutat. 1996;8(1):358-61

Oliva R. Protamines and male infertility. Hum Reprod Update. 2006;12(4):417-35

Olivera BM, Baine P, Davidson N. Electrophoresis of the nucleic acids. Biopolymers. 1964;2(3): 245-57

Opel KL, Chung D, McCord BR. A study of PCR inhibition mechanisms using real time PCR. J Forensic Sci. 2010;55(1):25-33

Panaccio M, Georgesz M, Hollywell C, Lew A. FoLT PCR: a simple PCR protocol for amplifying DNA directly from whole blood. Biotechniques. 1993;14(2):238-43

Panaccio M, Georgesz M, Hollywell C, Lew A. Direct PCR from solid tissues without DNA extraction. Nucleic Acids Res. 1993;21(19):4656

Pang BCM, Cheung BKK. Double swab technique for collecting touched evidence. Leg Med. 2007;9(4):181-4

Paradela ER, Glidewell D, Konotop F, de Carvalho EF, Crouse C. Feasibility of conducting PCR based DNA analysis at the crime scene.   https://www.promega.com/-/media/files/resources/conference-proceedings/ishi-11/oral-presentations/paradela.pdf?la=en. 2006.

Park MJ, Lee HY, Chung U, Kang SC, Shin KJ. Y-STR analysis of degraded DNA using reduced-size amplicons. Int J Legal Med. 2007;121(2):152-7

Park SJ, Kim JY, Yang YG. Direct STR amplification from whole blood and blood or saliva spotted FTA without DNA purification. J Forensic Sci. 2008;53(2):335-41

Pavlov AR, Pavlova NV, Kozyavkin SA, Slesarev AI. Recent developments in the optimization of thermostable DNA polymerases for efficient applications. Trends Biotechnol. 2004;22(5):253-60

Payseur BA, Jing P, Haasl RJ. A genomic portrait of human microsatellite variation. Mol Biol Evol. 2011;28(1):303-12

Pearson H. Genetics: what is a gene? Nature. 2006;441(7092):398-401.

Pierce BA. Genetics: A Conceptual Approach. W. H. Freeman. 2016; 6th Edition

Poon H, Elliott J, Modler J, Fregeau C. The Use of Hemastix® and the Subsequent Lack of DNA Recovery Using the Promega DNA IQTM System. J Forensic Sci. 2009;54(6):1278-86

President's DNA Initiative. Protocol 3.05 Chelex®100 non-differential extraction. DNA analyst training: Laboratory training manual

Promega. PowerPlex® 16 System. Technical Manual. 2016

Qiagen. For automated purification of DNA from forensic and biosecurity samples using EZ1 instruments. EZ1® DNA Investigator® Handbook. 2014

Qiagen. QIAamp® DNA Mini and Blood Mini Handbook. 2016

Qu BY, Wu ZY, Tian XX, Chen K, Fang F. Optimization of direct whole blood PCR amplification with applications on a static thermostat chip. Anal. Bioanal. Chem. 2007;389(1):1499-1504

Ravid-Amir O, Rosset S. Maximum likelihood estimation of locus specific mutation rates in Y-chromosome short tandem repeats. Bioinformatics. 2010;26(18):i440-5

Raymond JJ, van Oorschot RA, Gunn PR, Walsh SJ, Roux C. Trace evidence characteristics of DNA: A preliminary investigation of the persistence of DNA at crime scenes. Forensic Sci. Int. Genet. 2009;4(1):26-33

Redd AJ, Agellon AB, Kearney VA, Contreras VA, Karafet T, Park H, et al. Forensic value of 14 novel STRs on the human Y chromosome. Forensic Sci Int. 2002;130(2):97-111

Rodbard D, Chrambach A. Unified theory for gel electrophoresis and gel filtration. Proc. Natl. Acad. Sci. U.S.A. 1970;65(4):970-7

Romsos EL, Vallone PM. Rapid PCR of STR markers: Applications to human identification. Forensic Sci Int Genet. 2015;18:90-99

