Extraction, sequencing and bioinformatics analysis of DNA from dried blood spots

By

Phiwokuhle Bulelwa Mjoli

Submitted in fulfillment of the academic requirements for the degree of Master of Science in the School of Life Sciences, Department of Biological and Conservational Sciences, University of KwaZulu-Natal, Durban

March 2016

As the candidate's supervisor I have approved this thesis/dissertation for submission.

Name: **Dr. Paula Sommer** Signed:

(**Supervisor**) Date: **23 March 2016**

Name: **Prof Tulio de Oliveira** Signed:

(**Co-supervisor**) Date: **28 March 2016**

Extraction, sequencing and bioinformatics analysis of DNA from dried blood spots

By

Phiwokuhle Bulelwa Mjoli

Submitted in fulfillment of the academic requirements for the degree of Master of Science in the School of Life Sciences, Department of Biological and Conservational Sciences, University of KwaZulu-Natal, Durban

March 2016

As the candidate's supervisor I have approved this thesis/dissertation for submission.

Name: **Dr. Paula Sommer** Signed:

(**Supervisor**) Date: **23 March 2016**

Name: **Prof Tulio de Oliveira** Signed:

(**Co-supervisor**) Date: **28 March 2016**

ABSTRACT

The Africa Centre for Health and Population Studies has a demographic surveillance site 200km north of Durban at the Umkhanyakude district. In this setting, blood samples are routinely collected and used for the diagnosis of HIV infection. In this setting and in many settings in Africa, samples are normally collected and stored on filter paper as dried blood spots (DBS) as those are easy to transport and store. DNA can be isolated/extracted from DBS and used for viral and/or host genomic analysis. The aim of this work was to extract DNA from DBS of sufficient quality and yield that could be used for subsequent analysis. Specifically, we aimed to perform host HLA genotyping from DBS as HLA type has an impact on HIV-1 replication levels.

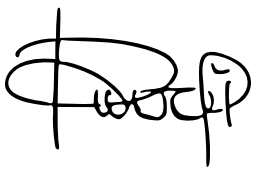
Dried blood spots were prepared from anonymous samples that are commonly used for the validation of new laboratory methods at the Africa Centre Virology Laboratory in Durban. The QIAamp DNA Mini Kit method was optimised to isolate DNA from DBS. DNA levels were quantified using the Qubit 2.0 Fluorometer (Qubit Assay). Polymerase chain reactions (PCR) were used to amplify the HLA Class 1A, 1B and 1C loci and the PCR products were purified using the Pure Link Purification Kit (Invitrogen Life Technologies). Sanger sequencing techniques were used to genotype the HLA's PCR products. Assign^{TM} ATF Softwere v1.5 was used to detect HLA allele variations in the consensus sequences produced through Sanger sequencing.

The DNA yield that could be extracted from the DBS was low. This was most probably due to the low quantity of blood that can be stored in one DBS. Despite the relatively low DNA yields sequencing of the target gene (HLA Class 1A, 1B and 1C) was successful using Sanger Sequencing and variations in the majority of the HLA alleles were detected. This MSc study shows that it is possible to sequence HLA loci directly from DBS.

PREFACE

The experimental work described in this dissertation was carried out in the School of Life Sciences, Department of Biological and Conservation Sciences, University of KwaZulu-Natal, Westville and at the Africa Centre for Health and Population studies Genomic Lab, Doris Duke Medical Research Institute at the Nelson R. Mandela Medical School of Medicine, Durban from April 2014 to March 2016, under the supervision of Dr. Paula Sommer and Prof Tulio de Oliveira.

These studies represent original work by the author and have not otherwise been submitted in any form for any degree or diploma to any tertiary institution. Where use has been made of the work of others it is duly acknowledged in the text.



23 March 2016

Phiwokuhle Bulelwa Mjoli

Date

PLAGIARISM DECLARATION

I, **Phiwokuhle Bulelwa Mjoli**, declare that:

1. The research reported in this thesis, except where otherwise indicated, is my original research.

2. This thesis has not been submitted for any degree or examination at any other university.

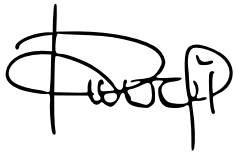
3. This thesis does not contain other persons' data, pictures, graphs or other information, unless specifically acknowledged as being sourced from other persons.

4. This thesis does not contain other persons' writing, unless specifically acknowledged as being sourced from other researchers. Where other written sources have been quoted, then:

   a. Their words have been re-written but the general information attributed to them has been referenced

   b. Where their exact words have been used, then their writing has been placed in italics and inside quotation marks, and referenced.

5. This thesis does not contain text, graphics or tables copied and pasted from the Internet, unless specifically acknowledged, and the source being detailed in the thesis and in the References sections.

Signed:  Date: **23 March 2016**

Phiwokuhle B. Mjoli

TABLE OF CONTENTS

ABSTRACT ..... **ii**
PREFACE ..... **iii**
PLAGIARISM DECLARATION ..... **iv**
TABLE OF CONTENTS ..... **v**
LIST OF TABLES ..... **vii**
LIST OF FIGURES ..... **ix**
ABBREVIATIONS ..... **x**
ACKNOWLEGDEMENTS ..... **xii**
1. INTRODUCTION **1**
1.1 Human Immunodeficiency Virus (HIV) 1
1.2 Role of host genome in HIV infection 3
1.3 Human Leukocyte Antigen (HLA) 3
1.3.1 Role of the HLA Class I in HIV-1 infection/progression 6
1.4 Dried Blood Spots 9
1.4.1 DNA extraction from Dried Blood Spots 11
1.5 DNA Sequencing 17
1.5.1 Sanger Sequencing 17
1.5.2 Sanger sequencing with capillary electrophoresis 18
1.6 Aim and objectives 19
2. MATERIALS AND METHODS **21**
2.1 Dried blood spots 21
2.2 DNA Extraction 21
2.3 DNA quantification 22
2.4 Polymerase chain reactions 23
2.4.1 PCR Step 1 23
2.4.2 PCR Step 2 25
2.5 Gel electrophoresis 26
2.6 PCR product purification 26
2.7 Sequencing 27
2.7.1 Sequencing reaction 27
2.7.2 Sequencing cleanup 28
2.8 Bioinformatics 29
2.8.1 Sequence Quality control 29
2.8.2 De novo Assembly 34
2.8.3 PubMed NCBI BLAST 35
2.8.4 HLA sequence analysis 36
3. RESULTS **37**
3.1 Evaluation of DNA Extraction 37
3.2 PCR Amplification and PCR product purification 38
3.3 Sequence quality control 40
3.3.1 Assembly 40
3.3.2 PubMed NCBI BLAST 42
3.3.3 HLA typing analysis 45
3.3.4 PubMed NCBI BLAST result comparison with the Assign^{TM} ATF Software v1.5 46
4. DISCUSSION **49**
4.1 DNA extraction 49
4.2 Amplification of the HLA Class 1A, 1B and 1C loci. 50
4.3 Sequence quality control and HLA typing analysis 51
4.3.1 Sequence quality control 51
4.3.2 HLA typing analysis 52
4.3.3 PubMed NCBI BLAST result comparison with the Assign^{TM} ATF Software v1.5 53
4.4 Conclusion and Future studies 55
5. APPENDIX **56**
A. Protocols used in the optimisation of DNA extraction from dried blood spots 56

A.1 DNA extraction using the Prepfiler Forensic DNA extraction Kit from cheek epithelium and DBS 56
A.2 DNA extraction using the QIAamp® DNA Mini Kit from whole blood- generation of a positive control for trouble shooting 58
A.3 DNA extraction from epithelial tissue using the QIAamp® DNA Mini Kit 59
A.4. Sequence Assembly and Sequence Quality control 59
A.4.1 Steps followed to construct consensus sequences using Geneious R8 v8.1.8 59
A.4.2 Sequence quality control using NCBI BLAST 61
B. General Recipes 62
C. PCR Optimization 64
C.1 PCR optimization performed at the School of Life Sciences Molecular Biology Lab 64
C.2 PRC optimization performed at the Africa Centre for Health and Population Studies Genomics Lab (Durban) 109
6. REFERENCES 125

LIST OF TABLES

Table 1.1: Results obtained from various publications based on the extraction of DNA from DBS using different extraction methods and downstream analyses…….13

Table 2.1: Primer sequences used to amplify the HLA Class 1 A, 1B and 1C loci in step 1 of the nested PCR……..……23
Table 2.2: Reagent volumes used for PCR amplification in step 1 of the nested PCR……24
Table 2.3: Programmed thermal cycles used for PCR amplification in step 1 of the nested PCR……...24
Table 2.4: Primer sequences used to amplify HLA Class 1 A, 1B and 1C loci in step 2 of the nested PCR……25
Table 2.5: Reagent volumes used for PCR amplification in step 2 of the nested PCR……25
Table 2.6: Programmed thermal cycles used for PCR amplification in step 2 of the nested PCR……...25
Table 2.7: Primer sequences used to generate the four fragments of the HLA Class 1 A, 1B and 1C loci used in Sanger sequencing…….27
Table 2.8: The volumes of each reagent used in preparing fresh sodium acetate and ethanol for sequence cleanup…….28

Table 2.9: Quality and length of HLA Class 1A Sanger sequencing reads before and after editing using Geneious R8 v8.1.8…….31

Table 2.10: Quality and length of HLA Class 1B Sanger sequencing reads before and after editing using Geneious R8 v8.1.8……..…32

Table 2.11: Quality and length of HLA Class 1C Sanger sequencing reads before and after editing using Geneious R8 v8.1.8…….....33

Table 3.1: The DNA yield (ng /μL) in samples extracted from Dried Blood spots using the QIAamp® DNA Mini Kit (QIAGEN) measured using the Qubit 2.0 Flourometer (Qubit Assay)……...37

Table 3.2: Quality of the consensus sequences produced by mapping the four sequencing primer reads from Sanger sequencing using Geneious R8 v8.1.8 for each of the 10 donor sample under the three HLA alleles (HLA Class 1A, 1B and 1C)……...41

Table 3.3: Details of the sequences that match the GenBank non redundant (NR) database in BLAST……..43

Table 3.4: Results from the Assign™ ATF Software v1.5 (Conexio Genomics, Australia) used to genotype and detect variation in HLA Class 1A, 1B and 1C alleles for the 10 sequenced samples. The numbers represent the allele and its subtype that each sequence matched to in the HLA Class 1 gene…...…….……45

Table 3.5: Details of the sequences that the NCBI BLAST matches were identical to the Assign™ ATF Software v1.5 classification…….47

LIST OF FIGURES

Figure 1.1: Anatomy of the Human Immunodeficiency Virus (HIV)……...2
Figure 1.2: Schematic drawing of the HLA locus on chromosome 6…….5
Figure 1.3: An example of how blood is stored as dried blood spots on specialized filter paper……..10
Figure 1.4: Diagram showing the process followed in Sanger sequencing when using dye-terminators……19

Figure 2.1: Programmed thermal cycles used for sequencing reaction……28
Figure 2.2: Image showing a read obtained from Sanger sequencing being edited by trimming off the ends, which are of bad quality. Bases from no. 62 to 925 form the section of the read that will be used in the analysis of interest…….. 30
Figure 2.3: Four reads of sample #1 (HLA Class 1A) mapped into a consensus sequence using the de novo option in Geneious R8 v 8.1.8……..34
Figure 2.4: Four reads of sample #1 (HLA Class 1C) mapped to a known reference gene as a guide in constructing the consensus sequence……..35
Figure 2.5: Image showing the evaluation specificity used in PubMed NCBI BLAST to align the newly constructed consensus sequences against the local sequence database from previous runs……36

Figure 3.1: Image of an agorose gel electrophoresis showing the presence of DNA in samples extracted from Dried Blood Spots and the positive control……..38

Figure 3.2: Image of an agarose gel showing PCR amplification of HLA Class 1A, 1B and 1C using a sample volume of 5 µl in step 1 and 2.5 µl in step 2 and a T_m of 63°C in step 1 and 65°C in step 2……39

Figure 3.3: Electrophoresis gel showing PCR amplification of HLA Class 1A, 1B and 1C using a sample volume of 2.5 µl in step 2 and a T_m of 63°C in step 1 and 65°C in step 2 after purification using the Pure Link Purification Kit (Invitrogen Life Technologies)……...40

ABBREVIATIONS

+ve:	positive control
-ve:	negative control
°C:	degrees celcius
µL:	microlitre(s)
AIDS:	acquired immunodeficiency syndrome
B cells:	B-lymphocytes
Bp:	Base pair
cHCMV:	Congenital human cytomegalovirus
CD (4 or 8):	Cluster of differentiation
CMV:	Cytomegalovirus
dNTP:	Deoxynucleosidetriphosphate
ddNTP:	Di-deoxynucleosidetriphosphate
DBS:	Dried blood spots
dH2O:	Distilled water
di.H2O:	Deionized water
DNA:	Deoxyribonucleic acid
EthOH:	Ethanol
gDNA:	Genomic DNA
GWAS:	Genomic-wide association study
HAART:	Highly Active Anti-Retroviral Therapy
HIV:	Human immunodeficiency virus
HLA:	Human leukocyte antigen
IFN:	Interferon
Kb:	Kilo base
LTNP:	Long-term non-progressors
MHC:	Major histocompatibility complex
Min:	Minute(s)
Mg:	Microgram(s)
MW:	Molecular weight marker
nPCR:	Nested polymerase chain reaction
NCBI:	National Centre for biotechnology information
Ng:	Nano gram(s)

NR:	Non redundant
pVL:	Plasma viral load
PCR:	Polymerase chain reaction
Sec:	Seconds(s)
SSCP:	Single-stranded conformational polymorphism
T-cells:	T lymphocytes
TNF:	Tumor necrosis factor
UV:	Ultraviolet
V:	Volts
VL:	Viral load
wgaDNA:	Whole genome amplified DNA

ACKNOWLEGDEMENTS

Firstly, I would like to thank my supervisor Dr. Paula Sommer for always believing in me and keeping me motivated. Without her continuous guidance in the lab and in the writing of this thesis I wouldn't have made it this far.

I would like to thank my co-supervisor Prof Tulio de Oliveira for granting me the opportunity of working on such an interesting project and guiding me through it till the end.

To my parents (Mr. S.V Mjoli and Mrs. R.N Mjoli) and my siblings, thank you for your endless support and encouragement. Your continuous belief in my potential made giving up a non-existent option.

Many thanks go to my colleagues at the School of Life Sciences Molecular Lab for their helpful input in my project and making my lab work a bearable experience.

Thank you to the lab staff at the Africa Centre for Health and Populations Studies Genomics lab for welcoming me into their working space and making my stay a pleasant one.

A special thanks is given to Siva Danaviah, for her assistance and supervision throughout my PCR optimization and Sanger sequencing process at the Africa Centre for Health and Population Studies Genomics Lab.

I thank Kamini Gouder for assisting me with the HLA sequence analysis.

I would like to thank the MRC Flagship Scholarship and the NRF for their support.

1. INTRODUCTION

1.1 Human Immunodeficiency Virus (HIV)

The Human immunodeficiency virus, better known as HIV-1 is a slowly replicating retrovirus that results in acquired immunodeficiency syndrome (AIDS). HIV infections are a global health problem, resulting in over 2 million AIDS-related deaths per year (Kulpa and Collins, 2011). AIDS is a condition in humans that progressively causes failure of the immune system and, in turn, allows life-threatening opportunistic infections and/or cancers to thrive in the human body (Kassutto *et al*., 2004). Depending on the subtype of the virus, without treatment, the average survival period after infection of HIV ranges between 9 to11 years. HIV is transferred (between people) mainly through direct contact with infected semen, pre-ejaculate, vaginal fluid, breast milk, or blood. In these bodily fluids the virus is present both in free virus particles and within the cells of the immune system (Kassutto *et al*., 2004).

HIV targets vital cells of the immune system such as the macrophages, dendritic cells and helper T cells (mainly the CD4+ T cells) (Kassutto *et al*., 2004). HIV infection results in lowered numbers of the CD4+ T cells through mechanisms such as direct viral killing of infected cells, killing of infected CD4+ T cells by the CD8 cytotoxic lymphocytes that keep an alert on infected cells within the body, and through apoptosis (programmed cell death) (Kassutto *et al*., 2004). As a result, cell-mediated immunity in the body is lost when the CD4+ T cell numbers decline sharply below the critical level therefore making the immune system prone and susceptible to infections that are opportunistic (Kassutto *et al*., 2004).

HIV is a complex virus as seen in Figure 1.1. The viral envelope of HIV is a derivative of the host cell's plasma membrane that is obtained through a process called budding (Clapham and McKnight, 2002). It consists of two layers in which two types of glycoproteins are embedded; glycoprotein (gp)120 in the outermembrane and gp41 attached on the transmembrane (Engelman and Cherepanov, 2012). In order for the virus to infect a host cell it needs to attach to a protein found on the surface of the T cell known as CD4 and it does this with assistance from the two abovementioned glycoproteins. The gp120 plays a role in co-ordination of attachment to the host cell and gp41 is essential for the cell fusion process (Clapham and McKnight, 2002). These two glycoproteins work together with co-receptors on the surface of

the cell, chemokine receptor 5 (CCR5) or chemokine (C-X-C motif) receptor 4 (CXC4) (Engelman and Cherepanov, 2012; Clapham and McKnight, 2002), allowing the virus to enter the host cell.

A viral capsid protein (p24 in Figure 1.1) is situated towards the viral core. The capsid surrounds the viral enzymes including the protease, reverse transcriptase, integrase and ribonuclease, which are all vital for HIV replication. It also houses nine genes, three of which (gag, pol and env) contain the information required for the production of structural proteins that give rise to new virus particles (Fiorentini *et al*., 2006). Another important HIV protein is the p17 matrix protein found between the viral core and the viral envelope. It is produced by the gag gene and serves a structural function in mature HIV particles (Fiorentini *et al*., 2006). The p17 protein lines the HIV envelope and anchors gp41 and gp120 to the envelope. p17 also carries localisation signals which are responsible for transferring the HIV core into the host nucleus once it has entered the host cell (Fiorentini *et al*., 2006).

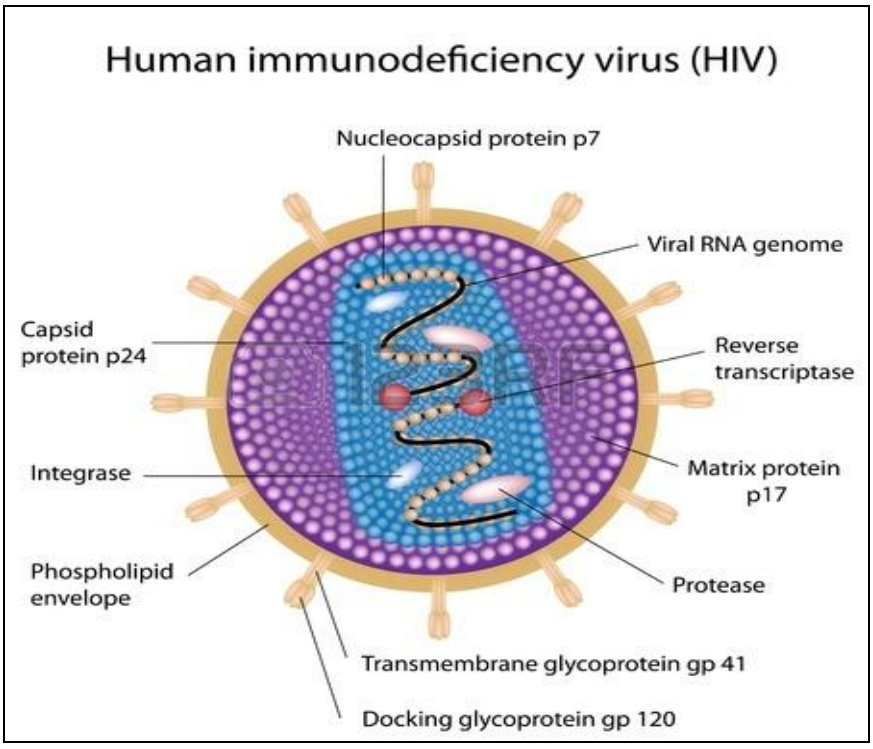


Figure 1.1: Anatomy of the Human Immunodeficiency Virus (HIV) (http://www.123rf.com/photo_12476223_structure-of-hiv.html).

To date the only effective treatment known for AIDS is Highly Active Anti-Retroviral Therapy (HAART). HAART is efficacious only in controlling the disease and complete eradication is yet to be achieved (Kulpa and Collins, 2011). In the past two decades, research has shown that the host genome plays a role in modifying disease outcome. As such, addressing this significant global health problem requires vast understanding on how the host's immune system responds to HIV and the mechanisms involved in viral takeover of the immune system in order to accelerate the development of preventative and therapeutic vaccines (Kulpa and Collins, 2011).

1.2 Role of host genome in HIV infection

While there is extensive knowledge on the genetics of HIV, there is less information on the role that host genome variants play in the course of HIV infection and disease progression. Multiple relationships occurring between specific gene variants and HIV-1 disease outcomes have been reported over the past 15 years (Fellay, 2009). Although most of them have not been confirmed or have been proven false positive, the identification of several definitive genotype–phenotype associations has shed new light on HIV-1 pathogenesis. This is particularly true with HIV-1 where a number of host genetic factors have been determined. Genes that have been identified and are believed to have an effect on the progression of HIV-1 include chemokine (C-C motif) receptor 5 (CCR5), Human leukocyte antigen B57 (HLA B57), Tripartite motif-containing protein 5α (TRIM5α), Apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like 3G (APOBEC3G) and Mannose-binding lectin 2 (MLB2) (Fellay, 2009).

1.3 Human Leukocyte Antigen (HLA)

The HLA gene family provides instructions for making a group of related proteins known as the human leukocyte antigen (HLA) complex. The HLA complex helps the immune system distinguish the body's own proteins from proteins made by foreign invaders such as viruses and bacteria (Bodmer, 1987; Frahm, 2005; Goulder and Watkins, 2008). The MHC complex in humans consists of more than 200 genes located in the same vicinity on chromosome 6 controlling the function of these antigens. The HLA region spans about 4 megabases of DNA on the short arm of chromosome 6 from 6p21.1 to p21.3 (Mehra, 2001). The class II (red),

class III (green) and class I (blue) genes are located from the centromeric (Cen) to the telomeric (Tel) end as indicated on Figure 1.2. The class I locus contains the classical genes, A, B and C, and the nonclassical genes E, F and G. Genes belonging to this complex are categorized into three basic groups: class I, class II, and class III (Figure 1.2) (Mehra, 2001). Humans have three main MHC class I genes, referred to as *HLA-A*, *HLA-B*, and *HLA-C* (Morandi *et al*., 2014). Class I antigens are cell surface glycoproteins composed of a highly polymorphic heavy chain (M_r 45,000) which is connected in a non-covalent manner to β_2-micro-globulin. The heavy chain consists of 2 peptide-binding domains, an Ig-like domain, and a transmembrane region with a cytoplasmic tail. The heavy chain of the class I molecule is encoded by genes at HLA-A, HLA-B, and HLA-C loci. These molecules (Class I) are widely distributed with the expression of the proteins produced from them present in most, but not all, nucleated cell surfaces (Morandi *et al*., 2014). These proteins are bound to protein fragments (peptides) that have been exported from within the cell and brought to the cell surface. MHC class I proteins display these peptides to the immune system and if recognized as foreign (viral or bacterial peptides), the infected cells are triggered to self-destruct as a response to the foreign invasion.

Class II antigens are composed of two glycosylated chains (M_r 32,000 and 29,000 respectively), and are predominantly expressed by cells involved in immunological responses namely B cells, antigen presenting cells and activated T lymphocytes (López-Nevot *et al*., 1989, Morandi *et al*., 2014). There are six main MHC class II genes in humans: *HLA-DPA1*, *HLA-DPB1*, *HLA-DQA1*, *HLA-DQB1*, *HLA-DRA*, and *HLA-DRB1*. These are usually present exclusively on professional Ag-presenting cells (B cells, macrophages, dendritic cells, Langerhans cells), thymic epithelium, and activated (but not resting) T cells. Most nucleated cells can be induced to express class II MHC molecules by interferon (IFN)-γ. Class II MHC molecules consist of 2 polypeptide (α and β) chains; each chain has a peptide-binding domain, an Ig-like domain, and a transmembrane region with a cytoplasmic tail. Both polypeptide chains are encoded by genes in the HLA-DP, -DQ, or -DR region of chromosome 6.

The MHC class III region of the genome encodes several molecules important in inflammation; they include complement components C2, C4, and factor B; tumor necrosis factor (TNF)-α; lymphotoxin-α; lymphotoxin-β; and three heat shock proteins. The proteins produced from MHC class III genes have somewhat different functions; they are involved mainly in inflammation and other immune system activities.

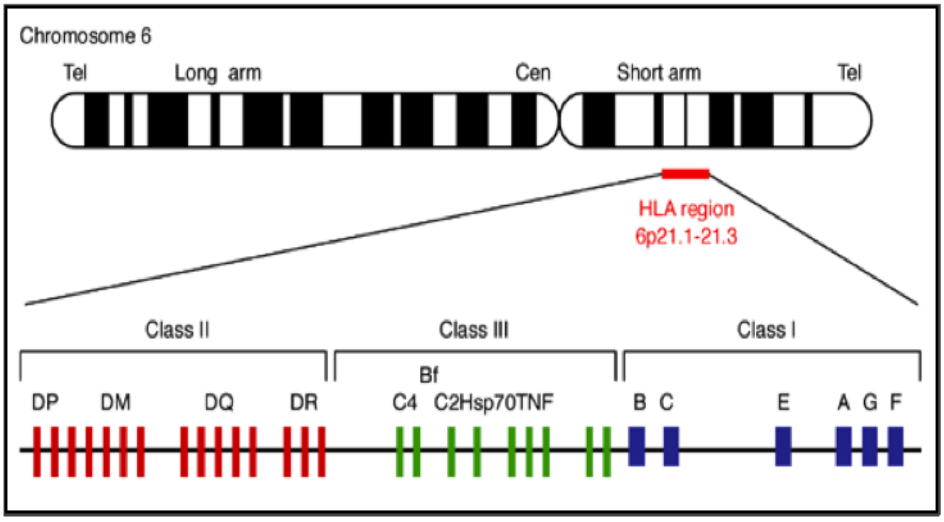


Figure 1.2: Schematic drawing of the HLA locus on chromosome 6 (http://neurowiki2012.wikispaces.com/).

HLA genes have many possible variations, which leads to each person's immune system exhibiting a wide range of reactions to overcome foreign invaders, which is probably a likeness of the natural selective influence by historic disease outbreaks that afflicted the ancestors of modern human ethnic groups (O'Brien and Nelson, 2004). Some HLA genes have hundreds of identified versions usually referred to as alleles, each of which is then given a particular number (such as *HLA-B27*) to identify it. Closely related alleles are thereafter categorized together; for example, at least 40 very similar alleles are subtypes of *HLA-B27* are known and these subtypes are classified as *HLA-B*2701* to *HLA-B*2743*. More than 100 diseases have been associated with different alleles of HLA genes. For example, the *HLA-B27* allele increases the risk of developing an inflammatory joint disease called ankylosing spondylitis (Klein and Sato, 2000). Many other disorders involving abnormal immune function and some forms of cancer have also been associated with specific HLA alleles. The abundant variation in HLA alleles provides a broad range for individual recognition of virus agents to which they had been exposed in the past, as well as those to which they had not (O'Brien and Nelson, 2004).

1.3.1 Role of the HLA Class I in HIV-1 infection/progression

In previous research it has been concluded that HIV-specific $CD8^+$ T cells play a central role in the primary amount of the virus found in the blood and the long-term suppression of viral replication (Leslie *et al*., 2010). This conclusion is supported by the observed correlation between possession of particular human leukocyte antigen (HLA) class I alleles and the control of HIV, measured mainly in two ways, both directly by time-to-AIDS and indirectly via clinical markers of disease progression such as viral load (VL) and CD4 count. Specific HLA class I alleles have been shown to be allied with satisfactory successful control of viral replication and in slowing down disease progression, with alleles such as HLA-B*57 and HLA-B*27 showing the most noticeable results. Some alleles however, such as B*35(Px), B*5802, and B*18, have demonstrated the opposite effect showing relatively ineffective control of viral replication and instead rapidly progressing the disease (Leslie *et al*., 2010). In addition, certain trends have also been described indicating a HLA class I heterozygote advantage such as B*3910-Cw*1203 and B*8101-Cw*0401 (Leslie *et al*., 2010). Associations with lower VL and higher CD4 count can only be observed when the two alleles (B*8101-Cw*0401) are expressed together; in the absence of either of the two their protective quality is lost. In the case of B*3910-Cw*1203, B*3910 is only protective when associated with Cw*1203. It is still however, not clear whether this information demonstrates a benefit of Cw*1203 or of the haplotype as a whole (Leslie *et al*., 2010).

Detectable HIV- specific $CD8^+$ T-cell responses amongst different HLA class I loci, responses that express a more effective "poly-functional" phenotype, $CD8^+$ T cells that exert the strongest selection pressure on the virus and the strongest HLA association with both slow and rapid progresssion are a few disease outcomes that can be associated with the HLA-B alleles. This, therefore, has led to the opinion that $CD8^+$ T-cell responses restricted by HLA-B alleles play a dominant role in most influential disease outcomes associated with HIV (Leslie *et al*., 2010).

The HLA-B allele most consistently associated with potent control of HIV-1 is B57, with B5701 observed almost exclusively in the Northern Hemisphere and B5703 mostly in African ancestry. HLA B57 is a HLA-B serotype, the sister serotype of B58. According to a study conducted by Descours *et al*. (2012), HLA B57 plays a large role in depressing the size of the

HIV reservoirs and also in the distribution of HIV by preserving the central memory cells (T_{CM}) compartment but at the same time not affecting other CD4 T-cell subsets. To determine whether these protective HLA alleles could impact the cell HIV-DNA distribution in resting CD4 T-cell subsets, 18 long-term non-progressors (LTNPs) from the LTNP cohort, which included 8 HLA-B27/B57–positive and 10 non–HLA-B27/B57 LTNPs were analysed (Descours *et al*., 2012). No significant differences were reflected of pVL or CCR5 expression due to this preservation, but instead a robust CD8 T-cell response against HIV Gag observed in HLA-B27/HLA-B57 LTNPs was seen (Descours *et al*., 2012). This could be explained by the fact that heterozygosity for human leukocyte antigen genes of class 1 loci (A, B, and C) have been shown to be associated with delayed disease progression to AIDS in HIV infected individuals while homozygosity of these loci has been associated with rapid progression to AIDS eventually leading to death (Descours *et al*., 2012). The advantage in being heterozygous is most likely because individuals with such loci for HLA are able to produce higher numbers of antigenic epitopes to cytotoxic T-lymphocytes than homozygotes, resulting in a stronger and more effective immune response when conquering HIV-1 infection.

In another study to further investigate the association between HLA type and the control it has on HIV, Leslie *et al*. (2010) used a cohort comprising 1211 subjects from Durban (South Africa) that were chronically infected with the HIV C-clade and were treatment-naïve.
The aim of the study was to determine whether polymorphisms in HLA-A and HLA-C can indeed influence disease progression or whether results are due to the dominant role that HLA-B is said to have on the progression of the disease (Leslie *et al*., 2010).
Variations in viral load (VL) in the cohort were measured using the nonparametric Kruskal-Wallis test of variance. This test does not only assess significance but also produces a Kruskal-Wallis statistic, *H*, which corresponds to the strength of the effect being tested. The effect tested was the differential contribution that the different HLA types (A, B and C) have on the variation in VL by grouping the individuals according to HLA type. They first grouped individuals by their HLA-A type and repeated the test, grouping individuals in turn by their HLA-B and HLA-C types. Comparisons of the VLs in subjects expressing the relevant allele with VLs in subjects not expressing that allele was the key concept used in analyzing HLA type associations with VL. This analysis was undertaken for a total of 65 alleles using the stringent Boneferroni correction, a method used to neutralize the problem of multiple comparisons, generating significance cut off P value of <0.0007. Associations between HLA class 1 types and the absolute CD4 count were assessed following the same steps as stated

above. The results showed that the HLA-B type was the strongest predictor of both VL and CD4 while HLA-A type was the weakest predictor of the three.
They further examined the possibility that the observed dominant effect of HLA-B on VL and CD4 count might be a coincidence of the fact that the main protective (B*57 and B*5801) and susceptible (B*5802 and B*18) alleles happen to belong to the HLA-B locus. Initially, all subjects possessing B*57, B*5801, B*5802, and B*18 from the data set were excluded and they repeated the analysis on the remaining 682 subjects. Even with these four alleles excluded, HLA-B remained the strongest predictor of both VL and CD4 while HLA-A type remained the weakest. A more sophisticated likelihood ratio test (LRT) was implemented to confirm this observation. This ratio works by allowing the top four alleles described above to be used as covariates instead of disregarding them from the analysis. The method also identified HLA-B as the strongest predictor of both VL and CD4 count as when all alleles were considered.

Taken together, these data suggest, firstly, that the dominant role of HLA-B in controlling HIV infection extends beyond the strongest four HLA-B associations and, secondly, that much of the apparent impact of HLA-A alleles on VL is related to linkage disequilibrium with HLA-B alleles, as grouping by HLA-A type no longer significantly explains any variation in VL or CD4 count when the top four HLA-B alleles are either excluded or used as covariates. Whether this apparent association between HIV immune control and HLA-B is a general and a causal trend or, rather, is biased by the coincidence that the strongest HLA associations with either extreme of disease control happen, by chance, to involve HLA-B alleles still remains uncertain as other studies state otherwise (Leslie *et al*., 2010). More downstream analysis of the gene need to be done in order to settle the uncertainty around it.

