Development, Validation and Applications of a Novel Multiplex Assay RM-Yplex Amplifying 13 Rapidly Mutating Y Chromosome Short Tandem Repeat Regions

Rashed Al Ghafri, MSc

December 2014

A thesis submitted to the University of Central Lancashire in partial fulfilment of the requirements for the degree of Doctor of Philosophy

DECLARATION

I declare that no material contained in the thesis has been used in any other submission for an academic award and is solely my own work. I declare that the thesis contains no materials previously published or written by another person except where due reference is made.

Date: ……

Signature: ……

ABSTRACT

A polymerase chain reaction (PCR) multiplex assay capable of amplifying 13 rapidly mutating Y chromosome short tandem repeats (RM Y-STRs) simultaneously was developed and optimised. This multiplex assay which was termed RM-Yplex is the first to include all 13 RM Y-STRs including DYF387S1, DYF399S1, DYF403S1a/b, DYF404S1, DYS449, DYS518, DYS526a/b DYS547, DYS570, DYS576, DYS612, DYS626 and DYS627. A developmental validation was performed following the Scientific Working Group for DNA Analysis Methods (SWGDAM) revised guidelines. Robustness and limitations of the assay were demonstrated through a range of studies including reproducibility, sensitivity, specificity, stability and mixture studies. Appropriate controls were used during the studies that included a number of male and female commercial controls including, 2800M, 9948 and Taqman male controls and 9947A female control. An allelic ladder was developed for the assignment of the alleles. This was done by choosing samples with different alleles, amplifying them and then adjusting the volumes of amplified products in a mixture. The developed mixtures were used to balance the composite ladder. Multiple alleles of the various loci included in the ladder were sequenced. Reference haplotypes were developed for the 5 male samples included in the Y chromosome Standard Reference Material 2395 (SRM2395) using RM-Yplex. The International Society of Forensic Genetics (ISFG) recommendations were followed for adopting allele nomenclature. As part of developmental validation, the assay was included in an external proficiency trial which was concluded successfully. An internal validation of RM-Yplex was carried out at the Department of Forensic Sciences and Criminology Laboratory, Dubai where apart from other studies; application of the assay was demonstrated using non-probative forensic casework samples. The value of RM-Yplex was demonstrated for differentiating close male relatives in a case where a

previously used Y-STR multiplex assay had shown identical haplotypes for those individuals. 1160 male individual samples were analysed in this study including UAE, other Arabian Peninsula populations as well as two South Asian populations residing in United Arab Emirates. RM-Yplex haplotypes have extremely high power of discrimination. The haplotype diversity for RM-Yplex haplotype is much more than the existing commercial Y-STR assays. Population studies have been carried out for the Arab, Indian and Pakistani populations. AMOVA was conducted for determining the apportionment of diversity and pairwise F_{ST}'s were estimated between populations. These have shown a marked homogeneity within the UAE Arab sub-populations. MDS plots of pairwise F_{ST}'s indicated that populations were not grouped significantly in accordance with the geographical locations. A network analysis showed the extent of distribution of haplotypes of various populations and their relationships. A highly sensitive and reliable RM-Yplex multiplex assay has been thus developed, which is expected to help genetic populations studies and forensic casework.

CONTENT

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENT

First and foremost I would like to thank God who have given me the strength to believe in myself and pursue my dreams, whom his mercy was upon me through my life and even more during this study and whom without I would have never done this research.

My wholehearted thanks goes to my parents who have always given me the strength and wisdom to be sincere in my research, for setting high moral standards, for giving me their unconditional trust and for their endless patience. In addition, I would like to thank all other members of my family.

I would like to humbly thank the Crown Prince of Dubai, His Highness Sheikh Hamdan Bin Mohammed Bin Rashed Al-Maktoum, for enabling Dubai Government grant support for this project.

I would like to thank United Arab Emirates Government in general and specifically Dubai Government where I have gotten most of the support for this project especially Dubai Police General Head Quarters, Dubai Health Authority and Dubai Municipality.

Deep appreciation goes to my director of study, Dr. Sibte Hadi, who have dedicated his time to support me 24/7 and who treated me as his younger brother rather than just a student. I would like to thank him for his, patient, and belief in me and of course teaching me.

Finally, I would like to thank all staff working at Department of Forensic Sciences and Criminology, members of Forensic Genetic Group at University of Central Lancashire and all students I have come to know since the beginning of my PhD journey for their endless support through my PhD.

LIST OF ABBREVIATIONS

DNA	Deoxyribonuceic Acid
STR	Short Tandem Repeat
Y-STR	Y chromosome STR
RM Y-STR	Rapidly Mutating Y-STR
PAR	Pseudo-Autosomal Region
NRY	Non-recombining Y chromosome Region
ISFG	International Society of Forensic Genetics
SWGDAM	Scientific Working Group on DNA Analysis Methods
NIST	National Institute of Standards and Technology
EDNAP	European DNA profiling group
DFSC	Depaertment of Forensic Sciencecs and Criminology
SRM	Standard Reference Materials
PCR	Polymerase Chain Reaction
MgCl2	Magnisium Chloride
dNTPs	deoxy Nucleotide TriPhosphate
EDTA	Ethylene Diamine Tetra Acetic Di-Sodium Salt
RFU	Relative Fluorescent Units
LOD	Limit of Detection
LOQ	Limit of Quantitation
DC	Discrimination Capacity
HD	Haplotype Diversity
HGDP	Human Genome Diversity Project
CEPH	Human Polymorphism Study Center

1. Introduction

1.1. Overview of Forensic DNA Analysis

Forensic science is the practice of scientific analysis of evidentiary materials leading to the development of justified explanations and presenting these to the courts to help them make appropriate decisions (Broeders, 2006). Forensic DNA analysis involves the analysis of the genetic material in biological materials such as blood, saliva, semen, muscles and bones. When biological materials are found at crime scenes, a forensic DNA analyst might be required to develop a DNA profile from them and compare this with a potential suspect's reference profile (Patzelt, 2004). The aim of forensic DNA analysis is to assess the biological evidence and link it to the probable source (Inman and Rudin, 2001, Saks and Koehler, 2008).

1.2. History of Biological Markers

Forensic biological analysis for human identification was limited to antigen based markers like the ABO, Rhesus, Gc and Km blood grouping systems in early part of the 20th century (Landsteiner, 1900, Landsteiner and Wiener, 1940). These markers had limited variability among individuals and hence the matching probability between random individuals was always high. In addition, these markers were not always able to withstand the environmental conditions to which such stains are usually subjected to (Budimlija et al., 2003). In 1955, the discovery of the proteins and enzymes markers had a significant improvement in forensic serology as they showed higher degree of polymorphism than the blood group systems (Smithies, 1955). Although several of these markers became available for forensic purposes, ubiquitous, more stable and more variable systems were required for human identification. DNA based markers have largely overcome the limitations and challenges of using blood antigen and enzyme systems owing to their nature and distribution.

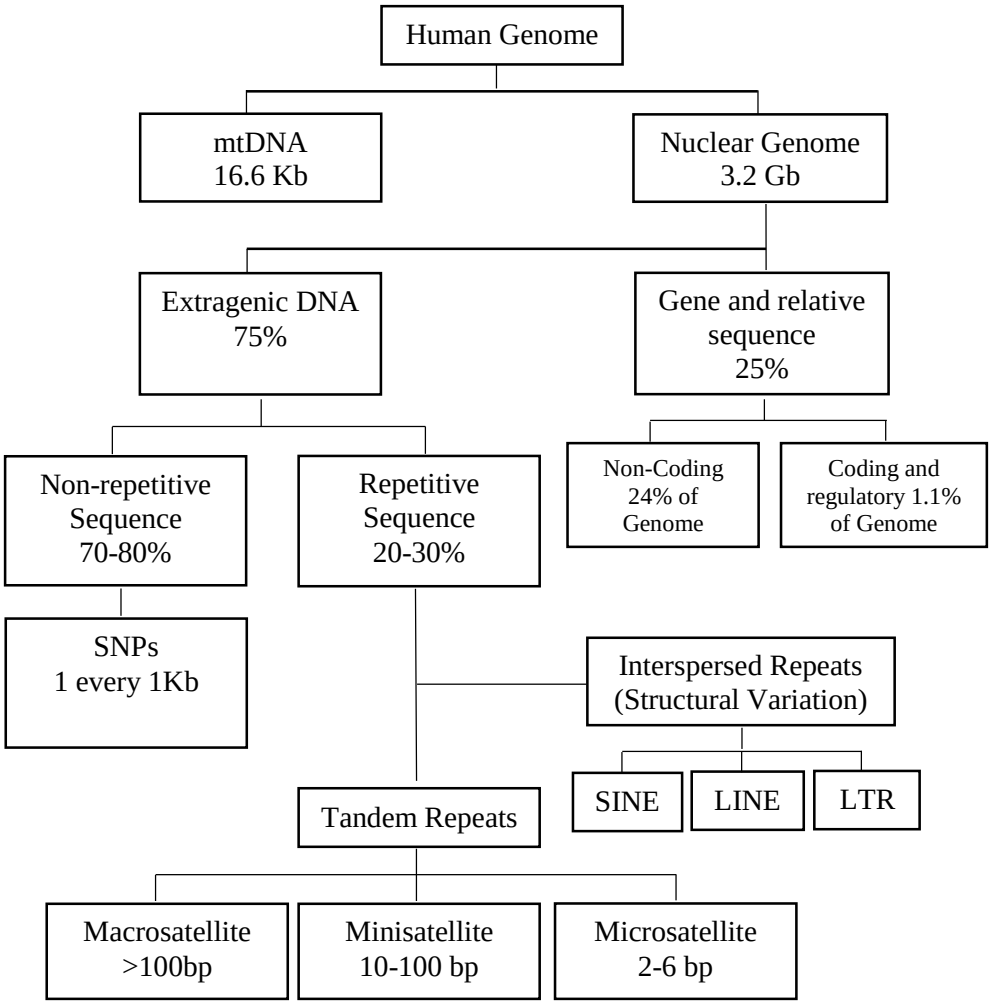

Figure 1.1: Descriptive chart of human genome structure adapted from Kashyap et al. (2004).

1.3. Human Genome

There are billions of cells in the human body. The nucleus of each cell contain nuclear deoxyribonucleic acid (DNA), this DNA is composed of 3 billion base pairs in a haploid cell (Li et al., 2007). DNA was first isolated by Friedrich Miescher in 1895 (reviewed in Dahm, 2005). DNA structure, as first identified by Watson and Crick (1953), is comprised of four chemical bases adenine (A), guanine (G), cytosine (C) and thymine (T). The nuclear human genome consists of nuclear DNA as 46 chromosomes (23 homogenous pairs). Normal body cells are therefore diploid. The sex cells however are haploid and a

human individual inherits 23 chromosomes from each parent. At fertilization the diploid number of chromosomes is restored. Generally, humans display 99.9% homology in DNA. The remaining 0.1% represents the genetic differences among individuals and hence every individual is genetically unique (Ullrich et al., 1980, Cooper et al., 1985, Pena et al., 1995). Although primary findings showed that monozygotic twins have exactly same DNA sequence (Miyake et al., 2013), recently, using a deep sequencing approach, the ability to distinguish monozygotic twins was successfully demonstrated (Weber-Lehmann et al., 2014). A smaller part of the genome consists of mitochondrial DNA, which is transmitted to all individuals, maternally.

1.3.1. Gene and Relative Sequences

The coding and non-coding DNA sequences comprise 25% of the human genome. Non-coding region within gene sequences contain pseudogenes and untranslated DNA sequences, which form approximately 24% of the whole genome. Coding regions of DNA include DNA sequences, which are transcribed for proteins synthesis, as well as regulatory elements which assist the process of protein synthesis such as enhancers, promoters and TATA boxes, and represent 1.1% of human genome (Strachan, 2004) (Figure 1.1). The regions of DNA coding for proteins are called genes. The genes in human cells as well as in all eukaryotes cells are made up of exons, introns and regulatory elements. The introns are excised during protein synthesise process. The exons are joined together during protein process and encode for messenger RNA (mRNA) (Strachan, 2004). DNA therefore holds the blue print for all proteins synthesized in the body (Figure 1.2). Some of these proteins are responsible for all biochemical processes in the cell as well as the phenotypic traits of an individual including, hair colour, eye colour and skin colour (Baltimore, 2001).

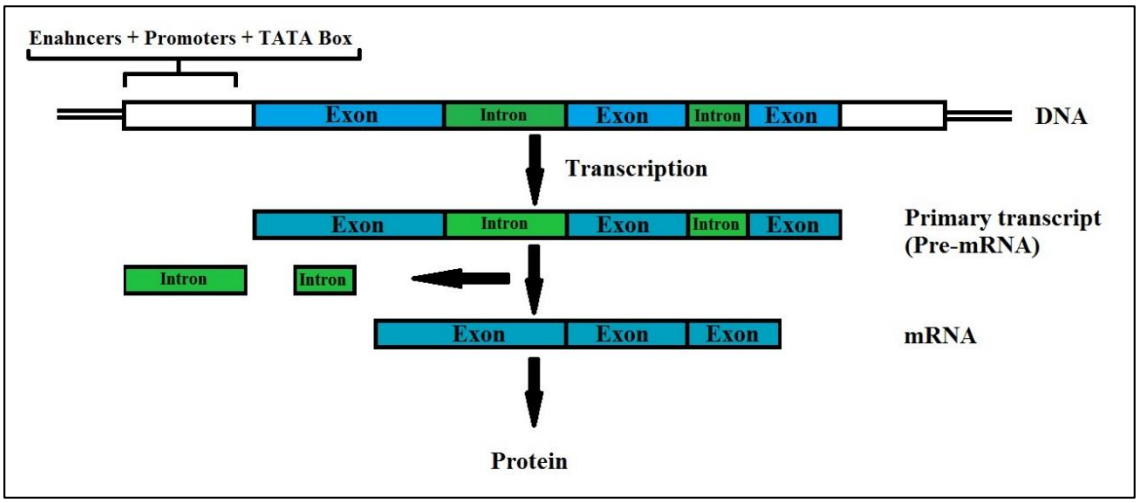

Figure 1.2: The description of the protein synthesis process showing the composition of gene DNA sequence and mRNA.

1.3.2. Extragenic DNA

The majority of the genetic differences are found within the non-coding region of DNA which represents 99% of the human genome (Sachidanandam et al., 2001, Venter et al., 2001, Hattori, 2005). Most of the non-coding regions of DNA sequences in human genome are called extragenic DNA. This region of DNA contains non-repetitive sequences and also repetitive sequences. Repetitive sequences, which form 25% to 30% of the extragenic DNA sequences, are of particular use in forensic DNA profiling (Batzer and Deininger, 2002). Such repetitive sequences are classified into two categories: interspersed repetitive sequences and tandem repeats sequences. Interspersed repetitive sequences represent the majority of the human DNA and are of two main types; the long interspersed nucleotide elements (LINEs) and the short interspersed nucleotide elements (SINEs). The high level of genetic variation in DNA sequence in these sequences are due to many factors such as insertion, deletion, duplications and other alterations that accumulate over time in successive generations of individuals (Batzer and Deininger, 2002). The tandemly repeating DNA sequences are highly variable among human individuals and have become the focus of DNA forensic profiling. Three main sub-classes of tandem repeats are macrosatellites, minisatellites and microsatellites. The latter have

been found to be the most useful for forensic DNA profiling mainly due to the size of the fragments, the absence of linkage with other microsatellite markers, and high variability among individuals (Butler, 2007).

1.3.3. DNA Polymorphisms

There are many types of polymorphic loci spread through the non-coding regions of the DNA. DNA polymorphism could be divided into length and sequence polymorphisms (Figure 1.3) (Kashyap et al., 2004). The three best known DNA length polymorphisms are restriction fragment length polymorphisms (RFLP), variable number of tandem repeats (VNTR) and short tandem repeats (STR) whereas DNA sequence polymorphisms are mtDNA sequences and single nucleotide polymorphisms (SNPs) (Lee et al., 1994).

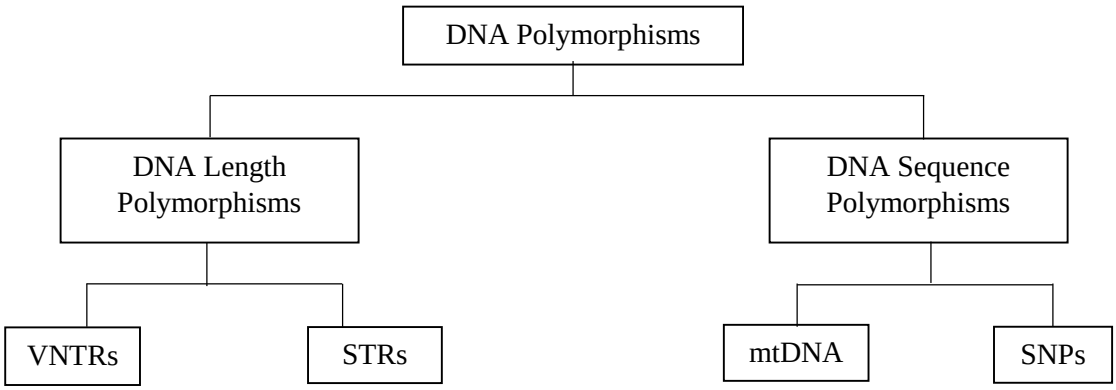

Figure 1.3: Classification of DNA polymorphisms in the non-coding region of the human genome.

1.3.3.1. Restriction Fragment Length Polymorphisms (RFLP)

In early part of 1980, RFLPs were the first type of DNA polymorphisms applied in forensic DNA analysis. It was helpful in linking evidence found in the crime scene to suspects and thus serving justice. However, the application was limited by the need of large amount of DNA being labour intensive and due to the complexity of the analyses (Lee et al., 1994).

1.3.3.2. Variable Number of Tandem Repeats (VNTR)

VNTRs are composed of tandemly repeating arrays of sequences consisting of 10 to 100 nucleotides (Botstein et al., 1980, Nakamura et al., 1989). As the number of repeats are highly variable among individuals, this technique ushered an era of forensic DNA profiling and was extensively used since its discovery in 1980s (Jeffreys et al., 1985). Although this technique was a great advancement, it had limitations similar to the RFLP technique (Goes et al., 2002).

1.3.3.3. Short Tandem Repeats (STR)

STRs were first reported as repetitive regions comprising of a repeat motif that was only 2-6 base pairs (bp) long (Litt and Luty, 1989). It was estimated that these occur once in every 10000 base pairs within human genome and were distributed over all the chromosomes (Edwards et al., 1991). An STR position on a chromosome is called a locus and as every individual inherits 23 chromosome from each of the parents, there would be two copies of each STR, called alleles, for each locus located on the autosomal chromosomes in each individual. An individual is termed a homozygote for a particular locus if the same alleles were inherited from each of the parents, and is a heterozygote if different alleles were inherited from each of the parents (Goodwin et al., 2011). An example of STR marker currently used in forensic analysis is D5S818 (Figure 1.4) (Butler and Hill, 2012). There are different types of STR, which are classified according to their repeat unit size and complexity.

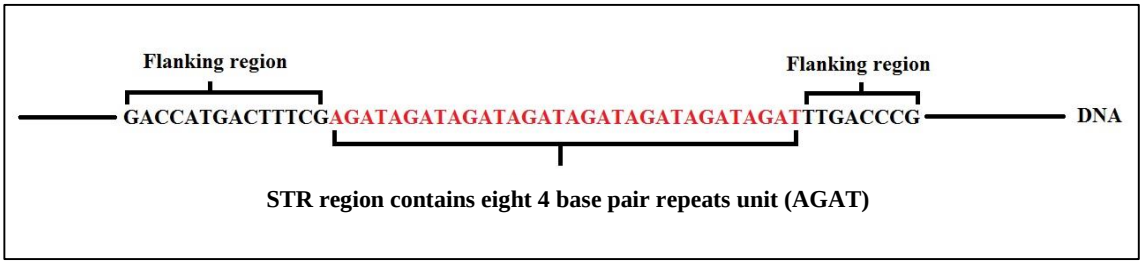

Figure 1.4: Illustration diagram of allele 8 at D5S818 STR marker containing AGAT repeat units.

1.3.3.3.1. Repeat Unit Sizes Classification

Depending on the size of the repeat unit, STRs are classified into di-, tri-, tetra, penta- and hexanucleotide (Table 1.1). The usefulness of each type of these STRs for DNA profiling is determined according to the level of variation between individuals and the unambiguous identification of an allele. Such variation is represented by the number of repeats found at the locus and hence alleles are designated according to the number of repeats detected. The presence of amplification artefacts like a stutter can hamper the determination of an allele. A stutter is mostly one repeat smaller polymerase chain reaction (PCR) product and this can cause difficulty in allele determination. It has been shown that longer repeat unit STRs like tetranucleotides and pentanucleotides are less prone to stuttering (Brookes et al., 2012). Therefore, tetranucleotides and a few pentanucleotides have become standard forensic loci.

Table 1.1: STR classifications according to the repeat unit size.

Repeat Unit Size of STR	**STR Sequence**
Dinucleotide $(TA)_{10}$	---TA TA TA TA TA TA TA TA TA TA---
Trinucleotide $(CTC)_8$	---CTC CTC CTC CTC CTC CTC CTC CTC---
Tetranucleotide $(GAAG)_5$	---GAAG GAAG GAAG GAAG GAAG---
Pentanucleotide $(CATTG)_5$	---CATTG CATTG CATTG CATTG CATTG---
Hexanucleotide $(GGATCC)_4$	---GGATCC GGATCC GGATCC GGATCC---

1.3.3.3.2. STR Complexity Classification

STRs are classified as simple, compound and hyper-variable/complex STR loci (Urquhart et al., 1994, Gill and Evett, 1995). Simple STR loci consist of alleles with identical length and sequence of repeat unit such as the autosomal locus D5S818 (Figure 1.4); however, there could also be simple STR loci with non-consensus repeat length as HUMTHO1 locus. Compound STR loci consist of two or more adjacent repeat motifs with different sequences or compound structure with non-consensus alleles (Urquhart et al., 1994). The

hyper/variable complex repeats comprise of non-consensus alleles having differences in both size and sequence of the repeat motifs, such as D21S11 and SE33, which were described as the most complex STR markers among the STRs used in forensic DNA profiling (Gill et al., 1997, Rolf et al., 1997, Butler, 2006) (Table 1.2).

Table 1.2 Summary and examples of STR categories as defined by Internal Society of Forensic Genetics (ISFG).

STR repeat category		Example locus	Repeat Structure	Reference
Simple STR	**consensus**	HUMFES/FPS	$(ATTT)_{8–1}$	(Polymeropoulos et al., 1991a)
	non-consensus	HUMTH01	$(TCAT)_4CAT(TCAT)_5$	(Polymeropoulos et al., 1991b)
Compound STR		HUMVWFA31	$(TCTA)(TCTG)_{3–4}(TCTA)_{8,10–17}$	(Kimpton et al., 1992, Moller et al., 1994, Barber et al., 1995)
Hyper variable/Complex STR		D21S11	$(TCTA)_{4–6}(TCTG)_{5-6}(TCTA)_3TA(TCTA)_3TCA$ $(TCTA)_2TCCATA(TCTA)_{8–16}$	(Sharma and Litt, 1992)
		SE33	$(AAAG)_2$ AG $(AAAG)_3$ AG $(AAAG)_{7-22}$ AA AAAG $(AAAG)_9$ G AAGG $(AAAG)_2$ AG	(Rolf et al., 1997)

1.3.3.3.3. Microvariant Alleles

The term "microvariant" is used to describe an allele of STR locus which consists of incomplete number of repeat motif. For instance, allele 9.3 at TH01 locus, illustrated in Figure 1.5, comprises nine tetranuceotide repeats and one incomplete repeat comprising of three nucleotides (Puers et al., 1993).

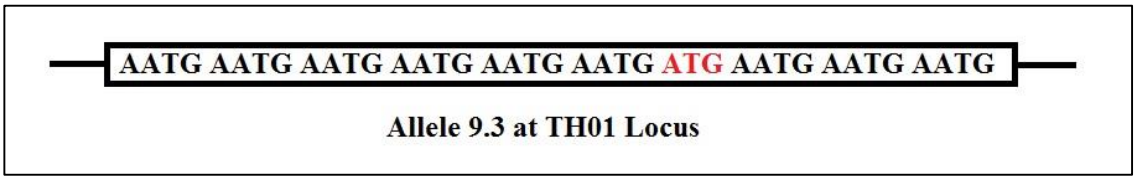

Figure 1.5: An example of microvariant allele 9.3 at TH01 simple repeat STR locus.

In the last two decades, the identification and the analysis of hundreds of STR loci had a great impact in human identification for forensic DNA analysis (Moretti et al., 2001). Hence, after extensive investigations of these markers, the utility and validity of STR typing for forensic applications such as human identity testing has been proven (Lygo et al., 1994, Kimpton et al., 1996, Budowle et al., 1997, Frégeau et al., 1999). Presently a DNA profile of an individual consisting of 16-24 STR loci can be developed from minute quantities of DNA from different types of forensic samples depending on the commercial multiplex kit used. Such a profile can result in a random match probability of over 1 in a trillion (Butler et al., 2012, Flores et al., 2014, Fujii et al., 2014).

1.4. STR Detection

1.4.1. Polymerase Chain Reaction (PCR)

Development of PCR has enabled the forensic scientists to detect DNA polymorphisms from minute amounts of DNA. The polymerase chain reaction is an enzyme based reaction by which replication of specific region of DNA yields large number of copies of the specific region (Saiki et al., 1985, Mullis et al., 1986). The process of replicating DNA is called amplification. PCR is highly sensitive and only few copies of DNA are needed to amplify DNA polymorphisms. This feature has made this method widely used for forensic applications (Schneider et al., 2004, Kline et al., 2005). There are several essential components needed for a successful PCR reaction including DNA polymerase enzyme, PCR buffer, PCR primer pair, dNTPs, $MgCl_2$ and DNA template. The quality of the primers has a great effect on the amplified products i.e. amplicons. In order for the PCR to work efficiently, both forward and reverse primers should be specific to the targeted region and both should have approximately the same annealing temperature (T_m). The PCR primer binding regions sequences should be conserved in order to produce desired amplicons, otherwise the reaction might show no results and “null alleles” may result for a specific locus (Saiki et al., 1988). A high fidelity preferably hot start enzyme

is needed for a PCR since the reaction is driven by the polymerase enzyme. Therefore, TaqGold enzyme has been included in most commercial STR assays (Butler, 2005).

1.4.2. Multiplex PCR

Multiplex PCR is used to describe the simultaneous amplification of multiple targeted DNA templates by adding several pairs of primers in one reaction. The concept of multiplex PCR was first described and applied by Chamberlain et al. in 1988 for medical genetics diagnostic purposes. Since the discovery of the PCR by Saiki et al. in 1988, the technique of multiplex PCR has been applied in different studies like the analysis of deletions in human DNA, STR mutation rate study, autosomal STR and Y chromosome STR analysis (Puers et al., 1993, Shuber et al., 1993, Henegariu et al., 1994, Kimpton et al., 1996, Prinz et al., 1997, Redd et al., 1997, Gusmaúo et al., 1999, Prinz and Sansone, 2001, Butler, 2005). Moreover, the wide accessibility of genetic information after the publication of the human genome sequence made the technique of PCR multiplexing more widely applied (Sachidanandam et al., 2001). Amplicons resulting from multiplex PCR amplification are different in sizes and can be differentiated using an appropriate technique such as electrophoresis (Kuhr, 1990). Initially, agarose gel electrophoresis (AGE) method based on UV absorbance was used in separating PCR amplicons according to their sizes, however this method had low resolution in separating PCR amplicons (Henegariu et al., 1997). AGE could also identify alleles monochromatically inhibiting separation and identification of alleles, which were of similar size. This limitation was overcome by attaching different fluorescent labels to the 5′ end of the forward primers prior to use in PCR amplification. Different fluorescent emission could be detected from same sized amplicons facilitating the detection of similar sized PCR amplicons using both UV absorbance and fluorescent induced laser (FIL) (Altria, 1999). The latter technique has been widely adopted in different applications due to its high level of precision and accuracy compared to the UV absorption method (Altria, 1999).

It is a challenge to design a multiplex assay as multiple primer annealing events are taking place, which need to occur under the same reaction conditions with no cross-reactivity appearing between a number of primer pairs. As a result, extensive PCR optimisation is needed to obtain similar concentrations of amplicons of different loci (Henegariu et al., 1997, Schoske et al., 2003). Previous studies have focused on studying the effect of critical PCR reaction components, such as $MgCl_2$ concentration, reaction buffer concentration, annealing temperature and DNA polymerase enzyme concentration, in order to establish guidance for designing a multiplex assay (Shuber et al., 1995, Henegariu et al., 1997).

Compatibility of primer pairs is compulsory in order for the multiplex reaction to work appropriately. Hence the primers properties should have similar characteristics such as annealing temperature. In addition primers should neither be self-complimentary nor complementary to the non-desired regions of DNA template (Schoske et al., 2003). Primers of high quality should be used for optimal multiplex results; high pressure liquid chromatography (HPLC) and mass spectroscopy (MS) are used for enhancing the quality of the primers (Butler et al., 1998, Devaney et al., 2000).

1.4.3. Capillary Electrophoresis (CE)

Capillary electrophoresis has become a widely used analytical technique since it was first introduced in 1980's (Kuhr, 1990). The technique is relatively simple and the migration of the fragments of different sizes and electrical charge occurs in a fused silica capillary as voltage is applied. The capillary is filled with a medium for separating the DNA fragments called 'polymer' (Chen et al., 1994). CE has been extensively studied in forensic DNA profiling and is specifically used to separate STR fragments with high precision and reproducibility (Landers et al., 1993, McCord et al., 1993, Butler et al., 1994, Budowle and Allen, 1998). The growth of commercially available multichannel CE was noticeable in the past 10 years and three such instruments ABI PRISM® 3100, ABI

PRISM® 3130 and ABI 3500 Genetic Analysers (Applied Biosystems) have been developed and used within the forensic and research laboratories. The analysis using such instruments begins with filling of the capillary array with the polymer where different sizes of molecules are separated based on their electrophoretic mobility using a conductive buffer. Electrophoretic mobility of a fragment is almost equal to the charge-to-mass ratio of the molecule (Altria, 1999). When an electrical current is applied, the negatively charged DNA fragments migrate towards the anode with variable speeds.

These instruments detect fluorescence, i.e. photons, emitted from the fluorescence molecules attached to the DNA fragments when a laser at the end of the capillary array excites them. When photons of light emitted hit the charged-coupled device (CDD) camera, it converts them to electrical signals. The collection of signals after each run is displayed as peaks in a graphical data called electropherogram (EPG) where each peak will have a height reflecting the relative fluorescent units (RFUs) versus time (Altria, 1999, Butler, 2005). The time between sample injection and detection is measured and the size of fragment estimated accordingly. Smaller DNA fragments migrate faster than the larger DNA fragments. Amplicons of the same size are differentiated by different fluorescent dye labels.

1.5. Y Chromosome Short Tandem Repeats (Y-STRs)

1.5.1. Y Chromosome

Y chromosome is the smallest chromosome in the human genome, with an average size of 60 Mb (Buehler, 1980). Most of the Y chromosome consists of a non-recombining region of the Y chromosome (NRY). Both ends of the Y chromosome are called pseudo autosomal regions (PARs) which make up about 5% of the Y chromosome (Quintana-Murci and Fellous, 2001). The PARs consist of DNA sequences which are homologous to sequences on X chromosome and are the only region of Y chromosome where an

exchange of genetic material with the X chromosome happens during meiotic recombination (Quintana-Murci and Fellous, 2001). NRY is made up of heterochromatin "tightly packed DNA" and euchromatin "loosely packed DNA" having a size of 30 Mb and 24Mb respectively. NRY exists in a haploid state. The NRY is inherited intact through paternal lineages unchanged unless mutations have occurred (Quintana-Murci et al., 2001). Due to its paternal inheritance, NRY region is extremely useful for genetic structure studies; deficient paternity testing; identification of disasters male victims; identification of male lineages for anthropology purposes and the identification of male perpetrators in sexual assault criminal cases (Kayser et al., 1997a, Underhill et al., 2000, Hammer et al., 2001, Oota et al., 2001, Jobling and Tyler-Smith, 2003, Roewer et al., 2005, Shi et al., 2010).

1.5.2. Y Chromosome Markers Mutations Rates

The inheritance of the intact NRY region from one generation to another suggests that similarities within the NRY regions between males individuals reflect shared ancestry among individuals, whereas differences within the NRY regions between males individuals indicate the absence of close paternal lineages (Ballantyne et al., 2014). However, such a conclusion will depend on the mutation rate of the markers used for the comparison. Hence, it is important to consider the mutation rate of the markers being used in order to reach a proper conclusion for the analysis (Fan and Chu, 2007).

For STR markers, the main mechanisms of mutation are; the slippage of the DNA polymerase enzyme during replication of the DNA and the unequal crossover during recombination, which is mainly due to the composition of the repetitive markers structure (Schlotterer and Tautz, 1992, Shriver et al., 1993). Compared to other DNA sequences, repetitive DNA sequences are more prone to dissociation (Kuhner et al., 2007). Because of its repetitive structure, there will be a possibility of misaligning event occurring for the dissociated strands of DNA during the process of hybridisation. This results in secondary

structures such as loops, hairpins or triplexes. Therefore, an additional repeat will be synthesised or the repeat unit will be skipped depending on the mismatch location (Shriver et al., 1993, Pearson and Sinden, 1998). During DNA replication there are two ways of recognising and repairing errors, these are the proofreading capability of exonucleases involved in the nascent strand synthesis and the DNA mismatch repair system. It has been shown that proofreading is ineffective in recognising repeat loss or addition, hence leaving the mismatch system only to correct any errors within STRs (Schlotterer and Tautz, 1992). The most likely explanation for the proofreading exonuclease deficiency in repetitive DNA sequences is repeat displacement existence of the mismatch by the 3′ terminus of replicated DNA strand, which eventually will not be recognised by proofreading exonuclease (Kroutil et al., 1996, Sia et al., 1997). The process of DNA mismatch repair involves a group of proteins which are essential to start the repairing process of DNA including MutSα, MutLα, Exo1, RPA, PARP-1, PCNA, RFC and DNA polymerase δ (Harfe and Jinks-Robertson, 2000). There are two ways of mismatch repairing processes. The simplest process is the process which is directed by a break in 5' strand. This process will begin with an activation of an exonuclease (Exo1) by MutSα and MutLα proteins. Then, RPA and PARP-1 proteins will bind to the stabilised and broken DNA strands respectively, where the latter will signal the broken strand. After that, in the presence of RPA, PCNA and RFC DNA polymerase δ will start strand displacement (Harfe and Jinks-Robertson, 2000). The second way is more complicated which starts the process of DNA displacement after endonuclease activation by skipping hydrolysis of the broken strand. In this type of mismatch repair the process may start at either 5′ or 3′ strand break (Harfe and Jinks-Robertson, 2000, Li, 2008). Due to the complexity involved in the pathway of DNA mismatch repair process as well as the fact that this is the only repairing process of DNA available for errors involved in DNA replication at STR regions, it is expected that such genomic areas will have high

mutation rates. It was also recognised that defects in DNA mismatch repair process will increase the spontaneous mutation rates (Tiraby and Fox, 1973).

The mutation rates of STR markers have been estimated to be between 10^{-5} and 10^{-2} per marker per generation (Edwards et al., 1992, Weber and Wong, 1993, Brinkmann et al., 1998). Currently autosomal STRs used in the forensic profiling have mutation rates between 1×10^{-3} and 6.4×10^{-3} as estimated using extracted DNA from proven family members samples (*Short Tandem Repeats Internet Database*, 2014). There are many factors affecting the mutation rate of STR markers, for example, the male germ line records more mutations than the female germ line which is most likely due to the higher number of mitotic divisions, which sperm cells are going through, compared to oocytes (Weber and Wong, 1993, Brinkmann et al., 1998). Also, it has been shown that there is a correlation between repeat unit length and mutation rates; for example, dinucleotide and trinucleotide repeat markers have shown higher mutation rates than tetranucleotide repeat markers and consequently will have higher allelic variability (Kruglyak et al., 1998, Sup Lee et al., 1999). However, in complex STR markers, it has been shown that the repeat number of the longest homogenous repeat strand has the highest effect on mutation rate (Klintschar and Wiegand, 2003), whereas, in a recent study it has been shown that the total number of repeats in the complex STR markers has more influence on mutation rate (Ballantyne et al., 2010). Another correlation study conducted between number of repeats and mutations events where model organisms were used and analysed parallels human STR mutation rate data published previously. It was realised that the mutation rate increases when the number of repeats increase, this was also noted in other organisms (Seyfert et al., 2008).

1.5.3. Y-STRs Applications

1.5.3.1. Population Genetics

The genetic markers included in the NRY can be divided into two categories depending on their mutation rates. The first category of markers with very low mutation rates such as SNPs are valuable in tracing and comparing male lineages haplogroups (Underhill et al., 2000, Van Oven et al., 2014). Mutation rates of Y-SNPs have been estimated to be 3 x 10^{-8} per nucleotide per generation (Xue et al., 2009, Poznik et al., 2013). The second category including markers with high mutation rates such as Y-STRs, have an average mutation rate of about 10^{-3} per locus per generation (Goedbloed et al., 2009, Ballantyne et al., 2010). These markers have proven to be useful for forensic applications in ascertaining the probability of origin of the sample and for paternal kinship patterns (Kayser et al., 2003, Kayser et al., 2005, Roewer et al., 2005, Coble et al., 2009, Van Oven et al., 2010, Van Oven et al., 2014).

1.5.3.2. Forensic Applications

It has been determined that Y-STR loci are as polymorphic as autosomal STRs and can help in individual differentiation (Roewer et al., 1992). When autosomal STR markers fail to be informative in forensic investigations involving male individuals, Y-STR analysis can add to the value of DNA analysis. Y-STR analysis is mostly used in forensic DNA profiling for sexual assault cases where vaginal swabs of the victim usually contain mixed DNA from the victim and the perpetrator. Although differential extraction has been proven to be effective (Vuichard et al., 2011), female DNA is also present and hence the results obtained using autosomal STRs are difficult to obtain or interpret as mixtures. In a situation where post-coital interval is extended, it was shown the successfulness of differential extraction is reduced and practically not useful after 48 hours (Hall and Ballantyne, 2003), whereas previous studies showed surviving time of sperm in the post-

coital period is up to 3 days and up to 7 days in the cervix (Willott and Allard, 1982). Spermatozoa can be lost due to vaginal inflammation, excessive douching and administering of spermicides (Sibille et al., 2002). Using male specific markers such as Y-STRs, male DNA is easily amplified without amplifying female DNA in a sample (Jobling et al., 1997, Gusmaúo et al., 2006).

Another application of Y-STR kits in forensic DNA profiling is kinship testing of male relatives or paternity testing of a male child. Autosomal STR kits are usually used for paternity cases however it has a limitation in some deficiency cases when the putative father is not available for comparison. In this situation, Y-STR analysis is more useful by analysing the closest available male relative of the putative father since same Y-STR haplotype is shared between close male relatives (Kayser and Sajantila, 2001, Stumpf and Goldstein, 2001). Due to the lack of recombination in NRY region, the individual Y-STR loci are used together and the resultant profile is called a "haplotype". Thus all the markers on a Y chromosome actually behave as one locus due to being linked. The power of such a test is greatly reduced as only the frequency of the haplotype can be represented in the relative population as the power of discrimination (Gusmaúo et al., 2006).

1.5.4. Y-STR Multiplex Assays

Since the mid of 1990's, there have been a number of Y-STR loci characterised and validated for forensic applications (Kayser et al., 1997a, Schneider et al., 1998). Based on the studied loci, the Europe an core set of Y-STR loci called "minimum haplotype" which comprised of 7 markers was developed (Table 1.3) (Kayser et al., 1997a). The Scientific Working Group on DNA Analysis Methods (SWGDAM) in USA extended the number of the loci in the minimum haplotype by adding the two loci DYS438 and DYS439 (Report on the Current Activities of the Scientific Working Group on DNA Analysis Methods Y-STR Subcommittee, 2004). Both minimal haplotype and extended haplotype could be developed using both Y-Plex 6™ and Y-Plex-5™ multiplex Y-STR

Kits (ReliaGene Technologies, Inc., New Orleans, LA). The Y-Plex 6™ kit included DYS19, DYS390, DYS391, DYS393, DYS389II and DYS385 whereas the Y-Plex 5™ kit included DYS389I/II, DYS392, DYS438 and DYS439 (Pascali et al., 1998).

A comprehensive study was conducted searching for novel Y-STR markers resulted in 166 novel Y-STR markers, which bringing the total number of Y-STRs with a potential value for use in research and forensic DNA profiling to 219 (Kayser et al., 2004). Subsequently, many commercial Y-STR multiplex assays have been developed and were made available for forensic DNA profiling. The most commonly used multiplex Y-STR assays are AmpFℓSTR Yfiler® PCR Amplification Kit (Applied Biosystems™), and the Powerplex Y® Kit (Promega™), which amplify 17 Y-STR and 12 Y-STR loci respectively. Both kits comprise the minimal haplotype and have been validated for forensic casework (Krenke et al., 2005, Mulero et al., 2006). However, recently PowerPlex Y23® (Promega™) has been developed which amplifies 23 Y-STR loci simultaneously (Thompson et al., 2013). In addition, the newly released, Yfiler plus® (Applied Biosystems™) has the ability to amplify 26 Y-STR loci as described in Table 1.3 (Ballantyne et al., 2013).

Table 1.3: Y-STR markers available in currently used multiplex PCR assays.

Minimal Haplotype	Powerplex Y®	Yfiler®	PowerPlex Y23®	Yfiler Plus®	RM-Yplex*
DYS19	DYS19	DYS19	DYS19	DYS19	
DYS389I	DYS389I	DYS389I	DYS389I	DYS389I	
DYS389II	DYS389II	DYS389II	DYS389II	DYS389II	
DYS390	DYS390	DYS390	DYS390	DYS390	
DYS391	DYS391	DYS391	DYS391	DYS391	
DYS392	DYS392	DYS392	DYS392	DYS392	
DYS393	DYS393	DYS393	DYS393	DYS393	
DYS385ab	DYS385ab	DYS385ab	DYS385ab	DYS385ab	
	DYS437	DYS437	DYS437	DYS437	
	DYS438	DYS438	DYS438	DYS438	
	DYS439	DYS439	DYS439	DYS439	
		DYS448	DYS448	DYS448	
		DYS456	DYS456	DYS456	
		DYS458	DYS458	DYS458	
		DYS635	DYS635	DYS635	
		YGATAH4	YGATAH4	YGATAH4	
			DYS481	DYS481	
			DYS533	DY533	
			DYS549		
			DYS570	DYS570	DYS570
			DYS576	DYS576	DYS576
			DYS643		
				DYS460	
				DYS627	DYS627
				DYS518	DYS518
				DYS449	DYS449
				DYF387S1	DYF387S1
					DYF399S1
					DYF403S1ab
					DYF404S1
					DYS612
					DYS626
					DYS526ab
					DYS547

*RM-Yplex was developed as a multiplex during this PhD work.

1.5.5. Limitations of Y-STR Assays

While the male specificity of Y-STRs is probably their greatest advantage in forensic DNA profiling, it still has a number of limitations. Y chromosome has low effective population size, which means an increased susceptibility to genetic drift, causing drastic fluctuation of alleles frequencies due to such random events. Such events can completely wipe-out a particular male lineage reducing the genetic variability in a population (Charlesworth, 2009).

It was shown previously that Y-STR loci are almost as polymorphic as autosomal STRs, however the lack of recombination in the Y chromosome makes the match of haplotypes in male individuals, especially in close relatives, more common. It was also shown that, autosomal and Y chromosome STRs have the same range of mutations rates (Roewer et al., 1992).While currently used Y-STRs are able to reliably differentiate between different male lineages, they still cannot resolve these lineages down to the individual level when assessing paternal relatives (Gusmaúo et al., 2003).

Thus in forensic DNA profiling, using currently used Y-STR kits means in the majority of cases paternally related males individuals will show identical Y-STR haplotype (Mulero et al., 2006, Thompson et al., 2012). In addition, identical haplotypes could be found between unrelated males individuals from the same community when there is considerable population substructure, arising as a result of non-random mating in that subpopulation. This leads to haplotype frequencies divergent from estimations based on the total population. Generally speaking, if few males contributed in founding a subpopulation then some haplotypes will be more common in the subpopulation than in the wider population. Highly sub-structured populations thus require their own database in order to get the true value of a match in case where individuals from such populations are incriminated (Lewontin and Hartl, 1991, Charlesworth, 2009).

Since most of populations exhibit substructure to different levels, many unrelated males might share the same Y-STR haplotypes (Gusmaúo et al., 2003, Hedman et al., 2004). The evidential value of Y-STRs is therefore very weak when compared to autosomal loci because of linkage and sub-structuring. Most of the times the evidential value of a matching Y-STR haplotype is simply expressed as "Haplotype Frequency". The Discrimination capacity of a Y-STR haplotype can also represent the evidential power of such a system. Discrimination capacity (DC) is an estimate of the number of unique haplotypes observed in the populations samples (Kayser et al., 1997a). The current batteries of Y-STR loci in various multiplex kits have been enhanced as far as the number of loci is concerned in order to increase this evidential value with little success (Purps et al., 2014). Clearly, there is a need to enhance the discrimination capacity of Y-STR systems.

1.5.6. Improving Y-STR Discrimination Capacity

Theoretically, there are two ways to increase the discrimination capacity of the Y-STR haplotypes. First is to increase the number of markers used in the analysis which then will increase the probability of detecting loci with different alleles. Hence it will help to identify male individuals. Such allele differences, which have been accumulated by mutation events, will lead to separating a lineage into two separate haplotypes. Prior to the release of Yfiler® amplification kit, such an approach was attempted previously in a study where 20 novel single copy Y-STR loci were examined, using single multiplex reaction, on seventy-four male samples from USA population and showed an increase in haplotype resolution compared to the 6 Y-STR loci available in Y-Plex™ 6 kit (Butler et al., 2002). In a subsequent comparative study 37 Y-STR loci, including 17 loci available in Yfiler® amplification kit, were analysed in 656 male samples from US population. Using Yfiler® kit, three samples had the same haplotype. These three could be differentiated by adding only one Y-STR locus, either DYS522 or DYS576. In addition,

out of the additional 20 Y-STR loci analysed, 7 loci only were able to resolve as much as the number of haplotypes resolved by 20 Y-STR loci (Decker et al., 2007). Another study analysed a worldwide sample consisting of 590 male samples, from 51 worldwide population belonging to 8 different regions, using 49 novel simple single copy Y-STRs (ssY-STR) parallel to 17 known Y-STRs. With this approach more than 95% of male individuals displayed a unique haplotype as compared to 90% of male's individuals when Yfiler® kit was used. It was determined that the markers with highest mutation rates, DYS570 and DYS576, were the ones which contributed the most in discriminating males individuals for this population sample. Although 67 Y-STR loci were used in this study there were still 20 haplotypes of males shared between 47 male individuals which were confirmed by autosomal STR that they do not share close biological relationship (Vermeulen et al., 2009).

In order to enhance the discrimination capacity, another approach is to use a panel of Y-STR markers that display a higher mutation rate than the markers that are usually combined in the multiplexes like Yfiler® or Powerplex® Y. A survey had been conducted for suitable Y-STR markers, however the mutation rates of these markers were not known (Kayser et al., 2004). In a subsequent study, mutation rates of 186 Y-STR loci were studied and estimated in about 2000 father-son pairs from Poland and Germany (Ballantyne et al., 2010). As a result, 13 markers were discovered to have mutation rates of 1×10^{-2} and higher, whereas the rest of the investigated Y-STR loci showed a mutation rate of 1×10^{-3} and lower. These comprised 12 tetranucleotide markers and 1 trinucleotide marker. Four out of the 13 Y-STRs were multi-copy markers and 9 were single copy markers. Therefore, 21 alleles can be typically amplified using these 13 RM-YSTR markers, were termed rapidly mutating Y-STRs (RM-Y-STRs) (Ballantyne et al., 2010). Table 1.4 summarises their properties along with their mutation rates.

Table 1.4: Properties of RM Y-STRs loci as adapted from Ballantyne et al. (2012) and Mulero et al. (2014).

Locus	Repeat type	Repeat motif (Variants in bold type)	Allele Ranges		Mutation rate
			Repeat number	Base pairs	
DYF387S1	Tetra, complex	$(AAAG)_3(GTAG)_1(GAAG)_4N_{16}(GAAG)_9$ $(\mathbf{AAAG})_{13}$	28-38	241-281	1.58×10^{-2}
DYF399S1	Tetra, complex	$(GAAA)_3\mathbf{N_{7–8}}(\mathbf{GAAA})_{10–23}$	10-23	261-313	7.73×10^{-2}
DYF403S1a	Tetra, complex	$(\mathbf{TTCT})_{10–17}\mathbf{N}_{2–3}(\mathbf{TTCT})_{3–17}$	12-39	310-438	3.10×10^{-2}
DYF403S1b	Tetra, complex	$(TTCT)_{12}N_2(\mathbf{TTCT})_8(\mathbf{TTCC})_9(\mathbf{TTCT})_{14}N_2(TTCT)_3$	40-59	414-490	$1.18 \times 10_{-2}$
DYF404S1	Tetra, complex	$(\mathbf{TTTC})_{10–20}N_{42}(TTTC)_3$	10-20	171-211	1.24×10^{-2}
DYS449	Tetra, complex	$(\mathbf{TTTC})_{12–18}N_{22}(TTTC)_3N_{12}(\mathbf{TTTC})_{12–18}$	24-37	309-361	1.22×10^{-2}
DYS518	Tetra, complex	$(AAAG)_3(GAAG)_1(\mathbf{AAAG})_{14–22}(GGAG)_1(AAAG)_4N_6(\mathbf{AAAG})_{11–19}$	23-35	243-291	1.84×10^{-2}
DYS526a	Tetra, complex	$(CCCT)_3N_{20}(\mathbf{CTTT})_{11–17}(\mathbf{CCTT})_{6–10}N_{113}(\mathbf{CCTT})_{10–17}$	10-17	138-166	2.72×10^{-3}
DYS526b	Tetra, complex	$(CCCT)_3N_{20}(\mathbf{CTTT})_{11–17}(\mathbf{CCTT})_{6–10}N_{113}(\mathbf{CCTT})_{10–17}$	29-42	345-397	1.25×10^{-2}
DYS547	Tetra, complex	$(\mathbf{CCTT})_{9–13}T(\mathbf{CTTC})_{4–5}N_{56}(\mathbf{TTTC})_{10–22}N_{10}(CCTT)_4(TCTC)_1(\mathbf{TTTC})_{9–16}N_{14}(TTTC)_3$	36-48	410-458	2.36×10^{-2}
DYS570	Tetra, simple	$(\mathbf{TTTC})_{14–24}$	10-21	246-286	1.24×10^{-2}
DYS576	Tetra, simple	$(\mathbf{AAAG})_{13–22}$	13-23	170-210	1.43×10^{-2}
DYS612	Tri, simple	$(CCT)_5(CTT)_1(TCT)_4(CCT)_1(\mathbf{TCT})_{19–31}$	14-31	187-255	1.45×10^{-2}
DYS626	Tetra, complex	$(\mathbf{GAAA})_{14–23}N_{24}(GAAA)_3N_6(GAAA)_5$ $(AAA)_1$ $(\mathbf{GAAA})_{2–3}(GAAG)_1(GAAA)_3$	11-23	221-269	1.22×10^{-2}
DYS627	Tetra, complex	$(AGAA)_3N_{16}(AGAG)_3(\mathbf{AAAG})_{12–24}N_{81}(AAGG)_3$	10-24	301-372	1.23×10^{-2}

1.5.7. Rapidly Mutating Y-STRs (RM Y-STRs)

Despite of the recent discovery of the 13 RM Y-STRs, there are already number of studies incorporating at least one or more of the single copy RM Y-STRs in multiplex assays and their high diversity was noted in all of the studies (Redd et al., 2002, Hanson and Ballantyne, 2007a, Hanson and Ballantyne, 2007b, Rodig et al., 2008, D'Amato et al., 2009, Vermeulen et al., 2009, D'Amato et al., 2010, Hedman et al., 2011). However, the multi-copy markers including DYF387S1, DYF399S1, DYF403S1, and DYF404S1 have been studied to a lesser extent than single-copy which could be due to the difficulty associated in genotyping and sequencing of these markers (Henson, 2005). The interpretation of multi-copy markers is often difficult for two reasons. Firstly, the imbalance in the amplified peaks can affect the correct genotyping. Secondly in mixed samples, where more than one male individual has contributed to the sample, peak height imbalance will make assigning particular alleles to various contributors, difficult (Butler, 2005). Consequently, multi-copy Y-STR markers have been avoided in forensic DNA profiling multiplexes. However, the advantage of using multi-copy Y-STRs is that each copy has an independent mutation probability thus increasing the chances of mutation in accordance with number of copies present in the marker. Therefore, these markers have high power of discrimination (Berger et al., 2003, Butler and Schoske, 2004, Butler and Schoske, 2005, Jacobs et al., 2009).

In a recent study, Ballantyne et al. (2012) had employed three multiplex panels amplifying 13 RM Y-STRs to investigate 604 unrelated male samples from 51 worldwide populations samples (HGDP-CEPH). Thirteen RM Y-STR loci showed higher haplotype diversity and haplotype discrimination capacity than that obtained using 17 Y-STRs available in Yfiler® multiplex kit (Applied Biosystems) (Ballantyne et al., 2012). In addition, this set of markers was extremely successful in differentiating between closely

and distantly related males with 4.4-fold increase of average male relatives differentiation relative to Yfiler® kit (Ballantyne et al., 2012).

Moreover, in a worldwide study of RM Y-STRs, more than 99% of 12,272 investigated unrelated male individuals were completely individualised. Haplotype diversity in such a large number of samples was found to be extremely high. Haplotype sharing between 111 worldwide populations was almost absent, apart from 6 haplotypes which were shared between populations, within the same geographical region (Ballantyne et al., 2014). Although, the value of RM Y-STRs has already been demonstrated, multiplexing the 13 RM Y-STRs was not attempted. Such a multiplex could be used in forensic casework as normal workflow alongside other human identification kits. The multiplex will also enable studying more populations for RM Y-STRs quickly and easily.

1.6. United Arab Emirates (UAE)

1.6.1. Arabian Peninsula

The Arabian Peninsula including the great desert is located in the south west of Asia and is considered one of the most thinly populated areas of the world. It comprises of seven countries namely, Saudi Arabia (the largest), Yemen, Oman, UAE, Kuwait, Qatar, and Bahrain (Figure 1.6). It has been known that the earliest events of migrations in Arabian history are migrations from the south west of peninsula, from Yemen into neighbouring areas (Serai, 1997, Al-Sayegh, 1998).

Figure 1.6: Arabian Peninsula map illustrating locations of different countries in the region (Robert and Neil, 2014).

1.6.2. UAE Population

United Arab Emirates (UAE) lies in the Arabian Gulf between Oman and Saudi Arabia. It was founded in 1971, and consists of seven Emirates, namely: Abu Dhabi, Dubai, Sharjah, Ajman, Ra's Al-Khaymah, Al-Fujayrah, and Umm Al-Qaywayn (Figure 1.7), with a total population of 8.26 million in 2010, where native Arabs represent approximately 11.5% of the total population, and the majority of other ethnicities are Indian and Pakistanis (*Demographic and Social Statistics*, 2012). The first group of people arrived at this region of the Arabian Peninsula around nine or ten thousand years

ago. The ancient history of UAE population stays largely hypothetical. It is evidential that a highly developed civilisation existed in that region but later it has been vanished (Miles, 2005). The term Arab has been used to describe the ancient people who inhabited the Arabian Peninsula which have been drifted from the south regions of Arabian Peninsula, which is now known as Yemen. Arabs were travelling into two different directions. First route was the migration toward the valleys of the Tigris and Euphrates, which are located at the northern region of the Arabian Peninsula. In present, his region is known as Iraq. Whereas the second route of migration was toward eastern and northern coasts of the Arabian Paninsuela. People followed the later route were emerged from two recent groups in Yemen, Kahtanites (South) and Adnanites (North) (Miles, 2005). When this expansion reached United Arab Emirates, previously called Turicial States in nineteenth century, there were only a few tribes who inhabited the desert. The best known such tribes are Bani-yas, Almanasir, AlAwamir and Almurar (Abed and Hellyer, 2001). These tribes represent the majority of native Arabs living in United Arab Emirates (Miles, 2005). At the very beginning of the twentieth century these tribes started to migrate in different directions seeking natural resources for life. Some of them have reached the south coast of the Arabian Gulf, some others have reached Hajar Mountains which are located in the far north part of United Arab Emirates in the present and some of them inhabited the dessert near the oases in the middle region of UAE. These migrations have generated three different groups of UAE native Arabs which are the Bedouin Arabs (people living in the dessert), Urban Arabs (people living on the South coast of the Arabian Gulf) and Rural Arabs (people living in north Hajar mountains) (Al-Sayegh, 1998, Abed and Hellyer, 2001). Overtime, each group has developed slightly different Arabic language accent and traditions. Although urbanisation had a significant effect on people living in UAE, the basic structure and pattern of their lives have remained intact. All of these three

groups have a high preference of consanguineous marriages, which might have resulted in population sub-structuring (Abed and Hellyer, 2001, Mohammed et al., 2001).

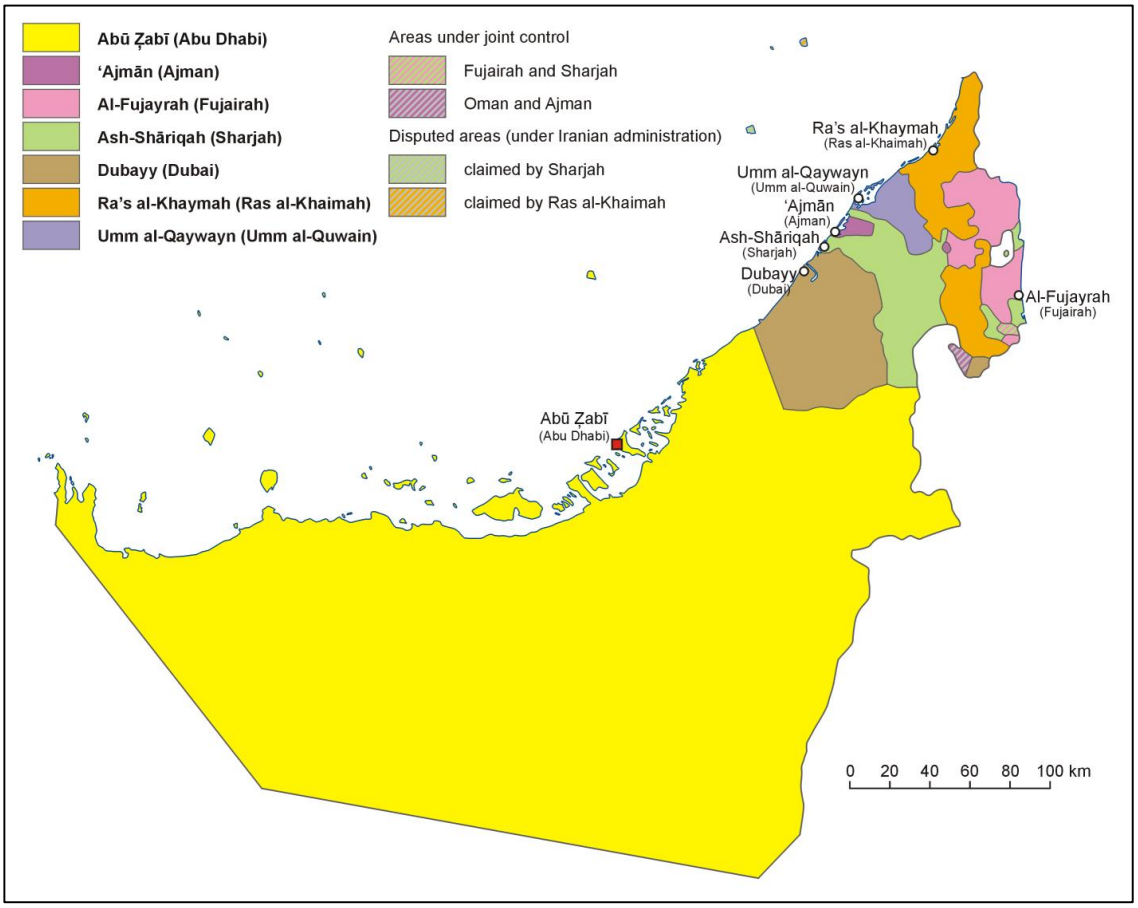

Figure 1.7: United Arab Emirates Map showing the boarders of each city in the country (Abed and Hellyer, 2001).