Rose DJ, Jorgenson JW. Characterization and automation of sample introduction methods for capillary zone electrophoresis. Anal. Chem. 1998;60(7):642-8

Ruitberg CM, Reeder DJ, Butler JM. STRBase: a short tandem repeat DNA database for the human identity testing community. Nucleic Acids Res. 2001;29(1):320-2

Rychlik W, Spencer WJ, Rhoads R.E. Optimization of the annealing temperature for DNA amplification in vitro. Nucl. Acids. Res. 1990;18(21):6409–12

Saferstein R, Hall P. An Introduction to Forensic Science. Criminalistics. Prentice Hall, Upper Saddle River, New Jersey. 2011;108-12

Sajantila A, Budowle B. Identification of individuals with DNA testing. Ann Med. 1991;23(6): 637-42

Sajantila A, Strom M, Budowle B, Karhunen PJ, Peltonen L. The polymerase chain reaction and post-mortem forensic identity testing: application of amplified D1S80 and HLA-DQ alpha loci to the identification of fire victims. Forensic Sci. Int.1991;51(1):23-34

Sanders CT, Sanchez N, Ballantyne J, Peterson DA. Laser Microdissection Separation of Pure Spermatozoa from Epithelial Cells for Short Tandem Repeat Analysis. J Forensic Sci. 2006;51(4):748-57

Sangeeta A, Sharma AK, Bhattacharyya CN, Chaudhuri K. Generating STR profile from "Touch DNA". J Forensic Leg Med. 2011;18(1):295-8

Satoh Y, Takasaka N, Hoshikawa Y, Osaki M, Ohfuji S, Ito H, Kaibara N, Kurata T, Sairenji T. Pretreatment with restriction enzyme or bovine serum albumin for effective PCR amplification of Epstein-Barr virus DNA in DNA extracted from paraffin-embedded gastric carcinoma tissue. J. Clin. Microbiol. 1998;36(11):3423-5

Schlötterer C. Evolutionary dynamics of microsatellite DNA. Chromosoma. 2009;109(6):365.

Schochetman G, Chin-Yih O, Jones WK. Polymerase Chain Reaction. J Infect Dis. 1988;158(6): 1154-7

Schoell WM, Klintschar M, Mirhashemi R, Pertl B. Separation of sperm and vaginal cells with flow cytometry for DNA typing after sexual assault. Obstet Gynecol. 1999;94(4):623-7

Schoell WM, Klintschar M, Mirhashemi R, Strunk D, Giuliani A, Bogensberger G, Pertl B. Separation of sperm and vaginal cells based on ploidy, MHC class I-, CD45-, and cytokeratin expression for enhancement of DNA typing after sexual assault. Cytometry. 1999;36(4):319-23

Schuelke M. An Economic Method for the Fluorescent Labeling of PCR Fragments. Nat. Biotechnol. 2000;18:233-4

Schumm JW. Why Use a Size Marker and Allelic Ladders in STR Analysis? Geneprint™ Products. 1997.

Scientific Working Group on DNA Analysis Methods (SWGDAM). SWGDAM Interpretation Guidelines for Autosomal STR Typing by Forensic DNA Testing Laboratories. 2010

Scientific Working Group on DNA Analysis Methods (SWGDAM). SWGDAM Validation Guidelines for DNA Analysis Methods. 2012

Shi Y, Anderson RC. High-resolution single-stranded DNA analysis on 4.5 cm plastic electrophoretic microchannels. Electrophoresis. 2003;24(19–20):3371–7

Shi Y. DNA sequencing and multiplex STR analysis on plastic microfluidic devices. Electrophoresis. 2006;27(19):3703–11

Sinha RP, Hader DP. UV-induced DNA damage and repair: a review. Photochem Photobiol Sci 2002;1(4):225-36

Sinha SK, Budowle B, Chakraborty R, Paunovic A, Guidry RD, Larsen C, Lal A, Shaffer M, Pineda G, Sinha SK, Schneida E, Nasir H, Shewale JG. Utility of the Y-STR typing systems Y-PLEX 6 and Y-PLEX 5 in forensic casework and 11 Y-STR haplotype database for three major population groups in the United States. J Forensic Sci. 2004 Jul;49(4):691-700