Stephens (2005) states that in the African, Caucasoid and Asian population HIV-1 is diversifying at a fast rate. This diversity is believed to run in parallel with the extensive polymorphism that is displayed by genes encoding HLA class I (Stephens, 2005). Imprinted mutations in HIV-1 mediated by immune responses linked to the HLA class I molecules affect HIV-1 diversity. Many studies based on intra- and inter-ethnic populations have shown reproducible HLA class I allele, haplotype and supertype links with HIV-1 infection and also with the development of AIDS (Stephens, 2005). It is therefore essential for this host gene (HLA class 1) to be analyzed so its role in the manner of development of the disease can be understood, predict risk and be used to personalize care.
McLaren *et al*. (2015) tested for associations with set point viral load in approximately 8

million variants. These variants included single nucleotide polymorphisms (SNPs), short insertions and deletions, usual HLA alleles, and variable amino acids in HLA proteins of the European ancestry that consisted of 6315 individuals. They did this by combining the majority of available genome-wide genotyping data in HIV-infected populations with clinical data as they believed that previous genome-wide association studies based on similar populations were underpowered to detect common variants with moderate impact on disease outcome and had not considered the phenotypic variance which can be explained by additive effects of genome–wide variation (McLaren *et al*., 2015). Their main aim in the study was to map the genomic regions with the greatest influence on the viral load of HIV and quantify the impact those regions have. The results revealed that 14.5% of the observed variations in the HIV-1 viral load were explained by the SNPs in the chemokine (C-C motif) receptor 5 better known as CCR-5 and by the amino acid positions located in the HLA proteins binding grooves specifically HLA-A and HLA-B. Additional to this, an estimated 5.5% of the variations were explained by additive genetic variation. This, therefore, validated that the majority of the host genetic components playing role in managing the HIV viral load are explained by common variants of large effect (McLaren *et al*., 2015).

It is important to note that the major obstacle in performing genomic studies is the requirement of a large amount of blood that is repeatedly needed from sample subjects of experiments or studies to extract the DNA from white blood cells (Zhou *et al*., 2006). Fortunately, the reawakening of a method of sampling, in existence for a number of years, known as dried blood spots (DBS), has regularized the storage of blood for these studies. DBS require only a small amount of blood from the subjects and these blood samples can be stored for a number of years (Hue *et al*., 2011; St. Julien *et al*., 2013).

1.4 Dried Blood Spots

Dried blood spots (DBS) refers to a blood sampling technique where small volumes of blood are spotted on an appropriate filter paper, dried, and taken to the laboratory for analysis (Figure 1.3). Collection of blood samples on filter paper can be dated to more than 40 years ago (Hue *et al*., 2012; Jahannessen, 2010). This simple and practicable collection and storage method has been adopted for broad use in analytical screening, drug monitoring, and genetic analysis, being particularly suitable for molecular epidemiologic studies in remote areas with

tropical climate, where transport and storage conditions are often not the most favorable (Chaisomchit *et al*., 2003; Chaisomchit *et al*., 2005). Most samples are stable at room temperature when dried on these filter papers for at least a week, thereby obviating the need for maintaining a cold chain for transportation of the sample (Lakshymy, 2008).

Archived dried blood spots are an important and precious resource for genomic research. Unlike venipuncture, the technique is minimally invasive (Lakshymy, 2008) and there is no need for centrifugation and separation of samples, making dried blood suitable for field applications with minimum involvement of laboratory-trained laborers (Lakshymy, 2008). In addition, it is also inexpensive and easy to analyse and automate. DBS samples are compatible with a large number of bio-analytical methods, among them chromatography, mass spectrometry, DNA fingerprinting, and immunoassays (Chaisomchit *et al*., 2005). DNA in the filter paper is entrapped into a chemically treated matrix that avoids degradation, and ambient storage is possible for years without deterioration (Pachot *et al*., 2007). Considering the growing knowledge of genetic polymorphisms that might be implicated in the pathogenesis of major diseases such as HIV-1 and AIDS, the advantages of DBS samples may facilitate polymorphism analysis in routine clinical practice and the creation of large DNA banks using minute amounts of blood (Pachot *et al*., 2007; Hue *et al*., 2012).

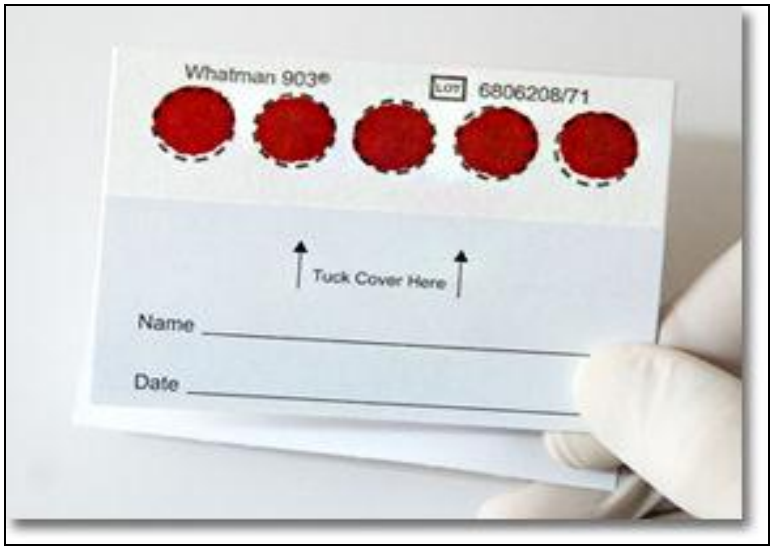


Figure 1.3: An example of how blood is stored as dried blood spots on specialized filter paper (http://www.biomerieux-diagnostics.com/hiv-diagnostics-management).

Polymerase chain reaction (PCR) is a requirement for amplification of the desired genes (Zhou *et al*., 2006). Over the years a shortfall has been reported in PCR results when DBS are used as the source of genomic DNA. This could likely be the consequence of low purity, stability, and integrity of the extracted genomic DNA from filter paper. Analysis of blood spots poses few challenges, mainly that of elution of red blood cells along with the analyte of interest (Lakshymy, 2008). This is due to cellular components rupturing when whole blood samples are dried on filter, which further decomposes into solutions when blood spots are reconstituted leading to additional extraction procedures being required for certain analytes to overcome this problem (Lakshymy, 2008). The majority of PCR inhibitors are integrated from the sample or introduced while samples are being processed or in nucleic acid extractions. Examples of known blood components that become problematic in PCR amplification are namely immunoglobin G, haemoglobin and lactoferrin (Al-Soud and Radstrom, 2000). These components reduce the amplification efficiency by decreasing sensitivity of the procedure or by giving false-negative results (Schrader *et al*., 2012, Al-Soud and Radstrom, 2000). They do so by binding to the DNA polymerase, which in turn prevents the assembling of nucleotides to form DNA strands (Al-Soud and Radstrom, 2000).

Another key problem with sample collection on filter paper is the efficiency of elution of the analyte of interest and relative sample volume limitations of sample collected that can be incorporated into the extraction procedure (Wong *et al*., 2008; Lakshymy, 2008). Therefore, an efficient method for recovering high quality genomic DNA from dried blood specimens is required. There are many commercially available kits for extracting DNA from dried blood samples such QIAGEN kits but the problem with these is that they are expensive and involve spin columns that provide low recovery rates. Furthermore, the spin column kits used in these kits always result in increased plastic waste demonstrating undesirability of this extractive process (Chaisomchit *et al*., 2003; Lin *et al*., 2005).

1.4.1 DNA extraction from Dried Blood Spots

Despite the known drawbacks, storage of blood as dried spots on filter paper is a practicable approach for genetic screening (Sirdah, 2014; Lakshymy, 2008). A number of DNA extraction kits have been available commercially for many years, but do not accommodate the use of DBS in their extraction protocols. Newer kits such as QIAamp DNA mini and

microkits can now be used to extract DNA from DBS (Sirdah, 2014). Automated DNA extraction methods have also been discovered that accommodate extraction of DNA from this sample type such as the automated liquid handling system that purifies DNA from FTA cards (Sirdah, 2014) and a DNA extraction method using the MagNA Pure LC, a magnetic particle based automated extraction discussed in de Vries *et al*. (2009).

Table 1.1 is a summary of different extraction methods described by numerous authors alleged to produce good quality DNA from DBS. Modifications have been implemented in some of these extraction methods with a purpose of increasing the DNA quality and yield they produce. The table also shows the different technologies that these authors used to validate the quality and purity of the extracted DNA samples.

Table 1.1: Results obtained from various publications based on the extraction of DNA from DBS using different extraction methods and downstream analyses (nested PCR- nPCR).

Reference	Source of blood	Extraction method	Type of card	Type of analysis	Size of amplicons/ Name of gene	Analysis output	Automated/ Non-automated
Zhou *et al*., 2006	Sheep	NaOH method	3MM	PCR	242 bp (DQA2)	Successful	Non-automated
					1.4 kb (DQA2l -dn)	Successful	
			FTA		242 bp (DQA2)	Successful	
					1.4 kb (DQA2l -dn)	Successful	
		FTA purification Reagent	3MM		242 bp (DQA2)	Failed	
					1.4 kb (DQA2l -dn)	Failed	
			FTA		242 bp (DQA2)	Successful	

					1.4 kb (DQA2l -dn)	Successful	
Chaisomchit *et al*., 2003	Human- volunteers with written consent	Not Specified	Filter paper (Schleicher & Schuell)	PCR	255 bp (PCR product)	Successful	Non-automated
					674 bp (PAX8)	Successful	
					1039 bp (β-globin)	Successful	
Hue *et al*., 2012	Human- clinical samples (new born screening program)	One-step (Phenol method- one lyses steps)	Not specified	PCR	241 bp (G6PD-E2-F)	Failed	Non-automated
				DNA yield	-	17.6 ng/ µl (Purity- 1.66)	
		Two-step (Phenol method- two lyses steps)		PCR	241 bp (G6PD-E2-F)	Successful	
				DNA yield	-	51.9 ng/ µl (Purity-1.66)	
Fischler *et al*., 1999	Infants - CMV infection study	Method 1 (4 extraction steps)	Guthrie cards	PCR	107 bp (OP1 and OP2)	Failed	Non-automated
				nPCR	167 bp (IEP3A and IEP3B)	Failed	
					723 bp (IEP2A and IEP4B)	Failed	

		Method 2 (1 extraction step)		PCR	107 bp (OP1 and OP2)	Successful	
				nPCR	167 bp (IEP3A and IEP3B	Successful	
					723 bp (IEP2A and IEP4B)	Successful	
Lin *et al*., 2005	Human-newborn screening tests	Automated method (HPLC-grade methanol)	Filter paper #903 (Schleicher & Schuell)	PCR	241 bp (Human β-globin)	Successful	Automated
				High throughput genotyping assay	Melting peak at 73° C was detected for all.	Successful amplification	
				DNA yield		0.5 ng/ µl	
Göhring *et al*., 2010	Human-native HCMV-seronegative EDTA-blood	QIAamp blood mini kit	Guthrie cards	PCR	gB-region	Successful	Non-automated
				nPCR	IE1-Exon4	Successful with highest sensitivity	
		Heat extraction method		PCR	gB-region	Failed	Non-automated
				nPCR	IE1-Exon4	Failed	
		Traditional phenol-		PCR	gB-region	Successful	Non-automated

		chloroform		nPCR	IE1-Exon4	Successful with highest sensitivity	
		Automated nucleic acid extraction system		PCR	gB-region	Failed	Automated
				nPCR	IE1-Exon4	Failed	

1.5 DNA Sequencing

DNA sequencing is the process used for determining the specific order of nucleotides within a DNA molecule. Any method or technology that can be utilized to determine the order of the four bases (adenine, guanine, cytosine, and thymine) in a strand of DNA is referred to as DNA sequencing (França *et al*., 2002). Applied fields such as biotechnology, virology, medical diagnosis, medicine and biology rely greatly on DNA sequencing, and the high speed of sequencing has been influential to the sequencing of complete DNA sequences and genomes of numerous types of species of life such as the human genome (França *et al*., 2002). The first DNA sequences were obtained in the early 1970s by academic researchers, Hamilton Smith and his coworkers, using laborious methods based on two-dimensional chromatography (Smith and Welcox, 1970). They discovered type II restriction enzymes, which can recognize and digest DNA at specific nucleotide sequences and making it possible to cut a large DNA fragment into numerous smaller pieces that could be separated according to their sizes by using gel electrophoresis (Smith and Welcox, 1970).

1.5.1 Sanger Sequencing

The classical chain-termination method, commonly referred to as Sanger sequencing, is a DNA sequencing method that requires a single-stranded DNA template, a DNA primer, DNA polymerase, normal deoxynucleosidetriphosphates (dNTPs), and modified di-deoxynucleosidetriphosphates (ddNTPs). The ddNTPs are responsible for the termination of the DNA strand elongation (Sanger *et al*., 1977; Karger and Guttman, 2009). These chain-terminating nucleotides lack a 3'-OH group required for the formation of a phosphodiester bond between two nucleotides, causing DNA polymerase to terminate extension of DNA when a modified ddNTP is integrated into a DNA strand. The resulting newly synthesized DNA chains will be a mixture of lengths, depending on how long the chain was when a ddNTP was randomly integrated into the chain. The ddNTPs may be radioactively or fluorescently labeled for detection in automated sequencing machines (Sanger *et al*., 1977; França *et al*., 2002; Karger and Guttman, 2009).

The DNA sample is divided into four separate sequencing reactions, containing all four of the standard deoxynucleotides (dATP, dGTP, dCTP and dTTP) and the DNA polymerase. To each reaction, only one of the four dideoxynucleotides (ddATP, ddGTP, ddCTP, or ddTTP) is added while the four other nucleotides are standard ones. To put it in a more sensible order, four separate reactions are needed in this process to test all four ddNTPs (França *et al*., 2002; Karger and Guttman, 2009). Following rounds of template DNA extension from the bound primer, the resulting DNA fragments are heat denatured and separated by size using gel electrophoresis and visualized by autoradiography or UV light reading them directly off X-ray film or gel image (Sanger *et al*., 1977; França *et al*., 2002; Karger and Guttman, 2009).

1.5.2 Sanger sequencing with capillary electrophoresis

The aftermath of the discovery of Sanger sequencing, saw Smith and his associates, introducing a dye-terminator alternative that is less time consuming than the original chain-terminator. In this method, the same principle as chain-termination though different size fluorescent DNA fragments are generated by incorporation of dye-labeled ddNTPs is used (Smith *et al*., 1986). Each of the four ddNTP chain terminators are integrated with a different fluorescent dye, allowing sequencing in one reaction rather than four (Figure 1.4). When DNA template, unlabeled primer, dNTPs, four fluorescently labeled ddNTPs and DNA polymerase are mixed together, different size fluorescent DNA fragments are generated by incorporation of dye-labeled ddNTPs (Smith *et al*., 1986).

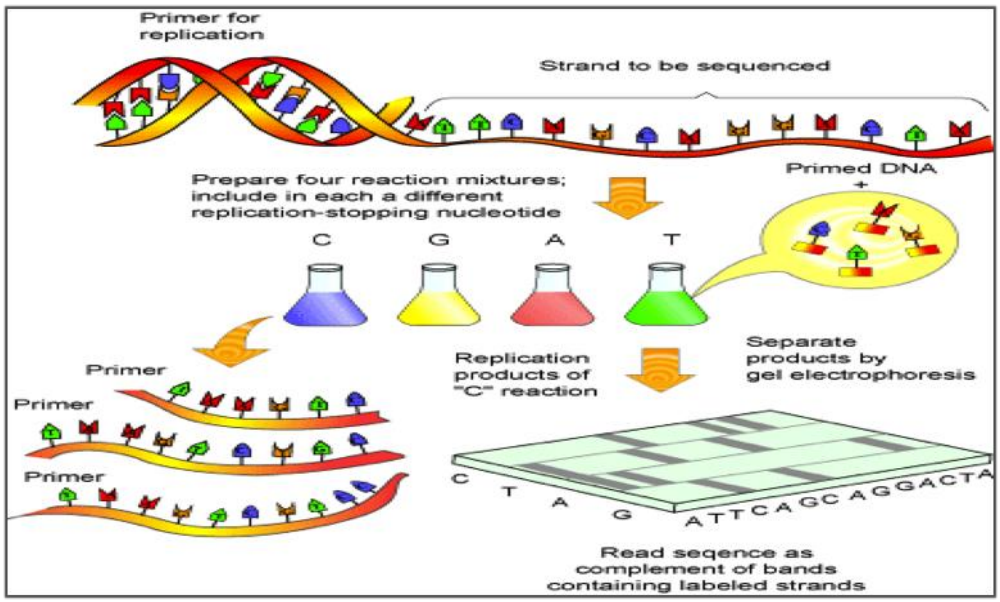


Figure 1.4: Diagram showing the process followed in Sanger sequencing when using dye-terminators (http://www.scq.ubc.ca/wp-content/uploads/2006/08/sequencing2.gif).

A breakthrough in DNA sequencing technology was demonstrated when viewing a sequence on the slab gel was replaced by viewing using capillary electrophoresis. Prior to this, the DNA sequencing products were separated using polyacrylamide gels, manually poured between two glass plates that was a tedious process. In capillary electrophoresis, the DNA sequencing products enter the capillary as a result of an electrokinetic injection (Drossman *et al*., 1990). In essence, a high voltage charge is applied to the sequencing reaction buffer and the voltage in turn pushes the negatively charged fragments into the capillaries leading to the DNA sequencing products being separated by size based on their total charge (Drossman *et al*., 1990). These detect and record dye fluorescence, and output the data as fluorescent peak trace chromatograms. Sequencing reactions (thermocycling and dye-labelling), cleanup and re-suspension of samples in a buffer solution are performed separately, before loading samples onto the sequencer. Sanger sequencing with obtainable sequence length 1000 bps using dye-terminators thereafter became the dominant sequencing technique until the introduction of so-called next-generation sequencing technologies (Drossman *et al*., 1990).

1.6 Aim and objectives

A significant increase in understanding host genetic and genomic determinants of susceptibility to HIV-1 infection and disease progression has been witnessed over recent years

mainly through identifying common variants in some host loci that influence disease progression, characterizing the scale and dynamics of gene and protein expression changes in response to infection and lastly providing pathways involved in in viral replication through comprehensive catalogs of genes (Telenti and Johnson, 2012). Candidate gene studies, genome-wide association studies, genome-wide transcriptome analyses, and large-scale in vitro genome screens mainly drive these discoveries (Telenti and Johnson, 2012).

As asserted in Lakshymy (2008) storing blood as dried blood spots is a powerful tool in screening programs and especially research based on a large population. This is because it is a method of blood collection that is less invasive and relatively painless, which gives it an advantage over other methods especially when working with elderly people or infants. Blood stored in this manner can also be transported without a cold chain which makes it suitable for research on developing countries where facilities are far from cohorts being studied and where cost is a major issue hindering the progress of research (Lakshymy, 2008).

It is imperative that DNA extraction methods producing good quality DNA with high yields from DBS are discovered for deeper understanding of susceptibility to HIV-1 infection and disease progression. The aim of this MSc research project is to produce high yield and good quality DNA from DBS that can be used for host genomic analysis, more specifically HLA genotyping.

Objective 1

Extract DNA of good quality and high yield from dried blood spots.

Objective 2

Amplify HLA Class 1A, 1B and 1C loci using nested PCR

Objective 3

Sequence PCR products

Analyse the HLA sequences using variant detection and genotyping sequence analysis software.

2. MATERIALS AND METHODS

2.1 Dried blood spots

Dry blood spots (DBS) were prepared from anonymous blood samples that are commonly used for the validation of new laboratory methods at the Africa Centre Lab in Durban. 50 µl of the blood sample were pipetted onto each circle on a filter paper card (Whatman 903™, UK) and the filter cards were allowed to dry at room temperature for 3 hours in a sterile environment (laminar airflow). Thereafter, the cards were placed in an envelope with a packet of dessicant and the envelope was placed in a zip lock bag before storing in an airtight container at -20°C.

2.2 DNA Extraction

Two DNA extraction methods, QIAamp® DNA Mini Kit (QIAGEN, Germany) and PrepFiler® Forensic DNA extraction Kit (Life Technologies, California) were used to determine which of the two would give the best quality and yield of DNA (summarized in Appendix A.1). The QIAamp® DNA Mini Kit was shown to generate best yields and is described below.

DNA from the DBS was extracted using QIAamp® DNA Mini Kit according to the manufacturer's instructions with a few modifications (as indicated below). The protocol was optimized in attempt to isolate good quality DNA from the samples. The following procedure was followed:

A metal one-hole paper punch was used to punch out two 3 mm circles from a dried blood spot. These were placed into a 1.5 ml microcentrifuge tube with 180 µl of Buffer ATL. The microcentrifuge tube was incubated at 85°C for 10 min in a shaking heating block (Eppendorf Thermomixer compact) and briefly centrifuged (Eppendorf 5415 R centrifuge). 20 µl Proteinase K was added to the microcentrifuge tube and mixed by vortexing using a FINE VORTEX (FINE PCR) at full speed, before incubation at 56°C for 60 min in a stationary heating block (TECHNE DRI-Block DB.2A). The tube was briefly centrifuged to collect the fluid at the bottom of the tube. 200 µl of Buffer AL was then added to the sample, mixed

thoroughly by vortexing at high speed, and incubated at 70°C for 10 min in a stationary heating block. All fluid was collected by brief centrifugation. The cell lysate was then passed through a QIAshredder (QIAGEN) at full speed for 2 min in an Eppendorf 5415 R centrifuge. The QIAshredder column was discarded and the collection tube containing the DNA and the lysate was retained.

200 µl of 100% ethanol was added to the lysate and mixed thoroughly by vortexing at high speed. The mixture was carefully applied to the QIAamp Spin Column (in a 2 ml collection tube) and centrifuged at 6000 x *g* for 1 min. The QIAamp Spin Column was then transferred into a clean 2 ml collection tube. 500 µl of Buffer AW1 was added to the spin column and the tube centrifuged at 6000 x *g* for 1 min. The filtrate was discarded and the collection tube replaced. Thereafter, 500 µl of Buffer AW2 was added to the spin column and centrifuged at full speed for 3 min. To eliminate any chance of Buffer AW2 carryover, the QIAamp Spin Column was transferred into a clean 2 ml collection tube and centrifuged at 20 000 x *g* for 1 min to dry the membrane. DNA was eluted by adding 100 µl Buffer AE to the QIAamp Spin Column placed in a clean 1.5 ml microcentrifuge. The sample was incubated at room temperature for 1 min before centrifugation at 6000 x *g* for 2 min. All extracted samples were stored in at -80°C.

2.3 DNA quantification

DNA concentrations of the extracted samples were quantified using the Qubit 2.0 Fluorometer (Qubit Assay). The extracted samples were also separated on a 1% agarose gel at 70V using a Bio-Rad Powerpac (Basic) for 45 min in 1x TBE buffer. For the purpose of DNA illumination under a UV light source 1 µl of Novel juice (Gene Direx) was mixed with 4 µl of the PCR product before running the gel. The gel was imaged using a GelDoc™ XR+ imaging system (Bio-Rad) and Image Lab 4.1 software package (Bio-Rad). This was done to confirm the presence of DNA particularly in the samples that had given invalid concentrations as seen in Table 3.1.

2.4 Polymerase chain reactions

Numerous attempts at the nested polymerase chain reaction (PCR) were conducted to find the most appropriate and conducive PCR conditions that amplified the expected product size (approximately 1000 bp) for the HLA Class 1A, 1B and 1C loci (summarized in Appendix C.1 and C.2). The KAPPA2G™ Robust HotStart ReadyMix 2X (KAPABIOSYSTEMS) kit was used to amplify the PCR product. PCR amplifications were performed using the Gene Amp PCR System 9700. A total of 12 samples underwent amplification; 10 samples from the anonymous donors, an HIV negative blood sample used as a positive control (+ve) and deionized water used as a negative control (-ve) as seen in Figure 3.2.

2.4.1 PCR Step 1

The nested polymerase chain reaction involves two sets of primers, used in two successive runs of PCR, the second set intended to amplify a secondary target within the first run product. It is done in this manner to attempt eliminating the commonly occurring problem of primers binding to incorrect regions of the DNA, giving unexpected products.

Table 2.1: Primer sequences used to amplify the HLA Class 1 A, B and C loci in step 1 of the nested PCR.

HLA Class 1	**Primer name**	**Primer sequence 5' 3'**	**Primer size**
A	5A.1	CCC AGA CGC CGA GGA TGR CSG	21 bases
	3A1.1	GCA GGG CGG AAC CTC AGA GTC ACT CTC T	28 bases
B	5B.1	TCC CAG TTC TAA AGT CCC CAC G	22 bases
	3B.1	TCC ATT CAA GGG AGG GCG AC	20 bases
C	5C.1-2	AGC GAG GKG CCC KCC CGG CGA	21 bases
	3C.1-2	GGA GAT RGG GAA GGC TCC CCA CT	23 bases

Table 2.2: Reagent volumes used for PCR amplification in step 1 of the nested PCR.

Reagents	**Volume / reaction** (µl)
Water	3.8
Master Mix	10
$MgCl_2$ (25mM)	0.4
Primer 1(10uM)	0.4
Primer 2 (10uM)	0.4
Sample	5
Total reaction volume	**20**

Table 2.3: Programmed thermal cycles used for PCR amplification in step 1 of the nested PCR.

Thermocycling Conditions (Step 1)		
Temperature (°C)	**Time**	**Cycle**
94	4 min	Hold
94	15 sec	X 40
63	30 sec	
72	30 sec	
72	7 min	Hold
4	∞	Hold

As mentioned above the second set of primers are intended to amplify a secondary target in the first run, therefore 2,5 µl of the PCR product in step 1 was used as the sample in step 2 of the PCR amplification as seen in Table 2.5 (refer to Appendix C2.3).

2.4.2 PCR Step 2

Table 2.4: Primer sequences used to amplify HLA Class 1A, 1B and 1C loci in step 2 of the nested PCR.

HLA Class 1	**Primer name**	**Primer sequence 5' 3'**	**Primer size**
A	5A1N1	CCT CTG YGG GGA GAA GCA A	19 bases
	3A1N3	GAC TCA GAA GTG CTG CTG GAC TC	23 bases
B	5B3	GGC TCC CAG TTC TAA AGT CCC CAC G	25 bases
	3B1	CCA TCC CCG GCG ACC TAT AGG AGA TG	27 bases
C	CEX1F-2	CCT GAY CGA GAC CTG GGC C	19 bases
	DMEX3	TGG GAG GCC ATS CCG GGA GAT	21 bases

Table 2.5: Reagent volumes used for PCR amplification in step 2 of the nested PCR.

Reagents	**Volume / reaction** (µl)
Water	7.5
Master Mix	12.5
Enhancer	1.0
$MgCl_2$ (25mM)	0.5
Primer 1(10uM)	0.5
Primer 2 (10uM)	0.5
Sample	2.5
Total reaction volume	**25**

Table 2.6: Programmed thermal cycles used for PCR amplification in step 2 of the nested PCR.

Thermocycling Conditions (Step 2)		
Temperature (°C)	**Time**	**Cycle**
94	4 min	Hold
94	15 sec	X 40
65	30 sec	
72	30 sec	
72	7 min	Hold
4	∞	Hold

PCR products of step 2 were subjected to gel electrophoresis to determine whether the expected product size (1000 bp) had been amplified.

2.5 Gel electrophoresis

PCR products were examined by gel electrophoresis in a 1% agarose gel run at 70V with a Bio-Rad Powerpac (Basic) for 45 min in 1x TBE buffer. 200bp O'RangeRuler™ DNA ladder (Thermo Scientific) was used to determine product size. For the purpose of DNA illumination under a UV light source 2 µl of Novel juice was mixed with 2 µl of the DNA ladder and 1 µl of Novel juice mixed with 4 µl of the PCR product before running the gel. The gel was imaged using a GelDoc™ XR+ imaging system (Bio-Rad) and Image Lab 4.1 software package (Bio-Rad).

2.6 PCR product purification

PCR products were purified using the Pure Link Purification Kit (Invitrogen Life Technologies) following the manufacturer's protocol.

Briefly, 4 volumes of PureLink® Binding Buffer (B2) were added to 1 volume of the PCR product and mixed by vortexing. The samples were thereafter transferred into a PureLink® Spin Column in a collection tube. An amount of 650 µl of Wash Buffer was added to each column. The columns were centrifuged at 10 000 x *g* for 1 min at room temperature. The flow through was discarded and the spin column placed into a new collection tube. It was centrifuged at maximum speed for 3 min at room temperature. The spin column was then placed in a clean 1.7 ml PureLink® Elution Tube. The PCR products were eluted by adding 50 µl of elution buffer (10 mM Tris-HCl, pH 8.5) into the center of the spin column. The column was incubated at room temperature for 1 min before centrifuging at maximum speed for 2 min. Purified PCR products (contained in the collection tube) were run on a 1% agarose gel.

2.7 Sequencing

2.7.1 Sequencing reaction

The PCR products were sequenced using the ABI Prism BigDye Terminator kit v3.1 Cycle Sequencing Kit (USA) and 4 primers for each sample.

Table 2.7: Primer sequences used to generate the four fragments of the HLA Class 1A, 1B and 1C loci used in Sanger sequencing.

HLA Class 1	Primer name	Primer Sequence
A	INT2R	GGA TCT CGG ACC CGG AG (exon 2)
	A1N110	AGC CGC GCC KGG AGG AGG GTC G (exon 2)
	5AE3.2	GTT TAG GCC AAA AAT YCC CC (exon 3)
	3AE3.2	TGT TGG TCC CAA TTG TCT CCC CTC (exon 3)
B	INT2R	GGA TCT CGG ACC CGG AG (exon 2)
	3B1	TCC ATT CAA GGG AGG GCG AC (exon 2)
	5BE2.2	GGG AGG AGM GAG GGG ACC SCA G (exon 3)
	3BE3.2	CCA TCC CCG GCG ACC TAT (exon 2)
C	INT2C	CGA CCC GGG CCG TC (exon 2)
	BC3N1P3	GGC TCC CCA CTG CCC CTG GTA C (exon 2)
	5CE2.2	GGC TCC CCA CTG CCC CTG GTA C (exon 3)
	3CE3.2	AGA TGG GGA AGG CTC CCC ACT (exon 3)

The amplification reaction consisted of 5.1 µl water, 0.4 µl Ready Reaction mix, 0.5 µl primer (5µM) (as seen in Table 2.7), 2 µl 5X sequencing buffer and 2 µl of the template (20 pmol/µl) making a total volume of 10 µl for each primer. PCR was performed using the Gene Amp PCR System 9700. The thermal cycles were programmed as follows:

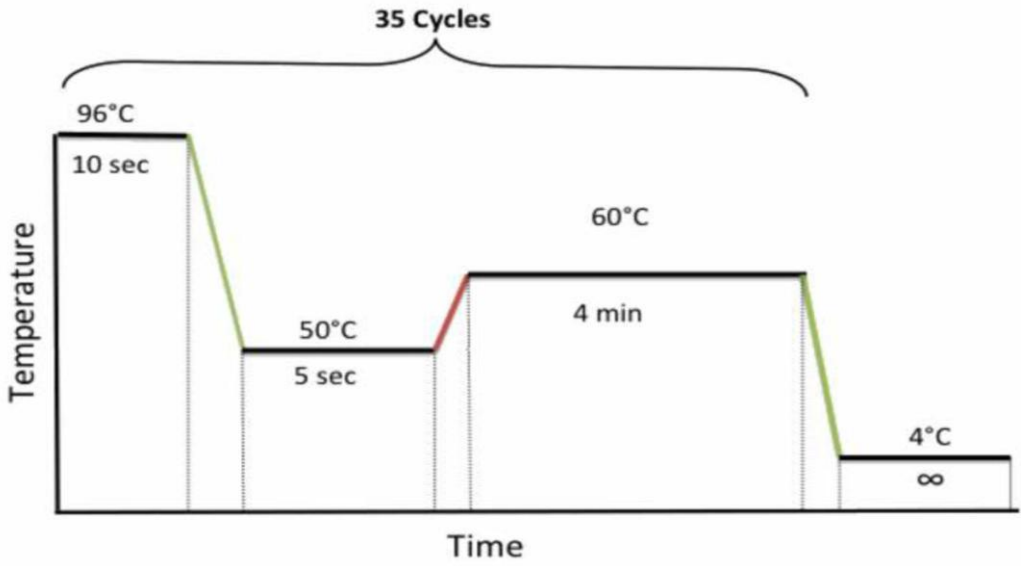


Figure 2.1: Programmed thermal cycles used for the sequencing reaction (Adapted from Manasa *et al*., 2014).

2.7.2 Sequencing cleanup

The sequences were immediately cleaned after PCR amplification following the SATuRN/ Life Technologies Genotyping Manual procedure (ABI).

Table 2.8: The volumes of each reagent used in preparing fresh sodium acetate and ethanol for sequence cleanup.

Reagent	**X1 (µl)**	**X120 (for on full plate)**
3M sodium acetate	5	240
100% EtOH	50	6000
Total	**55**	**6240**

55 µl of sodium acetate solution (Table 2.8) was added to each well. The wells were then carefully sealed with an adhesive foil cover and mixed by vortexing at high speed. Thereafter, the samples were centrifuged at 3000 x g for 20 min. The foil cover was then removed and the plate was inverted onto a folded 'kimwipe' to remove supernatant but cautiously enough not to dislodge the pellet. For further drying, the plate was centrifuged while still inverted on 'kimwipes' at 150 x g for 2 min. 150µl of 70% ice cold EtOH was immediately added to each well, it is vital that the addition of EtOH is immediate, as dye blobs will form if there is any delay. The samples were centrifuged at 3000 x g for 5 min and thereafter the plate was inverted over folded 'kimwipes' followed by centrifugation at 150 x g for 1 min still inverted. The uncovered plate was then dried at 50°C for 5 min in a thermocycler and, once dry, the plate was thereafter reconstituted with 10µl of formamide and denatured at 95°C for 2 min before sequencing in a 3130 X1 Genetic Analyzer using a standard 35-cycle protocol.