1.6.3. Genetic Polymorphisms in United Arab Emirates Population

Allele frequencies can be different from one population to another because of several reasons. Mutations, migration, random process in the transmission of alleles and the inbreeding rate, might result in distinctive frequencies of genetic markers within a specific population (Charlesworth, 2009). It was therefore important to estimate the allele frequencies of different RM Y-STR markers in UAE populations. Similarly, paternity and forensic parameters need studying for the purposes of gauging the value of these markers for use in forensic and paternity casework. According to Mohammed et al (2001), there was a significant difference in frequencies between Arabs and South Asian populations,

Indian and Pakistanis, in a study using Powerplex® 1.2 with eight autosomal loci. However, there was no significant difference detected between Indian and Pakistani subgroups residing in UAE. Also, there was no significant difference in frequencies between two of the UAE cities populations, Al Sharjah and Abu Dhabi. This made the author to conclude that a single database for UAE population could be used for forensic DNA profiling. However, neither all populations of the UAE were included in that study nor other loci, such as new autosomal loci, that have been introduced in the newer multiplex kits and Y chromosome STR loci. Another study conducted by Triki-Fendri et al. (2010) on samples of Kuwaiti population indicated a close genetic relationship between Kuwait and all other populations in Arabian Peninsula which support the theory of Arab expansion from Yemen and the principle of shared ancestry and reciprocal gene flow.

Using three different genetic DNA polymorphism panels, autosomal STR (Identifiler®), Y-STR (Yfiler®) and Y-SNPs (96 Y-SNP kit), Mohammad et al. (2009) were able to detect the genetic isolation and drift in nomadic Bedouin population from Kuwait by analysing 153 samples from their population. However, this study showed no evidence of segregation into the two original male descent groups from Yemen, Kahtanites (South) and Adnanites (North).

It has been shown in a previous study of populations from Dubai and Oman using Powerplex Y®, that the two populations are considerably clustered as one group and significantly differentiated from another cluster consisting of Yemen and Saudi Arabia (Alshamali et al., 2009). The haplotype diversity was found to be 0.989 in 217 male individuals (Alshamali et al., 2009). In a second study on 164 male individuals the haplotype diversity was found to be 0.9 (Cadenas et al., 2007). When these samples were typed using Yfiler® kit the haplotype diversity rose to 0.965. The higher value of haplotype diversity using Yfiler® kit is obvious due to the higher number of Y-STR loci

analysed in the samples. The haplotype diversities in UAE population in addition to two other populations in Arabian Peninsula, Qatar and Yemen, obtained using Yfiler® and Powerplex® Y are summarised in Table 1.5.

Table 1.5: Haplotype Diversities of three Arab populations using Yfiler® and Powerplex® Y multiplex kits (Cadenas et al., 2007).

Country	Haplotype Diversity (Yfiler®)	Haplotype Diversity (Powerplex® Y)
UAE	0.965	0.900
Qatar	0.947	0.834
Yemen	0.993	0.933

1.7. Objectives of Research

The general aim of this research is to develop, optimise and validate a multiplex assay, which can amplify 13 RM Y-STR loci simultaneously. Such multiplex assay will aid human population's studies as well as forensic applications. Several objectives of this research were to:

- Develop and optimise a multiplex assay amplifying 13 RM Y-STR loci simultaneously.
- Developing a sequenced allelic ladder for the multiplex assay.
- Perform developmental validation of the multiplex assay according to SWGDAM criteria.
- Perform internal validation of the multiplex assay at Dubai Police Forensic Laboratory.
- Investigate the RM Y-STR loci in United Arab Emirates populations.
- Develop RM Y-STR haplotype databases for United Arab Emirates, Indian and Pakistani populations.
- Conducting a population genetic study of UAE, Indian and Pakistani population using RM Y-STRs allele frequency data.

2. Materials and Methods

2.1. Materials

2.1.1. Population Samples

All samples were collected as liquid blood in UAE and stored as stains on FTA® cards. The UAE population samples comprised of 200 samples of unrelated male individuals from each of the three major native groups (Bedouin Arabs, Urban Arabs and Rural Arabs). In addition, some male samples were obtained from countries neighbouring UAE. These were 34 samples from Qatar, 13 samples from Kingdom of Saudi Arabia (KSA), 12 samples from Bahrain, 6 samples from Kuwait and 10 samples from Yemen populations. These samples were collected from candidates at Dubai Police Academy where more than 1000 native candidates are enrolled for different bachelor and master degrees. Information regarding native groups were obtained from each individual. Samples were collected up to submission aiming for 200 samples per native group of UAE population.

Two sets including, 242 Indian males samples and 243 Pakistani males samples were obtained from each of Indian and Pakistani populations residing in UAE. These samples were provided by Dubai Health Authority according to their records. The total number of population samples analysed in this project was 1160 male individual samples.

Samples from two typical Arab families from UAE were obtained. Family 1 consist of 18 male individuals and family 2 consist of 12 male Individuals. Both families represent samples from three different generations.

All samples used in this study were consented for the purpose of the research and were obtained from the Dubai Health Authority. Ethical approval was granted by both Medical Research Committee at Dubai Health Authority "Ref MRC-SR-03/2012_04" and Science, Technology, Engineering and Medicine (STEM) Committee at University of

Central Lancashire "Ref STEM034". All samples as stains on FTA® cards (Whatman® Bioscience, UK) were shipped to UK using DHL® courier.

2.1.2. DNA Control Samples

Male and female DNA control samples including 9948, 9947A and 2800M were obtained from Promega®, whereas 007 DNA controls and TaqMan® DNA controls were obtained from Life Technologies®.

Standard reference materials, SRM2395, male control DNA samples for Y-STR DNA profiling were obtained from National Institute of Standards and Technology (NIST).

2.1.3. Chemicals

All chemicals used in this study were obtained from SIGMA-ALDRICH® Co. UK, unless otherwise indicated.

2.1.4. General Preparatory Procedures

All experiments related to this study were carried out in the DNA forensic profiling laboratory at University of Central Lancashire unless otherwise indicated. The design of the laboratory allows unidirectional workflow (Figure 2.1) in order to prevent the occurrence of contamination. All Pre-PCR work and Post-PCR analyses were carried out in separate rooms for the same purpose.

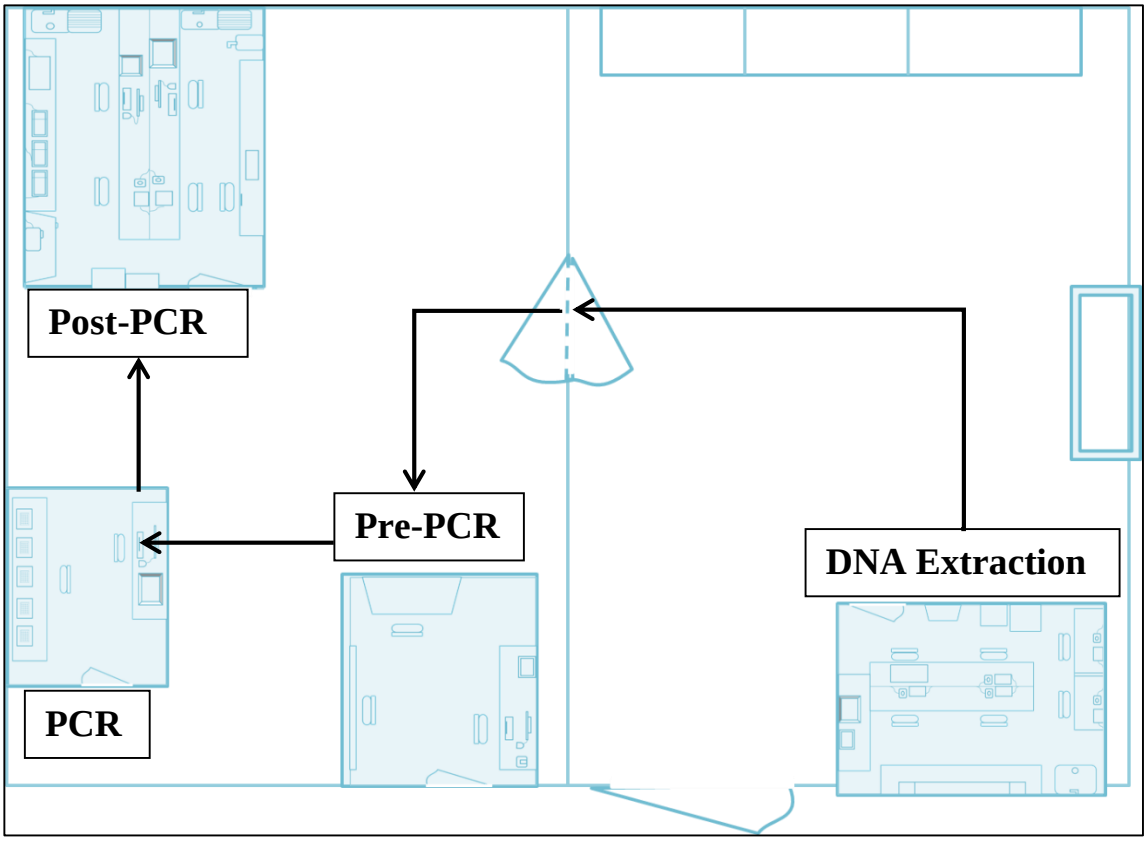

Figure 2.1: DNA profiling laboratory design at the University of Central Lancashire, Preston, UK.

2.2. Methods

2.2.1. RM Y-STR Multiplex Development and Optimisation

2.2.1.1. Selecting RM Y-STR Loci

13 RM Y-STR markers published by Ballantyne et al. (2010) were selected for this project. Primer sequences were obtained either from published resources or were designed to amplify the targeted loci (Table 2.1). Since primer sequences for these loci were developed for singleplex PCR purposes, it was essential to develop reference DNA sequences for each locus to help redesigning primers for the purpose of multiplexing. The loci needed to be amplified as a multiplex so appropriate primers were required. In a separate consideration appropriate fluorescent molecules to be used for each locus were selected to allow all the loci to be well located within the multiplex.

Table 2.1: Forward and reverse primers of 13 RM Y-STR loci as published by Ballantyne et al. (2010).

Locus	Forward Primer 5' - 3'	Reverse Primer 5' - 3'
DYF387S1	GCCTGGGTGACAGAGCTAGA	GCCACAGTGTGAGAAGTGTGA
DYF399S1	GGGTTTTCACCAGTTTGCAT	CCATGTTTTGGGACATTCCT
DYF403S1a/b	CAAAATTCATGTGGATAATGAG	ACAGAGCAGGATTCCATCTA
DYF404S1	GGCTTAAGAAATTTCAACGCATA	CCATGATGGAACAATTGCAG
DYS449	TGGAGTCTCTCAAGCCTGTTC	CCATTGCACTCTAGGTTGGAC
DYS518	GGCAACACAAGTGAAACTGC	TCAGCTCTTACCATGGGTGAT
DYS526a/b	TCTGGTGAACTGATCCAAACC	GGGTTACTTCGCCAGAAGGT
DYS547	TCCATGTTACTGCAAAATACAC	TGACAGAGCATAAACGTGTC
DYS570	GAACTGTCTACAATGGCTCACG	TCAGCATAGTCAAGAAACCAGACA
DYS576	TTGGGCTGAGGAGTTCAATC	GGCAGTCTCATTTCCTGGAG
DYS612	CCCCCATGCCAGTAAGAATA	TGAGGGAAGGCAAAAGAAAA
DYS626	GCAAGACCCCATAGCAAAAG	AAGAAGAATTTTGGGACATGTTT
DYS627	CTAGGTGACAGCGCAGGATT	GGATAATGAGCAAATGGCAAG

2.2.1.2. Study of Reference Sequences

BLAST (Basic Alignment Search Tool) was used to obtain the sequence of each locus in GeneBank® (www.ncbi.nlm.nih.gov) (Altschul et al., 1990). Primer sequences listed in Table 2.1 were used to search targeted sequences of the loci. BLAST search resulted in a list of sequences, which were similar to the query sequence and arranged in order of alignments scores. Alignment scores a measurement of similarity between primers sequences and query sequences being compared; the higher the alignment score the more likely that the listed sequence was related (Karlin and Altschul, 1993). Sequences for each locus with a high alignment scores were obtained and the structure of the repeats was compared with previously published repeat structure for confirmation (Ballantyne et al., 2010). The reference sequences for 13 RM Y-STR loci were thus obtained (Appendix 1).

2.2.1.3. Polymerase Chain Reaction (PCR) Optimisation

2.2.1.3.1. Selection of Primers

All unlabelled primers were ordered from Eurofins MWG Operon® (Germany), in order to test them in singleplex PCRs and analyse the products using agarose gel electrophoresis. Also, unlabelled primers were used to estimate the range of melting temperature (T_m) for various loci.

2.2.1.3.2. Primers Redesign

The published primers sets for RM Y-STR loci generated PCR products for DYF387S1, DYF399S1, DYS570, DYS626 and DYS518 loci that were overlapping each other. New primer sets for four loci, DYF387S1, DYS570, DYS576 and DYS612 were designed using DNA sequences generated from GeneBank® for each locus. Primer 3 software version 0.2 (www.genome.wi.mit.edu/cgi-bin/primer/primer3_www.cgi) was used to design the primers. This software allows options to choose certain parameters which are critical for PCR reactions to work such as, length of primers, melting temperatures and

expected product size. These parameters were kept as default while redesigning the primers. Primers sets for DYS576 and DYS612 were slightly modified considering the targeted range of primers melting temperature. Both DYF387S1 and DYS570 were completely redesigned in order to assemble all markers into one multiplex assay taking into account the allele size ranges as well as melting temperatures. These primers were obtained from Eurofins MWG Operon® (Germany) as unlabelled primers, for use in singleplex PCRs and analysed on agarose gel electrophoresis (Table 2.2).

Table 2.2: Sequences of the redesigned forward and reverse primers for four RM Y-STR loci, DYF387S1, DYS570, DYS576 and DYS612 and the PCR products size ranges for both Ballantyne et al. (2010) and redesigned primers.

Locus	Primers Sequences		PCR Product Size Range (bp)	
			Original	Modified
DYF387S1	**Forward 5' - 3'**	ACAGAGCTAGATTCCATTTTACCC	241-281	232-272
	Reverse 5' - 3'	GCCACAGTGTGAGAAGTGTGA		
DYS570	**Forward 5' - 3'**	CTGGCTGTGTCCTCCAAGTT	246-286	123-163
	Reverse 5' - 3'	GGCAACCTAAGCTGAAATGC		
DYS576	**Forward 5' - 3'**	GTTGGGCTGAGGAGTTCAATC	170-210	171-211
	Reverse 5' - 3'	GGCAGTCTCATTTCCTGGAG		
DYS612	**Forward 5' - 3'**	CCCCATGCCAGTAAGAATA	187-255	187-255
	Reverse 5' - 3'	GTGAGGGAAGGCAAAAGAAAA		

2.2.1.3.3. Development of Labelled Primers

Forward primers for each locus can be labelled with a fluorescent molecule to generate a labelled PCR product. This is required for detection purposes when using capillary electrophoresis based instruments such as ABI 3500, 3130 and 3100 Genetic Analysers. 6-FAM, Yakima Yellow, ATTO 550 and ATTO 565 dyes (Schubbert and Rittler, 2010), were chosen considering their emission spectral ranges and covalently linked to the 5′ end of the forward primers of various loci (Table 2.3). A decision was made to use GeneScan™ 600 LIZ® (Applied Biosystems, USA) as an internal standard taking into consideration the overall alleles sizes ranges for the 13 RM Y-STR loci. Care was taken

to select the dye panel to have the same absorption and emission wavelengths as the dyes used in G5 matrix standard supplied by Applied Biosystems. All primers as well as EF-01™ standard matrix were ordered from Eurofins MWG Operon® (Germany), whereas The DS-33 standard matrices, which includes G5 dye set, was supplied by Life Technologies®. Both standard matrix were used to generate data for the data collection software, installed on the computer connected to the ABI 3500 Genetic Analyser, to automatically analyse the different coloured dye-labelled DNA fragments in each single capillary. The forward primer fluorescent dyes for each of the 13 RM Y-STRs loci ensured adequate spread, separation and detection of different loci when amplified as a multiplex (Table 2.3).

Table 2.3 Fluorescent labels attached to each of the forward primer designed for RM Y-STR loci.

Locus	Forward Primer 5' - 3'	Modification
DYF387S1	TCCATTTTACCCCTAACAAGAAAA	**ATTO 565**
DYF399S1	GGGTTTTCACCAGTTTGCAT	**6-FAM**
DYF403S1a/b	CAAAATTCATGTGGATAATGAG	**Yakima Yellow**
DYF404S1	GGCTTAAGAAATTTCAACGCATA	**Yakima Yellow**
DYS449	TGGAGTCTCTCAAGCCTGTTC	**ATTO 565**
DYS518	GGCAACACAAGTGAAACTGC	**ATTO 550**
DYS526a/b	TCTGGTGAACTGATCCAAACC	**6-FAM**
DYS547	TCCATGTTACTGCAAAATACAC	**6-FAM**
DYS570	CTGGCTGTGTCCTCCAAGTT	**ATTO 565**
DYS576	TTGGGCTGAGGAGTTCAATC	**ATTO 550**
DYS612	CCCCCATGCCAGTAAGAATA	**6-FAM**
DYS626	GCAAGACCCCATAGCAAAAG	**Yakima Yellow**
DYS627	CTAGGTGACAGCGCAGGATT	**ATTO 550**

2.2.1.3.4. Quality Control for PCR Primers

All primers used in this project were ordered and quality control tested by Eurofins MWG Operon® (Germany) using HPLC purification prior to further use to confirm proper synthesis according to the STR multiplex development strategy.

2.2.1.4. 13 RM Y-STR Loci Multiplex Plan

The allele size spread for each locus, as observed in 604 unrelated male samples from 51 worldwide populations samples (HGDP-CEPH) (Ballantyne et al., 2012), was used to develop individual colour panels to be used for the multiplex reaction. A schematic diagram of the allele size ranges of each locus along with the fluorescent dye attached to the forward primer of each locus shows the position of each marker in the designed assay (Figure 2.2).

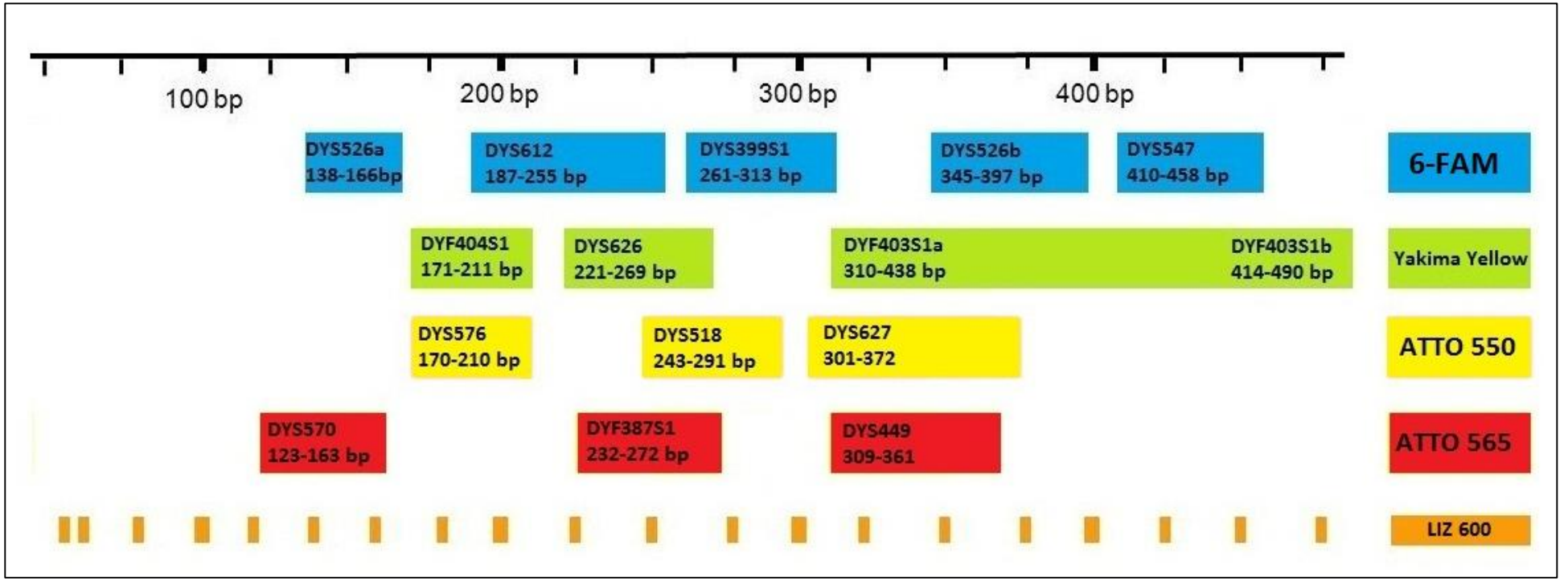

Figure 2.2: Various fluorescent colour panels' arrangement for developing RM-Yplex multiplex assay.

2.2.2. Quality Control

Good laboratory practices and precautions were adopted, during the project, to minimize contamination within the laboratories. Various controls were used during different experiments for the purposes of quality control. In the optimization stage of this project both TaqMan positive male DNA control and 9947A female DNA control were used in all experiments along with two extracted male DNA (coded R and K). TaqMan positive male DNA control, 2800M positive male DNA control, 9947A negative female DNA control and extracted female DNA control were included during allelic range determination experiments. During developmental validation, internal validation and population studies; TaqMan positive male DNA control, 2800M positive male DNA control, 9947A negative female DNA control, and 3 male DNA samples (R, K and M), were used in addition of the developed allelic ladder. Three extracted female DNA samples were used for various studies as well. Male DNA controls, 007 and 9948 were typed for quality assurance purposes during proficiency testing conducted for allelic nomenclature alignment with the RM Y-STR study group.

2.2.3. DNA Extraction and Purification

DNA was extracted from FTA® cards using an organic solvent extraction method followed by purification using a Microcon® centrifugal device (Fast Flow Filters). FTA® discs were also used for PCR after purification.

2.2.3.1. DNA Extraction from FTA® Cards

Using a sharp sterilised blade, a small piece of FTA® card sized approximately 3 mm^2 was cut and placed into a 1.5 ml microcentrifuge tube. 480 µl of freshly made digestion buffer (10 mM Tris-HCl, 10 mM EDTA, 50 mM NaCl, 2%(w/v) SDS, pH 7.5) was added to each sample in the 1.5 ml microcentrifuge tube along with 15 µl of proteinase K

(Sigma-Aldrich Inc. USA). The tube was briefly vortexed and centrifuged to bring all the contents to bottom of the tube which then was incubated at 56 °C for 24 hours.

After incubation, 400 µl of phenol:chloroform:isoamyl alcohol (25:24:1 Sigma-Aldrich Inc. USA) was added to the mixture, mixed, vortexed and then centrifuged for 5 minutes at 13000 g. The aqueous phase (upper layer) was then transferred, without disturbing the interphase, into a new 1.5 ml microcentrifuge tube.

To each tube, 1 ml of cold absolute ethanol was added. The tubes were inverted few times and then incubated at -20 °C for 1 hour. Then, the tubes were centrifuged for 30 minutes at 13000 g. Ethanol was removed by pipetting out from the tubes. The tubes caps were left opened to allow complete evaporation of the ethanol. DNA was resuspended by adding 30 µl of TE buffer (10 mM Tris-HCL, 0.1 mM EDTA, pH 8.0) and was incubated at room temperature for 10 minutes.

2.2.3.2. DNA Purification

Extracted DNA was purified and concentrated by transferring 30 µl extracted DNA, to the Microcon® centrifugal filter unit and centrifuged at 8000 g for 20 minutes. 30 µl of TE buffer was added to the filter which was then inverted into new sterilised 1.5 ml microcentrifuge and centrifuged at 1000 g for 2 minute to collect the purified DNA.

2.2.3.3. FTA® Cards (Whatman® Bioscience, UK) Purification

For each PCR reaction a disc sized 1.2 mm of FTA® card stained with blood was punched and transferred to a clean PCR tube. 200 µl of FTA® purification reagent (Whatman® Bioscience, UK) was added to the disc in the PCR tube and incubated at RT for 15 minutes, then the reagent was discarded and the disc washed with 200 µl of TE buffer for 5 minutes. The wash step was repeated once. Finally the disc was kept to dry at RT for 30 minutes before adding 14 µl of the PCR master mix. FTA® disc processing steps were carried out in a PCR hood to minimize the incidence of contamination.

2.2.4. DNA Quantification

The DNA concentrations were determined using Quantifiler® Duo Human DNA Quantification kit (Applied Biosystems, USA) and ran on an AB 7500 real-time PCR system (Applied Biosystems, USA). Quantification standards were prepared by serial dilution in TE buffer of the 200 ng/µl Human DNA Standard included in the kit. Eight serial dilutions were prepared which are 50 ng/µl, 16.7 ng/µl, 5.56 ng/µl, 1.85 ng/µl, 0.620 ng/µl, 0.210 ng/µl, 0.068 ng/µl and 0.023 ng/µl.

The procedure was carried out according to the manufacturer's protocol, but instead of using full volume reactions, the total volume of this assay was reduced (Westring et al., 2007). The reduced volume reaction consisted of a 5.25 µl of Quantifiler® Duo Human PCR Reaction Mix, 6.25 µl of Quantifiler® Duo Human Primer Mix and 1 µl of the DNA sample to a final reaction volume of 12.5 µl.

Samples, including the non-template control (NTC) and the DNA standards were loaded into MicroAmp® Optical 96-well reaction plate and sealed using optical adhesive covers (Applied Biosystems, USA). Each standard dilution was ran in duplicate, with replicates placed in wells adjacent to each other, followed by DNA samples.

The sample sheets were set up and the samples were run using the standard PCR cycle of 95 °C for 10 min; 40 cycles of 95 °C for 15 seconds and 60 °C for 1 min. Standard curves were generated and amplification plots were used to compare the results. Internal positive control (IPC) results were monitored for the presence of PCR inhibitors.

From the results obtained, the samples with low DNA concentrations (< 0.01 ng) were re-quantified to confirm DNA quantitation value thus obtained. The samples, which still showed very low DNA amounts, were re-extracted from the original FTA® card using the organic solvent (phenol/chloroform/isoamyl alcohol) method (Section 2.2.3.1) and then quantified using Quantifiler® Duo Human DNA Quantification kit again.

2.2.5. Unlabelled Primers

2.2.5.1. Singleplex PCR Optimisation

In order to minimize the number of steps involved in preparing the PCR reaction, Platinum® Master Mix (Applied Biosystems, USA) was chosen for optimization experiments as it includes PCR buffer, dNTPs, $MgCl_2$ and a hot start DNA polymerase enzyme mixed together in one tube. This was thought to reduce the preparation time, errors in the preparation and also reduce the chance of contamination compared to separately adding each component of the PCR reaction mix. Each primer set of the 13 RM Y-STR was tested in singleplex PCR amplification in reaction volume of 10 µl using 5 µl of Platinum® multiplex PCR Master Mix (Applied Biosystems, USA), 0.5 µl (0.5 µM) of each forward and reverse primers, 3 µl of PCR graded water and 1 µl (1 ng) of the respective DNA template. Quantities of PCR reaction mix components were chosen as per the manufacturer recommendations but with reduced volume by maintaining the proportion of each component. Each locus primer pairs were tested on 4 DNA samples including: TaqMan male DNA control; 9947A female DNA control and two organically extracted male DNA samples (K and R), over a range of different annealing temperatures in order to determine the optimal annealing temperature for each pair. Two types of PCR cycles were tested for optimization purposes, touch-down (TD) PCR cycles and a standard PCR cycle (Table 2.4).

TD thermal cycling PCR was performed on GeneAmp® 9700 (Applied Biosystems, USA) (Table 2.4). TD PCR allows for the annealing temperature of primers to be set higher than the melting temperature and then gradually decreased in every cycle by 1 °C or 2 °C degrees until it reaches the optimum annealing temperature. By the time the temperature reaches optimum temperature, specific products are present in excess and will be preferentially amplified. Hence, only a specific target will be amplified. However compared to standard PCR cycle, TD PCR takes longer (Korbie and Mattick, 2008). As

suggested by Ballantyne et al. (2010), 13 loci were separated into two sets with respect to TD PCR amplification protocol. The first set of loci comprised of DYF387S1, DYF399S1, DYS570 and DYS576. Each of these loci was amplified in a singleplex PCR reaction using TD 70-50 °C PCR protocol. The second set comprised of DYF403S1ab, DYF404S1, DYS449, DYS518, DYS526ab, DYS547, DYS612, DYS626 and DYS627. Each of these loci was amplified in a singleplex PCR reaction using TD 65-55 °C PCR protocol. Keeping the aim of multiplexing all the loci, amplification of the four loci included in the first set was conducted using TD 65-55 °C PCR protocol in singleplex reaction as well as the amplification of the second set was conducted using TD 70-50 °C.

In addition, for the purpose of comparison and shortening the amplification time a standard PCR cycle was also tested and optimal annealing temperatures determined. Initially, the 32 PCR cycles were chosen according to previously published PCR amplifications protocols (Hanson and Ballantyne, 2007b) for some of the loci used in this project, including DYS449, DYS570, DYS576 and DYS627. All primer sequences were the same as the ones used previously (Table 2.1). Using same composition of PCR components used in TD PCR experiment, each primer set was tested in a 32 cycles of standard PCR cycling conditions over a range of annealing temperatures from 54 ºC to 59 ºC. This range of temperatures was chosen from the results obtained in the TD PCR experiment.

Table 2.4: Touch down (TD) and standard PCR cycling conditions used for 13 RM Y-STR loci for singleplex reactions (Hanson and Ballantyne, 2007b, Ballantyne et al., 2010).

<table>
<tr><th colspan="4">TD 65-55 °C</th><th colspan="4">TD 70-50 °C</th><th colspan="3">Standard PCR</th></tr>
<tr><td>95 °C</td><td colspan="3">10 min</td><td>95 °C</td><td colspan="3">10 min</td><td>95 °C</td><td colspan="2">10 min</td></tr>
<tr><td>94 °C</td><td>30s</td><td rowspan="3">x10 cycles</td><td>94 °C</td><td>30s</td><td rowspan="3">x20 cycles</td><td>94 °C</td><td>30s</td><td rowspan="5">x32 cycles</td></tr>
<tr><td>65-1 °C</td><td>45s</td><td>70-1 °C</td><td>45s</td><td rowspan="3">58 °C</td><td rowspan="3">45s</td></tr>
<tr><td>72 °C</td><td>1 min</td><td>72 °C</td><td>1 min</td></tr>
<tr><td>94 °C</td><td>30s</td><td rowspan="3">x25 cycles</td><td>94 °C</td><td>30s</td><td rowspan="3">x15 cycles</td></tr>
<tr><td>50 °C</td><td>45s</td><td>50 °C</td><td>45s</td><td rowspan="2">72 °C</td><td rowspan="2">1 min</td></tr>
<tr><td>72 °C</td><td>1 min</td><td>72 °C</td><td>1 min</td></tr>
<tr><td>60 °C</td><td colspan="3">45 min</td><td>60 °C</td><td colspan="3">45 min</td><td>60 °C</td><td colspan="2">45 min</td></tr>
<tr><td>15 °C</td><td colspan="3">Forever</td><td>15 °C</td><td colspan="3">Forever</td><td>15 °C</td><td colspan="2">Forever</td></tr>
</table>

2.2.5.2. Multiplex PCR Optimisation

Firstly, 13 RM Y-STR loci were amplified in three multiplex PCR reactions in 10 µl according to Ballantyne et al. (2010) in order to assess primer suitability and the ability of the PCR master mix for multiplexing. The composition of each multiplex reaction is described in Table 2.5. All three reactions were amplified using TD 65-55 °C PCR amplification protocol (Table 2.4).

Table 2.5: The composition of three multiplexes for amplifying 13 RM-Y STRs using unlabelled primers.

Content	Set 1		Set 2		Set 3	
	Locus	Volume	Locus	Volume	Locus	Volume
Primers (forward and reverse)	**DYF387S1**	0.1µl (0.1 µM)	**DYS518**	0.3 µl (0.3 µM)	**DYF403S1**	0.6 µl (0.6 µM)
	DYF399S1	0.36 µl (0.36 µM)	**DYS626**	0.3 µl (0.3 µM)	**DYF404S1**	0.1µl (0.1 µM)
	DYS570	0.1µl (0.1 µM)	**DYS627**	0.5 µl (0.5 µM)	**DYS449**	0.1µl (0.1 µM)
	DYS576	0.1µl (0.1 µM)	**DYS526ab**	0.5 µl (0.5 µM)	**DYS547**	0.6 µl (0.6 µM)
					DYS612	0.2 µl (0.2 µM)
Sub-total	1.32 µl		3.2 µl		3.2 µl	
Water	2.68 µl		0.8 µl		0.8 µl	
Master Mix	5 µl		5 µl		5 µl	
DNA template	1 µl (1 ng)		1 µl (1 ng)		1 µl (1 ng)	
Final Volume	10 µl		10 µl		10 µl	

2.2.6. Labelled Primers

2.2.6.1. Singleplex PCR Optimisation

To confirm the amplification of each locus and the detection of fluorescent dyes, each of the 13 RM Y-STR loci were amplified in a singleplex PCR reaction using the conditions describe above and a TD 65-55 °C PCR amplification protocol (Section 2.2.5.1; Table 2.4).

2.2.6.2. 13 RM Y STR Loci Multiplex PCR Optimisation

The primary multiplex PCR amplification of 13 RM Y-STR loci was performed using equimolar concentrations of primers and the results were used as a guide for further experimentation. Primer concentrations were altered in order to achieve balanced peaks heights in the electropherograms (EPGs), following the PCR multiplex development strategy as suggested by Schoske et al. (2003). Starting concentration for each forward and reverse primer was 0.13 µM, bringing the total volume of primer mix to 5.2 µl. Since the recommended volume of master mix was half the volume of the reaction, 15 µl PCR

reaction volume was chosen to consist of 7 µl master mix, 5.2 µl primer mix, 1.8 µl PCR grade water and 1 µl of male DNA template (Table 2.6). TD 65-55 °C PCR amplification protocol was used (Table 2.4).

Table 2.6: Composition of multiplex PCR of 13 RM Y-STR loci tested in an initial experiment.

Content	Locus	Volume (µl)	
		F Primer	R Primer
Primers pairs	DYF387S1	0.2	0.2
	DYF399S1	0.2	0.2
	DYF403S1a/b	0.2	0.2
	DYF404S1	0.2	0.2
	DYS449	0.2	0.2
	DYS518	0.2	0.2
	DYS526a/b	0.2	0.2
	DYS547	0.2	0.2
	DYS570	0.2	0.2
	DYS576	0.2	0.2
	DYS612	0.2	0.2
	DYS626	0.2	0.2
	DYS627	0.2	0.2
	Total	5.2 µl	
Water		1.8 µl	
Master Mix		7 µl	
DNA Template		1 µl (1 ng)	

The optimisation of multiplex PCR of the 13 RM Y-STR loci was performed in a 15 µl total reaction volume. Various optimisation experiments established the optimal amounts of various components including the concentrations of each primer. The optimal PCR consisted of 7 µl of Platinum® Multiplex PCR Master Mix (Applied Biosystems, USA), 5 µl of the 13 RM Y-STR primers mix (see Table 2.7 for optimal volumes and concentrations of each primer), 2 µl of PCR grade water and 1 µl (1 ng) of DNA template.

Thermal cycling was performed using a GeneAmp® 9700 (Applied Biosystems, USA) using TD 65-55 °C PCR amplification protocol (Table 2.4).

Table 2.7: Optimal amounts of various components for multiplex PCR amplification reaction of 13 RM Y-STR loci using TD 65-55 °C PCR amplification protocol

Content	Locus	Volume (µl)	Concentration (µM)
Primers pairs	DYF387S1	0.07	0.05
	DYF399S1	0.1	0.07
	DYF403S1a/b	0.3	0.20
	DYF404S1	0.15	0.10
	DYS449	0.09	0.06
	DYS518	0.12	0.08
	DYS526a/b	0.6	0.40
	DYS547	0.3	0.20
	DYS570	0.06	0.04
	DYS576	0.05	0.03
	DYS612	0.1	0.07
	DYS626	0.06	0.04
	DYS627	0.5	0.33
	Total	5 µl	
Water		2 µl	
Master Mix		7 µl	
DNA Template		1 µl (1 ng)	

2.2.6.3. PCR Cycles Optimisation

For multiplex PCR a standard (Table 2.4) and a TD 65-55 °C PCR cycles (Table 2.8) were used. DYF403S1a/b and DYS526a/b loci showed inefficient amplification and further experiments were conducted by altering primer pairs concentrations. However, the results obtained were not satisfactory.

A new PCR cycle was therefore conceptualized and tested. Thirty-two PCR cycles were split into two stages; first stage, 16 cycles, with annealing temperature of 58 °C, and second stage, 16 cycles, with annealing temperature of 55 °C. The number of cycles were altered in each stage by keeping the total number of PCR cycles at 32 cycles until the optimised combination of cycles number was reached (Table 2.8). Upon successful optimisation of the developed multiplex assay, it was given a name of RM-Yplex assay.

Table 2.8: Optimised PCR cycling conditions for RM-Yplex multiplex assay.

Temperature	**Duration**	
95 °C	10min	
94 °C	30s	12 Cycles
58 °C	45s	
72 °C	1 min	
94 °C	30s	20 Cycles
55 °C	45s	
72 °C	1 min	
60 °C	45 min	
15 °C	Forever	

2.2.7. Detection and Analysis of RM-Yplex Multiplex PCR Products

2.2.7.1. Gel Electrophoresis

The PCR products of singleplex PCRs were prepared for AGE. The amplified products (5 µl) of single PCR reaction were mixed with 2 µl of 6x loading buffer and electrophoresed in a 2.5% agarose gel. The gel was stained using 5µl of SafeView™ (NBS Biologicals Ltd). 1 kb ladder (Thermo Scientific™) was loaded in a separate lane in the gel. The voltage was kept at 100 volts.

The PCR products of the three multiplex PCRs were prepared for AGE as well. 5 µl of the amplified product were mixed with 2 µl of 6x loading buffer and electrophoresed in a 3.0% agarose gel. The gel was stained with 5 µl of SafeView™ (NBS Biologicals Ltd). 1 kb ladder (Thermo Scientific™) was loaded in a separate lane in the gel. The voltage was kept at 50 volts.

2.2.7.2. Capillary Electrophoresis

The detection and separation of the 13 RM Y-STR multiplex products was done on ABI 3500 Genetic Analyser (Applied Biosystems, USA). Both G5 and EF-01™ standard matrix filter set were used to determine the five dyes FAM, Yakima Yellow, ATTO 550, ATTO 565 and LIZ. Before analysis was conducted, a spectral matrix was prepared using matrices standard set DS-33 (Applied Biosystems, USA) and EF-01™ matrix standard. PCR products from both singleplex and multiplex PCRs were prepared for loading. A master mix was prepared using 9.6 µl Hi-Di™ formamide (Applied Biosystems, USA) and 0.4 µl GeneScan™ 600 LIZ® size standard. For each sample, 10 µl of this mastermix was pipetted into each well and 1 µl of PCR product from each sample was added. Eight samples thus prepared were injected at a time for 10 seconds at 3,000 volts and separated at 15,000 volts for 45 minutes on run temperature of 60 °C. Separation of PCR product was performed using GeneScan POP-6 (Applied Biosystems, USA) polymer, 1x Genetic

Analyser Buffer with EDTA and in 50 cm capillary arrays. Samples were analysed using GeneMapper® *ID-X* Software v1.1.1 (Applied Biosystems, USA) after data collection.

2.2.8. Allele Range Determination

192 random male individuals' samples were selected from the set of population samples in order to maximise the capture of different alleles for each locus. These samples were extracted organically and quantitated using Quantifiler® Duo Human DNA Quantification kit. All of these samples were amplified using 1 ng of male DNA template and electrophoresed on ABI 3500 Genetic Analyser. In each run, 4 samples, TaqMan Male control, 9947A female control, male samples R and K. were used for quality control since allele sizes for each locus in each male control sample were established during optimisation experiments. The results from all 192 samples was analysed and allele ranges were determined.

2.2.9. Allelic Ladder Construction

2.2.9.1. Selection of Alleles

A number of alleles were determined during allelic range determination study. Appendix A1.1-A1.4 show samples selected for the development of the allelic ladder. The samples thus identified were used for developing the allelic ladder. 197 alleles for various RM Y-STR loci were chosen for inclusion in the allelic ladder (Table 2.9).

Table 2.9: Numbers of alleles identified and used for the development of allelic ladders for various RM Y-STR loci.

Locus	No. of alleles
DYF387S1	9
DYF399S1	27
DYF403S1	a: 15 b:13
DYF404S1	15
DYS449	12
DYS518	13
DYS526	a: 8 b:13
DYS547	12
DYS570	12
DYS576	11
DYS612	11
DYS626	12
DYS627	14

2.2.9.2. RM Y-STR Allele Sequencing

2-10 alleles at each marker were sequenced in order to confirm the sequence of PCR amplification product of each locus using BigDye® Terminator v3.0. 66 alleles included in the allelic ladder were thus sequenced (Table 2.10). Prior to sequencing reactions, the multi-allelic markers were separated by electrophoresing PCR products on polyacrylamide gel (FastCast™ BIO-RAD®), extracted from the gel and purified using Qiagen™ MiniElute spin column..

Table 2.10: Number of successfully sequenced alleles at each of the 13 RM Y-STR loci.

Locus	Number of Sequenced Alleles
DYF387S1	5
DYF399S1	3
DYF403S1a	2
DYF403S1b	3
DYF404S1	3
DYS449	4
DYS518	6
DYS526ab	8
DYS547	4
DYS570	6
DYS576	10
DYS612	3
DYS626	4
DYS627	5

2.2.9.2.1. Singleplex PCR

Using unmodified primers, PCR amplification was performed in reaction volume of 10 µl with 5 µl of Platinum® Multiplex PCR Master Mix (Applied Biosystems, USA), 0.5 µl (0.5 µM) of each forward and reverse primers, 3 µl of PCR graded water and 1 µl (5 ng) of DNA template. For multi-allelic markers, reaction volume and components were doubled (20 µl) in order to increase the total yield of the reaction. Thermal cycling was performed using a GeneAmp® 9700 (Applied Biosystems, USA) (Table 2.8).

2.2.9.2.2. PCR Products Purification

Following the PCR amplification process, PCR products were purified using two methods; microCLEAN™ reagent and MinElute™ Gel Extraction Kit. For single allele loci 10 µl of microCLEAN™ was added to the PCR product, mixed by pipetting and left at RT for 5 minutes. The PCR tubes were centrifuged at 13000 g for 7 minutes, supernatant was removed and then the PCR product pellets were resuspended in 10 µl of 10 mM Tris-Cl buffer pH 8.0 (TE) and left to rehydrate for 5 minutes at RT. For multi-allelic loci, DYF399S1, DYF387S1, DYF403S1ab and DYF404S1 separation was done by running 20 µl of amplified PCR product on precast polyacrylamide gel (FastCast™ BIO-RAD®). Each allele was separated from the gel using a sterilised sharp blade and purified separately using MinElute™ spin column as per manufacturer protocol with 10 µl final elution volume.

2.2.9.2.3. Sanger Sequencing of RM-Y STR Amplified Alleles

Sequencing reactions were performed in a reaction volume of 10 µl with 0.75 µl of BigDye® Reaction Mix (Applied Biosystems), 1.7 µl BigDye® Sequencing Buffer, 0.32 µl (3.2 pmoles) of forward or reverse primer, 2.23 µl of PCR graded water and 5 µl of purified PCR products. Each batch of sequencing reactions was prepared in conjunction with pGEM™ positive control in order to ensure that both instrument and DNA

sequencing kit are working properly. Thermal cycling was performed on a GeneAmp® 9700 (Applied Biosystems, USA) (Table 2.11).

Table 2.11: Sanger Sequencing PCR cycling conditions for sequencing amplified alleles of various RM Y STR loci.

Temperature	Duration	
96°C	1min	
96°C	10 seconds	25 Cycles
50°C	5 seconds	
60°C	4 minutes	
12°C	Forever	

2.2.9.2.4. Sanger Sequencing PCR Product Purification

Purification of PCR product before sequencing is essential for removing excess dNTPs, primers and other nonspecific PCR products. This prevents base line noise and the data obtained is clean enough for analysis. Following sequencing reaction, 1 µl of cold 3 M NaoAc, pH 4.6, 1 µl of Glycogen (20 µg/µl), 1 µl of EDTA (100 mM) and 30 µl cold ethanol (96%) were added and left to incubate for 24 hours at RT. After incubation, the samples were centrifuged for 30 minutes at 4 °C, supernatants were carefully removed and the pellets were washed twice with freshly prepared 70% ethanol. These were left to dry in the PCR machine at 50 °C for 10 minutes.

2.2.9.2.5. RM Y-STR Allele Sequence Detection and Analysis

ABI 3500 Genetic Analyser (Applied Biosystems, USA) was used to detect the sequencing reaction products. Before analysis, a spectral matrix was prepared and ran in the sequencer using BigDye® Terminator v3.1 sequencing standard matrix (Applied Biosystems, USA). Eight samples were injected at a time for 10 seconds at 3,000 volts and separated at 15,000 volts for 120 minutes. The run temperature was 60 °C. Samples

were prepared by adding 11 µl Hi-Di™ Formamide (Applied Biosystems, USA) to each sample in the PCR tube. The prepared sample was left for 10 minutes at RT to rehydrate and finally transferred to the corresponding well in the 96-well plate. Separation was done using GeneScan POP-6 (Applied Biosystems, USA), 1x Genetic Analyser Buffer with EDTA and 50 cm capillary arrays. After completion of data collection, samples were analysed using Sequencing Analysis Software v5.4 (Applied Biosystems, USA).

2.2.9.3. Development of RM-Yplex Allelic Ladder Cocktail

Samples having different alleles for each locus were co-amplified in 50 µl singleplex reaction using 2.5 µM of each pair of locus-specific primers, 1 µl (5 ng/µl) of male DNA template representing each allele, 25 µl of Platinum® Master Mix and PCR grade water was used to make up the final reaction volume. PCR reactions were performed using optimised PCR conditions (Table 2.8). Amplification products were electrophoresed in an ABI 3500 Genetic Analyser to determine the allelic peak heights for different alleles. Alleles Peak heights were balanced, focusing on inter-loci rather than intra-locus, by adjusting the input of DNA template according to allelic ladder development strategy shown in Figure 2.3. Upon reaching the final adjustment of inter-loci allele peaks balance, each locus-specific PCR reaction was purified using microCLEAN™ method described earlier (Section 2.2.9.2.2), however the volume of microCLEAN™ reagent added was 50 µl and the elution was done also in 50 µl. After purification eluted allelic ladder of each locus were pooled together in single tube and subjected to lyophilisation and stored at -20 ºC. Prior to use, ladder was reconstituted in 50 µl of TE buffer.

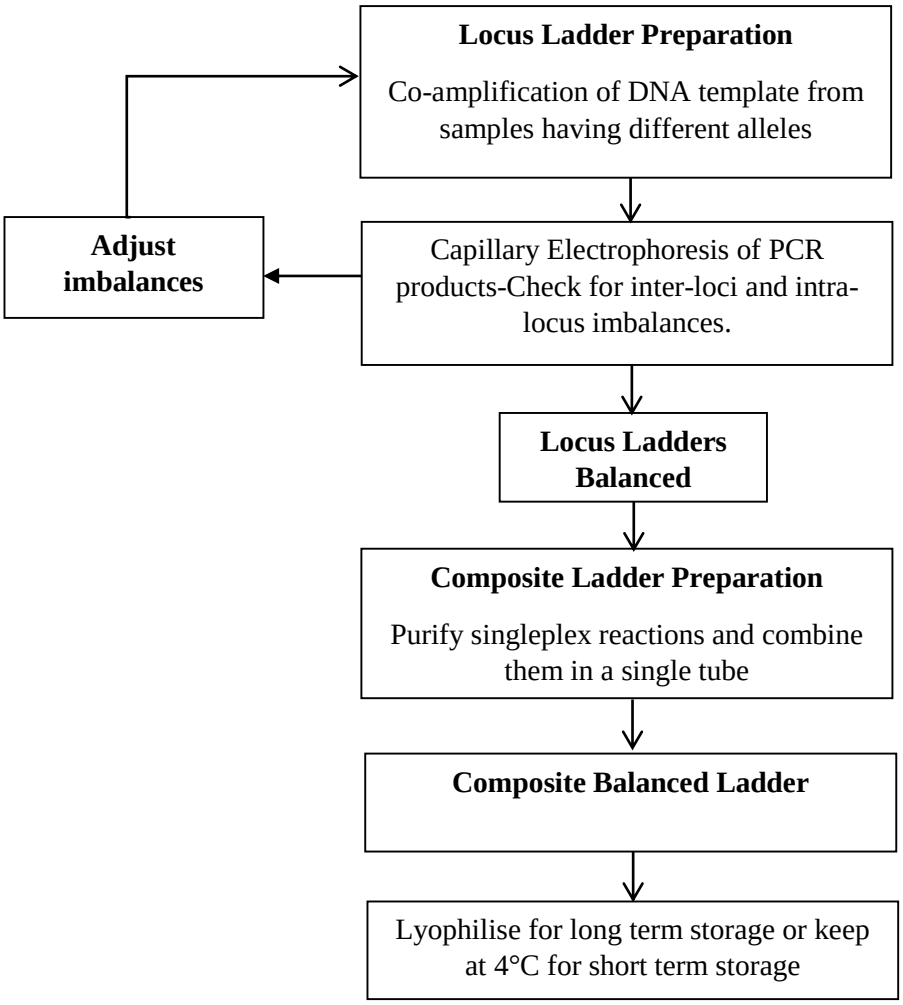

Figure 2.3: Flow chart outlining steps of development of the composite allelic ladder for RM-Yplex multiplex assay.

2.2.9.4. Precision Study

8 composite allelic ladders were prepared using 1 µl of the ladder mixed with 0.5 µl GeneScan LIZ600® and 9 µl of Hi-Di™ Formamide. The prepared allelic ladders were injected 6 times in a 8 capillary ABI 3500 Genetic DNA Analyser twice with a gap of 3 months. The first experiment was done on 1st of March 2014 and was repeated on the 1st of June 2014, in order to study the precision of the analysis. Preparation of the ladder was carried out twice and immediately before each experiment. The data from 48 injections in each experiment was analysed. The mean allele size and standard deviations for each allele were calculated and used to prepare the bins for each allele in the RM-Yplex multiplex assay.

2.2.9.5. Inter-laboratory Comparison of the Developed Allelic Ladder

Three allelic ladders were received from RM Y-STR study group from Manfred Kayser's laboratory in Rotterdam. The ladders were re-amplified. Each ladder was reconstituted in 7 µl of 1x TE buffer and then serial dilutions were prepared for each of the three allelic ladders at 1:100, 1:500 and 1:1000 ratios. For singleplex and multiplex reactions, 1 µl of the corresponding diluted allelic ladder was amplified in 10 µl and 15 µl PCR respectively (Table 2.12). Both singleplex and multiplex reactions were performed using optimal PCR cycling conditions (Table 2.8).

Table 2.12: Singleplex and multiplex PCR reactions composition for the amplification of RM Y-STR allelic ladders.

Singleplex PCR		**Multiplex PCR**	
Component	**Volume**	**Component**	**Volume**
Primers (F and R)	0.5 µl each	Primers Mix	2.5 µl
Allelic Ladder	1 µl	Allelic Ladder 1	1 µl
Platinum® MasterMix	5 µl	Allelic Ladder 2	1 µl
PCR Grade Water	3 µl	Allelic Ladder 3	1 µl
Total	10 µl	Platinum® MasterMix	7 µl
		PCR Grade Water	2.5 µl
		Total	15 µl

2.2.10. RM-Yplex Multiplex Developmental Validation Study

In order to maintain high quality standards, the Scientific Working Group on DNA Analysis Method (SWGDAM) has published revised guidelines for the use of STR multiplex PCR genotyping systems in forensic analysis (NIST, 2012). The developmental validation was conducted in DNA profiling laboratory at University of Central Lancashire in accordance with the revised SWGDAM guidelines.

2.2.10.1. Analytical Threshold

Analytical threshold was determined using 25 negative female control. The samples were analysed at 1 RFU for each dye in GeneMapper *ID-X* software. The table of peak heights was transferred to Microsoft® Excel software for analysis. The highest non-specific peaks from each electropherogram were recorded for each dye. The peak heights were averaged over the total number of analysed samples. For all these peaks, standard deviation was calculated for each of the five dyes, blue, green, yellow, red and orange. The analytical threshold was then determined by adding 3x standard deviation to the average peak height value calculated for each dye (Gilder et al., 2007).

2.2.10.2. Reproducibility Study

Six male control samples consisting of 3 commercial standards 2800M, TaqMan, 9948; 3 previously tested male samples R, M and K were analysed by five experienced DNA analysts independently as blind samples. These controls were extensively used during the optimization stage of this project so the genotypes had been established. Each analyst was trained with a minimum of 2 years' experience in a forensic DNA profiling laboratory. Results were compared to each other as well as to reference genotypes. In addition, the Y-STR standard SRM2395 (National Institute of Standard Technology), which is a set of 6 samples including 5 male samples was genotyped first during this project. The set was then genotyped by the five analysts mentioned above. The genotyping results were compared for these samples.

2.2.10.3. External Proficiency Testing

In order to demonstrate the ability of the assay to produce accurate genotypes of male samples, an analysis was conducted and confirmed by inter-laboratory collaborative study. Eight quality control samples were received from Prof. Manfred Kayser, Erasmus MC, Rotterdam, Netherland as part of an exercise, which were amplified using RM-

Yplex. Positive male control 007, 9948 male control, negative female control 9947A and allelic ladder were ran aside the 8 control samples the samples for quality control. Results of the 8 control samples, as presented in Appendix 1 (Figures A1.1 to A1.9). These results were sent to Erasmus MC laboratory for assessment.

2.2.10.4. Precision and Accuracy Studies

A precision study was conducted using in-house developed allelic ladder (Section 2.2.9.4). TaqMan positive DNA control sample was used additionally for the precision and accuracy study. Analysis of this sample was conducted at different occasions along with the allelic ladder. Results from three different runs were analysed and compared to allelic ladder. Mean allele size and standard deviation was determined over three different runs and compared with the allelic ladder in order to illustrate the precision and accuracy of sample analysis.

2.2.10.5. Sensitivity Study

A set of six different male DNA samples (TaqMan, 9948, 2800M, R, K and M) were used to test the sensitivity of the assay, serial dilutions of these samples were prepared ranging from 31.25 pg to 1000 pg. Each sample was analysed using the optimal PCR cycling conditions (Table 2.8). Three different numbers of PCR cycles; 28, 30 and 32 cycles were used to assist the sensitivity of the assay.

2.2.10.6. Mixture Studies

Different male/female mixture ratios were prepared by holding the amount of 2800M male DNA control input DNA constant at 1 ng and varying the female portion. Three female DNA samples were extracted from female buccal swab sample and quantitated using Quantifiler® Duo quantification kit for this purpose. Mixture ratios tested were 1:250, 1:500, 1:750 and 1:1000 male to female ratios. Samples were diluted according to the mentioned ratios.

Male/male mixtures samples were prepared using 3 samples 2800M (M 1), TaqMan (M 2), R (M 3). Mixtures were prepared for Male 1/Male 2 and Male 2/Male 3 mixed in different ratios. Those ratios are 19:1, 9:1, 3:1, 1:1, 1:3, 9:1, and 1:19. Two male DNA control samples 2800M and Taqman, as well as sample R were diluted and prepared according to the quantitation results obtained using Quantifiler® Duo quantification kit.

2.2.10.7. Species Specificity Study

Non-human samples were provided by Dubai Zoo as per Dubai Municipality Authorisation. Buccal swabs were obtained from four higher primates, male Gorilla, male Chimpanzee, unknown gender Gibbon and male Orang-utan. The swabs were extracted organically and purified using Microcon® filters. DNA extractions were quantitated using NanoDrop 2000®. These samples were amplified using optimal PCR conditions in a 15 µl PCR reaction (Table 2.7 and 2.8) and electrophoresed on ABI 3500 Genetic Analyser. Samples which showed amplifications at various loci were re-amplified in a 10 µl singleplex reaction using the corresponding primer pair (Section 2.2.5.1).

2.2.10.8. Stability Study

The effect of common inhibitors including haematin, humic acid, and tannic acid were studied over a range of different concentrations of each. Stock solutions were prepared by dissolving the appropriate amount of inhibitor in order to produce three different concentrations including 1500 ng/µl, 3000 ng/µl and 4500 ng/µl for each. 1 µl from each stock was added in the 15 µl optimised PCR reaction in order to reach a final required concentrations; 100 ng/µl, 200 ng/µl and 300 ng/µl respectively. Three male DNA controls, 2800M, TaqMan and 9948 were analysed in triplicate at three different occasions.

2.2.10.9. PCR Parameters

To evaluate the effect of magnesium chloride ($MgCl_2$) on the multiplex PCR reaction, a titration experiment was conducted by amplifying 1 ng of 2800M, 9948 and TaqMan male positive controls DNA separately with $MgCl_2$ being added to the mix at different input amount. Since the amounts of $MgCl_2$ and EDTA were not known in the Platinum® multiplex master mix, an additional concentration of each additive to PCR reaction was achieved by using stock solution of each as described in Table 2.13. In each experiment a positive control, containing same amount of DNA template without the addition of $MgCl_2$ and EDTA, was prepared and analysed parallel with serial dilution reactions.