Smit RN. Accurate Size Comparison of Short Tandem Repeat Alleles Amplified by PCR. BioTechniques.1995;18:122-8

Smith LM, Burgoyne LA. Collecting, archiving and processing DNA from wildlife samples using FTA® databasing paper. BMC Ecol. 2004;4(1):4

Srinivasan K, Morris SC, Girard JE, Kline MC, Reeder DJ. Enhanced detection of PCR products through use of TOTO and YOYO intercalating dyes with laser induced fluorescence--capillary electrophoresis. Appl Theor Electrophor. 1993;3(5):235-9

Stryer L, Berg JM, Tymoczko JL. Biochemistry. 2002; 5th edition

Suetonius, Julius, c. 82.

Sullivan KM, Mannucci A, Kimpton CP, Gill P. A rapid and quantitative DNA sex test: fluorescence-based PCR analysis of X-Y homologous gene Amelogenin. Biotechniques. 1993;15(4):636-8, 640-1

Swerdlow H, Gesteland R. Capillary gel electrophoresis for rapid, high resolution DNA sequencing. Nucleic Acids Res. 1990;18(6):1415-9

Templeton J. Genetic profiling from challenging samples: Direct PCR of touch DNA. Forensic Sci. Int. Genet. 2013;4(1):224-5

Thacker C, Phillips CP, Syndercombe-Court D. Use of FTA cards in small volume PCR reactions. Progress in Forensic Genetics 8 (1999). 2000;8:473-5

Thomas CD, Jacobson SC, Ramsey JM. Strategy for repetitive pinched injections on a microfluidic device. Analytical chemistry. 2004;76(20):6053-7

Thompson R, Zoppis S, McCord B. An overview of DNA typing methods for human identification: past, present, and future. DNA Electrophoresis Protocols for Forensic Genetics. Humana Press 2012; 3-16

Tie J, Uchigasaki S, Haseba T, Ohno Y, Isahai I, Oshida S. Direct and rapid PCR amplification using digested tissues for the diagnosis of drowning. Electrophoresis. 2010;31(1):2411-5

Tilstone WJ, Savage KA, Clark LA. Forensic science: An encyclopedia of history, methods, and techniques. ABC-CLIO. 2006

Urquhart A, Kimpton CP, Downes TJ, Gill P. Variation in short tandem repeat sequences--a survey of twelve microsatellite loci for use as forensic identification markers. Int J Legal Med. 1994;107 (1):13-20

Vallone PM, Butler JM. AutoDimer: A screening tool for primer-dimer and hairpin structures. Biotechniques. 2004;37(2):226–31

Vallone PM, Hill CR, Butler JM. Demonstration of rapid multiplex PCR amplification involving 16 genetic loci. Forensic Sci Int Genet. 2008;3(1):42-5

Vallone PM, Hill CR, Podini D, Butler JM. Rapid amplification of commercial STR typing kits. Forensic Sci Int Genet Suppl Ser. 2009;2(1):111-2

Vandewoestyne M, Deforce D. Laser capture microdissection in forensic research: a review. Int J Legal Med. 2010;124(6):513-21

Verheij S, Harteveld J, Sijen T. A protocol for direct and rapid multiplex PCR amplification on forensically relevant samples. Forensic Sci Int Genet. 2012;6(2):167–75

Virkler K, Lednev IK. Analysis of Body Fluids for Forensic Purposes: From Laboratory Testing to Non-Destructive Rapid Confirmatory Identification at a Crime Scene. Forensic Sci. Int. 2009;188(1-3):1-17

Walker JA, Hedges DJ, Perodeau BP, Landry KE, Stoilova N, Laborde ME, Shewale J, Sinha SK, Batzer MA. Multiplex polymerase chain reaction for simultaneous quantitation of human nuclear, mitochondrial, and male Y-chromosome DNA: application in human identification. Anal. Biochem. 2005; 337(1):89-97