2.8 Bioinformatics

Geneious R8 v8.1.8 was used to edit and map the four sequences from Sanger sequencing into a consensus sequence. The reads were mapped into a consensus sequence using the de novo assembly option. The PubMed NCBI BLAST (Basic Local Alignment Search Tool) was used to align these new sequences. They were aligned against the non redundant (NR) database for sequence quality control purposes. This was done in order to evaluate if the consensus sequences fall under the HLA Class 1 loci. The Assign^{TM} ATF Softwere v1.5 (Conexio Genomics, Australia) was also used to detect variation of HLA Class 1A, 1B and 1C alleles in the reads of the 10 sequenced samples.

2.8.1 Sequence Quality control

Reads produced from Sanger sequencing were edited using Geneious R8 v8.1.8. Normally the first 30-50 bp and the last 100-200 bp of sequence reads are of low quality and are therefore trimmed off leaving only the section of the reads that are of good quality (as seen in Figure 2.2). Bases with a quality score below 30 are considered low quality.

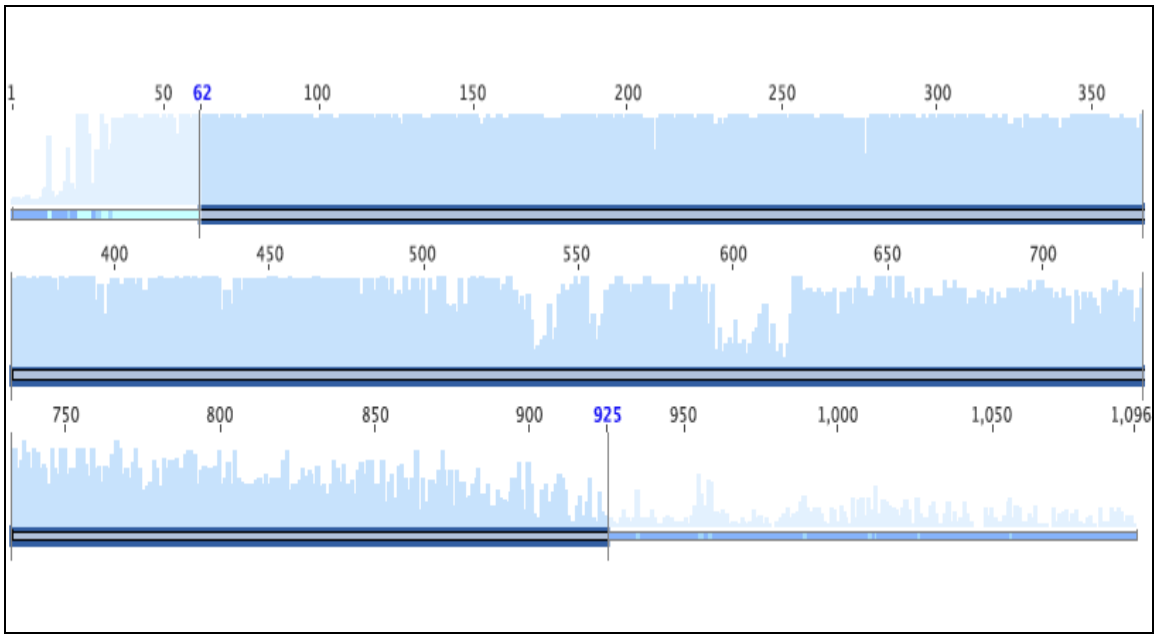


Figure 2.2: Image showing a read obtained from Sanger sequencing being edited by trimming off the ends, which are of bad quality. Bases from no. 62 to 925 form the section of the read that will be used in the analysis of interest (http://www.geneious.com).

The tables that follow (Table 2.9 to Table 2.11) shows all the primer reads that underwent trimming in order to extract the section of the reads that were of good quality. The tables show the length (bp), quality (%) and GC content (%) present in these reads before and after the trimming making it possible to see how the editing assists in increasing the quality of the reads.

Table 2.9: Quality and length of HLA Class 1A Sanger sequencing reads before and after editing using Geneious R8 v8.1.8.

Primer sequence	Before Editing			After Editing		
	Length (bp)	Quality (%)	GC conc. (%)	Length (bp)	Quality (%)	GC conc. (%)
HLA1_Class 1A_INT2R	828	54.3	70.7	569	79.1	70.8
HLA1_Class 1A_A1N110	1111	55.4	66.4	912	67.3	66.6
HLA1_Class 1A_5AE3.2	793	78.4	61.7	663	93.5	61.5
HLA1_Class 1A_3AE3.2	1096	65.3	68.9	880	81.3	68.9
HLA2_Class 1A_INT2R	1181	35.9	68.8	608	68.8	70.2
HLA2_Class 1A_A1N110	1107	51.9	67.7	886	63.1	68.2
HLA2_Class 1A_5AE3.2	1032	41.2	62.6	719	58.7	62.6
HLA2_Class 1A_3AE3.2	1111	58.4	69	887	72.3	69
HLA3_Class 1A_INT2R	752	50.5	69.8	509	73.1	69.9
HLA3_Class 1A_A1N110	1111	47	65.3	942	55.4	65.9
HLA3_Class 1A_5AE3.2	697	62.1	61.3	573	75.4	61.3
HLA3_Class 1A_3AE3.2	1101	52.3	67.2	858	67	67.2
HLA4_Class 1A_INT2R	927	42.0	70.5	718	56.5	70.5
HLA4_Class 1A_A1N110	1098	54.5	67.9	780	74.6	67.9
HLA4_Class 1A_5AE3.2	1110	50.5	60.2	649	86	60.2
HLA4_Class 1A_3AE3.2	1148	59.8	68.7	891	76.8	68.5
HLA5_Class 1A_INT2R	716	47.2	69.3	492	67.9	69.3
HLA5_Class 1A_A1N110	1095	52.6	67.3	888	64.2	67.3
HLA5_Class 1A_5AE3.2	685	76.8	61.4	643	81.2	61.3
HLA5_Class 1A_3AE3.2	1088	59.7	68.8	861	74.7	68.8
HLA6_Class 1A_INT2R	992	32.2	67.6	481	63.2	67.6
HLA6_Class 1A_A1N110	1080	51.3	66	835	63.4	66.7
HLA6_Class 1A_5AE3.2	859	40.7	62	685	50.9	62
HLA6_Class 1A_3AE3.2	819	67.6	68.3	769	71.8	68.3
HLA7_Class 1A_INT2R	825	53.9	70.4	563	78.7	70.3
HLA7_Class 1A_A1N110	1087	42.5	66.3	846	54.1	67
HLA7_Class 1A_5AE3.2	697	52.4	65.7	435	81.1	65.7
HLA7_Class 1A_3AE3.2	1113	59.7	68.1	842	77.4	68.1
HLA8_Class 1A_INT2R	655	65.2	70.2	521	77.4	70.2
HLA8_Class 1A_A1N110	1110	50.4	67.2	725	72.1	67.2
HLA8_Class 1A_5AE3.2	699	62.8	61	554	77.4	61.0
HLA8_Class 1A_3AE3.2	1134	53.5	67.8	832	70.6	67.8
HLA9_Class 1A_INT2R	1114	31.4	62	467	74.1	69.6
HLA9_Class 1A_A1N110	1105	54	67.9	809	70.8	67.9
HLA9_Class 1A_5AE3.2	1068	34.4	64.9	419	84.5	64.9
HLA9_Class 1A_3AE3.2	1049	66	68.2	853	80.2	68.2
HLA10_Class 1A_INT2R	656	45.7	69.6	460	63.7	69.6
HLA10_Class 1A_A1N110	1070	46.2	67.1	800	59.8	68.1
HLA10_Class 1A_5AE3.2	698	68.9	61.7	551	84.9	61.7
HLA10_Class 1A_3AE3.2	1093	59.7	68.8	856	76.1	68.8

Table 2.10: Quality and length of HLA Class 1B Sanger sequencing reads before and after editing using Geneious R8 v8.1.8.

Primer sequence	Before Editing			After Editing		
	Length (bp)	Quality (%)	GC conc. (%)	Length (bp)	Quality (%)	GC conc. (%)
HLA1_Class 1B_INT2R	1203	36.2	69.3	561	77.4	69.3
HLA1_Class 1B_3B1	1174	46.5	66.7	775	70	68,8
HLA1_Class 1B_5BE2.2	1163	32.8	69.1	474	79.3	69.2
HLA1_Class 1B_3BE3.2	1112	43.3	69.2	859	56	69.6
HLA2_Class 1B_INT2R	1244	24.8	67.5	416	72.1	69
HLA2_Class 1B_3B1	1059	24.1	67.1	386	65.5	68.9
HLA2_Class 1B_5BE2.2	1146	19.7	66.7	387	56.3	70
HLA2_Class 1B_3BE3.2	1021	21.3	66.8	392	55.1	67.7
HLA3_Class 1B_INT2R	980	21.9	68.2	321	67	67
HLA3_Class 1B_3B1	1148	30.3	65.8	457	74.4	70.5
HLA3_Class 1B_5BE2.2	913	17.2	65.6	232	66.8	66.4
HLA3_Class 1B_3BE3.2	1161	26.6	63.3	448	67.9	69.9
HLA4_Class 1B_INT2R	1111	39.2	68.6	558	78.1	68.6
HLA4_Class 1B_3B1	1157	19.4	65.7	643	34.7	65.8
HLA4_Class 1B_5BE2.2	514	70.2	68.8	478	75.1	68.8
HLA4_Class 1B_3BE3.2	1161	20.5	64	616	38.3	65.4
HLA5_Class 1B_INT2R	581	78.8	68.8	560	81.4	68.6
HLA5_Class 1B_3B1	1156	32.6	65.5	532	69.9	69.5
HLA5_Class 1B_5BE2.2	514	69.3	69.8	477	74.2	69.8
HLA5_Class 1B_3BE3.2	1125	31.2	66.4	434	80.6	69.8
HLA6_Class 1B_INT2R	1200	48.8	68	625	92.8	68
HLA6_Class 1B_3B1	1158	67.0	70.5	928	82.9	70.6
HLA6_Class 1B_5BE2.2	1044	49.5	67.9	542	94.3	68.1
HLA6_Class 1B_3BE3.2	1135	68.8	70.5	986	79.2	70.5
HLA7_Class 1B_INT2R	1094	51.4	67	634	88.5	67
HLA7_Class 1B_3B1	1125	71.7	70.4	932	85.7	70.4
HLA7_Class 1B_5BE2.2	775	63.1	68.5	487	97.5	69
HLA7_Class 1B_3BE3.2	1143	66.1	70.5	958	78.6	70.7
HLA8_Class 1B_INT2R	1003	35.2	66	472	73.8	69.7
HLA8_Class 1B_3B1	1175	53.6	69.6	909	68.9	69.5
HLA8_Class 1B_5BE2.2	501	57.1	69.1	472	60.6	69.1
HLA8_Class 1B_3BE3.2	1157	53.3	69.6	904	67.9	69.9
HLA9_Class 1B_INT2R	1188	32.3	68.4	556	67.4	68.3
HLA9_Class 1B_3B1	1016	23.6	64.9	345	69	70.1
HLA9_Class 1B_5BE2.2	1134	26.6	67.8	480	62.3	67.9
HLA9_Class 1B_3BE3.2	1004	23.3	65.5	326	70.9	70.9
HLA10_Class 1B_INT2R	677	47.9	69	560	57.4	69
HLA10_Class 1B_3B1	1113	21.3	65.7	420	54.3	68.1
HLA10_Class 1B_5BE2.2	575	41	66.1	310	71.9	72.9
HLA10_Class 1B_3BE3.2	972	21.4	67.2	321	64.2	70.4

Table 2.11: Quality and length of HLA Class 1C Sanger sequencing reads before and after editing using Geneious R8 v8.1.8.

Primer sequence	Before Editing			After Editing		
	Length (bp)	Quality (%)	GC conc. (%)	Length (bp)	Quality (%)	GC conc. (%)
HLA1_Class 1C_INT2C	553	62.6	70.1	435	79.3	69.7
HLA1_Class 1C_BC3N1P3	1128	23.9	65.2	478	54.2	66.7
HLA1_Class 1C_5CE2.2	1188	20.3	62	705	34.1	68.3
HLA1_Class 1C_3CE3.2	1107	0,7	63.6	435	0.7	63
HLA2_Class 1C_INT2C	456	80.7	70.6	435	84.6	70.6
HLA2_Class 1C_BC3N1P3	908	27	65.1	300	81.7	71
HLA2_Class 1C_5CE2.2	597	42.4	69.6	391	64.7	69.6
HLA2_Class 1C_3CE3.2	271	0	63.2	425	1.2	66.1
HLA3_Class 1C_INT2C	456	73.9	69.4	435	77.5	69.4
HLA3_Class 1C_BC3N1P3	951	31.7	64.6	582	50.5	66.5
HLA3_Class 1C_5CE2.2	401	1	58.7	188	1.6	58.5
HLA3_Class 1C_3CE3.2	957	2.7	65.1	422	5.5	67.5
HLA4_Class 1C_INT2C	527	71.7	70.6	432	87	70.6
HLA4_Class 1C_BC3N1P3	941	76.6	69.8	893	80.6	69.8
HLA4_Class 1C_5CE2.2	871	37.3	66.7	801	40.4	66.7
HLA4_Class 1C_3CE3.2	946	10.8	69.2	853	12	69.4
HLA5_Class 1C_INT2C	458	83.6	70.3	435	88	70.3
HLA5_Class 1C_BC3N1P3	908	28.4	64.8	350	72.9	67.7
HLA5_Class 1C_5CE2.2	560	45	69.8	460	51.5	69.8
HLA5_Class 1C_3CE3.2	951	1.3	63	424	1.9	65.6
HLA6_Class 1C_INT2C	717	53.7	68.6	542	71	68.5
HLA6_Class 1C_BC3N1P3	941	76.8	69.9	898	80.4	70.3
HLA6_Class 1C_5CE2.2	932	42.5	68.5	844	46.9	68.5
HLA6_Class 1C_3CE3.2	1140	5.1	68	836	6.9	67.9
HLA7_Class 1C_INT2C	5	-	100	5	-	0
HLA7_Class 1C_BC3N1P3	325	0	51.1	11	0	51.1
HLA7_Class 1C_5CE2.2	5	-	100	5	-	0
HLA7_Class 1C_3CE3.2	210	0	57.1	58	0	57.1
HLA8_Class 1C_INT2C	531	62.5	70.8	425	77.9	71.1
HLA8_Class 1C_BC3N1P3	955	24.3	65.4	352	65.9	66.8
HLA8_Class 1C_5CE2.2	560	50.2	66.2	491	56.4	66.2
HLA8_Class 1C_3CE3.2	947	1.4	66	413	2.4	62.0
HLA9_Class 1C_INT2C	454	74.9	69.9	434	78.3	69.6
HLA9_Class 1C_BC3N1P3	950	32.3	63.8	553	54.8	67.3
HLA9_Class 1C_5CE2.2	637	42.2	69.5	622	43.2	69.5
HLA9_Class 1C_3CE3.2	955	3.7	63.7	587	5.8	66.3
HLA10_Class 1C_INT2C	462	72.9	71.9	427	78.9	71.9
HLA10_Class 1C_BC3N1P3	939	70.8	69.6	888	74.9	69.9
HLA10_Class 1C_5CE2.2	480	48.7	69.5	452	48.9	69.5
HLA10_Class 1C_3CE3.2	949	10.9	67.3	850	12.1	68

In all three tables above it is evident that the editing/trimming of the reads is beneficial to their quality (Table 2.9 to 2.11). This is because the quality percentage increased in all of the primer reads that were trimmed except those of HLA7_Class1C (Table 2.11). This sample failed to sequence; some of its sequenced primer reads were too short in length making it impossible to edit e.g. HLA7_Class1C_INT2C, HLA7_Class 1C_BC3N1P3 and HLA7_Class 1C_5CE2.2 (Table 2.11).

2.8.2 De novo Assembly

Sequence assembly refers to a process where fragments (reads) of a longer DNA sequence are aligned and merged to reconstruct the original sequence. A reconstructed sequence is called a consensus sequence, which is the calculated order of most nucleotides that are found in each position in a sequence alignment. The de novo assembly option was used to assemble the 4 sequencing primer reads into a consensus sequence (as seen in Figure 2.3). Steps followed in mapping the sequences using the de novo option into a consensus sequence are summarized in Appendix A.4.1.

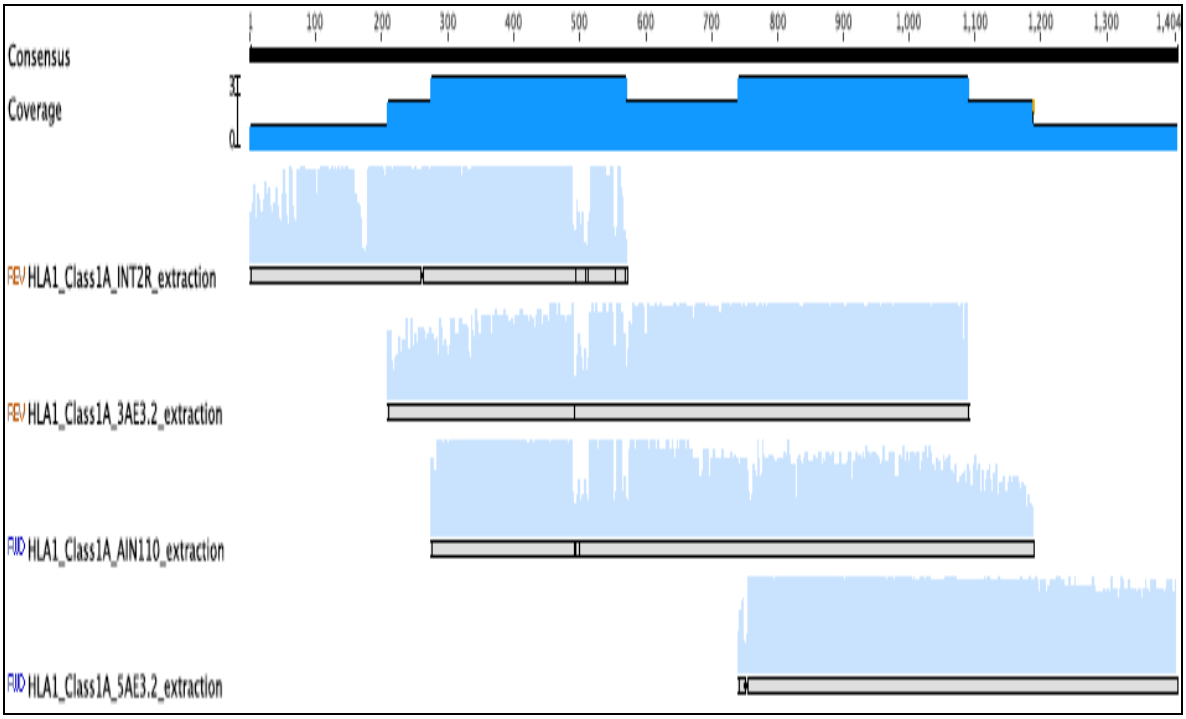


Figure 2.3: Four reads of Sample #1 (HLA Class 1A) mapped into a consensus sequence using the de novo option in Geneious R8 v 8.1.8 (http://www.geneious.com).

The map to reference option is used when you have a known sequence and you wish to compare a number of reads of the same sequence with it to locate differences. Some reads could not be mapped using the de novo option and therefore had to be mapped to a known reference gene as a guide to obtain a consensus sequence (Figure 2.4).

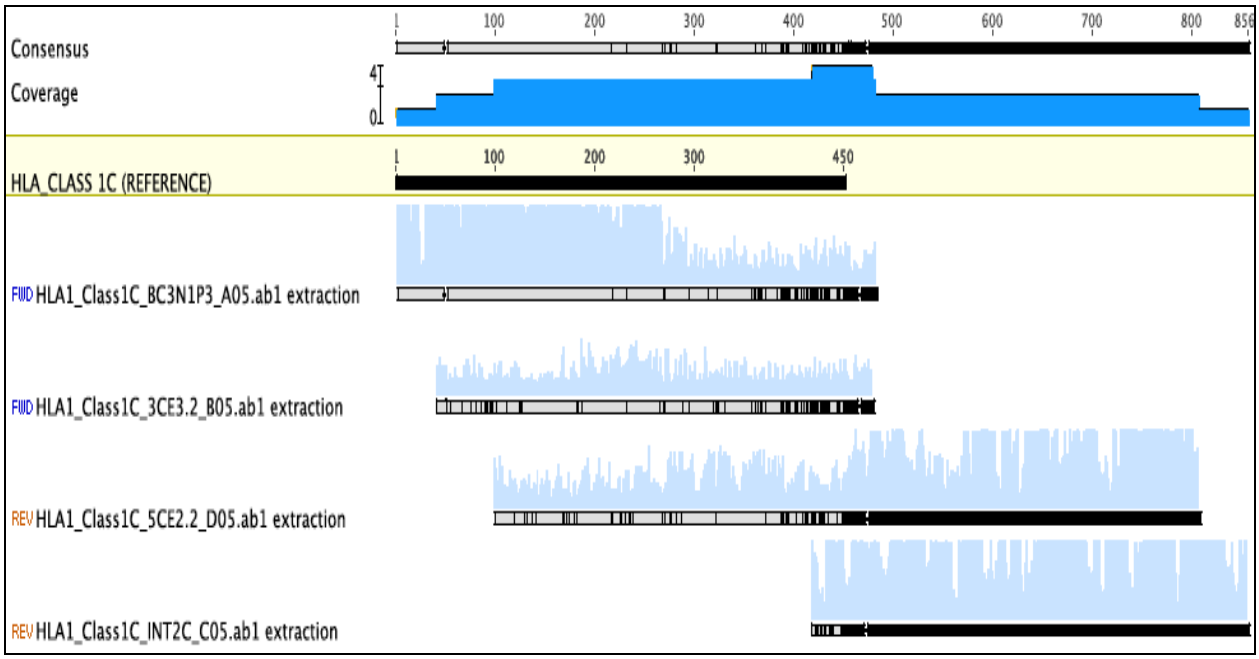


Figure 2.4: Four reads of Sample #1 (HLA Class 1C) mapped to a known reference gene as a guide in constructing the consensus sequence (http://www.geneious.com).

The newly constructed sequences were thereafter exported into PubMed NCBI BLAST for evaluation.

2.8.3 PubMed NCBI BLAST

To evaluate whether the newly constructed sequences fall under the HLA Class 1 loci, the consensus sequences were exported into PubMed NCBI BLAST where they were aligned against the non redundant (NR) database of NCBI. The specificity of the consensus sequences was evaluated by blasting them against the database (www.ncbi.nlm.nih.gov) using the standard nucleotide BLAST (Figure 2.5).

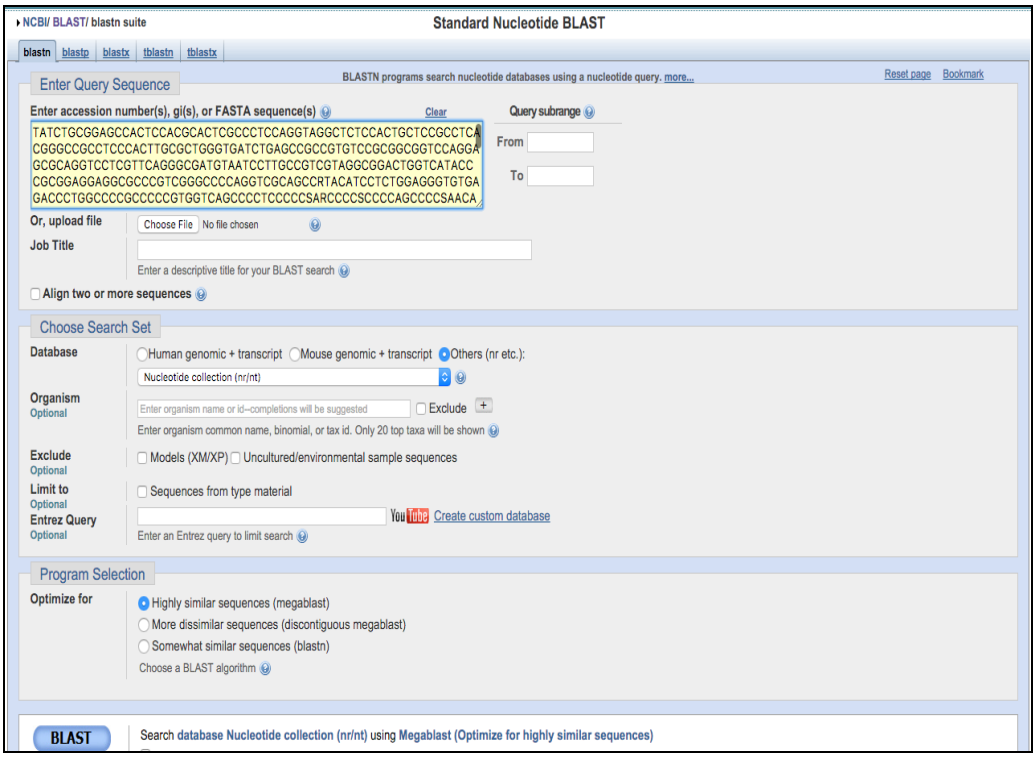


Figure 2.5: Image showing the evaluation specificity used in PubMed NCBI BLAST to align the newly constructed consensus sequences against the non redundant (NR) database of NCBI (www.ncbi.nlm.nih.gov).

2.8.4 HLA sequence analysis

Kamini Gounder, a research Scientist at the UKZN HIV Pathogenesis Program (HPP) situated in the Doris Duke Medical Research Institute (DDMRI) at the University of KwaZulu-Natal's Nelson R. Mandela School of Medicine Campus performed the HLA sequence analysis on the Sanger sequences. The Assign^{TM} ATF Softwere v1.5 (Conexio Genomics, Australia) was used to detect variation of HLA Class 1A, 1B and 1C alleles in the reads of the 10 sequenced samples.

3. RESULTS

3.1 Evaluation of DNA Extraction

Prior to PCR amplification, the quality of all DNA samples was validated by measuring the DNA concentration using the Qubit 2.0 Flourometer (Table 3.1) and by DNA gel electrophoresis (Figure 3.1).

Of the 10 samples that were evaluated, DNA yield could only be measured in 6 using the Qubit 2.0 Flourometer (Qubit Assay). In total, the 6 samples with a positive result had DNA yields mean concentration of 1.85 ng/µL (varying from 1.20 ng/µL to 3.40 ng/µL). The Qubit 2.0 Flourometer was not able to measure DNA concentrations in samples 4, 5, 7 and 9. Samples that showed invalid concentrations were denoted by a (-) in Table 3.1.

Table 3.1: The DNA yield (ng /µL) in samples extracted from Dried Blood spots using the QIAamp® DNA Mini Kit (QIAGEN) measured using the Qubit 2.0 Fluorometer (Qubit Assay).

Sample no.	Original Name	DNA [] ng/µL	Kit used
1	#1	3.40	QIAamp DNA Mini Kit
2	#2	2.56	QIAamp DNA Mini Kit
3	#3	1.49	QIAamp DNA Mini Kit
4	#4	-	QIAamp DNA Mini Kit
5	#5	-	QIAamp DNA Mini Kit
6	#6	1.57	QIAamp DNA Mini Kit
7	#7	-	QIAamp DNA Mini Kit
8	#8	1.20	QIAamp DNA Mini Kit
9	#9	-	QIAamp DNA Mini Kit
10	#10	1.62	QIAamp DNA Mini Kit
11	+ve (HIV –ve blood)	51.0	QIAamp DNA Mini Kit

One microlitre (µl) of DNA extracted using the QIAamp® DNA Mini Kit (QIAGEN) method was run on a 1% agarose gel in order to determine if DNA bands could be visualized. Although no DNA could be measured for samples 4, 5, 7 and 9 by the Qubit 2.0 Flourometer, the gel image (Figure 3.1) shows that there was DNA in those samples, but in low amounts. The low quantity was most probably due to the low quantity of DNA that can be stored in one dried blood spot.

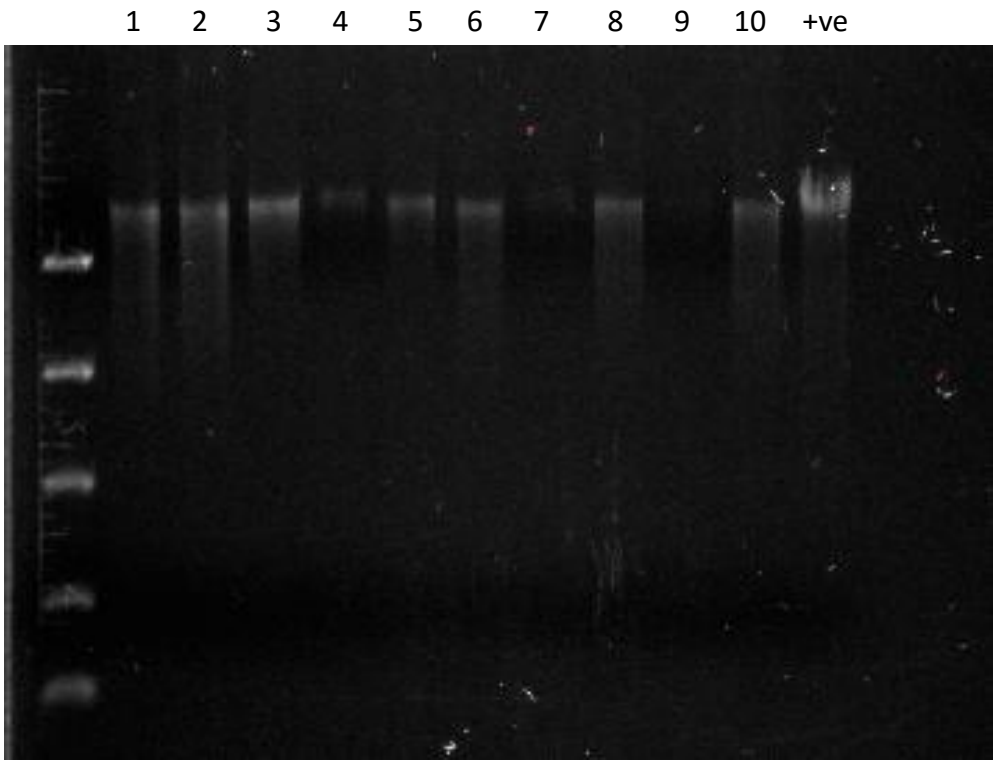


Figure 3.1: Image of an agarose gel electrophoresis showing the presence of DNA in samples extracted from Dried Blood Spots and the positive control. 1-10: extracted samples; +ve: positive control (HIV negative blood).

3.2 PCR Amplification and PCR product purification

Nested polymerase chain reactions (nPCR) were used to amplify HLA Class 1A, 1B and 1C loci using the KAPPA2G™ Robust HotStart ReadyMix 2X (KAPABIOSYSTEMS) kit. The PCR products for each class were separated on 1% agarose gel a with a 200bp O'RangeRuler™ DNA ladder mix molecular marker. Amplification of good quality DNA is denoted by the presence of one distinct band (1000 bp) on the gel representing either HLA

Class 1A, 1B or 1C as seen on (Figure 3.2). From the 10 samples that were amplified, amplification could be seen in all samples except sample 2 for HLA Class 1A, in all samples for HLA Class 1B, and in all except sample 7 in HLA Class 1C.

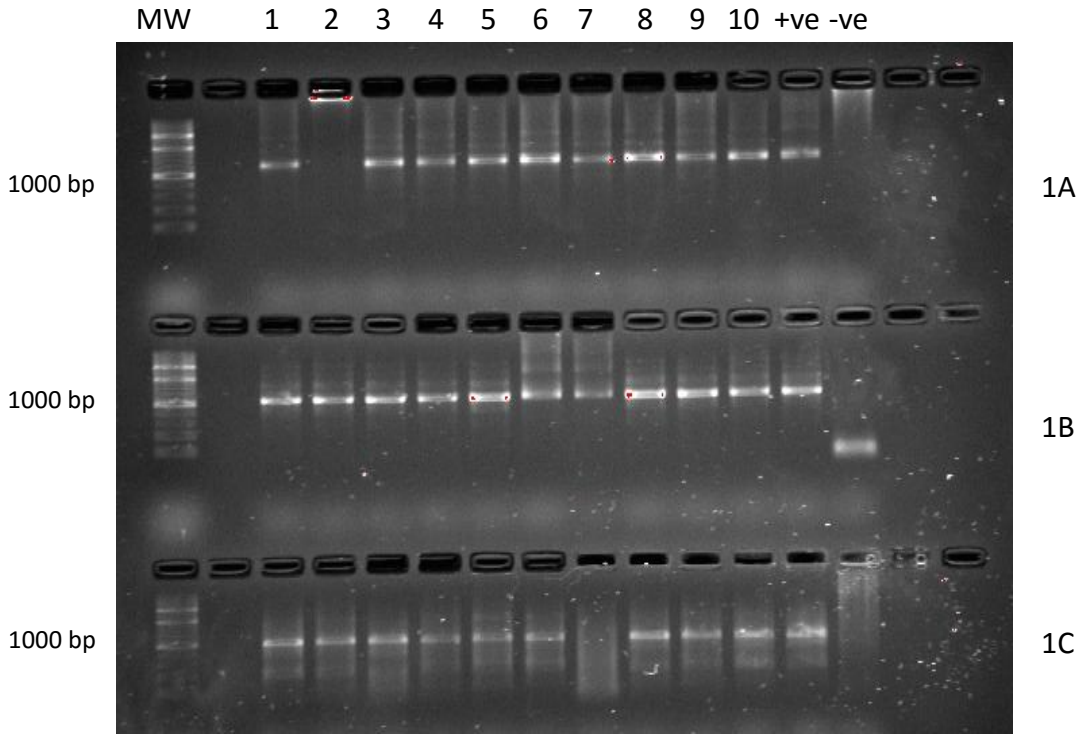


Figure 3.2: Image of an agarose gel showing PCR amplification of HLA Class 1A, 1B and 1C using a sample volume of 5 µl in step 1 and 2.5 µl in step 2 and a T_m of 63°C in step 1 and 65°C in step 2. MW: O'RangeRuler^{TM} DNA Ladder molecular marker; 1-10: extracted samples; bp: base pairs of the product size; 1A-1C: HLA Class 1 alleles; +ve: positive control (HIV negative blood); -ve: negative control (water).