Table 2.13: Volumes of $MgCl_2$ and EDTA added for preparing a serial dilution of each.

Required concentration of $MgCl_2$	Volume added to 15 µl PCR reaction using 50mM $MgCl_2$ stock	Required concentration of EDTA	Volume added to 15 µl PCR reaction using 10mM EDTA stock
2.5 mM	0.75 µl	0.1 mM	0.15 µl
5 mM	1.5 µl	0.25 mM	0.37 µl
7.5 mM	2.25 µl	0.5 mM	0.75 µl
10 mM	3.0 µl	1.0 mM	1.5 µl

2.2.10.10. Stutter Analysis

In order to calculate stutters percentage at each locus, previously analysed 192 samples used for allele range determination experiment were employed. Stutter percentages were calculated for each locus by dividing the stutter peak height by the parent peak height and multiplying it by 100.

2.2.11. RM-Yplex Multiplex Internal Validation

An internal validation was conducted at Department of Forensic Sciences and Criminology (DFSC) laboratory in Dubai (United Arab Emirates). Internal validation

establishes the parameters of a method within a laboratory prior to its regular use. This is essential so that the reliability and the limitations of the method can be assessed. Revised guidelines for the use of Y-STR multiplex PCR genotyping systems for forensic DNA analysis published by SWGDAM were followed in this validation (NIST, 2012).

2.2.11.1. Differential Extraction

Differential extraction method was used to separate spermatozoa from other cell types as per the standard procedures followed at DFSC. Using a clean disposable scalpel for each sperm positive sample, half of the cotton swab was cut and transferred to sterile 1.5 ml micro-centrifuge tube. 450 µl of stain extraction buffer and 20 µl of proteinase-K were added to the swab and vortexed for 5 seconds. The tube was centrifuged briefly and was then incubated for 2 hours at 56 °C to lyse non-sperm cells. After incubation, cotton swab was removed and the solution filtered using spin-basket in a microcentrifuge tube at 13000 rpm for 5 minutes. The supernatant, containing DNA from epithelial cells was pippeted into a new sterile 1.5 µl micro-centrifuge tube.

2.2.11.1.1. Epithelial Cell Fraction

This fraction of the sample was then processed using the organic extraction methods starting from phenol chloroform addition step as detailed previously (Section 2.2.3).

2.2.11.1.2. Sperm Fraction

Sperm fraction was then washed by adding 400 µl of digestion buffer to the pellet and centrifuged again at 13000 for 5 minutes. All supernatant, but 50 µl, was removed without disturbing the pellet. 20 µl of proteinase-K and 20 µl of dithiothreitol (DTT) were added to the cell pellet and vortexed for 5 seconds. The sample was then incubated at 56 °C overnight. The sample was then processed following the traditional organic solvent extraction methods starting from the step of adding the organic phase to the sample (Section 2.2.3).

2.2.11.2. DNA Quantification

All male DNA samples used for internal validation were quantitated using Quantifiler® Duo quantitation kit (Section 2.2.4). Casework samples were previously quantitated using Quantifiler® Duo quantitation kit at DFSC. The quantitation data was accessed and appropriate quantity of samples were used for amplification reactions.

2.2.11.3. PCR Amplification and Genotyping using RM-Yplex

All samples were amplified according to the optimised PCR cycles conditions using GeneAmp® 9700 PCR thermal cycler (Section 2.2.6.3) and genotyped using ABI 3500 Genetic Analyser with POP-4 configuration (Section 2.2.7.2). Alleles were assigned according to the latest nomenclature update for the corresponding loci (Ballantyne et al., 2012, Mulero et al., 2014). A detection threshold of 50 RFU was used during all experiments which was determined during prior to perform validation experiments.

2.2.11.4. Capillary Electrophoresis

All analyses during the internal validation study were conducted on ABI 3500 Genetic Analyser using POP-4 and 36 cm capillaries arrays. Before analysis was conducted, a spectral matrix was prepared using matrix standard set EF-01™ Matrix Standard. PCR products from both singleplex and multiplex PCR amplifications were prepared with 9.6 µl Hi-Di™ formamide (Applied Biosystems, USA), 0.4 µl GeneScan™ 600 LIZ® size standard and 1 µl of PCR product. Eight samples were injected at a time for 10 seconds at 3,000 volts and separated at 15,000 volts for 45 minutes with a run temperature of 60 °C. Separation of PCR product was performed using GeneScan POP-4 polymer (Applied Biosystems, USA), 1x Genetic Analyser Buffer with EDTA and 36 cm capillary arrays. After the data collection was complete, samples were analysed using GeneMapper® *ID-X* Software v1.1.1 (Applied Biosystems, USA).

2.2.11.5. Analytical Threshold

Minimum threshold study was established using a total of 27 amplification reactions using 3 sperm-negative vaginal swabs amplified and analysed 9 times. Same procedures were followed in determining the limit of detection (LOD) and limit of quantitation (LOQ) previously (Section 2.2.10.1) as per the guidelines provided by Gilder et al. (2007) and compared with the results obtained in developmental validation.

2.2.11.6. Accuracy and Precision

For reproducibility study, 6 samples including 2800M, 9948, TaqMan and 3 unrelated male samples, namely R, K and M, were amplified and genotyped by three analysts along with SRM2395 set of samples, blindly. All DNA analysts had at least two years' experience in the field. After completion of analysis by each analyst, the sample genotypes were collected and assessed. Precision study was done using 96 male DNA samples. These samples were amplified using purified 1.2 mm FTA® disc and electrophoresed twice on ABI 3500 Genetic Analyser. Average allele sizes and their standard deviations were calculated for each locus for all samples.

2.2.11.7. Specificity and Sensitivity

Two male DNA control samples, 2800M and TaqMan, were used in the sensitivity study. A male DNA sample R, which was extracted using organic solvent methods and quantified using Quantrifiler® Duo system. Serial dilutions of the three different male DNA samples were prepared including, 1000 pg, 750 pg, 500 pg, 250 pg, 125 pg, 62.5 pg and 31.25 pg of each. These samples were passed to three different DNA analysts with at least two years' experience in the field.

In specificity study, 1 ng of 9948 male DNA control was amplified in the presence of 2 µg of different female DNA samples. Male DNA control, 9948, was tested three times, where each time different female DNA was added to the PCR reaction.

2.2.11.8. Stutter Analysis

Stutter study was conducted using same set of 96 samples used for the precision and accuracy study. Stutter ratio established was used for creating stutter file for GeneMapper *ID-X* software.

2.2.11.9. Male/Male Mixture

Three male DNA samples 2800M (M1), TaqMan (M 2) and male DNA sample R (M 3) were used to prepare two sets of male/male mixtures. Two sets of mixtures including M1/M2 and M2/M3 were mixed at different ratios including 19:1, 9:1, 3:1, 1:1, 1:3, 1:9 and 1:19. Mixtures were analysed and interpereted independently and then results were collected and assessed.

2.2.11.10. Non-probative Casework Samples

25 casework samples from 14 cases received by DFSC in the past were provided to assess the application of RM-Yplex assay in forensic casework samples. Casework samples included 8 vaginal swabs, 3 perineal swabs, sets of 4 underwear, 1 beddings, 1 cigarette butts, 2 nails cuttings and 1 tooth sample which had been previously analysed using Identifiler® and Yfiler® kits. All samples had been previously extracted using organic solvent extraction method and differential extraction methods and quantified using Quantifiler® Duo quantification system.

2.2.11.11. General Precautions

Samples were processed taking precautions to avoid all contamination. Negative controls were employed during the different steps of sample processing in order to assess

contamination. Unidirectional movement from pre to post-PCR areas was observed, as per DFSC standards, to avoid cross-contamination across sections (Figure 2.4). Samples were only transferred through specific hatches from pre-PCR to post- PCR areas.

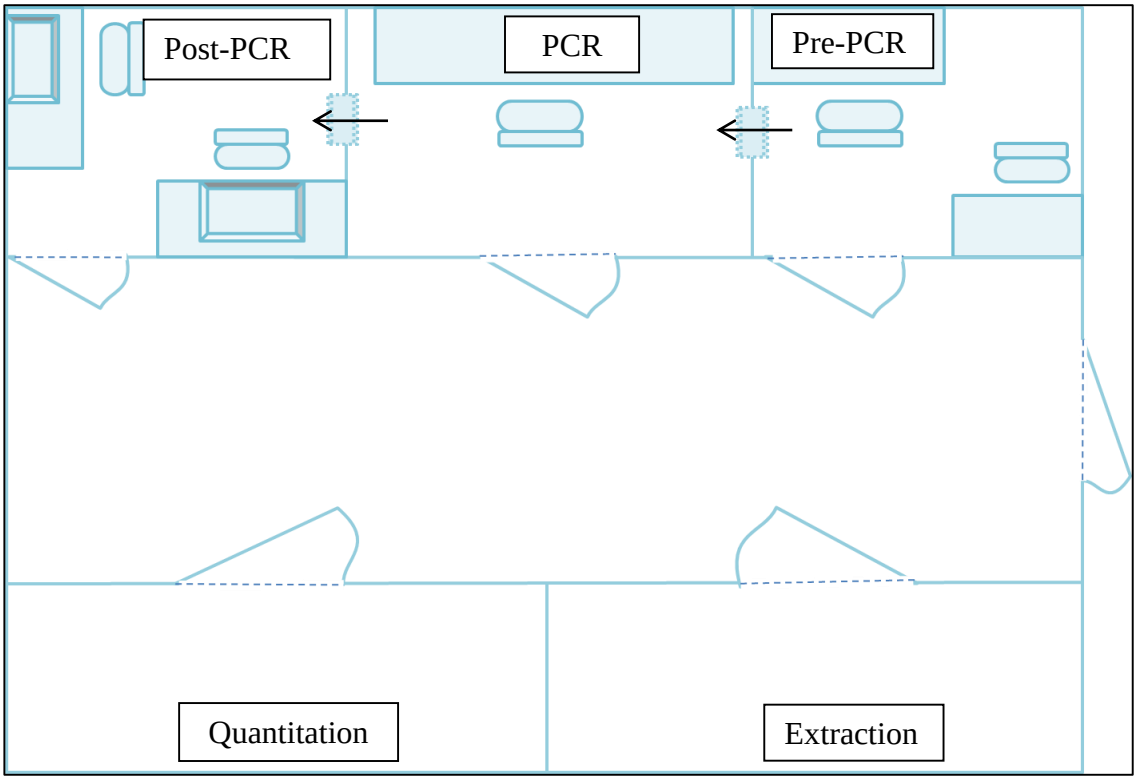

Figure 2.4: Illustration of DNA profiling Laboratory design in the Department of Forensic Sciences and Criminology (DFSC) Dubai Police.

2.2.12. Statistical Analysis of Data

2.2.12.1. General Data Analysis

All general statistical analyses were performed using Microsoft® Excel 2013 software. SPSS software v21 was also used to perform statistical analyses such as independent t-test with a significance level of 0.05.

2.2.12.2. Population Data Analysis

2.2.12.2.1. Allele Ranges and Frequencies

Calculation of allele frequencies were done for each locus, using Microsoft® Excel, by counting the number of individual alleles and dividing by the total number of alleles in

the relevant population. However, for multi-allelic loci, since the region represented by each allele cannot be differentiated using any of the primers described herein, were analysed as haplotype. For example, if genotype at DYF387S1 was 35 and 37 repeats, it is not known which region contains the respective allele. That means this sample can be either 35-37 or 37-35 (Redd et al., 2002).

2.2.12.2.2. Haplotypes and Loci Diversity Values

The locus and haplotype diversities were calculated using the following formula:

$$Haplotype\ Diversity/Locus\ Diversity\ (H) = \left(\frac{n}{n-1}\right)\left(1 - \sum_{i=1}^{k} p_i^2\right)$$

Where n is the number of samples, pi is the allelic frequency or haplotype frequency (Bosch et al., 2002). The probability of obtaining an identical haplotype in a pair of random unrelated male was estimated as (1- Haplotype Diversity). Discrimination capacity was determined according to the following formula (Redd et al., 2002):

$$Discrimination\ Capacity = \frac{Number\ of\ different\ Haplotypes\ observed}{Number\ of\ Individuals}$$

Microsoft® Excel software was used to perform all locus and haplotype diversities in the generated set of data.

2.2.12.2.3. Population Sub-structuring

The amount of genetic structure was estimated by calculating the F_{ST}, fixation index, values for the total population in the datasets as well as for pairwise comparison between regional groups.

The fixation index (F_{ST}) illustrates the correlation between randomly chosen gametes from a sub-population compared to the entire population. It could also be explained as genetic differences between populations because of the differences in alleles and

haplotype frequencies. F_{ST} value ranges between 0 and 1. Values close to zero mean no significant difference between populations and that most genetic variations are within a population i.e. between individuals in the same population. However, higher F_{ST} value means there is significant genetic variation between the populations (Wright, 1978, Holsinger and Weir, 2009).

F_{ST} values for all datasets analysed in this project were calculated using analysis of molecular variance (AMOVA) which was performed using Arlequin v3.5 software (Excoffier and Lischer, 2010).

In order to simplify the results obtained from AMOVA analysis, F_{ST} values were plotted in Multidimensional Scaling figure using SPSS v21 software.

NETWORK 4.6.1.1, was used also to draw a network of the generated population data to investigate the clustering possibility, which has been used previously to illustrate population sub-structuring.

3. Development and Optimisation of RM-Yplex Multiplex Assay

3.1. Overview

Multiplex PCR has revolutionised the field of forensic DNA profiling. The development of multiplex assay should be tested and optimised to the point where it meets stringent performance standards developed by International Society of Forensic Science (ISFG) and Scientific Working Group for DNA Analysis Method (SWGDAM). An optimised multiplex PCR assay should be carefully designed in order to amplify all targeted loci and result in balanced signals under different conditions that a crime stain might undergo. In addition, optimised multiplex PCR reaction should result in a specific profile free of non-specific PCR products and artefacts. Targeted fragments of different loci, labelled with the same fluorescent dye, should be distinguishable from one another i.e. allele ranges of loci within the same colour fluorescence panel should not overlap. The aim of this chapter was to develop an optimised multiplex PCR assay to amplify the 13 RM Y-STR loci simultaneously.

3.2. Results

3.2.1. RM-Yplex Multiplex Development

A simple strategy was followed to design, develop and optimise the multiplex assay amplifying all the 13 RM Y-STR loci in single PCR reaction (Figure 3.1 A and B).

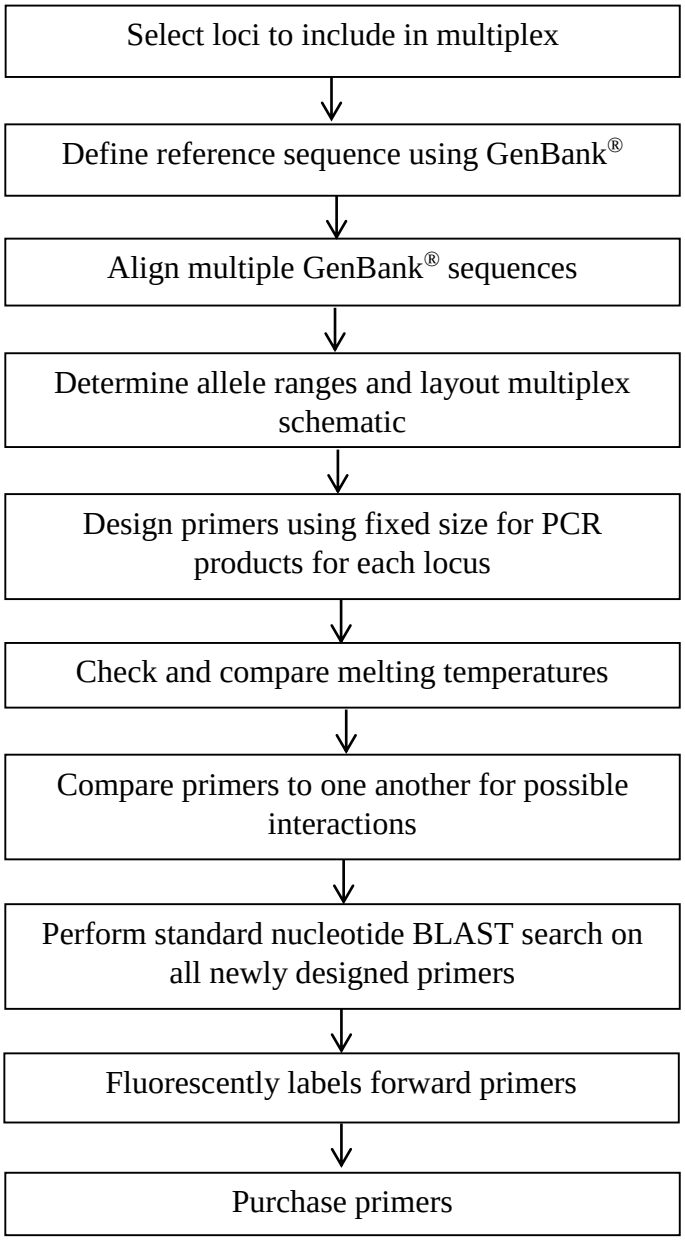

Figure 3.1 A: First part of developing new multiplex PCR panel as adapted from Schoske et al. (2003)

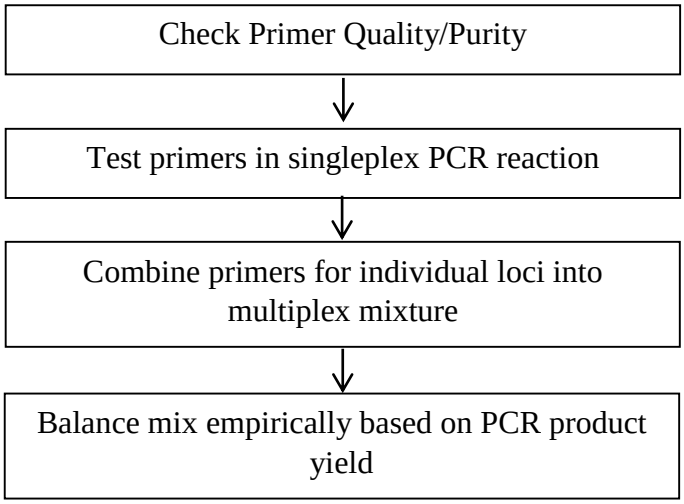

Figure 3.1 B: Second part of developing new multiplex PCR panel as adapted from Schoske et al. (2003).

3.2.1.1. Unlabelled Primer Sets

Three male samples were used to optimise the multiplex PCR. These were TaqMan® male control DNA, R male sample and K male sample. Unmodified primer sequences as published by Ballantyne et al. (2010) were used in a 10 µl singleplex PCR reactions with TD 65-55 °C cycle conditions. All loci were amplified successfully with 1 ng input for each of the 3 male DNA samples used as control with TD 65-55 °C PCR cycle condition, apart from DYF387S1, which showed many unspecific products with the same cycle condition. However, using the TD 70-50 °C PCR cycling conditions, according to Ballantyne et al. (2010), DYF387S1 was successfully amplified showing two specific bands on an agarose gel (Figure 3.2). Although specific bright bands were observed using TD 70-50 °C PCR cycles condition, there was still an observable smear appearing on the gel before and after the real band which was most likely caused by low yield of unspecific PCR products.

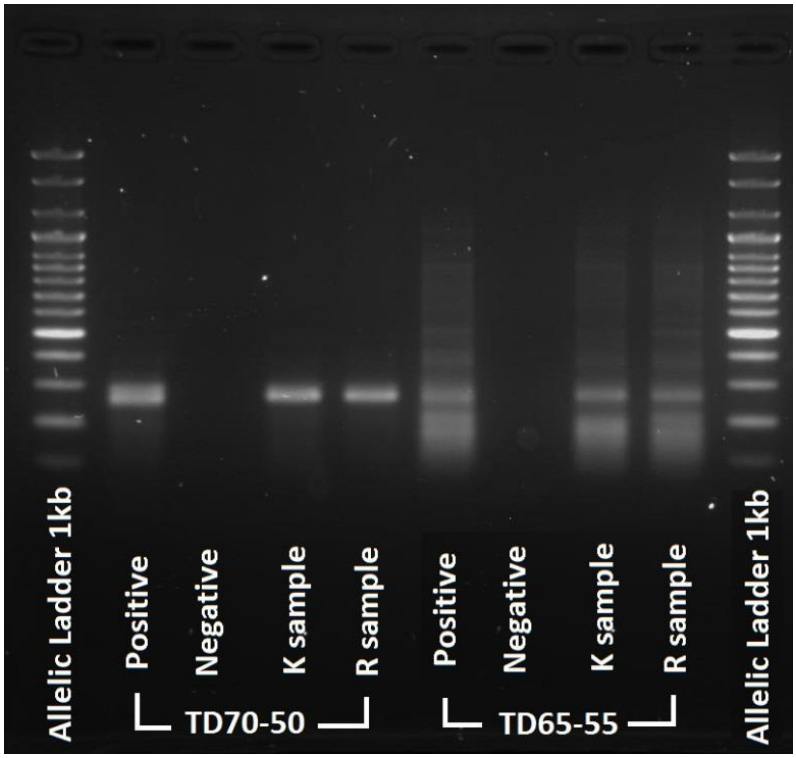

Figure 3.2: Agarose gel electrophoresis showing DYF387S1 locus amplification of positive control, negative control, male sample K and male sample R using two different PCR cycles conditions.

Since the aim of this experiment was to amplify the 13 RM Y-STRs in a multiplex reaction, all loci were tested with TD 70-50 °C PCR cycles condition. Most loci showed no amplification bands on an agarose gel apart from DYF387S1, DYF399S1, DYS576 and DYS570. As an alternative a standard PCR cycling conditions were used to estimate the optimal annealing temperature for all primers over the range of 54 °C to 59 °C temperatures with 1 °C increment. All primer pairs showed optimal amplification results between 55 °C and 58° C apart from DYF387S1 locus. Over such range, amplifications of DYF387S1 were showing unspecific peaks constantly and hence annealing temperature was tested up to 65 °C. As a result, DYF387S1 showed optimal amplified products over annealing temperatures range 60 °C to 65 °C. It was determined that it might not be possible to include all the set of 13 primers published by Ballantyne et al. 2010 in a multiplex reaction. Also using these primers in one reaction would need a 6-dyes chemistry multiplex assay in order to resolve overlapping sizes of alleles resulting from DYF387S1, DYF399S1, DYS570, DYS518 and DYS626 loci (Ballantyne et al., 2010). 6-dye chemistry was not validated at the time when this project started. A special matrix standard would have been required for such purpose, which might result in a need for budget beyond the limits of this project. Also it was anticipated that more dyes in the PCR might increase the artefacts such as dye bulbs and overlap of absorbance and emissions of fluorescence light.

Using Primer 3 software, primer pairs for DYF387S1 were redesigned in order to bring annealing temperatures down to the range of other loci. Also, primer pairs were redesigned for DYS570 using same software in order to bring the size of PCR products from 230-290 base pairs to 120-170 base pairs. Upon receiving primer pairs, they were tested in TD 65-55 °C PCR cycle condition. Both sets of primers showed specific product band at the expected size range (Figure 3.3).

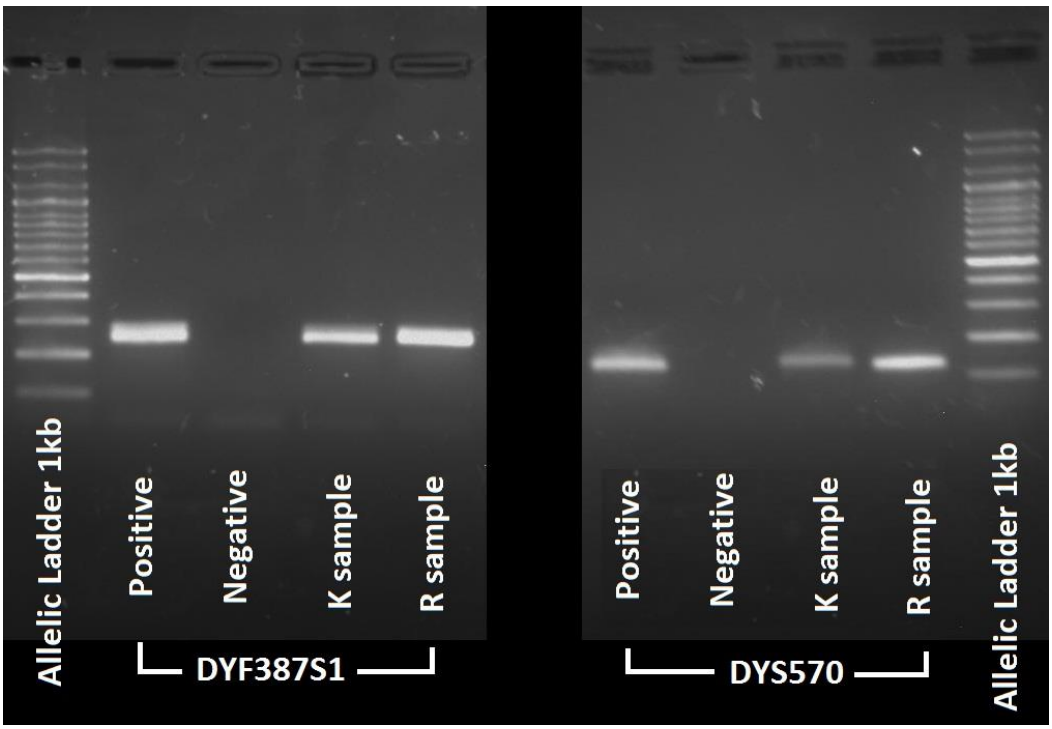

Figure 3.3: Agarose gel electrophoresis showing DYF387S1 and DYS570 redesigned primer sets amplification products using TD 65°-55° C cycling conditions.

3.2.1.2. Three Multiplexes

The PCR primers for the 13 loci were divided into three multiplexes according to Ballantyne et al. (2010), (Table 3.1). Primer concentrations were kept as the published ones; however instead of preparing a master mix using separate PCR buffer, $MgCl_2$, dNTPs and DNA polymerase enzyme, three different commercial master mixes were used in order to simplify the experiment by reducing preparation time and choosing the most efficient master mix for amplification of 13 RM Y-STR loci. These master mixes were ReddyMix™ (Thermo Scientific), AmpliTaq® Gold (Life Technologies) and Platinum® (Life Technologies). As per manufacturer recommendation, the volume of the master mix added to PCR reactions was half the volume (5 µl) of the total reaction volume for each master mix. All master mixes were compatible with touch down PCR cycling conditions and all annealing temperatures were within the recommended range of each master mix. The results (Figure 3.4) of this experiment were studied and compared. The specificity and intensity of PCR product amplified using Platinum® PCR Master Mix (Life

Technologies) was consistently better among the three master mixes, as it showed the brightest bands compared to other master mixes results which reflect the high amount of yield results out of PCR amplification. Hence it was chosen to be used in subsequent experiment for development of the multiplex PCR assay.

Table 3.1: RM Y-STR loci contained within the three multiplexes developed by Ballantyne et al. (2010).

RM1	**RM2**	**RM3**
DYF387S1	DYS518	DYF403S1ab
DYF399S1	DYS526ab	DYF404S1
DYS570	DYS626	DYS449
DYS576	DYS627	DYS547
		DYS612

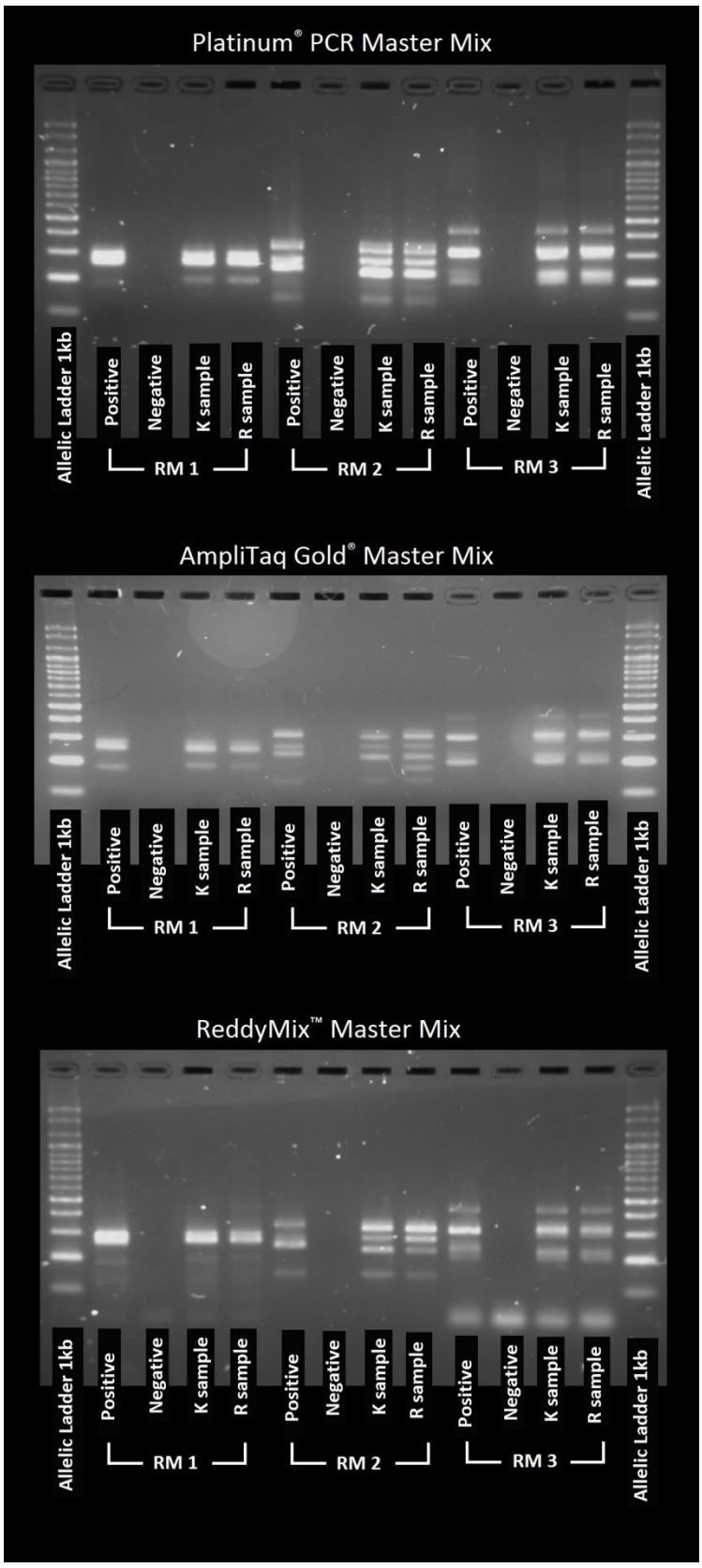

Figure 3.4: Agarose gel electrophoresis of PCR amplification products using ReddyMix™ Master Mix (Thermo Scientific), AmpliTaq Gold® Master Mix (Life Technologies) and Platinum® PCR Master Mix (Life Technologies) used to amplify 13 RM Y-STR in three different PCR multiplex reactions.

3.2.2. Labelled Primer Sets

Amplification products of 13 RM Y-STR loci using modified primer sets in singleplexes with TD 65-55 °C cycling conditions were analysed on ABI 3500 Genetic Analyser. This showed only expected peaks for the four samples; TaqMan® positive male control DNA and two unrelated male samples (K and R). 9947A female DNA control sample showed no amplification for all primer sets confirming the male specificity and appropriate primer design. The results of singleplexes confirmed the quality of primer sets in amplifying the target products using the new combination of 4-dye chemistry (Schubbert and Rittler, 2010).

Initially, PCR multiplex of 13 RM Y-STR loci was tested by combining all primers with equal concentrations, 0.13 µM, and amplified using TD 65-55 °C cycle condition. The resultant profile was not balanced and one drop-out was observed at DYS627, which was recovered after increasing the primer concentration to double, 0.26 µM. Then, all primer concentrations were optimised by holding DNA input constant at 1 ng in order to obtain a balanced profile of 13 RM Y-STR loci (Section 2.2.6.2; Table 2.7). However, DYF403S1b, with an average peak height ~500 relative fluorescent unit (RFU), was always below the average peak height ~4000 RFU among all other loci in the three samples, an example is shown in Figure 3.5.

DYF403S1a and DYF403S1b are amplified using the same primers set, hence increasing the concentration of the primers will result in unbalancing the DYF403S1a part of the locus with other loci in the profile. Since optimal results obtained from singleplex experiment on gel electrophoresis for DYF403S1 was at 58°C and the fact that all other loci showed a great level of amplification at 58°C based on gel electropherasis, PCR cycles were changed to standard PCR and performed at 58°C using 32 cycles. Resultant profiles were better in balancing DYF403S1b allele with an average height peaks ~1000 RFU compared to ~4000 RFU for all other loci. However, the average peak height of

allele at DYS526a has dropped down to ~500 RFU as shown in Figure 3.6. Again, DYS526a is amplified using same primers set as DYS526b and hence increasing the primer concentration will result in an unbalanced profile as the proportion between DYS526a and DYS526b will always remain the same. Therefore, in order to obtain balanced allele peak heights between DYF403S1b and DYS526a, a standard PCR cycle condition was performed over two stages at 58 °C and 55 °C with 32 total number of PCR cycles. Different numbers of cycles were tested at each PCR stage until the optimal profiles have been achieved. Finally, balanced electropherograms, for example Figure 3.7, were generated using optimal PCR amplification conditions (Section 2.2.6.3; Table 2.8). Peak height ratio between highest peak height at DYF403S1a and DYF403S1b peak height was improved when results from three samples were analysed and compared between to PCR cycles conditions (Figure 3.8).

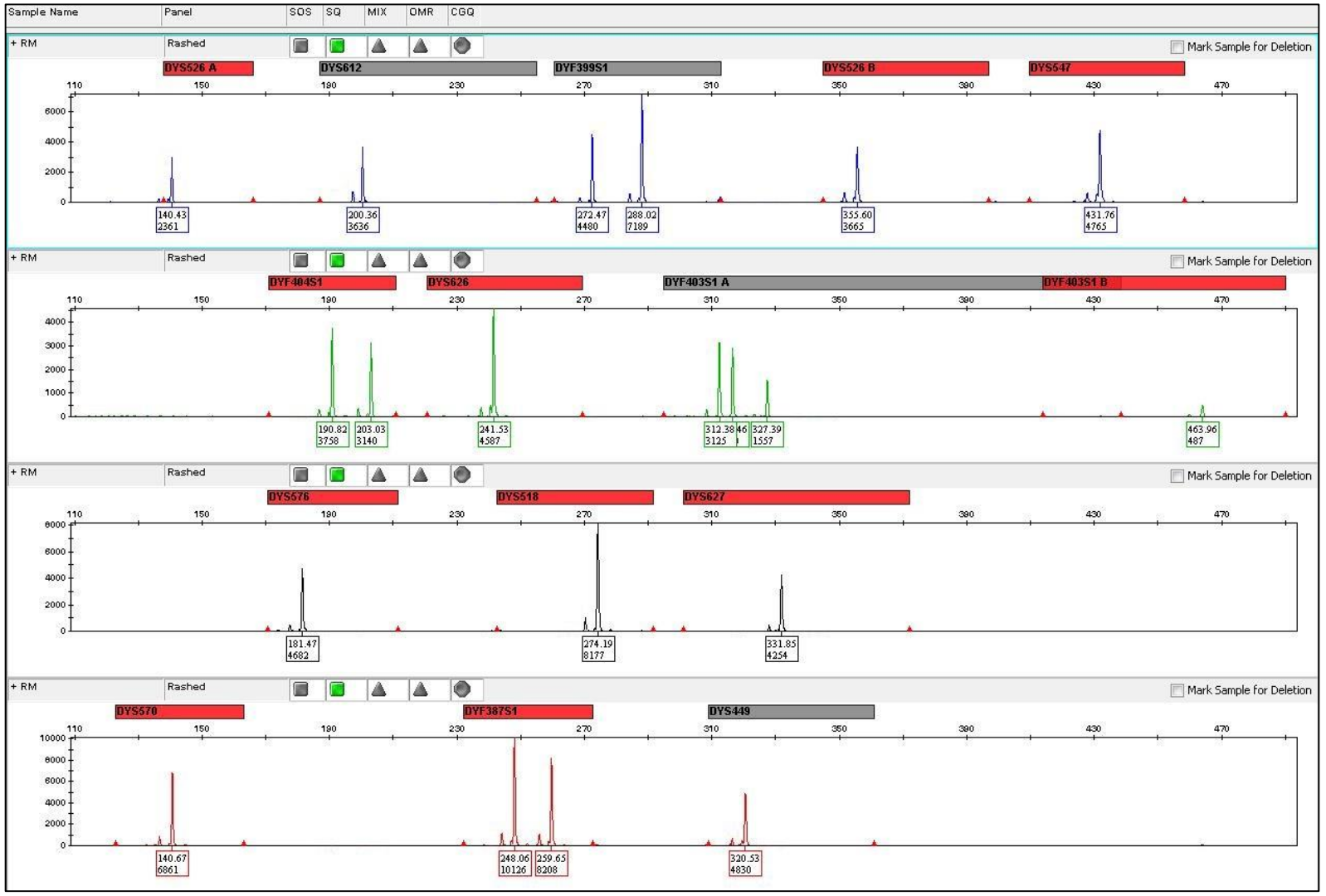

Figure 3.5: 13 RM Y-STR multiplex electropherogram (EPG) of TaqMan positive male control using TD 65-55 °C.

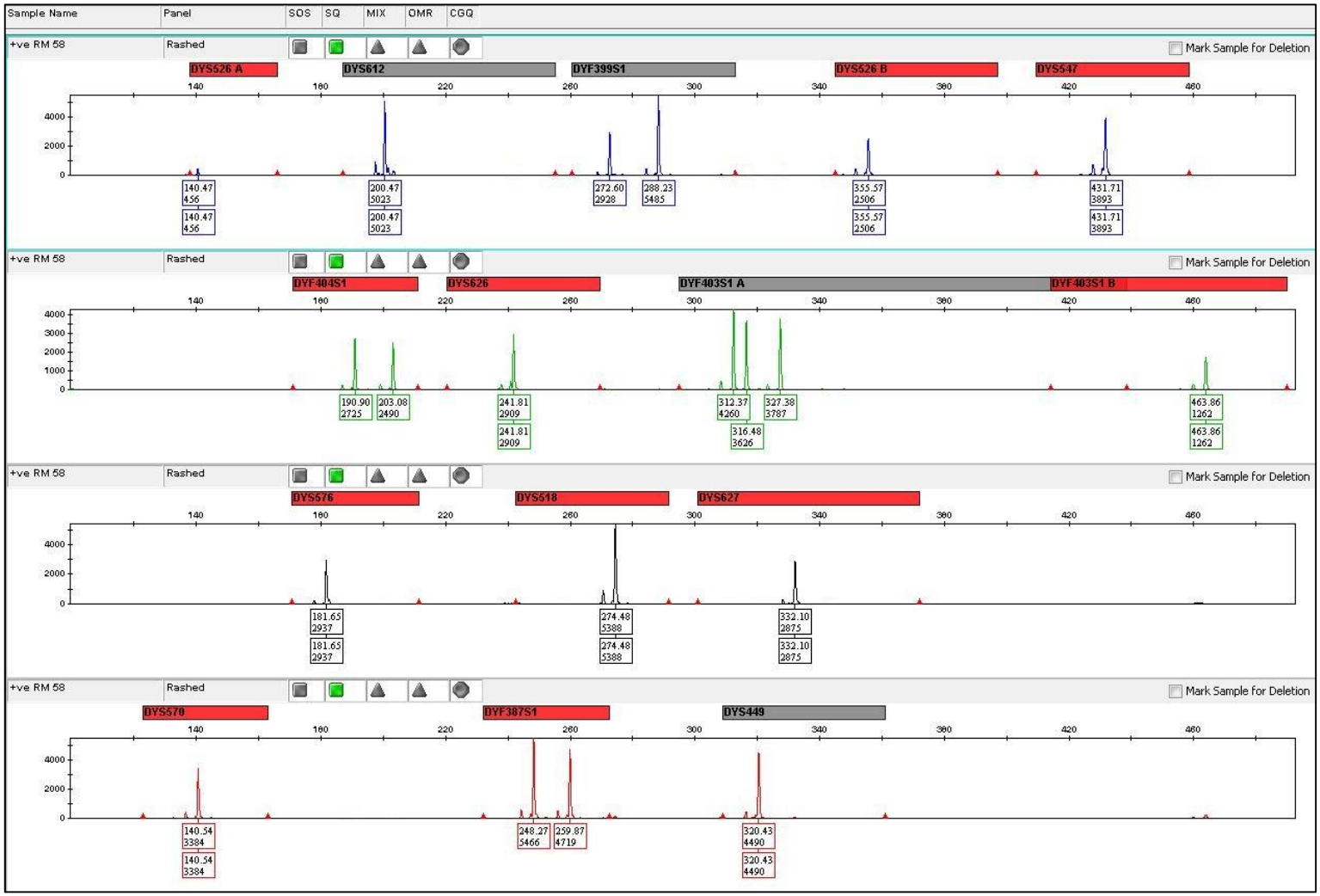

Figure 3.6: 13 RM Y-STR multiplex electropherogram (EPG) of TaqMan positive male control sample using standard PCR with annealing temperature at 58 °C.

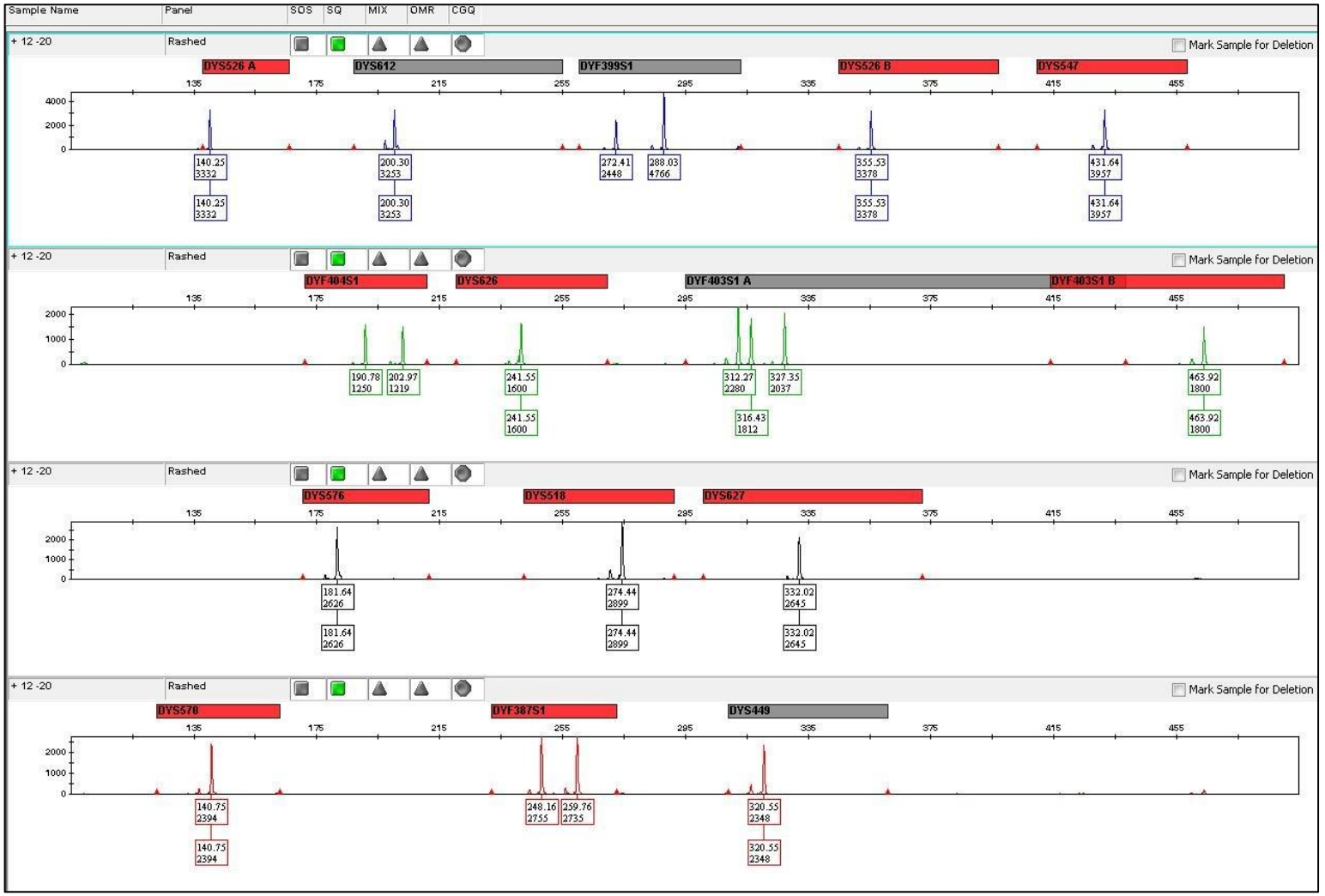

Figure 3.7: 13 RM Y-STR multiplex electropherogram (EPG) of TaqMan positive sample using standard PCR with two different annealing temperatures (58 °C and 55 °C).

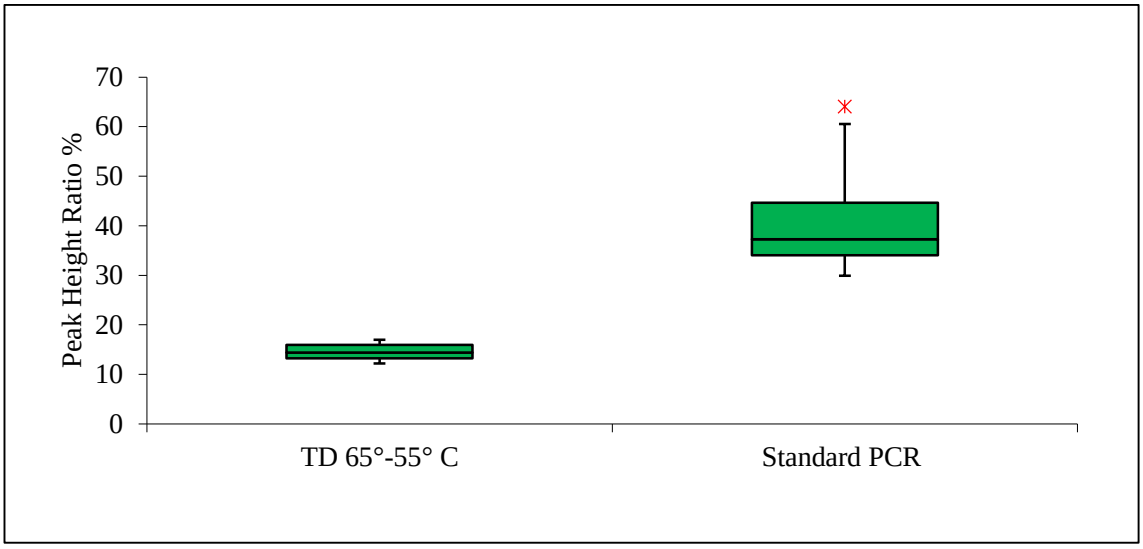

Figure 3.8: Box-plot illustrating the difference in peak height ratio at DYF403S1a/b locus between TD 65-55 °C and the optimised standard PCR cycle conditions. Intersecting line representing the median and red star showing the highest outlier point. This figure generated using data obtained from 3 different male DNA samples replicated 3 times. Upper and lower whiskers represent the greatest and lowest values excluding outliers respectively. Maximum outlier is presented in red star. Outliers indicating values more than 3/2 times of upper quartile or less than 3/2 times of lower quartile.

A comparison between three different PCR cycles conditions was done using three males DNA samples with an input of 1 ng of each in the amplification reaction (Figure 3.9). Peak highest using standard split PCR cycles conditions showed more consistent results compared to 32 PCR cycles at 58°C and TD 65-55 °C PCR cycles conditions.

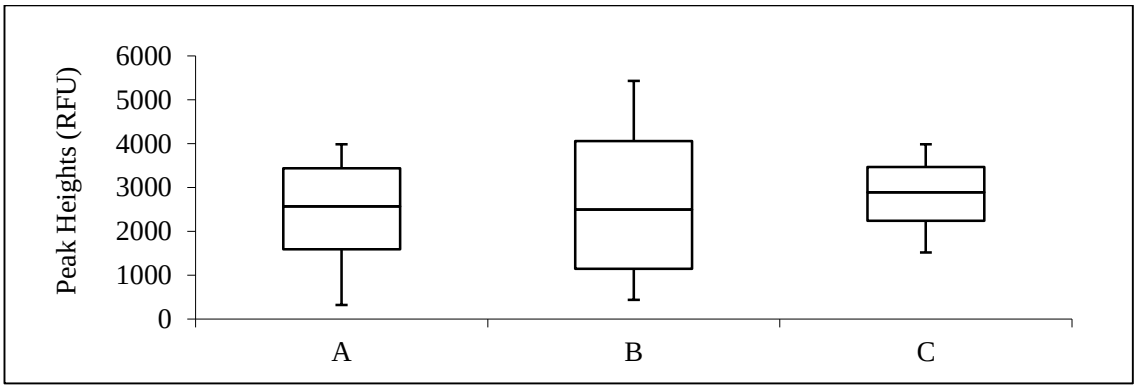

Figure 3.9: Box-plot illustrating the difference in peaks heights over all alleles in RM-Yplex assay at three different PCR cycle conditions. A: TD 65-5 °C; B: standard PCR with 32 cycles at 58 °C; C: standard split PCR with 12 cycles at 58 °C and 20 cycles at 55 °C. Intersecting line representing the median. Upper and lower whiskers represent the greatest and lowest values excluding outliers respectively.

3.2.3. Matrix Standard

A set of eight samples were prepared and amplified using optimised PCR reaction parameters, and then run on ABI 3500 Genetic Analyser using two different matrix standards, G5 and EF-01™. Each standard contains different type of fluorescence dye with slightly different absorption and emission of light values (Table 3.2).

Table 3.2: Absorption and emissions of fluorescent labels included in G5 and EF-01 matrix standards.

G5	Absorption (nm)	Emission (nm)	EF-01™	Absorption (nm)	Emission (nm)
FAM	494	525	**FAM**	495	520
VIC	538	554	**Yakima Yellow**	530	549
NED	546	575	**ATTO550**	554	576
PET	558	595	**ATTO565**	563	562

As a result, when using the G5 matrix there were many pull-ups, peaks appearing in the wrong dye having the same fragment size as parent peak detected in another dye, detected in both blue and red panels caused by peaks detected in green panels (Figure 3.10A and 3.10B). Also, there was same effect observed at red panel cause by peaks detected at yellow (Figure 3.10C). The same prepared plate was re-injected again using EF-01™ matrix standard. As a result, pull-ups peaks were completely disappeared (Figure 3.11 A, B and C).

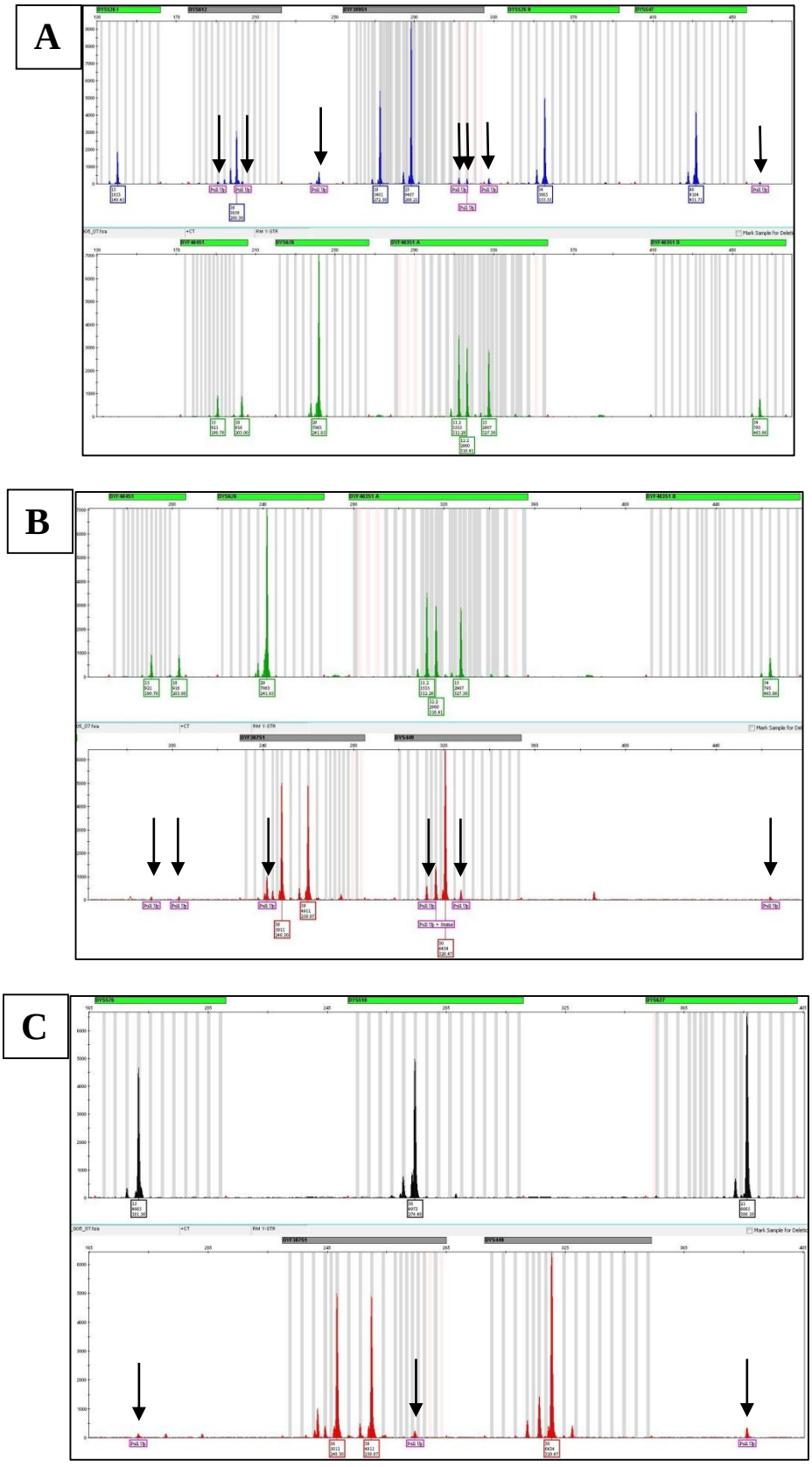

Figure 3.10: Electropherograms illustrating the pull-ups effect using G5 matrix standard in: (A) Blue panel, (B) Red panel cause by green panel, (C) Red panel cause by black panel. Pull-ups are indicated by pink boxes as well as black arrows.

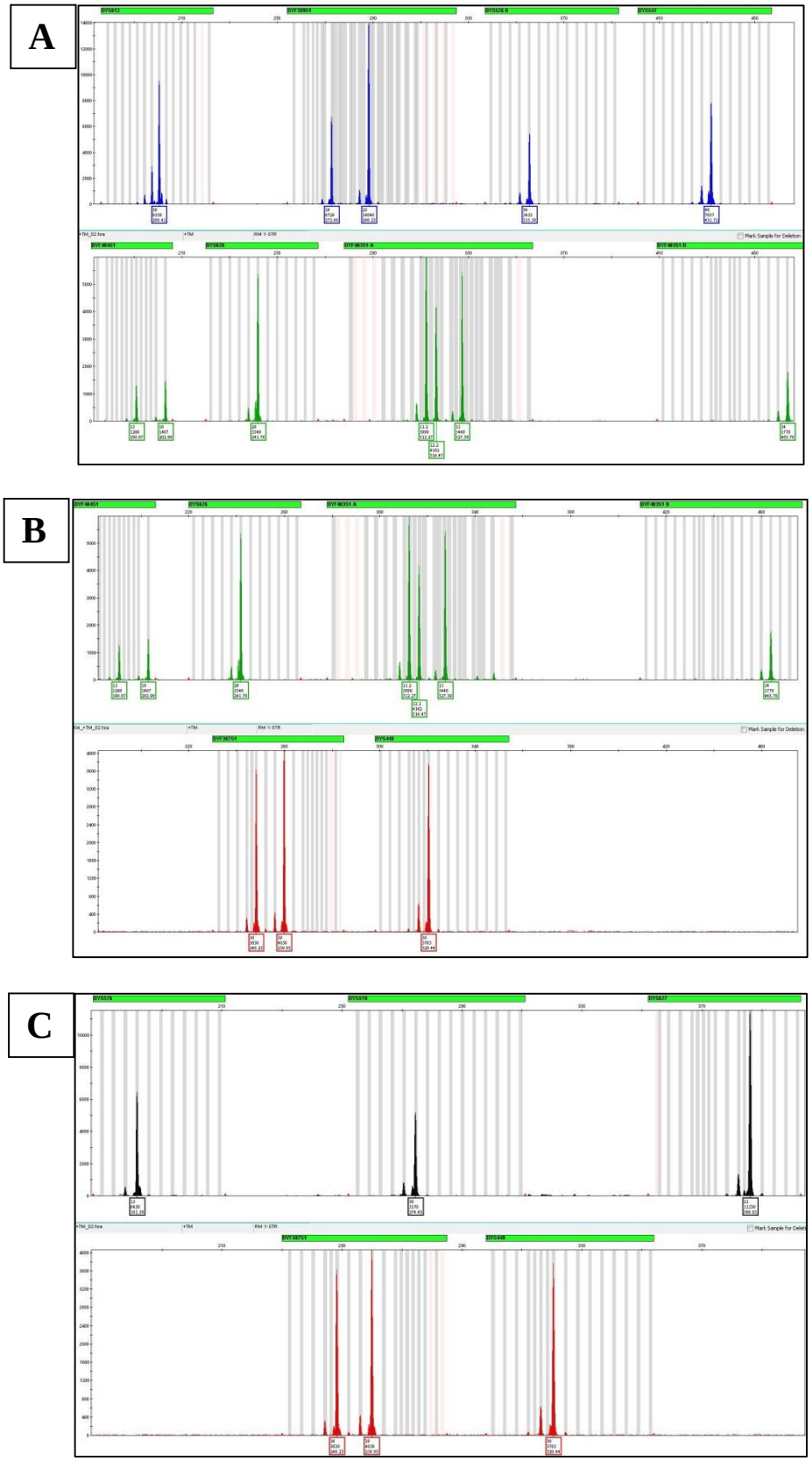

Figure 3.11: Electropherograms illustrating the pull-ups effect using G5 matrix standard in: (A) Blue panel, (B) Red panel cause by green panel, (C) Red panel cause by black panel.

3.2.4. RM-Yplex Electropherogram Profile

192 male DNA samples were amplified using an optimised PCR reaction conditions with constant male DNA template input of 1 ng per reaction, which have been quantitated using Quantifiler® Duo quantification kit (Applied Biosystems). Average peak heights ranged from ~1000 to ~4000 RFU (Figure 3.12). DYS526a resulted in relatively low average peak heights, with a maximum outlier around ~3000, compared to other loci in the same fluorescent dye panel, although the second part DYS526b was balanced through the 6-FAM fluorescent dye panel. DYF403S1a showed a high level of variation in peak heights compared to other loci in the assay, on the other hand, DYF403S1b recorded the lowest averaged peaks heights through all loci. Two peaks at DYF404S1 locus and DYS626 locus showed almost same averaged peak heights of ~1500 RFU. The complete yellow panel comprise of DYS576, DYS518 and DYS627 showed consistent average peak heights of ~3000 RFU. In addition, the red panel comprise of DYS570, DYF387S1 (two alleles) and DYS449 resulted in a balanced averaged peaks height as well.