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):506-13

Walsh PS, Fildes NJ, Reynolds R. Sequence analysis and characterization of stutter products at the tetranucleotide repeat locus vWA. Nucleic Acids Res. 1996;24(14):2807-12

Wambaugh J. The Blooding. Bantam. 1995; 60334th edition

Wang Y, Prosen DE, Mei L, Sullivan JC, Finney M, Vander Horn PB. A novel strategy to engineer DNA polymerases for enhanced processivity and improved performance in vitro. Nucleic Acids Res. 2004;32(3):1197–207

Wang DY, Chien CW, Hennessy LK. Rapid STR analysis of single source DNA samples in 2h. J Forensic Sci. 2009;2(1):115-6

Wang DY, Chang CW, Lagace RE, Oldroyd NJ, Hennessy LK. Development and validation of the AmpFLSTR® Identifiler® Direct PCR Amplification Kit: A multiplex assay for the direct amplification of single source samples. J Forensic Sci. 2011;56(4):835-845

Watson JD, Crick FH. Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. Nature. 1953;171(4356):737-8

Watts D, MacBeath JRE. Automated Fluorescent DNA Sequencing on the ABI PRISM 310 Genetic Analyzer. DNA Sequencing Protocols. Methods in Molecular Biology. Humana Press. 2001;167.

Weinberger R. Practical Capillary Electrophoresis. Academic Press, 2000.

Willard JM, Lee DA, Holland MM. Recovery of DNA for PCR amplification from blood and forensic samples using a chelating resin. Methods Mol Biol. 1998;98: 9-18

Williams PE, Marino MA, Del Rio SA, Turni LA, Devaney JM. Analysis of DNA restriction fragments and polymerase chain reaction products by capillary electrophoresis. J. Chromatogr. A 1994;680(2):525-40

Wilson IG. Inhibition and facilitation of nucleic acid amplification. Appl. Environ. Microbiol. 1997;63(10):3741-51

Witt S, Neumann J, Zierdt H, Gébel G, Röscheisen 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):539-47

Wittwer CT, Fillmore GC, Hillyard DR. Automated polymerase chain reaction in capillary tubes with hot air. Nucleic Acids Res. 1989;17(11):4353-57

Wittwer CT, Garling DJ. Rapid cycle DNA amplification: Time and temperature optimization. Biotechniques. 1991;10(1):76-83

Wolfe KA, Breadmore MC, Ferrance JP, Power ME, Conroy JF, Norris PM Landers JP. Toward a microchip-based solid-phase extraction method for isolation of nucleic acids. Electrophoresis. 2002;23(5):727-33

Woolley AT, Lao K, Glazer AN, Mathies RA. Capillary electrophoresis chips with integrated electrochemical detection. Anal Chem.1998;70(4):684–8

Wu DY, Ugozzoli L, Pal BK, Qian J, Wallace RB. The effect of temperature and oligonucleotide primer length on the specificity and efficiency of amplification by the polymerase chain reaction. DNA Cell Biol. 1991;10(3):233-8

Wyman AR, White R. A Highly Polymorphic Locus in Human DNA. Proc. Natl. Acad. Sci. 1980;77:6754-8

Yang YG, Kim JY, Kim SW, Lee SG, Inventors. Methods for performing direct enzymatic reactions involving nucleic acid molecules. Patent WO2006090987 (Rep. of Korea PCT ∕ KR 000457) 2006.

Yeung ES, Kuhr WG. Indirect detection methods for capillary separations. Anal. Chem. 1991;63(5):275A-82A

Young CC, Burghoff RL, Keim LG, Minak-Bernero V, Lute JR, Hinton SM. Polyvinylpyrrolidone-agarose gel electrophoresis purification of polymerase chain reaction-amplifiable DNA from soils. Appl. Environ. Microbiol. 1993;59(6):1972-4

Zhang X, Zhang W, Fan X. Genetic polymorphism of 15 Y chromosomal STR loci and haplotypes of Henan Han population. Hereditas. 2014;151(6):201–8

ZyGEM. PDQeX DNA extraction device instruction manu

APPENDICES

First Copyright Permission

Extract from 'Letters of Note' complied by Shaun Usher published by Canongate Books in 2016. ISBN 978-1-78211-928-9 First published in 2013 © Shaun Usher – permission granted Feb 2018.