PCR purification was performed on the PCR product to purify the DNA in the completed PCR reaction by removing the remaining dNTPs, primers, Taq, and Mg^{2+} ions. These components are all required for the PCR reaction, but post amplification, they are considered contaminants or impurities because they may interfere with subsequent manipulations such as DNA sequencing or restriction digests. Extracted samples were purified using the Pure Link Purification Kit (Invitrogen Life Technologies) and were thereafter subjected to 1% agarose gel electrophoresis to check if the samples still produced bands of the expected size (1000 bp).

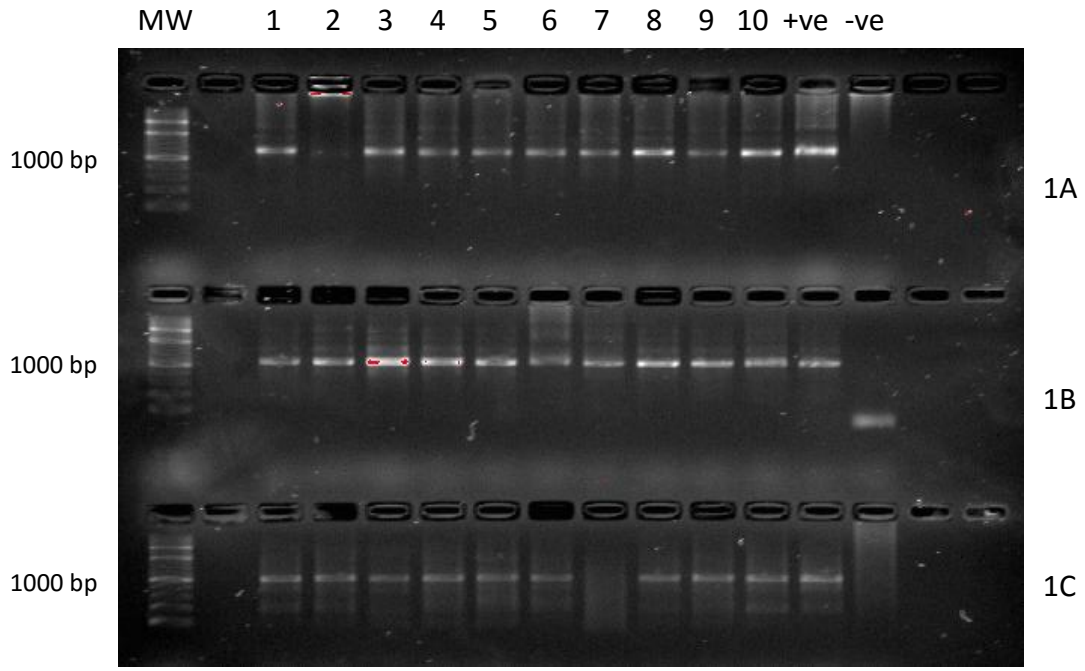


Figure 3.3: Electrophoresis gel showing PCR amplification of HLA Class 1A, 1B and 1C using a sample volume of 2.5 µl in step 2 and a T_m of 63°C in step 1 and 65°C in step 2 after purification using the Pure Link Purification Kit (Invitrogen Life Technologies). MW: O'RangeRuler^{TM} DNA Ladder molecular marker; 1-10: extracted samples; bp – base pairs of the product size; 1A-1C: HLA Class 1 alleles; +ve: positive control (HIV –ve blood); -ve: negative control (water).

The extraction worked for all samples because as seen in Figure 3.3 after PCR purification all the PCR products still produced the expected product size for all HLA Class 1 loci. Sample 2 which had previously failed to produce a band for HLA Class 1A now produced a band even though it was very faint. However, sample 7 failed to produce the expected band for HLA Class 1C.

3.3 Sequence quality control

3.3.1 Assembly

The PCR products were sequenced using the ABI Prism BigDye Terminator kit v3.1 Cycle Sequencing Kit (USA) using 4 primers for each sample in a Sanger ABI 3100 Genetic Analyzer. Geneious R8 v8.1.8 was used to map the four sequences into a consensus sequence using the de novo assembly option (Figure 2.3). The following results were obtained:

Table 3.2: Quality of the consensus sequences produced by mapping the four sequencing primer reads from Sanger sequencing using Geneious R8 v8.1.8 for each of the 10 donor samples under the three HLA alleles (HLA Class 1A, 1B and 1C).

Donor Sample	**Consensus sequences**	**After sequence mapping**		
		Length (bp)	**Quality (%)**	**GC conc. (%)**
#1	HLA1_Class 1A	1404	91	66
	HLA1_Class 1B	1016	92.2	69.8
	HLA1_Class 1C	853	80	70.4
#2	HLA2_Class 1A	1521	76.1	66.6
	HLA2_Class 1B	1093	73.1	67.0
	HLA2_Class 1C	863	74.7	67.5
#3	HLA3_Class 1A	1258	87.4	65.5
	HLA3_Class 1B	784	80.9	65.4
	HLA3_Class 1C	863	74.4	65.1
#4	HLA4_Class 1A	1553	77.8	66.2
	HLA4_Class 1B	1065	69.4	66.3
	HLA4_Class 1C	914	96.1	69.9
#5	HLA5_Class 1A	1321	87	65.4
	HLA5_Class 1B	1060	87.1	67.3
	HLA5_Class 1C	901	72.8	65.8
#6	HLA6_Class 1A	1357	72	64.9
	HLA6_Class 1B	1119	97.9	69.1
	HLA6_Class 1C	968	94.3	69.0
#7	HLA7_Class 1A	1170	90.7	68.3
	HLA7_Class 1B	1115	94.1	69
	HLA7_Class 1C	**SEQUENCING FAILED**		
#8	HLA8_Class 1A	1273	89.9	67
	HLA8_Class 1B	1233	84.7	68.7
	HLA8_Class 1C	901	78.9	69.3
#9	HLA9_Class 1A	1088	90	67.7
	HLA9_Class 1B	1042	80.2	66.3
	HLA9_Class 1C	871	85.6	65.7
#10	HLA10_Class 1A	1219	85.2	66.2
	HLA10_Class 1B	847	70.8	66.9
	HLA10_Class 1C	885	90.0	69.6

As mentioned in section 2.8.2 of the Materials and Methods, some of the primer reads could not be mapped into a consensus sequence using the de novo option and therefore had to be mapped to a known reference gene. All consensus sequences constructed with assistance of a known reference gene are highlighted (light grey) in the table (Table 3.2). Looking at the table (Table 3.2), it is evident that most of the constructed consensus sequences were of good

quality as they all had a quality score above 70 with exception of 1 sequence (#4 HLA4_Class 1B).

3.3.2 PubMed NCBI BLAST

The newly constructed sequences were exported into PubMed NCBI BLAST (Basic Local Alignment Search Tool). NCBI BLAST was used to align these new sequences against the non redundant (NR) database of NCBI. This was done in order to evaluate if the consensus sequences fall under the HLA Class 1 loci that have been previously described.

Table 3.3: Details of the sequences that match the GenBank non redundant (NR) database in BLAST.

Donor Sample	Consensus sequences	Name of matched variant	Ident (%)
#1	HLA1_Class 1A	Homo sapiens HLA-A gene for MHC class I antigen, allele HLA-A*32:01:01	99
		Homo sapiens HLA-A gene for MHC Class I antigen, HLA-A*74:01 allele, exons 1-8	99
	HLA1_Class 1B	Homo sapiens HLA-B gene for MHC class I antigen, allele HLA-B*58	97
		Homo sapiens HLA-B gene for MHC class I antigen, cell line DKMS-LSL-B-691, allele HLA-B*15:new	97
	HLA1_Class 1C	Homo sapiens MHC class I antigen (HLA-Cw) gene, HLA-Cw*0202var allele, exons 2, 3 and partial cds	94
		Homo sapiens HLA-C gene for MHC class I antigen, allele HLA-C*06:02:01var	95
#2	HLA2_Class 1A	Homo sapiens HLA-A gene, HLA-A*0205 allele, exons 1-8	98
		Homo sapiens HLA-A gene for MHC Class I antigen, HLA-A*66:01:01 allele, exons 1-8	98
	HLA2_Class 1B	Homo sapiens MHC class I antigen (HLA-B) gene, HLA-B*58:01:01 allele, complete cds	92
		Homo sapiens HLA-B gene for MHC class I antigen, allele HLA-B*35:01:01:02	92
	HLA2_Class 1C	Homo sapiens HLA-C gene for MHC class I antigen, allele HLA-C*12:03:01:01	95
		Homo sapiens MHC class I antigen (HLA-C) gene, HLA-C*17:01:01:03 allele, complete cds	97
#3	HLA3_Class1A	Homo sapiens, HLA-A gene for MHC class I antigen, allele HLA-A*30:02:01:01	98
		Homo sapiens HLA-A gene for MHC class I antigen, allele HLA-A*03:01:01:01	97
	HLA3_Class 1B	Homo sapiens HLA-B gene for MHC class I antigen, allele HLA-B*18:01:01:02	93
		Homo sapiens partial HLA-B gene for MHC class I antigen, allele HLA-B*27:05:02	89
	HLA3_Class 1C	Homo sapiens HLA-C gene for MHC class I antigen, cell line DKMS-LSL-560, allele HLA-C*07:new	96
		Homo sapiens isolate NT01266 MHC class I antigen (HLA-C) gene, HLA-C*07:01:01:01V allele, complete cds	95
#4	HLA4_Class 1A	Homo sapiens HLA-A gene for MHC class I antigen, HLA*0201 allele, exons 1-8	99
		Homo sapiens HLA-A gene for MHC class I antigen, HLA-A*6802 allele	98
	HLA4_Class 1B	Homo sapiens HLA-B gene for MHC class I antigen, allele HLA-B*14:01:01	92
		Homo sapiens HLA-B gene for MHC class I antigen, allele HLA-B*45:01	93
	HLA4_Class 1C	Homo sapiens HLA-C gene for MHC class I antigen, HLA-C*08:02:01:02, exons 1-8	98
		Homo sapiens HLA-C gene for MHC class I antigen, allele HLA-C*05:01:01:02	95
#5	HLA5_Class 1A	Homo sapiens HLA-A gene for, A*2901 allele, exons 1-8	97
		Homo sapiens HLA-A gene for MHC Class I antigen, HLA-A*66:01:01 allele, exons 1-8	96
	HLA5_Class 1B	Homo sapiens HLA-B gene for MHC class I antigen, allele HLA-B*44:03:01	96
		Homo sapiens HLA-C gene for MHC class I antigen, allele HLA-C*06:02:01var	96
	HLA5_Class 1C	Homo sapiens HLA-C gene for MHC class I antigen, allele HLA-C*06:02:01var	96
		Homo sapiens isolate BY01269 MHC class I antigen (HLA-C) gene, HLA-C*12:59 allele, complete cds	95

#6	HLA6_Class 1A	Homo sapiens HLA-A gene for MHC class I antigen, allele HLA-A*29:01:01:01	95
		Homo sapiens MHC class I antigen (HLA-A) gene, HLA-A*68:01:01V allele, complete cds	95
	HLA6_Class 1B	Homo sapiens MHC class I antigen (HLA-B) gene, HLA-B*58:01:01 allele, complete cds	99
		Homo sapiens HLA-B gene for MHC class I antigen, HLA-B*5711 allele, exon 1-7	98
	HLA6_Class 1C	Homo sapiens HLA-C gene for MHC class I antigen, allele HLA-C*06:02:01var	99
		Homo sapiens HLA-C gene for MHC class I antigen, allele HLA-C*06:02:01var	99
#7	HLA7_Class 1A	Homo sapiens HLA-A gene for MHC class I antigen, HLA-A*01 null variant allele	99
		Homo sapiens HLA-A gene for MHC class I antigen, allele HLA-A*03:01:01:01	98
	HLA7_Class 1B	Homo sapiens HLA-B gene for MHC class I antigen, allele HLA-B*08:01:01	99
		Homo sapiens HLA-B gene for MHC class I antigen, allele HLA-B*08:01:01	99
	HLA7_Class 1C	**CONSENSUS SQUENCE MAPPING FAILED**	
#8	HLA8_Class 1A	Homo sapiens HLA-A gene for MHC class I antigen, allele HLA-A*01:01:01:01	99
		Homo sapiens MHC class I antigen (HLA-A) gene, HLA-A*36:01 allele, complete cds	98
	HLA8_Class 1B	Homo sapiens MHC class I antigen (HLA-B) gene, HLA-B*15:10:01 allele, complete cds	97
		Homo sapiens HLA-B gene for MHC class I antigen, HLA-B*07:05:01 allele, exon 1-7	96
	HLA8_Class 1C	Homo sapiens MHC class I antigen (HLA-C) gene, HLA-C*03:02:02 allele, complete cds	94
		Homo sapiens MHC class I antigen (HLA-C) gene, HLA-C*03:05 allele, complete cds	94
#9	HLA9_Class 1A	Homo sapiens HLA-A gene for MHC class I antigen, allele HLA-A*03:01:01:01	98
		Homo sapiens MHC class I antigen (HLA-A) gene, HLA-A*74:02V allele, complete cds	97
	HLA9_Class 1B	Human MHC class I antigen (HLA-B) mRNA, allele HLA-B*1522, complete cds	97
		Homo sapiens HLA-B gene, exons 2 and 3, B*1805	99
	HLA9_Class 1C	Homo sapiens HLA-C gene for MHC class I antigen, complete cds, allele: HLA-C*07:02:01:05	94
		Homo sapiens HLA-C gene for MHC class I antigen, allele HLA-C*07:22	96
#10	HLA10_Class 1A	Homo sapiens HLA-A gene for MHC class I antigen, allele HLA-A*23:01:01	98
		Homo sapiens HLA-A gene for MHC class I antigen, allele HLA-A*24:02:01:01	98
	HLA10_Class 1B	Homo sapiens HLA-B gene for MHC class I antigen, HLA-B*4802 allele	96
		Homo sapiens HLA-B gene, exons 2 and 3, B*1805	99
	HLA10_Class 1C	Homo sapiens HLA-C gene for MHC class I antigen, allele HLA-C*06:02:01var	96
		Homo sapiens MHC class I antigen (HLA-C) gene, HLA-C*17 allele, complete cds	96

3.3.3 HLA typing analysis

The Assign™ ATF Software v1.5 (Conexio Genomics, Australia) was used for HLA typing analysis of the sequenced reads. This software can be used for an extensive range of sequencing applications and also for producing information in a unique manner for quality control. In this study, the software was used to assemble sequences according to the HLA alleles and detect variation of HLA Class 1A, 1B and 1C alleles within the reads of the 10 sequenced samples. The following HLA results were extracted:

Table 3.4: Results from the Assign™ ATF Software v1.5 (Conexio Genomics, Australia) used to genotype and detect variation in HLA Class 1A, 1B and 1C alleles for the 10 sequenced samples. The numbers represent the allele and its subtype that each sequence matched to in the HLA Class 1 gene.

Sample no.	HLA-A	HLA-A	HLA-B	HLA-B	HLA-C	HLA-C
1	32:01	74:01	15	58	02:02	06:02
2	02:05	66:01	39:01	58:01	07:01	12:03
3	26:01	30:02	18:01	27:05	02:02	07:04
4	02:01	68:02	14:01	45	**FAILED**	
5	29:01	66:01	44:03	58:02	06:02	07:01
6	29:02	68:01	58:02	58:02	06:02	06:02
7	01:01	03:01	08:01	08:01	**FAILED**	
8	01:01	68:02	15:10	81:01	**FAILED**	
9	03:01	74:02	08:01	15:03	02:10	07:02
10	23:01	29:02	42:01	58:02	06:02	12/17

The HLA gene has many possible variations; hundreds of identified versions of a specific gene (known as alleles) are all given a particular number to be identified by. Alleles closely related to eachother are grouped together as subtypes of that allele. For example looking at the results under the HLA-A gene for sample 1 (Table 3.4), the results show that the sequences of that sample matched to HLA-A*3201; subtype 01 of allele 32 of the HLA-A gene.

The above results show that majority of the Sanger sequenced sequences were of good quality as variants could classify the major and minor HLA classification. For example, we could classify all samples for HLA Class 1A; for HLA Class 1B 8 out of 10 of the samples could identify the major and minor allele; and 50% for HLA Class 1C. Sequences that could not

identify the minor allele were highlighted in light grey in Table 3.4. Three of the sequences from HLA Class 1C could not identify any HLA typing (denoted by '**FAILED'** in Table 3.4).

3.3.4 PubMed NCBI BLAST result comparison with the Assign^{TM} ATF Software v1.5

The ATF classification and BLAST results were compared. This was done in order to evaluate if the consensus sequences fall under the HLA Class 1 loci and further see if they match the variant results produced in the HLA typing analysis using the Assign^{TM} ATF Software v1.5 (Table 3.4).

All BLASTed sequences were variants of the HLA Class 1 loci (either HLA Class 1A, 1B or 1C). As seen in Table 3.5, the majority of the variants were common between the PubMed NCBI BLAST and the ATF Software v1.5. The variants for some samples however, did not align. These samples are denoted by a dash (-) in the table.

Table 3.5: Details of the sequences that the NCI BLAST matches were identical to the Assign™ ATF Software v1.5 classification.

Donor Sample	Consensus sequences	Name of matched variant	Ident (%)
#1	HLA1_Class 1A	Homo sapiens HLA-A gene for MHC class I antigen, allele HLA-A*32:01:01	99
		Homo sapiens HLA-A gene for MHC Class I antigen, HLA-A*74:01 allele, exons 1-8	99
	HLA1_Class 1B	Homo sapiens HLA-B gene for MHC class I antigen, allele HLA-B*58	97
		Homo sapiens HLA-B gene for MHC class I antigen, cell line DKMS-LSL-B-691, allele HLA-B*15:new	97
	HLA1_Class 1C	Homo sapiens MHC class I antigen (HLA-Cw) gene, HLA-Cw*0202var allele, exons 2, 3 and partial cds	94
		Homo sapiens HLA-C gene for MHC class I antigen, allele HLA-C*06:02:01var	95
#2	HLA2_Class 1A	Homo sapiens HLA-A gene, HLA-A*0205 allele, exons 1-8	98
		Homo sapiens HLA-A gene for MHC Class I antigen, HLA-A*66:01:01 allele, exons 1-8	98
	HLA2_Class 1B	Homo sapiens MHC class I antigen (HLA-B) gene, HLA-B*58:01:01 allele, complete cds	92
		-	-
	HLA2_Class 1C	Homo sapiens HLA-C gene for MHC class I antigen, allele HLA-C*12:03:01:01	95
		-	-
#3	HLA3_Class1A	Homo sapiens, HLA-A gene for MHC class I antigen, allele HLA-A*30:02:01:01	98
		-	-
	HLA3_Class 1B	Homo sapiens HLA-B gene for MHC class I antigen, allele HLA-B*18:01:01:02	93
		Homo sapiens partial HLA-B gene for MHC class I antigen, allele HLA-B*27:05:02	92
	HLA3_Class 1C	-	-
		-	-
#4	HLA4_Class 1A	Homo sapiens HLA-A gene for MHC class I antigen, HLA*0201 allele, exons 1-8	99
		Homo sapiens HLA-A gene for MHC class I antigen, HLA-A*6802 allele	98
	HLA4_Class 1B	Homo sapiens HLA-B gene for MHC class I antigen, allele HLA-B*14:01:01	92
		Homo sapiens HLA-B gene for MHC class I antigen, allele HLA-B*45:01	93
	HLA4_Class 1C	**VARIANT ANALYSIS FAILED**	
#5	HLA5_Class 1A	Homo sapiens HLA-A gene for, A*2901 allele, exons 1-8	97
		Homo sapiens HLA-A gene for MHC Class I antigen, HLA-A*66:01:01 allele, exons 1-8	96
	HLA5_Class 1B	Homo sapiens HLA-B gene for MHC class I antigen, allele HLA-B*44:03:01	96
		-	-
	HLA5_Class 1C	Homo sapiens HLA-C gene for MHC class I antigen, allele HLA-C*06:02:01var	96
		-	-
	HLA6_Class 1A	Homo sapiens HLA-A gene for MHC class I antigen, allele HLA-A*29:01:01:01	95

#6		Homo sapiens MHC class I antigen (HLA-A) gene, HLA-A*68:01:01V allele, complete cds	95
	HLA6_Class 1B	-	-
		-	-
	HLA6_Class 1C	Homo sapiens HLA-C gene for MHC class I antigen, allele HLA-C*06:02:01var	99
		Homo sapiens HLA-C gene for MHC class I antigen, allele HLA-C*06:02:01var	99
#7	HLA7_Class 1A	Homo sapiens HLA-A gene for MHC class I antigen, HLA-A*01 null variant allele	99
		Homo sapiens HLA-A gene for MHC class I antigen, allele HLA-A*03:01:01:01	98
	HLA7_Class 1B	Homo sapiens HLA-B gene for MHC class I antigen, allele HLA-B*08:01:01	99
		Homo sapiens HLA-B gene for MHC class I antigen, allele HLA-B*08:01:01	99
	HLA7_Class 1C	**VARIANT ANALYSIS FAILED**	
#8	HLA8_Class 1A	Homo sapiens HLA-A gene for MHC class I antigen, allele HLA-A*01:01:01:01	99
		-	-
	HLA8_Class 1B	Homo sapiens MHC class I antigen (HLA-B) gene, HLA-B*15:10:01 allele, complete cds	97
		-	-
	HLA8_Class 1C	**VARIANT ANALYSIS FAILED**	
#9	HLA9_Class 1A	Homo sapiens HLA-A gene for MHC class I antigen, allele HLA-A*03:01:01:01	98
		Homo sapiens MHC class I antigen (HLA-A) gene, HLA-A*74:02V allele, complete cds	97
	HLA9_Class 1B	-	-
		-	-
	HLA9_Class 1C	-	90
		-	
#10	HLA10_Class 1A	Homo sapiens HLA-A gene for MHC class I antigen, allele HLA-A*23:01:01	98
		-	-
	HLA10_Class 1B	-	-
		-	-
	HLA10_Class 1C	Homo sapiens HLA-C gene for MHC class I antigen, allele HLA-C*06:02:01var	96
		Homo sapiens MHC class I antigen (HLA-C) gene, HLA-C*17 allele, complete cds or Homo sapiens partial HLA-C gene for MHC class I antigen, allele HLA-C*12:new	96 96

4. DISCUSSION

Logistical constraints associated with the collection and analysis of biological samples in community-based settings have been a significant hindrance to the development of biodemographic and biobehavioral research. Recent methodological developments, in particular the use of dried blood spot (DBS) samples which only require a drop of whole blood collected on filter paper, have overcome many of these constraints (Adawaye *et al*., 2013; McDade *et al*., 2007). They have also expanded the options for including biomarkers in population-based health research. Not only does this sample collection method provide a relatively non-invasive and painless method of obtaining blood, it also provides access to physiological information that would not otherwise be attainable in non-clinical settings by allowing accessibility to analytes that are accessible only through blood (McDade *et al*., 2007).

4.1 DNA extraction

The predominant aim of this work was to find an extraction method that could isolate both high yield and good quality DNA from dried blood spots that can be used for host genomic analysis. After much evaluation (summarized in Appendix A), the QIAamp DNA Mini Kit method was optimized to isolate good quality DNA from DBS. The DNA yield from the DBS using this extraction kit was very low with a mean of 1.85ng/µl as compared to the 51ng/µl (+ve) that was achieved when whole blood was used to extract the DNA using the same kit (Table 3.1). The low concentrations of DNA extracted from DBS could be explained by the small amount of blood that is used when storing the blood samples. One drop of blood on a DBS contains approximately 50 µl of whole blood and results in a DBS sample approximately 12 mm in diameter (McDade *et al*., 2007).

There are many difficulties encountered in laboratories when extracting the maximum possible material from DBS. When whole blood is extracted from subjects of research, the blood incorparates liquid and cellular fractions, in which the cellular components are removed by centrifugation to yield serum or plasma in samples extracted through processes such as venipuncture. A difficulty faced in blood storage on DBS is that the centifugation step prior storage is omitted and it has been reported that when whole blood samples are dried on filter paper, cellular components rupture and are released into solution when blood

spots are reconstituted. This leads to additional extraction procedures being required for extraction of certain analytes to overcome the problem of inhibitors (cellular components) (Lakshmy, 2008), and in turn the additional extraction procedures open a doorway for possible contamination and also decrease DNA yield.

Although low concentrations of DNA were extracted, we proceeded to determine the quality of the extracted samples by PCR.

4.2 Amplification of the HLA Class 1A, 1B and 1C loci.

In order to amplify a specific target site using any type of PCR, good quality DNA is critical for specificity and efficiency (Sairkar *et al*., 2013). Much effort was expended on optimizing the PCR (summarized in Appendix C).

Despite the low yield of DNA that was measured in the extracted samples (Table 3.1), amplification of the HLA Class 1A, 1B and 1C loci through nested PCR was successful, which in turn confirms that the extracted DNA was sufficient to amplify HLA alleles. Amplification of DNA was denoted by the presence of one distinct band (1000 bp) on the electrophoresis gel representing either HLA Class 1A, 1B or 1C. As seen in Figure 3.2, all samples extracted from DBS amplified the targeted site of 1000 bp in the HLA Class 1 locus except sample 2 in amplification of HLA Class 1A and sample 7 in the amplification of HLA Class 1C.

The PCR products were purified post PCR to remove all remaining dNTPs, primers, Taq, and Mg^{2+} ions in the PCR products. These components after amplification are considered to be contaminants as they could possibly interfere with downstream analyses such as DNA sequencing. Sample 2 in the amplification of HLA Class 1A after PCR purification did produce the expected target site of 1000 bp even though the band was not as bold as the bands produced by the other samples (Figure 3.3). The absence of the targeted site of this sample in Figure 3.2 could have been due to one of the contaminants stated above that probably remained in the PCR product after amplification and hindered the amplification of the HLA Class 1A expected site. On the other hand sample 7 in amplification of HLA Class 1C still did not produce a band on the agarose electrophoresis gel even after the sample

underwent PCR purification (Figure 3.3). The absence of the expected band size in this sample can only be due to poor quality and/or low quantity of the sample DNA.

4.3 Sequence quality control and HLA typing analysis.

One of the most valuable platforms that can be used to study biological systems is DNA sequencing; this commonly involves the use of dideoxy chain termination technology (Ronaghi, 2001). In this study Sanger sequencing was performed on the purified PCR products to sequence the HLA Class 1 loci (HLA Class 1A, 1B and 1C) from the 10 amplified samples. The sequences attained from Sanger sequencing were analysed using Geneious R8 v8.1.8 and PubMed NCBI BLAST and also underwent HLA typing analysis using the Assign^{TM} ATF Software v1.5 for sequence quality control purposes.

4.3.1 Sequence quality control

To investigate the quality of the Sanger sequenced fragments of the HLA Class 1 loci, the four sequences produced by sequencing for each sample (4 per allele) were mapped together to construct consensus sequences using the Geneious R8 v8.1.8 de novo assembly option. Samples that could not be mapped using the de novo option were mapped to a reference gene in order to obtain the consensus sequence (Table 3.3). Of all sequenced samples only sample 7 for HLA Class 1C failed to construct a consensus sequence. This is because the sequenced fragments for this HLA Class 1C sample were of poor quality (before and after trimming) as seen in Table 2.11. These results confirm that the DNA extracted from DBS was of good quality because irrespective of the low DNA concentration yields, Sanger sequencing was successful in majority (all except one) of the samples leading to the construction of consensus sequences being achievable.

When the newly constructed sequences (consensus sequences) were BLASTed using PubMed NCBI BLAST to align these new sequences against the non redundant (NR) database of NCBI for quality control purposes, all the sequences fell under the HLA Class 1 loci (Table 3.4).

4.3.2 HLA typing analysis

It is essential to exploit the potential of the quantitative information obtained from these sequencing traces in order to accurately detect the variations of HLA allele sequences. In this study variations in the HLA allele sequences were detected through comparison of the HLA Class 1 alleles to a single reference sequence using the Assign^{TM} ATF Software v1.5. This software is an automated program of high sophistication that is used for analysis of DNA sequence electropherograms to produce rapid, high throughput, quality controlled results from automated DNA sequencers.

Results from this analysis show that host genomic analysis can be performed on DNA extracted from DBS as HLA major and minor allele variations were detected in majority of the samples. In Table 3.4 it is evident that the HLA Class 1A allele had the best sequence quality as the software was able to detect variations in all 10 sequenced samples. For HLA Class 1B, variations were seen in all the samples except for sample 1 and sample 3. Auto editing is an in-built base call algorithm that is used by this program when the quality of a sequence peak is poor. Assign™ ATF v1.5 uses prior base calling information at a certain position as a guide to the most likely base. Even after editing the sequences using the software (auto editing), the variants matched to these samples (sample 1 and sample 3) were questionable. Sanger sequencing of these two samples would need to be repeated in order to obtain more accurate results. On the other hand HLA Class 1C did not only have questionable variations (sample 3 and sample 10) but the software could not assemble some of the sequences and therefore variations in sample 4, 7 and 8 could not be detected. This was due to the Sanger sequenced fragments being of poor quality as the software has a limitiation of only being able to produce results from sequences that are of good quality. The lengths of the 4 Sanger sequencing fragments for these samples (especially sample 7 and sample 8) were quite short as compared to the other samples (Table 2.11). Irrespective of the high quality percentange that each fragment had, the shorter the fragments are in length, the less specific those fragments become and, in turn, that makes it difficult to get specific results using them.

4.3.3 PubMed NCBI BLAST result comparison with the Assign™ ATF Software v1.5

The BLAST (Table 3.3) and the ATF software v1.5 classification (Table 3.4) results were compared. This was done in order to evaluate if the consensus sequences fall under the HLA Class 1 loci and further see if they match the variant results produced in the HLA typing analysis using the Assign™ ATF Software v1.5. As seen in Table 3.5, for majority of the samples the variants from the PubMed NCBI BLAST were the same as those produced in ATF Software v1.5.

Rajatileka and associates (2013) performed a similar study to ours using DBS. In their study they compared the efficiency of gDNA from three different blood sources (whole blood, umbilical cord blood and DBS). Amongst other aims they focused on the suitability of newborn DBS and umbilical cord tissue for PCR and DNA sequencing. They however, used a next generation sequencing method known as pyrosequencing (Rajatileka *et al*., 2013). Pyrosequencing is based on the principle of sequencing by synthesis, which differs from Sanger sequencing in that its detection depends on the detection of pyrophosphate release in the incorporation of the nucleotides whereas Sanger sequencing relies on dideoxynucleotide chain termination (Ronaghi, 2001).

Not surprisingly, and similar to our findings, the mean yield in the DNA extraction from DBS was lower than that of the two other blood sources. A significant difference was observed between the three groups; newborn DBS had a p-value of $p < 0.001$ as compared to $p < 0.01$ for whole blood versus umbilical cord blood. For the purpose of validating the quality of the extracted DNA, a 325 bp fragment of a house keeping gene (β-actin) was amplified. This showed a clear specific band with the expected size. All tested samples produced an amplicon at the expected size (Rajatileka *et al*., 2013). In our study, we used larger fragments of 1000 bp belonging to the HLA Class 1 loci (HLA Class 1A, 1B and 1C). In their study, all of the amplified samples were thereafter used in the detection of two unrelated SNPs by pyrosequencing. While all of the samples from whole blood or umbilical cord produced conclusive pyrograms, 6% of rs1835740 and 14% of rs4354668 were unsuccessful in the detection of conclusive pyrograms for the DBS samples (Rajatileka *et al*., 2013). Their results prove that DBS has certain limitations to be used as a source of DNA. However, they showed that despite limitations it could be used in successful down stream analyses of host genes.

In 2007, McNulty *et al*. assessed the use of DBS in HIV-1 drug resistance testing. They did so by investigating the efficiency of amplifying a 1 023 bp HIV-1 pol fragment (approximately the same size as the HLA Class 1 fragments used in our study) from viral DNA extracted from DBS collected in the field and of that collected and stored under conditions that were defined comparing it to the amplification efficiency of blood plasma. The evaluated samples collected were from two different sources. The first set was from Virology Quality Assessment program (VQA). They classified those samples into two groups (A and B), with each group consisting of a mixture of samples with a low, medium, or high virus load. Group A was stored at a temperature of - 30°C for 6 years and contained samples with resistance-associated mutations. Group B was stored at either room temperature (Group B1) or -70°C (Group B2) for 5 years and contained samples with wild-type viruses. The second group of samples incorporated matched plasma and DBS samples from HIV-1-infected donors collected at Cameroonian blood banks. These samples were prepared from blood that was HIV positive. The DBS specimens had been stored for approximately 3 years at - 20°C, which was the same temperature our DBS samples were stored. The amplification was performed using an in-house reverse transcriptase (RT) nested PCR method (McNulty *et al*., 2007).