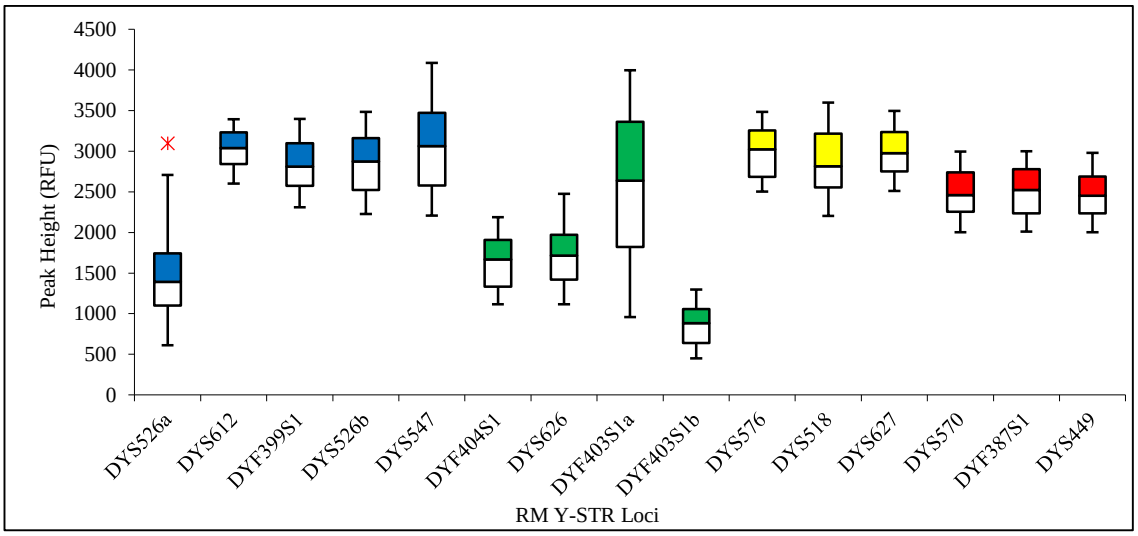

Figure 3.12: Distribution of locus peak height generated as a result of amplifying 192 male DNA samples using optimised RM-Yplex multiplex assay. Results are presented using box-plot, where the boxes for each locus represent the range of values from bottom quartile (25%) to the top quartile (75%) and the median is depicted an intersecting line within the boxes. Upper and lower whiskers represent the greatest and lowest values excluding outliers respectively. Maximum outlier is presented in red star. Outliers indicating values more than 3/2 times of upper quartile or less than 3/2 times of lower quartile.

Peak heights for multi-allelic loci, DYF387S1, DYF399S1, DYF403S1 and DYF404S1, were estimated in samples comprise typical number of peaks which are 2 peaks, 3 peaks, 3 peaks and 2 peaks respectively. However, samples in which fewer numbers of peaks appeared at each of the multi-allelic loci were analysed separately. The number of peaks detected at DYF399S1 varied from one to three peaks. Three peaks resulted in balanced peak heights with an average of ~3000 RFU. There was a very marked imbalance between DYF399S1 allelic peaks in most of the cases when only two peaks were detected. Although a couple of peak height outliers were detected, the majority of samples, in two allelic peak situations, showed the highest peak height is almost double the peak height of the shortest allelic peak (Figure 3.7 and Figure 3.13). Whereas when only one peak was detected, most of these samples showed a peak height of ~9000 RFU, though there were still some outliers which showed peak heights of ~6000 RFU and ~3000 RFU (Figure 3.13). The same pattern was observed with DYF387S1 and DYF404S1 as illustrated in Figure 3.14 and Figure 3.15 respectively. However, DYF403S1a showed no specific pattern and peak heights were completely random ranging between ~1000 to ~4000 RFU. Peak height ratio percentages were calculated for all multi-allelic markers, where typical number of alleles was resulted, by dividing the lowest peak height over the highest peak height multiplied by 100, results are shown in Figure 3.16.

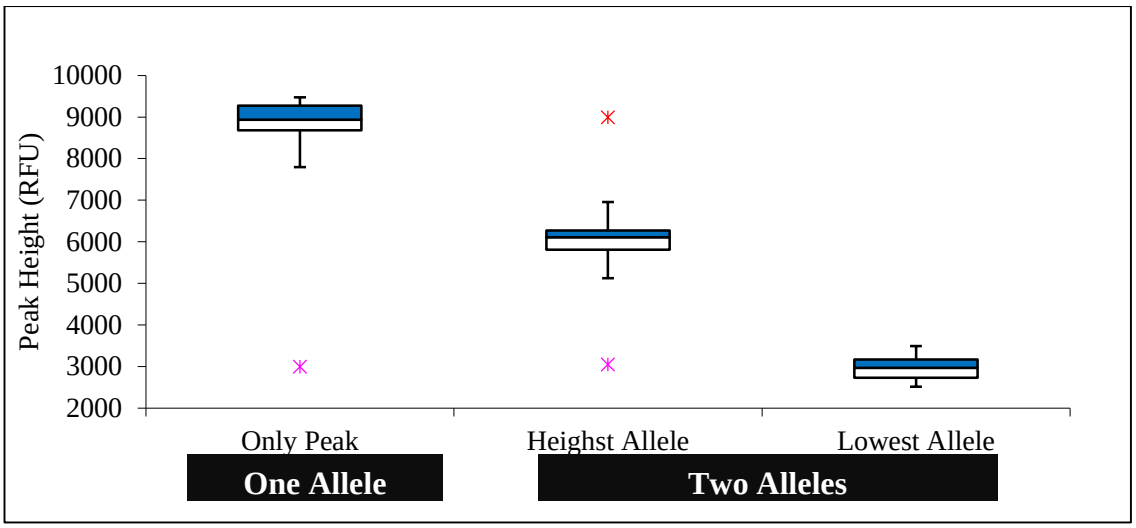

Figure 3.13: Peak heights at DYF399S1 locus with one allele and two allele's situations. Pink star represent the lowest outlier and red star represent the highest outlier.

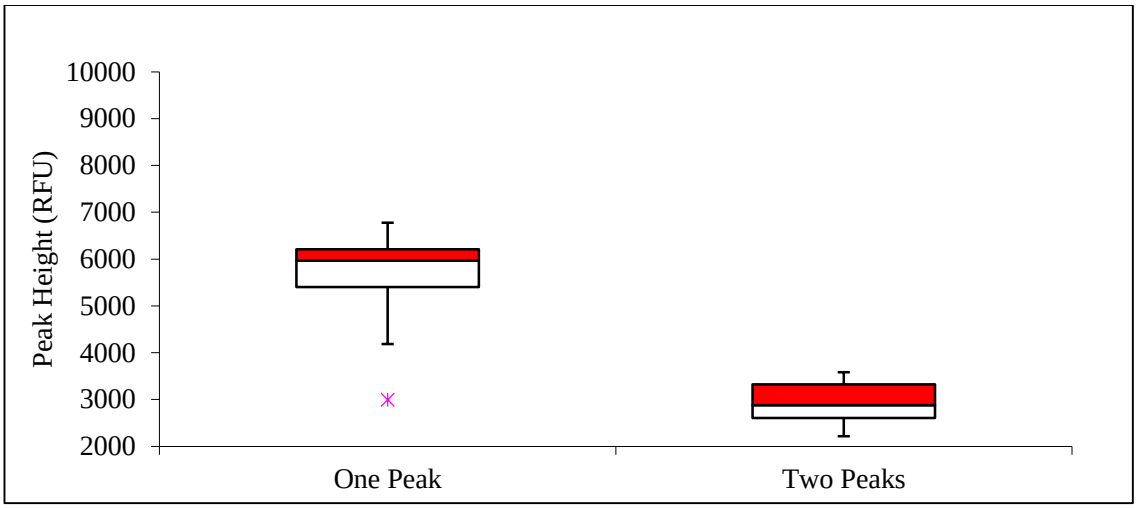

Figure 3.14: Peak heights at DYF387S1 locus with one allele and two allele's types of the resulted genotypes. Pink star represent the lowest outlier.

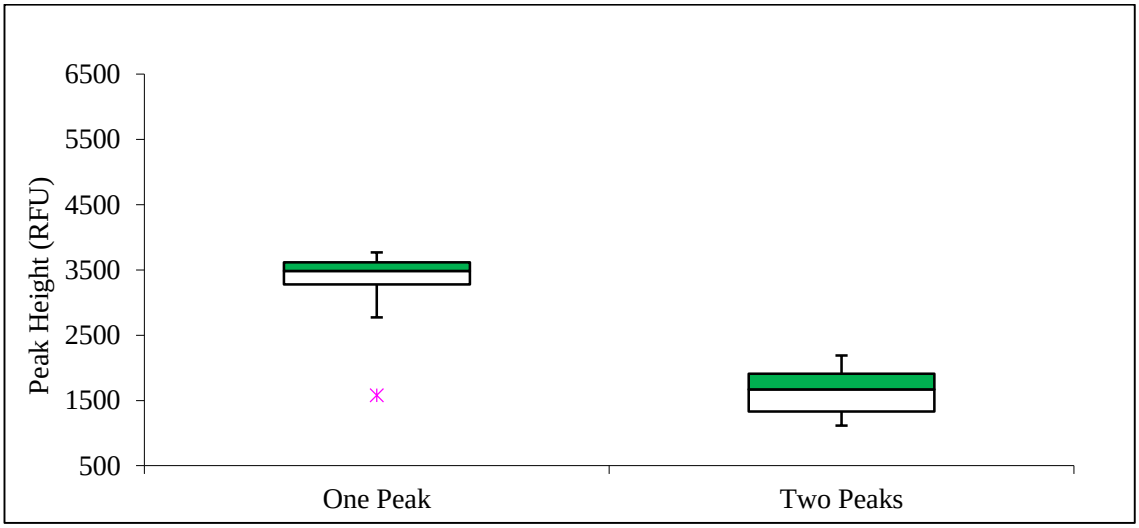

Figure 3.15: Peak heights at DYF404S1 locus with one allele and two allele's types of the resulted genotypes. Pink star represent the lowest outlier.

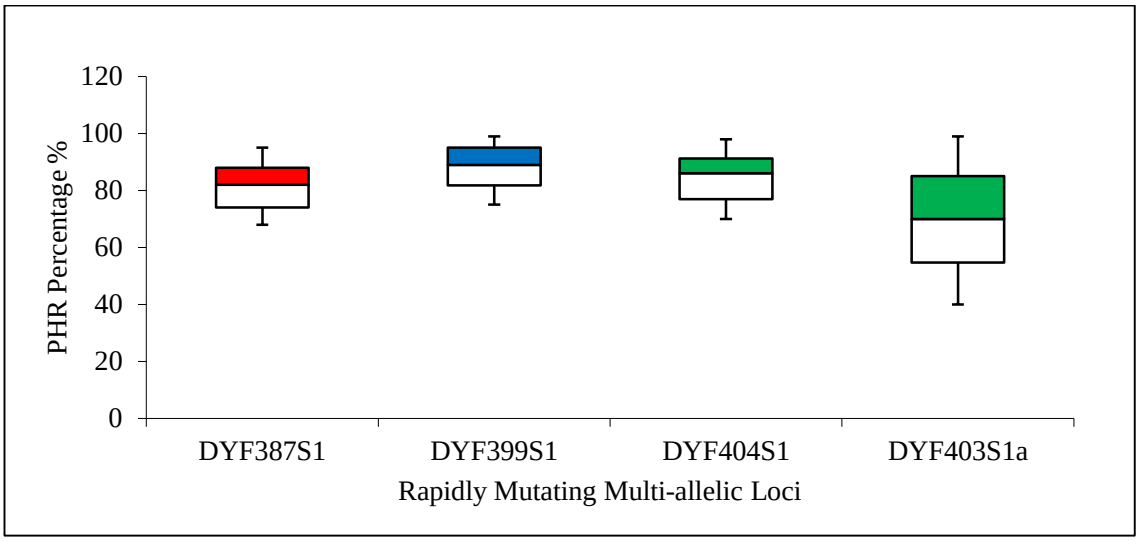

Figure 3.16: Peak height ratio at DYF387S1, DYF399S1, DYF404S1 and DYF403S1a loci in situations where 2 peaks, 3 peaks, 2 peaks and 3 peaks results at each locus respectively.

Overall, all profiles showed sharp peaks and exhibited stutter peaks with different heights. Minus stutter was most commonly observed at each marker, where the resultant peak is one repeat shorter than parent allele, whereas plus stutter peaks were rarely observed in generated profiles. At DYS612, minus one and minus two stutter peaks were more pronounced than all other peaks. A total of four different female DNA samples showed no amplification at all markers. Detailed analysis of stutter and specificity will be discussed in developmental validation chapter.

3.2.5. Allele Size Range Determination

A comprehensive evaluation of the data at 13 RM Y-STR loci was carried out to determine the number of alleles for each locus and the size, in base pairs, of each allele. Moreover, the precision of measurements, standard deviation from the measured mean allele size, was assessed for each locus. The allele size ranges for all loci were estimated using 192 samples set, which has been analysed earlier in this chapter to investigate allelic peak heights. In addition to the extensive literature searches (Butler et al., 2002, Henson,

2005, Hanson and Ballantyne, 2007a, Hanson and Ballantyne, 2007b, Rodig et al., 2008, D'Amato et al., 2009, Vermeulen et al., 2009, D'Amato et al., 2010, Hedman et al., 2011).. Table 3.3 describe the allele sizes ranges found in this study in comparison with literature values. Genotyping at this stage was dependant on the sizes of the detected allele where alleles that were not presented within the determined allele ranges were considered off-ladder products (OL) allele or microvariant. Allele sizes within calculated standard deviation were grouped together and given a designation accordingly.

The repeat number ranges were determined using the GenBank® reference sequences for each locus. For example, results of blasting DYS570 primers retrieved a reference sequence sized of 133 base pairs (bp) which contains a simple tetranucleotide repeat motif consisting of 17 TTTC repeats. So all amplicons with less or more nucleotides calling for DYS570 were compared to the reference sequence in the literature and named accordingly. The same strategy was used with all other loci. However, some markers are expected to have microvariants such as DYS576 and DYS627. For example, using primers for the locus DYS576, the GenBank® reference sequence retrieved contained a simple tetranucleotide repeat motif consisting of 16 AAAG repeats and sized 191 bp. Hence, detection of amplicon sized 196 bp means an addition of one repeat motif and 1 extra nucleotide. Then such an allele will be called 17.1 because of the microvariant occurrence as it is illustrated in Figure 3.17 (A). Similarly with DYS627, the reference sequence in GenBank® represented an amplicon sized 337 bp containing a complex tetranucleotide repeat motif consisting of 18 AAAG repeats. Detecting of amplicon sized 335 bp and 340 bp will be considered as microvariant alleles 17.2 and 18.3 respectively as it is shown in Figure 3.17 (B).

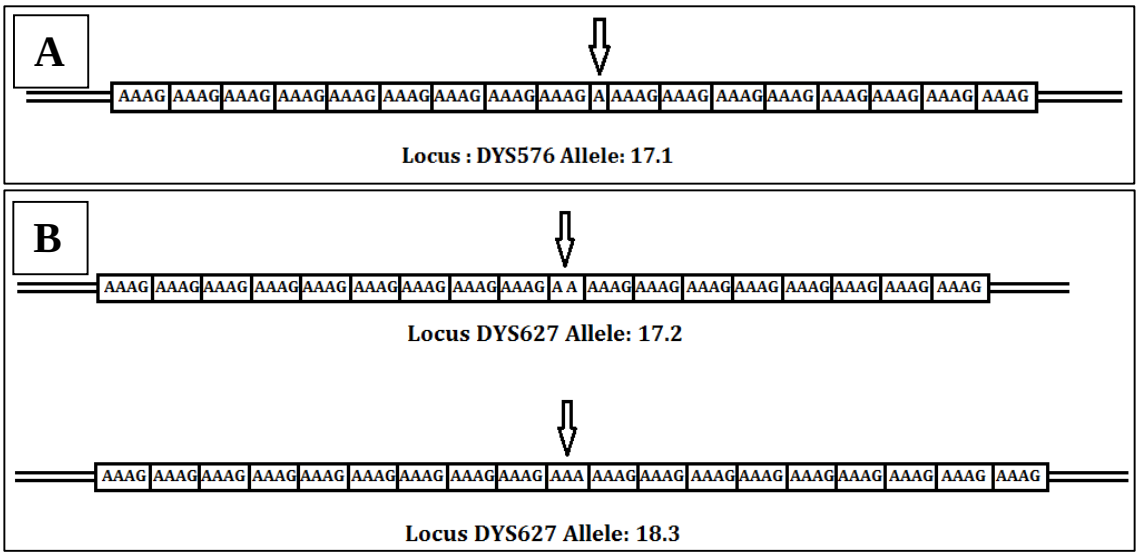

Figure 3.17: Illustration of microvariant alleles occurred at DYS576 locus (A) and DYS627 locus (B).

Table 3.3: Alleles and estimated size ranges in literature and in this project modified from Ballantyne et al. (2012). Red highlights indicating overlapping fragment size of different loci.

Locus	Repeats number		Allele size in base pairs	
	Literature Alleles (repeats)	UAE population alleles (repeats)	Literature Allele sizes (bp)	UAE population allele sizes (bp)
DYF387S1	28-38	32-44	241-281	232-280
DYF399S1	10-23	15-32	261-313	256-320
DYF403S1a	12-39	7-22	310-438	280-355
DYF403S1b	40-59	41-57	414-490	411-475
DYF404S1	10-20	11-18	171-211	175-203
DYS449	24-37	28-39	309-361	300-344
DYS518	23-35	33-47	243-291	**255-309**
DYS526a	10-17	11-18	138-166	132-160
DYS526b	29-42	30-43	345-397	339-391
DYS547	36-48	41-54	410-458	403-455
DYS570	10-21	13-23	246-286	116-156
DYS576	13-23	13-23	170-210	170-210
DYS612	14-31	29-43	187-255	178-221
DYS626	11-23	24-35	221-269	221-265
DYS627	10-24	12-19	301-372	**301-347**

3.2.6. Overlapped Markers

There was an overlapping allele size between DYS518 and DYS627 loci, which have been placed in the same panel using same fluorescent dye label, detected in the 192 sample set analysis, which has not been reported yet in the literature to date. The overlapped size length was 8 bp long (Figure 3.18; Table 3.3). Two markers were reconsidered in order to replace them within the multiplex without overlapping in fragment sizes occurring. A new set of primers were designed for DYS627 for such purpose. As it was the case for other redesigned primers, the latest information about chromosome homology or polymorphic nucleotides in the primer binding regions was applied to avoid regions homology to DYS627 locus or impact PCR amplification. Both previously designed and newly designed primer sequences are presented in Table 3.4.

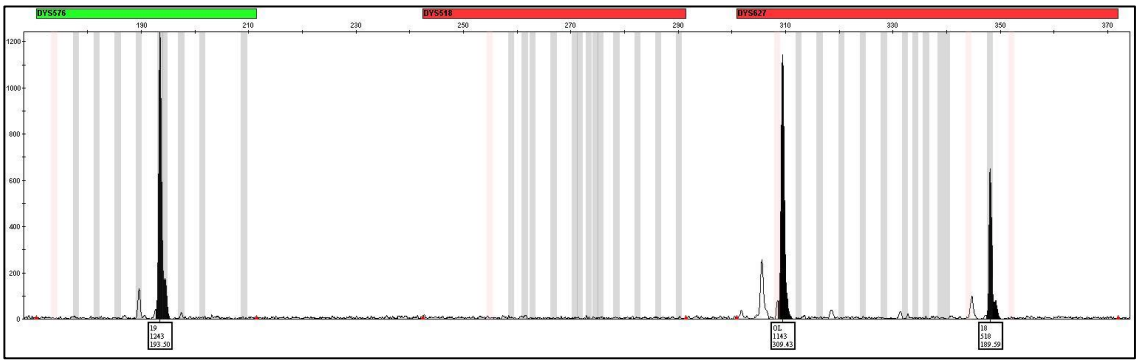

Figure 3.18: Allele fragments sizes overlapping between DYS518 and DYS627 loci in RM-Yplex assay.

Table 3.4: Previously published and redesigned primer sequences for DYS627 marker.

	Forward Primer	**Reverse Primer**
Published	CTAGGTGACAGCGCAGGATT	GGATAATGAGCAAATGGCAAG
Redesigned	GATGGGGAGGTTGCAGTAAG	TCTGTGAGTCCACTGGAGACC

3.2.7. DYF403S1a Locus Genotyping

There was a notable number of samples exhibiting split peaks at DYF403S1a locus. The amplification of these samples was repeated in multiplex as well as in singleplex reactions. However, peaks were still detected as split peaks. A singleplex with extra extension time showed no improvements. Since the difference between the pseudo-split peaks was consistently ~0.6 base pair in all the samples under investigation, a concern has been highlighted about whether these are two real peaks or not and whether this defect resulting out of the high resolution feature of the polymer being used (POP-6). In order to test this hypothesis, eight samples were amplified and then run on ABI 3500 Genetic Analyser using POP-6 with 50 cm capillary array as well as using POP-4 with 36 cm capillary array. As a result, pseudo-split peaks disappeared and instead a noticeable higher peak was observed (Figure 3.19). In deeper investigation of each allele's nucleotide sequences, there was a significant difference between nucleotide sequences of each allele copy at this locus. Taking into consideration three facts, first there is a negligible mobility difference between each of the three possible nucleotides, which makes up the DNA sequence. Such difference would hardly be observed in short fragments. Second, allele copies at DYF403S1ab locus are more than 300 base pairs long. Third, it has been proven previously POP-4 polymer is relatively lower in resolution than POP-6 polymer. Therefore, it is most likely the resulted pseudo-split peaks are representing two different copies of alleles at DYF403S1a given that changing the polymer to POP-4 resulted in only one peak for the same PCR product tested. These results were consistent with all samples used in this experiment (Figure 3.19).

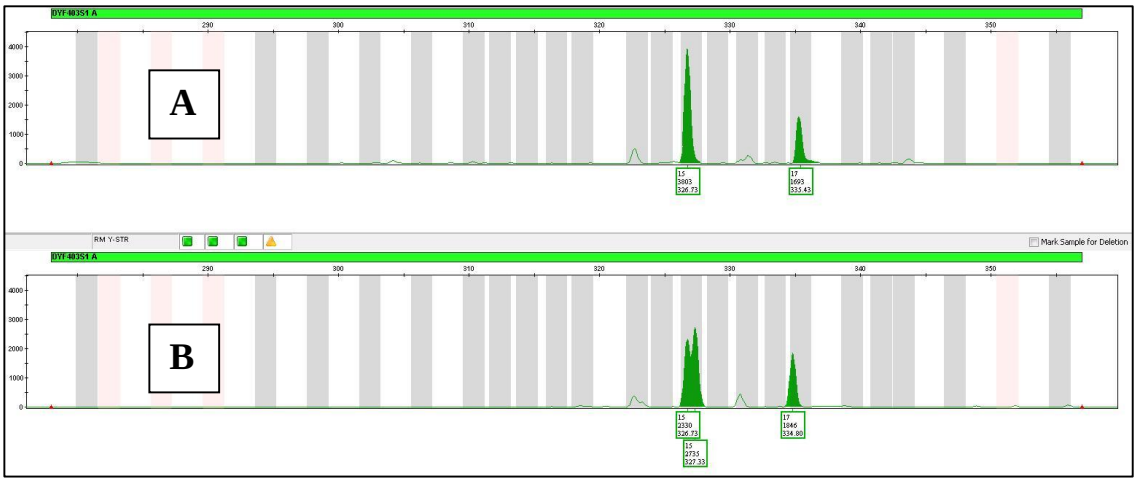

Figure 3.19: Amplified product of male DNA samples run on ABI 3500 Genetic Analyser using both (A) POP-4 and (B) POP-6.

3.3. Discussion

The PCR multiplex development strategy (Figure 3.1), has been proven effective in developing and optimising a multiplex assay which can amplify 13 RM Y-STRs loci simultaneously. Both published and redesigned primers have shown a high level of efficiency as they have proved, in singleplexes, the ability to efficiently amplify a targeted region of DNA over different range of annealing temperatures. These capabilities of primers along with the DNA polymerase enzyme efficiency and robustness have facilitated the development of the RM-Yplex multiplex assay. All three different self-contained DNA polymerase master mixes, ReddyMix™ Master Mix (Thermo Scientific), AmpliTaq Gold® Master Mix (Life Technologies) and Platinum® PCR Master Mix (Life Technologies), were able to amplify 13 RM Y-STR loci divided into three different multiplexes. However, only Platinum® PCR Master Mix was able to amplify the 21 allele fragments which make up the 13 RM Y-STR loci in single 15 µl PCR reaction simultaneously. All other enzymes showed no amplification at all. The deficiency of the other enzymes in this experiment could be due to the high number of different fragments of DNA needed to be amplified in one reaction. Other parameters could have played a role in the deficiency of the reaction such as $MgCl_2$ concentrations and enzyme

concentrations. Since Platinum® PCR master mix has shown best results with fewer preparation steps for PCR reaction, it has been chosen to be used for the rest of the optimisation experiments.

3.3.1. Balanced Profile

Primer concentration had a great role in balancing the peak heights of all loci apart from DYF403S1a/b and DYS526a/b. This is expected because of the fact DYF403S1a/b and DYS526a/b are multi-allelic loci and hence varying the concentration will result in same peak height ratio across each locus. There are many factors which can affect the multiple alleles amplification and hence results in different peak heights such as the number of copies of that region available in the DNA template, primers sequence mismatch in some of the binding site and the size of the fragment being amplified. Although changing PCR cycles conditions has shown an improvement in balancing peaks across the profile, intra-locus peaks height imbalance was seen most of the time.

DYF403S1 locus typically has 4 alleles which could be separated according to allele sizes. DYF403S1a showed three alleles with overlapping sizes in the range of 280 to 355 bp, whereas DYF403S1b showed only one allele in the range of 411 to 475 bp. Although, DYF403S1a has imposed a large variation of peak height ratio, DYF403S1b was most of the times unbalanced with the average peak height in DYF403S1a. This is most likely because of the fact that DYF403S1b is a relatively large DNA fragment and PCR reaction will always shows preferential amplification to short fragments of DNA (Schoske et al., 2003). Although, DYF403S1a/b was studied previously to a lesser extent, the situation of imbalance was always showing (Victoria, 2010, Ballantyne et al., 2012, Ballantyne et al., 2014)

DYS526a/b locus resulted in almost the same imbalance issue between DYS526a and DYS526b. However, in this case primer concentration helped only the largest fragment, DYS526b, to impose an inter-locus peak height balance through 6-FAM dye panel. Whereas the shortest fragment has shown a reduced peak height most of the time compared to the other peaks in the blue panel even after the PCR cycle optimisation. The multi-allelic nature of this marker is not presented because of the duplication on Y chromosome however; the shortest fragment is being amplified because of the duplication of the forward primer attachment site within the DYS526b marker itself (Figure 3.20).

TCTGGTGAACTGATCCAAACCTTTACTTATTTTCTTCCTTTTTTCCTCTCTCCTTCCCTCCC
TCCCTCTTTCTTCCTTTCTTTCTTCCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT
CTTTCTTTCTTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCTTTCCTTCTTTCCTT
CATTTCTTCTTTCTTTCTCTTTACCTTCTGGGAAACTGATCCAAACCTTTACTTACTTTCTA
TTTTTCTCTATTTTTCTTTCTCTATTTTTTTCTCTCCTTCCTTCCTTCCTTCCTTCCTTCCTTC
CTTCCTTCCTTCCTTCCTTCCTTCCTTCTCTTTACCTTCTGGCGAAGTAACCC

Figure 3.20: DYS526ab locus sequence retrieved from GenBank®. Sequence highlighted in green colour is forward primers attachment site and sequence highlighted in red is reverse primers attachment site, while blue highlighted sequence colour represents a mismatch in primer sequences in the second binding site (underlined).

The ambiguous results of the peak height ratio between DYS526a and DYS526b could be regarded to the fact that there is a mismatch in two nucleotides within the secondary annealing site at DYS526a (Figure 3.20) which prevent primers from functioning efficiently at this binding site. However, there were some samples which have shown balanced peaks height between DYS526a and DYS526b which suggest that this mismatch might not be consistent through all the samples, the reason why random results were obtained regarding peak height at DYS526a.

3.3.2. Multi-allelic Loci

In order to balance the blue panel which include the multi-allelic marker DYF399S1 the sum of peak heights were averaged over the typical number of expected alleles at this

marker which is 3 alleles. This average was then considered for balancing the panel. When two peaks are resulted at DYF399S1 there always has been intra-locus imbalance with 2:1 or 1:2 ratios. This could be considered as pseudo-homozygous which means, it is more likely that there are two alleles at this locus having the same fragment size which then resulted in doubling-up the peak height of the other distinct allele in the same locus. Also, it has been realised when there is only one allele detected at DYF399S1, this peak will always be as three times higher than typical single peak height in three peaks pattern. Although there were some outliers, the heights of peaks were always showed consistent results by keeping the male DNA template input constant. In addition, DYF387S1 and DYS404S1 have shown the same pattern in the cases when only one or two distinct alleles resulted. Although these markers have shown high level of consistency, DYF403S1a locus have shown completely random peak height ratio and hence number of alleles, when less than three distinct peaks are presented, could not be determined. This also can introduce an issue in determining the true allele in case where the peak height ratio is very low between the peaks at this locus. Moreover, such issue will add more complication to the already existing with the multi-allelic Y-STR markers in cases when there is more than one male DNA source contributing to the generated profiles. Therefore, it is highly recommended to exclude this marker when analysing mixtures with more than one male contributor.

3.3.3. Matrix Standard

Although, a full profile was detectable using G5 matrix standard, there was a consistent pull-ups peaks presented in blue and red panel caused by Yakima-Yellow dye in the green panel in addition to the pull-up peaks appeared in the red panel caused by the ATTO550 dye in the yellow panel. It has been known that these peaks usually appear because of either a high amount of DNA template was amplified or there is an overlap between dyes used in the multiplex assay (Gill et al., 1996, Gill et al., 1997). The presence of these

peaks might cause interpretation issues specially when there are more than one male individual contributing to the source of DNA in forensic DNA profiling. These peaks could also cause peaks height imbalance as well as increasing stutter peak heights. In order to maximise the efficiency of the assay, EF-01™ matrix standard has been calibrated on the ABI 3500 Genetic Analyser before running the same amplified product tested with G5 matrix standard. The results confirmed the source of pull-up peaks was the matrix standard since all pull-ups peaks were disappeared in the run using EF-01™ Matrix standards have been made specifically for the fluorescent dyes used in the developed multiplex assay.

3.3.4. Allele Size Range Determination

When a new multiplex assay is being designed it is essential to determine the allele ranges in all populations and therefore taking this into consideration in order to locate the markers within a specific size position. Since it is impossible to profile all populations in the world, randomly selected samples are usually used to represent the ranges of allele sizes available at each marker in addition to the literature investigation. The allele determination experiment showed all marker positions were appropriate within the assay apart from DYS518 and DYS627 which have shown an overlapped DYS518 allele size calling at DYS627 marker range. Amplifying the samples in a singleplex was used to confirm the overlapping issue and assign the peak to a DYS627 marker. Although, these markers were studied in separate multiplexes in the past, these alleles were not detected in the analysed sample sets previously. Since allele size overlapping events within the same dye panel is problematic for interpretation of the generated profiles, it was important to allow some extra room for undiscovered alleles between the two markers by relocating DYS627 marker. The newly designed primer set showed male specific amplification, no cross reactivity in the multiplex and no unspecific amplifications. Overall, there is always extra rooms for at least two novel alleles between any two markers in the assay, making

this assay useable for worldwide populations. To summarise, a multiplex assay amplifying 13 RM Y-STR loci simultaneously, using PCR mediated analysis and capillary electrophoresis technology, has been developed and optimised successfully following a logical strategy. This multiplex assay was given a unique name “RM-Yplex” to be differentiated from other STR assays developed previously. An alternative four fluorescent dye combination proven efficient as the commonly used ones in the field in addition to its compatibility with currently used instruments and methods of DNA profiling analysis. This developed multiplex will need to be developmentally validated and tested in forensically relevant applications for suitability.

4. Development of Allelic Ladder for RM-Yplex

4.1. Overview

The allelic ladder is a balanced mixture of all common alleles found in human populations in the loci included in a multiplex STR assay. Employment of a sequenced allelic ladder for the accurate determination of an unknown allele is an established standard (Gusmão et al., 2006). It had been established early on that DNA fragment migration during electrophoresis is dependent not only on the length of the DNA fragments in the samples but also on the nucleotide sequence of the fragment (Frank and Köster, 1979). During capillary electrophoresis, factors such as the polymer being used, type of buffer, the voltage of electrophoresis and the coupling of a fluorescent molecule to the sample for detection purposes affect the migration of DNA fragments (Puers et al., 1993). Therefore, the use of allelic ladder is required in STR based DNA profiling to avoid the potential of a shift during migration of marker fragments as a response to changing factors and electrophoretic conditions. As allele fragments in an allelic ladder and alleles amplified in a sample have the same length and nucleotide sequence, both will migrate similarly during a capillary electrophoretic run. Hence assignments of unknown alleles within the samples could be concluded confidently (Schumm, 1997).

Genotyping studies using only internal lane standards (ILS) have been conducted considering the high level of precision of current genetic analysers and ILS (Butler et al., 2002, Schoske, 2003, Schoske et al., 2003). Since currently tetranucleotides STRs are most commonly employed in forensic DNA profiling, the size variation between adjacent full alleles is 4 bp. It has been shown that a particular ILS results in precise allele sizing with a maximum variation not exceeding +/-0.5 bp, the level which has been determined to be acceptable in genotyping analysis (Frank and Köster, 1979, Schoske, 2003).

In forensic DNA profiling both accuracy and precision are needed in sizing the alleles. ILS can size the alleles precisely but not accurately. A sequenced allelic ladder ran in the same run as amplified samples adds accuracy in genotyping the unknown alleles. The

allelic ladder also adds a great precision in confirming the microvariants which can be as close as 1 to 3 bp in tetranucelotide STR loci (Puers et al., 1993), and also out of marker (OMR) range alleles (Butler et al., 2012). Microvariants can be included within a ladder but it is not always possible to add all of them. Allelic ladder alleles are generated by amplifying them using the same primers sets and DNA polymerase enzyme used for STR analysis. Hence, it provides an accurate reference for each of the common alleles being analysed (Figure 4.1). RM Y-STR markers that were multiplexed in the RM-Yplex assay included highly diverse markers which have many microvariants such as DYF399S1 e.g. alleles 20.1, 20.2 and 20.3 (Ballantyne et al., 2012).

The objectives of the studies presented in this chapter included:

1- Identifying and sequencing alleles of the 13 RM Y-STR loci included in the RM-Yplex assay.
2- Developing a balanced allelic ladder of 13 RM Y-STR loci included in the multiplex assay.
3- Assessing the precision of the developed allelic ladder.
4- Comparing the developed sequenced allelic ladder with those developed by other research groups for RM Y-STR loci.

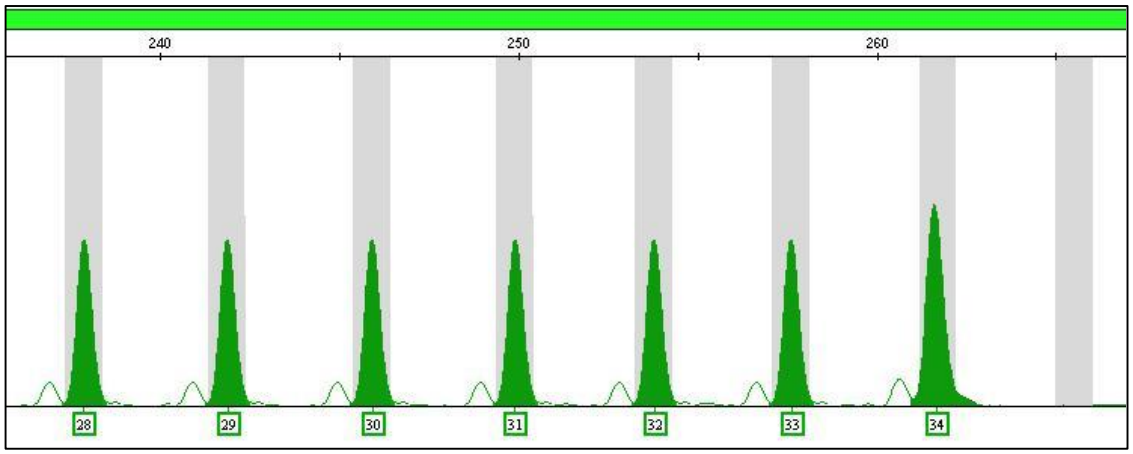

Figure 4.1: In-house developed DYS626 locus allelic ladder comprising alleles 28 to 34. The alleles had 28 to 34 repeats and were called in accordance with ISFG recommendations in this regard (Ballantyne et al., 2010).

4.2. Results

4.2.1. Allelic Ladder Development

The strategy followed in developing an allelic ladder has been proven effective in producing an allelic ladder for all loci included in the RM-Yplex multiplex assay developed earlier which can be used for the analysis. Electrophoresing 1 µl of the co-amplified alleles resulted in sharp peaks with slightly different peak heights (Figure 4.2), even though an equal amount of male DNA template input of each samples was used. For multi-allelic markers, samples were chosen carefully in order to avoid adding the same copy of an allele twice; however, where it was needed a sample with a single allele at this marker was used for balancing purposes. An example of multi-allelic locus-specific allelic ladder for DYF387S1 is presented in Figure 4.3. As an example, part A of Figure 4.3 shows the initial results of co-amplification of different alleles using equal amount of DNA template (5 ng). Part B shows the adjusted locus-specific allelic ladder by increasing the amount of DNA sample representing allele 34 to the double (10 ng) and reducing the initial amount of DNA sample representing allele 38 to the half (2.5 ng). This strategy was followed for all other multi-allelic loci whereas the single-allelic loci showed always intra-locus allele's peak heights balance using equal amount of DNA template.

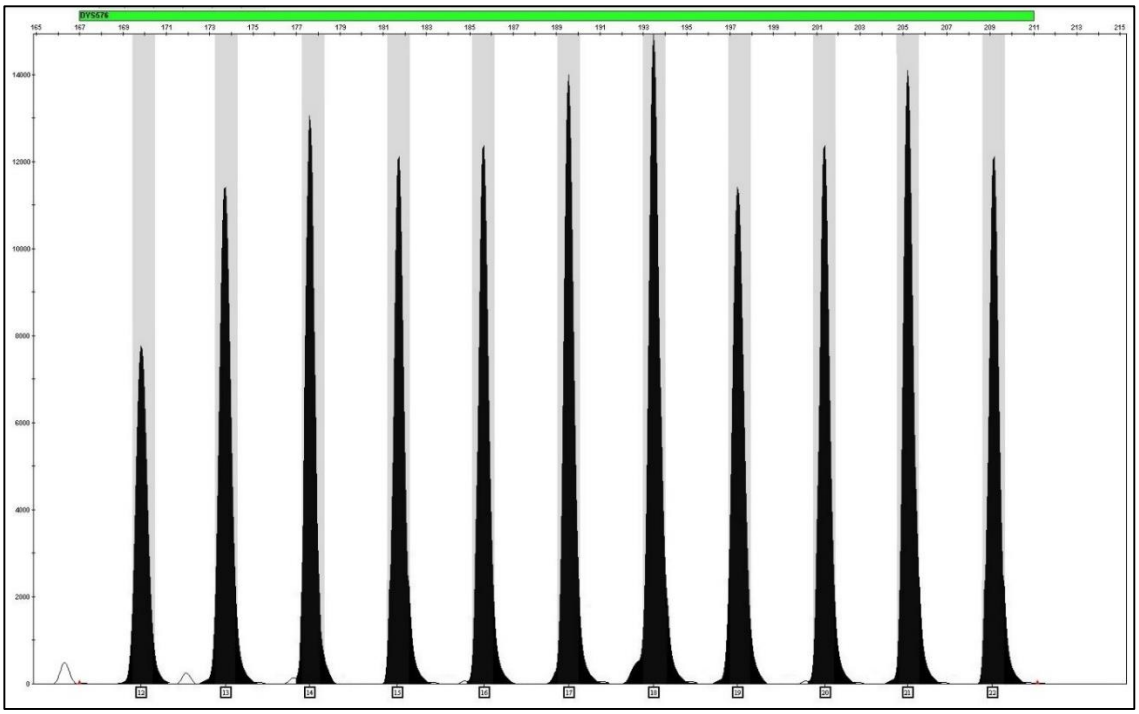

Figure 4.2: DYS576 locus-specific allelic ladder demonstrating results of co-amplifying equal amount of male DNA templates, each with different allele size.

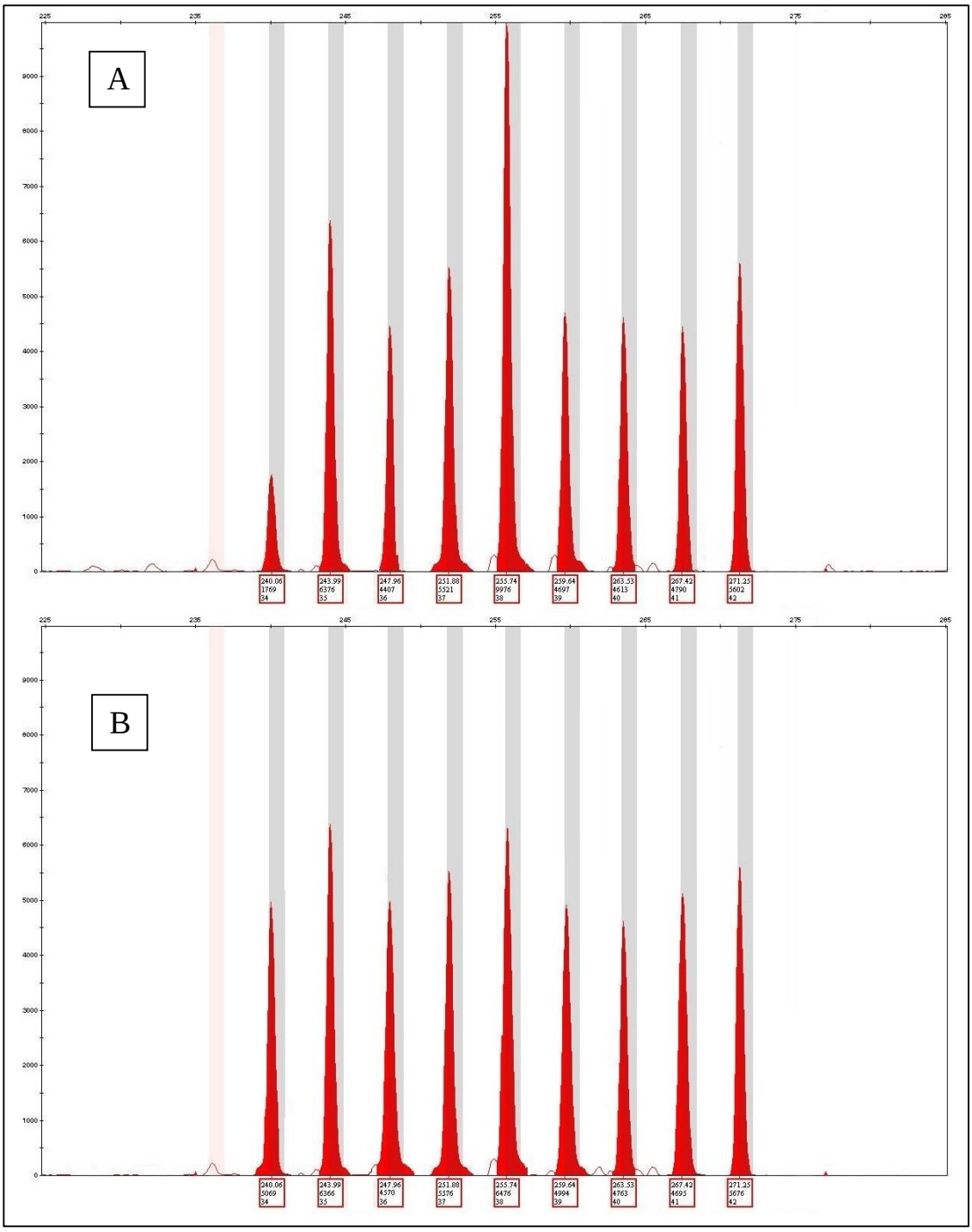

Figure 4.3: DYF387S1 locus-specific allelic ladder having 9 alleles. Part A showing initial results of the co-amplification of different alleles. Part B showing the co-amplification with adjusted amount of DNA template for each allele.

A composite ladder was developed by mixing the locus amplifications. This resulted in a largely balanced composite allelic ladder. The procedure is detailed in Section 2.2.9.3. All alleles were successfully called when 1 µl of the eluted full allelic ladder was electrophoresed on ABI 3500 Genetic Analysed (Figure 4.4). DYF403S1b and DYS526a loci showed the lowest allele peaks height across all loci with approximately 1000 RFU. The decreasing pattern of allelic peak heights as the size of repeats increase within a locus was more pronounced in some of the loci including DYS612, DYS547, DYF404S1, DYS626, DYF403S1 and DYS449.

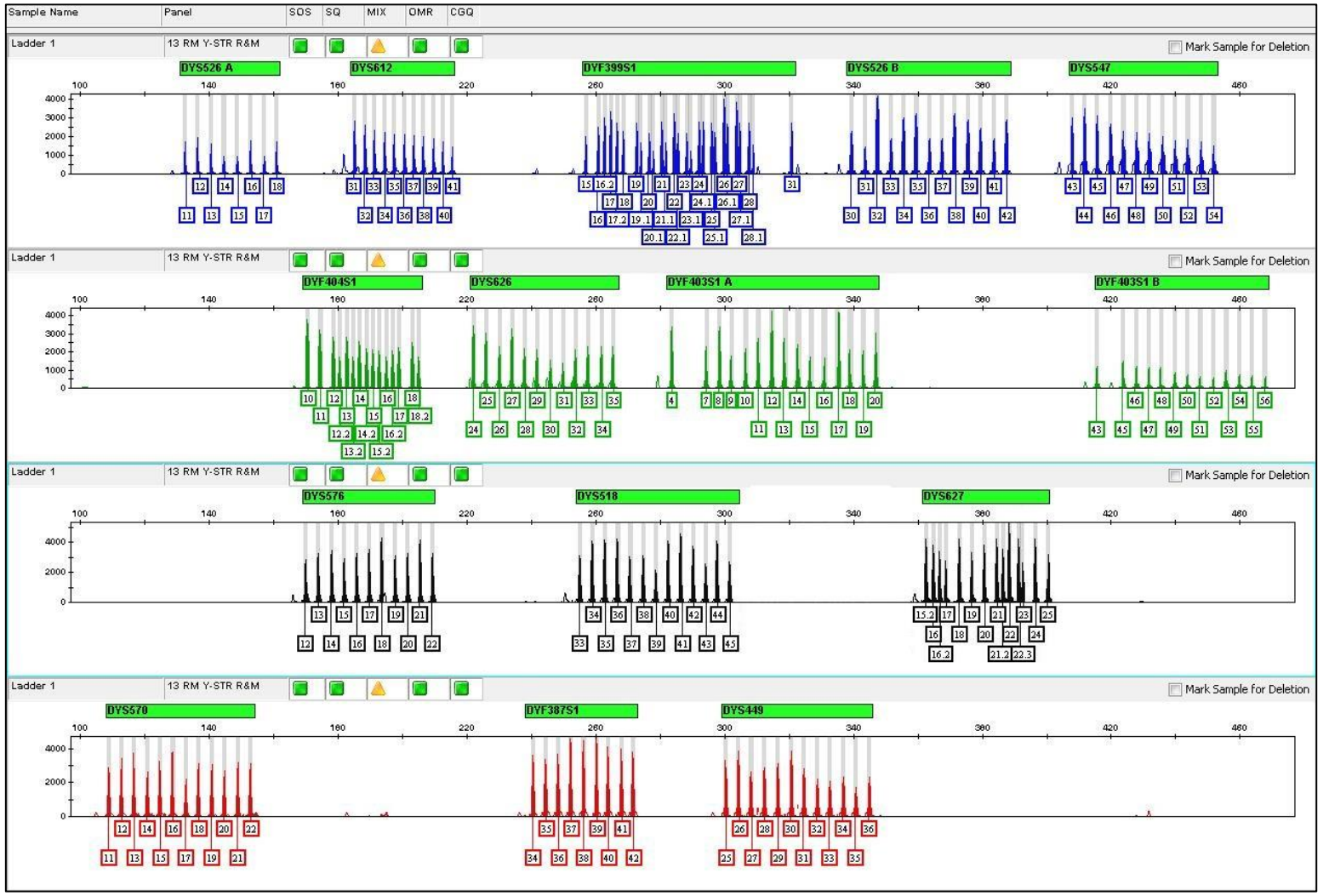

Figure 4.4: Electropherogram as a results of analysing 1 µl of the developed allelic ladder for RM-Yplex multiplex assay

4.2.2. Allele Sequencing

Sequences of alleles at each locus were successfully obtained using BigDye® Terminator v3.0 and analysed using Sequencing Analysis Software v5.4 (Applied Biosystems). All sequences showed consistent results with alleles sequences retrieved from GenBank® as well as previously published sequences of these markers (Ballantyne et al., 2010). However, one out of three allele sequences obtained at DYF403S1b, has shown a substitution of one base pair, cytosine instead of thymine, in the first repeat motif (Figure 4.5). An example of allele sequences results at each locus are presented in Appendix 2 (Figures A2.1-A2.14). Each allele was sequenced in two reactions, using forward or reverse primers in each. The published and modified primers, used in the developed multiplex assay, worked efficiently for sequencing purposes and produced sharp and intense peaks (Figure 4.6). Peak heights among all sequences were balanced apart from alleles at DYF403S1b, DYS526ab and DYS547, which showed decreasing slope pattern toward the end of the sequence.

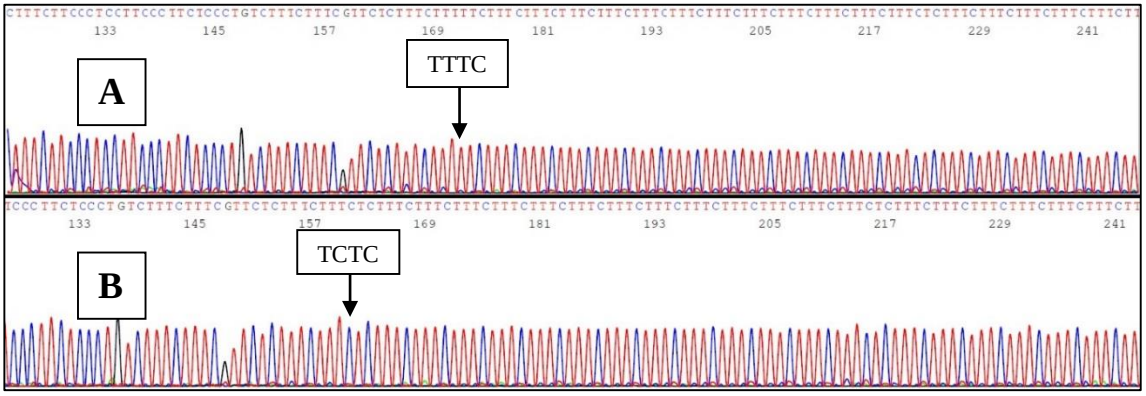

Figure 4.5: Single nucleotide substitution at DYF403S1b locus. Black arrows showing the substitution detected in the sequence of different alleles. A: reference allele sequence; B: allele sequence with substitution.

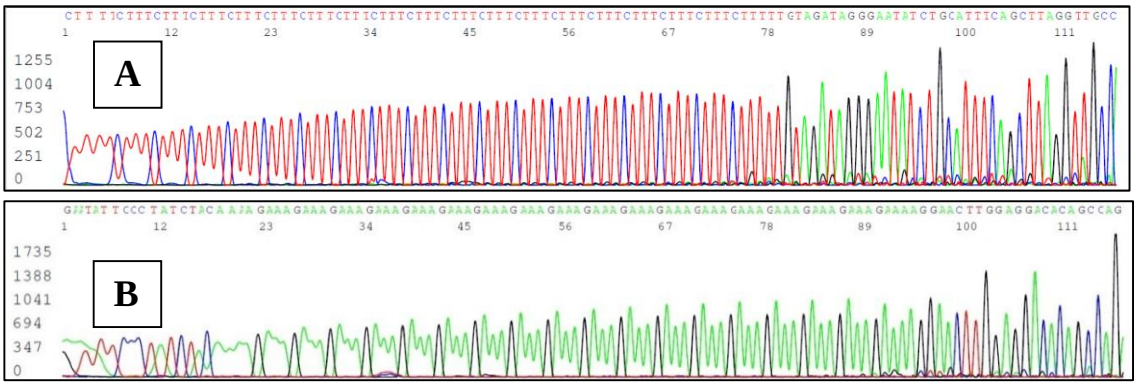

Figure 4.6: Sequencing electropherogram for DYS570 allele 18. Panel A shows forward strand simple STR repeat region. Panel B shows the reverse strand sequence of the repeat region.

Alleles with microvariants were also selected in the sequencing analysis in order to confirm the location of microvariant sequence. Allele 17.2 at locus DYF399S1 was successfully sequenced after gel extraction and purification which showed an addition of two base pairs (G and A) (Figure 4.7).

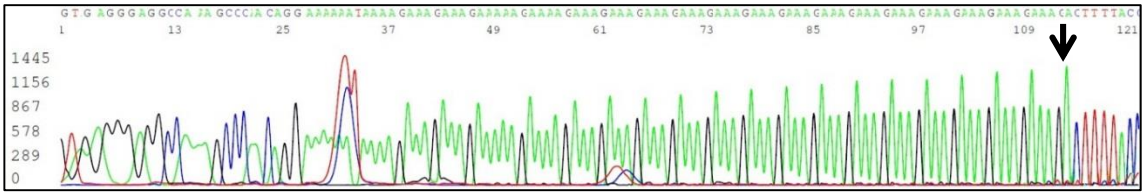

Figure 4.7: Sequencing electropherogram for allele 17.2 at DYF399S1 locus using forward primer showing the repeat unit GAAA and the GA microvariant. The black arrow indicates the microvariant.

An analysis of same fragment size, from two different male individuals, at locus DYS449 has resulted in two different orders of repeat regions within the locus however both had the same total number of repeats (Figure 4.8). Both alleles were sequenced using both forward and reverse primers for confirmation of the sequence.

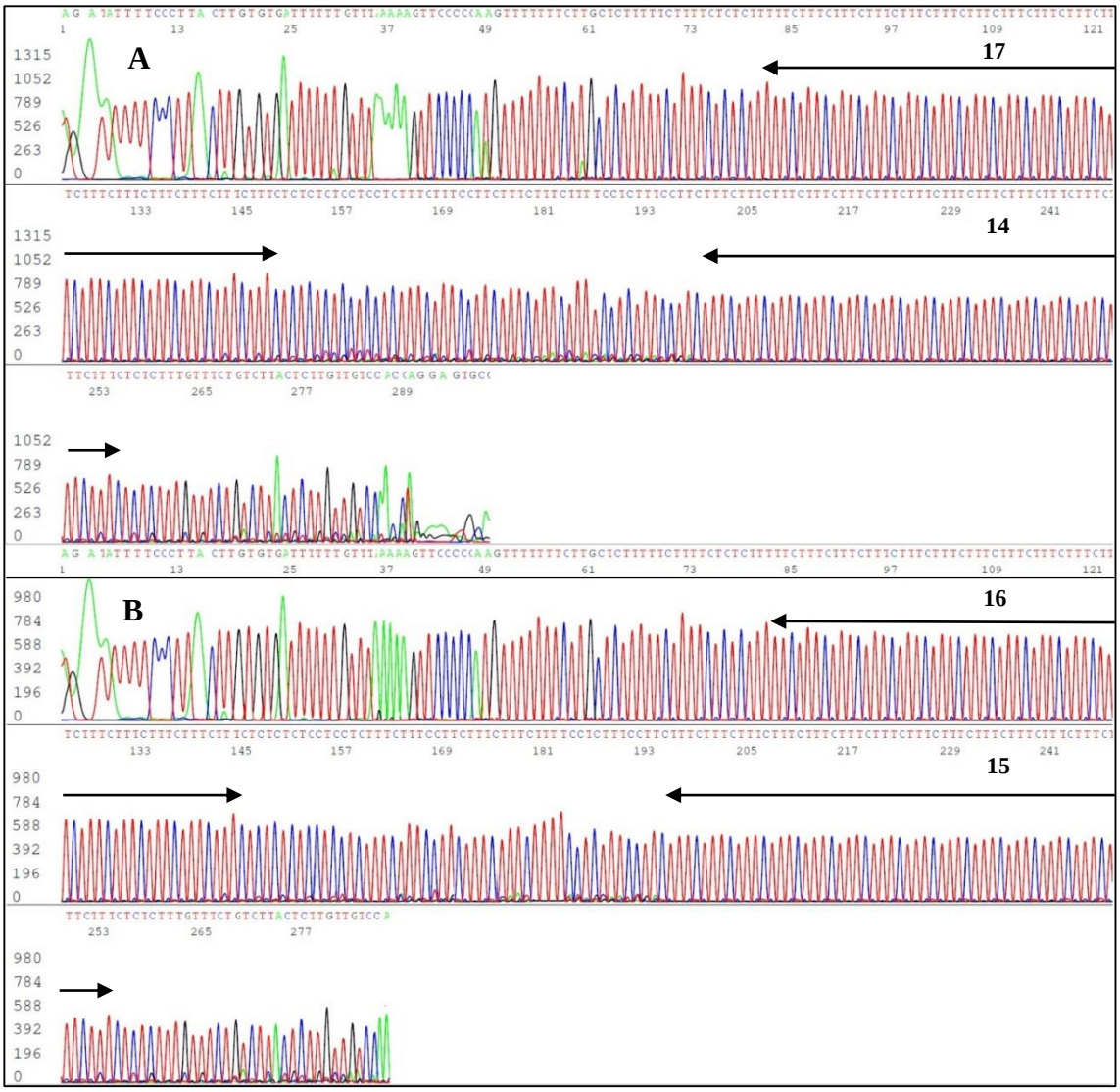

Figure 4.8: Two sequences electropherograms of the same allele size from different individuals A and B showing different order of repeat unit [TTTC] at DYS449 locus. Individual A has allele 31 with the combination of 17 and 14 repeats while individual B has allele 31 with the combination of 16 and 15.