Second Copyright Permission

Original letter written by Francis Crick (1916- 2004) on 19th March 1953. © Family Member - Francis Crick – permission granted Feb 2018.

Third Copyright Permission

Circulation/distribution	>50,000
Author of this Springer	yes
Nature content	
Title	Dr.
Instructor name	Bruce McCord
Institution name	Florida International University
Expected presentation date	Mar 2018
Portions	Full article
Requestor Location	Florida International University
	11200 SW 8th St
	MIAMI, FL 33199
	United States
	Attn: Florida International University
Billing Address	Florida International University
	11200 SW 8th St
	MIAMI, FL 33199
	United States
	Attn: Florida International University
Total	0.00 USD

Terms and Conditions – see later

Forth Copyright Permission

Author of this Springer Nature content	yes
Title	Dr.
Instructor name	Bruce McCord
Institution name	Florida International University
Expected presentation date	Mar 2018
Portions	Full article
Requestor Location	Florida International University
	11200 SW 8th St
	MIAMI, FL 33199
	United States
	Attn: Florida International University
Billing Address	Florida International University
	11200 SW 8th St
	MIAMI, FL 33199
	United States
	Attn: Florida International University
Total	0.00 USD

Springer Nature Terms and Conditions for RightsLink Permissions

Springer Customer Service Centre GmbH (the Licensor) hereby grants you a nonexclusive, world-wide licence to reproduce the material and for the purpose and requirements specified in the attached copy of your order form, and for no other use, subject to the conditions below:

1. The Licensor warrants that it has, to the best of its knowledge, the rights to license reuse of this material. However, you should ensure that the material you are requesting is original to the Licensor and does not carry the copyright of another entity (as credited in the published version).

    If the credit line on any part of the material you have requested indicates that it was reprinted or adapted with permission from another source, then you should also seek permission from that source to reuse the material.

2. Where print only permission has been granted for a fee, separate permission must be obtained for any additional electronic re-use.

3. Permission granted **free of charge** for material in print is also usually granted for any electronic version of that work, provided that the material is incidental to your work as a whole and that the electronic version is essentially equivalent to, or substitutes for, the print version.

4. A licence for 'post on a website' is valid for 12 months from the licence date. This licence does not cover use of full text articles on websites.

5. Where **'reuse in a dissertation/thesis'** has been selected the following terms apply:

Print rights for up to 100 copies, electronic rights for use only on a personal website or institutional repository as defined by the Sherpa guideline (www.sherpa.ac.uk/romeo/).

6. Permission granted for books and journals is granted for the lifetime of the first edition and does not apply to second and subsequent editions (except where the first edition permission was granted free of charge or for signatories to the STM Permissions Guidelines http://www.stm-assoc.org/copyright-legal-affairs/permissions/permissions-guidelines/), and does not apply for editions in other languages unless additional translation rights have been granted separately in the licence.

7. Rights for additional components such as custom editions and derivatives require additional permission and may be subject to an additional fee. Please apply to Journalpermissions@springernature.com/bookpermissions@springernature.com for these rights.

8. The Licensor's permission must be acknowledged next to the licensed material in print. In electronic form, this acknowledgement must be visible at the same time as the figures/tables/illustrations or abstract, and must be hyperlinked to the journal/book's homepage.

9. Use of the material for incidental promotional use, minor editing privileges (this does not include cropping, adapting, omitting material or any other changes that affect the meaning, intention or moral rights of the author) and copies for the disabled are permitted under this licence.