Additional to this, the proviral DNA was also assessed examining similarities between pol sequences from paired plasma and DBS samples. Sequence analysis of HIV-1 pol was done in an ABI 3100 capillary sequencer and the Vector NTI program (suite 8) was used to analyze the data and calculate amino acid and nucleotide similarities (McNulty *et al*., 2007).

The foremost findings from their study was the amplification efficiency of the large HIV pol fragments from DBS stored for approximately 3 years at - 20°C, suggesting that this may be a suitable temperature for long-term storage of DBS. For collection sites based in less-developed areas, storing DBS at room temperature or 4°C was presented as feasible alternative for short-term storage (McNulty *et al*., 2007). A high concordance between the plasma and the DBS was seen as the mean similarity between the nucleotides of pol sequences from plasma and those of DBS ranged from 97% to 100%. Similarities between the amino acids were also witnessed to be very high, as they ranged from 96% to 100% (McNulty *et al*., 2007). The results mentioned above further confirm that DBS, when well stored, can be utilized in viral drug resistance analyses.

4.4 Conclusion and Future studies

This study shows that, in spite of the small amount of blood that is used when storing the blood samples in DBS, it is possible to sequence the HLA locus directly from DBS. DBS can also yield valuable information such as somatic mutations in tumor tissue, emerging mutations in viral genomes that confer drug resistance, or evaluating the amount of methylation possessed by a particular CpG locus, which are all vital outcomes in the progression of research. In the near future, Africa Centre may apply for ethics permission to genotype its sample repository in order to determine the effect of host HLA on HIV-1 disease progression within our African context.

5. APPENDIX

A. Protocols used in the optimisation of DNA extraction from dried blood spots

A.1 DNA extraction using the Prepfiler Forensic DNA extraction Kit from cheek epithelium and DBS

Two DNA extraction kits, QIAamp® DNA Mini Kit (QIAGEN, Germany) and PrepFiler® Forensic DNA extraction Kit (Life Technologies, California) were explored to determine which of the two would give the best quality and yield of DNA. The two highlighted sample types in Table 1 (below) are the sample types that were used for the optimization of the Nested PCR using the PrepFiler® Forensic DNA extraction Kit. The DNA was extracted following the manufacturer's instructions. The protocol was as follows:

The sample (cheek swab or DBS) was placed in a PrepFiler® Spin Tube and a mixture of 300 µl PrepFiler® Lysis Buffer and 3 µl DTT (1.0 M) was also added to the tube making sure that the fluids covered the substrate. The tube was capped and vortexed using a VORTEX GENIE 2 (Scientific Industries) for 5 seconds, then centrifuged (Eppendorf 5418 R centrifuge) briefly. The tube was thereafter placed in a thermal shaker and incubated at 70°C at a speed of 900 rpm for 40 min.

The sample was then centrifuged for 2 sec. A PrepFiler® Filter Column was inserted into a new 1.5ml PrepFiler® Spin Tube and the sample was carefully transferred into the filter column using a pipette to transfer the liquid and a pipette tip to transfer the substrate. The tube was then centrifuged at maximum speed (14 000 rpm) for 2 min. The filter column was disposed and the sample lysate was incubated at room temperature for 5 min. 15 µl of magnetic particles were pipetted into the tube containing the sample lysate successive to vortexing the PrepFiler® Magnetic Particles tube for 5 sec and inverting the tube making sure that no visible pellet remained at the bottom of the tube. After addition of the magnetic particles, the tube with the sample lysate was vortexed at a low speed (850 rpm) for 10 sec before being centrifuged briefly. After centrifuging, 180 µl of isopropanol was added to the sample lysate tube and the tube was vortexed at low speed (850 rpm) for 5 sec before being

briefly. The sample lysate tube was placed in a thermal shaker and was allowed to mix at room temperature (25°C) at a speed of 1000 rpm for 10 minutes.

The sample DNA tube was then placed in the magnetic stand until the size of the pellet of magnetic particles on the back of the tube stopped increasing in size. While the sample DNA tube remained in the magnetic stand, a pipette was used to carefully remove and discard all visible liquid phase making sure that the magnetic particle pellet was not aspirated or disturbed. After binding the DNA to the magnetic particles, the magnetic particles were washed to remove impurities and inhibitors. 600 µl Wash Buffer A was added to the DNA tube. The tube was removed from the magnetic stand and vortexed at high speed (10 000 rpm) for 5 sec and thereafter centrifuged briefly for the mixture to collect at the bottom of the tube. The tube was placed back on the magnetic stand and was left untouched for 60 sec to allow the magnetic particle pellet to grow in size on the back of the tube. With the sample DNA tube remaining in the magnetic stand again a pipette was used to carefully remove and discard all visible liquid phase. The abovementioned procedure (from the addition of the wash buffer step to the discarding of all visible liquid phase) was repeated twice, for the second wash using 300 µl Wash Buffer A and 300 µl Wash Buffer B for the third wash. After the third wash the sample DNA tube still standing in the magnetic stand was left open allowing the magnetic particles-bound DNA to air-dry for 10 min.

Thereafter an addition of 50 µl PrepFiler® Elution Buffer was added to the sample DNA tube prior to the tube being vortexed at maximum speed until no visible magnetic particle pellet could be seen. The tube was then incubated in a thermal shaker at a speed of 900 rpm for 5 min at 70°C. The sample DNA tube was again vortexed at maximum speed until there was no visible magnetic particle pellet on the side of the tube then briefly centrifuged. The tube was placed back on the magnetic stand and left undisturbed until the magnetic particle pellet on the side of the tube stopped increasing in size. All the liquid in the sample DNA tube (which contained the isolated genomic DNA) was pipetted into a 1.5-mL microcentrifuge and was stored at –20°C.

A.2 DNA extraction using the QIAamp® DNA Mini Kit from whole blood- generation of a positive control for trouble shooting

In scientific research one must have a control group that gives reliable standard data to compare the results with. A positive control receives a treatment with a known response, so that this positive response can be compared to the unknown response of the treatment. For this research DNA extracted from whole blood was used as a positive control. The blood was extracted using the QIAamp® DNA Mini Kit following the manufacturer's instructions. The protocol was as follows:

20 µl Proteinase K (QIAGEN) was pipetted into the bottom of a 1.5 ml microcentrifuge tube. An addition of 200 µl of the sample (blood) and 200 µl Buffer AL was also added to the microcentrifuge tube. The content of the tube were mixed by pulse-vortexing for 15 sec using a FINE VORTEX (FINE PCR) making sure the sample and Buffer AL were mixed thoroughly into a homogeneous solution to ensure efficient lysis. The tube was thereafter incubated in a heating block (TECHNE DRI-Block DB.2A) for 10 min at 56°C and after incubation the microcentrifuge was briefly centrifuged (Eppendorf 5415 R centrifuge) to remove drops from the inside of the lid. 200 µl of ethanol (100%) was added to the sample, and the tube was mixed by pulse-vortexing for 15 sec. The mixture was then transferred into a QIAamp Spin column placed in a 2 ml collection tube. The spin column was centrifuged at a speed of 6000 x g for 1 min and thereafter transferred into a clean collection tube. The spin column was carefully opened and 500 µl Buffer AW1 was added before the tube was centrifuged at 6000 x g for another 1 min. After centrifugation, the spin column was transferred into another clean collection tube and 500 µl Buffer AW2 was added to it. The spin column was then centrifuged at highest speed for 3 min. For the elimination of any Buffer AW2 carryover, the spin column was placed in a clean collection tube and centrifuged at the highest speed for 1 min, as Buffer AW2 may become an inhabitant in downstream applications if carried over to the eluting stage. The QIAamp Spin Column was then placed in a clean 1.5 ml microcentrifuge tube and 200 µl Buffer AE was added. The tube was incubated at room temperature for 1 min, thereafter centrifuged at 6000 x g for 1 min. The newly extracted DNA samples were then stored at -20°C.

A.3 DNA extraction from epithelial tissue using the QIAamp® DNA Mini Kit

The tissue (25 mg) of the tissue sample was cut into small pieces and placed in a 1.5 ml microcentrifuge tube. 180 µl of Buffer ATL and 20 µl Proteinase K was added to the microcentrifuge. The sample was then mixed by vortexing and was thereafter incubated at 56°C in a shaking heating block until the tissue was completely lysed. The sample was briefly centrifuged to collect all fluid at the bottom. 200 µl Buffer AL was added to the sample followed by pulse-vortexing for 15 sec assuring that the Buffer AL mixes thoroughly with the sample before incubating in a heating block at 70°C for 10 min. Proceeding this step, 200 µl ethanol (100%) was added to the sample and pulse-vortexed to mix for 15 sec.

The sample with its precipitate was thereafter transferred into a QIAamp Spin Column placed in a 2 ml collection tube and was centrifuged at 6000 x g for 1 min. The QIAamp Spin Column was placed in a clean 2 ml collection tube and the tube containing the filtrate was discarded. 500 µl Buffer AW1 was added to the QIAamp Spin Column and the spin column was centrifuged at 6000 x g for 1 min. The QIAamp Spin Column was again placed in a clean 2 ml collection tube and was topped with 500 µl Buffer AW2 and centrifuged at full speed for 3 min. To avoid carryover of buffer AW2 into the eluting step, the QIAamp Spin Column was placed in a new 2 ml collection tube and was centrifuged for 1 min at full speed. Thereafter the QIAamp Spin Column was placed in a clean 1.5 ml microcentrifuge tube and 200 µl Buffer AE was added to the microcentrifuge. The tube was incubated at room temperature for 1 min then centrifuged at 6000 x g for 1 min. The eluting step was repeated before the DNA was stored as 20 µl aliquots at -20°C.

A.4. Sequence Assembly and Sequence Quality control

A.4.1 Steps followed to construct consensus sequences using Geneious R8 v8.1.8

Geneious R8 v8.1.8 was used to map the four sequences for Sanger sequencing into a consensus sequence using the De novo Assembly option. The step carried out in constructing the consensus sequences were as follows:

a) Launch the program **Geneious.**

b) Import the **ABI files** generated by the sequencing machine to a working folder using the import tool

c) Extract sequence using the cursor by selecting the mid-section of the sequence leaving out the ends, which are usually of poor quality.

d) Click on the extract button to extract the region with good quality sequence

e) Select all four extracted sequences for each sample

f) Click on **Align/Assemble**

g) Click on **De novo Assemble**

h) Select the assembled sequence

i) Right click on it and select **Generate consensus**

j) Copy the base pairs of the consensus sequence

k) Open NCBI and **BLAST** the sequence

Some extracted sequences had gaps and therefore a consensus sequence could not be achieved. Those sequences were mapped together by doing the following**:**

a) Find your closest reference and copy it

b) Go to **FASTA** on **NCBI** and retrieve the nucleotides of the closest reference

c) Go to Sequence at the top of **Geneious**

d) Select **New sequence**

e) Paste sequence from **FASTA** and save as **REFERENCE SEQUENCE**

f) Select all 4 extracted sequences **PLUS** the **REFERENCE SEQUENCE**

g) Click on **Align/Assemble**

h) Select **Map to Reference**

i) Proceed with the generating consensus sequence as stated on the procedure above

A.4.2 Sequence quality control using NCBI BLAST

The PubMed NCBI BLAST (Basic Local Alignment Search Tool) was used to align these new sequences against the non redundant (NR) database of NCBI as a sequence quality control. This was done in order to evaluate if the consensus sequences fall under the HLA Class 1 loci. The specificity of the consensus sequences was evaluated by blasting them in the NCBI database (www.ncbi.nlm.nih.gov) using the standard nucleotide BLAST.

B. General Recipes

i. 70% Ethanol

Reagent	50 ml volume (ml)
100% Ethanol	35
Distilled water	15

- Distilled water was combined with 100% ethanol to yield a 70% volume.

ii. 1% Agarose Gel with ethidium bromide

Reagent	100 ml volume
Agarose	1g
1x TBE running buffer	100 ml
0.05 µl/ml Ethidium bromide	30 µl

- The agarose powder was melted in the 1x TBE running buffer and allowed to cool before the ethidium bromide was added.

To make Ethidium bromide stock solution:

Reagent	1 ml volume (ml)
Ethidium bromide	10 mg
Di.H_20	1ml

- Ethidium bromide was dissolved in di.H_20 to make a stock solution of 10mg/ml

Ethidium bromide working solution:

The ethidium bromide stock solution was diluted 1:200 in di.H_20 to create a 0.05mg/ml working solution

Ethidium bromide solution:

Reagent	100 ml volume (ml)
Ethidium bromide	500 µl
1x TBE buffer	100ml

- The ethidium bromide stock solution was diluted 1:200 in 1xTBE buffer to create a 0.05 mg/ml working solution.

iii. 1x TBE running buffer

To make up 10x TBE running buffer:

Reagent	**400 ml volume**
Tris-base	43.2g
Boric acid (Merck)	22 g
EDTA	3.72 g
Distilled water	320 ml

- Tris-base, boric acid and EDTA were all dissolved in 320 ml of distilled water.

To make up 1x TBE running buffer:

Reagent	**1L volume**
10x TBE running buffer	100 ml
Distilled water	900 ml

- 10x TBE buffer was diluted with distilled water.

iv. 1% agarose electrophoresis without ethidium bromide

Reagent	**80 ml volume**
Agarose DNase/Rnase free tablet	2 tablets
1x TBE running buffer	80 ml

- Two Agarose DNase/Rnase free tablets (BIOLINE) were melted into 80 ml of TBE buffer to make the 1% agarose gel.

C. PCR Optimization

Experiments were performed in two different places, the Molecular Biology laboratory in the School of Life Sciences, UKZN Westville Campus and the Africa Centre for Health and Population Studies Genomics Lab at UKZN Medical School Campus.

C.1 PCR optimization performed at the School of Life Sciences Molecular Biology Lab

C.1.1 DNA quantification

Table 3: DNA yield (ng /µL) in samples extracted from Dried Blood spots and cheek epithelium using the QIAamp DNA Mini Kit and the PrepFiler® Forensic DNA extraction Kit.

Sample no.	Original Name	DNA (ng/uL)	260/280	Kit used
1	1	4,5	2,69	QIAamp DNA Mini Kit
2	2	8,4	1,45	QIAamp DNA Mini Kit
3	3	6,2	1,04	Forensic Kit (PrepFiler)
4	4	3,6	1,66	Forensic Kit (PrepFiler)
5	N(1)	2,5	2,08	QIAamp DNA Mini Kit
6	10	11,2	1,48	QIAamp DNA Mini Kit
7	11	5,5	1,72	QIAamp DNA Mini Kit
8	12	3,2	1,86	QIAamp DNA Mini Kit
9	13	3,3	1,78	QIAamp DNA Mini Kit
10	14 (6)	4	1,39	QIAamp DNA Mini Kit
11	15 (6)	2,6	1,52	QIAamp DNA Mini Kit
12	16 (6)	1,5	19,05	QIAamp DNA Mini Kit
13	17	26,5	1,9	QIAamp DNA Mini Kit
14	18	5,9	1,3	QIAamp DNA Mini Ki
15	9	28,1	1,97	QIAamp DNA Mini Kit

16	21(2)	10,1	2,5	QIAamp DNA Mini Kit
17	Prep (PS)	26,9	2,69	Forensic Kit (PrepFiler)
18	Fresh (QIAamp)	6,9	1,5	QIAamp DNA Mini Kit

Table 4: DNA yield (ng /µL) in samples extracted from cheek epithelium cells using the QIAamp DNA Mini Kit.

Sample no.	Sample Donor	DNA (ng/uL)	260/280	Kit used
2	Black-SA (Donor 1)	10,4	1,9	QIAamp DNA Mini Kit
3	Black-ZM (Donor 2)	6,2	1,04	QIAamp DNA Mini Kit
4	Coloured (Donor 3)	12,4	1,66	QIAamp DNA Mini Kit
5	Indian (Donor 4)	21,5	2,08	QIAamp DNA Mini Kit

The tabulated DNA quantities in Table 3 and 4 were quantified through spectrophotometry using the **Nanodrop Spectrophotometer (ND-1000)** at 260/280 absorbance ratios. The extracted samples (Table 3) were from dried blood spots made using my own blood **(15th of July 2014)** except the two that are highlighted, which were received from Africa Centre Genomics Lab **(made on the 7th of May 2010).** The cheek epithelium samples (Table 4) were from anonymous donors from the School of Life Sciences given on the **6th of May 2015.** All gels done at the School of Life Sciences were made following the method stipulated under Appendix B (ii).

C1.2 First attempt at amplification of HLA using a protocol received from the Africa Centre

DNA Extraction			
Name	Phiwa Mjoli	Sample	DBS
Date	11-08-14	Extraction Method	QIAGEN Mini Kit

Nested PCR			
Name	Phiwokuhle Mjoli	Name of Kit	KappaRobust
Date	09-12-14	Kit Lot #	
Study	HLA Typing	Expiry Date	08/2013
No. of Samples	7	Run #	Optimization

Reagents	**Volume / reaction (µl)**
Water	6.3
10x Buffer (with $MgCl_2$)	4.0
Enhancer	1.0
$MgCl_2$ (25mM)	0.4
dNTP	0.4
Primer 1 (10µM)	0.4
Primer 2 (10µM)	0.4
KappaRobust Taq	0.1
Sample	7.0
Total reaction volume	20.0

Thermocycling Conditions		
Temperature (°C)	**Time**	**Cycle**
94	4 min	Hold
94	15 sec	X 40
55	30 sec	
72	30 sec	
72	7 min	Hold
4	∞	Hold

(Used the same cycles for Step 1 and 2)

The following table is a summary of the results seen on the electrophoresis gel that follows (Figure 1), showing the success/failure of the attempted HLA Class 1B and HLA Class 1C amplification.

No.	Class 1B	Class 1C
1	✗	✗
2	✓	✗
3	✓	✗
4	✗	✗
5	✗	✗
6	✗	✗
7	✗	✗

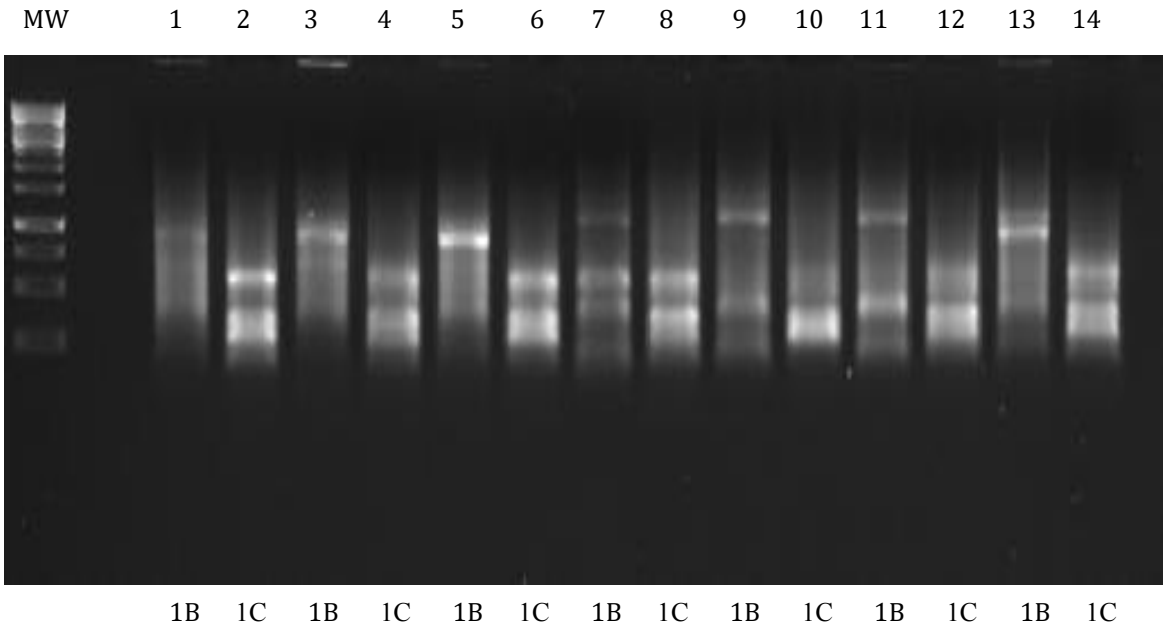


Figure 1: Image of an electrophoresis gel showing amplification of samples 1-7 using 7 µl sample volume after second round of PCR using primers for HLA class 1B and 1C alleles.

In Figure 1 it can be seen that there is non-specific amplification (PCR doublets) for both HLA Class 1B and 1C. The non-specific amplification could be the consequence of using too much DNA template in the PCR. The input sample volume (for both step 1 and step 2) was decreased to 4µl to see if the smaller volume would eliminate the non-specific amplification.

C1.3 Attempt of HLA amplification using 4 µl of sample volume

DNA Extraction			
Name	Phiwa Mjoli	Sample	DBS
Date	11-08-14	Extraction Method	QIAGEN Mini Kit

Nested PCR			
Name	Phiwokuhle Mjoli	Name of Kit	KappaRobust
Date	05-01-15	Kit Lot #	
Study	HLA Typing	Expiry Date	08/2013
No. of Samples	7	Run #	Optimization

Reagents	**Volume / reaction (µl)**
Water	9.3
10x Buffer (with $MgCl_2$)	4.0
Enhancer	1.0
$MgCl_2$ (25mM)	0.4
dNTP	0.4
Primer 1 (10µM)	0.4
Primer 2 (10µM)	0.4
KappaRobust Taq	0.1
Sample	4.0
Total reaction volume	20.0

Thermocycling Conditions		
Temperature (°C)	**Time**	**Cycle**
94	4 min	Hold
94	15 sec	X 40
55	30 sec	
72	30 sec	
72	7 min	Hold
4	∞	Hold

(Used the same cycles for Step 1 and 2)

No.	**Class 1B**	**Class 1C**
1	✗	✗
2	✓	✗
3	✗	✗
4	✓	✗
5	✓	✗
6	✓	✗
7	✓	✗

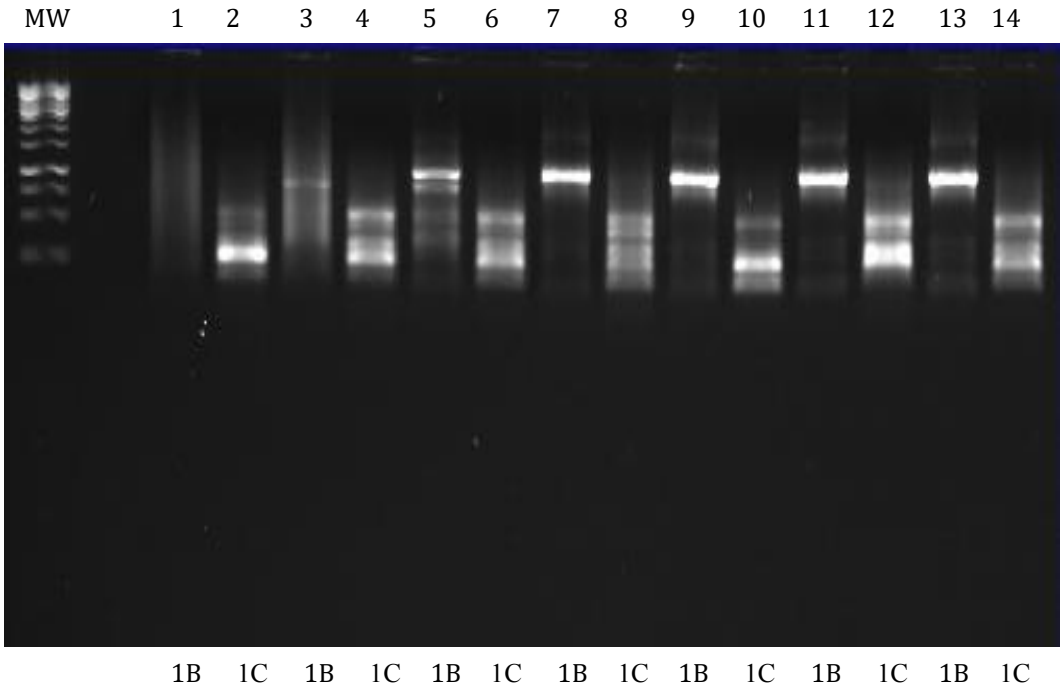


Figure 2: Image of an electrophoresis gel showing amplification of samples 1-7 using 4 µl sample volume using primers for HLA Class 1B and 1C alleles.

When 4 µl of the sample volume was used in the PCR bands of the expected size were seen in samples 3 to 7 for HLA Class 1B. The non-specific amplification persisted for in HLA Class 1C and also in samples 1 and 2 for HLA Class 1B. Since HLA Class 1C had consistently failed and some of the HLA Class 1B samples failed, we increased the input volume to 10 µl to see if this increase would assist amplification of HLA Class 1C.

C1.4 Attempt of HLA amplification using 10 µl of sample volume

DNA Extraction			
Name	Phiwa Mjoli	Sample	DBS
Date	11-08-14	Extraction Method	QIAGEN Mini Kit

Nested PCR			
Name	Phiwokuhle Mjoli	Name of Kit	KappaRobust
Date	03-02-15	Kit Lot #	
Study	HLA Typing	Expiry Date	08/2013
No. of Samples	7	Run #	Optimization

Reagents	**Volume / reaction (µl)**
Water	3.3
10x Buffer (with $MgCl_2$)	4.0
Enhancer	1.0
$MgCl_2$ (25mM)	0.4
dNTP	0.4
Primer 1 (10µM)	0.4
Primer 2 (10µM)	0.4
KappaRobust Taq	0.1
Sample	10.0
Total reaction volume	20.0

Thermocycling Conditions		
Temperature (°C)	**Time**	**Cycle**
94	4 min	Hold
94	15 sec	X 40
55	30 sec	
72	30 sec	
72	7 min	Hold
4	∞	Hold

(Used the same cycles for Step 1 and 2)

No.	**Class 1B**	**Class 1C**
1	✗	✗
2	✗	✗
3	✗	✗
4	✗	✗
5	✗	✗
6	✗	✗
7	✗	✗

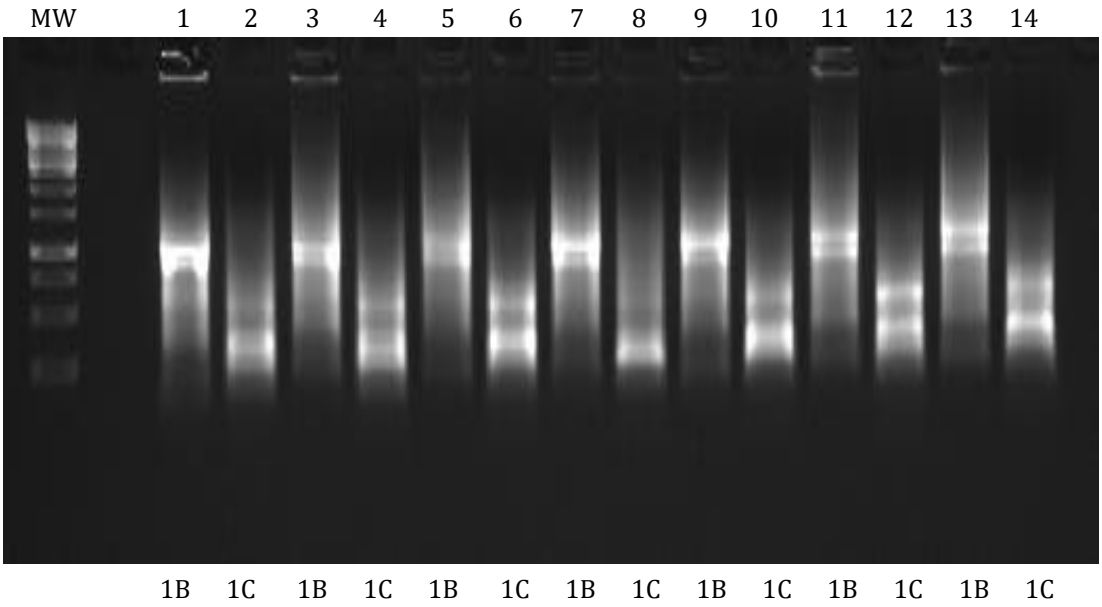


Figure 3: Image of an electrophoresis gel showing amplification of samples 1-7 using 10 µl input sample volume and primers specific for HLA class 1B and 1C alleles.

Using the volumes and concentrations of the reagents provided by Africa Centre, non-specific amplification persisted. A single specific band for HLA Class 1B was achieved when the sample volume was decreased from 7 µl to 4 µl, keeping all the concentrations and volumes of the other reagents the same (Figure 2). This however did not work on all the 14 samples amplified (amplification of sample 8 to 14 not shown). The PCR conditions that resulted in a specific band for HLA Class 1B resulted in non-specific amplification for HLA Class 1C when the sample volume was raised to 10 µl (Figure 3) suggesting that the input DNA concentration is not the sole reason for PCR failure.

We then attempted to optimize PCR conditions. Our first attempt was to manipulate $MgCl_2$ concentration. $MgCl_2$ in PCR is a co-factor for Taq enzyme. It also helps in addition of correcct DNTPs complementary to the sequence in newly synthesizing strand by binding to the dNTPs. In short, if the $MgCl_2$ concentration is too low, primers fail to anneal to the target DNA and if the concentration is too high, the base pairing becomes too strong and fails to denature when heated. It is therefore imperative to make sure that a balanced $MgCl_2$ concentration is maintained in the reaction mixture for effective amplification.

C1.5 Attempt of HLA amplification using a concentration of 0.2 and 0.8 $MgCl_2$

DNA Extraction			
Name	Phiwa Mjoli	Sample	DBS
Date	11-08-14	Extraction Method	QIAGEN Mini Kit

Nested PCR			
Name	Phiwokuhle Mjoli	Name of Kit	KappaRobust
Date	15-02-15	Kit Lot #	
Study	HLA Typing	Expiry Date	08/2013
No. of Samples	1	Run #	Optimization

Reagents	**Volume / reaction (µl)**
Water	6.1
10x Buffer (with $MgCl_2$)	4.0
Enhancer	1.0
$MgCl_2$ (25mM)	0.2
dNTP	0.4
Primer 1 (10µM)	0.4
Primer 2 (10µM)	0.4
KappaRobust Taq	0.1
Sample	7.0
Total reaction volume	20.0

Thermocycling Conditions		
Temperature (°C)	**Time**	**Cycle**
94	4 min	Hold
94	15 sec	X 40
55	30 sec	
72	30 sec	
72	7 min	Hold
4	∞	Hold

(Used the same cycles for Step 1 and 2)

No.	**Class 1B**	**No.**	**Class 1C**
6 (1)	✗	6 (4)	✗
6 (2)	✗	6 (5)	✗
6 (3)	✓	6 (6)	✗

Note: Numbers in the brackets correspond with the labels on the gel wells

DNA Extraction			
Name	Phiwa Mjoli	Sample	DBS
Date	11-08-14	Extraction Method	QIAGEN Mini Kit

Nested PCR			
Name	Phiwokuhle Mjoli	Name of Kit	KappaRobust
Date	15-02-15	Kit Lot #	
Study	HLA Typing	Expiry Date	08/2013
No. of Samples	1	Run #	Optimization

Reagents	**Volume / reaction (µl)**
Water	5.9
10x Buffer (with $MgCl_2$)	4.0
Enhancer	1.0
$MgCl_2$ (25mM)	0.8
dNTP	0.4
Primer 1 (10µM)	0.4
Primer 2 (10µM)	0.4
KappaRobust Taq	0.1
Sample	7.0
Total reaction volume	20.0

Thermocycling Conditions		
Temperature (°C)	**Time**	**Cycle**
94	4 min	Hold
94	15 sec	X 40
55	30 sec	
72	30 sec	
72	7 min	Hold
4	∞	Hold

(Used the same cycles for Step 1 and 2)

No.	**Class 1B**	**No.**	**Class 1C**
6 (7)	✗	6 (10)	✗
6 (8)	✗	6 (11)	✗
6 (9)	✗	6 (12)	✗

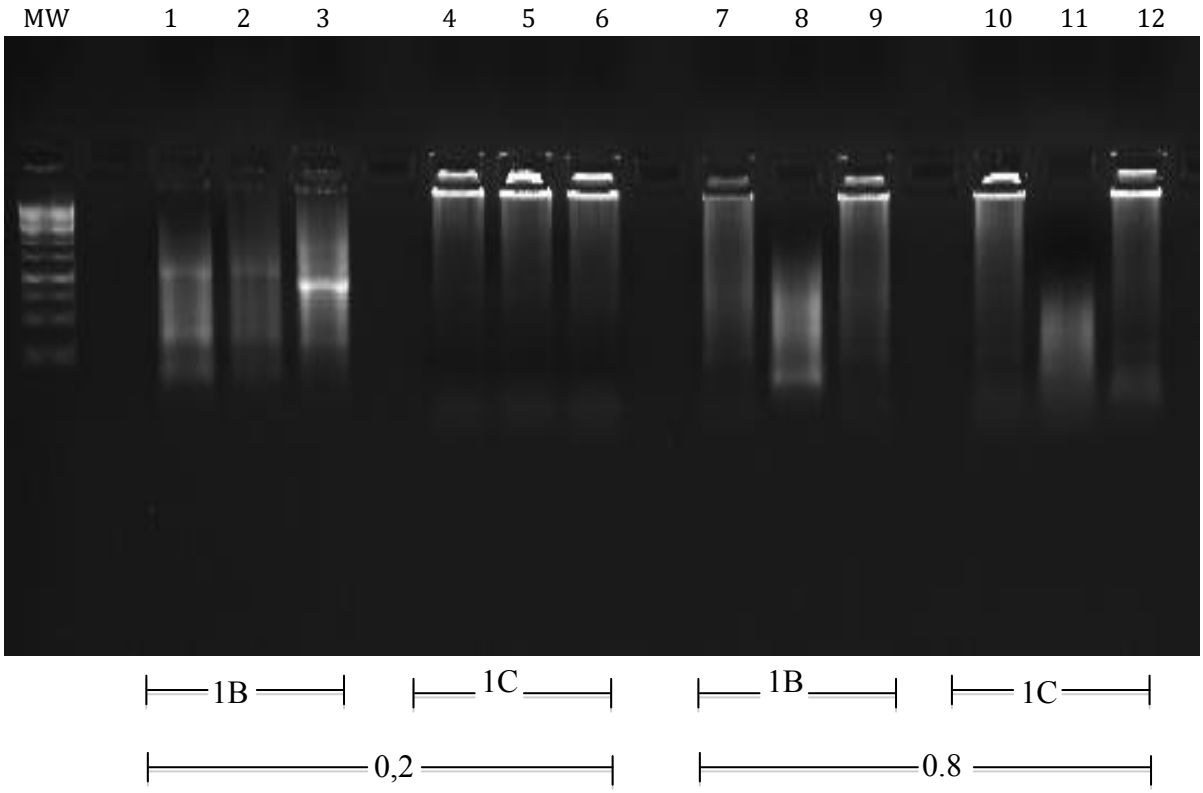


Figure 4: Image of an electrophoresis gel showing amplification of sample 6 using a concentrations of 0,2 and 0,8 MgCl2 (25 mM) and 7 µl sample input volume using primers specific for HLA class 1B and 1C alleles.