4.2.3. Inter-laboratory Comparison of Allelic Ladder

Re-amplification of allelic ladder provided by RM Y-STR study group showed best result at the dilution ratio of 1:1000, in comparison to 1:100 and 1:500 dilutions, in singleplex reactions as well as in multiplex. All markers where successfully re-amplified in both singleplex and multiplex reactions (Figure 4.9). However, DYS576 and DYS627 loci were interrupted by non-specific peaks (Figure 4.10 and Figure 4.11).

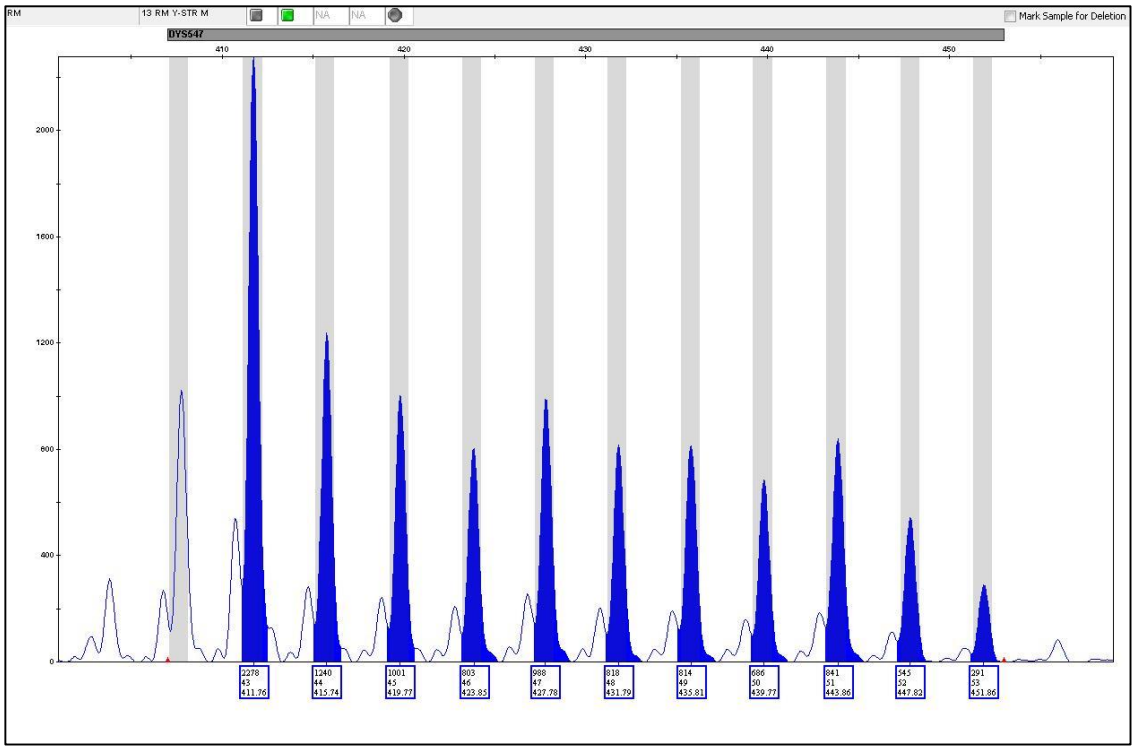

Figure 4.9: Singleplex re-amplification of DYS547 locus-specific allelic ladder received from RM Y-STR study group.

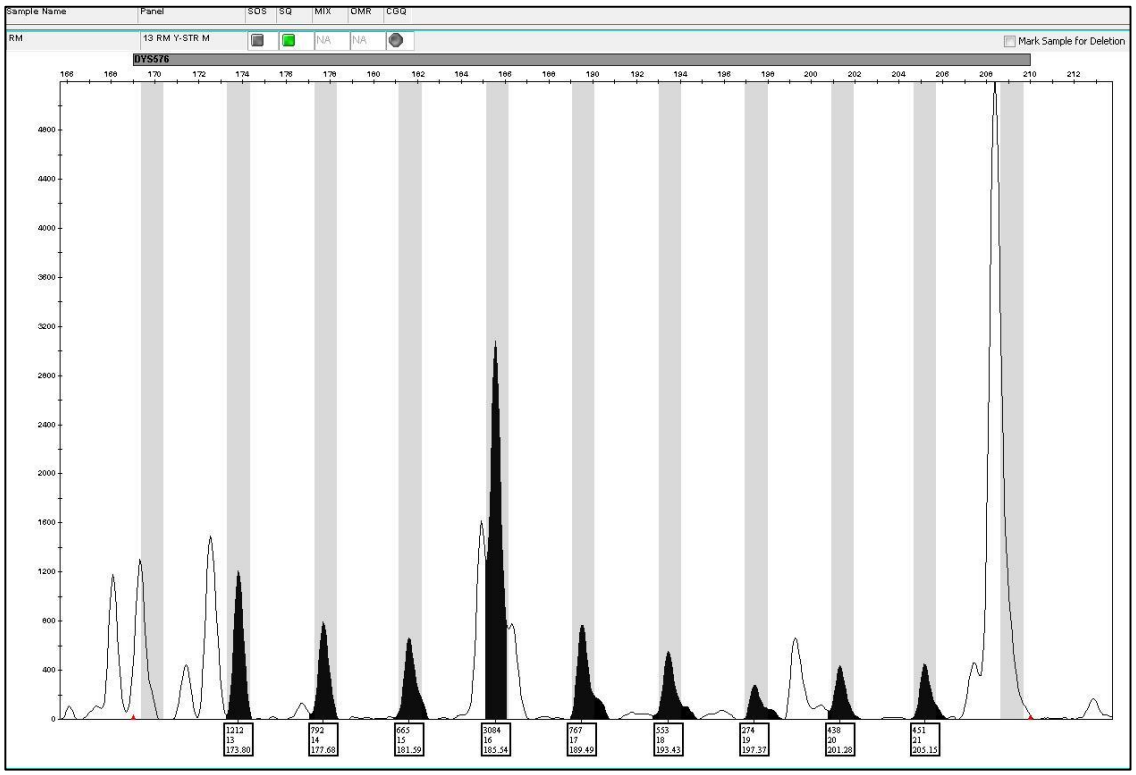

Figure 4.10: Re-amplified DYS576 locus allelic ladder received from RM Y-STR studygroup.

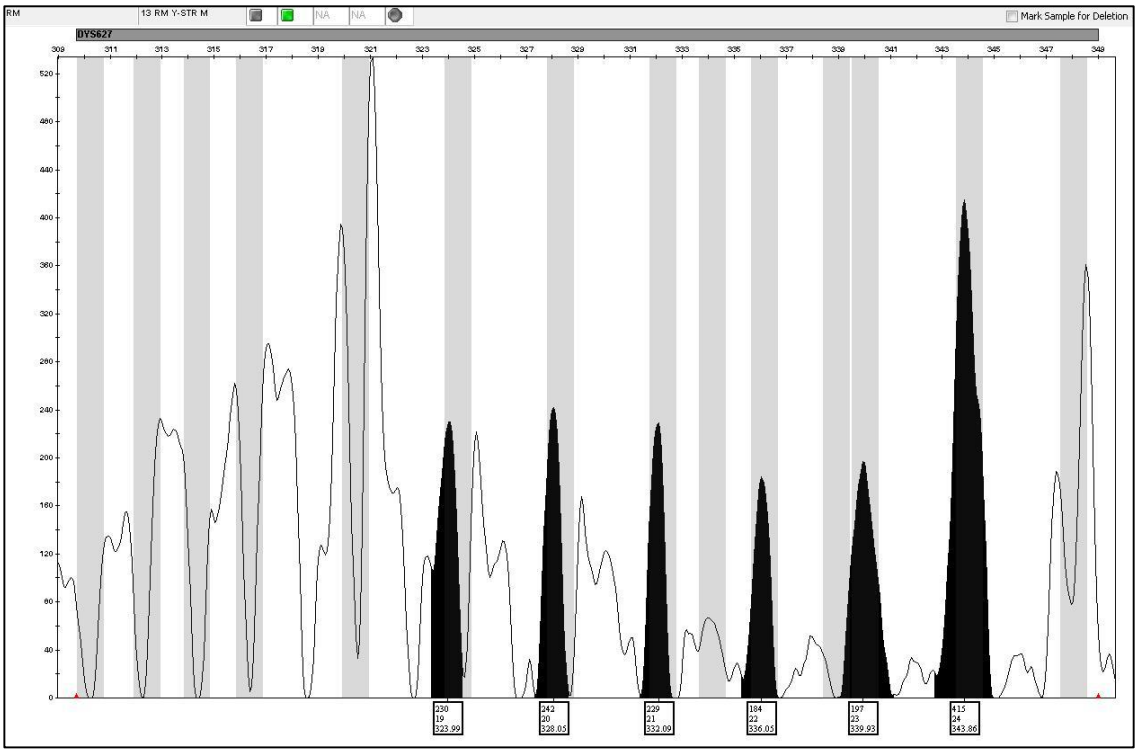

Figure 4.11: Re-amplified DYS627 locus allelic ladder received from RM Y-STR study group.

1 µl of each allelic ladder provided by the RM Y-STR group was amplified in one multiplex reaction. This resulted in an allelic ladder for all the markers with sharp peaks detected with 50 RFU threshold, however, there were many unspecific peaks interrupting the yellow panel comprising of DYS576, DYS518 and DYS627 (Figure 4.12 and 4.13). A positive male DNA control sample of known genotypes, 007 (Applied Biosystems) as well as 9948 positive male DNA control, were used as quality control. Genotypes of these two alleles were already provided by the RM Y-STR study group to aid designating alleles available in the re-amplified allelic ladder and from this to align in-house developed allelic ladder for correct genotyping of samples during the subsequent studies.

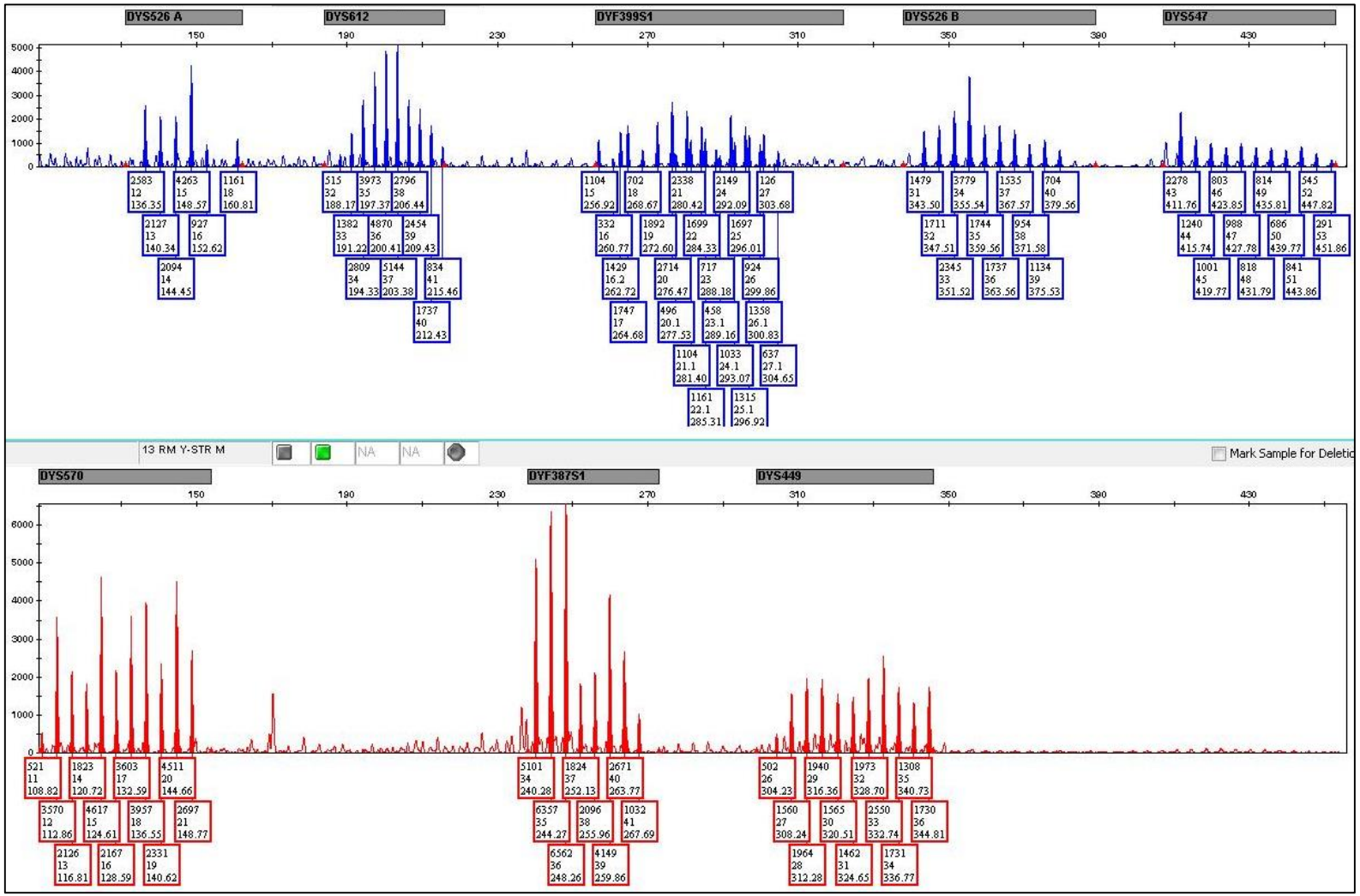

Figure 4.12: Blue and red panels of re-amplified allelic ladder using optimised multiplex reaction.

Figure 4.13: Green and yellow panels of re-amplified allelic ladder using optimised multiplex reaction.

4.2.4. Precision Study of the Developed RM Y-STR Composite Ladder

In order to investigate the size precision of the developed allelic ladder and also setting up the bin files for Genemapper® *ID-X* software, the allelic ladder was analysed 48 times. The mean allelic sizes and the standard deviations for all the 197 alleles included were calculated (Tables A1.1-A1.4). Overall, the average standard deviation was calculated to be 0.05.The highest standard deviation calculated was 0.096 while the lowest was 0.021 (Figure 4.14; Table 4.1). In order to determine the upper limit and lower limit of the size variation for each allele, +/-3 standard deviations were used to size all slightly differing fragment sizes representing the same allele. Thus a window was created for the size of each allele which was then were used to prepare the bins for different loci. In order to investigate the robustness of the analysis, the composite allelic ladder was prepared as before and was injected after a period of three months. The mean allelic sizes and standard deviations for all alleles were calculated for the second time. These results were compared with the results obtained before. In a statistical comparison analysis using an independent t-test, there was no statistical significant difference between the means of standard deviations calculated in each experiment (t-stat = 1.227, df =196, p = 0.221) (Figure 4.15).

Table 4.1: A sample of precision study results showing the highest standard deviation calculated for an allele at each locus.

Locus	Allele	Mean Size (bp)	Standard Deviation
DYF387S1	41	267.46	0.087507
DYF399S1	18	267.94	0.089929
DYF403S1a	15	326.80	0.084499
DYF403S1b	50	447.89	0.095541
DYF404S1	13	181.87	0.093132
DYS449	29	317.26	0.085749
DYS518	36	266.37	0.091975
DYS526a	18	159.61	0.082694
DYS526b	39	375.56	0.087381
DYS547	50	440.13	0.084875
DYS570	14	121.33	0.086873
DYS576	20	201.18	0.080616
DYS612	37	202.89	0.090077
DYS626	25	223.99	0.072009
DYS627	20	381.96	0.057911

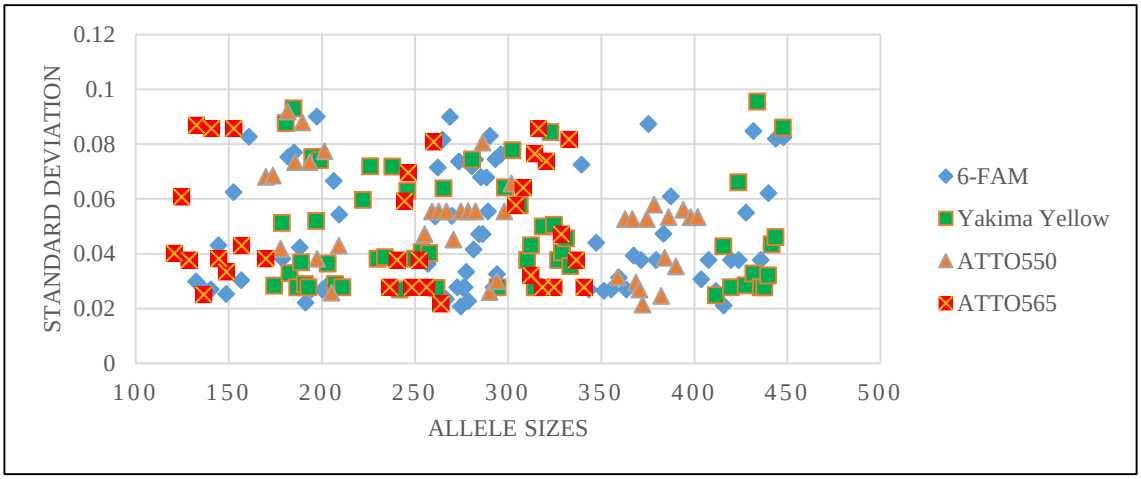

Figure 4.14: Standard deviation analysis of 197 fragments sizes included in the allelic ladder.

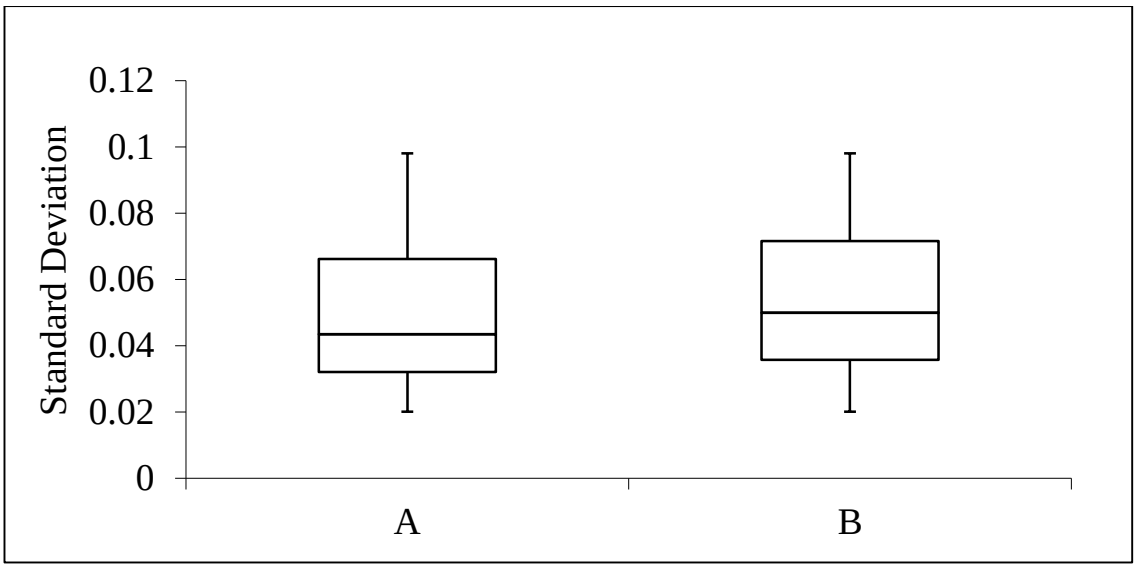

Figure 4.15: Box plot of standard deviations of allele sizes generated during the precision study. A and B shows data of the ladders that were run with 3 months gap on the same Genetic Analyser 3500 instrument using POP-6™ polymer.

4.3. Discussion

4.3.1. Development of Allelic Ladder

The allelic ladder is essential in forensic DNA profiling and is a required quality control addition which is used to call the correct alleles in STR based DNA profiling. The strategy of developing an allelic ladder followed in this project has been proven to work effectively for short term usage in addition to the low cost involved compared to other techniques such as cloning (Anushika et al., 2010, Wang et al., 2014). The strategy used in this experiment was adapted from previous studies where allelic ladder was simply developed by mixing different alleles of each marker with equal proportions in order to obtain a composite allelic ladder (Butler et al., 2003, Anushika et al., 2010, Cloete et al., 2010, D'Amato et al., 2010). The developed allelic ladder helped to confirm the size of microvariants, which were detected earlier when only the internal size standard was used to size the alleles (Chapter 3, Section 3.2.5). All microvariant alleles included in the initial ladders were confirmed initially when they were co-amplified and electrophoresed as fragments with microvariants were sized between alleles with full repeat number (Figure

4.16). The allele sizing was quite precise showing the robustness of the size standard. This also, reflects the high accuracy and precision of ABI 3500 Genetic Analyser. Due to the shortage of sample supply, microvariant samples were not used in further developing allelic ladder and instead they were saved for population genotyping.

Overall, allelic ladders for six loci, namely DYS612, DYS547, DYF404S1, DYF403S1b, DYS626 and DYS449, demonstrated an increase in allelic peak heights from longest to shortest alleles. This was probably due to the efficient amplification of smaller alleles and also due to the stutter peaks of adjacent alleles enhancing the size of the shorter alleles. However, this pattern was minimally observed in other loci. This could be due to the low stutter peaks for these markers. Locus DYF403S1ab multi-allelic marker showed random amplification of the three different copies part of the marker on Y chromosome and hence the pattern of reverse peak height increasing was not expected.

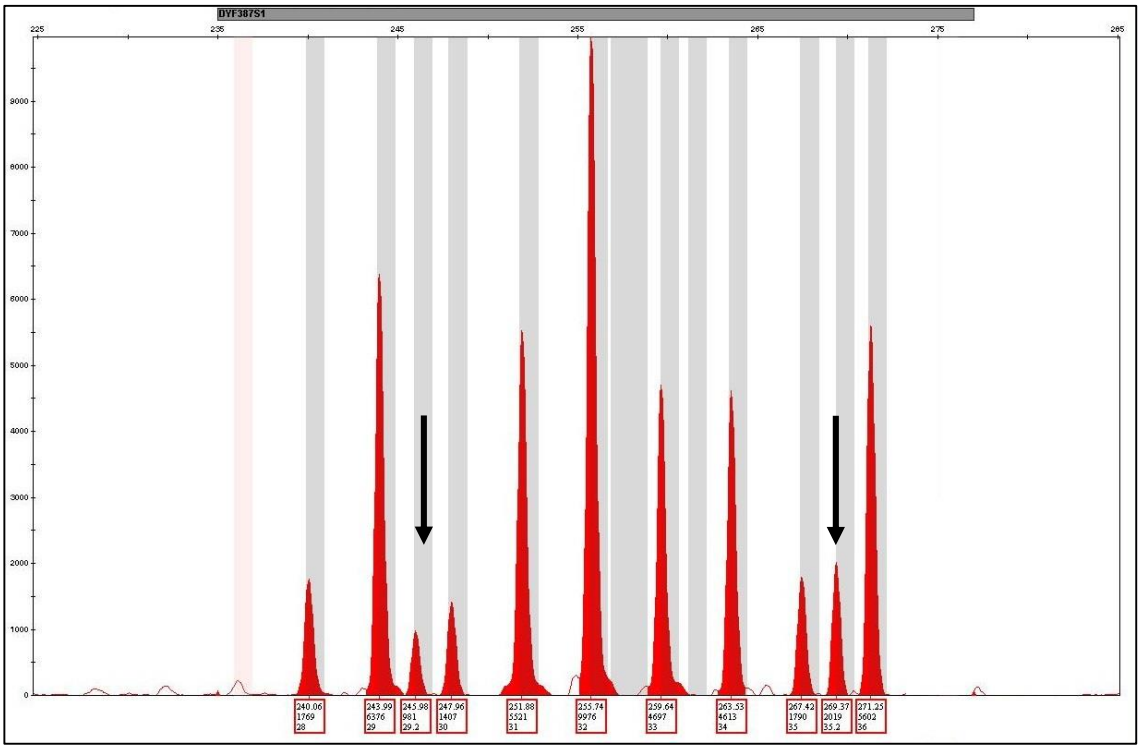

Figure 4.16: Electropherogram demonstrating the confirmation of the presence of microvariant at DYF387S1 by co-amplification of alleles. Black arrows indicate microvariant alleles.

4.3.2. Re-amplification of Allelic Ladder

Re-amplification of an allelic ladder received from RM Y-STR group was successfully achieved at the dilution of 1:1000 rather than 1:100 and 1:500. The interruption of allelic ladder at DYS576 and DYS627 complicated the analytical process in assigning the real allelic peaks as some of the artefacts presented were showing as same morphological features as the real allele peaks (Figures 4.10 and 4.11). The reason behind this interruption was that the primer fluorescent labels in the received allelic ladder were completely different than the ones being used in the RM-Yplex multiplex assay developed during this project. Hence this interruption is most likely arose from original fluorescent labels used by the RM Y-STR group. Another interrupting peak with same morphological features as a real peak was detected at DYF404S1 within the position of microvariant allele 16.2 which has been confirmed to be an artefact by developer (Figure 4.12). The same situation has been noticed at DYS547 as shown in Figure 4.9, there was a peak which is 4 base pairs less than the shortest allele at this marker with equal peak intensity to other alleles in the marker. This peak was also confirmed to be an artefact peak by the developer. The red dye channel was the least interrupted by non-specific peaks as this dye was not used in developing allelic ladder for the three multiplexes developed earlier (Victoria, 2010). The results of the re-amplifying the allelic ladder showed the robustness of the developed RM-Yplex assay. Electrophoretic detection of sharp peaks suggested the optimal components of the master mix and also the optimal final extension time in PCR cycles conditions. Overall, the re-amplified allelic ladder was used to designate the alleles included in the developed allelic ladder for the RM-Yplex multiplex assay correctly.

4.3.3. Precision Study

Size precision is vital for accurate genotyping of unknown amplified STR products. The developed allelic ladder was used to study the size precision of the ladder alleles which depends on the instrument performance and internal size standard efficiency during

electrophoresis. This study was conducted twice at different time of the year in order to investigate the effects of the analysis conditions and parameter over a period of time. The analysis of 6 injections (48 samples) of allelic ladder on ABI 3500 Genetic Analyser using POP-6 polymer with 3 months gap between each analysis demonstrated a great high level of precision. Both analyses showed no significant difference between standard deviations calculated at each experiment. Standard deviation for all averaged allele sizes fluctuated between 0.02 and 0.1. These findings are slightly different than previous studies when comparing average standard deviations. However, the range of standard deviation was still similar to other studies performed using different STR multiplex assay on 3500 Genetic Analyser (French, 2011, Thompson et al., 2013). This could be due to the efficiency of the internal size standard being used as well as the existence of genetic analyser's performance differences across different laboratories.

4.3.4. Allele Sequencing

Sequencing of both the forward and the reverse strands served as a method to confirm the DNA sequences obtained. All alleles from multi-allelic loci were extracted, run on polyacrylamide gel and purified apart from DYS526ab locus. For DYS526ab the forward primer binding site is duplicated, thus fragment DYS526a resides within the fragment DYS526b and therefore, sequencing could be achieved without separation of alleles as overlapping in sequence peaks is not possible in such situation. Sequencing for this locus was successfully achieved with a noticeable increase in peak intensity at the second attachment site of the forward primer. This was expected as when the DNA polymerase at the sequencing reaction reaches the second attachment site, there would be as double amount of PCR product templates and two attachment sites, available for the reaction to occur.

The sequence data helped to serve a number of purposes. First, it confirmed the allele calls for the alleles included in the allelic ladder. A complete concordance was observed

between sequenced alleles and the corresponding allele calls identified during the optimisation of the multiplex assay (Chapter 3, Section 3.3.4) and sequences retrieved from GenBank® database. Secondly, no polymorphisms were detected in the flanking region of all the RM Y-STR loci alleles that were sequenced. This means that primer binding sites are located within stable regions and null alleles are not expected in UAE populations. Also, sequences generated out of this study were compared with previously published sequences by Ballantyne et al. (2010). This showed a complete concordance. Thirdly, sequencing data allowed for the identification of internal sequence variation that might exist within the same locus, i.e. homoplasmy. This was detected in analysing two alleles at locus DYS449 that showed same fragment size and calling allele 31 according to the recent recommended nomenclature by ISFG (Mulero et al., 2014). The two alleles have identical total repeat numbers but in different order (Figure 4.8). This phenomenon has been reported previously in a sequencing study of this and also seen in DYS448 (Schoske, 2003, D'Amato et al., 2010). Internal sequence variations due to micro-heterogeneity found in compound loci can lead to an underestimation of STR diversity of such markers as such variation is masked due to fragment length analysis that is currently employed. Finally, one polymorphism was observed at one out of three alleles sequences at DYF403S1b in the first repeat unit. Such a single nucleotide polymorphisms (SNP) can add further discrimination to this allele from other same size alleles when using sequencing technology. Such polymorphisms were also observed previously in many other STR markers (Rockenbauer et al., 2014).

5. Developmental Validation of RM-Yplex Assay

5.1. Overview

In order to employ a new STR multiplex assay for use in forensic investigations it needs to be validated to a level where it meets high performance standards. The limitations of the assay should be clearly defined as well. All data acquired during such studies can be of immense value for subsequent users of such an assay. Organisations like the Scientific Working Group on DNA Analysis Methods (SWGDAM), European DNA Profiling Group (EDNAP), International Society for Forensic Genetics (ISFG) and International Union of Pure and Applied Chemists (IUPAC) have defined criteria for the validation of multiplex PCR assays. This ensures that high genotyping and analysis standards are maintained within the sector. The SWGDAM validation guidelines have been widely used over the last decades in forensic genetics to validate a new DNA profiling technique. Since forensic samples could be obtained from a number of different conditions, it was necessary to assess the reliability of the RM-Yplex assay in producing accurate results and its limitations. To achieve these goals revised SWGDAM guidelines (NIST, 2012), for Y-STR based DNA profiling was followed during this study. The PCR conditions used had been established earlier (Table 2.7; Table 2.8).

5.2. Results

5.2.1. Analytical Threshold

Using all 25 electropherograms generated for negative female DNA control amplifications, the highest noise peak detected in the baseline were recorded for each dye to calculate the analysis limit of detection (LOD) and limit of quantitation (LOQ) for each dye channel as described by Gilder et al. (2007) (Table 5.1). Based on this study, 50 RFU threshold was used for all subsequent studies in this chapter as well as population study chapter.

Table 5.1: Dye-specific limit of detection (LOD) and limit of quantitation (LOQ) analysis results.

Channel	Maximum RFU	Average RFU	Standard Deviation (SD)	Averaged RFU +3 SD (LOD)	Averaged RFU +10 SD (LOQ)
Blue	25	15.57	5.39	31.74	69.47
Green	23	15.06	5.67	32.07	71.74
Yellow	24	14.94	6.13	33.32	76.21
Red	26	15.41	5.97	33.34	75.16
Orange	15	10.29	3.08	19.53	41.10

5.2.2. Reproducibility

To investigate the consistency of the developed multiplex assay, each of the 6 male DNA samples, TaqMan, 9948, 2800M, R, K and M, were typed on at least at 6 different occasions by five different DNA analysts. Consistent profiles were obtained for all samples genotyped using the multiplex assay (Table 5.2). In addition, assessment result of 8 quality control samples genotyping results which were sent to Erasmus MC laboratory, was confirmed to be correct (Figures A1.1 to A1.9). Moreover, six samples provided in the SRM2395 standard sample set for Y-STR analysis were successfully genotyped and results were compared to reference genotyping results (Thompson et al., 2013) (Table 5.3).

Table 5.2: RM-Yplex profiles for different human male DNA samples. The profiles were developed by injecting the amplified products in an 8 capillary ABI 3500 genetic analyser and analysing the data using Genemapper® *ID-X*.

Locus	2800M	9948	TaqMan	R	M	K
DYF387S1	37:38	35:38	36:39	38:39	37	37:38
DYF399S1	24:25.1:26.1	21:22:25.1	19:23	18:24:25.1	23.1:26:26.1	21:21.1:23
DYF403S1a	11:14:17	10:15:16	11.2:12.2:15	8:14:17	11:14:17	10:14:17
DYF403S1b	46	49	54	53	45	50
DYF404S1	13:16	12:14	15:18	14:15	12:15	14:16
DYS449	34	30	29	31	35	29
DYS518	40	38	42	42	39	46
DYS526a	12	14	13	13	12	15
DYS526b	36	36	34	33	34	38
DYS547	45	48	48	48	45	38
DYS570	17	18	19	18	17	20
DYS576	18	16	15	18	16	17
DYS612	35	37	36	36	33	38
DYS626	29	28	29	25	31	33
DYS627	22	22	21	18	22	17

Table 5.3: RM-Yplex profiles for six samples available in SRM2395 standard set of samples used for Y-STR analysis.

Locus	SRM A	SRM B	SRM C	SRM D	SRM E	SRM F
DYF387S1	34:36	39:40	38:39	36: 39	36: 37	-
DYF399S1	20: 24:27.1	21:24	20:24: 26	18.2: 21.1:24	20:23:25	-
DYF403S1a	10:15:17	10:11:15	12:15	12:13:17	12:17:19	-
DYF403S1b	50	49	52	45	45	-
DYF404S1	14:16	15:16	13:15	14.2:15	14:15	-
DYS449	29	33	31	29	28	-
DYS518	39	40	38	36	40	-
DYS526a	14	13	14	13	12	-
DYS526b	36	36	34	33	32	-
DYS547	48	47	52	49	45	-
DYS570	17	18	18	17	18	-
DYS576	18	16	17	18	17	-
DYS612	38	38	34	34	39	-
DYS626	31	34	28	31	31	-
DYS627	25	18	21	22	20	-

5.2.3. Precision and Accuracy

The principle of STR genotyping is based on the ability of the method to measure the size of the amplified product precisely. Hence, size precision facilitates the reliability of the assay (Krenke et al., 2005, Mulero et al., 2006). In developmental validation, the precision study determined the reproducibility of the size measured for the amplified product and

also the ability of the analysis method to generate a correct genotype (Daniels et al., 2004, Krenke et al., 2005). In this experiment TaqMan male DNA control was amplified and genotyped on three different occasions, with minimum of one month lapse between each genotyping occasion. Precision of each allele at the corresponding locus was calculated and expressed as standard deviation, whereas accuracy was expressed as the variation in allele size (bp) between sample alleles compared to those of the allelic ladder in the same run.

The precision experiment indicated that among all loci the standard deviation did not exceed 0.1 base pairs (Table 5.4). The highest standard deviation of 0.081 base pairs was detected at DYF387S1 and DYF404S1 loci, whereas the lowest standard deviation of 0.023 was observed at locus DYS576. The size of alleles of all loci fell within the ±0.5 bp range (Table 5.5). In each set of the analysis performed, TaqMan DNA control did not show the presence of any significant variant alleles.

Table 5.4: Precision study showing average allele size and standard deviation of TaqMan male DNA control for 13 RM Y-STR loci.

Locus	Mean Size Allele	Standard Deviation
DYF387S1	248.25	0.081
	259.85	0.069
DYF399S1	272.56	0.035
	288.15	0.026
DYF403S1a	312.33	0.057
	316.46	0.040
	327.37	0.026
DYF403S1b	463.86	0.066
DYF404S1	190.84	0.081
	203.00	0.038
DYS449	320.50	0.035
DYS518	274.46	0.028
DYS526a	140.37	0.029
DYS526b	355.57	0.035
DYS547	431.75	0.044
DYS570	140.64	0.029
DYS576	181.60	0.023
DYS612	200.36	0.020
DYS626	241.72	0.078
DYS627	382.11	0.026

Table 5.5: Variation of allele sizes of the TaqMan male DNA control sample in comparison to the allelic ladders for each of the RM Y-STR loci. Variation of size (bp) = allelic ladder size – allele size of sample. Each run of TaqMan sample represent the mean of 8 replications

Locus	Allele sizes of TaqMan sample (bp)			Allele sizes of Allelic Ladder (bp)			Variation of size (bp)		
	Run 1	Run 2	Run 3	Run 1	Run 2	Run 3	Run 1	Run 2	Run 3
DYF387S1	248.19	248.34	248.21	248.22	248.35	248.28	0.03	0.01	0.07
	259.81	259.93	259.81	259.80	259.91	259.85	-0.01	-0.02	0.04
DYF399S1	272.56	272.59	272.52	272.51	272.57	272.53	-0.05	-0.02	0.01
	288.16	288.17	288.12	288.11	288.15	288.14	-0.05	-0.02	0.02
DYF403S1a	312.39	312.31	312.28	312.40	312.32	312.28	0.01	0.01	0
	316.48	316.41	316.48	316.50	316.38	316.42	0.02	-0.03	-0.06
	327.36	327.35	327.4	327.37	327.31	327.45	0.01	-0.04	0.05
DYF403S1b	463.92	463.79	463.87	463.99	463.74	463.81	0.07	-0.05	-0.06
DYF404S1	190.91	190.75	190.85	190.95	190.71	190.88	0.04	-0.04	0.03
	203.03	202.96	203.02	203.09	202.96	203.05	0.06	0	0.03
DYS449	320.47	320.50	320.50	320.44	320.52	320.57	-0.03	0.02	0.07
DYS518	274.35	274.60	274.43	274.36	274.61	274.42	0.01	0.01	-0.01
DYS526a	140.39	140.34	140.39	140.37	140.36	140.40	-0.02	0.02	0.01
DYS526b	355.61	355.55	355.55	355.63	355.59	355.60	0.02	0.04	0.05
DYS547	431.8	431.73	431.72	431.82	431.73	431.71	0.02	0	-0.01
DYS570	140.66	140.61	140.66	140.65	140.62	140.67	-0.01	0.01	0.01
DYS576	181.59	181.63	181.59	181.60	181.61	181.60	0.01	-0.02	0.01
DYS612	200.36	200.34	200.38	200.31	200.33	200.39	-0.05	-0.01	0.01
DYS626	241.68	241.81	241.67	241.69	241.85	241.69	0.01	0.04	0.02
DYS627	381.99	382.24	382.09	382.02	382.27	382.11	0.03	0.03	0.02

5.2.4. Sensitivity and Stochastic Studies

The sensitivity of the RM-Yplex assay was investigated over seven different input amounts of male DNA template including, 1000 pg, 750 pg, 500 pg, 250 pg, 125 pg, 62.5 pg, 31.25 pg using 6 different male individuals samples used for the reproducibility study. In addition, each set of the prepared serial dilution for each sample was tested with 28, 30 and 32 cycles (Figure 5.1). Overall, input DNA <125 pg the multiplex assay showed preferential amplification of the smaller alleles and some allelic dropouts. Optimised PCR cycle of 32 cycles produced complete profile consistently for all 6 samples, down to 62.5 pg. Allelic dropouts started to appear earlier at 125 pg when 30 PCR cycles and 28 PCR cycles number were used. Median peak heights were calculated along with the 25% and 75% quartiles. These were plotted in box plot (Figure 5.2 A, B and C). The peak heights of larger loci kept detected with decreasing DNA template amount for all numbers of cycles.

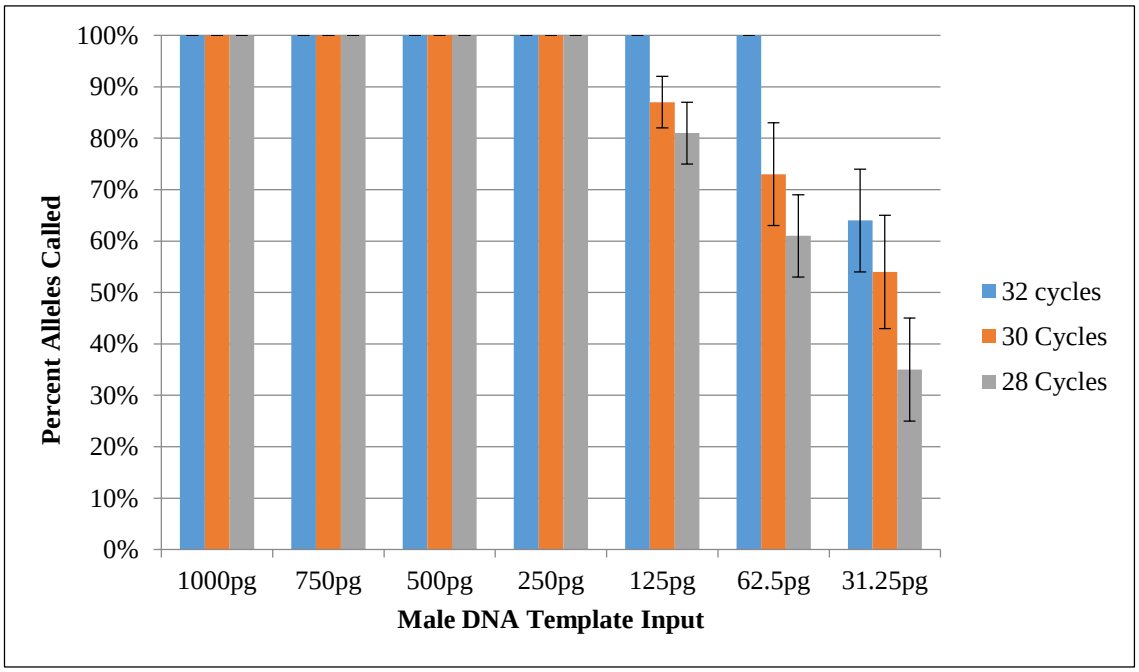

Figure 5.1: 15 µl RM-Yplex multiplex assay allele call percentage when tested using 6 different male DNA samples for different input DNA amounts using 28, 30 and 32 PCR cycles.

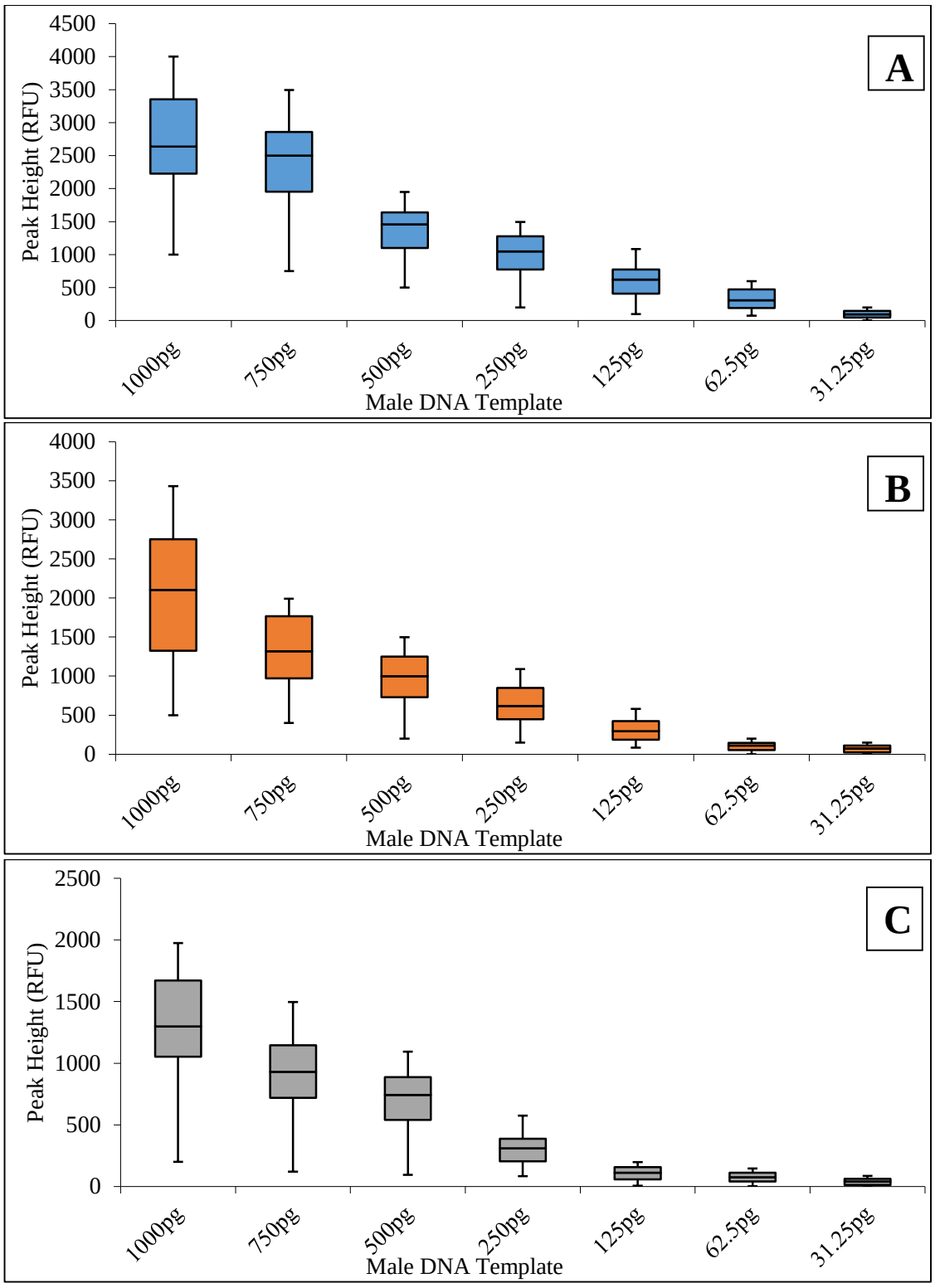

Figure 5.2: Distribution of peak height intensity of RM-Yplex multiplex assay profiles generated using (A) 32 PCR cycles, (B) 30 PCR cycles and (C) 28 PCR cycles.

5.2.5. Human and Species Specificity

5.2.5.1. Female DNA Template Titration

Three different female DNA templates were used to prepare mixtures with 2800M male DNA control. The female DNA concentration was increased from 250 ng to 1000 ng while the male DNA concentration was kept at 1 ng. These mixtures were amplified in triplicate and injected in the genetic DNA analyser ABI 3500 Genetic Analyser. Full male profiles were obtained for all mixture ratios (Figure 5.3 and 5.4). Although, male DNA was constant at 1 ng only a marginally reduction in overall peak heights was observed at extremely high female content mixtures.

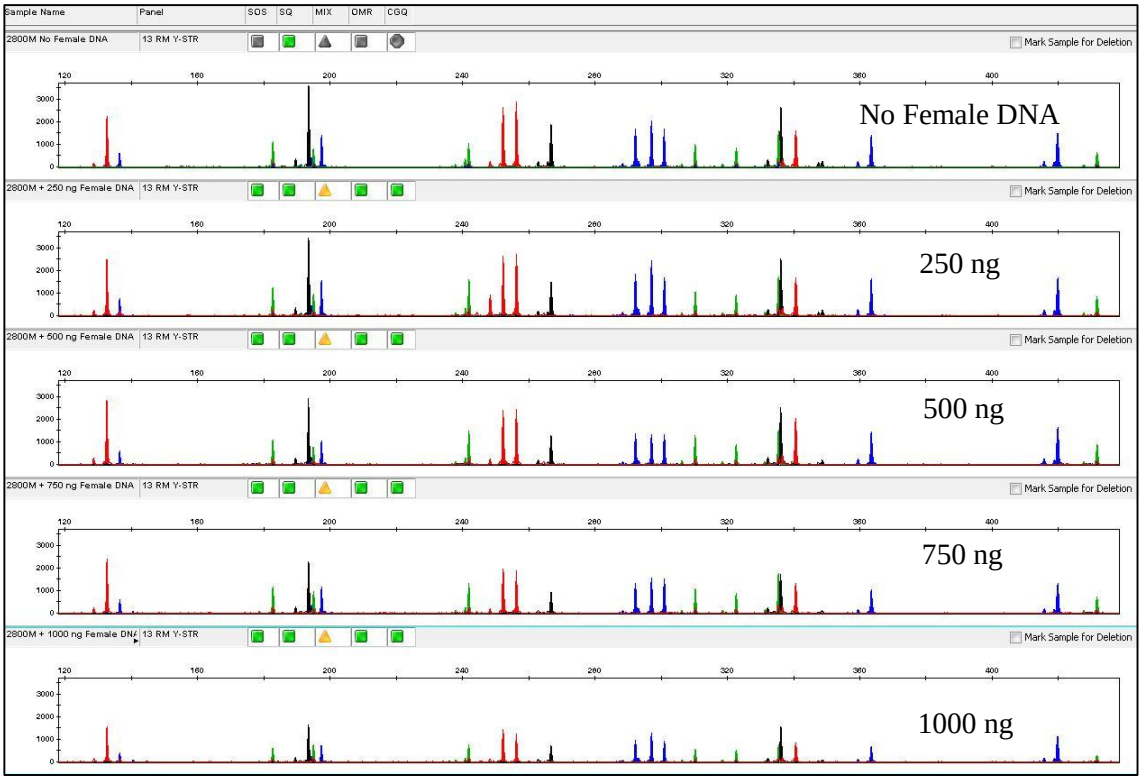

Figure 5.3: RM Y STR profiles generated from amplification of various ratios of male:female mixtures. The male sample used was the male control DNA 2800. All panels are in scale. Amount of female DNA is typed in each panel.

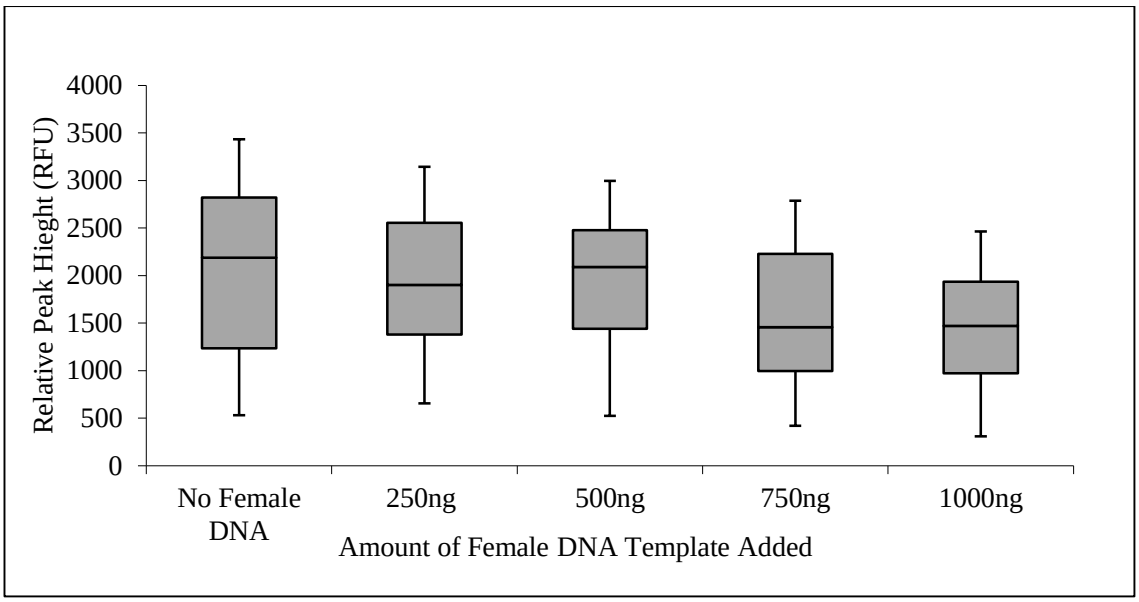

Figure 5.4: Allelic peak height comparison of profiles generated using male/female mixture. Results are presented using a box-plot where each box represents 9 replicates, where values range from bottom quartile (25%) to the top quartile (75%) and the median is depicted as an intersecting line within the boxes.

5.2.5.2. Species Specificity

Four samples were provided by the Dubai Zoo for male gorilla, male chimpanzee, unknown gender gibbon and male orang-utan. The ability of the RM-Yplex multiplex assay to amplify non-human DNA was tested using these samples. Chimpanzee DNA sample amplification showed three single-peaks at each of DYS576, DYS626 and DYS612 loci. Amplification of gorilla DNA sample resulted in two single-peaks at DYS626 and DYS612. Orang-utan DNA samples showed only a single peak at DYS449. However, gibbon showed no amplification at any locus (Table 5.6). All resulted peaks were confirmed by singleplex PCR reactions.

Table 5.6: Results from amplification of RM Y-STR loci in four male primate DNA samples.

Species	**Amplified markers**
Chimpanzee	DYS576 – DYS626 – DYS612
Gorilla	DYS626 – DYS612
Orang-utan	DYS449
Gibbon	No Detection

5.2.6. Stability Study

Three common PCR inhibitors in biological stains, were tested over ranges of different concentrations in order to determine the limitation of the assay. These inhibitors were haematin, humic acid and tannic acid. Three different male DNA samples were amplified in the presence of different inhibitors concentrations. Each amplification reaction was replicated thrice at three different occasions. The percentages of averaged detected alleles were determined for each inhibitor along with the standard deviations for all replicates (Figure 5.5). Across three inhibitors, haematin had the greatest effect on PCR inhibition where allelic dropout was observed at 200 ng/µl. For humic acid allelic dropout was observed at 200 ng/µl concentration. However more than 85% of alleles were called at this concentration. Tannic acid showed inhibitory effects at 300 ng/ µl.

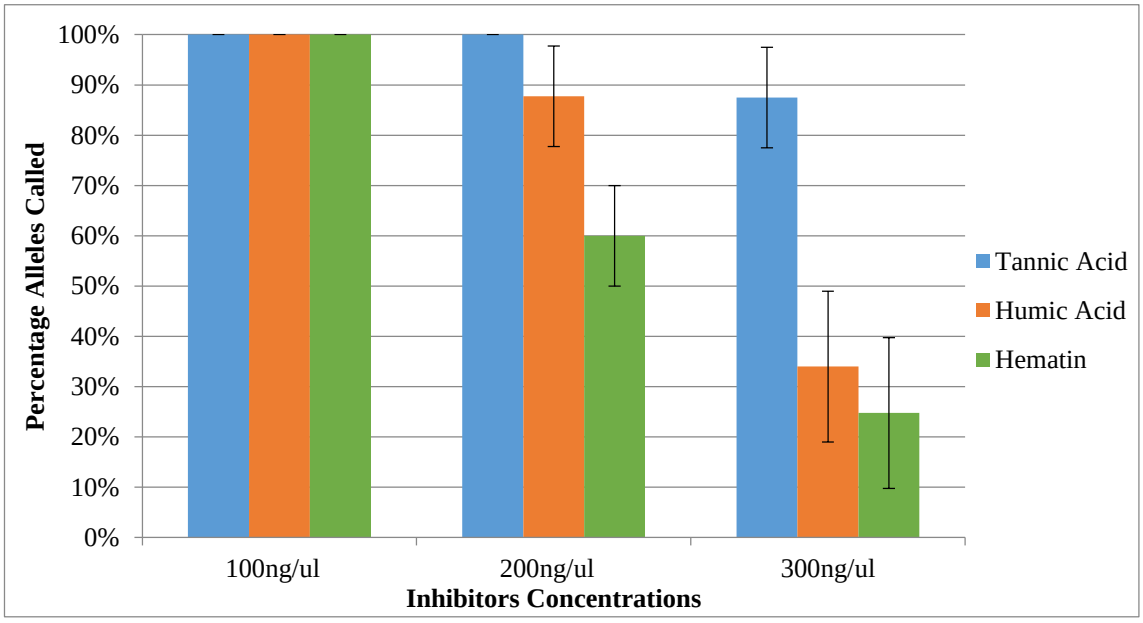

Figure 5.5: Effect of haematin, humic acid and tannic acid on PCR amplification. Three different concentrations of each were amplified using 1 ng three different human male samples. The average percentage of loci called for each inhibitor concentration is represented. Error bars show the standard deviation of 9 replicate PCRs for each concentration of inhibitor analysed.

5.2.7. Mixture Study

In this study, ability of the RM-Yplex assay to reliably detect male/male mixtures was evaluated. Two male/male DNA mixture sets (M1M2 and M2M3) were prepared including mixture ratios of 19:1, 9:1, 3:1, 1:1, 1:3, 1:9, and 1:19. Each mixture was amplified in triplicate using total of 1 ng DNA and analysed on ABI 3500 Genetic Analyser. The percentage of alleles detected in unique minor profile was calculated for each of the samples. Locus DYF403S1a was excluded from this study since it had been realised that it had constant imbalances peak heights. The results showed all of the unique minor profiles were called for ratios of 3:1, 1:1 and 1:3 (Figure 5.6). Some of the alleles in the minor profile dropped out in the 9:1and 1:9 ratio samples, however, greater than 70% of the unique minor profile alleles were called. When the mixture ratio was increased to 19:1 or 1:19, an average of greater than 55% of the unique minor profile alleles were called. In the 1:1 mixture samples the peak heights of the two profiles were similar and consistent with this type of mixture (Figure 5.7).

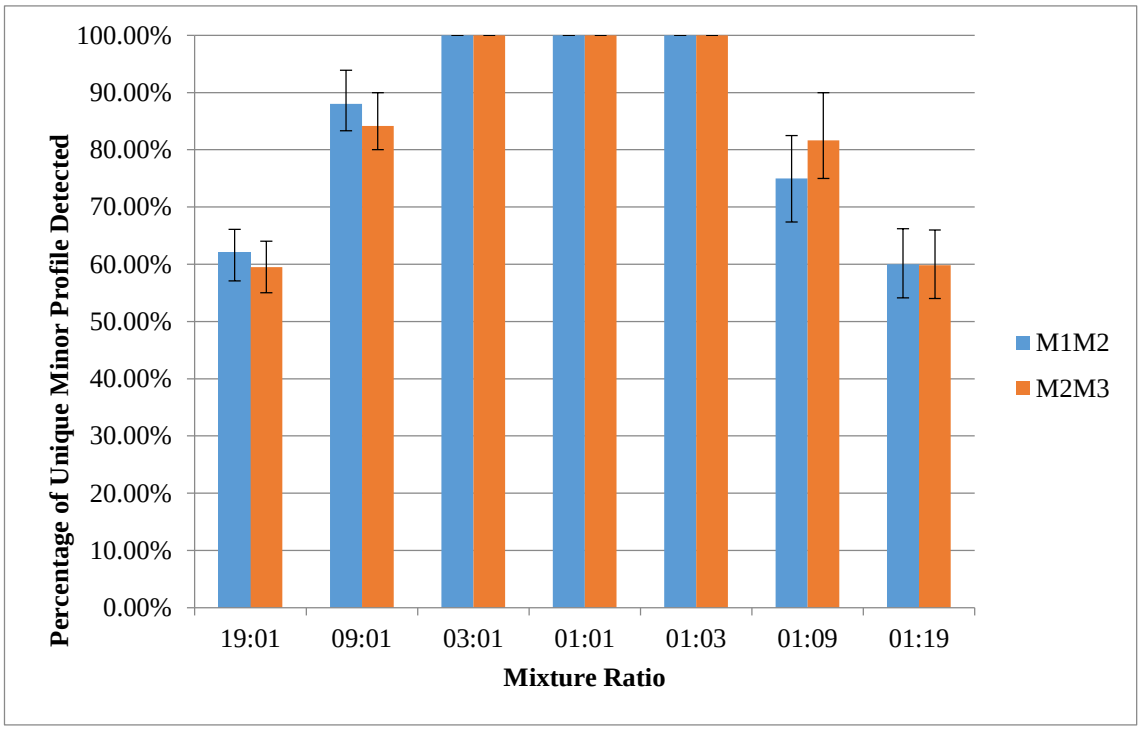

Figure 5.6: Amplification of two male/male DNA mixture sets using 1 ng of total DNA. This graph represents the average percentage of unique minor profile detected with ratios ranging from 19:1 to 1:19 for each mixture set. Error bars show the standard deviation between the three replicate amplification reactions for each male/male mixture at each ratio analysed.