10. Minor adaptations of single figures (changes of format, colour and style) do not require the Licensor's approval.

VITA

GEORGIANA CHARLOTTE GIBSON-DAW

Born, Northampton, England

Education

Fall 2006 - Spring 2008 Baxillerato (Spanish High school) ending in Selectividad Exam

IES No.1 Secondary Institute, Javea Spain. Subjects included Chemistry, Physics, and Biology.

Fall 2009 – Summer 2013 BS/MChem (Masters of Chemistry with Forensic Science) (Integrated masters and undergrad program) University of East Anglia, Norwich, England.

MSProject: Detection of secondary transfer of touch DNA using the ABI 310 Genetic Analyser and MALDI-TOF.

Fall 2010- Summer 2013 Student representative School of Chemistry school board, UEA.

Summer of 2011Nuffield bursary awarded for summer research project.
Summer Project: Synthesis of a polymer resin used to separate Strontium from Lead in forensic samples, then be analysed by stable isotopes, or to allow Strontium dating. This was performed using 18-C-6 crown ether as the metal ion selective ligand.

Spring 2012 Exchange program to FIU in Florida US.

Fall 2013 – Spring 2018 PhD/MSFS Chemistry.

Spring 2015-Summer 2016 President of Graduate Student Organization (GSO).

Spring 2016- Fall 2017 NIJ Grant, Award number: 2015-R2-CX-K038.

Spring 2018- Fall 2018 FIU Dissertation Year Fellowship.

Technical Skills

- Agilent Bioanalyzer 2100 and associated software
- ABI 310, ABI 3130 and associated software
- Thermocyclers both normal and Rapid (Philisa)
- Thermal Gradient thermocycler
- Performing DNA extractions by PCIA, EZ1 extraction robot, Chelex and FTA paper

- DNA quantification by real time PCR
- Knowledge of Microsoft Office

Teaching Experience

Peer assisted learning (PAL) mentor	Fall 2012- Summer 2013
General chemistry 1 and 2 Teaching Assistant. Undergraduate level/FIU	Fall 2013 - Summer 2014
Forensic Chemistry Teaching Assistant. Undergraduate level/FIU	Fall 2014- Fall 2015

PUBLICATIONS AND PRESENTATIONS

- Gibson-Daw, Georgiana. Masters thesis "Forensic Analysis of Loci Obtained From Touch DNA Using the ABI Prism 310 Genetic Analyser". University of East Anglia. 2013
- Lee, S.B.; Aboud, M.; Hye Hyun Oh; Cooper, D.; Duncan, D.; Gibson-Daw, Georgiana; McCord, B. The development of new enzymes for rapid direct amplification of DNA and its application in presumptive testing. AAFS, Orlando, Fl, USA 2015 (Poster Presentation)
- Gibson-Daw, Georgiana MChem; Albani, P.; Gassmann, M.; McCord, B. Rapid PCR of an YSTR multiplex as a screening tool for presence of male DNA. AAFS, Las Vegas,NV, USA, 2016 (Poster Presentation)
- Gibson-Daw, Georgiana MChem, MSFS; Bruce McCord. Ultrahigh speed PCR: method for obtaining STR based genotypes in 8 minutes or less. AAFS, New Orleans, LA, USA, 2017 (Oral Presentation)
- Gibson-Daw, Georgiana MChem, MSFS; Albani, P.; Gassmann, M.; McCord, B. Rapid microfluidic analysis of a Y-STR multiplex for screening of forensic samples. Anal. Bioanal. Chem. February 2017, Volume 409, Issue 4, pp 939–947
- Thomas K, Yu E, Garrison K, Gibson-Daw, Georgiana, MSFS, Roig M, McCord B, Lee SB. Rapid, microwave-accelerated DNA extraction from saliva, semen and hair for downstream PCR typing. ISHI, Seattle, WA, USA 2017 (Poster Presentation)
- Gibson-Daw, Georgiana MChem, MSFS; Crenshaw, K.; McCord, B. Optimization of ultrahigh speed multiplex PCR for forensic analysis. Anal. Bioanal. Chem. January 2018, Volume 410, Issue 1, pp 235-245. doi: 10.1007/s00216-017-0715-x.