The original concentration suggested for $MgCl_2$ was 0,4 per 20 µl reaction. We attempted to optimize using lower and higher $MgCl_2$ concentration. At a concentration of 0.2 there was non-specific of HLA Class 1B while HLA Class 1C did not amplify at all. Neither amplified at a concentration of 0.8 (Figure 4).

We then decided to experiment with different annealing temperatures.

C1.6 Attempt of HLA amplification using a range of annealing temperatures (55° to 65 °C) and different sample volumes with primers specific for HLA Class 1B

DNA Extraction			
Name	Phiwa Mjoli	Sample	DBS
Date	11-08-14	Extraction Method	QIAGEN Mini Kit

Nested PCR			
Name	Phiwokuhle Mjoli	Name of Kit	KappaRobust
Date	20-02-15	Kit Lot #	
Study	HLA Typing	Expiry Date	08/2013
No. of Samples	1	Run #	Optimization

Reagents	**Volume / reaction (µl)**
Water	9.3
10x Buffer (with $MgCl_2$)	4.0
Enhancer	1.0
$MgCl_2$ (25mM)	0.4
dNTP	0.4
Primer 1 (10µM)	0.4
Primer 2 (10µM)	0.4
KappaRobust Taq	0.1
Sample	4.0
Total reaction volume	20.0

Thermocycling Conditions		
Temperature (°C)	**Time**	**Cycle**
94	4 min	Hold
94	15 sec	X 40
55-65	30 sec	
72	30 sec	
72	7 min	Hold
4	∞	Hold

(Used the same cycles for Step 1 and 2)

No.	**Temperature**	**Class 1B**
6 (1)	55	✓
6 (4)	57	✗
6 (7)	59	✗
6 (10)	61	✗
6 (13)	63	✗
6 (16)	65	✗

DNA Extraction			
Name	Phiwa Mjoli	Sample	DBS
Date	11-08-14	Extraction Method	QIAGEN Mini Kit

Nested PCR			
Name	Phiwokuhle Mjoli	Name of Kit	KappaRobust
Date	20-02-15	Kit Lot #	
Study	HLA Typing	Expiry Date	08/2013
No. of Samples	1	Run #	Optimization

Reagents	**Volume / reaction (µl)**
Water	6.3
10x Buffer (with $MgCl_2$)	4.0
Enhancer	1.0
$MgCl_2$ (25mM)	0.4
dNTP	0.4
Primer 1 (10µM)	0.4
Primer 2 (10µM)	0.4
KappaRobust Taq	0.1
Sample	7.0
Total reaction volume	20.0

Thermocycling Conditions		
Temperature (°C)	**Time**	**Cycle**
94	4 min	Hold
94	15 sec	X 40
55-65	30 sec	
72	30 sec	
72	7 min	Hold
4	∞	Hold

(Used the same cycles for Step 1 and 2)

No.	**Temperature**	**Class 1B**
6 (2)	55	✓
6 (5)	57	✗
6 (8)	59	✗
6 (11)	61	✗
6 (14)	63	✗
6 (17)	65	✗

DNA Extraction			
Name	Phiwa Mjoli	Sample	DBS
Date	11-08-14	Extraction Method	QIAGEN Mini Kit

Nested PCR			
Name	Phiwokuhle Mjoli	Name of Kit	KappaRobust
Date	20-02-15	Kit Lot #	
Study	HLA Typing	Expiry Date	08/2013
No. of Samples	1	Run #	Optimization

Reagents	**Volume / reaction (µl)**
Water	3.3
10x Buffer (with $MgCl_2$)	4.0
Enhancer	1.0
$MgCl_2$ (25mM)	0.4
dNTP	0.4
Primer 1 (10µM)	0.4
Primer 2 (10µM)	0.4
KappaRobust Taq	0.1
Sample	10.0
Total reaction volume	20.0

Thermocycling Conditions		
Temperature (°C)	**Time**	**Cycle**
94	4 min	Hold
94	15 sec	X 40
55-65	30 sec	
72	30 sec	
72	7 min	Hold
4	∞	Hold

(Used the same cycles for Step 1 and 2)

No.	**Temperature**	**Class 1B**
6 (3)	55	✗
6 (6)	57	✗
6 (9)	59	✗
6 (12)	61	✗
6 (15)	63	✗
6 (18)	65	✗

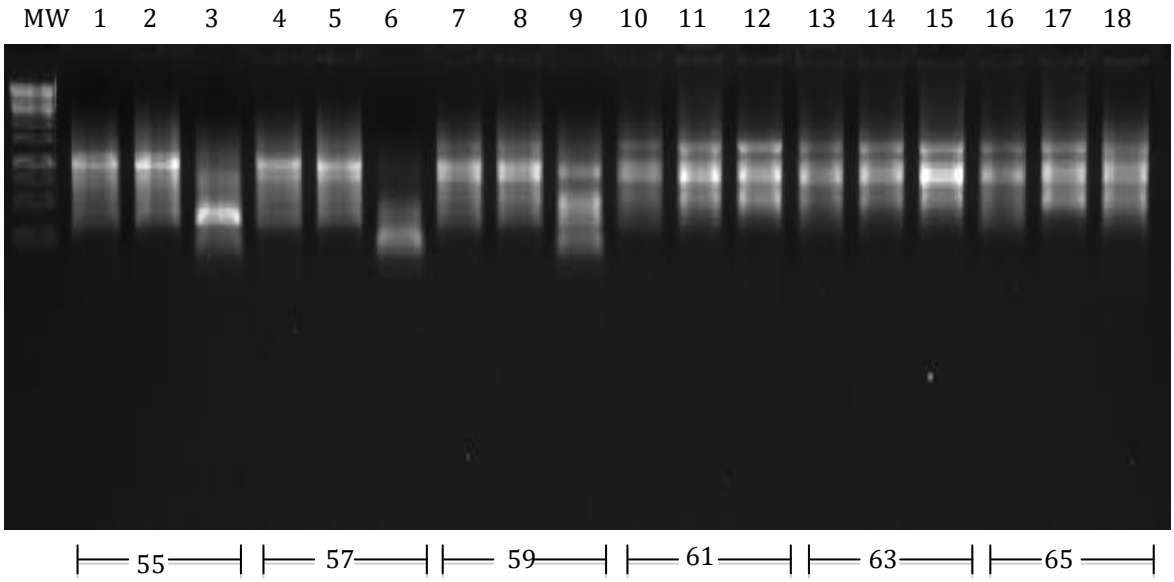


Figure 5: Image of an electrophoresis gel showing amplification of sample 6 with annealing temperatures ranging from 55° to 65 °C using 4, 7 and 10 µl sample volumes and primers specific for the HLA class 1B allele. (Well 1, 4, 7, 10, 13 and 16 = 4 µl; well 2, 5, 8, 11, 14 and 17 = 7 µl; well 3, 6, 9, 12, 15 and 18 = 10 µl)

PCR was performed under different annealing temperatures ranging from 55°C to 65°C for both alleles (HLA Class 1B and 1C). In conjunction with this, the sample volume was also varied to see which sample volume would work better with which temperature.

For HLA Class 1B allele, a single specific band could be seen using both 4 µl and 7 µl input when the annealing temperature was set at 55°C and 57°C. As the temperature increased, non-specific amplification was evident. As previously seen, non-specific amplification was observed when using 10 µl input sample volume (Figure 5).

C1.7 Attempt of HLA amplification using a range of annealing temperatures (55° to 65 °C) and different sample volumes to amplify the HLA Class 1C allele

DNA Extraction			
Name	Phiwa Mjoli	Sample	DBS
Date	11-08-14	Extraction Method	QIAGEN Mini Kit

Nested PCR			
Name	Phiwokuhle Mjoli	Name of Kit	KappaRobust
Date	09-02-15	Kit Lot #	
Study	HLA Typing	Expiry Date	08/2013
No. of Samples	1	Run #	Optimization

Reagents	**Volume / reaction (µl)**
Water	9.3
10x Buffer (with $MgCl_2$)	4.0
Enhancer	1.0
$MgCl_2$ (25mM)	0.4
dNTP	0.4
Primer 1 (10µM)	0.4
Primer 2 (10µM)	0.4
KappaRobust Taq	0.1
Sample	4.0
Total reaction volume	20.0

Thermocycling Conditions		
Temperature (°C)	**Time**	**Cycle**
94	4 min	Hold
94	15 sec	X 40
55-65	30 sec	
72	30 sec	
72	7 min	Hold
4	∞	Hold

(Used the same cycles for Step 1 and 2)

No.	**Temperature**	**Class 1C**
8 (1)	55	✗
8 (4)	57	✗
8 (7)	59	✗
8 (10)	61	✗
8 (13)	63	✗
8 (16)	65	✗

DNA Extraction			
Name	Phiwa Mjoli	Sample	DBS
Date	11-08-14	Extraction Method	QIAGEN Mini Kit

Nested PCR			
Name	Phiwokuhle Mjoli	Name of Kit	KappaRobust
Date	09-02-15	Kit Lot #	
Study	HLA Typing	Expiry Date	08/2013
No. of Samples	1	Run #	Optimization

Reagents	**Volume / reaction (µl)**
Water	6.3
10x Buffer (with $MgCl_2$)	4.0
Enhancer	1.0
$MgCl_2$ (25mM)	0.4
dNTP	0.4
Primer 1 (10µM)	0.4
Primer 2 (10µM)	0.4
KappaRobust Taq	0.1
Sample	7.0
Total reaction volume	20.0

Thermocycling Conditions		
Temperature (°C)	**Time**	**Cycle**
94	4 min	Hold
94	15 sec	X 40
55-65	30 sec	
72	30 sec	
72	7 min	Hold
4	∞	Hold

(Used the same cycles for Step 1 and 2)

No.	**Temperature**	**Class 1C**
8 (2)	55	✗
8 (5)	57	✗
8 (8)	59	✓
8 (11)	61	✗
8 (14)	63	✗
8 (17)	65	✗

DNA Extraction			
Name	Phiwa Mjoli	Sample	DBS
Date	11-08-14	Extraction Method	QIAGEN Mini Kit

Nested PCR			
Name	Phiwokuhle Mjoli	Name of Kit	KappaRobust
Date	09-02-15	Kit Lot #	
Study	HLA Typing	Expiry Date	08/2013
No. of Samples	1	Run #	Optimization

Reagents	**Volume / reaction (µl)**
Water	3.3
10x Buffer (with $MgCl_2$)	4.0
Enhancer	1.0
$MgCl_2$ (25mM)	0.4
dNTP	0.4
Primer 1 (10µM)	0.4
Primer 2 (10µM)	0.4
KappaRobust Taq	0.1
Sample	10.0
Total reaction volume	20.0

Thermocycling Conditions		
Temperature (°C)	**Time**	**Cycle**
94	4 min	Hold
94	15 sec	X 40
55-65	30 sec	
72	30 sec	
72	7 min	Hold
4	∞	Hold

(Used the same cycles for Step 1 and 2)

No.	**Temperature**	**Class 1C**
8 (3)	55	✗
8 (6)	57	✓
8 (9)	59	✗
8 (12)	61	✗
8 (15)	63	✗
8 (18)	65	✗

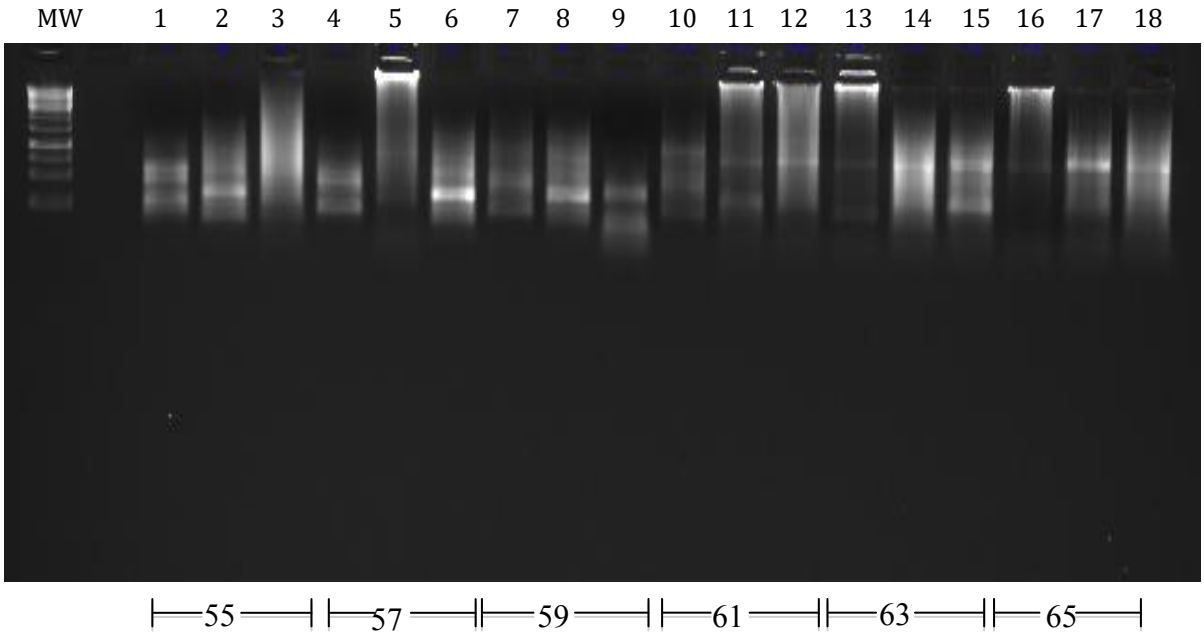


Figure 6: Image of an electrophoresis gel showing amplification of sample 8 with annealing temperature ranging from 55° to 65 °C using 4, 7 and 10 µl sample volume and primers specific for HLA class 1C allele. (Well 1, 4, 7, 10, 13 and 16 = 4 µl; well 2, 5, 8, 11, 14 and 17 = 7 µl; well 3, 6, 9, 12, 15 and 18 = 10 µl)

Non-specific amplification was seen at all annealing temperatures and volumes except 57°C when 10 µl of the sample was used for amplification of the HLA Class 1C allele (Figure 6).

We then decided to try a different Taq.

C1.8 Attempt of HLA amplification using a different PCR kit (Ex Takara)

DNA Extraction			
Name	Phiwa Mjoli	Sample	DBS
Date	11-08-14	Extraction Method	QIAGEN Mini Kit

Nested PCR			
Name	Phiwokuhle Mjoli	Name of Kit	Ex Takara
Date	24-03-15	Kit Lot #	
Study	HLA Typing	Expiry Date	
No. of Samples	2	Run #	Optimization

Reagents	**Volume / reaction (µl)**
Water	17.25
10x Buffer	2.5
dNTP	2.0
Forward Primer	1.0
Reverse Primer	1.0
Taq	0.25
Sample	1.0
Total reaction volume	25.0

Thermocycling Conditions		
Temperature (°C)	**Time**	**Cycle**
94	4 min	Hold
94	15 sec	X 40
55	30 sec	
72	30 sec	
72	7 min	Hold
4	∞	Hold

(Used the same cycles for Step 1 and 2)

No.	**Class 1C**
7 (1)	✗
7 (2)	✗
9 (4)	✗
9 (5)	✗

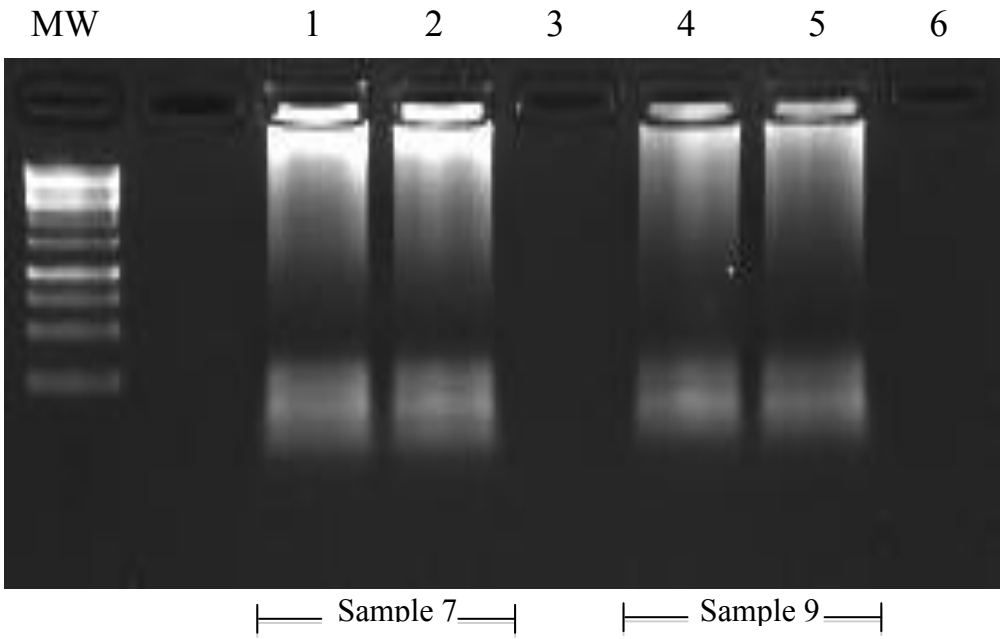


Figure 7: Image of an electrophoresis gel showing amplification of sample 7 and 9 using 7 µl sample input volume for HLA Class 1C allele. (Well 1 and 2 = sample 7; well 4 and 5 = sample 9)

An attempt of amplifying HLA Class 1C allele using a different PCR kit (Ex Takara) was made to see if the problem was with the samples that had been extracted and stored or the reagents of the PCR kit being used. Two samples were used; one that had previously given a single band for HLA Class 1B (sample 7) and one that gave doublets (sample 9). As seen in Figure 7 (above) amplification was not a success for both samples.

We were unsure whether the problem lay with the extraction kit being used to extract the DNA sample. We therefore decided to perform these experiments with freshly extracted DNA. We extracted DNA from stored DBS using a specifically designed forensic kit.

C1.9 Attempt of HLA amplification using a fresh DBS sample extracted with the PrepFiler Forensic DNA extraction Kit

DNA Extraction			
Name	Phiwa Mjoli	Sample	DBS
Date	06-04-15	Extraction Method	PrepFiler Forensic Kit

Nested PCR			
Name	Phiwokuhle Mjoli	Name of Kit	KappaRobust
Date	06-04-15	Kit Lot #	
Study	HLA Typing	Expiry Date	08/2013
No. of Samples	1	Run #	Optimization

Reagents	**Volume / reaction (µl)**
Water	12.3
10x Buffer (with $MgCl_2$)	4.0
Enhancer	1.0
$MgCl_2$ (25mM)	0.4
dNTP	0.4
Primer 1 (10µM)	0.4
Primer 2 (10µM)	0.4
KappaRobust Taq	0.1
Sample	1.0
Total reaction volume	20.0

Thermocycling Conditions		
Temperature (°C)	**Time**	**Cycle**
94	4 min	Hold
94	15 sec	X 40
55	30 sec	
72	30 sec	
72	7 min	Hold
4	∞	Hold

(Used the same cycles for Step 1 and 2)

No.	**Class 1C**
15 (1)	✗
15 (2)	✗

DNA Extraction			
Name	Phiwa Mjoli	Sample	DBS
Date	06-04-15	Extraction Method	PrepFiler Forensic Kit

Nested PCR			
Name	Phiwokuhle Mjoli	Name of Kit	KappaRobust
Date	06-04-15	Kit Lot #	
Study	HLA Typing	Expiry Date	08/2013
No. of Samples	1	Run #	Optimization

Reagents	**Volume / reaction (µl)**
Water	11.3
10x Buffer (with $MgCl_2$)	4.0
Enhancer	1.0
$MgCl_2$ (25mM)	0.4
dNTP	0.4
Primer 1 (10µM)	0.4
Primer 2 (10µM)	0.4
KappaRobust Taq	0.1
Sample	2.0
Total reaction volume	20.0

Thermocycling Conditions		
Temperature (°C)	**Time**	**Cycle**
94	4 min	Hold
94	15 sec	X 40
55	30 sec	
72	30 sec	
72	7 min	Hold
4	∞	Hold

(Used the same cycles for Step 1 and 2)

No.	**Class 1C**
15 (3)	✗
15 (4)	✗

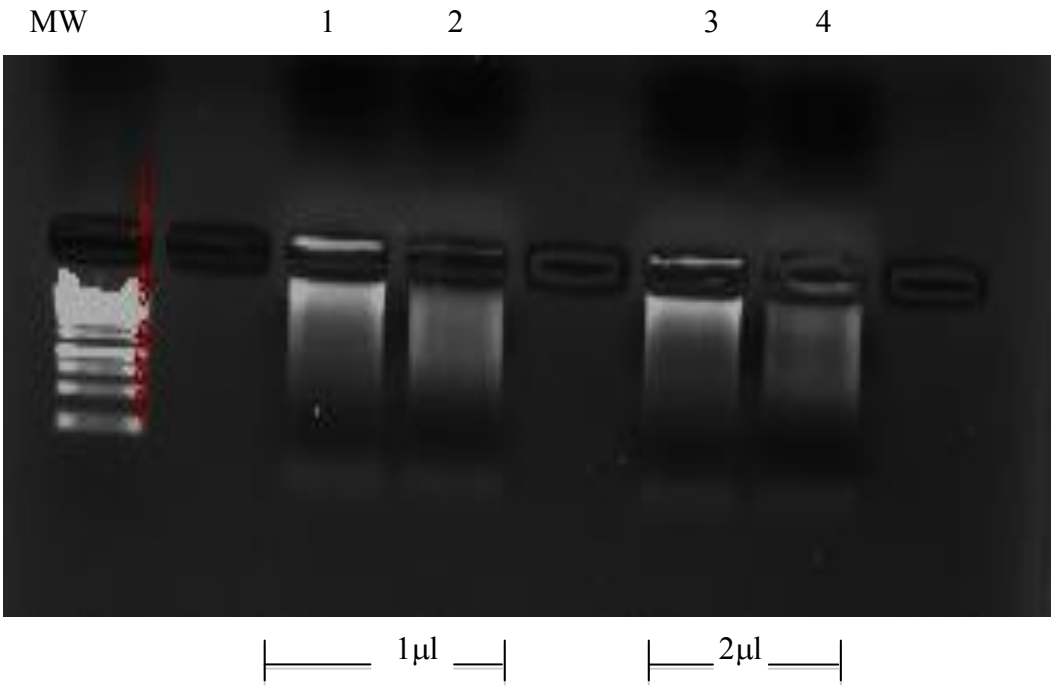


Figure 8: Image of an electrophoresis gel showing amplification of freshly extracted DNA (sample 15) extracted using the PrepFiler Forensic DNA extraction Kit extracted from stored DBS; sample volume at 1 µl and 2 µl using primers specific for HLA class 1C allele.

A PrepFiler Forensic DNA extraction kit was used to extract a fresh sample of DNA from the stored dried blood spots performing PCR on the extracted sample the same day. No amplification was achieved as seen in Figure 8.

It seemed, therefore, that we needed to optimize the PCR reaction. To this end, we tried using different buffers.

C1.10 Attempt of HLA amplification using two different buffers (GC and 5x buffer) at different annealing temperatures

DNA Extraction			
Name	Phiwa Mjoli	Sample	DBS
Date	11-08-14	Extraction Method	QIAGEN Mini Kit

Nested PCR			
Name	Phiwokuhle Mjoli	Name of Kit	KappaRobust
Date	09-04-15	Kit Lot #	
Study	HLA Typing	Expiry Date	08/2013
No. of Samples	2	Run #	Optimization

Reagents	**Volume / reaction (µl)**
Water	6.3
5x Buffer (with $MgCl_2$)	4.0
Enhancer	1.0
$MgCl_2$ (25mM)	0.4
dNTP	0.4
Primer 1 (10µM)	0.4
Primer 2 (10µM)	0.4
KappaRobust Taq	0.1
Sample	7.0
Total reaction volume	20.0

Thermocycling Conditions		
Temperature (°C)	**Time**	**Cycle**
94	4 min	Hold
94	15 sec	X 40
57 - 65	30 sec	
72	30 sec	
72	7 min	Hold
4	∞	Hold

(Used the same cycles for Step 1 and 2)

No.	**Temperature (°C)**	**Class 1C**
6 (1)	57	✗
6 (2)	59	✗
6 (3)	61	✗
6 (4)	63	✗
6 (5)	65	✗

DNA Extraction			
Name	Phiwa Mjoli	Sample	DBS
Date	11-08-14	Extraction Method	QIAGEN Mini Kit

Nested PCR			
Name	Phiwokuhle Mjoli	Name of Kit	KappaRobust
Date	09-04-15	Kit Lot #	
Study	HLA Typing	Expiry Date	08/2013
No. of Samples	2	Run #	Optimization

Reagents	**Volume / reaction (µl)**
Water	6.3
GC Buffer (with $MgCl_2$)	4.0
Enhancer	1.0
$MgCl_2$ (25mM)	0.4
dNTP	0.4
Primer 1 (10µM)	0.4
Primer 2 (10µM)	0.4
KappaRobust Taq	0.1
Sample	7.0
Total reaction volume	20.0

Thermocycling Conditions		
Temperature (°C)	**Time**	**Cycle**
94	4 min	Hold
94	15 sec	X 40
57- 65	30 sec	
72	30 sec	
72	7 min	Hold
4	∞	Hold

(Used the same cycles for Step 1 and 2)

No.	**Temperature (°C)**	**Class 1C**
7 (8)	57	✗
7 (9)	59	✗
7 (10)	61	✗
7 (11)	63	✗
7 (12)	65	✗

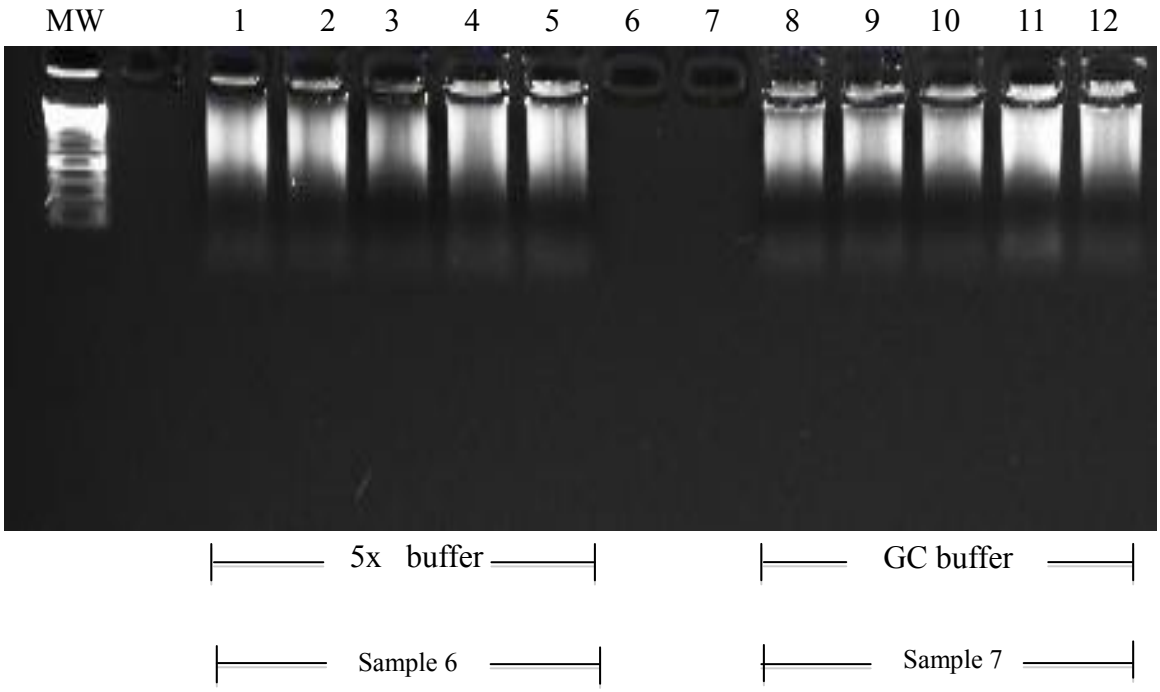


Figure 9: Image of an electrophoresis gel showing amplification of sample 6 and 7 using two different buffers (GC and 5x buffer); sample volume at 7 µl with annealing temperature ranging from 57° to 65°C for HLA Class 1C allele. (Well 1 and 8 = 57°C; well 2 and 9 = 59°C; well 3 and 10 = 61°C; well 4 and 11 = 63°C; well 5 and 12 = 65°C)

Sample that had previously given the expected product (6 and 7) were used in the PCR amplification using different buffers (5x buffer and GC buffer), as the GC content of the HLA Class 1C primers was higher than 65% (requirement for using GC buffer). Both buffers showed no success in amplifying the desired target at all temperatures (Figure 9).

The above amplification protocol was repeated but instead of using DNA that had been extracted and stored, a new sample was extracted and amplified on the same day.

C1.11 Attempt of HLA amplification using two different buffers (GC and 5x buffer) at different annealing temperatures on a freshly extracted sample

DNA Extraction			
Name	Phiwa Mjoli	Sample	DBS
Date	09-04-15	Extraction Method	QIAGEN Mini Kit

Nested PCR			
Name	Phiwokuhle Mjoli	Name of Kit	KappaRobust
Date	09-04-15	Kit Lot #	
Study	HLA Typing	Expiry Date	08/2013
No. of Samples	1	Run #	Optimization

Reagents	**Volume / reaction (µl)**
Water	6.3
5x Buffer (with $MgCl_2$)	4.0
Enhancer	1.0
$MgCl_2$ (25mM)	0.4
dNTP	0.4
Primer 1 (10µM)	0.4
Primer 2 (10µM)	0.4
KappaRobust Taq	0.1
Sample	7.0
Total reaction volume	20.0

Thermocycling Conditions		
Temperature (°C)	**Time**	**Cycle**
94	4 min	Hold
94	15 sec	X 40
55	30 sec	
72	30 sec	
72	7 min	Hold
4	∞	Hold

(Used the same cycles for Step 1 and 2)

No.	**Class 1C**
16 (5)	✗
16 (6)	✗

DNA Extraction			
Name	Phiwa Mjoli	Sample	DBS
Date	09-04-15	Extraction Method	QIAGEN Mini Kit

Nested PCR			
Name	Phiwokuhle Mjoli	Name of Kit	KappaRobust
Date	09-04-15	Kit Lot #	
Study	HLA Typing	Expiry Date	08/2013
No. of Samples	1	Run #	Optimization

Reagents	**Volume / reaction (µl)**
Water	6.3
GC Buffer (with $MgCl_2$)	4.0
Enhancer	1.0
$MgCl_2$ (25mM)	0.4
dNTP	0.4
Primer 1 (10µM)	0.4
Primer 2 (10µM)	0.4
KappaRobust Taq	0.1
Sample	7.0
Total reaction volume	20.0

Thermocycling Conditions		
Temperature (°C)	**Time**	**Cycle**
94	4 min	Hold
94	15 sec	X 40
55	30 sec	
72	30 sec	
72	7 min	Hold
4	∞	Hold

(Used the same cycles for Step 1 and 2)

No.	**Class 1C**
16 (2)	✗
16 (3)	✗

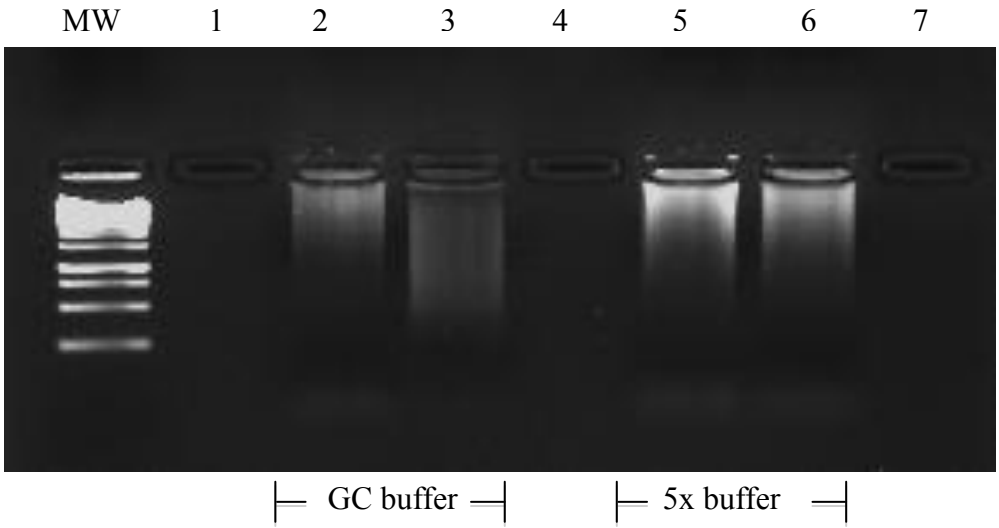


Figure 10: Image of an electrophoresis gel showing amplification of freshly extracted DNA (sample 16) comparing two different buffers (GC and 5x buffer) using a sample volume of 7 µl for HLA class 1C allele.