Figure 5.7: Electropherograms of male/male mixtures at different ratios amplified using the RM-Yplex multiplex and analysed using an ABI 3500 Genetic Analyser. The ratio of M1 to M2 is presented in each panel.

5.2.8. Magnesium Chloride ($MgCl_2$) Titration

Three different male DNA samples were amplified, in triplicates, using 1 ng of DNA template at different concentrations of $MgCl_2$ (Figure 5.8). No artefacts observed with increasing concentrations of $MgCl_2$. Taking into consideration overall peak heights resulted in this experiment, PCR amplifications of male DNA samples have shown the best results without any additional $MgCl_2$ added to the reaction compared to PCR amplifications performed with extra amount of $MgCl_2$ except those that already existed in amplifications using the Platinum® multiplex master mix. A slight reduction in overall peak heights of the generated profile was detected at higher concentrations.

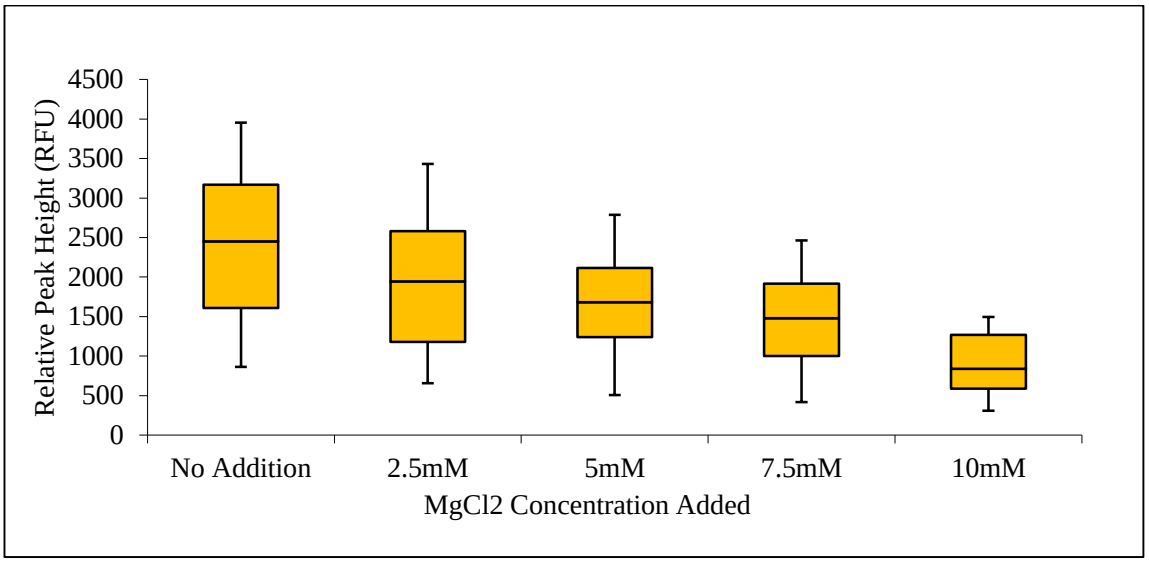

Figure 5.8: Distribution of average peak heights from RM-Yplex assay for various $MgCl_2$ concentrations added to the amplification reaction with 1 ng of male DNA template. Each box plot represents each set of 9 replicates of three samples. , The range of values from bottom quartile (25%) to the top quartile (75%) and the median is depicted an intersecting line within the boxes.

5.2.9. EDTA Titration

Three different male DNA samples were amplified, in triplicates each, using 1 ng of DNA template in presence of different concentrations of EDTA, additional to what is already available in the Platinum® multiplex master mix (Figure 5.9). There were no artefacts observed in the generated profiles with additional EDTA in the PCR amplification

reaction. A slight reduction in averaged peak heights of the generated profile was detected as the concentration of the EDTA was increased from 0.1 mM to 0.5 mM, whereas a significant reduction was observed at 1mM of additional EDTA concentration.

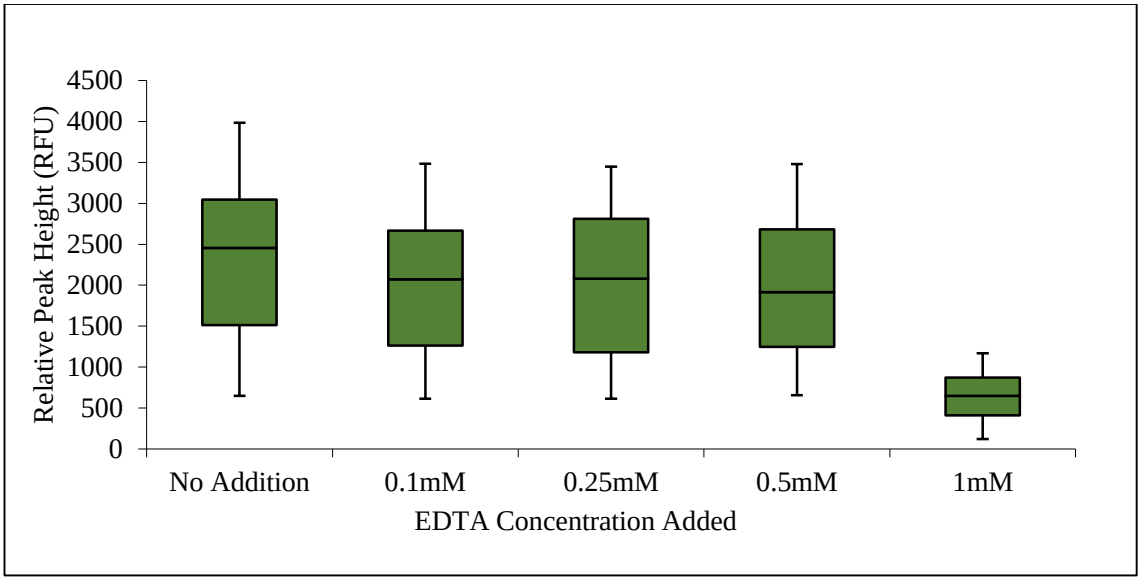

Figure 5.9: Distribution of average peak heights from RM-Yplex assay for various EDTA concentrations added to the amplification reaction with 1ng of male DNA template. The EDTA concentration increases from left to right. Results are presented using a box-plot where the boxes represent for each set of samples, 9 replicates, where the range of values from bottom quartile (25%) to the top quartile (75%) and the median is depicted an intersecting line within the boxes.

5.2.10. Stutter Analysis

Stutter peak heights were calculated from 192 samples, analysed earlier in optimisation chapter, which were electrophoresed on the ABI 3500 genetic analyser using a 3kV injection voltage and 10 seconds injection time. The averaged stutter plus three standard deviations, maximum minus one stutters, were used for the stutter filter in GeneMapper® *ID-X* (Table 5.7). All markers exhibited minus stutter peak whereas forward stutter was very rarely observed and hence was not calculated in this study DYS612 the only locus where minus two stutter was detected. Minus one stutter at DYS612 showed significantly higher stutter ratio than all other loci (Figure 5.10).

Table 5.7: Stutter characteristics for the 13 RM Y-STR loci. Stutter percentages are expressed as the percentage of the stutter peak height relative to that of the parent peak height. Maximum minus one and maximum minus two stutters were calculated by adding 3 times standard deviations to the average stutter percentage for each locus. (ND: Not Detected).

Locus	Minus One Stutter (%)	Minus Two Stutter (%)	Maximum Minus One Stutter (%)	Maximum Minus Two Stutter
DYF387S1	15	ND	36.39	ND
DYF399S1	15	ND	32.62	ND
DYF403S1a	17	ND	31.84	ND
DYF403S1b	12	ND	25.53	ND
DYF404S1	16	ND	28.76	ND
DYS449	13.5	ND	23.36	ND
DYS512	14	ND	30.26	ND
DYS526a	14	ND	29.81	ND
DYS526b	12	ND	26.19	ND
DYS547	16.5	ND	34.02	ND
DYS570	11	ND	25.67	ND
DYS576	15	ND	32.46	ND
DYS612	27	7	41.00	10.9
DYS626	12	ND	33.01	ND
DYS627	12	ND	26.86	ND

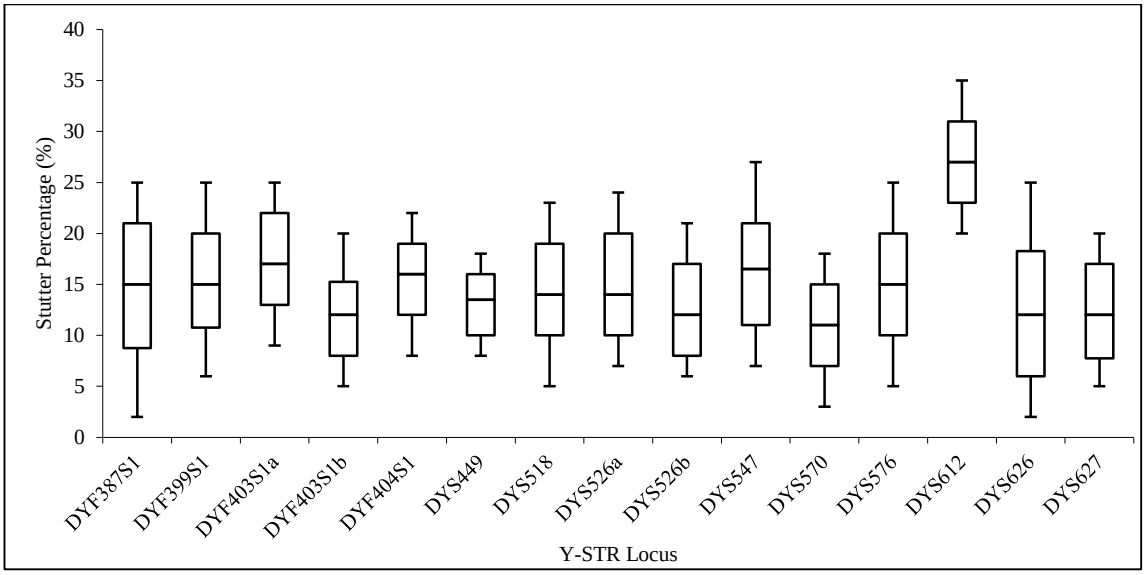

Figure 5.10: Distribution of stutter peak percentage from analysed using 192 RM Y-STR profiles generated using the developed multiplex assay. The range of values are from bottom quartile (25%) to the top quartile (75%) and the median is depicted as an intersecting line within the boxes.

5.3. Discussion

Analytical threshold determination is an important experiment, which should be conducted prior to validation. It has been shown previously, using different limits of detection might lead to wrong conclusion specially when analysing STR markers with low DNA template input (Gilder et al., 2007, Butts et al., 2011). Therefore, it was important to essential to determine the analytical threshold for RM-Yplex analysis in ABI 3500 Genetic Analyser.

5.3.1. Reproducibility

The developed multiplex assay results were completely reproducible for the internally developed QC samples and the external quality control samples. The ability of the developed multiplex to produce complete and interpretable genetic profile data was proven. Hence the multiplex can be used for forensic purposes. Data generated using RM-Yplex, therefore, would be comparable with the largest database generated for the 13 RM Y-STR loci to date. (Ballantyne et al., 2014).

The SRM2395 set of samples were successfully amplified. The F component of the set showed no amplification at all loci. This was expected as this sample was confirmed to be female DNA sample used as negative control within the SRM2395 set. This set of samples is usually used to standardise the results of amplifying same Y-STR markers using different multiplex assays. Such an experiment is valuable in order to make sure that the multiplex under investigation is working efficiently and producing consistent results. For quality assurance and standardisation purposes, haplotype references for different Y-STR markers in these samples are only analysed and reported by NIST using sequencing approach and therefore provide references to validation experiments of different Y-STR kits. The genotyping nomenclature used for these samples are follow the SWGDAM recommendations. RM Y-STR loci were recently discovered and yet there is

no validated multiplex assay to analyse these markers, apart from the one developed in this project. Moreover, there is no previous publication regarding genotypes of these loci for this set of samples apart from DYS570 and DYS576 loci. These two loci are already available in Powerplex® Y23 kit. Resulted genotypes for the two loci generated in this experiment and previously reported genotypes were consistent (Thompson et al., 2013). Therefore, genotypes resulted for SRM2395 samples in this experiment should work as a reference for any future studies of these markers. Especially, after the release of Yfiler® Plus kit which includes 4 extra RM Y-STR loci than DYS570 and DYS576.

5.3.2. Precision and Accuracy

RM-Yplex multiplex assay analysis has shown a high level of precision. The results were not significantly different when compared to the large-scale precision study done using allelic ladder (Chapter 4; Section 4.2.4). The standard deviation was between 0.02 and 0.1. Comparing this result with a previously published precision study using ABI 3500, it was found that the highest standard deviation ranged from 0.02 to 0.085 (Thompson et al., 2013). In another study the highest standard deviation was up to 0.25 (Schoske, 2003). The difference in precision in these works is most likely due to the different type of polymer, different internal size standard and the performance of the instrument. Although, the highest standard deviation in this study found to be 0.08, this is well below the accepted norm of 0.2 standard deviation (NIST, 2012). Thus RM-Yplex multiplex assay was found to be capable of generating reproducible results with high precision. Alleles were accurately called within ±0.5 bp of the same allele in the allelic ladder (Table 5.5). In each set of analyses performed, the same genotype profile was generated accurately.

5.3.3. Sensitivity and Stochastic Effects

It was noted that in a number of the Y-STR assays tested previously, the number of cycles differs from 28 to 38 cycles (Prinz et al., 1997, Thomas et al., 1999, Prinz and Sansone,

2001, Sinha et al., 2003). Even though an increase in cycle number can improve sensitivity it can result in the formation of non-specific products, which might have been absent at lower cycle numbers (Wu et al., 2010). The results of the sensitivity experiment demonstrated that the multiplex assay was sensitive to 62.5 pg of male DNA template, using 32 PCR cycles. Such sensitivity level was noted by other Y-STR multiplex systems using 30 PCR cycles (Thompson et al., 2013). It must be noted that direct comparison cannot be made to other studies because of the differences in instrumentation, cycle number, and the total number of loci simultaneously examined. However, this sensitivity study does indicate that the developed multiplex assay is as sensitive as commercially available Y-STR multiplex assays (Mulero et al., 2006, Ballantyne et al., 2013, Thompson et al., 2013). Good allele peak heights and absence of artefacts at the examined amount of male DNA template at different numbers of cycles reflect the high specificity and robustness of RM-Yplex assay.

5.3.4. Species Specificity

Because of the genetic distance from humans, markers developed for human identification will not be amplified in most other species. However, any potential cross reactivity with non-human species should be known so that profiles can be analysed correctly and considerations should be taken if necessary. Therefore, examining the human specificity of a new marker is a standard part of developmental validation of newly developed DNA profiling method. It is known that many of loci, commercially available autosomal and Y-STR multiplex assay, amplify in non-human primate species (Crouse and Schumm, 1995, Lazaruk et al., 2001, Krenke et al., 2005, Mulero et al., 2006, Mayntz-Press and Ballantyne, 2007). Hence it was reasonable to expect some of the RM Y-STR loci to be amplified in non-human primates. Four of these markers have already been evaluated for amplification in primates those are DYS576, DYS626, DYS612 and DYS449 (Erler et al., 2004, Douadi et al., 2007, Hanson and Ballantyne, 2007a).

Comparing the results of this experiment with previously published data, all markers previously found to amplify in chimpanzee, gorilla and orang-utan male DNA samples were confirmed in this study as well. In addition, DYS626 locus showed amplification in gorilla DNA sample which was not reported previously. Gibbon DNA sample showed no amplification at any locus. These results were confirmed by re-amplifying DNA samples in singleplex amplification reactions.

5.3.5. Stability Study

Forensic samples are rarely found in a prime shape and inhibitors may be presented in extracted DNA samples. In order to examine how environmental or chemical insults affect amplification of male DNA samples using RM-Yplex assay, samples were tested with the addition of three common inhibitors found in recovered biological stains in forensic case work, those are haematin, humic acid and tannic acid. Haematin is abundant in blood samples whereas humic acid and tannic acid are usually found in soil. Full profiles were obtained with 1 ng of male DNA with the presence of at most 100 ng/µl, 100 ng/µl and 200 ng/µl from each of haematin, humic acid and tannic acid respectively. The ability of resisting such amount of inhibitors by RM-Yplex assay reflects the robustness of the assay. When the concentration of haematin was increased to 200 ng/µl and 300 ng/µl many loci dropped out resulting in an average of 60% and 25% of loci called respectively. When 200 ng/µl of humic acid was added to the amplification reaction, partial profiles were detected in all replicates resulting in 88% of loci called. Addition of 300 ng/µl of humic acid resulted in partial profiles, with an average of 34% of loci called. Complete profiles were detected with the addition of 100 ng/µl and 200 ng/µl of tannic acid. Increasing the concentration to 300 ng/µl of tannic acid resulted in partial profiles for all replicates; with 88% of loci called. DYF403S1b was always showing drop-out in all partial profiles, this must be due to the fact that this marker comprises long fragment alleles which are not favoured by DNA polymerase especially

in such challenging conditions. In addition to the fact that this locus has shown low amplification level, earlier in optimisation, compared to other loci in RM-Yplex assay. On the other hand, alleles at DYF399S1 and DYF404S1 loci were always detected in all partially generated profiles. This reflects the high sensitivity of PCR conditions toward these markers and the stability of these loci. The resistance level of inhibitors established by RM-Yplex multiplex assay was slightly less than commercially available Y-STR multiplex assay (Thompson et al., 2013). Such resistance could be improved by adding PCR enhancer cocktails which will prevent inhibitors from distressing the function of the DNA polymerase enzyme and therefore such feature could be introduced to the level of amplifying DNA samples directly from biological stains which eventually will shorten the time of the analysis (Zhang et al., 2010).

5.3.6. Mixture Study

5.3.6.1. Male/Female Mixture

Mixed biological stains originating from different individuals is a common phenomenon in forensic casework. A vaginal swab, which is usually been taken from sexual assault female victim, contains a combination of female epithelial cells and male semen. Extracting DNA from mixed samples can be difficult and is not always successful (Dekairelle and Hoste, 2001, Shewale et al., 2003). In the case of a male committing the sexual assault being azospemic, doesn't ejaculate sperms, then differential extraction is not effective in separating sperms from epithelia cells (Shewale et al., 2003). Hence studying the male specificity of RM-Yplex in presence of excess female DNA template is important. Complete profiles were obtained in all male/female mixtures at different ratios, ranging from 1:250 to 1:100. A faint drop in overall allele peak heights, of RM Y-STR loci alleles, was observed as the concentration of female DNA template increase. This effect was also noted in previous studies (Mulero et al., 2006, Thompson et al.,

2013). Artefacts or non-specific alleles were not observed with high concentrations of female DNA. This reflects the high specificity of the reaction for male samples.

5.3.6.2. Male/Male Mixture

In some instances, forensic casework samples contain DNA from more than one male individual. Hence it is important to study the ability of the multiplex assay to distinguish between the major and minor profiles. Decrypting the genotypes of each individual in the mixed samples can be very challengeable especially in instances where contributors share common alleles. In such case analysis of peak height ratios becomes an important tool in differentiating between the DNA originating from minor and major contributors (Chung et al., 2004). The advantages of Y-STR has been known over years as tranquil tool in analysing males mixtures, this was based on the fact that each locus will produce only one peak in the profile and hence it is easy to detect mixture by simply looking at extra peaks appearing in some of the detected loci. However, in multiplex assays in which multi-allelic markers are used, the difficulty of decrypting the genotype of each individual is expected to be increased. The consistent balance at the multi-allelic markers, DYF399S1, DYF387S1 and DYF404S1 was an advantage of the RM-Yplex in this study as peaks height imbalance was clearly reflecting the contribution coming for another contributor. Two limitations have been observed in using RM-Yplex assay on males DNA mixtures which are, multi-allelic interpretation has become very difficult for mixture ratios above 1:3 or 3:1, whereas the other 11 single-allelic loci were successfully interpretate up to male mixture ratio of 1:19 or 19:1. A second limitation was observed earlier in the optimisation chapter when DYF403S1a was shown to exhibit inconsistent intra-locus peak heights and hence was excluded from this analysis as it will be impossible to interpretate such locus.

5.3.7. Magnesium and EDTA Titration

Magnesium is an important co-factor for proper function of DNA polymerase. Variations in the magnesium concentration may vary the performance of the DNA polymerase and hence affecting the overall performance of the amplification reaction. Magnesium is included in the Platinum® multiplex master mix and hence it can only be changed if the master mix is not properly mixed or if the magnesium is chelated. EDTA is a commonly encountered reagent that will chelate magnesium. Pipetting error and insufficient mixing could account for some fluctuations in magnesium. These studies were performed to evaluate the impact of additional magnesium or the addition of EDTA to the reaction on RM-Yplex analysis.

An addition of up to 10 mM of $MgCl_2$ in the PCR reaction did not show any drop out and all alleles were genotyped correctly in all replicates. A study of allele peak heights generated was plotted and a noticeable reduction in overall allele's peak heights was observed as the concentration of $MgCl_2$ was increasing. Therefore, the amount of magnesium already included in the platinum® master mix should be enough for optimal results; however, these results confirm the fact that presence of an $MgCl_2$ contaminant in the reaction should not affect the amplification starkly and all alleles will still be correctly genotyped.

Four different concentrations of EDTA were added to the reaction mix for a final concentration of 0.1 mM, 0.25 mM, 0.5 mM, and 1 mM. The most significant difference to EDTA was observed in the 1 mM EDTA samples as the overall peak heights dropped. The most affected loci were DYF399S1, DYF387S1 and DYF403S1. The absence of alleles drop-out at such range of EDTA demonstrated the sensitivity and robustness of

the multiplex assay. Consequently, it is recommended to store extracted DNA samples in buffers with no or low EDTA such as sterile water or TE buffer.

The ability of the RM-Yplex multiplex assay to work over big range of such components reflect the robustness of the assay which is one of the favourable features for forensic DNA profiling applications.

5.3.8. Stutter Study

A stutter is normally one repeat shorter than the parent allele and is an amplification artefact. Stutter production might depend on the number of consecutively repeated homogenous units in the locus and the size of the repeat subunit (Klintschar and Wiegand, 2003, Shinde et al., 2003). Since biological stains recovered for forensic analysis purposes might be a mixture of DNA hence presence of stutter can complicate the analysis of such mixtures. Consequently, it is essential to determine the stutter ratio for each locus in a multiplex assay destined to be used in forensic casework to establish guidelines and thresholds for analysis.

In this study, DYS612 which is a trinucleotide repeat locus, showed significantly higher average minus one stutter peak ratio of 27% when compared to other loci in the assay. All other loci have shown average ratios in the range of 10% to 20%. This result is considered slightly higher compared to previously published stutter ratios for different tetranucleotide Y-STR loci currently used in forensic DNA profiling (Mulero et al., 2006, Gross et al., 2008, Thompson et al., 2013). In another study, it was shown that typically, tetranucleotide STR loci are expected to show a stutter in the range of 5% to 15% (Brookes et al., 2012). In this study three loci including DYF403S1a, DYF404S1 and DYS576 have shown an average stutter ratio percentage slightly higher than 15%. In addition, minus-two stutter peak only appeared at DYS612 locus. Comparably, in the

previously three published multiplexes amplifying RM Y-STR, plus one stutter peak was very pronounced at some of the loci such as DYF403S1a (Victoria, 2010, Ballantyne et al., 2012). This is probably because of the differences in DNA polymerase efficiency used in each study.

It was noticed previously, the longer allele sizes will tend to result in higher stutter (Mulero et al., 2006). However, such pattern was not observed in this study. This could be explained in two ways. First, this might be a result of high number of PCR cycles used in RM-Yplex multiplex assay compared to an average of 28-30 PCR cycles mostly used for STR amplifications. Such increase in PCR cycles is expected to increase chances in the DNA polymerase slippage. Secondly, this could be a result of the Platinum® master mix efficiency especially since the composition of each component is unknown. As it has been shown previously an increase in $MgCl_2$ concentration in PCR amplification reactions will eventually increase the stutter ratio (Viguera et al., 2001). The latter suspected reason is more likely as it was shown in this validation study an increase in $MgCl_2$ concentration in RM-Yplex multiplex assay will affect amplification negatively, whereas it was shown previously increasing the concentration of $MgCl_2$ gradually will increase the efficiency of amplification before it drops down at very high concentrations (Mulero et al., 2006). This might suggest that the $MgCl_2$ content in Platinum® master mix is more than what have been used in the study which concluded the stutter range of 5% to 15%. To sum up, a stutter threshold was established for each locus by taking the average value and adding three standard deviations for each locus.

6. Forensic Applications of RM-Yplex

6.1. Overview

Y-STR DNA profiling has become an invaluable tool in forensic casework. DNA analysis methods are improving in sensitivity and precision, enhancing the evidential value of DNA profiling. Statistics at the Department of Forensic Sciences and Criminology (DFSC) at Dubai Police show that males commit the vast majority of violent crimes including murder and sexual assaults. DNA samples collected in most rape cases contain mixtures of DNA from both the male perpetrator and the female victim. In these samples, the female DNA is often in vast excess, and as a result, male DNA may not be detected using an autosomal STR amplification assay or the sperm fraction might not be successfully separated. In some cases there would be no sperm fraction available due to the perpetrator being azoospermic. In cases where autosomal STR analysis fails to give probative results, Y-STR analysis would be appropriate as it will only amplify male DNA in the sample. This study was conducted as part of the validation process in order to evaluate the RM-Yplex assay if it meets the standards for use in a forensic DNA profiling laboratory. According to SWGDAM validation guidelines 2012, internal validation could be defined as a set of experiments carried out at specific laboratory in order to show that a developed DNA analysis method is performing as expected prior to being applied in forensic casework. Based on the results of these experiments, quality assurance parameters, limitations and interpretation guidelines will be established. In this chapter the study of non-probative casework samples is also presented as well as analysis of close relative male individuals from two typical Arab families from UAE.

All PCR amplifications were performed according to optimized conditions (Table 2.7; Table 2.8). Experiments, recommended by SWGDAM, were conducted in order to demonstrate the performance of RM-Yplex in the DFSC. Sensitivity, specificity and mixture studies were repeated by three different experienced analysts. The results were

analysed and compared with the same type of experiments conducted in developmental validation.

6.2. Results

6.2.1. Analytical Threshold

Minimum threshold determination was done using the data generated for 9 replicates of 3 sperm-negative vaginal swabs obtained from DFSC for validation purposes. Highest noise peak resulted at each dye within alleles detection range (100 bp to 500 bp) was recorded in each profile generated; the peak heights were averaged and standard deviation was calculated for each dye. The orange channel showed the lowest average 9.18 RFU and the highest was obtained for the green channel (Table 6.1). An example of profiles generated using sperm-negative vaginal swabs samples used in calculating the LOD and LOQ values is shown in Figure 6.1. Comparing the heights of highest noise peaks detected for each dye in this study with the ones estimated in developmental validation study using independent t-test, there was no significant difference detected according to P-value obtained (Table 6.2). By comparing means calculated in each experiment using independent t-test, there was also no significant difference between means generated in the two experiments (t-stat = 1.397, df = 4, P = 0.2349).

Table 6.1: Dye-specific limit of detection (LOD) and limit of quantitation (LOQ) analysis results.

Channel	Maximum RFU	Average RFU	Standard Deviation (SD)	Averaged RFU +3 SD (LOD)	Averaged RFU +10 SD (LOQ)
Blue	26	15.18	7.09	36.44	86.05
Green	23	15.24	6.81	35.66	83.30
Yellow	24	15.12	6.68	35.17	81.96
Red	29	14.85	6.02	32.91	75.06
Orange	14	9.18	3.14	18.59	40.55

Table 6.2: Dye-specific limit of detection statistical comparison results using independent t-test between average heights of non-specific peaks estimated in this study (Int.) compared to developmental validation study (Dev.).

Channel	Average RFU Dev.	Average RFU Int.	t-stat	P-value
Blue	15.57	15.18	0.012	0.991
Green	15.06	15.24	1.082	0.279
Yellow	14.94	15.12	0.198	0.843
Red	15.41	14.85	0.022	0.983
Orange	10.29	9.18	1.219	0.230

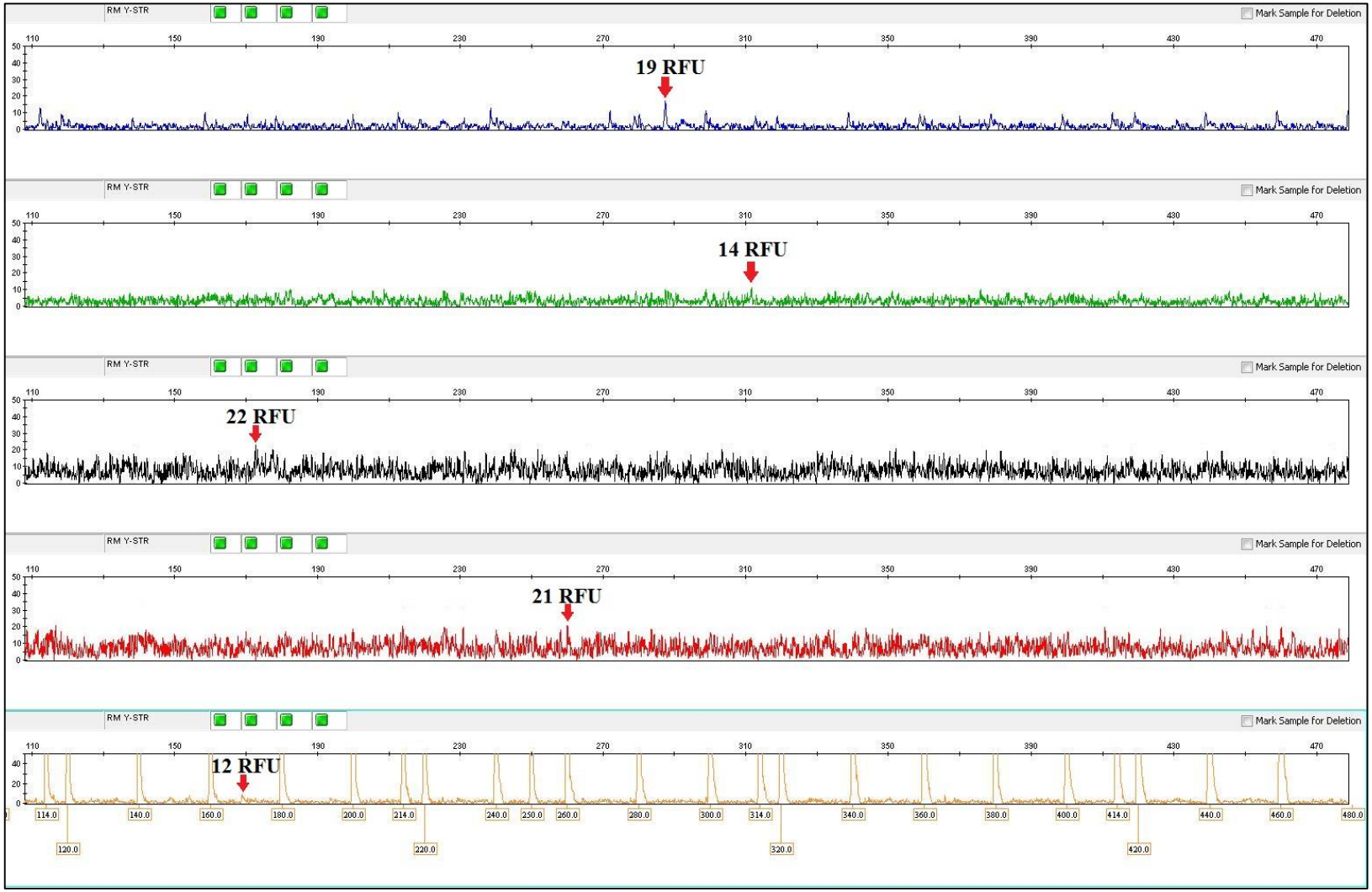

Figure 6.1: Baseline noise as appeared in the profile generated from PCR amplifying negative vaginal swab samples.

6.2.2. Accuracy and Precision

Six samples were provided to each of the three analysts along with the SRM2395 sample set in order to evaluate the reproducibility of the assay. All samples, of which female DNA were included as negative controls were provided blind to each analyst. Results of all samples were consistent among three analysts and also consistent with results obtained in the developmental validation experiment previously described in (chapter 5; Section 5.3.2).

In the precision study, allele sizes were averaged and a standard deviation was calculated for each distinct allele at each locus for all samples. An average standard deviation was determined to be 0.046, whereas the minimum and maximum standard deviations recorded were 0.021 at DYS570 and 0.078 at DYF403S1a respectively. There was no correlation found between allele's sizes and standard deviation when the values were plotted against each other (Figure 6.2). Sizing precision of each allele was less than ±0.3 base pairs from the average fragment size determined. POP-4 was used in the DFSC DNA profiling laboratory using 3500 Genetic Analyser instead of POP-6 that was used in the UCLan laboratory. A comparison between two sets of standard deviations values, the ones generated in this study and the ones generated in Chapter 5, showed that there was a significant difference between the precision study conducted using POP-6 (Mean = 0.049, SD = 0.021) and the precision study conducted using POP-4 (Mean = 0.046, SD = 0.016) according to the P-value obtained using independent t-test statistical analysis (t-stat = -2.62, df = 179, P-value = 0.009386). Higher precision was realised in the analysis of RM-Yplex assay using POP-4 compared to POP-6 (Figure 6.3).

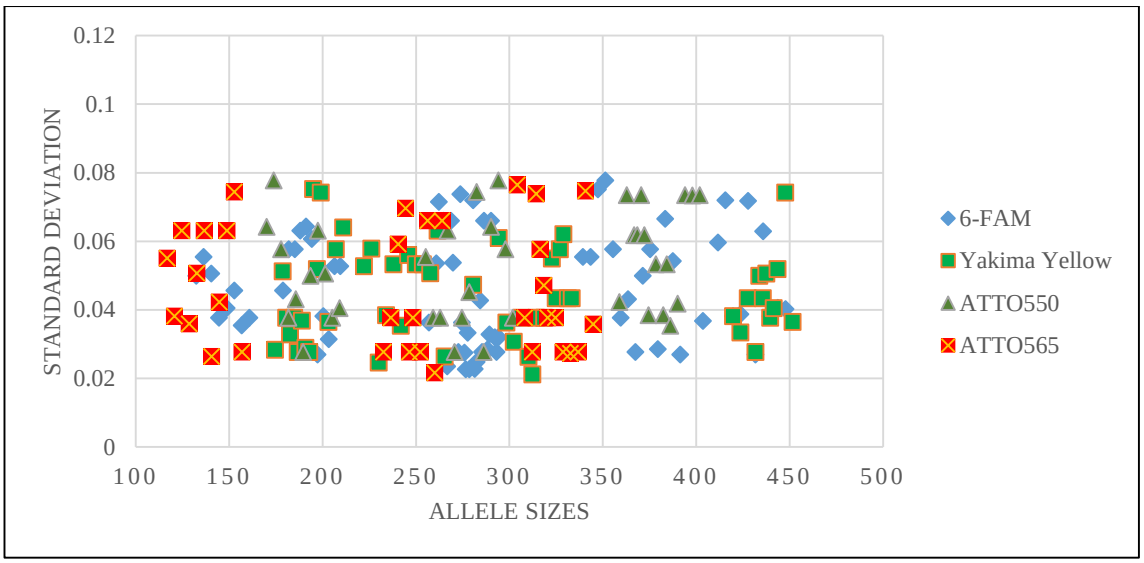

Figure 6.2: Standard deviations all fragments sizes generated through the analysis of the 96 male DNA samples. 180 data points were plotted in this graph.

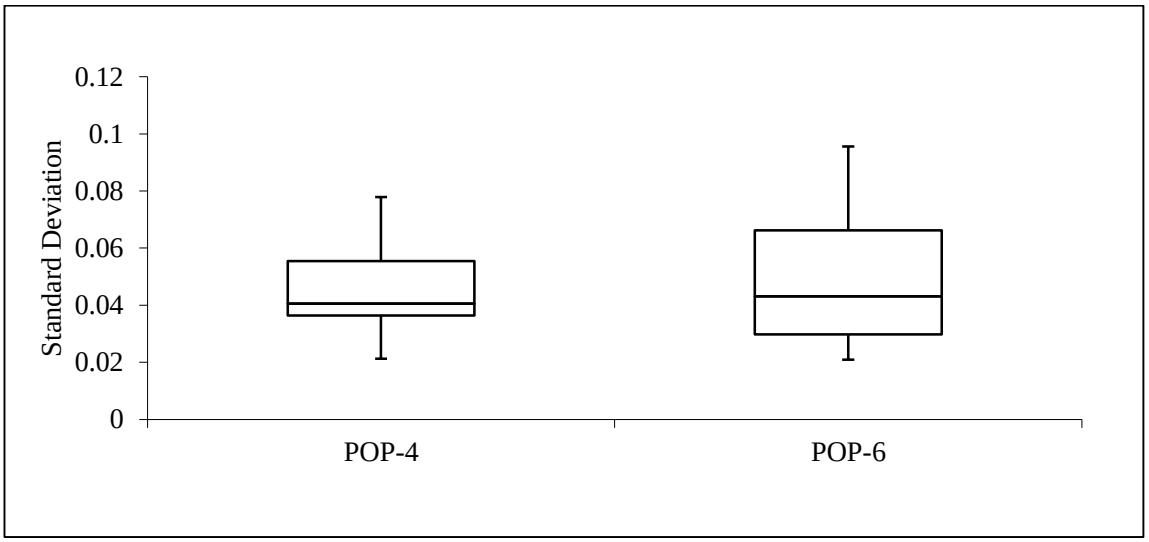

Figure 6.3: Box-plot showing a comparison of the standard deviation of two precision studies conducted using POP-4 and POP-6 polymers. POP-6 was used in developmental validation and POP-4 was used in casework study application study at DFSC DNA profiling laboratory using 3500 Genetic Analyser.

6.2.3. Sensitivity and Specificity

In the sensitivity study, the average number of successfully called alleles for the 3 replicates of each sample set was calculated along with standard deviations (Figure 6.4). A full profiles were generated consistently for all replicates of samples using 1000 pg down to 125 pg of male DNA template. Two replicates of the same male DNA sample,

analysed by two different analysts showed allelic drop-outs at DYF403S1b when 62.5 pg male DNA template input was used. A maximum of 13 alleles and a minimum of 5 out of 21 alleles, from different loci, dropped-out when 31.25 pg of male DNA input was used in the PCR amplification reaction.

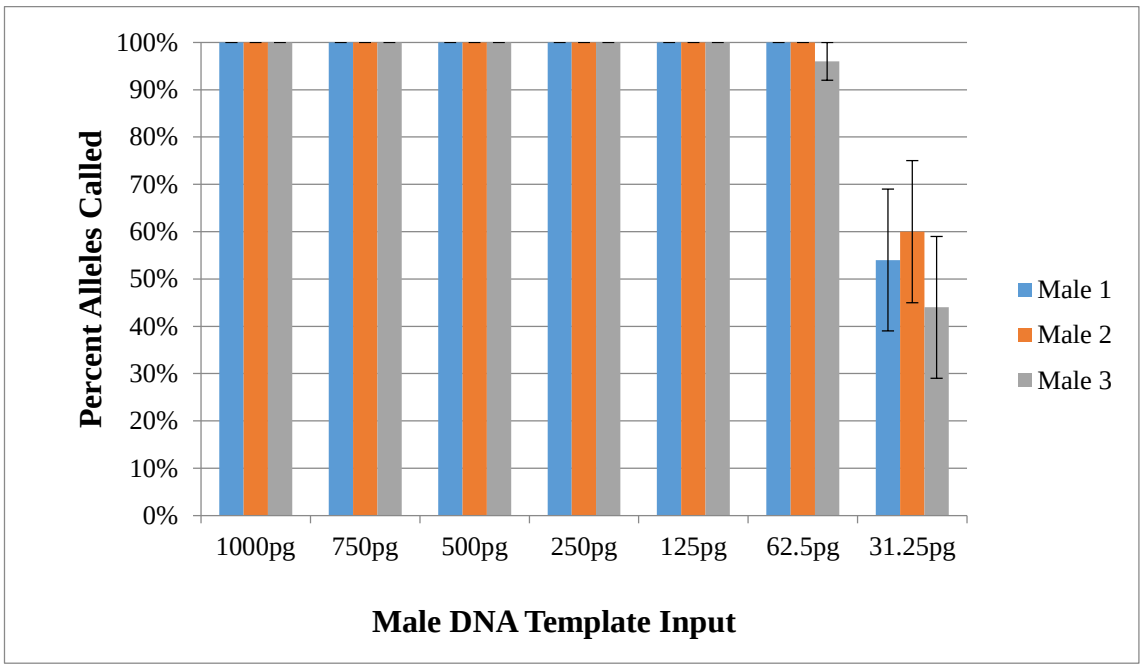

Figure 6.4: Human male DNA amplification reactions containing different amounts of DNA template from different individuals. The percentage of alleles called for male 1, male 2 and male 3 are represented in blue, orange and grey bars colours respectively. Error bars show the standard deviation of the three replicate amplification reactions.

Male DNA sample, 9948 was successfully amplified in an excess amount of each of the three female DNA samples at different concentrations. Non-specific amplification was not observed from the mixtures analysed during this study. There were also no artefacts detected in any of the profiles generated. A full profile resulted from amplification of 1 ng of 9948 male DNA control in presence of 2000 ng of female DNA (Figure 6.5).

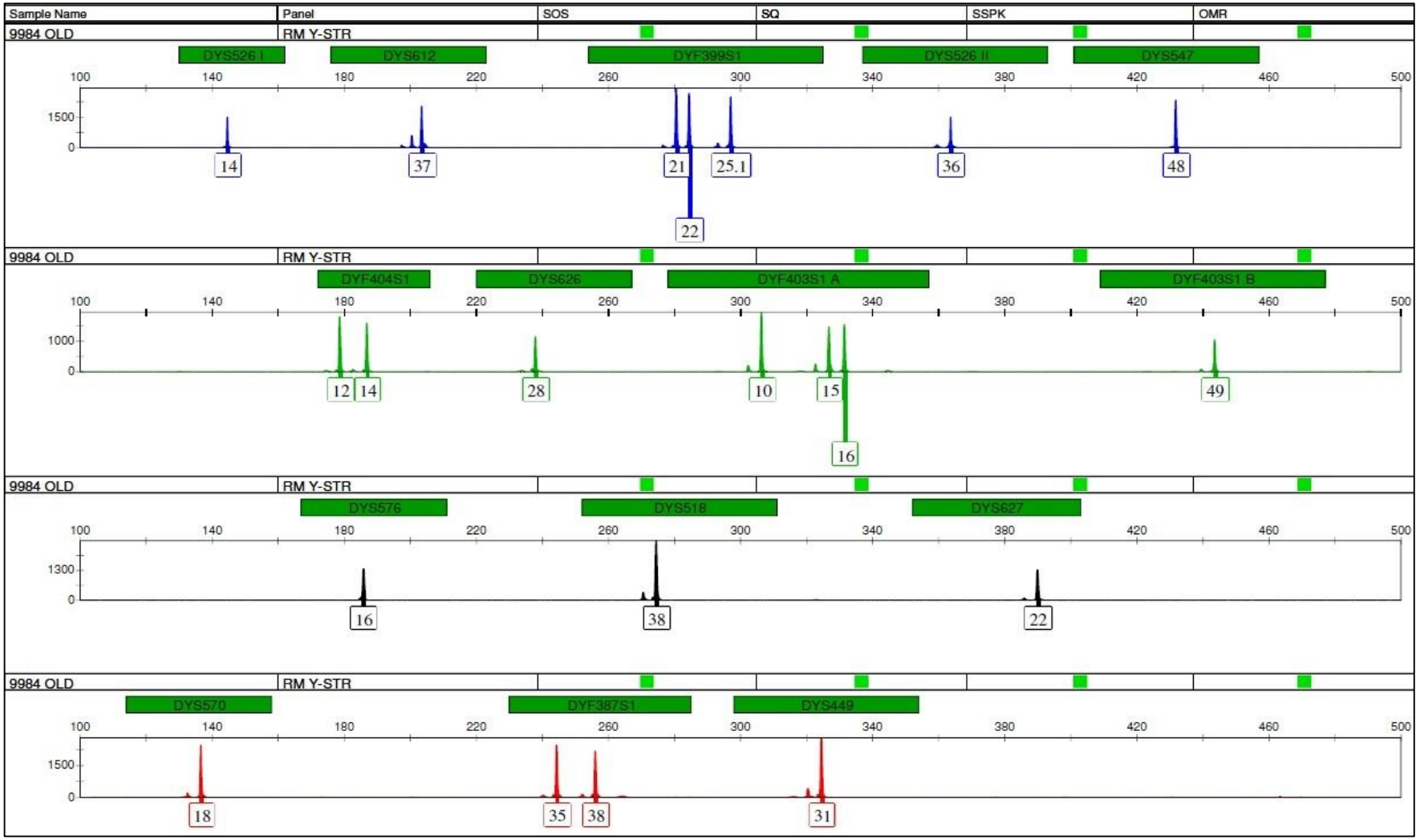

Figure 6.5: RM-Yplex profile generated using ABI 3500 Genetic Analyser for 9948 male DNA control in presence of 2000ng of female DNA.

6.2.4. Stutter Analysis

Stutters are usually caused by the slippage of the DNA polymerase enzyme during amplification of repetitive region and most commonly appear as 1 repeat smaller than the allelic peak. Table 6.3 represent the range of stutter percentages for each locus resulting from the analysis of 96 male DNA samples, injected twice. The highest average stutter of 25.6% was calculated for DYS612 which is a trinucleotide marker and hence stutter is expected to be significantly higher than tetranucleotide loci which have shown stutter percentages in the range of 7.9 % to 13.6 %. Maximum stutter limit was established by adding 3 standard deviations to the average stutter percentages results. This value was used as a cut-off for expected stutter. A forward stutter was hardly observed for any of the loci.

Table 6.3: Percentage stutter values obtained for the RM-Yplex assay.*Analysis based on at least 192 data point at each locus.

Locus	Average Stutter (%)*	Standard Deviation	Stutter Cut-off Ratio (%)
DYF387S1	10.2	0.53	11.8
DYF399S1	12.5	0.47	13.9
DYF403S1a	9.10	0.67	11.1
DYF403S1b	10.1	0.47	11.5
DYF404S1	13.6	0.60	15.4
DYS449	11.4	0.67	13.4
DYS518	11.6	0.50	13.1
DYS526a	12.1	0.50	13.6
DYS526b	10.8	1.10	14.1
DYS547	8.5	0.80	10.9
DYS570	7.9	0.77	10.2
DYS576	8.10	0.73	10.3
DYS612	25.6	0.50	27.1
DYS626	12.7	0.50	14.2
DYS627	11.3	1.13	14.7

6.2.5. Male/Male Mixtures

Three different sources of male DNA were used to develop two sets of mixed male DNA samples (M1:M2 and M2:M3) at different ratios to evaluate mixture interpretation. DYF403S1a locus was excluded from the analysis based on the optimisation results which showed a high level of peak height imbalance at this locus. A mixture was detected in all samples tested by different analysts by the presence of one extra allele at least in two of the 11 single allelic markers. All alleles of both sets of mixtures analysed for both minor and major contributors at 1:3 and 3:1 ratios were detected by all analysts as illustrated in Figure 6.6. Only at such ratios, profiles of major and minor contributes were resolvable at least using the single allelic loci. Alleles from both contributors were also detectable in all replicates at 1:1 ratio. At this ratio, even though all alleles were detected, alleles remained unresolvable into minor and major contributors. DYF399S1 and DYF404S1 were the most sensitive loci in the assay and hence, alleles were always detectable at all different ratios tested, whereas the least sensitive loci were DYF403S1a/b and DYS547. DYF403S1b minor allele was always dropping-out in both sets at the ratios of 19:1, 9:1, 1:9 and 1:19.

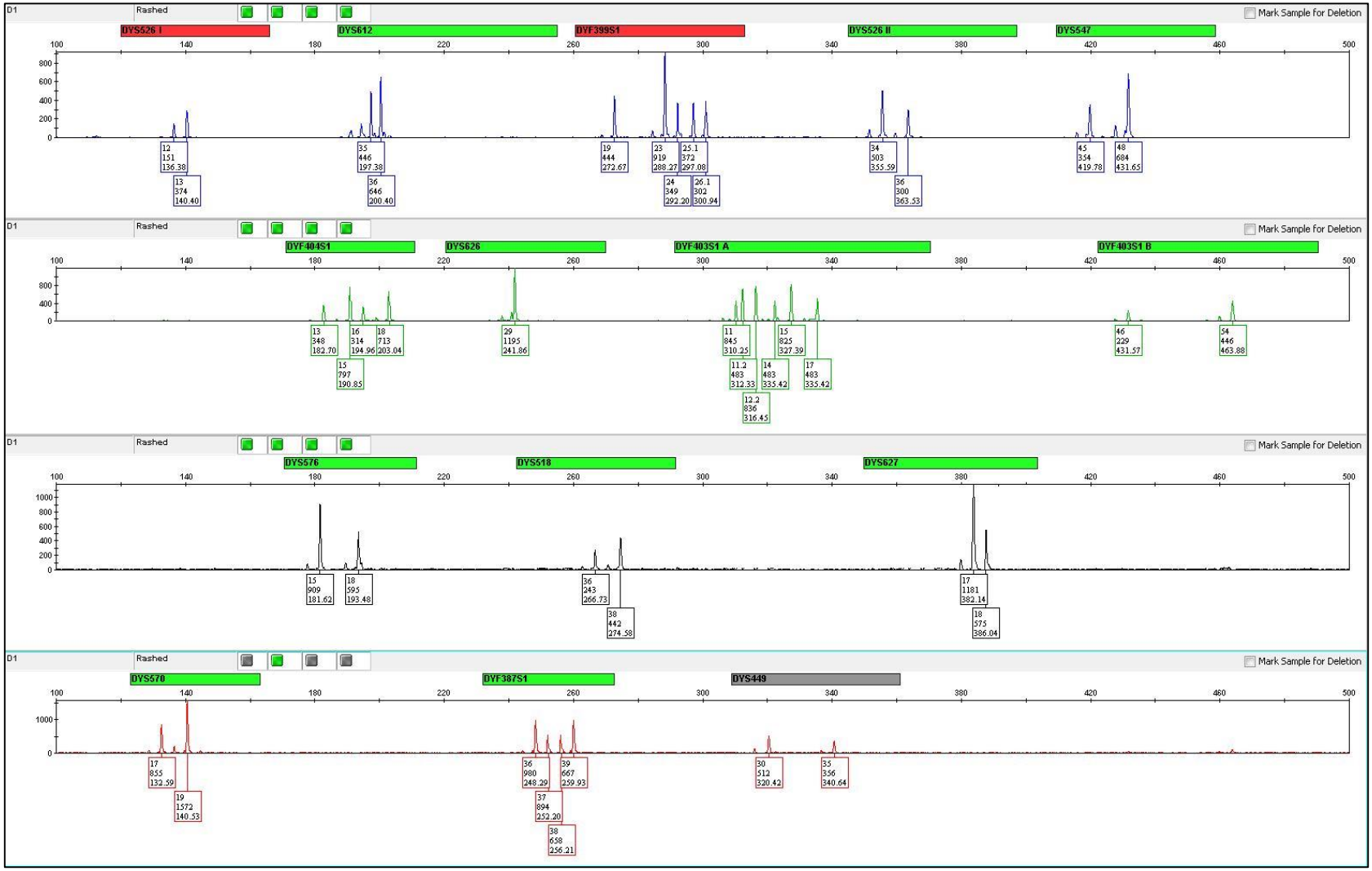

Figure 6.6: Mixture profile of two male DNA (M1and M2) co-amplified with a 3:1 ratio.

6.2.6. Non-probative Casework Samples

25 samples from 14 different cases were analysed during this study (Table 6.4). Results had been obtained earlier using Identifiler® and Yfiler® (Applied Biosystems) amplification kits. Quantitation ratios of Male/Female were determined from results obtained for each fraction using Quantifiler® Duo human quantification assay.

Optimised conditions were used for RM-Yplex amplification (Chapter 2; Table 2.7 and Table 2.8). In sexual assault cases, full profiles were obtained when sperm fractions of 8 vaginal swabs were analysed using RM-Yplex assay with male DNA input ranging from 120 pg to 2000 pg whereas one full profile and two partial profiles were detected out of 5 epithelial fractions of vaginal swab samples and the rest showed no amplification at all.

Perineal swabs samples from three different sexual assault cases were successfully amplified in the presence of excess amount of victim female DNA. In addition, 4 underwear samples from 4 different sexual assault cases were analysed using RM-Yplex. 3 of them were successfully amplified and generated full profiles whereas one of them resulted in 6 alleles dropping-out when 1 ng of male DNA was used as a template in the PCR reaction in the presence of 8000 ng of female DNA. Moreover, semen deposited on bedding cotton sheets were recovered from crime scene of sexual assault case was successfully amplified using RM-Yplex assay and a full profile was generated.

A remarkable sexual assault case (No. 2) was analysed, where two male relatives were suspected for the rape of female victim. Results obtained using Identifiler® showed a mixture of at least 3 contributors (Figure 6.7), due to the complications involved in this case, it was deemed to be inconclusive. Yfiler® had shown a full male profile which had matched both suspects' reference samples indicating a close paternal relationship (Figure 6.8). RM-Yplex amplification result analysis showed two balanced peaks at DYF403S1b one repeat apart from each other and three peaks detected at DYF404S1 showed intra-

locus imbalance indicating a presence of mixture as shown in Figure 6.9. Reference samples from both suspects were analysed and a unique profile was generated for each individual. Comparison of the mixture profile with reference samples from two suspects, indicated that the possible source of mixture was a mixture of both males DNA samples.

In a missing person case, a tooth was found during search and recovery, which was suspected belonged to the missing person. DNA was extracted and amplified previously using both Identifiler® and Yfiler®, along with reference sample from the brother of the missing person and brotherhood was confirmed. In an attempt to show the performance of RM-Yplex, a full profile was generated using previously organically extracted DNA from the tooth, which was 6 months old stored at -20°C, for RM-Yplex and matched the reference sample of the brother.

In a manslaughter case, two samples were analysed including nails of the female victim and a cigarette butt where perpetrator DNA was expected to be found. Both nails and cigarette butt were tainted with victim's blood. Quantitation resulted in detection of low amount of male DNA and in an attempt to amplify the sample using Identifiler® both samples showed full victim profile with Y Amelogenin peak just above detection threshold. Using the same extracted samples, RM-Yplex amplification showed full RM-Yplex profile.

Table 6.4: Results obtained from non-probative casework samples analysed by RM-Yplex. *(E: Epethilia Fraction; S:Sperm Fraction; NP: No Profile; FP: Full Profile; PP: Partial Profile) Deficiency of RM-Yplex indicated in Bold.

Case #	Case Type	Sample Type*	Identifiler® Results	Amount of male DNA amplified	Male:Female ratio	Yfiler® Results	RM-Yplex Results
1	Sexual Assault	Vaginal Swab - E	Victim Profile	420pg	1:2100	N/A	FP
		Vaginal Swab - S	Victim and suspect profiles (resolved)	1000pg	1:2	N/A	FP
		Perineal Swab	Victim Profile	760pg	1:494	N/A	FP
2	Sexual Assault	Vaginal Swab - E	Victim Profile	45pg	1:6300	FP	**NP**
		Vaginal Swab - S	Mixture of at least 3 contributors	1200pg	1:1	FP	Mixture of at least two males
		Perineal Swab	Victim Profile	840pg	1:1360	N/A	FP
3	Sexual Assault	Vaginal Swab - E	Victim Profile	No detection	-	N/A	**NP**
		Vaginal Swab - S	Victim and suspect profiles (resolved)	120pg	1:240	N/A	FP
4	Sexual Assault	Vaginal Swab - E	Victim Profile	70pg	1:1000	N/A	**PP(4 loci drop out)**
		Vaginal Swab - S	Victim and suspect profiles (resolved)	900pg	1:20	N/A	FP
5	Sexual Assault	Vaginal Swab - S	Suspect profile	2000pg	All male DNA	N/A	FP
		Perineal Swab	Not attempt	1500pg	1:10	N/A	FP
6	Sexual Assault	Vaginal Swab - E	Victim Profile	160pg	1:9500	N/A	**PP(10 loci drop out**)
		Vaginal Swab - S	Victim Profile	1000pg	1:15	N/A	FP
		Underwear	Victim Profile	1000pg	1:8000	N/A	**PP(6 loci drop out)**
7	Sexual Assault	Vaginal Swab - S	Mixture with partial suspect profiles	1000pg	1:3	N/A	FP
		Underwear-S	Victim Profile	240pg	1:90	N/A	FP
8	Sexual Assault	Vaginal Swab - S	Suspect profile	1000pg	All Male DNA	N/A	FP
9	Sexual Assault	Underwear	Not attempt	250pg	1:500	N/A	FP
10	Sexual Assault	Underwear	Mixture victim and suspect profiles (resolve)	1400pg	1:3	N/A	FP
11	Sexual Assault	Semen on beddings (cotton)	Only suspect profile detected	1000pg	All male DNA	N/A	FP
12	Manslaughter	Nails (Mixed with victim blood)	Only Victim Profile detected	1000pg	1:500	N/A	FP
		Cigarette butt (Mixed with victim blood)	Only Victim Profile detected	1000pg	1:700	N/A	FP
13	Missing person	Tooth	Full Male Profile	1000pg	All male DNA	FP	FP
14	Sexual Assault	Nails	Mixture victim and suspect profiles (resolved)	1000pg	1:100	N/A	FP

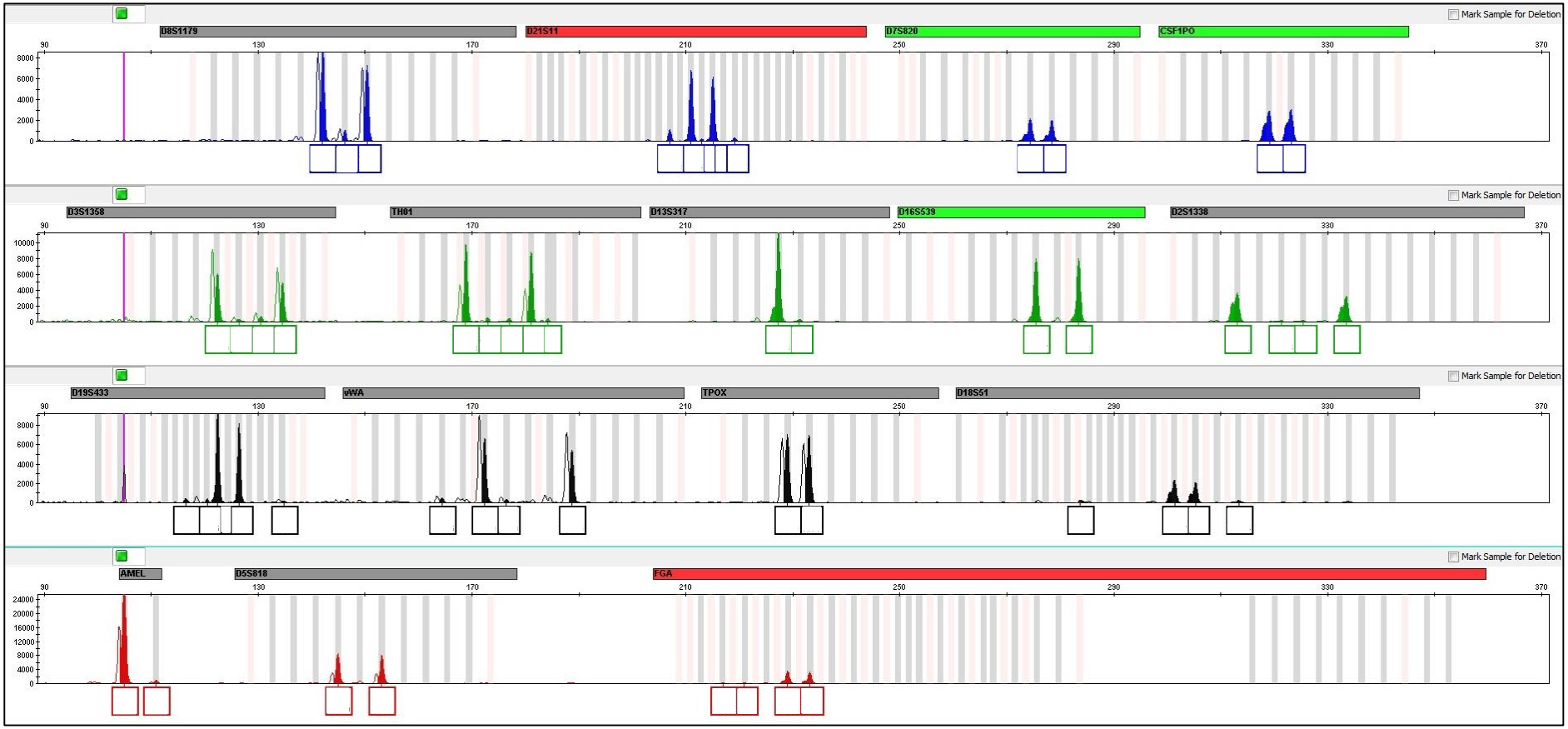

Figure 6.7: A profile showing a mixture of at least 3 individuals from sample recovered in case no. 2, generated using Identifiler® amplification kit. (Alleles calling are blank for data protection).

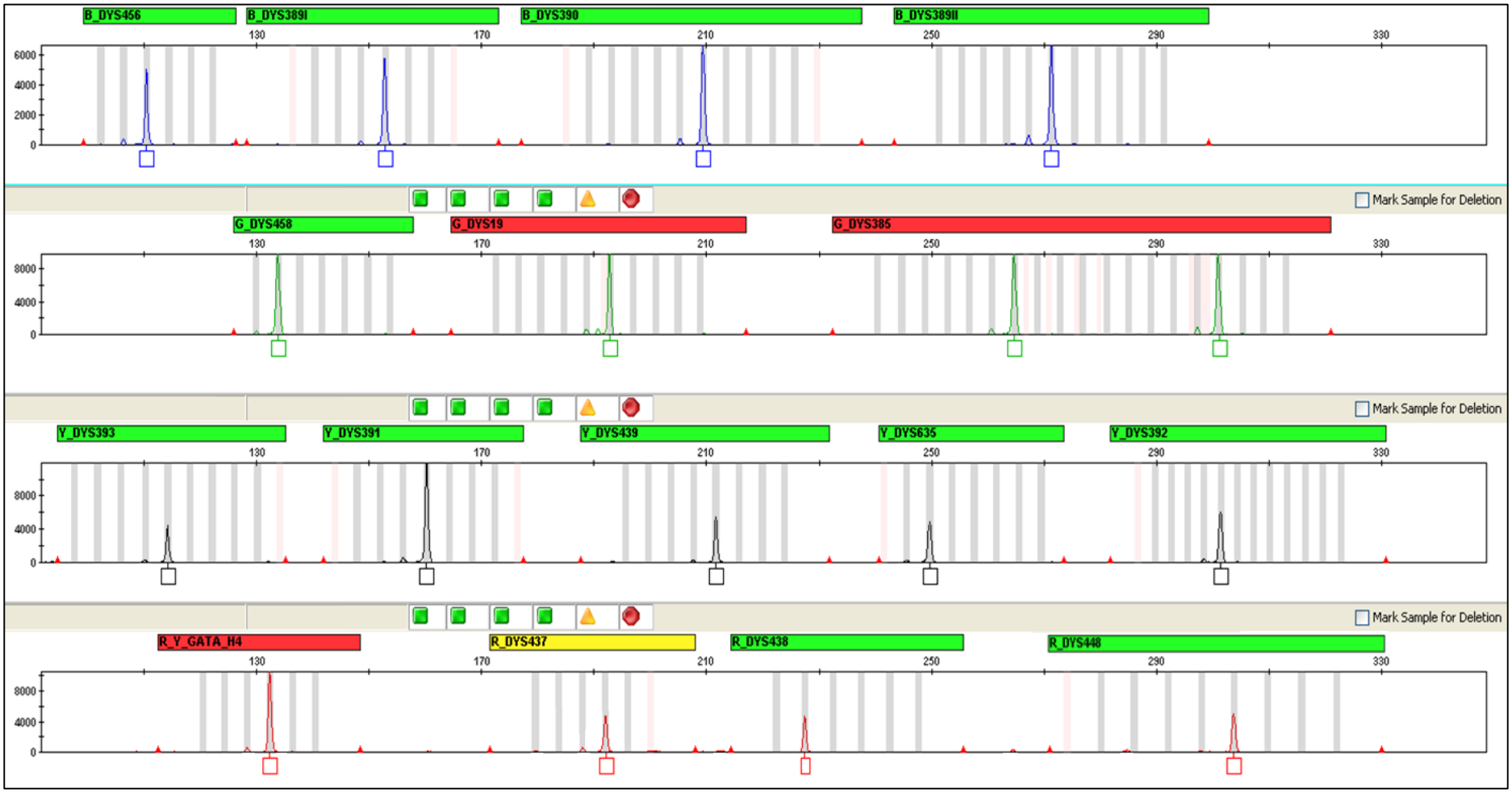

Figure 6.8: A Yfiler® profile showing single source male DNA amplification results Case No. 2. (Alleles calling are blank for data protection).

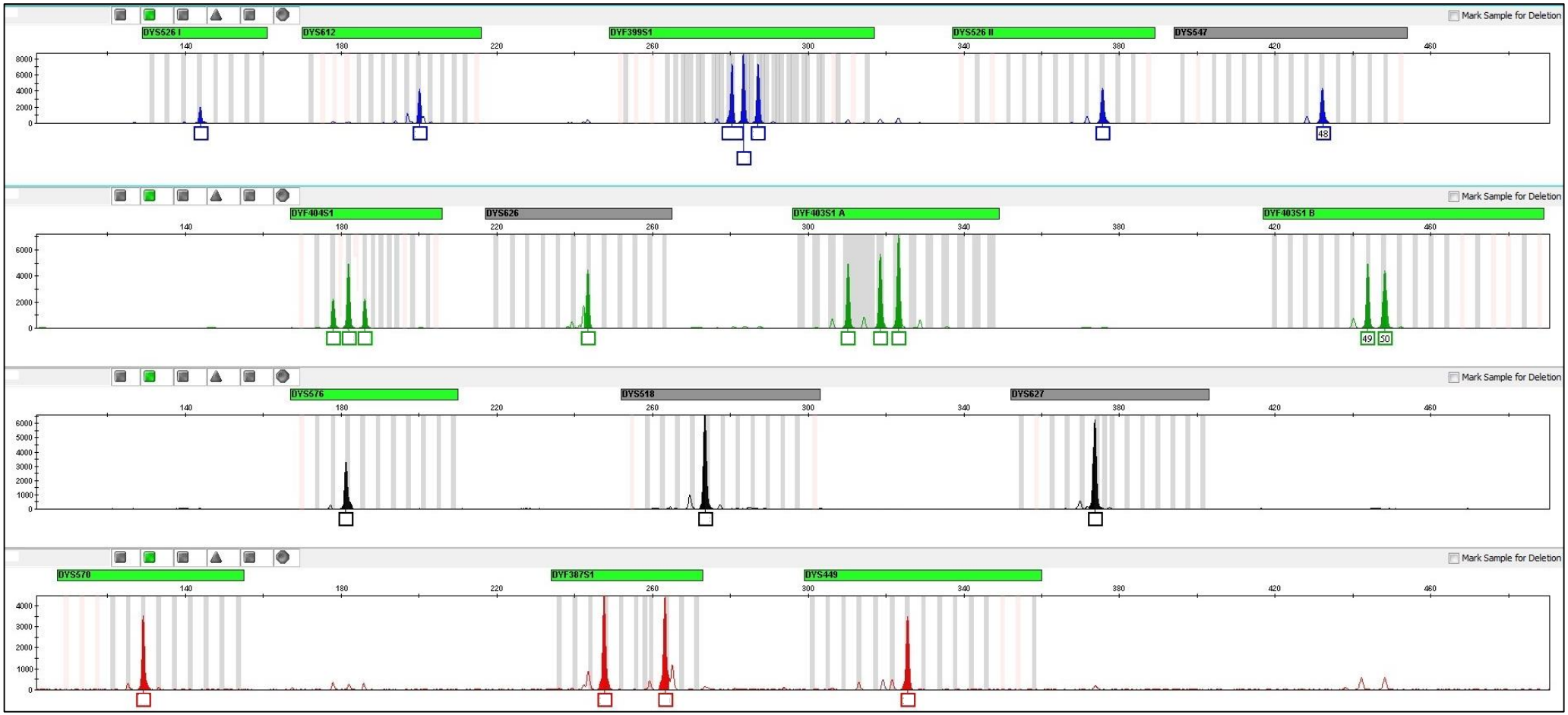

Figure 6.9: A RM-Yplex profile showing a male:male mixture generated from the sperm fraction of vaginal swab obtained from a casework sample analysed in case No. 2 . (Alleles calling are blank for data protection).

6.2.7. Families Pedigree Analysis

Two families pedigree were analysed in order to demonstrate the ability of the RM-Yplex assay to differentiate between close male relatives (Figure 6.10 and 6.11).

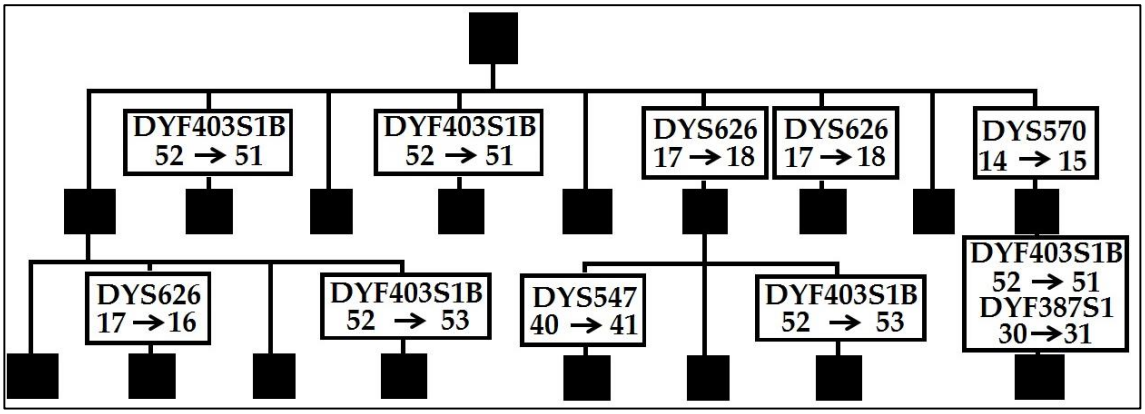

Figure 6.10: RM Y-STR mutation pattern and abilities to differentiate closely related males in an UAE Arab family 1 analysed with the RM-Yplex assay targeting 13 RM Y-STR markers. Only RM Y-STRs with observed mutations are shown with the mutated alleles.

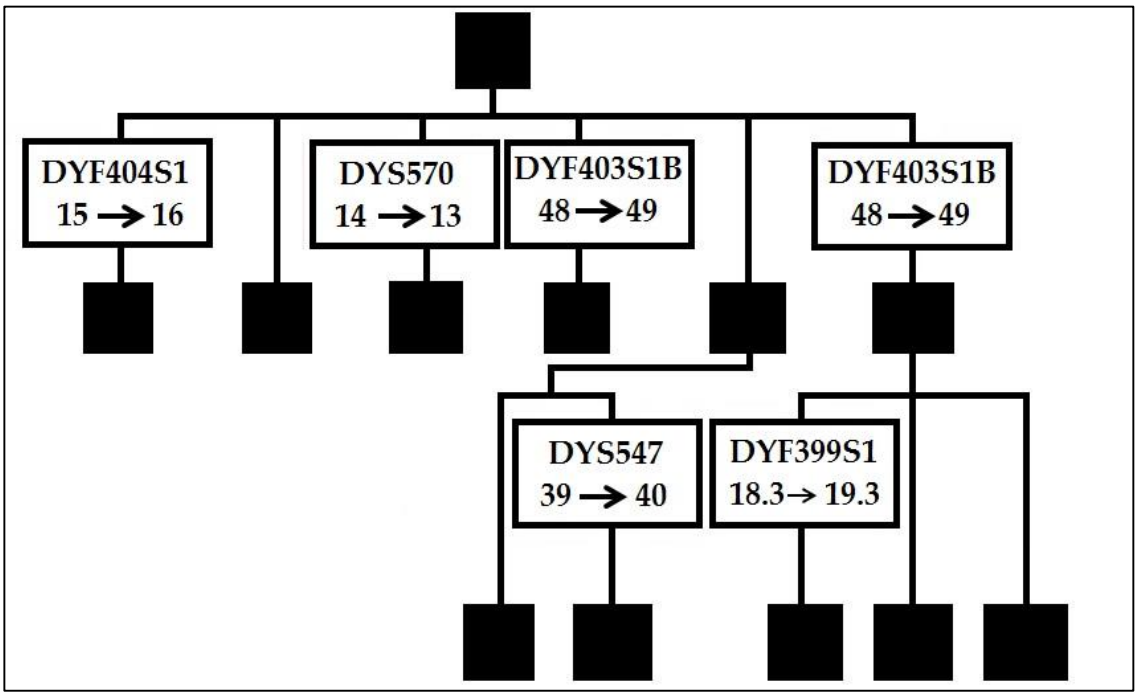

Figure 6.11: RM Y-STR mutation pattern and abilities to differentiate closely related males in an UAE Arab family 2 analysed with the RM-Yplex assay targeting 13 RM Y-STR markers. Only RM Y-STRs with observed mutations are shown with the mutated alleles.

In family no. 1, out of 18 male individuals, there were 11 unique haplotypes found which means 60% of the males were differentiated. On the other hand, in family 2, out of 12

closely relative males individuals, there were 7 unique haplotypes which means 58% of the males were differentiated. DYF403S1b marker recorded the highest number of mutations in both families. In family no. 1 there were 5 mutations detected at DYF403S1b whereas family no. 2 there were 2 mutations detected. DYS626 marker was the second highest mutating marker in family one. All other mutations detected at mutated markers were observed once only in both families.

6.3. Discussion

Prior to the implementation of a new multiplex assay into routine casework analysis, forensic laboratories are required to perform an internal validation study (NIST, 2012). This validation study is important in establishing the limitations of the analysis method and also to assist in the development of the specific interpretation guidelines accordingly. Internal validation helps to demonstrate particular laboratory's capability in successfully using a multiplex assay in its own environment which is suitable for casework (NIST, 2012).

6.3.1. Analytical Threshold

Analytical threshold has been determined following the validation guidelines provided by SWGDAM using sperm-negative vaginal swab samples. As suggested by Gilder et al. (2007), two types of analytical threshold were calculated in this study those are limit of detection (LOD) and limit of quantitation (LOQ). Limit of detection is the threshold above which any peak could be confidently interpreted as a true allelic peak whereas limit of quantitation (LOQ) is a limit below which measurements of signal strength cannot be reliably used in interpretation of the profile, especially peak height imbalances (Gilder et al., 2007). LOD and LOQ were set up to be 50 RFU and 100 RFU respectively for all dye channels after analysing the results of the samples used for determining the analytical

thresholds. Results found in this study showed no statistical difference in means values for each dye compared to those obtained in developmental validation study.

These results are significantly lower than previously published analytical thresholds for other assays such as Identifiler® and Identifiler® plus amplification kits where the thresholds were found to be 120 and 140 RFU respectively using 3500 Genetic Analyser instrument (Butts et al., 2011). In the same study it had been realised previously that using a single threshold will result in more allelic drop-outs than dye-specific thresholds (Butts et al., 2011). However in this study there was no statistically significant difference between average limits of detection for each of dyes included in the multiplex assay (Ellison and Barwick, 2009). Differences between the two studies may be due to the PCR amplification reaction conditions and the sensitivity fluctuations across different 3500 Genetic Analyser instruments.

6.3.2. Accuracy and Precision

The study of accuracy is very important in forensic DNA profiling. In this experiment, a total of 12 samples blindly provided to three different analysts working at DFSC. The reproducibility of accurate results has been demonstrated across analysts, which also were consistent with previous analysis of the samples done during the developmental validation. Hence the reliability of producing accurate genotyping results using RM-Yplex was achieved.

Data generated from twice-injected 96 male DNA samples was used to illustrate the size precision resulting of the RM-Yplex analysis on 3500 Genetic Analyser. There was a significant difference found between standard deviations calculated in this study compared to those calculated in developmental validation of RM-Yplex using POP-6 with 50 cm capillary array. These results showing the importance of internal validation for any newly implemented method in forensic DNA profiling laboratory as precision might

differ across different analytical instruments such as ABI 3500 Genetic Analysers and also between different type of polymers. Such difference is most likely due to the different type of polymer used in each experiment. In previous study, as well as in optimisation stage of RM-Yplex, POP-6 had demonstrated higher resolution than POP-4 and hence differences in sizes representing distinct allele will be detected more often in POP-6 than POP-4 (Detwiler et al., 2004). Using different male DNA samples in precision study will take into account not only variations resulting out of electropherometric migrations of different alleles with same length and nucleotide sequence, however, it also has the advantages of taking into account variations resulting out of alleles with same length but different nucleotide sequence. Such variation was noticed in previous studies (Frank and Köster, 1979). Therefore, more reliable precision will be estimated in case were such DNA sequence variation exists.

6.3.3. Specificity and Sensitivity

High specificity was observed in this experiment as an amplification of 1 ng of male DNA in presence of 2000 ng of female DNA successfully generated a full RM-Yplex profile. Sensitivity on the other hand, showed less efficiency at 62.5 pg for two males DNA with a drop-out at DYF403S1b, the longest allele available in the RM-Yplex assay. This could be due to many reasons, including pipetting fluctuations which do exist across different analysts and the efficiency of the PCR thermal cyclers. Hence the need of such validation is important in every laboratory in order to determine the limitations of analysis methods.

6.3.4. Stutter Analysis

Stutter cut-off limit determination is very important in interpreting the profiles generated using STR based multiplex assays. In male/male mixtures analysis where minor peak might be presented at the same fragment size as the stutter of another allele, detection of peak above the determined stutter cut-off for specific locus will be suspected of true allele.

Minus one and minus two stutters pattern in trinucleotide loci, such as DYS612, should be also considered in interpreting generated profiles and both should be determined. However, in this study minus two stutter was rarely observed and hence there was not enough observations available to produce a reliable stutter cut-off. Stutter ratios for both tetranucleotide loci and trinucleotide loci were reasonably within expected range of ratios determined previously in developmental validation chapter. Data generated in this study was used to establish specific interpretation guidelines for RM-Yplex analysis and also was used to prepare the stutter ratio file for GeneMapper® *ID-X* software.

6.3.5. Male/Male Mixture

Although all alleles for major and minor contributors were detected in both sets of male/male mixture with 1:3 and 3:1 ratios prepared for the purpose of this experiment, resolving full profiles into minor and major contributors was not achievable with 1:19, 1:9, 9:1, 19:1 and 1:1 ratios of mixtures. All alleles were detected also at 1:1 ratio for both sets of samples. However due to almost equal amplification at each locus, it is hard to resolve minor and major contributors especially when the samples genotypes are unknown. Whereas at such ratio with a reference sample of at least one contributor is available, the analysis would be easier by comparing mixture profiles to reference samples directly. However, the use of such approach is a controversial issue since some will argue that this approach will show some bias by analysts. Even though 1:3 and 3:1 ratios were resolved, interpreting multi-allelic loci will not be as easy as it is demonstrated here. This is so because of the fact that the number of the excepted alleles cannot be determined, rather expected, in unknown samples and also the fact that it has been found that allele number might differ from 1 to 6 alleles at DYF399S1 and from 1 to 4 alleles at DYF403S1a, DYF404S1 and DYF387S1 as it has been shown recently in the analysis of worldwide populations samples (Ballantyne et al., 2014). Not mentioning the fact that in

the case of male/male mixture, the ratios will be unknown and hence this will increase the complication toward correct interpretation of RM Y-STRs.

6.3.6. Non-probative Casework Samples

One of the most critical evidential material recovered from cases of sexual assault is the traces of semen. Currently, autosomal STR multiplex assays including the gender-profiling marker, amelogenin, represent the most popular approach to DNA profiling in forensic investigations. However, there were numbers of situations where autosomal STR multiplex assays fail to amplify the DNA profile of the male DNA fraction, represented by the semen, even though the Y chromosome peak at amelogenin indicates the presence of male DNA in the samples. Not mentioning the fact that these situation including some mixture samples containing semen from azospermic individuals (Prinz and Sansone, 2001, Sibille et al., 2002, Shewale et al., 2003). In addition, another critical situation includes a failure to amplify amelogenin from the Y chromosome of the male DNA which will result in a male STR profile being mistaken as that of female (Steinlechner et al., 2002, Thangaraj et al., 2002). Accordingly, the number of Y-STR analyses has been increased in order to evaluate the detection of such cases. Y-STR DNA profiling analysis has become a powerful assay for analysing vaginal swabs, from sexual assault cases, which include mixture of female and male fractions (Prinz et al., 1997, Gusmaúo et al., 1999, Honda et al., 1999, Prinz and Sansone, 2001).

Since confidence had been established, during validation experiments, in the performance of RM-Yplex assay, the potential of RM-Yplex in producing reliable evidence in actual forensic casework samples has been investigated. RM-Yplex was tested on 7 different types of biological stains deposited on different substrates usually encountered in forensic DNA cases previously reported to DFSC at Dubai police. The advantages of using Y chromosome STR loci in sexual assault cases, as illustrated previously with the commercially available Y-STR multiplex assays, has been demonstrated for RM-Yplex

in this study. Detection of small amount of male DNA has been proven as well as in presence of high amount of female DNA.

The potential advantages of the RM Y-STR loci were also shown in analysing sexual assault case (No.2) where the perpetrators were close relatives and the autosomal STR amplification kit failed to produce evidential profiles even after differential extraction was performed. Although Yfiler® amplification kit was successful in providing a full single profile of the perpetrator; it could not prove that there were two perpetrators according to information provided by victim. The results of RM-Yplex has introduced the possibility of that two closely related males contributed to the semen found in vaginal swab sample.

In another case (No.14) where a tooth, which was found during searching operations for a missing person, was analysed and compared with a reference sample provided from the brother of the missing person in order to produce evidence of brotherhood, a DNA extract from missing person tooth was amplified successfully using Identifilier®, Yfiler® and RM-Yplex. All assays produced full profiles comparable to the reference sample provided by missing person's brother.

While analysis of both cases (No.2 and No.14) using RM-Yplex produced invaluable evidence, using RM-Yplex in investigating close male relatives might not always be evidential as a mutation in meiosis transfer from father to son is highly expected according to the mutation rate of these loci and information about close relationship will be doubtful then. Therefore, in cases where RM-Yplex would be used for such purposes, mutation rates of these loci should be taken into consideration. On the other hand, due to the fact that a mutation will not always occur between father and son hence there will still be male individuals sharing the same RM-Yplex profiles and therefore, such limitation should also be taken into consideration when using RM-Yplex in casework samples.

6.3.7. Families Pedigree Analysis

In order to support the detection of mutation in close relative male individuals, two families pedigrees were analysed and the mutation of the RM Y-STR loci were recorded. It was evidential that such tool can differentiate close male relatives as almost 60% of male were individualised in both families. Detection of such high mutations support the findings in the forensic casework analysis of case number 2 (Section 6.3.6). Such ability of the RM Y-STR in differentiating close male relatives was demonstrated previously on a larger scale study where 2528 pairs of paternal relatives connected by one to 20 generations were analysed (Ballantyne et al. 2014). This analysis was conducted using both RM Y-STR and Yfiler® panels. The results showed 29% of these pairs were differentiated by RM Y-STR loci compared to 5.5% differentiated by Yfiler® loci (Figure 6.12). In the same study, it was also evidential as the number of generations, i.e. meiosis, increasing between relative pairs the chance of having different RM Y-STR haplotypes increased (Ballantyne et al. 2014). However, for Yfiler® the correlation between the number of meiosis and the differentiation between haplotypes was not strong as relative pairs separated by 20 meiosis were never differentiated (Figure 6.12). It is therefore, more likely for such mutation to be detected between cousins than between brothers, as the number of meiosis connecting cousins is higher. The high number of mutations detected for DYF403S1b could be explained because of the allele size, which is more than 400 bp with repeat numbers ranging from 40 to 59, and therefore it is highly prone to mutations compared to alleles with fewer repeats number (Ballantyne et al. 2012).

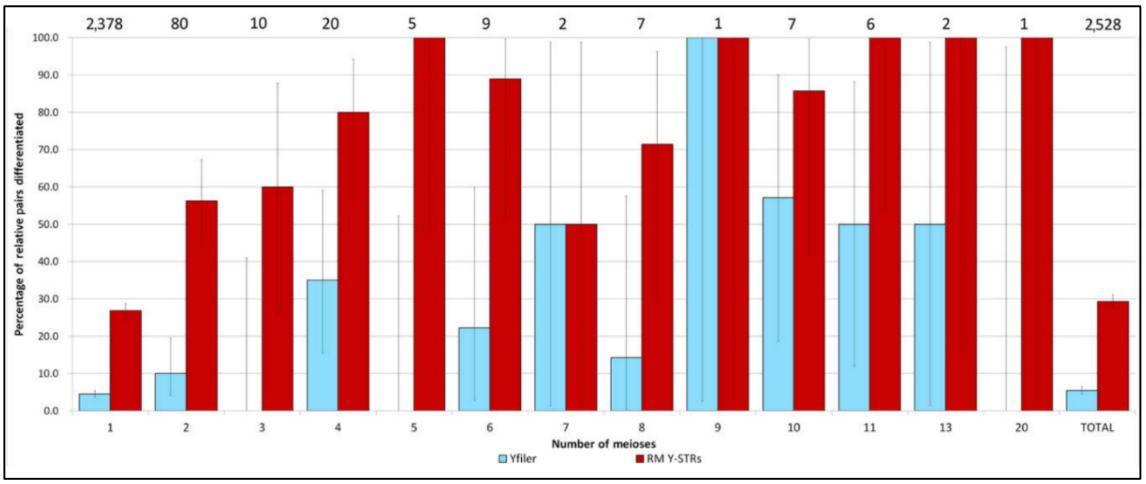

Figure 6.12: Parallel male relative differentiation using 2528 paternal relative pairs separated by one to 20 meiosis for Yfiler® and RM Y-STR adapted from Ballantyne et al. (2014).

To sum up, RM-Yplex multiplex assay has proven the potential efficiency in analysing casework, it should be noted that such analysed markers are not intended to supplant standard autosomal STR markers nor Y-STR markers currently used, however, they should be applied in cases where autosomal and Y-STRs fail to produce a satisfactory result.

7. Population Genetic Studies on Populations Residing in the Arabian Peninsula Using RM Y-STRs Loci

7.1. Overview

Evaluating the ability of the STR markers in differentiating individuals in specific populations is important in the practice of forensic DNA profiling. While autosomal STR loci can help differentiate any two individuals, excluding identical twins, with high statistical probability, Y-STR profiling is known to be less helpful in differentiating between unrelated males individuals. It has been known, that the primary shortcoming of Y-STR profiling in forensic applications is the lack of independence of Y chromosome markers. However, RM Y-STR loci have attracted the forensic DNA profiling community due to their higher power of discrimination mainly attributable to the higher mutation rates exhibited by these loci. Although these markers were investigated in worldwide male DNA samples by HGDP-CEPH as well as in a large-scale study of 111 worldwide male populations (Ballantyne et al., 2014), they were not investigated in Arabian Peninsula populations. The aim of this chapter was to develop a database of these markers comprising of an appropriate number of haplotypes and then comparing the results with available published data to investigate the RM Y-STR loci for their diversity in populations from different countries and ethnicities (Ballantyne et al., 2014). The resulting allele frequency data was also aimed to perform a population study for some of the populations residing in the Arabian Peninsula.

It is important to study genetic differentiation for forensic and genealogical purposes. If significant genetic differentiation were detected within the populations residing within a geographical area then population specific databases would be required for a set of markers. On the other hand, if no structure was detected within a population then a single database would probably be justified.

Additionally it is important to convey the weight of evidence when using a specific set of genetic markers. Haplotype diversity is a standard measure for estimating the power of

discrimination of a Y STR haplotype. This was calculated for different populations as was discrimination capacity (Chapter 2, Section 2.2.12.2.2).

Population substructure was investigated in this study using a component of F-statistics called F_{ST} which usually referred to as fixation index. In 1940s and 1950s Sewall Wright and Gustave Malécot introduced F_{ST} as a tool to illustrate the genetic diversity within and among populations (Malécot, 1948, Wright, 1949). Since then, F_{ST} value has been used as a useful measure of population's differentiation and structuring which usually results due to migration, mutation and genetic drift (Holsinger and Weir, 2009). F_{ST} is defined as a correlation of randomly chosen alleles within a subpopulation relative to the entire population (Holsinger and Weir, 2009). F_{ST} is directly related to the diversity in allele frequencies among populations and conversely to the degree of genetic resemblance among individuals within populations. F_{ST} value ranges from 0 to 1 depending upon the similarities or differences between allele frequencies of two populations, however due to the algorithm incorporated within genetic analysis software, estimated F_{ST} value could therefore be negative. This is usually happens when there are more homozygous genotypes than heterozygous genotypes in analysed population samples. Since all Y-STR considered at homozygous status it is possible that negative F_{ST} value results when comparing homogenous populations. Such negative value does not have any biological meanings however, it does give a hint of the gametic phase of DNA sequences. Therefore in this project negative F_{ST} value would be interpretated as zero (Balloux and Lugon-Moulin, 2002).

In addition to the F_{ST} estimation, two estimates are calculated in order to determine the level of variance between different populations groups these are F_{SC} and F_{CT}. F_{SC} is an estimate of the genetic variation among subpopulations within groups. On the other hand, F_{CT} is an estimation of the genetic variation among group relative to the total variations (Excoffier and Lischer, 2010).

Analysis of Molecular Variance (AMOVA) allows estimation of the partitioning of genetic variability within and among the populations. Arlequin software version 3.5, was used for genotyping data analysis, in order to estimate the within and between population genetic diversities. Pairwise F_{ST} and AMOVA were performed for Arabs sub-populations within UAE, and some Arabian Peninsula populations in comparison to 111 populations data (Ballantyne et al., 2014).

7.2. Results

7.2.1. UAE Population

7.2.1.1. Arab Sub-Populations

Wright's pairwise F_{ST} comparison of three sub-populations of UAE (Arabs, Bedouin, Urban and Rural), resulted in no significant genetic difference. Wright's F_{ST} values were in negative for all pairs with a very high p-value equal to 1.0. AMOVA analysis resulted in 100% of variation accounted for within populations compared to 0% variation among populations (Table 7.1).

Table 7.1: AMOVA analysis of 3 sub-populations of UAE using 13 RM Y-STR haplotypic data.

Source of Variation	Degree of freedom	Sum of Squares	Variance components	Percentage of Variation
Among populations	2	4.603	-0.02233 Va	-0.33
Within populations	597	4040.435	6.76790 Vb	100.33
Total	599	4045.038	6.99276	
F_{ST} Global	-0.00331 (P-value= 1.00000+-0.00000)			

In order to further investigate the genetic structure of the three sub-populations within UAE, a median-joining network analysis was performed, using 11 single allelic loci included in the RM-Yplex assay with the help of Network software (Figure 7.1). The results showed one major cluster. Outlier nodes were not detected in this analysis.

However, a number of haplotypes of the Bedouin sub-population clustered together forming a sub-cluster within the major cluster. This represents a close relationship between these haplotypes. Expectedly, five haplotypes were shared among ten individuals within this cluster of Bedouin sub-population. One haplotype at this cluster was shared among three individuals; two from Bedouin sub-population and one from Urban sub-population. All other Bedouin haplotypes were evenly distributed within major cluster where one haplotype was shared between two individuals was plotted away from Bedouin sub-cluster. In the Rural UAE sub-population, no haplotypes were shared within this population or within the urban and Bedouin UAE populations. Haplotypes generated for Urban sub-population were evenly distributed in the major cluster with only one haplotype shared between two individuals within this sub-population. Out of the 18 most distant nodes, 9 were plotted for Rural and 9 were plotted for the Urban sub-population haplotypes. This again indicates similarity of the genetic structure of these sub-populations.

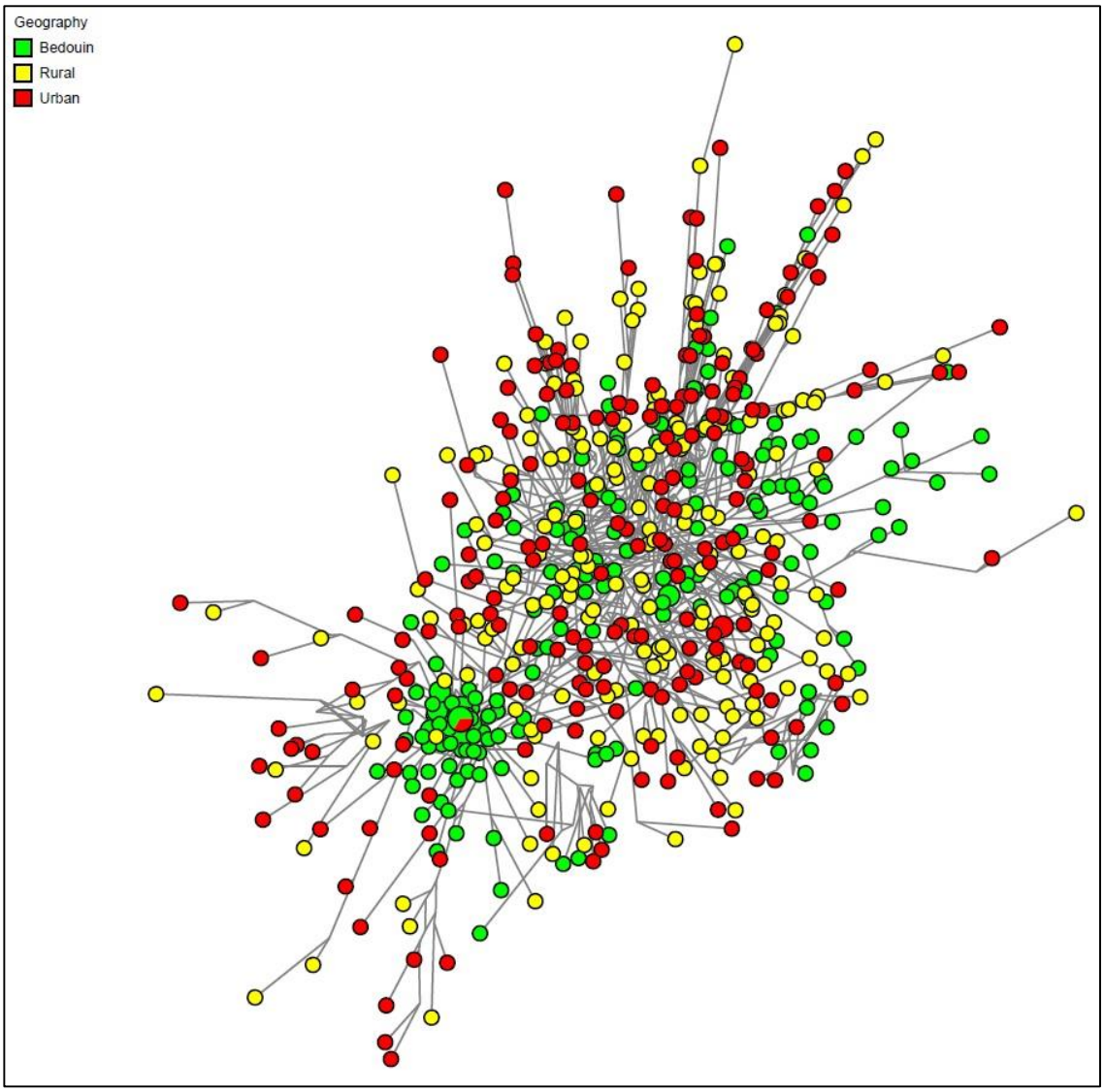

Figure 7.1: Median-joining network of 11 locus Rapidly Mutating Y-STR haplotypes generated using RM-Yplex assay for three UAE Arab sub-populations. The size of the node is proportional to the number of shared haplotypes presented at each node. Each colour signifies a sub-population.

In a separate analysis, haplotype diversities for 13 RM Y-STR loci were calculated as 0.99985 for both Bedouin and Rural sub-populations of UAE, whereas a haplotype diversity of 1 was calculated for Urban UAE sub-population. Haplotype sharing was only detected within Bedouin and Rural UAE sub-population. There was no shared haplotypes among populations. Haplotype diversity for all 600 samples from UAE, was 0.99997.

Considering the absence of genetic sub-structure indicated by these studies, the UAE sub-population data was combined together and analysed as one population against other populations in subsequent analyses.

Table 7.2: Haplotype diversity and discrimination capacity within three UAE sub-populations.

Forensic Parameter	**Bedouin Arabs**	**Urban Arabs**	**Rural Arabs**
Number of Samples	200	200	200
Number of Haplotypes	199	200	199
Haplotype Diversity	0.99985	1.00000	0.99985
Discrimination Capacity (%)	99.5%	100%	99.5%

7.2.1.2. Arab and South Asian Populations

AMOVA was conducted for regional groups including UAE Arab populations and two South Asian populations residing in UAE. The analysis showed that a significant genetic differentiation of 2.65% exists among geographic regions populations (F_{CT}). However, no genetic differentiation was detected among populations within same region (F_{SC}). Vast majority of genetic differentiation was found within populations (97.54%) (F_{ST}) (Table 7.3).

Table 7.3: AMOVA analysis between Arabs and South Asian populations residing in UAE data for 1087 samples using 13 RM Y-STRs haplotypic data.

Source of Variation	Degree of Freedom	Sum of Squares	Variance Components	Percentage of Variation
Among groups	1	103.768	0.18636 Va	2.65
Among populations within groups	3	11.788	0.01369 Vb	-0.19
Within populations	1080	7409.675	6.86081 Vc	97.54
Total	1084	7252.231	7.03349	
F_{ST}	0.02455 (P-value= 0.00000+-0.00000)			
F_{SC}	-0.00200 (P-value= 1.00000+-0.00000)			
F_{CT}	0.002650 (P-value= 0.0.010557+-0.00920)			

7.2.2. Arabian Peninsula Populations

7.2.2.1. Population Sub-Structure

Six populations, including UAE, located in the Arabian Peninsula as well as two South Asian populations, Indian and Pakistani, residing UAE were analysed using AMOVA approach on Arlequin software version 3.5. AMOVA Analysis resulted in a significant variation of 2.25% was found among regional populations (F_{CT}). In addition, there was a significant, but very low, genetic differentiation of 0.75% found among populations within same region (F_{SC}). Finally, the vast majority of differentiation was found to be significant within populations (97.01%) (F_{ST}) (Table 7.4).

Table 7.4: AMOVA analysis of 8 populations residing within Arabian Peninsula for 13 RM Y-STRs.

Source of Variation	Degree of Freedom	Sum of Squares	Variance Components	Percentage of Variation
Among groups	1	176.087	0.15793 Va	2.25
Among populations within groups	6	60.962	0.05248 Vb	0.75
Within populations	1152	7857.704	6.82092 Vc	97.01
Total	1159	8033.791	7.03134	
F_{ST}	0.02993 (P-value= 0.00000+-0.00000)			
F_{SC}	0.02993 (P-value= 0.00000+-0.00000)			
F_{CT}	0.02246 (P-value= 0.00163+-0.00000)			

In the population comparison analysis, values of Wright's pairwise F_{ST} analysis between populations were calculated (Figure 7.2; Table 7.5). The highest pairwise F_{ST} value was 0.08740 between Indian and Yemeni populations whereas the lowest Wright's pairwise F_{ST} value was -0.03482 between Kuwaiti and Qatari populations. In addition, Slatkin's pairwise linearized F_{ST} values were also estimated in the same analysis which indicated same pattern of results between populations.

Table 7.5: Pairwise F_{ST} values calculated for 13 RM Y-STRs frequency data computed using Arlequin v3.5. Bold numbers are significant (p-value less than 0.05).

	UAE N=600	Qatar N=34	Bahrain N=12	KSA N=13	Kuwait N=6	Yemen N=10	Indian N=242	Pakistani N=243
UAE	**0**							
Qatar	**0.02210**	**0**						
Bahrain	**0.01674**	0.00164	**0**					
KSA	**0.01814**	0.00144	-0.01722	**0**				
Kuwait	0.01446	-0.03482	-0.0129	-0.01151	**0**			
Yemen	**0.03798**	**0.00818**	**0.029**	0.00238	-0.00376	**0**		
Indian	**0.02796**	**0.06276**	**0.03544**	**0.04905**	**0.05859**	**0.0874**	**0**	
Pakistani	**0.02381**	**0.05504**	**0.03139**	**0.04484**	**0.04804**	**0.07838**	**0.00012**	**0**

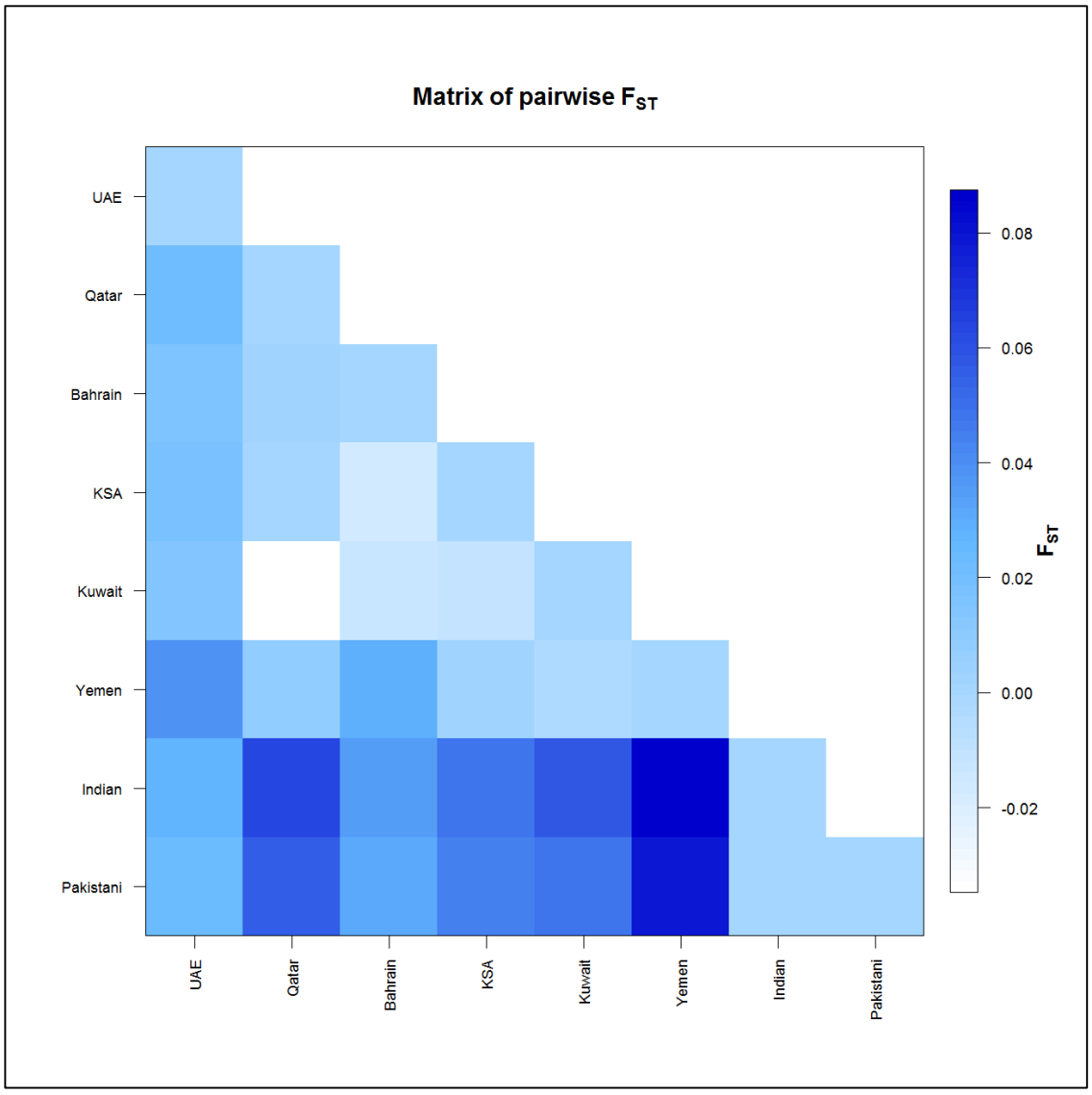

Figure 7.2: Pairwise F_{ST} values matrix describing the differences between population samples. The differences in F_{ST} expressed in blue colour intensity as indicated in the key map of the figure.

In order to represent the magnitude of haplotypes differences between populations and geographical areas, multidimensional scaling (MDS) analysis was performed using Slatkin's linearized F_{ST} values by SPSS software version 21 (Figure 7.3). MDS results demonstrated three clusters despite populations appearing to be loose within each cluster. UAE population was located away from rest of the populations. The remaining populations in the Arabian Peninsula were formed a loose cluster distantly located from UAE and South Asian populations, whereas two South Asian populations were clustered at significant distance from other clusters. This result is largely in agreement with the

history of various populations that were included in the analysis (Table 7.5 and Figure 7.2). The number of the samples from non UAE Arab populations were small and better results might be obtained when further samples are added to the analysis.

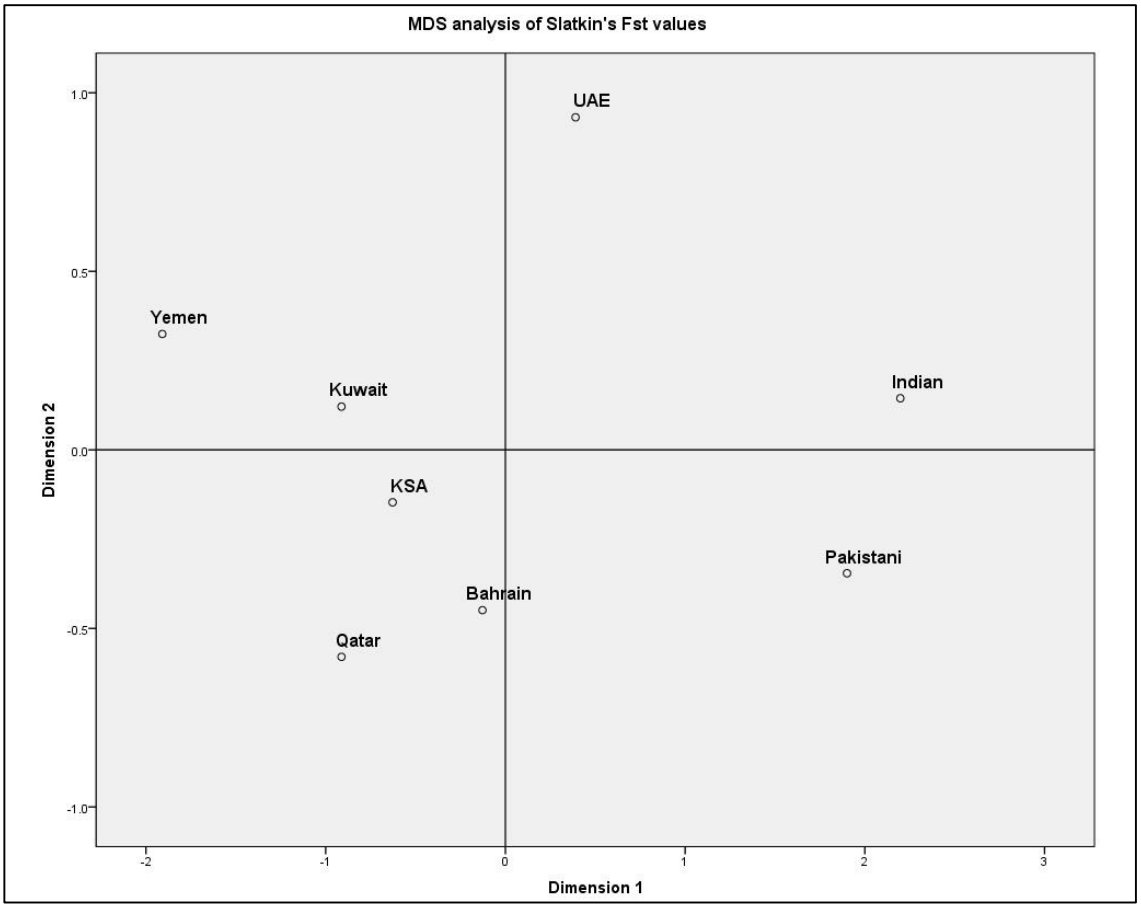

Figure 7.3: Two-dimensional plot of multidimensional scaling (MDS) analysis of Slatkin's linearized F_{ST} value for RM Y-STR haplotypes in Arabian Peninsula samples of 1160 male individuals from 8 populations (Stress 0.05176).

Using only 11 RM Y-STR single-allelic markers, counting DYS526a, DYS526b and DYF403S1b each as one marker, the set of 1160 male samples analysed resulted in a total of 1129 unique haplotypes, i.e. 97.3% of total samples were resolved. The median-joining network was constructed using Network version 4.6.1.1 software (Figure 7.4). The network showed one main diamond shaped cluster with few haplotypes emerging out of the cluster. Two distinct haplotypes were located far away from the main cluster connected through long strings representing very long time distance involving multiple mutations. Further investigation of these haplotypes, indicated that these were the only

two haplotypes, from South Asian populations, which had a drop out at DYF403S1b marker in the whole set of samples. On the other hand, haplotypes which were at short-distance and emerging from the main clusters were the haplotypes having microvariant alleles at least at one locus. According to geographic region of sampling, drawn network showed no specific cluster for regional haplotypes.

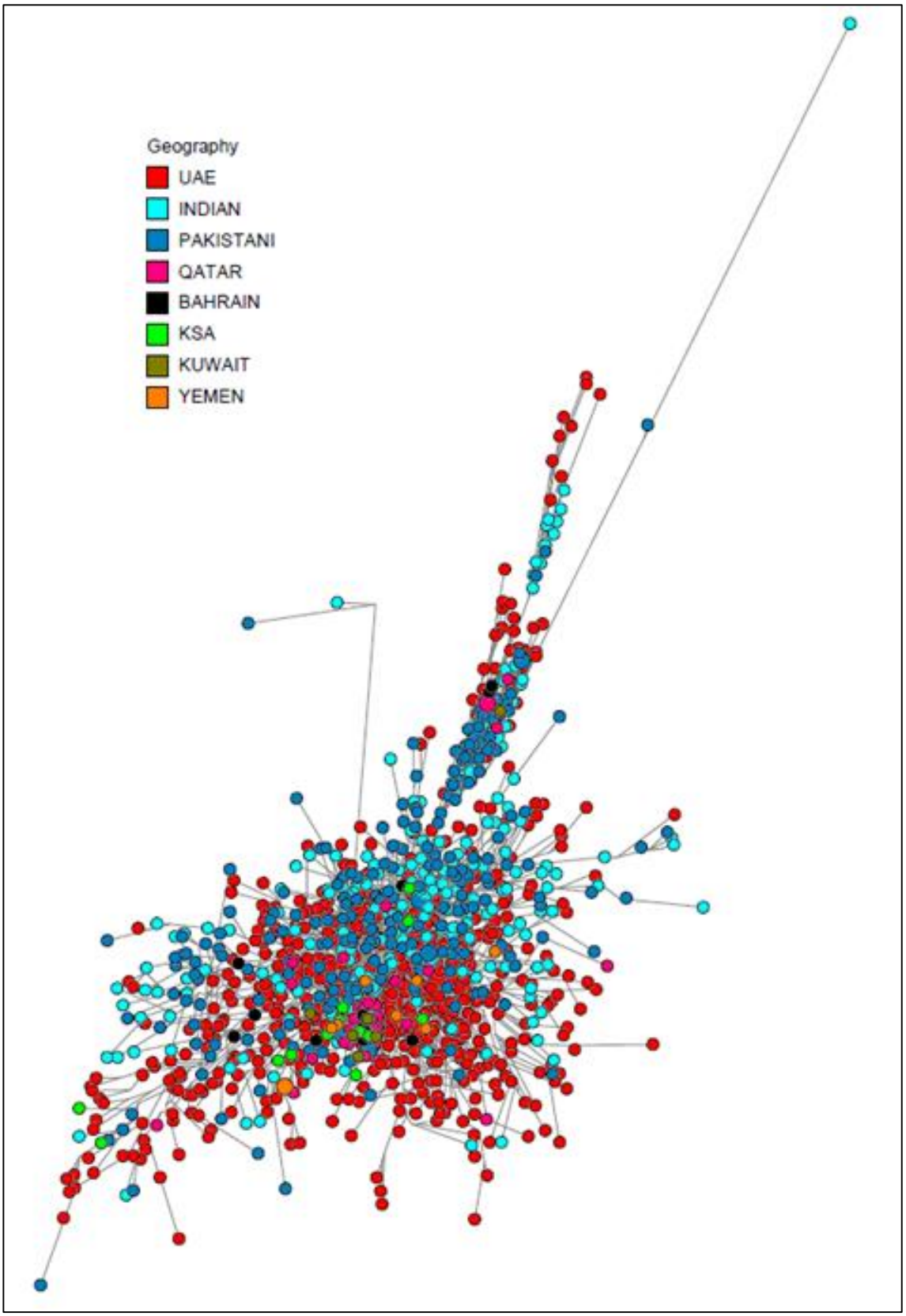

Figure 7.4: A network of 11 single-allelic RM Y-STR markers using data from various Arab, Indian and Pakistani populations. The node represents the haplotype found in the sample with the size of the node representing the number of shared haplotypes.

7.2.2.2. Haplotype Resolution

Haplotype diversity and discrimination capacity were calculated for each population using Microsoft® Excel software (Table 7.6). No shared haplotype was found among different populations, though two shared haplotypes between four individuals were detected in UAE population. In addition, the analysis of haplotypes consisting of single-allelic RM Y-STR markers also showed no shared haplotypes among populations. Allele and haplotypes frequencies in this study were generated using simple counting method for UAE Arabs, Indian and Pakistani populations, using Arlequin software v3.5 (Appendix 3, Figures A3.1-A3.15). All allele frequencies at single-allelic loci were approximately normally distributed with an exception to DYS449, which showed that the distribution is disrupted at the low repeat numbers alleles. Allele 26 was present at a significantly high frequency in UAE population when compared to Indian and Pakistani populations (Figure 7.5). Multi-allelic markers did not result in normal distribution shaped bar graph; however, the haplotype frequencies were almost evenly spread over the entire range. Alleles with microvariants were rarely observed at single-allelic loci whereas multi-allelic locus DYF399S1 scored the highest number of microvariants including 0.1, 0.2 and 0.3 bp variants almost at each allele. Loci DYF404S1 and DYF387S1 also showed a considerable number of microvariants alleles.

Population specific alleles were detected at three loci namely, DYF403S1b, DYS518 and DYS547. Allele 56 at DYF403S1b was only seen at in Qatar population whereas allele 57 at the same locus was only detected in Bahrain population. At DYS518 locus, allele 47 was only detected in Bahraini population. Allele 53 at DYS547 was only detected in Yemeni population samples.

Table 7.6: Haplotype diversities and discrimination capacities in Arabian Peninsula Arabs populations and South Asian populations residing in UAE.

	Individuals	Haplotypes	Haplotype Diversity	Discrimination Capacity (%)
UAE	600	598	0.99997	99.67
Qatar	34	33	0.99997	99.67
Bahrain	12	12	1	100
KSA	13	13	1	100
Kuwait	6	6	1	100
Yemen	10	8	0.97222	87.50
Indian	243	243	1	100
Pakistani	242	240	0.99993	99.17
Global	1160	1153	0.99999	99.40

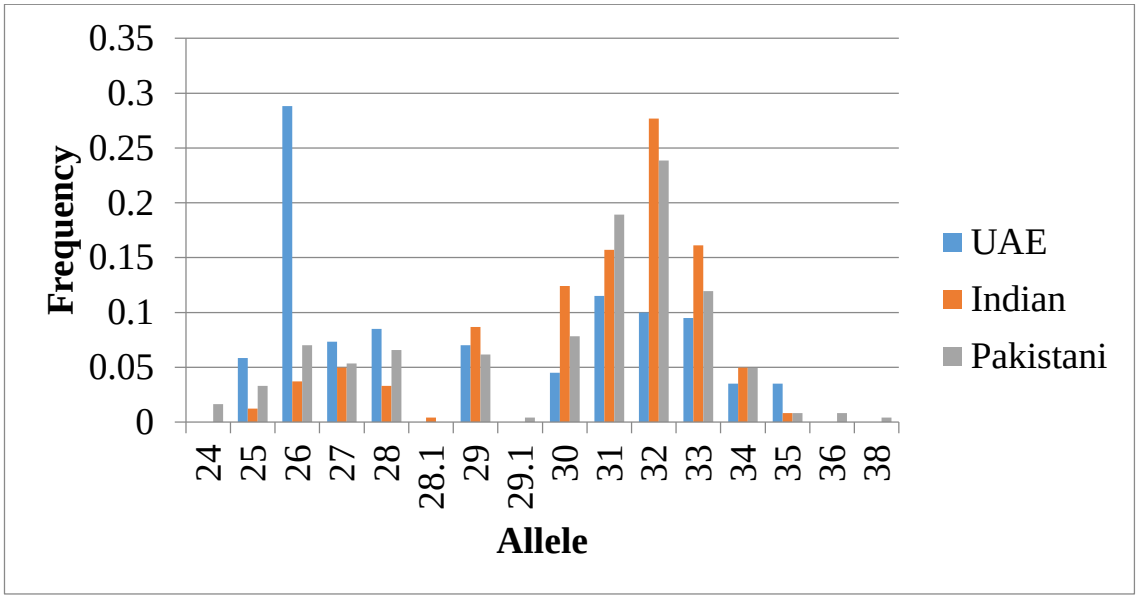

Figure 7.5: Alleles frequencies at DYS449 in three populations, UAE Arabs, Indian and Pakistani.