A fresh DNA sample using the stored DBS was extracted with the QIAamp extraction kit and PCR was performed on the sample (same day) comparing two different buffers (GC and 5x buffer) as performed in Figure 9. No amplification was accomplished for both samples as seen in Figure 10.

We were unsure whether the problem lay with the quality of the DNA sample. We therefore decided to perform these experiments with fresh tissue. We extracted DNA from cheek swabs using a specifically designed forensic kit.

C1.12 Attempt of HLA amplification from fresh DNA extracted from cheek epithelium using the PrepFiler Forensic DNA extraction kit at two different sample volumes

DNA Extraction			
Name	Phiwa Mjoli	Sample	Cheek epithelium
Date	20-04-15	Extraction Method	PrepFiler Forensic Kit

Nested PCR			
Name	Phiwokuhle Mjoli	Name of Kit	KappaRobust
Date	20-04-15	Kit Lot #	
Study	HLA Typing	Expiry Date	08/2013
No. of Samples	1	Run #	Optimization

Reagents	**Volume / reaction (µl)**
Water	12.3
10x Buffer (with $MgCl_2$)	4.0
Enhancer	1.0
$MgCl_2$ (25mM)	0.4
dNTP	0.4
Primer 1 (10µM)	0.4
Primer 2 (10µM)	0.4
KappaRobust Taq	0.1
Sample	1.0
Total reaction volume	20.0

Thermocycling Conditions		
Temperature (°C)	**Time**	**Cycle**
94	4 min	Hold
94	15 sec	X 40
55	30 sec	
72	30 sec	
72	7 min	Hold
4	∞	Hold

(Used the same cycles for Step 1 and 2)

No.	**Class 1C**
17 (1)	✓
17 (2)	✓

DNA Extraction			
Name	Phiwa Mjoli	Sample	Cheek epithelium
Date	20-04-15	Extraction Method	PrepFiler Forensic Kit

Nested PCR			
Name	Phiwokuhle Mjoli	Name of Kit	KappaRobust
Date	20-04-15	Kit Lot #	
Study	HLA Typing	Expiry Date	08/2013
No. of Samples	1	Run #	Optimization

Reagents	**Volume / reaction (µl)**
Water	11.3
10x Buffer (with $MgCl_2$)	4.0
Enhancer	1.0
$MgCl_2$ (25mM)	0.4
dNTP	0.4
Primer 1 (10µM)	0.4
Primer 2 (10µM)	0.4
KappaRobust Taq	0.1
Sample	2.0
Total reaction volume	20.0

Thermocycling Conditions		
Temperature (°C)	**Time**	**Cycle**
94	4 min	Hold
94	15 sec	X 40
55	30 sec	
72	30 sec	
72	7 min	Hold
4	∞	Hold

(Used the same cycles for Step 1 and 2)

No.	**Class 1C**
17 (3)	✗
17 (4)	✗

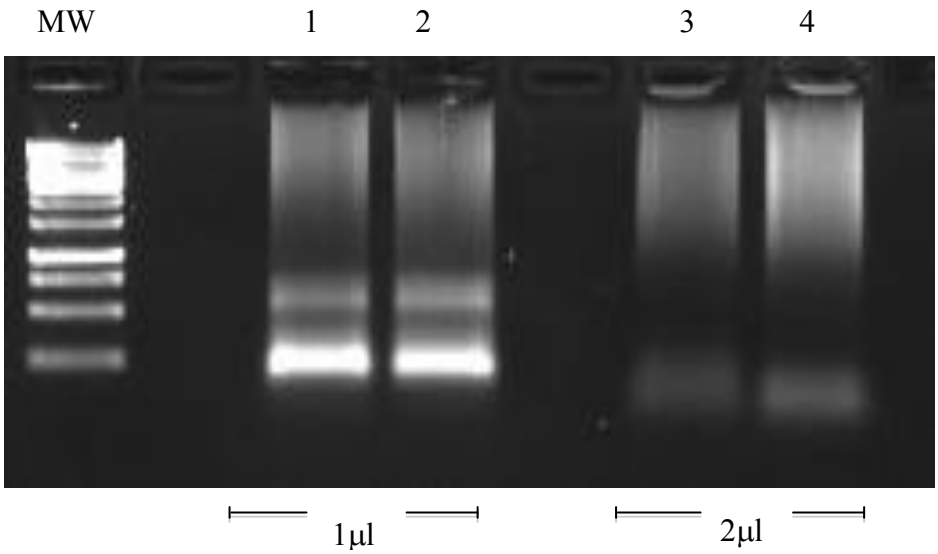


Figure 11: Image of an electrophoresis gel showing amplification of freshly extracted DNA from cheek epithelium (sample 17) using the PrepFiler Forensic DNA extraction kit; amplified using 1 µl and 2 µl sample volume for HLA class 1C allele.

A fresh sample was extracted using the PrepFiler Forensic DNA extracting kit but instead of extracting the DNA from dried blood spots, cheek epithelium was used instead. This was done to assess whether failure of PCR amplification was due to the use of DNA extracted from DBS. The sample volume used for the amplification was 1 µl and 2 µl; and the PCR was performed on the same day as the extraction. As seen on Figure 11, using 1 µl sample volume there was amplification (but doublets). No amplification was seen when the sample volume was at 2 µl.

Since some form of amplification was observed when the sample volume was decreased to 1 µl, a stored sample was then used for amplification using sample volumes below 4 µl to see if decreasing the sample volume further would have an effect on amplification.

C1.13 Attempt of HLA amplification performed on a stored extracted sample using sample input below 4 µl (Ranging between 1-3 µl)

DNA Extraction			
Name	Phiwa Mjoli	Sample	DBS
Date	11-08-14	Extraction Method	QIAGEN Mini Kit

Nested PCR			
Name	Phiwokuhle Mjoli	Name of Kit	KappaRobust
Date	21-04-15	Kit Lot #	
Study	HLA Typing	Expiry Date	08/2013
No. of Samples	1	Run #	Optimization

Reagents	**Volume / reaction (µl)**
Water	12.3
10x Buffer (with $MgCl_2$)	4.0
Enhancer	1.0
$MgCl_2$ (25mM)	0.4
dNTP	0.4
Primer 1 (10µM)	0.4
Primer 2 (10µM)	0.4
KappaRobust Taq	0.1
Sample	1.0
Total reaction volume	20.0

Thermocycling Conditions		
Temperature (°C)	**Time**	**Cycle**
94	4 min	Hold
94	15 sec	X 40
55	30 sec	
72	30 sec	
72	7 min	Hold
4	∞	Hold

(Used the same cycles for Step 1 and 2)

No.	**Class 1C**
7 (1)	✓
7 (2)	✓

DNA Extraction			
Name	Phiwa Mjoli	Sample	DBS
Date	11-08-14	Extraction Method	QIAGEN Mini Kit

Nested PCR			
Name	Phiwokuhle Mjoli	Name of Kit	KappaRobust
Date	21-04-15	Kit Lot #	
Study	HLA Typing	Expiry Date	08/2013
No. of Samples	1	Run #	Optimization

Reagents	**Volume / reaction (µl)**
Water	11.3
10x Buffer (with $MgCl_2$)	4.0
Enhancer	1.0
$MgCl_2$ (25mM)	0.4
dNTP	0.4
Primer 1 (10µM)	0.4
Primer 2 (10µM)	0.4
KappaRobust Taq	0.1
Sample	2.0
Total reaction volume	20.0

Thermocycling Conditions		
Temperature (°C)	**Time**	**Cycle**
94	4 min	Hold
94	15 sec	X 40
55	30 sec	
72	30 sec	
72	7 min	Hold
4	∞	Hold

(Used the same cycles for Step 1 and 2)

No.	**Class 1C**
7 (3)	✗
7 (4)	✗

DNA Extraction			
Name	Phiwa Mjoli	Sample	DBS
Date	11-08-14	Extraction Method	QIAGEN Mini Kit

Nested PCR			
Name	Phiwokuhle Mjoli	Name of Kit	KappaRobust
Date	21-04-15	Kit Lot #	
Study	HLA Typing	Expiry Date	08/2013
No. of Samples	1	Run #	Optimization

Reagents	**Volume / reaction (µl)**
Water	10.3
10x Buffer (with $MgCl_2$)	4.0
Enhancer	1.0
$MgCl_2$ (25mM)	0.4
dNTP	0.4
Primer 1 (10µM)	0.4
Primer 2 (10µM)	0.4
KappaRobust Taq	0.1
Sample	3.0
Total reaction volume	20.0

Thermocycling Conditions		
Temperature (°C)	**Time**	**Cycle**
94	4 min	Hold
94	15 sec	X 40
55	30 sec	
72	30 sec	
72	7 min	Hold
4	∞	Hold

(Used the same cycles for Step 1 and 2)

No.	**Class 1C**
7 (5)	✗
7 (6)	✗

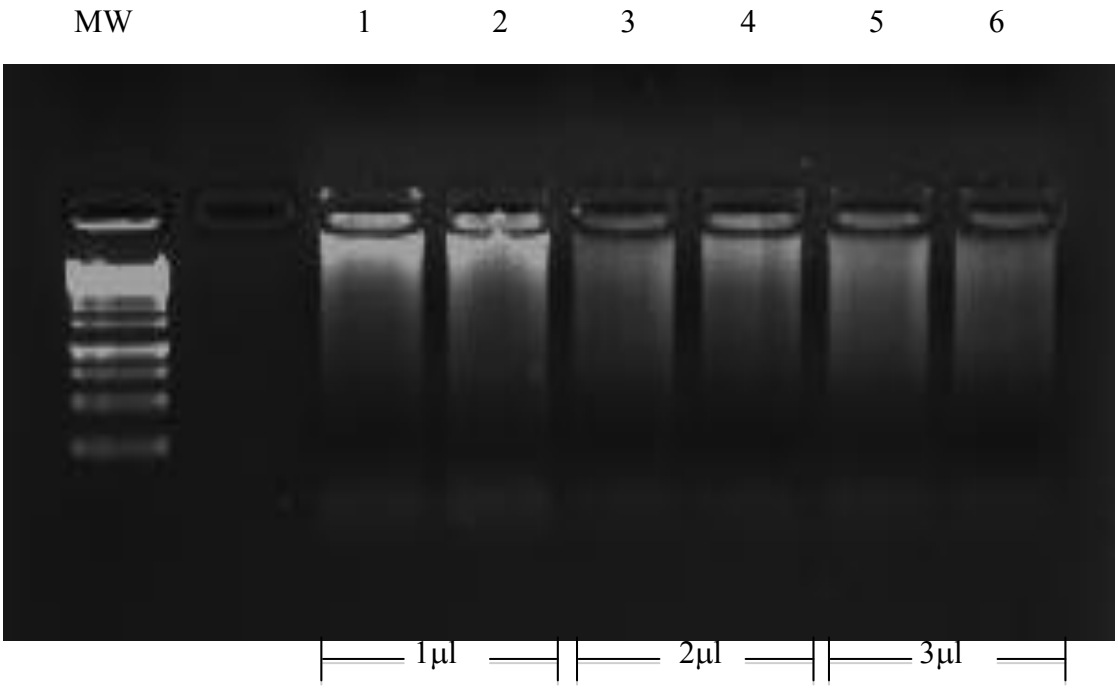


Figure 12: Image of an electrophoresis gel showing amplification of sample 7 with sample volume ranging from 1 to 3 µl using primers specific for the HLA class 1C allele.

A stored DNA sample (sample 7) was used to perform PCR using a decreased input sample volume - below 4 µl (ranging between 1-3 µl) as 1 µl (Figure 11) had showed some form of amplification even though it was non-specific. However, using lower input volume resulted in no amplification (Figure 12).

C1.14 Attempt of HLA amplification using a freshly extracted DNA sample extracted using the QIAamp extraction kit from stored dried blood spots

DNA Extraction			
Name	Phiwa Mjoli	Sample	DBS
Date	23-04-15	Extraction Method	QIAGEN Mini Kit

Nested PCR			
Name	Phiwokuhle Mjoli	Name of Kit	KappaRobust
Date	23-04-15	Kit Lot #	
Study	HLA Typing	Expiry Date	08/2013
No. of Samples	1	Run #	Optimization

Reagents	**Volume / reaction (µl)**
Water	12.3
10x Buffer (with $MgCl_2$)	4.0
Enhancer	1.0
$MgCl_2$ (25mM)	0.4
dNTP	0.4
Primer 1 (10µM)	0.4
Primer 2 (10µM)	0.4
KappaRobust Taq	0.1
Sample	1.0
Total reaction volume	20.0

Thermocycling Conditions		
Temperature (°C)	**Time**	**Cycle**
94	4 min	Hold
94	15 sec	X 40
55	30 sec	
72	30 sec	
72	7 min	Hold
4	∞	Hold

(Used the same cycles for Step 1 and 2)

No.	**Class 1C**
18 (1)	✓
18 (2)	✓

DNA Extraction			
Name	Phiwa Mjoli	Sample	DBS
Date	23-04-15	Extraction Method	QIAGEN Mini Kit

Nested PCR			
Name	Phiwokuhle Mjoli	Name of Kit	KappaRobust
Date	23-04-15	Kit Lot #	
Study	HLA Typing	Expiry Date	08/2013
No. of Samples	1	Run #	Optimization

Reagents	**Volume / reaction (µl)**
Water	11.3
10x Buffer (with $MgCl_2$)	4.0
Enhancer	1.0
$MgCl_2$ (25mM)	0.4
dNTP	0.4
Primer 1 (10µM)	0.4
Primer 2 (10µM)	0.4
KappaRobust Taq	0.1
Sample	2.0
Total reaction volume	20.0

Thermocycling Conditions		
Temperature (°C)	**Time**	**Cycle**
94	4 min	Hold
94	15 sec	X 40
55	30 sec	
72	30 sec	
72	7 min	Hold
4	∞	Hold

(Used the same cycles for Step 1 and 2)

No.	**Class 1C**
18 (4)	✗
18 (5)	✗

DNA Extraction			
Name	Phiwa Mjoli	Sample	DBS
Date	23-04-15	Extraction Method	QIAGEN Mini Kit

Nested PCR			
Name	Phiwokuhle Mjoli	Name of Kit	KappaRobust
Date	23-04-15	Kit Lot #	
Study	HLA Typing	Expiry Date	08/2013
No. of Samples	1	Run #	Optimization

Reagents	**Volume / reaction (µl)**
Water	10.3
10x Buffer (with $MgCl_2$)	4.0
Enhancer	1.0
$MgCl_2$ (25mM)	0.4
dNTP	0.4
Primer 1 (10µM)	0.4
Primer 2 (10µM)	0.4
KappaRobust Taq	0.1
Sample	3.0
Total reaction volume	20.0

Thermocycling Conditions		
Temperature (°C)	**Time**	**Cycle**
94	4 min	Hold
94	15 sec	X 40
55	30 sec	
72	30 sec	
72	7 min	Hold
4	∞	Hold

(Used the same cycles for Step 1 and 2)

No.	**Class 1C**
18 (7)	✗
18 (8)	✗

DNA Extraction			
Name	Phiwa Mjoli	Sample	DBS
Date	23-04-15	Extraction Method	QIAGEN Mini Kit

Nested PCR			
Name	Phiwokuhle Mjoli	Name of Kit	KappaRobust
Date	23-04-15	Kit Lot #	
Study	HLA Typing	Expiry Date	08/2013
No. of Samples	1	Run #	Optimization

Reagents	**Volume / reaction (µl)**
Water	9.3
10x Buffer (with $MgCl_2$)	4.0
Enhancer	1.0
$MgCl_2$ (25mM)	0.4
dNTP	0.4
Primer 1 (10µM)	0.4
Primer 2 (10µM)	0.4
KappaRobust Taq	0.1
Sample	4.0
Total reaction volume	20.0

Thermocycling Conditions		
Temperature (°C)	**Time**	**Cycle**
94	4 min	Hold
94	15 sec	X 40
55	30 sec	
72	30 sec	
72	7 min	Hold
4	∞	Hold

(Used the same cycles for Step 1 and 2)

No.	**Class 1C**
18 (11)	✗
18 (12)	✗

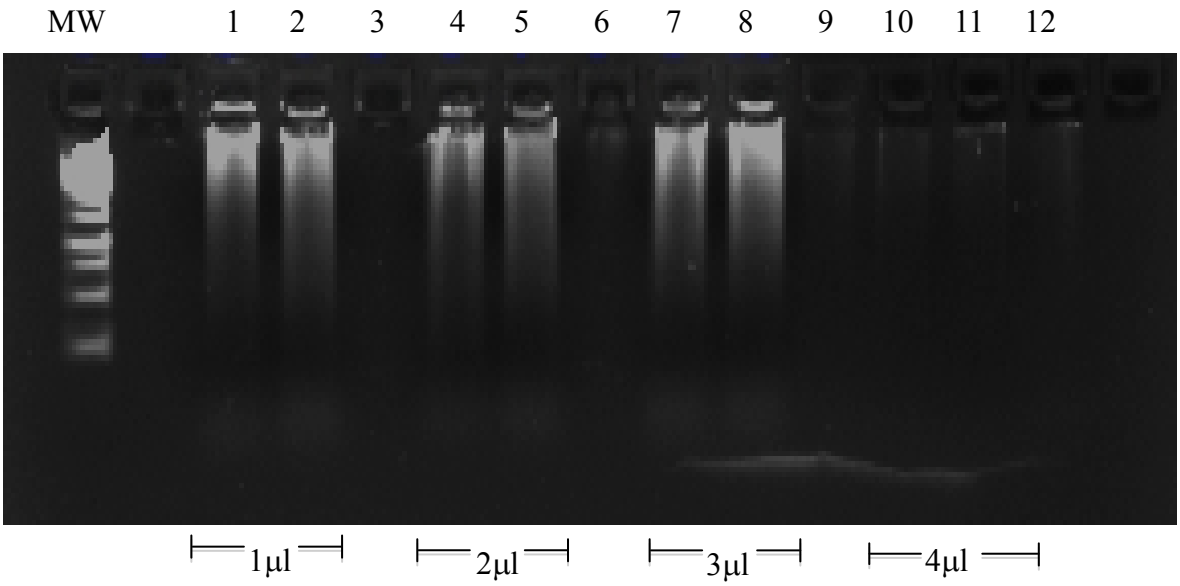


Figure 13: Image of an electrophoresis gel showing amplification of freshly extracted DNA (sample 18) with sample volumes ranging from 1 µl to 4 µl and using primers specific for the HLA class 1C allele.

Since no success seen using a stored sample (Figure 12), a freshly extracted DNA sample extracted using the QIAamp extraction kit from the stored dried blood spots were used to perform PCR (same day) with sample volume ranging from 1µl to 4µl. PCR amplification was not a success as seen in Figure 13.

To determine whether the problem lay in primer specificity i.e. primers specific for different racial groups, we performed a small study to investigate this.

C1.15 Attempt of HLA amplification using DNA extracted from cheek epithelium of individuals belonging to different ethnic groups

DNA Extraction			
Name	Phiwa Mjoli	Sample	Cheek epithelium
Date	06-05-15	Extraction Method	QIAGEN Mini Kit

Nested PCR			
Name	Phiwokuhle Mjoli	Name of Kit	KappaRobust
Date	06-05-15	Kit Lot #	
Study	HLA Typing	Expiry Date	08/2013
No. of Samples	4	Run #	Optimization

Reagents	**Volume / reaction (µl)**
Water	12.3
10x Buffer (with $MgCl_2$)	4.0
Enhancer	1.0
$MgCl_2$ (25mM)	0.4
dNTP	0.4
Primer 1 (10µM)	0.4
Primer 2 (10µM)	0.4
KappaRobust Taq	0.1
Sample	1.0
Total reaction volume	20.0

Thermocycling Conditions		
Temperature (°C)	**Time**	**Cycle**
94	4 min	Hold
94	15 sec	X 40
55	30 sec	
72	30 sec	
72	7 min	Hold
4	∞	Hold

(Used the same cycles for Step 1 and 2)

No.	**Class 1C**
2 (1)	✗
3 (4)	✗
4 (7)	✗
5 (10)	✗

DNA Extraction			
Name	Phiwa Mjoli	Sample	Cheek epithelium
Date	06-05-15	Extraction Method	QIAGEN Mini Kit

Nested PCR			
Name	Phiwokuhle Mjoli	Name of Kit	KappaRobust
Date	06-05-15	Kit Lot #	
Study	HLA Typing	Expiry Date	08/2013
No. of Samples	4	Run #	Optimization

Reagents	**Volume / reaction (µl)**
Water	11.3
10x Buffer (with $MgCl_2$)	4.0
Enhancer	1.0
$MgCl_2$ (25mM)	0.4
dNTP	0.4
Primer 1 (10µM)	0.4
Primer 2 (10µM)	0.4
KappaRobust Taq	0.1
Sample	2.0
Total reaction volume	20.0

Thermocycling Conditions		
Temperature (°C)	**Time**	**Cycle**
94	4 min	Hold
94	15 sec	X 40
55	30 sec	
72	30 sec	
72	7 min	Hold
4	∞	Hold

(Used the same cycles for Step 1 and 2)

No.	**Class 1C**
2 (2)	✗
3 (5)	✗
4 (8)	✗
5 (11)	✗

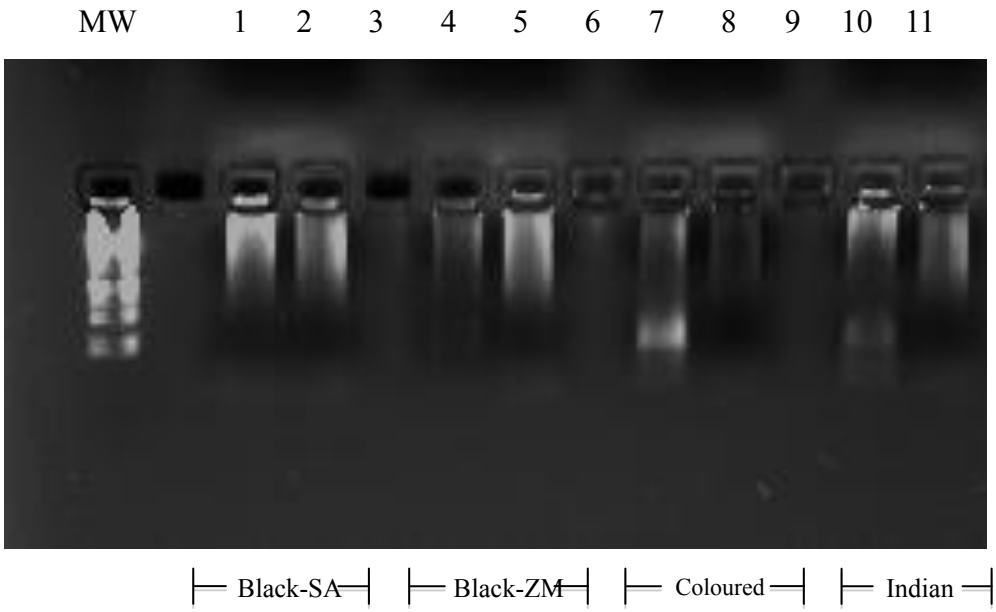


Figure 14: Image of an electrophoresis gel showing amplification of freshly extracted DNA (sample 2-5 on Table 2) with sample volume ranging from 1µl to 2µl using primers specific for the HLA class 1C allele.

After being able to amplify DNA with PCR using cheek epithelium even though it was non-specific (Figure 11) it was thought that it would be interesting to see if there is a difference in amplification using DNA from individuals belonging to different ethnic groups. Three ethic groups were compared namely Indian, Coloured and Black. The "Black" group was further split into two, namely Black-Zimbabwe and Black-SA.

The results in Figure 14 showed better amplification for the Coloured group followed by the Indian group; Black-Zimbabwe had a very faint band and there was no amplification for Black-SA. It was expected that more PCR amplification would be seen in the Black-SA individual as the cohort that the PCR primers being used in the troubleshooting had previously worked perfectly on, is based in the rural areas of a Black-SA dominant area.

This suggested that either there were inhibitors in the samples or the PCR reactions required more optimization. It was decided to continue this work at the Africa Centre laboratory at the UKZN Medical School where this protocol had been developed.

C.2 PRC optimization performed at the Africa Centre for Health and Population Studies Genomics Lab (Durban).

C2.1 DNA extraction

DNA was extracted and quantified as stated in section 2.2 and 2.3 of Materials and methods. The samples used for this section of optimization are the samples shown in Table 3.1 of the Results section under ‘Evaluation of DNA Extraction’. The 1% agarose gels were made following the instructions in Appendix B (iv.).

C2.2 Attempt of amplification of HLA Class 1A, 1B and 1C; comparing the efficiency of two PCR extraction Kit (Individual reagents vs. ReadyMix (RM))

DNA Extraction			
Name	Phiwa Mjoli	Sample	DBS
Date	06-07-15	Extraction Method	QIAGEN Mini Kit

Nested PCR			
Name	Phiwokuhle Mjoli	Name of Kit	KappaRobust
Date	07-07-15	Kit Lot #	
Study	HLA Typing	Expiry Date	08/2013
No. of Samples	12	Run #	Optimization

Thermocycling Conditions		
Temperature (°C)	**Time**	**Cycle**
94	4 min	Hold
94	15 sec	X 40
63	30 sec	
72	30 sec	
72	7 min	Hold
4	∞	Hold

(Used the same cycles for Step 1 and Step 2)

Reagents	**Volume / reaction (µl)**	**Mater mix for 12 reactions (µl)**
Water	8.3	99,6
10x Buffer (with $MgCl_2$)	4.0	48
Enhancer	1.0	12
$MgCl_2$	0.4	4,8
dNTP	0.4	4,8
Primer 1 (10uM)	0.4	4,8
Primer 2 (10uM)	0.4	4,8
KappaRobust Taq	0.1	1,2
Sample	5.0	-
Total reaction volume	**20.0**	**180 (15µl/tube)**

The following table is a summary of the results seen on the electrophoresis gel that follows (Figure 1), showing the success/failure of the attempted HLA Class 1B and HLA Class 1C amplification. (This applies to all tables of similar format to follow).

No.	Sample ID	Class 1A	Class 1B	Class 1C
1	#1	✗	✓	✓
2	#2	✗	✓	✓
3	#3	✓	✓	✓
4	#4	✗	✓	✓
5	#5	✓	✓	✓
6	#6	✓	✓	✓
7	#7	✓	✓	✓
8	#8	✓	✓	✓
9	#9	✓	✓	✓
10	#10	✓	✓	✓
11	PC (+ve)	✓	✓	✓
12	NTC (-ve)	✓	✓	✓

DNA Extraction			
Name	Phiwa Mjoli	Sample	DBS
Date	06-07-15	Extraction Method	QIAGEN Mini Kit

Nested PCR			
Name	Phiwokuhle Mjoli	Name of Kit	KappaRobust (RM)
Date	07-07-15	Kit Lot #	
Study	HLA Typing	Expiry Date	08/2013
No. of Samples	12	Run #	Optimization

Thermocycling Conditions		
Temperature (°C)	**Time**	**Cycle**
94	4 min	Hold
94	15 sec	X 40
63	30 sec	
72	30 sec	
72	7 min	Hold
4	∞	Hold

(Use the same cycles for Step 1 and Step 2)

Reagents	**Volume / reaction**	**Volume for 12 reactions**
Water	3.8	45.6
Master Mix	10	120
$MgCl_2$	0.4	4.8
Primer 1(10uM)	0.4	4.8
Primer 2 (10uM)	0.4	4.8
Sample	5	-
Total reaction volume	**20**	**180 (15µl/tube)**

No.	**Sample ID**	**Class 1A**	**Class 1B**	**Class 1C**
1	#1	✓	✓	✓
2	#2	✓	✓	✓
3	#3	✓	✓	✓
4	#4	✓	✓	✓
5	#5	✓	✓	✓
6	#6	✓	✓	✓
7	#7	✓	✓	✗
8	#8	✓	✓	✓
9	#9	✓	✓	✓
10	#10	✓	✓	✓
11	PC (+ve)	✓	✓	✓
12	NTC (-ve)	✓	✓	✓

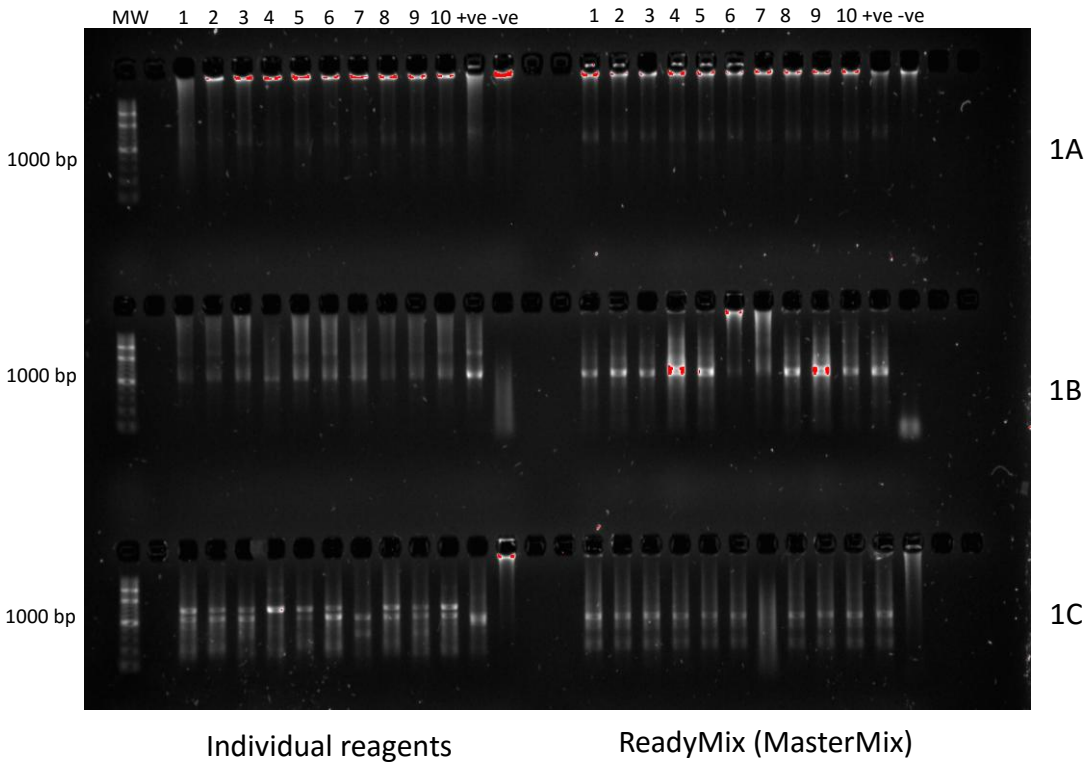


Figure 1: Electrophoresis gel showing amplification of HLA Class 1A, 1B and 1C (following the original protocol) comparing efficiency of individual reagents compared to that of the ready mix (MasterMix).

From the image above it can be seen that the ReadyMix (MasterMix) is more efficient in HLA amplification when compared to the individual reagents. This decision was based on that HLA Class 1B had less non-specific amplification using this kit and the bands produced were bolder compared to the individual reagents bands.

The protocol used showed efficiency in amplification of the HLA class 1B allele. However, for HLA Class 1A the bands were present but very faint therefore optimizing was needed for this class. Optimizing was also needed for HLA Class 1C as there was non-specific in all the samples amplified.

All protocol optimization were subsequently carried out using the ReadyMix (RM) due to reasons stated above.

C2.3 Attempt of HLA amplification after repeating Step 2 of the PCR reaction for HLA Class 1A at a lower Tm (57°C) comparing 3 different sample volumes

Extraction			
Name	Phiwa Mjoli	Sample	DBS
Date	06-07-15	Extraction Method	QIAGEN Mini Kit

Nested PCR			
Name	Phiwokuhle Mjoli	Name of Kit	KapaRobust (RM)
Date	08-07-15	Kit Lot #	
Study	HLA Typing	Expiry Date	08/2013
No. of Samples	12	Run #	Optimization

Thermocycling Conditions		
Temperature (°C)	**Time**	**Cycle**
94	4 min	Hold
94	15 sec	X 40
57	30 sec	
72	30 sec	
72	7 min	Hold
4	∞	Hold

Reagents	**Volume / reaction**	**Volume for 12 reactions**
Water	7.5	97.5
Master Mix	12.5	162.5
Enhancer	1.0	13
$MgCl_2$	0.5	6.5
Primer 1(10uM)	0.5	6.5
Primer 2 (10uM)	0.5	6.5
Sample	2.5	-
Total reaction volume	**25**	

No.	**Sample ID**	**Class 1A**
1	#1	✓
2	#2	✓
3	#3	✓
4	#4	✓
5	#5	✓
6	#6	✓
7	#7	✓
8	#8	✓
9	#9	✓
10	#10	✓
11	PC (+ve)	✓
12	NTC (-ve)	✗

DNA Extraction			
Name	Phiwa Mjoli	Sample	DBS
Date	06-07-15	Extraction Method	QIAGEN Mini Kit

Nested PCR			
Name	Phiwokuhle Mjoli	Name of Kit	KapaRobust (RM)
Date	08-07-15	Kit Lot #	
Study	HLA Typing	Expiry Date	08/2013
No. of Samples	12	Run #	Optimization

Reagents	**Volume / reaction**	**Volume for 12 reactions**
Water	5	65
Master Mix	12.5	162.5
Enhancer	1.0	13
$MgCl_2$	0.5	6.5
Primer 1(10uM)	0.5	6.5
Primer 2 (10uM)	0.5	6.5
Sample	5	-
Total reaction volume	**25**	

Thermocycling Conditions		
Temperature (°C)	**Time**	**Cycle**
94	4 min	Hold
94	15 sec	X 40
57	30 sec	
72	30 sec	
72	7 min	Hold
4	∞	Hold

No.	**Sample ID**	**Class 1A**
1	#1	✓
2	#2	✓
3	#3	✓
4	#4	✓
5	#5	✓
6	#6	✓
7	#7	✓
8	#8	✓
9	#9	✓
10	#10	✓
11	PC (+ve)	✓
12	NTC (-ve)	✓

DNA Extraction			
Name	Phiwa Mjoli	Sample	DBS
Date	06-07-15	Extraction Method	QIAGEN Mini Kit

Nested PCR			
Name	Phiwokuhle Mjoli	Name of Kit	KapaRobust (RM)
Date	08-07-15	Kit Lot #	
Study	HLA Typing	Expiry Date	08/2013
No. of Samples	12	Run #	Optimization

Thermocycling Conditions		
Temperature (°C)	**Time**	**Cycle**
94	4 min	Hold
94	15 sec	X 40
57	30 sec	
72	30 sec	
72	7 min	Hold
4	∞	Hold

Reagents	**Volume / reaction**	**Volume for 12 reactions**
Water	2.5	32.5
Master Mix	12.5	162.5
Enhancer	1.0	13
$MgCl_2$	0.5	6.5
Primer 1(10uM)	0.5	6.5
Primer 2 (10uM)	0.5	6.5
Sample	7,5	-
Total reaction volume	**25**	

No.	**Sample ID**	**Class 1A**
1	#1	✓
2	#2	✓
3	#3	✓
4	#4	✓
5	#5	✓
6	#6	✓
7	#7	✗
8	#8	✓
9	#9	✓
10	#10	✓
11	PC (+ve)	✓
12	NTC (-ve)	✗

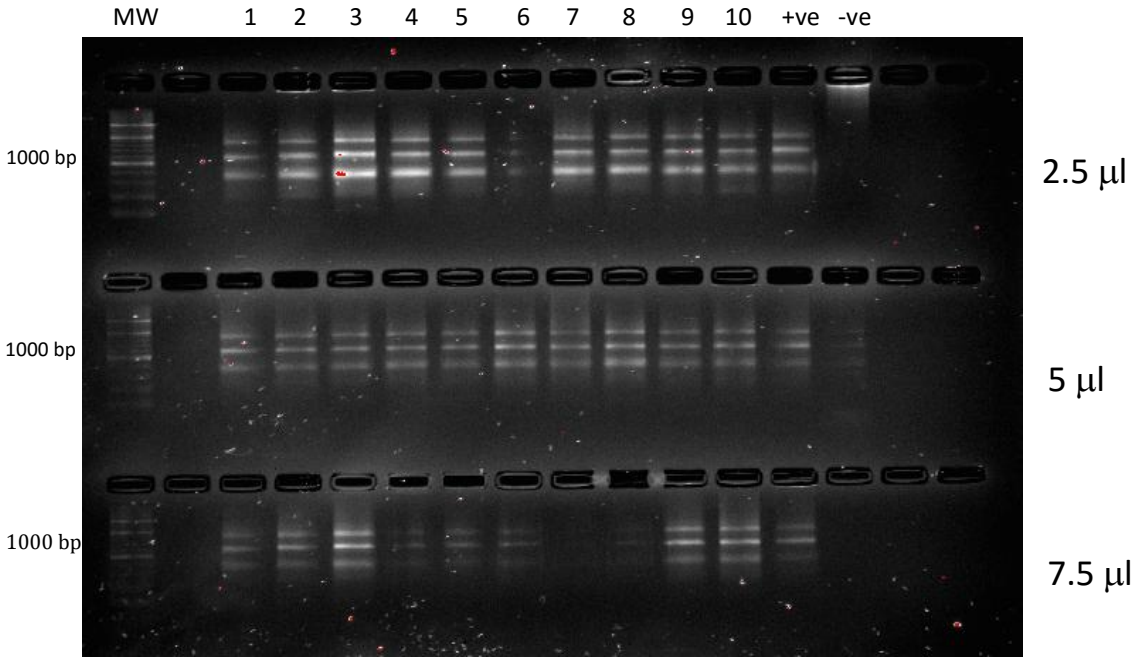


Figure 2: Electrophoresis gel showing amplification after repeating Step 2 of the PCR reaction for HLA Class 1A at a lower T_m (57°C) comparing 3 different sample volumes (2.5, 5 and 7.5 µl).

From the above gel image (Figure 2) it can be seen that 2.5 µl is the most efficient sample volume to be used compared to the other two sample volumes. This is because even though there is non-specific amplification, the desired region (1000 bp) is bold and is present in all samples (with exception of sample #6 - band is faint but present). We therefore decided to optimize the annealing temperature that would produce only the expected bands without the non-specific ones.

C2.4 Attempt of Step 2 amplification with a temperature gradient (65°C- 55°C) for HLA Class 1A using two different sample volumes

DNA Extraction			
Name	Phiwa Mjoli	Sample	DBS
Date	06-07-15	Extraction Method	QIAGEN Mini Kit

Nested PCR			
Name	Phiwokuhle Mjoli	Name of Kit	KapaRobust (RM)
Date	16-07-15	Kit Lot #	
Study	HLA Typing	Expiry Date	08/2013
No. of Samples	Positive control	Run #	Optimization

Thermocycling Conditions (Step 2)		
Temperature (°C)	**Time**	**Cycle**
94	4 min	Hold
94	15 sec	X 40
65-55	30 sec	
72	30 sec	
72	7 min	Hold
4	∞	Hold

Reagents	**Volume / reaction**	**Volume for 9 reactions**
Water	7.5	67.5
Master Mix	12.5	112.5
Enhancer	1.0	9
$MgCl_2$	0.5	4.5
Primer 1(10uM)	0.5	4.5
Primer 2 (10uM)	0.5	4.5
Sample	**2.5**	-
Total reaction volume	**25**	

No.	**Temperature**	**Class 1A**
1	65.0	✓
2	64.5	✓
3	63.3	✓
4	61.4	✓
5	59.0	✗
6	57.0	✗
7	55.7	✗
8	55.0	✗

DNA Extraction			
Name	Phiwa Mjoli	Sample	DBS
Date	06-07-15	Extraction Method	QIAGEN Mini Kit

Nested PCR			
Name	Phiwokuhle Mjoli	Name of Kit	KapaRobust (RM)
Date	16-07-15	Kit Lot #	
Study	HLA Typing	Expiry Date	08/2013
No. of Samples	Positive Control	Run #	Optimization

Thermocycling Conditions (Step 2)		
Temperature (°C)	**Time**	**Cycle**
94	4 min	Hold
94	15 sec	X 40
65-55	30 sec	
72	30 sec	
72	7 min	Hold
4	∞	Hold

Reagents	**Volume / reaction**	**Volume for 9 reactions**
Water	5	45
Master Mix	12.5	112.5
Enhancer	1.0	9
$MgCl_2$	0.5	4.5
Primer 1(10uM)	0.5	4.5
Primer 2 (10uM)	0.5	4.5
Sample	**5**	-
Total reaction volume	**25**	

No.	**Temperature**	**Class 1A**
1	65.0	✓
2	64.5	✓
3	63.3	✓
4	61.4	✗
5	59.0	✗
6	57.0	✗
7	55.7	✗
8	55.0	✗

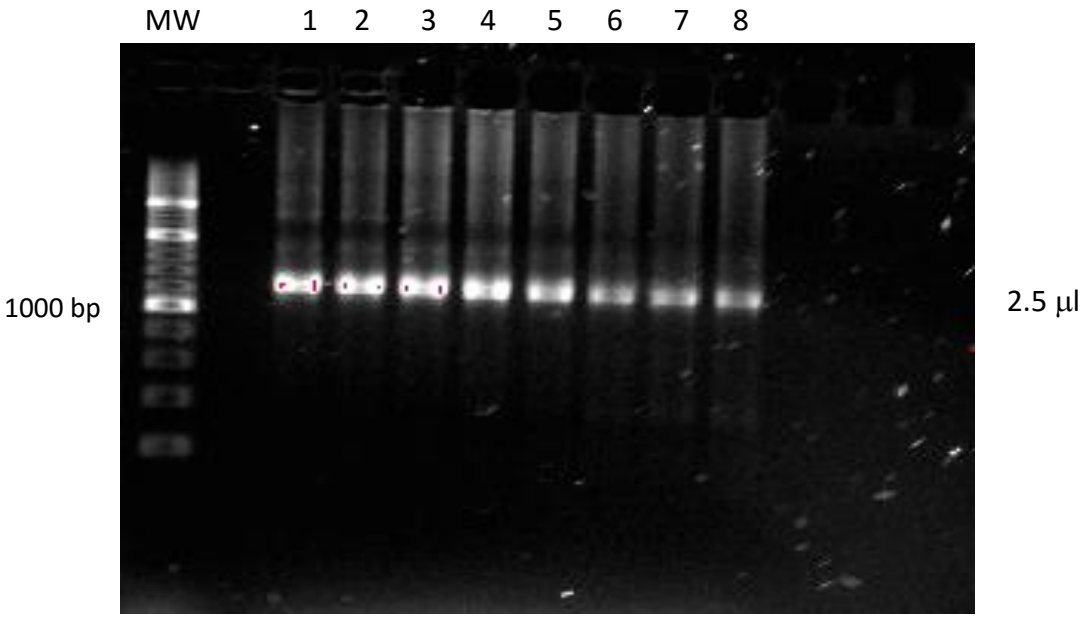


Figure 3: Electrophoresis gel showing Temperature Gradient (65°C- 55°C) amplification of HLA Class 1A using 2.5 µl of sample volume.

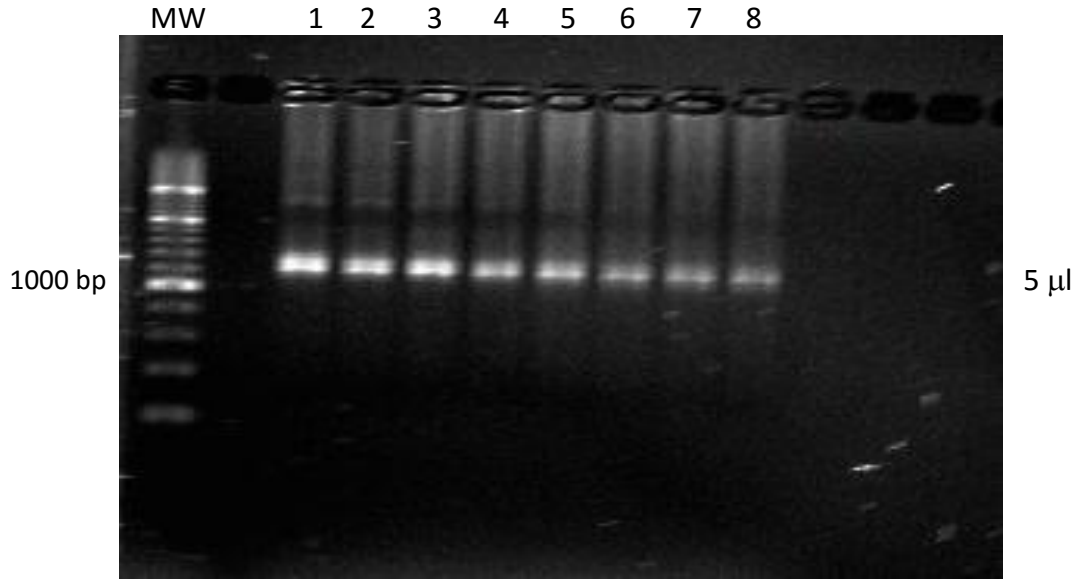


Figure 4: Electrophoresis gel showing Temperature Gradient (65°C-55°C) amplification of HLA Class 1A using 5 µl of sample volume.

The results show that 2.5 µl sample input and 65°C annealing temperature are the most favorable conditions as the boldest expected band is produced under them. We therefore repeated the same experiment using primers specific for HLA Class 1C.

C2.5 Attempt of Step 2 amplification with a temperature gradient (65°C- 55°C) for HLA Class 1C using two different sample volumes

DNA Extraction			
Name	Phiwa Mjoli	Sample	DBS
Date	06-07-15	Extraction Method	QIAGEN Mini Kit

Nested PCR			
Name	Phiwokuhle Mjoli	Name of Kit	KapaRobust (RM)
Date	21-07-15	Kit Lot #	
Study	HLA Typing	Expiry Date	08/2013
No. of Samples	Positive control	Run #	Optimization

Thermocycling Conditions (Step 2)		
Temperature (°C)	**Time**	**Cycle**
94	4 min	Hold
94	15 sec	X 40
65-55	30 sec	
72	30 sec	
72	7 min	Hold
4	∞	Hold

Reagents	**Volume / reaction**	**Volume for 9 reactions**
Water	7.5	67.5
Master Mix	12.5	112.5
Enhancer	1.0	9
$MgCl_2$	0.5	4.5
Primer 1(10uM)	0.5	4.5
Primer 2 (10uM)	0.5	4.5
Sample	**2.5**	-
Total reaction volume	**25**	

No.	**Temperature (°C)**	**Class 1C**
1	65.0	✓
2	64.5	✓
3	63.3	✓
4	61.4	✗
5	59.0	✗
6	57.0	✗
7	55.7	✗
8	55.0	✗

DNA Extraction			
Name	Phiwa Mjoli	Sample	DBS
Date	06-07-15	Extraction Method	QIAGEN Mini Kit

Nested PCR			
Name	Phiwokuhle Mjoli	Name of Kit	KapaRobust (RM)
Date	21-07-15	Kit Lot #	
Study	HLA Typing	Expiry Date	08/2013
No. of Samples	Positive Control	Run #	Optimization

Thermocycling Conditions (Step 2)		
Temperature (°C)	**Time**	**Cycle**
94	4 min	Hold
94	15 sec	X 40
65-55	30 sec	
72	30 sec	
72	7 min	Hold
4	∞	Hold

Reagents	**Volume / reaction**	**Volume for 9 reactions**
Water	5	45
Master Mix	12.5	112.5
Enhancer	1.0	9
$MgCl_2$	0.5	4.5
Primer 1(10uM)	0.5	4.5
Primer 2 (10uM)	0.5	4.5
Sample	**5**	-
Total reaction volume	**25**	

No.	**Temperature (°C)**	**Class 1C**
1	65	✓
2	64.5	✓
3	63.3	✓
4	61.4	✗
5	59.0	✗
6	57.0	✗
7	55.7	✗
8	55.0	✗

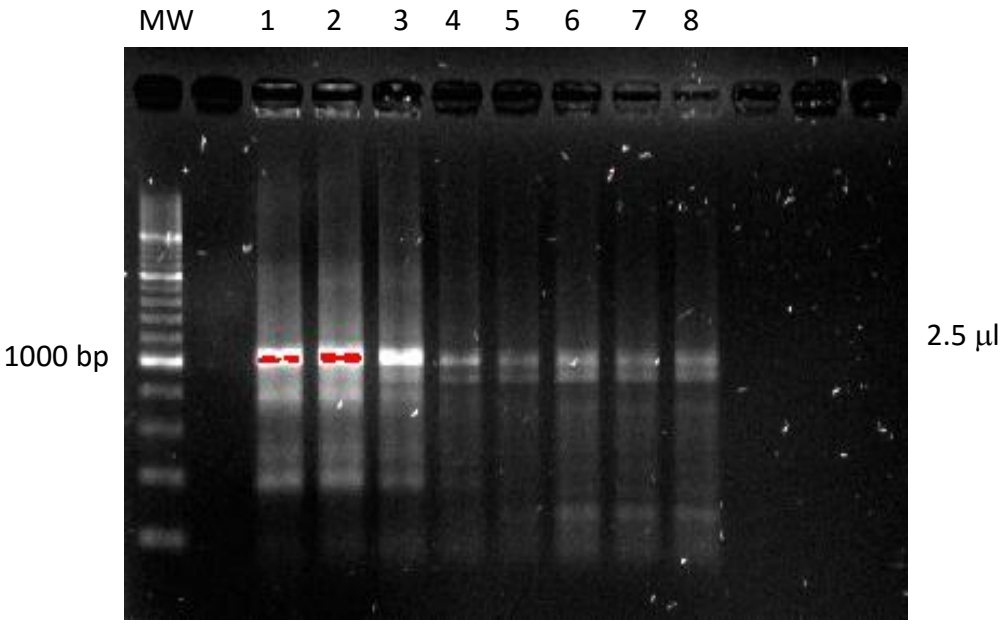


Figure 5: Electrophoresis gel showing Temperature Gradient (65°C-55°C) amplification of HLA Class 1C using 2.5 µl of sample volume.

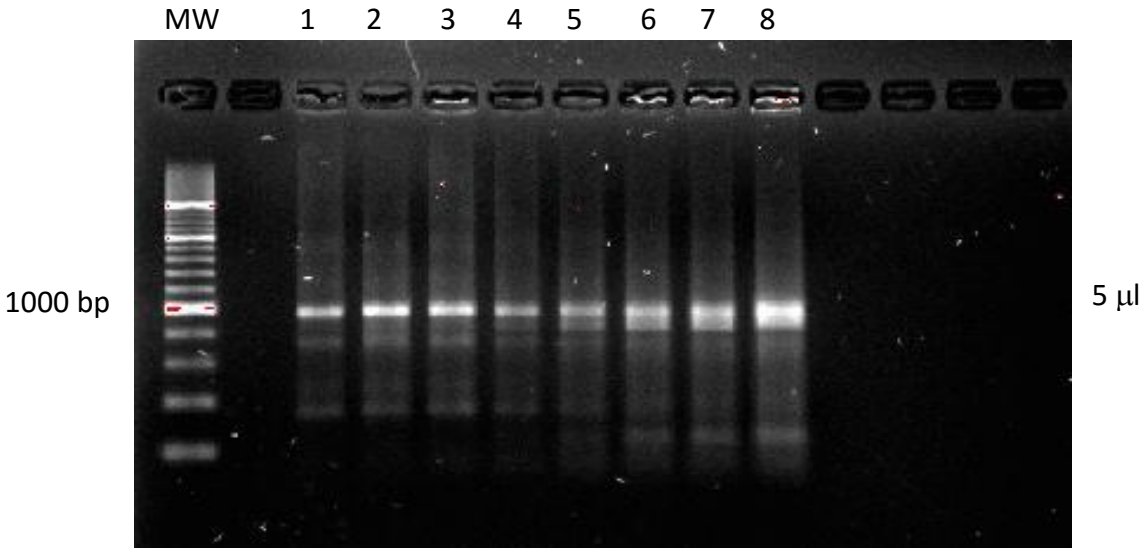


Figure 6: Electrophoresis gel showing Temperature Gradient (65°C-55°C) amplification of HLA Class 1C using 5 µl of sample volume.

Looking at Figure 3-6 it can be seen that the most effective annealing temperatures for both HLA typing Class 1A and HLA Class 1C in Step 2 range between 65°C-63.3°C (best being 65°C) and the best sample input volume to use for these alleles is 2.5 µl compared to 5µl.

The above observations are based on the size of the expected (target) region and how bold the bands are compared to each other. The conclusion was that amplifying the HLA Class 1 loci is best performed at 63°C (T_m) using 5 µl of sample volume for step 1 and at 65°C (T_m) using 2.5 µl sample volume for step 2. The above observations were used to amplify HLA Class 1A, 1B and 1C as seen Table 2.5 and 2.6 of Materials and Methods.

6. REFERENCES

Adawaye, C., Kamangu, E., Moussa, A.M., Tchoumbou B., Vaira, D., Moutschen, M., 2013. Use of Dried Blood Spots to improve the diagnosis and management of HIV in resource-limited settings. World Journal of AIDS 3, 251-256.

Al-Soud, W., Radstrom, P., 2000. Effects of Amplification Facilitators on Diagnostic PCR in the Presence of Blood, Feces, and Meat. Journal of Clinical Microbiology 38(12), 4463–4470.

Barbi, M., Binda, S., Primache, V., Caroppo, S., Dido, P., Guidotti, P., 2000. Cytomegalovirus DNA detection in Guthrie cards: a powerful tool for diagnosing congenital infection. Journal of Clinical Virology 17, 159–65.

Bodmer, W.F., 1987. The HLA system: structure and function. Journal of Clinical Pathology 40, 948-958.

Chaisomchit, S., Wichajam, R., Chowpreecha, S., Chareonsiriwatana, W., 2003. A simple method for extraction and purification of genomic DNA from dried blood spots on filter paper. Southeast Asian J Trop Med Public Health 34 (3), 64-645.

Chaisomchit, S., Wichajam, R., Chowpreecha, S., Chareonsiriwatana, W., 2005. Stability of genetic DNA in dried blood spots stored on filter paper. Southeast Asian J Trop Med Public Health 36 (1), 270-273.

Clapham, P.R., McKnight, A., 2002. Cell surface receptors, virus entry and tropism of primate lentiviruses. Journal of General Virology 83, 1809–1829.

Descours, B., Avettand-Fenoel, V., Blanc, C., Samri, A., Me'lard A., Supervie, V., Theodorou, I., Caecelain, G., Rouzioux, C., Autran, B., 2012. Immune Responses Driven by Protective Human Leukocyte Antigen Alleles From Long-term Nonprogressors Are Associated With Low HIV Reservoir in Central Memory CD4 T Cells. HIV/AIDS 54, 1495-1503.

De Vries, J.J.C., Claas, E.C.J., Kroes, A.C.M., Vossen, A.C.T.M., 2009. Evaluation of DNA extraction methods for dried blood spots in the diagnosis of congenital cytomegalovirus infection. Journal of Clinical Virology, 1-6.

Drossman, H., Luckey, J.A., Kostichka, A.J., D'Cunha, J., Smith, L.M., 1990. High-speed Separations of DNA Sequencing Reactions by Capillary Electrophoresis. Analytical Chemistry 62 (9), 900-903.

Engelman, A., Cherepanov, P., 2012. The structural biology of HIV-1: mechanistic and therapeutic insights. Nat Rev Microbiol 10(4), 279–290.

Fellay, J., 2009. Host genetics influences on HIV type-1 disease. Antiviral Therapy 14, 1-8.

Fiorentini, S., Marini, E., Caracciolo, S., Caruso, A., 2006. Functions of the HIV-1 matrix protein p17. The New Microbiologica 29, 1-10.

Fischler, B., Rodensjö, P., Nermet, A., Forsgren, M., Lwensohn-Fuchs, I., 1999. Cytomegalovirus DNA detection on Guthrie cards in patients with neonatal cholestasis. Arch Dis Fetal Ed 80, F130-F134.

Frahm, N., Adams, S., Kiepiela, P., Linde, C.H., Hewitt, H.S., Lichterfeld, M., Sango, K., Brown, N.V., Pae, E., Wurcel, A.G., Atfield, M., Feeney, M.E., Allen, T.M., Roach, T., John, M.A., Daar, E.S., Rosenburg, E., Korber, B., Marincola, F., Walker, B.D., Goulder, P.J.R., Brander, C., 2005. HLA-B63 Presents HLA-B57/B58-Restricted Cytotoxic T-Lymphocyte Epitopes and Is Associated with Low Human Immunodeficiency Virus Load. Journal of Virology 79(16), 10218- 10225.

França, L.T.C., Carrilho, E., Kist, T.B.L., 2002. A review of DNA sequencing techniques. Quarterly Reviews of Biophysics 35(2), 169–200.

Göhring, K., Dietz, K., Hartleif, S., Gerhard Jahn, G., Hamprecht, K., 2009. Influence of different extraction methods and PCR techniques on the sensitivity of HCMV-DNA detection in dried blood spot (DBS) filter cards. Journal of Clinical Virology 8, 278-281.

Goulder P.J.R., Watkins, D.I., 2008. Impact of MHC class I diversity on immune control of immunodeficiency virus replication. *Nat Rev Immunol* 8(8), 619–630.

Hamprecht, K., Steinmassl, M., Einsele, H., Jahn, G., 1998. Discordant detection of human cytomegalovirus DNA from peripheral blood mononuclear cells, granulocytes and plasma: correlation to viremia and HCMV infection. Journal of Clinical Virology 11, 125-36.

Hue, N. T., Phong, P.T., Chan, N.D.H., Hoan, N.K.H., Thuy, H.T.T. 2011. An Effieciency Human Genomic DNA Extraction from Dried Blood Spots. Procedia Environmental Sciences 8, 179-185.

Hue, N. T., Phong, P.T., Chan, N.D.H., Hoan, N.K.H., Thuy, H.T.T., Linh, N.T.T, Giang, N.D.T., 2012. Extraction of Human Genomic DNA from Dried Blood Spots and Hair Roots. International Journal of Bioscience 2(1), 21-26.

Johanessen, A., 2010. Dried blood spots in HIV monitoring: applications in resource-limited settings. Bio analysis 2(11), 1893-1908.

Karger, B.L., Guttman, A., 2009. DNA sequencing by capillary electrophoresis. Electrophoresis 30(Suup1), 1-11.

Kassutto, S., Rosenberg, E.S., 2004. Primary HIV Type 1 Infection. Clinical Infectious Diseases 38,1447-1453.

Klein, J., Sato, A., 2000. The HLA System (Second of Two Parts). New England Journal of Medicine 11, 782-786

Kulpa, D.A., Collins, K.L., 2011. The emerging role of HLA-C in HIV-1 infection. Immunology 134, 116-22.

Lakshmy, R., 2008. Analysis of the Use of Dried Blood Spot Measurements in Disease Screening. Journal of Diabetes Science and Technology 2(2), 242-243.

Leslie, A., Matthews, P.C., Listgarten, J., Carlson, J.M., Kadie, C., Ndung'u, T., Brander, C., Coovadia, H., Bruce D. Walker, B.B., Heckerman, D., Goulder, P.J.R., 2010. Additive Contribution of HLA Class I Alleles in the Immune Control of HIV-1 Infection. Journal of Virology 84 (19), 9879–9888.

Lin, Z., Suzow, J.G., Fontaine, J.M., Naylor, E.W., 2005. A simple Automated DNA Extraction Method for Dried Blood Specimens Collected on Filer Paper. Journal of Laboratory Automation 10, 310- 314.

López-Nevot, M.A., Esteban, F., Ferrón, A., Gutiérrez, J., Oliva, M.R., Romero, C., Huelin, C., Ruiz-Cabello, F., Garrido, F., 1989. HLA class I gene expression on human primary tumours and autologous metastases: demonstration of selective losses of HLA antigens on colorectal, gastric and laryngeal carcinomas. Br. J. Cancer 59 (2), 221-226.

Makalowski, W., 2001. The human genome structure and organization. Acta Biochimica Polonica 48(3), 587-598.

Manasa, J., Danaviah, S., Pillay, S., Padayachee, P., Mthiyane, H., Mkhize, C., Lessells, R.J., Seebregts, C., Rinke de Wit, T.F., Viljoen, J., Katzenstein, D., De Oliveira, T., 2014. An Affordable HIV-1 Drug Resistance Monitoring Method for Resource Limited Settings. Journal of Visualized Experiments, 1-49.

McDade, T.W., Williams, S., Snodgrass, J., 2007. What a drop can do: Dried blood spots as a minimally invasive method for integrating biomarkers into population-based research. Demography 44(4), 899-925.

McLaren, P.J., Coulonges, C., Bartha, I., Lenz, T.L., Deutsch, A.J, Bashirova, A., Buchbinder, S., Carrington, M.N., Cossarizza, A., Dalmau, J., De Luca, A., Goedert, J.J., Gurdasani, D, Haas, D.W., Herbeck, J.T., Johnson, E.O., Kirk, G.D., Lambotte, O., Luo, M., Mallal, S., van Manen, D., Martinez-Picado, J., Meyer, L., Mirohh, J.M., Mullins, J.I., Obel, N., Poli, G., Sandhu, M.S., Schuitemaker, H., Shea, P.R., Theodoroud, I., Walkeri, B.D, Weintrob, A.C., Winkler, C.A., Wolinsky, S.M., Raychaudhurig, S., Goldstein, D.B., Telenti, A., de Bakker, P.I.W., Zagury, J., Fellay, J., 2015. Polymorphisms of large effect explain the majority of the host genetic contribution to variation of HIV-1 viral load. PNAS 112(47), 14658–14663.

McNulty, A., Jennings, C., Bennett, D., Fitzgibbon, J., Bremer, J.W., Ussery, M., Kalish, M.L., Heneine, W., Garcı´a-Lerma, J.G., 2007. Evaluation of Dried Blood Spots for Human Immunodeficiency Virus Type 1 Drug Resistance Testing. Journal of Clinical Microbiology 45(2), 517-521.

Mehra, N.K., 2001. Histocompatibility Antigens. Encyclopedia of Life Sciences, 1-6.

Morandi, F., Fainardi, E., Rizzo, R., Rouas-Freiss, N., 2014. The Role of HLA-Class 1B Molecules in Immune-Related Diseases, Tumors, and Infections. Journal of Immunology Research 2014, 1-2.

O'Brien, S. J., Nelson, G.W., 2004. Human genes that limit AIDS. Nature Genetics 36(6), 565-574.

Pachot, A., Barbalat, V., Marotte, H., Diasparra, J., Gouraud, A., Mougin, B., Miossec, P., 2007. A rapid automated method for DNA extraction from dried-blood spots: Application to the HLA-DR shared epitope analysis in rheumatoid arthritis. Journal of Immunological Methods 328, 220-225.

Rajatileka, S., Luyt, K., El-Bokle, M., Williams, M., Kemp, H., Molnár, E., Váradi, A., 2013. Isolation of human genomic DNA for genetic analysis from premature neonates: a comparison between newborn dried blood spots, whole blood and umbilical cord tissue. BMC Genetics 14(105), 1-9.

Ronaghi, M., 2001. Pyrosequencing Sheds Light on DNA Sequencing. Genome Research11, 3-11.

Sairkar, P., Chouhan, S., Batav, N., Sharma, R., 2013. Optimization of DNA isolation process and enhancement of RAPD PCR for low quality genomic DNA of Terminalia arjuna. Journal of Genetic Engineering and Biotechnology 11, 17–24.

Sanger, F., Nicklen, S., Coulson, A. R., 1977. DNA sequencing with chain-terminating inhibitors. Biochemistry 74(12), 5463-5467.

Shibata, M., Takano,H., Hironaka,T., Hirai, K., 1994. Detection of human cytomegalovirus DNA in dried newborn blood filter paper. Journal of Virology Methods 46, 285–97.

Schrader, C., Schielke, A., Ellerbroek, L., Johne, R., 2012. PCR inhibitors – occurrence, properties and removal. Journal of Applied Microbiology 113, 1014-1026.

Sirdah, M.M., 2014. Superparamagnetic-bead Based Method: An Effective DNA Extraction from Dried Blood Spots (DBS) for Diagnostic PCR. Journal of Clinical and Diagnostic Research, 8(4), 1-4.

Smith, H.O., Wilcox, K.W., 1970. A restriction enzyme from Hemophilus influenza: I. Purification and general properties. Journal of Molecular Biology 51, 379-391.

Smith, L.M., Sanders, J.Z., Kaiser, R.J., Hughes, P., Dodd, C., Connell, C.R., Heiner, C., Kent, S.B.H., Hood, L.E., 1986. Fluorescence detection in automated DNA sequence analysis. Nature 321, 674-679.

Stephens, H.A.F., 2005. HIV-1 diversity versus HLA class I polymorphism. TRENDS in Immunology 26 (1), 41-47.

St. Julien, K.R., Jelliffe-Pawlowski, L.L., Shaw, G.M., Stevensons, D.K., O'Brodovich, H.M., Krasnow, M.A., 2013. High quality genome-wide genotyping from archived dried blood spots without DNA amplification. PLOS ONE 8(5), 1-7.

Telenti, A., Johnson, W.E., 2012. Host Genes Important to HIV Replication and Evolution. Cold Spring Harbor in Perspective Medicine 2, 1-21.

Wong, N.C., Moorley, R., Saffrey, R., Craig, J.M., 2008. Archived Guthrine blood spots as a novel quantitative DNA methylation analysis. Bio Techniques 45, 423-430.

Zhou, H., Hickford, J.G.H., Fang, Q., 2006. A two-step procedure for extracting genomic DNA from dried blood spots on filter paper for polymerase chain reaction amplification. Analytical Biochemistry 354, 159-161.

Websites

http://www.123rf.com/photo_12476223_structure-of-hiv.html, “date accessed 13/12/2015”

http://neurowiki2012.wikispaces.com/Theories+of+Origin+and+Risk+Factors#cite_note-4, “date accessed 13/12/2015”

http://www.biomerieux-diagnostics.com/hiv-diagnostics-management, “date accessed 13/12/2015”

http://www.seq.ubc.ca/wp-content/uploads/2006/08/sequencing2.gif, “date accessed 13/12/2015”