The average locus diversity was each locus included in the RM-Yplex were calculated (Table 7.7). Highest average diversity of 0.8721 was observed in the Pakistani population, whereas the lowest average diversity of 0.7259 was observed in Yemeni population. Locus DYF399S1 showed the highest diversity whereas the lowest calculated diversity was at locus DYS576. The multi-allelic markers DYF399S1 and DYF403S1a, each of which has three alleles, showed the high diversity values among all loci followed by the

multi-allelic markers DYF387S1 and DYF404S1, which each has two alleles, and the least diversity values were detected in single-allelic markers. In comparison to DYF387S1 and DYF404S1, the two alleles resulted from the multi-allelic marker DYS526ab were treated separately each as locus as the region of each allele is already defined and hence alleles could be assigned confidently.

Table 7.7: 13 RM Y-STR loci diversities across 8 populations.

Locus	Global	UAE	Qatar	Bahrain	KSA	Kuwait	Yemen	Indian	Pakistani
DYF387S1	0.9341	0.9106	0.8503	0.8788	0.8590	0.7333	0.7500	0.9416	0.9427
DYF399S1	0.9953	0.9826	0.9697	0.9848	0.9872	0.8000	0.9722	0.9964	0.9969
DYF403S1a	0.9903	0.9782	0.9715	0.9697	0.9615	0.9333	0.9167	0.9882	0.9885
DYF403S1b	0.8877	0.8897	0.8556	0.9394	0.8718	0.8667	0.8333	0.8502	0.8521
DYF404S1	0.9191	0.8154	0.7558	0.9394	0.8846	0.6000	0.5833	0.8963	0.8992
DYS449	0.8854	0.8607	0.8610	0.8788	0.7308	0.8667	0.8611	0.8457	0.8695
DYS518	0.8431	0.8293	0.8093	0.7879	0.8333	0.8000	0.8611	0.8508	0.8657
DYS526a	0.8210	0.8176	0.7291	0.7424	0.7821	0.8000	0.4167	0.7278	0.7751
DYS526b	0.8808	0.8711	0.8574	0.9091	0.8462	0.8667	0.5556	0.8750	0.8819
DYS547	0.8164	0.7766	0.7843	0.7879	0.6410	0.8667	0.8056	0.8499	0.8478
DYS570	0.7802	0.7291	0.6542	0.6515	0.7564	0.6000	0.6389	0.8155	0.8391
DYS576	0.7571	0.7232	0.7451	0.7727	0.6154	0.7333	0.7778	0.7842	0.7981
DYS612	0.8124	0.8061	0.8111	0.7879	0.7436	0.8000	0.7778	0.8295	0.8176
DYS626	0.8375	0.8084	0.6007	0.8939	0.7179	0.7333	0.3889	0.8492	0.8536
DYS627	0.8488	0.8209	0.7968	0.7424	0.7821	0.8000	0.7500	0.8615	0.8542
Average	**0.8828**	**0.8413**	**0.8034**	**0.8444**	**0.8009**	**0.7867**	**0.7259**	**0.8641**	**0.8721**
SD	**0.0677**	**0.0765**	**0.1007**	**0.0983**	**0.1058**	**0.0950**	**0.1741**	**0.0706**	**0.0630**

7.2.2.3. Unusual Alleles Observed During This Project

There were a significant number of samples which showed more alleles than expected at different loci. All of the additional alleles detected were amplified in singleplex PCR in order to confirm all such alleles. 58 samples exhibited an extra allele at one, two or three loci (Table 7.8). In general, multi-allelic loci showed higher number of duplication events

than the single-allelic loci. DYF399S1 usually has 3 alleles. In a particular sample 5 or possibly 6 alleles were amplified which was an unusual event (Figure 7.6). On the other hand, some samples exhibited fewer alleles than expected at specific loci as well as allelic drop-outs. All fewer alleles were observed at multi-allelic loci only with exception of single-allele locus DYS612 at which allele drop-out was detected in 24 (2.1%) samples. All drop-outs at DYS612 were detected in South Asian populations.

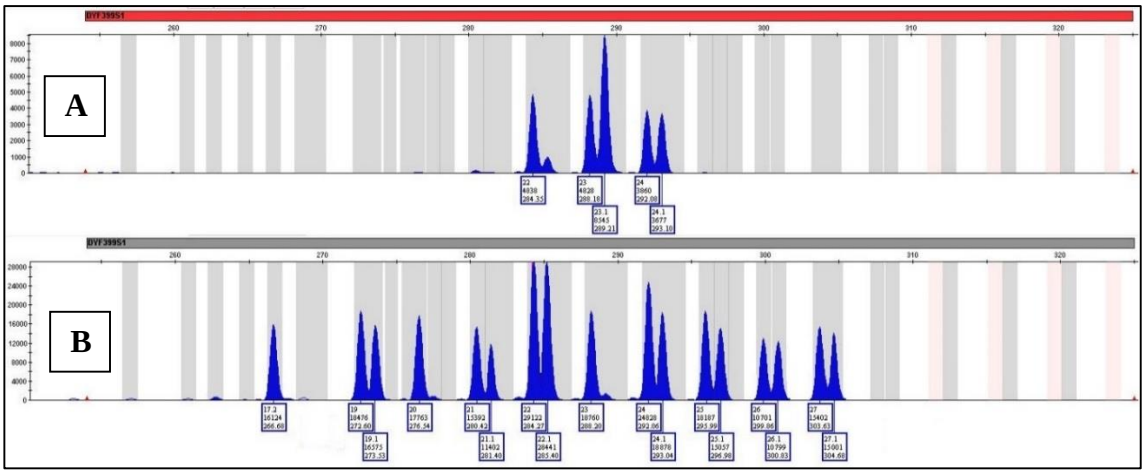

Figure 7.6: An example of duplication event at DYF399S1 locus showing two extra alleles than typical number of alleles expected. A: male individual sample (Sample B063), B: Allelic ladder for DYF399S1.

Table 7.8: Additional alleles and missing alleles at 13 RM Y-STR loci in 1160 male individual samples.

Locus	One Additional Allele		Two Additional Alleles		One missing Allele		Two Missing Alleles	
	Count	Frequency	Count	Frequency	Count	Frequency	Count	Frequency
DYF387S1	9	0.00776	3	0.00261	369	0.31810	-	-
DYF399S1	25	0.02155	3	0.00261	170	0.14655	17	0.014655
DYF403S1a	7	0.00603	-	-	286	0.24655	7	0.00603
DYF403S1b	0	0	-	-	2	0.00172	-	-
DYF404S1	7	0.00603	-	-	308	0.26552	-	-
DYS612	-	-	-	-	24	0.02068	-	-
DYS626	2	0.00172	-	-	-	-	-	-

7.2.3. Worldwide Population Analysis

In order to demonstrate the magnitude of haplotype differences between a 111 worldwide population database published recently (Ballantyne et al., 2014) and the data generated during this project for 8 populations, all data was analysed using Arlequin. The sample number was reduced to a maximum of 100 in each population in order to minimise the difference in sample numbers among populations and also to stay within the limit of characters allowed in the input file for Arlequin software. A percentage of 94.78% variation was seen for within population variation. 2.34% the variation existed among regional groups and 2.88% the variation among population within regional groups. A global F_{ST} value of 0.05219 was resulted with a p-value of 0. The regional groups were as determined in Ballantyne et al., 2014. A multidimensional scaling analysis was performed on Slatkin's linearized F_{ST} values resulted from analysing 119 populations using Arlequin software version 3.5 and SPSS version 21 (Figure 7.5 A and B).

Overall, populations formed a looser cluster with small level of dispersion across both dimensions. UAE, Indian and Pakistani populations were gathered in the centre of the cluster where the majority of the populations were overlapping over each other; however, UAE was located distantly from both South Asian populations, which were at closer distance from each other proportionally. Qatar, Bahrain and KSA populations were gathered at moderate level away from the centre of the clusters. Out of the populations analysed in this project, Yemen populations was the furthest from the centre of the cluster and Kuwait population was equidistant between Yemen and three populations, Qatar, Bahrain and KSA. Overall populations, Lhokpu population within Bhutan geographic region was the furthest located from the centre of the cluster. It was noticed that most populations from the same geographic area (apart from Qatar, Bahrain and KSA) were not clustering together either in this study, nor in the previously published worldwide populations data analysed (Ballantyne et al., 2014).

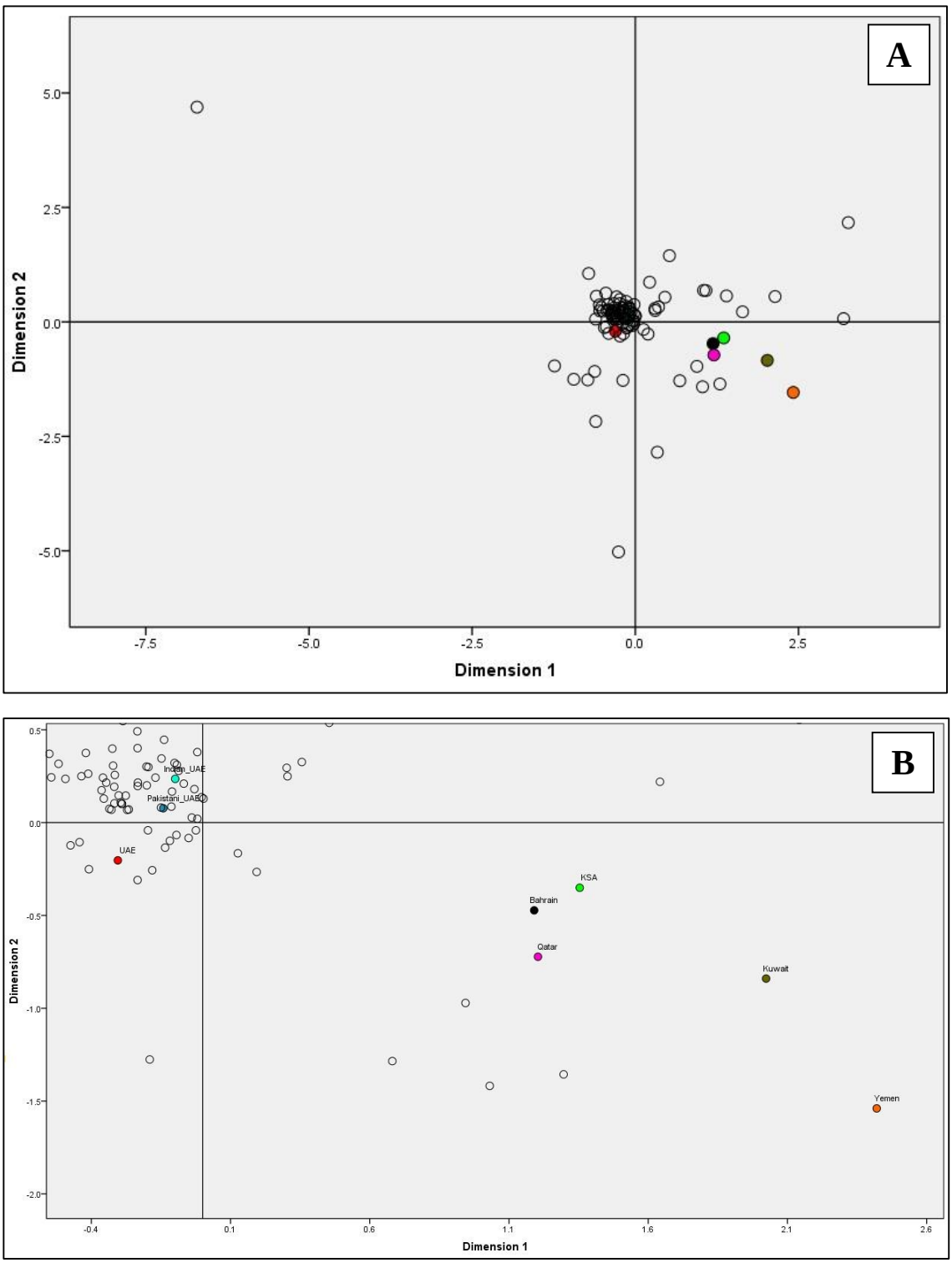

Figure 7.7: Two-dimentional plot of multidimentional scaling (MDS) analysis of Slatkin's linearized F_{ST} values for RM Y-STR haplotypes in male samples from 119 populations (Stress 0.03071). A: full scale MDS, B: Large-Scaled of the cluster centre. Colours codes: Red: UAE population; Light Blue: Indian population; Dark Blue: Pakistani population; Pink: Qatar population; Black: Bahrain population; Green: KSA population; Olive-Green: Kuwait population; Orange: Yemen populations.

7.3. Discussion

Absence of population substructure, as indicated by the F_{ST} values and AMOVA results, between different Arab sub-populations in UAE population was not surprising because of two reasons. First reason is that, although these three Arab's sub populations of UAE are located at different places within UAE , they all share a common recent ancestor (Al-Sayegh, 1998). Second reason is related to the high mutation rate of the used markers in this study which have already been seen to break down population substructure (Victoria, 2010, Ballantyne et al., 2012, Ballantyne et al., 2014). Hence, 600 samples from UAE were considered as one population to overcome the limitation of the analysis software in analysing large number of populations and also simplify the analysis of the larger number of samples.

The 6 populations of Arabs residing on the Arabian Peninsula as well as two South Asian populations residing in UAE showed considerably high genetic variation within the populations. The variation component within the populations was only 0.023%. This result was also confirmed by the analysis of haplotype resolution as a full or almost full haplotype resolution was achieved at each population, with an exception of Yemen population where haplotype resolution was found to be 87.50%. The low discrimination at Yemen is probably due to the smaller number of samples analysed.

Significant genetic distances were detected between all Arabs in the Arabian Peninsula and both South Asian populations using population's pairwise F_{ST} analysis. These results were illustrated by both plots of Wright's F_{ST} values as well as Slatkin's linearized F_{ST} values. This was expected as these two groups were separated geographically as well as the fact that there has been significantly less gene flow between these populations in the past. The significant difference between UAE and other populations in the Arabian Peninsula is most likely due to the low number of samples analysed in each population other than UAE. The resulted furthest genetic distance observed between Yemen and

other populations in this study is consistent with two previous study based on 10 widely used Y-STR loci, which indicted that Yemeni population was at furthest distance from UAE population (Alshamali et al., 2009, Mohammad et al., 2009). In the current study, this was indicated by the allele frequency analysis where population specific alleles were found in Qatar, Bahrain and Yemen populations and hence the frequency of such alleles was over estimated due to the small number population samples analysed. Although there was no evidence of population specific alleles at Kuwait and KSA populations, alleles at 13 loci analysed were not very diverse within each locus at these populations. In Kuwaiti population only three alleles were detected at DYS570 and DYS576 loci. On the other hand, three alleles only were found at locus DYS547 for KSA population. Over all, AMOVA results in three different analyses, UAE sub-populations, Arabian Peninsula and worldwide population, showed a significantly higher within population variation (>95%) compared to within populations variations (Barbujani, 2005, Alshamali et al., 2009, Mohammad et al., 2009).

The limitation in this analysis of calculating only F_{ST} value is that this analysis takes into account the different number of alleles detected between haplotypes at each locus in each population but not the repeats number. It is important in such analysis to take into account the repeat number as the difference in allele number might also give a clue regarding the common ancestor sharing between two haplotypes. R_{ST} value, which could be defined as the similarity between randomly chosen alleles within the same population relative to the entire population, was introduced by Slatkin's (1995). This value takes into account not only allele differences but also the magnitude of repeats numbers, whether two alleles are varying by few repeat number, representing closely related individuals, or many repeats number, representing distantly related individuals. R_{ST} values were not calculated, as overlapping alleles at multi-allelic markers could not be assigned to a specific region on Y chromosome using fragment analysis method.

Using another approach to test the population substructure, a median-joining network was constructed and analysed. Overall, the result of the network analysis showed no clustering of haplotypes using only 11 single-allelic markers. This result reflects the high level of within population variation compared to between population variation as detected by AMOVA results. This is in agreement with the normal and linearized F_{ST} plots, detected in Arabian Peninsula populations with South Asian residing UAE. No clusters resulted reflecting significant regional separation as indicated by the MDS plot. This could be due to the low global F_{ST} value of 0.02457 which indicate a very little genetic variation (Balloux and Lugon-Moulin, 2002). Hence, the genetic structure appeared to be broken down by the rapidly mutating loci as suggested previously (Victoria, 2010, Ballantyne et al., 2012, Ballantyne et al., 2014). Another reason could be due to absence of four multi-allelic loci, which are DYF387S1, DYF399S1, DYF403S1a and DYF404S1, in this network analysis. Multi allelic loci were prevented by the limitation of the software. These loci could have contributed more in separating Arabian Peninsula populations.

RM-Yplex haplotype diversity was found to be very high in all populations. In UAE the diversity of RM-Yplex haplotypes was found to be 0.99997 (600 samples) compared to 0.965 (164 samples) and 0.989 (217 samples) which were estimated previously using Yfiler® and Powerplex Y® assays respectively (Cadenas et al., 2007, Alshamali et al., 2009). RM-Yplex haplotype diversity in Qatar population was found to be 0.99997 which is also higher than previously estimated haplotype diversity of 0.834 and 0.947 using Yfiler® and Powerplex® assays respectively in 72 male individuals analysis (Cadenas et al., 2007). Although 62 male individuals from Yemen population demonstrated quite high haplotype diversity of 0.993 and 0.933 using both Yfiler® and Powerplex Y® assays respectively, the haplotype diversity using RM-Yplex assay was still higher (Cadenas et al., 2007). Even though, comparison of haplotype diversities within UAE, Qatar and

Yemen populations were based on different male samples, the results still show high diversity using RM-Yplex assay in all populations.

The allele frequency distribution in UAE, Indian and Pakistani populations at all loci followed the normal distribution except for alleles at DYS449 locus, which showed a skewed distribution for the three populations. Allele 26 at DYS449 had highest frequency in UAE population compared to allele 32, which had highest frequency in both Indian and Pakistani populations. The skewed distribution in such randomly collected samples was unexpected but this could be explained in two ways. Firstly, this marker has two different stretches of STR repeats and hence allele frequencies do not represent the repeats at each stretch but the total repeats number as a combination of both stretches (Ballantyne et al., 2010). Therefore, such loci would be prone to underestimate the diversity using fragment analysis and also can affect population's substructure analysis. This was confirmed earlier in the allele sequencing analysis and population studies (Schoske, 2003, D'Amato et al., 2010). Secondly, allele 26 & 32 at DYS449 locus could be population specific alleles, which reflects the stability of these alleles in these populations. There is no mutation analysis data of this marker in UAE/South Asian populations to support such a suggestion.

There was a considerable number of new alleles at different Y-chromosome STR loci detected in previous studies. This is most likely results because of the high content of repetitive and palindromic sequences spread over Y-chromosome which facilitate intra-chromosomal recombination and gene conversion events. A previous analysis of 647 male individuals using different Y-STR markers showed 1% of male individuals with extra alleles at different Y-STR markers (Kayser et al., 1997a, Schoske et al., 2004, Butler et al., 2005). However, in this study 4.2% of male individuals showed extra alleles at different RM Y-STR loci which reflecting the high mutation rate of such markers as estimated previously (Ballantyne et al., 2010). Extra alleles at Y-STR markers

traditionally will be interpreted as a mixture in forensic DNA profiling analysis and hence it is not favourable in the forensic applications. This means that the number of samples for a population study using RM Y-STRs needs to be much greater than traditional Y-STRs and autosomal STRs to capture as many alleles as possible. This result also indicates that the profile assessment criteria when using RM Y-STRs need to be further developed. The most discriminating loci, favoured in forensic DNA profiling, as shown in this study were the multi-allelic markers.

On the other hand, 2% of male individual samples analysed in this study showed an allelic dropout at locus DYS612. This could be a result of deletion event or more likely a mutation at primer binding site, which prevents the initiation of amplification process by DNA polymerase in PCR. However, allelic dropout at multi-allelic markers can be explained in two ways. First explanation is that some alleles would be amplified preferentially in PCR reaction and hence alleles with low amplification might not be detected in fragment analysis. Another reason would be the pseudo-homozygous situation where two alleles will have the same fragment size and hence only one signal will be detected in fragment analysis. Even though intra-locus peak highest imbalance study earlier in the optimisation chapter has shown steady amplification at all multi-allelic markers in RM-Yplex assay, with exception to DYF403S1a. However, such a conclusion remains unconfirmed as these alleles were not sequenced.

Haplotypes developed during this study were compared with the RM Y-STR worldwide data using Arlequin software (Ballantyne et al., 2014). 1153 RM Y-STR haplotypes developed during this study were searched against the 12,272 RM Y-STR haplotypes developed in the worldwide sample. No haplotype was shared among the populations, further confirming that haplotypes sharing among populations were almost absent. It was expected, that some of the South Asian population samples studied in this project, would show a higher level of haplotype sharing; however this was not detected. The most

reasonable explanation is that the origin of Indian and Pakistani population samples studied in this group was recorded based on the country only. These samples might represent male individuals from different sub-populations within India and Pakistan. This might have maximised the diversity within these samples so a greater number of unique haplotypes were detected in the South Asian samples.

Results of AMOVA showed reduced genetic variation within populations analysed during this project as compared to RM Y-STR study group. This is might be due to the smaller number of samples used in this study. The results of Slatkin's linearized F_{ST} values were illustrated by creating MDS plot of the populations. The results showed only one cluster for 119 populations plotted in this study with most of the populations located in the centre of the cluster. Although, there were few outliers, including populations analysed in this study, the overall result did not show any patterns comparable to the geographic origin of populations. This result is in agreement with the AMOVA results of the 119 populations performed in this study as well as in agreement with the population structure analysis of UAE sub-populations, Arabian Peninsula Arabs populations and the Indian & Pakistani populations. Hence, estimating the random match or rarity of specific RM Y-STR haplotype would be confidently estimated using the haplotypes database developed during this project.

8. General Discussion, Conclusion and Future Work

8.1. General Discussion

8.1.1. RM-Yplex Development and Performance

The strategy of designing and optimising a multiplex assay led to successful development of the RM-Yplex multiplex assay, which was found suitable for forensics DNA profiling applications as well as population genetic studies. The free access to full human genome sequences on the World Wide Web, through http://www.ncbi.nlm.nih.gov/genbank/, also facilitated the searching for and designing of new primer pairs used in RM-Yplex assay. The ability of RM-Yplex to amplify very small amounts of DNA as well as amplifying purified FTA® discs, stained with saliva or blood, will help the implementation of the assay for routine casework in most forensic and databasing laboratories. The utility of RM-Yplex in male/female mixtures was demonstrated by recovering full male DNA profiles in the presence of excess amounts of female DNA. In addition, the resistance of the assay to high concentrations of inhibitors, commonly encountered in forensic samples, shows its utility in forensic casework. In comparison, such performance characteristics of the RM-Yplex assay are equivalent to the standards of the commercially available Y-STR multiplexes assays such as Yfiler® and Powerpelx Y23® (Mulero et al., 2006, Thompson et al., 2013). However, a male/male mixture study of RM-Yplex indicated that the interpretation of such mixtures is difficult and needs care. This difficulty is compounded by the fact that new alleles are expected at the multi-allelic markers due to high mutation rates. For example, the number of alleles at DYF399S1 locus varies from 1 to 6 alleles as demonstrated during this project as well as in recently published study (Figure 7.6), (Ballantyne et al., 2014). The same situation occurs with multi allelic markers DYF387S1 and DYF404S1 loci. DYF403S1a was excluded from mixture studies during this project as it showed high level of intra-locus allele peak height imbalance in samples as well male controls, which is not clearly understood at the moment. However, 11 single allelic markers available in RM-Yplex and could be effectively used for mixture resolution in

cases where multi-allelic loci were difficult to assess. Nonetheless, the multi-allelic markers were found to be the most diverse markers and contribute to a high haplotype diversity & discrimination capacity.

8.1.2. Potential Haplotypes Resolution

Resolution of male haplotypes in the current study as well as previously published studies, using RM Y-STR loci, proved to be extremely high (Ballantyne et al., 2012, Ballantyne et al., 2014). Such resolution of Y-chromosome STR loci is preferred in forensic DNA profiling. According to the findings of this project, there is a potential to increase male haplotype resolution using the same set of RM Y-STR loci. This can be done by taking into consideration factors like duplication of multi-allelic loci, overlapping allele sizes analysis and improving the detection of loci with more than one repeat stretch.

8.1.2.1. Duplications at Multi-Allelic Loci

Duplication events in the Y chromosome has been shown to be significantly more common than autosomal chromosomes due to its high content of repetitive and palindromic sequences which facilitate such events (Kayser et al., 1997b). In 13,431 male individual samples, a duplication event at least at one locus has been detected in 676 (5%) samples (Ballantyne et al., 2014). More interestingly, 5% of the 676 samples showed at least one extra allele at each of the DYF387S1, DYF399S1 and DYF404S1 loci. On closer investigation of the haplotypes developed during this study, intra-locus peaks imbalances were noted in few samples at multi-allelic loci DYF387S1, DYF399S1 and DYF404S1 consistently. Samples showing such imbalances were re-extracted and amplified to confirm that such imbalance was not a result of preferential PCR amplification nor random errors. Considering the positions of RM Y-STR loci on Y chromosome (Figure 8.2), there is one copy of each multi-allelic markers, DYF387S1, DYF399S1 and DYF404S1 located on the long arm of Y chromosome adjacent to each other. A fair

assumption would be that a duplication event affects these loci simultaneously. Such duplication events are the most logical explanation for the consistent intra-locus imbalance at these three loci in some samples (Figure 8.1). As a final point, the occurrence of extra allele at DYF399S1 only, along with peak heights imbalance is possibly due to a locus mutation occurring, following the duplication at this marker. This hypothesis is also supported by the higher mutation rate of DYF399S1 locus.

Nevertheless, all of these different scenarios of unusual allele numbers or unusual intra-locus peak heights imbalance would be only detected when PCR amplification conditions and electrophoresis analysis conditions were optimised and kept constant. Hence, also duplication of all allele copies at multi-allelic markers, apart from DYF403S1a, detected as peak height imbalance will be more pronounced among loci within corresponding dye panel. Therefore, 5% of samples showing consistent duplication at each of the three multi-allelic loci might be an underestimation of the real number of duplication events at these loci. Confirmation of such duplication events is not achievable using fragment analysis; however, sequencing approach would be more appropriate to confirm these different scenarios. Peak heights tend to show allelic dosage. This is evidenced in and is used to analyse mixtures. Taking into account peak heights in multi-allelic loci will increase the resolution of the generated haplotypes using RM-Yplex assay, especially when less number of alleles detected in the profile as a result of two or more alleles having the same size.

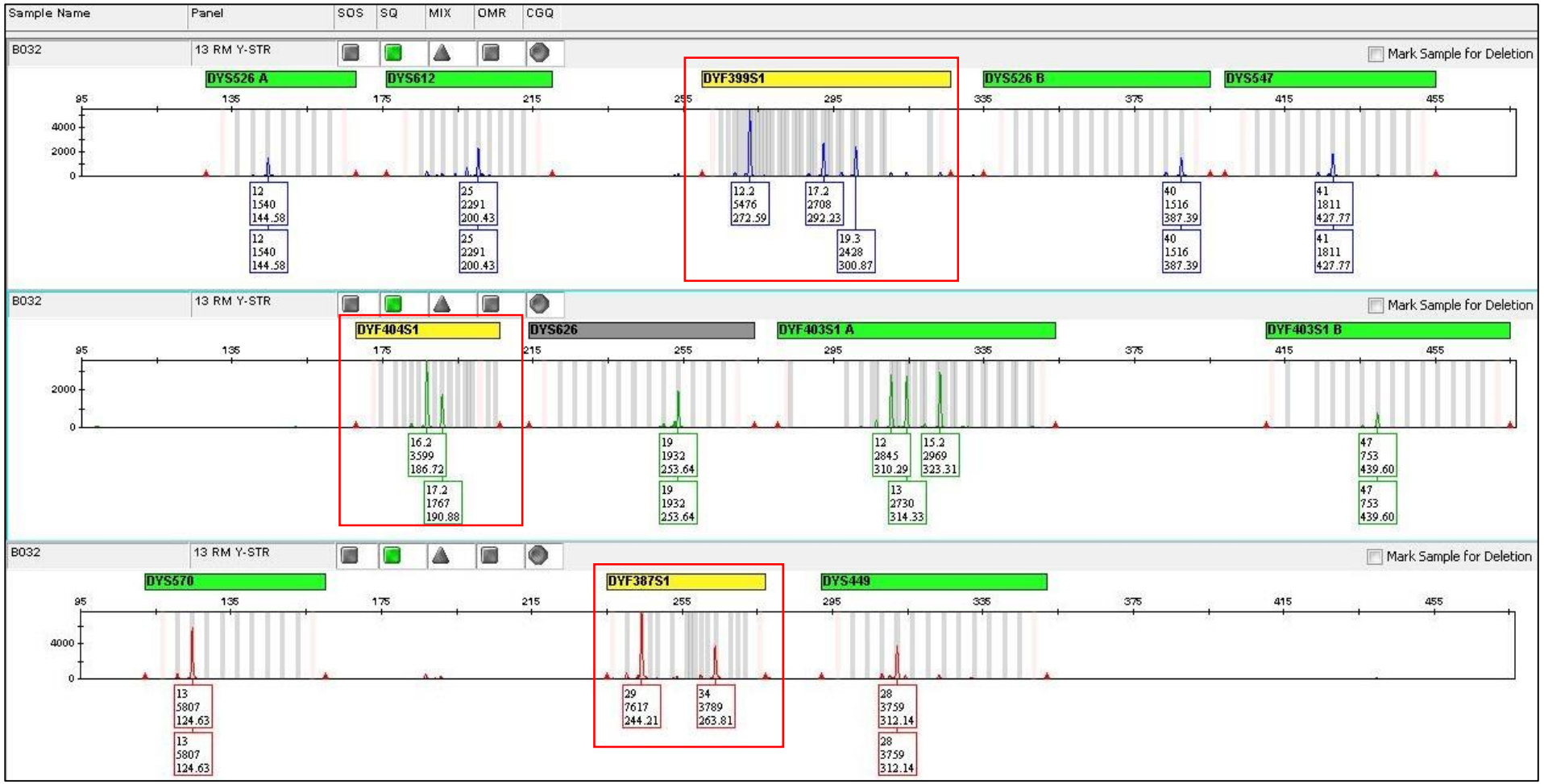

Figure 8.1: Electropherogram illustrating the consistent imbalance events at DYF387S1, DYF399S1 and DYF404S1 in a sample. Red boxes indicating intra-locus peak height imbalance.

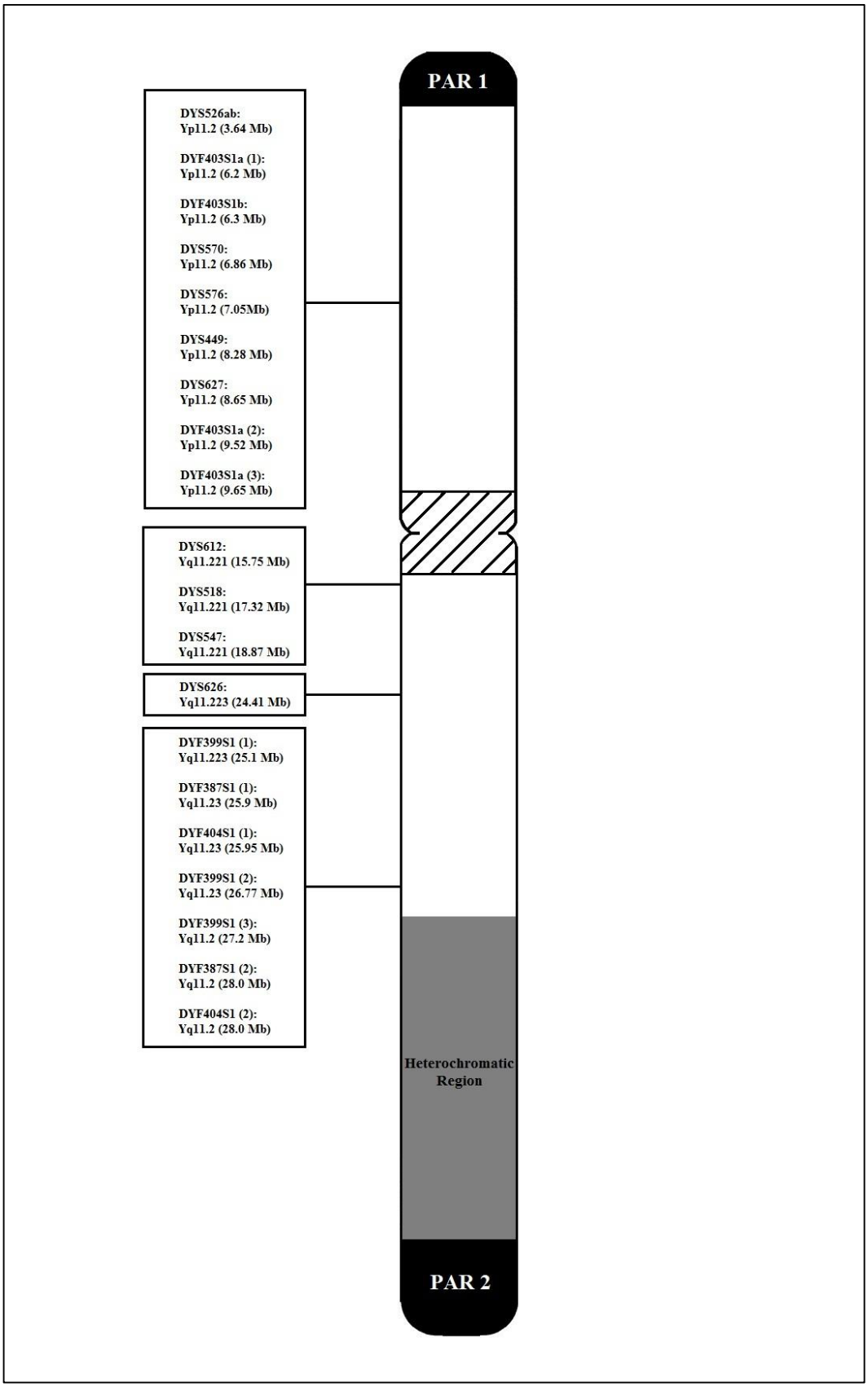

Figure 8.2: Simplified Y chromosome structure illustrating the positions of RM Y-STR loci along with two regions: pseudo-autosomal (PAR1 and PAR2) and heterochromatic region.

8.1.2.2. Issues Relating to the Locus DYF403S1a/b

While POP-6 polymer was used in the development and optimisation of the RM-Ypelx multiplex assay, it was also validated at Forensic DNA Profiling laboratory in DFSC using POP-4 on 3500 Genetic Analyser. The advantages of such applicability of RM-Yplex assay would not require extra modification to the current routine followed in forensic DNA profiling laboratories where POP-4 is widely used as standard polymer (Butler, 2007). However, as published previously (Detwiler et al., 2004), this project also established that POP-6 has higher resolution than POP-4, especially in analysing large sized overlapping STR fragments with different DNA sequence. This was realised during the analysis of DYF403S1a locus. The loss of such valuable information could be compensated by the use of POP-6 which will facilitate detection of same size alleles with different DNA sequence and hence increasing the resolution of the generated profiles. Looking closer at the sequences of the four alleles at DYF403S1ab and positions of such alleles on Y chromosome (Figure 8.2), it was realised that one of the alleles at DYF403S1a and DYF403S1b allele have the identical DNA sequence in flanking region. Whereas the other two alleles at DYF403S1a have identical DNA sequence in flanking region but as slightly different than the other two alleles with a number of SNPs appearing in this sequence which can facilitate the redesigning of two different primer pairs. Each pair will allow the detection of shaded alleles easily using POP-4. Such redesigning of primers pairs was attempted and analysed in 400 samples of male individuals from Korean population and has shown an improvement in haplotype diversity from 0.9971 to 0.9978 (not published). However, the sizes of PCR products produced using redesigned pairs of primers did not support the incorporation of all loci in a single multiplex panel.

8.1.2.3. Multi-STR Stretches Within RM Y-STRs

DNA sequences of RM Y-STR loci available in GenBank® as well as allele sequences obtained in this study confirm the fact that there is more than one STR stretch at DYF403S1a, DYS449, DYS518 and DYS547 loci. Each of the different STR stretches within a locus has many different number of repeats. Hence the different combination of different repeats number at each stretch may result in the same allele calling using fragment analysis method (Figure 8.3) (Ballantyne et al., 2010).

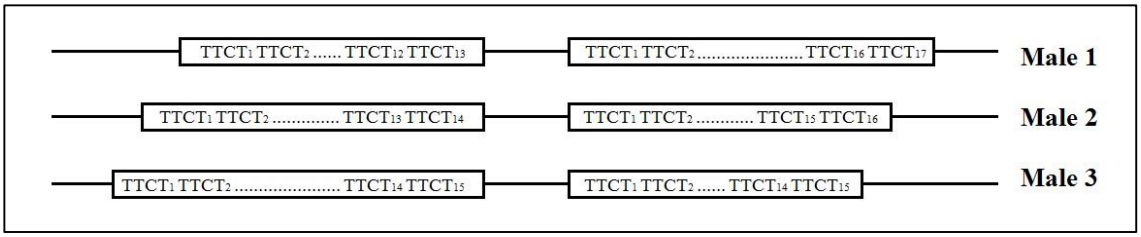

Figure 8.3: Possible combinations of Allele 30 at DYS449 locus in different male individuals.

Therefore, fragment analysis will always underestimate the variability of alleles due to its inability to detect different combinations of repeat numbers at loci with multi STR repeats stretches. Such limitation of fragment analysis method could be overcome by two approaches. First approach would employ designing different primers pairs for each STR stretch. Such approach is limited due to the available conserved region between STR stretches in the loci. At DYF403S1 and DYS518 this approach is impossible as there are less than 20 base pairs between two STR stretches and hence designing specific pairs of primers is not achievable. However, such approach is possible for DYS449 locus whereas at DYS547 locus it is only possible if two out of three stretches available at this locus were split into two fragments. Another approach is sequencing, however, traditional Sanger sequencing is time consuming and needs a lots of effort especially when it comes to large number of samples as well as the need of high amount of DNA to start with

(Sanger et al., 1977). Probably, next generation sequencing would be more appropriate for RM Y-STR analysis as it is able to give as same results as STR fragment analysis in addition to the sequence of the STR repeats (Rockenbauer et al., 2014).

8.1.3. Potential Linkage of Locus DYF399S1 to Azoospermia Factor C (AZFc)

Peak height reflects the number of alleles at a locus. The multi-allelic loci included in RM-Yplex where three alleles might be overlapping in the detected fluorescent signal in confusing fragment analysis, such an observation can help assessing the sample peaks. However, in some situations a single peak at DYF399S1 might show a peak height almost equal to the peak height of a sample where DYF399S1 was showing three different size of alleles. Moreover, that single peak can be in balance with other loci in the same dye channel. This situation would be interpreted as an event of deletion at two allele positions rather than allele size overlap. Deletion events at DYF399S1 locus were previously detected as this marker is located in the Azoospermia Factor C (AZFc) region on Y chromosome (Figure 8.4) (Henson, 2005). AZFc region is a sub-region out of the three regions of AZF gene which is responsible for sperm production in male individuals. AZFc region comprises of a large region of palindromic DNA sequence caused by common recombination events and it has been shown in some cases, that deletion in these genes was responsible for deficient sperm production (Ioulianos et al., 2002, McElreavey et al., 2006). An example of unaffected sperm production in male individuals was seen previously when DNA sequence of 1.8 Mb length, which belongs to Y chromosome haplogroup N (HgN), was deleted in AZFc removing both segments of g2 and g3 (Figure 8.4). This deletion is widely distributed in Northern Eurasian populations (Fernandes et al., 2004, Repping et al., 2004).

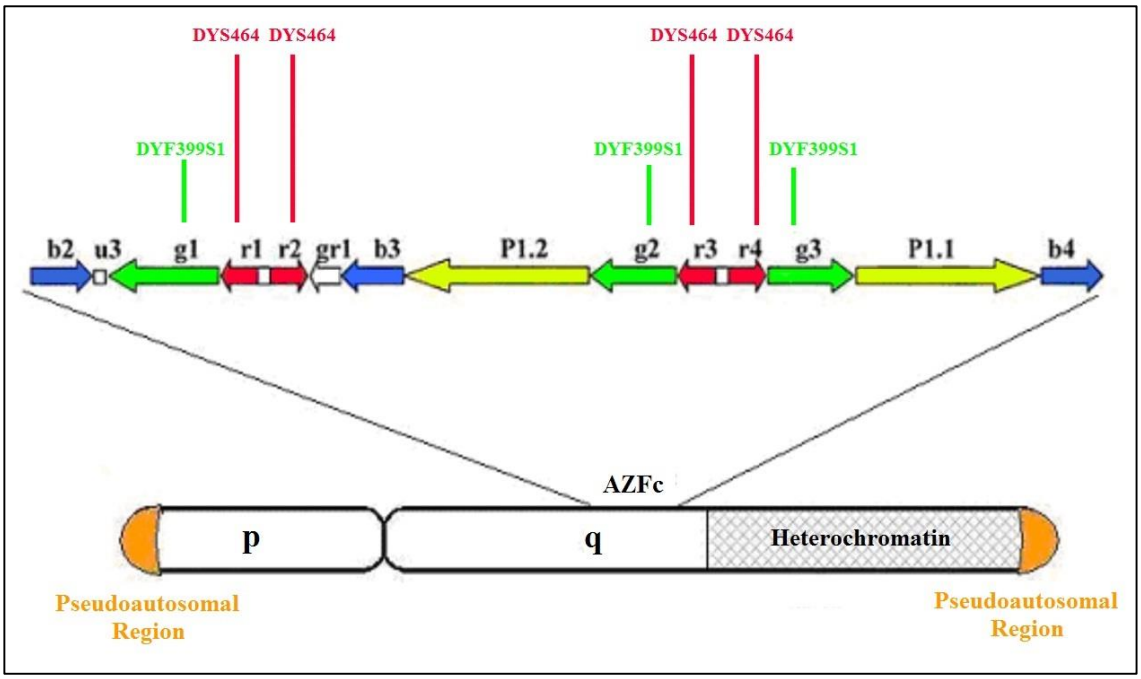

Figure 8.4: A subsection of Y chromosome showing the position of DYF399S1 locus within AZFc region (Adapted from Henson, 2005).

In the worldwide data for RM YSTRs, 308 of European male individuals showed a single allele at DYF399S1 including 88 male individuals from Finland, which was the largest population among the 111 populations, analysed. Y chromosome haplogroup N was found to be more frequent in Finns male individuals previously and hence it might be possible that single-allelic pattern at DYF399S1 is a result of a deletion (Zerjal et al., 1997, Derenko et al., 2007). DNA markers used for forensic DNA profiling and populations studies need to be unlinked to each other and other genetic loci. Despite of the fact that Y chromosome does not contain many genes, it has male specific genes which are responsible for production of sperm in males and hence markers included in such regions of the Y chromosomes, such as DYF399S1, might give a hint about male individual's ability of producing sperm (King et al., 2005). Due to such linkage, DYF399S1 locus might be excluded as it was the case with DYS464, a multi-allelic loci in close association with DYF399S1, which was recommended not to be used in DNA profiling testing previously (King et al., 2005, Rodig et al., 2008). However, this would be an unfortunate for forensic DNA profiling as DYF399S1 is the most diverse and

therefore most discriminating locus in RM-Yplex assay as determined in this study and previously published work (Henson, 2005, Ballantyne et al., 2012, Ballantyne et al., 2014)

8.1.4. Weight of Evidence of RM Y-STRs

The non-recombinant characteristic of Y chromosome prevents the application of product rule in order to estimate the weight of the evidence. Therefore, when a match is detected between evidence sample and suspected reference in casework investigation, the rarity of profile will be estimated typically using direct count method which is simply counting the number of times suspected profile was observed in the database divided by the total number of haplotypes in the database (Buckleton et al., 2011). This method has been known to be simple, robust, conservative and is used by many laboratories. In addition, there is no agreement by the forensic DNA profiling community on how to estimate the match probability when using RM-Yplex multiplex assay (Krawczak, 2001, Willuweit et al., 2011). This is because of the extremely high diversity of these loci included in this assay compared to commercially available Y-STR amplifications assays. A particularly large database will be required before using the counting method in estimating the rarity of the RM Y-STR haplotype or possibly a new approach might need developing, using special statistical measures to estimate their weight of evidence in forensic applications.

Considering the high mutation rate of the 13 RM Y-STR loci, such loci should be used in relationship testing with care and more stable Y-STR loci such as the ones included in Yfiler® and Powerplex® Y should be normally used. When RM Y-STR loci were used for relationship testing purposes, the mutation rate of each marker should be taken into consideration when interpreting generated profiles. The detection of mutations at three Y-STR loci is sufficient to exclude paternity (Kayser and Sajantila, 2001, Gjertson et al., 2007). However, this suggestion cannot be applied when using RM-Yplex considering the higher mutation rates of RM Y- STR loci. Recently the possibility of detecting up to

4 mutations in confirmed paternity was shown (Ballantyne et al., 2012). Thus new standards and methods need to be developed in order to fully utilize the potential of RM Y-STR loci.

8.1.5. Forensic Concerns

There are two main concerns when using RM Y-STR loci in forensic applications. Firstly, due to the high mutation rate of these markers, such mutations might be detected as an extra allele in the profile generated from the evidence, hence false exclusion might be possible or mixture suspicion could arise, complicating the analysis. The current study on RM Y-STRs have shown that the chance of two unrelated male individuals having the same profiles at all 13 RM Y-STR loci is extremely rare as has a published study (Ballantyne et al., 2014). Such mutations in germ lines would be pose difficulties in mixture resolution especially if male relative suspects are involved.

Secondly, somatic mutation rate study for RM Y-STR loci has not been done yet. Since the mechanisms of mutations in somatic cells are similar to the mutation mechanisms in germ line, it is expected, that the mutation rate at this region of DNA in somatic cells will be as high as mutation rate in germ lines (Strachan, 2004). If the 13 RM Y-STR loci mutation rates in somatic cells were as high, then the suitability of these loci for forensic casework application will need re-assessing.

8.2. Conclusion

During this research project, a novel Y-STR multiplex assay was developed, optimised and validated to amplify 13 RM Y-STR loci simultaneously. Four-dye chemistry was used following a systematic approach, which depends on careful primer design, location of loci within the multiplex and amplification efficiency. A sequenced allelic ladder was constructed for use in allele calling and for quality control purposes. The validation experiments conducted during this project established the robustness of the multiplex

assay in amplifying minute amounts of male DNA as well as amplifying male DNA in presence of different concentrations of inhibitors usually encountered in forensic DNA profiling. Application of RM-Yplex multiplex assay in the analysis of casework samples, required no additional modification to methods or instruments used currently in forensic DNA profiling routine. This proved the applicability of the assay for forensic DNA profiling. The combination of the 13 loci included in RM-Yplex has high power of Y-STR haplotype resolution, which is greater than all currently used Y-STR multiplex assays in forensic DNA profiling. The combination of both RM Y-STR haplotype database developed in this study provides a reference database to estimate a rarity of the profiles generated in forensic applications. Data analysis has shown that negligible population substructure was detected for UAE population. Considering the value of the assay, application of RM-Yplex multiplex assay in forensic laboratories would be strongly welcomed. However, there is a need to develop larger haplotype databases for RM Y-STRs compared to other Y-STR multiplex assays. This probably would be the only constraint that might face such application. In addition, RM-Yplex will replace the current three multiplexes assays available to amplify 13 RM Y-STR loci for population studies and other research purposes.

8.3. Future Work

RM-Yplex multiplex assay will introduce an easier and robust way to amplify the 13 RM Y-STR loci. These loci have attracted researchers due to their high mutation rates and therefore high diversity. There are many areas that need improving application of this assay in the forensic field and maximise haplotype resolution in male individuals. For this purpose, following future works are recommended:

Cloning of allelic ladder alleles should be attempted using standard methodology (Bai et al., 2010, Wang et al., 2014). Such work will overcome the limitations of sample

provision for the construction of allelic ladder and short-term storage allelic ladder developed during this project.

Methods to improve the haplotype resolution of the 13 RM Y-STR loci need to be developed further. First approach is to develop special interpretation guidelines for RM Y-STR haplotype. A large-scale study should be carried out for all multi allelic loci in order to explain and resolve the imbalance of peaks height.

A comprehensive mutation rate study of 13 RM Y-STR loci should be conducted to establish the mutation rates in somatic cells in order to evaluate the usefulness of these markers in forensic application.

The usefulness of these markers would be increased as different population databases grow, hence it is imperative that new populations should be profiled for RM Y-STRs using the assay developed and large databases are constructed for these markers.

9. Reference List

Abed, I. & Hellyer, P. 2001. *United Arab Emirates, a New Perspective*, Trident Press.

Al-Sayegh, F. 1998. *UAE from tribes to country*, Dubai, UAE, Al-Khaleej centre for books.

Alshamali, F., Pereira, L., Budowle, B., Poloni, E. S. & Currat, M. 2009. Local Population Structure in Arabian Peninsula Revealed by Y-STR Diversity. *Human Heredity,* 68**,** 45-54.

Altria, K. D. 1999. Overview of capillary electrophoresis and capillary electrochromatography. *Journal of Chromatography A,* 856**,** 443-463.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. 1990. Basic local alignment search tool. *Journal of Molecular Biology,* 215**,** 403-410.

Anushika, G., Dinuka, M., Priyani, H., Neil, F. & Ruwan, I. 2010. Development of a Low Cost In-House Method for the Analysis of Human Y-Chromosomal Mini STR Loci DYS19, DYS390 and DYS388. *Sri Lanka Journal of Forensic Medicine & Law,* 1**,** 3.

Bai, X., Li, S., Cong, B., Li, X., Guo, X., He, L., Ye, J. & Pei, L. 2010. Construction of two fluorescence-labeled non-combined DNA index system miniSTR multiplex systems to analyze degraded DNA samples in the Chinese Han Population. *Electrophoresis,* 31**,** 2944-2948.

Ballantyne, J., Hanson, E., Green, R., Holt, A. & Mulero, J. 2013. Enhancing the sexual assault workflow: Testing of next generation DNA assessment and Y-STR systems. *Forensic Science International: Genetics Supplement Series,* 4**,** e228-e229.

Ballantyne, K. N., Goedbloed, M., Fang, R., Schaap, O., Lao, O., Wollstein, A., Choi, Y., van Duijn, K., Vermeulen, M., Brauer, S., Decorte, R., Poetsch, M., von Wurmb-Schwark, N., de Knijff, P., Labuda, D., Vezina, H., Knoblauch, H., Lessig, R., Roewer, L., Ploski, R., Dobosz, T., Henke, L., Henke, J., Furtado, M. R. & Kayser, M. 2010. Mutability of Y-Chromosomal Microsatellites: Rates, Characteristics, Molecular Bases, and Forensic Implications. *The American Journal of Human Genetics,* 87**,** 341-353.

Ballantyne, K. N., Keerl, V., Wollstein, A., Choi, Y., Zuniga, S. B., Ralf, A., Vermeulen, M., de Knijff, P. & Kayser, M. 2012. A new future of forensic Y-chromosome analysis: Rapidly mutating Y-STRs for differentiating male relatives and paternal lineages. *Forensic Science International: Genetics,* 6**,** 208-218.

Ballantyne, K. N., Ralf, A., Aboukhalid, R., Achakzai, N. M., Anjos, M. J., Ayub, Q., Balažic, J., Ballantyne, J., Ballard, D. J., Berger, B., Bobillo, C., Bouabdellah, M., Burri, H., Capal, T., Caratti, S., Cárdenas, J., Cartault, F., Carvalho, E. F., Carvalho, M., Cheng, B., Coble, M. D., Comas, D., Corach, D., D'Amato, M. E., Davison, S., de Knijff, P., De Ungria, M. C. A., Decorte, R., Dobosz, T., Dupuy, B. M., Elmrghni, S., Gliwiński, M., Gomes, S. C., Grol, L., Haas, C., Hanson, E., Henke, J., Henke, L., Herrera-Rodríguez, F., Hill, C. R., Holmlund, G., Honda, K., Immel, U.-D., Inokuchi, S., Jobling, M. A., Kaddura, M., Kim, J. S., Kim, S. H., Kim, W., King, T. E., Klausriegler, E., Kling, D., Kovačević, L., Kovatsi, L., Krajewski, P., Kravchenko, S., Larmuseau, M. H. D., Lee, E. Y., Lessig, R., Livshits, L. A., Marjanović, D., Minarik, M., Mizuno, N., Moreira, H., Morling, N., Mukherjee, M., Munier, P., Nagaraju, J., Neuhuber, F., Nie, S.,

Nilasitsataporn, P., Nishi, T., Oh, H. H., Olofsson, J., Onofri, V., Palo, J. U., Pamjav, H., Parson, W., Petlach, M., Phillips, C., Ploski, R., Prasad, S. P. R., Primorac, D., Purnomo, G. A., Purps, J., Rangel-Villalobos, H., Rębała, K., Rerkamnuaychoke, B., Gonzalez, D. R., Robino, C., Roewer, L., Rosa, A., Sajantila, A., Sala, A., Salvador, J. M., Sanz, P., Schmitt, C., Sharma, A. K., Silva, D. A., Shin, K.-J., et al. 2014. Toward Male Individualization with Rapidly Mutating Y-Chromosomal Short Tandem Repeats. *Human Mutation,* 35**,** 1021-1032.

Balloux, F. & Lugon-Moulin, N. 2002. The estimation of population differentiation with microsatellite markers. *Molecular Ecology,* 11**,** 155-165.

Baltimore, D. 2001. Our genome unveiled. *Nature,* 409**,** 814-816.

Barber, M. D., Piercy, R. C., Andersen, J. F. & Parkin, B. H. 1995. Structural variation of novel alleles at the Hum vWA and Hum FES/FPS short tandem repeat loci. *International Journal of Legal Medicine,* 108**,** 31-35.

Barbujani, G. 2005. Human races: Classifying people vs understanding diversity. *Current Genomics.* Elsevier Science.

Batzer, M. A. & Deininger, P. L. 2002. Alu repeats and human genomic diversity. *Nat Rev Genet,* 3**,** 370-379.

Berger, B., Niederstätter, H., Brandstätter, A. & Parson, W. 2003. Molecular characterization and Austrian Caucasian population data of the multi-copy Y-chromosomal STR DYS464. *Forensic Science International,* 137**,** 221-230.

Bosch, E., Lee, A. C., Calafell, F., Arroyo, E., Henneman, P., de Knijff, P. & Jobling, M. A. 2002. High resolution Y chromosome typing: 19 STRs amplified in three multiplex reactions. *Forensic Science International,* 125**,** 42-51.

Botstein, D., White, R. L., Skolnick, M. & Davis, R. W. 1980. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *American Journal of Human Genetics,* 32**,** 314-331.

Brinkmann, B., Klintschar, M., Neuhuber, F., Huhne, J. & Rolf, B. 1998. Mutation rate in human microsatellites: Influence of the structure and length of the tandem repeat. *American journal of human genetics,* 62**,** 1408-1415.

Broeders, A. P. A. 2006. Of earprints, fingerprints, scent dogs, cot deaths and cognitive contamination brief look at the present state of play in the forensic arena. *Forensic Science International,* 159**,** 148-157.

Brookes, C., Bright, J. A., Harbison, S. & Buckleton, J. 2012. Characterising stutter in forensic STR multiplexes. *Forensic Science International: Genetics,* 6**,** 58-63.

Buckleton, J. S., Krawczak, M. & Weir, B. S. 2011. The interpretation of lineage markers in forensic DNA testing. *Forensic Science International: Genetics,* 5**,** 78-83.

Budimlija, Z., Prinz, M., Zelson-Moundorff, A., Wiersema, J., Bartelink, E., MacKinnon, G., Nazzaruolo, B., Estacio, S., Hennessey, M. & Shaler, R. 2003. World Trade Center human identification project: experiences with individual body identification cases. *Croatian Medical Journal,* 44**,** 259-263.

Budowle, B. & Allen, R. 1998. Analysis of Amplified Fragment-Length Polymorphisms (VNTR/STR Loci) for Human Identity Testing. *In:* LINCOLN, P. & THOMSON, J. (eds.) *Forensic DNA Profiling Protocols.* Humana Press.

Budowle, B., Moretti, T. R., Keys, K. M., Koons, B. W. & Smerick, J. B. 1997. Validation studies of the CTT STR multiplex system. *Journal of Forensic Sciences,* 42**,** 701-707.

Buehler, E. M. 1980. A synopsis of the human Y chromosome. *Human Genetics,* 55**,** 145-175.

Butler, J. M. 2005. *Forensic DNA typing: biology, technology, and genetics of STR* ***markers***, Elsevier Academic Press.

Butler, J. M. 2006. Genetics and genomics of core short tandem repeat loci used in human identity testing. *Journal of Forensic Sciences,* 51**,** 253-265.

Butler, J. M. 2007. Short tandem repeat typing technologies used in human identity testing. *BioTechniques,* 43**,** ii-iv.

Butler, J. M., Decker, A. E., Kline, M. C. & Vallone, P. M. 2005. Chromosomal duplications along the Y-chromosome and their potential impact on Y-STR interpretation. *Journal of Forensic Sciences,* 50**,** 853-859.

Butler, J. M. & Hill, C. R. 2012. Biology and genetics of new autosomal STR loci useful for forensic DNA analysis. *Forensic Science Review,* 24**,** 16-26.

Butler, J. M., Hill, C. R. & Coble, M. D. 2012. *Variability of new STR loci and kits in US population groups* [Online]. Promega Corpotaion. Available: http://www.promega.co.uk/resources/profiles-in-dna/2012/variability-of-new-str-loci-and-kits-in-us-population-groups/ [Accessed 05/09/2014 2014].

Butler, J. M., Li, J., Shaler, T. A., Monforte, J. A. & Becker, C. H. 1998. Reliable genotyping of short tandem repeat loci without an allelic ladder using time-of-flight mass spectrometry. *International Journal of Legal Medicine,* 112**,** 45-49.

Butler, J. M., McCord, B. R., Jung, J. M. & Allen, R. O. 1994. Rapid analysis of the short tandem repeat HUMTH01 by capillary electrophoresis. *BioTechniques,* 17.

Butler, J. M. & Schoske, R. 2004. Forensic value of the multicopy Y-STR marker DYS464. *International Congress Series,* 1261**,** 278-280.

Butler, J. M. & Schoske, R. 2005. U.S. population data for the multi-copy Y-STR locus DYS464. *Journal of Forensic Sciences,* 50**,** 975-977.

Butler, J. M., Schoske, R., Vallone, P. M., Kline, M. C., Redd, A. J. & Hammer, M. F. 2002. A novel multiplex for simultaneous amplification of 20 Y chromosome STR markers. *Forensic Science International,* 129**,** 10-24.

Butler, J. M., Shen, Y. & McCord, B. R. 2003. The development of reduced size STR amplicons as tools for analysis of degraded DNA. *Journal of Forensic Sciences,* 48**,** 1054-1064.

Butts, E. L. R., Kline, M. C., Duewer, D. L., Hill, C. R., Butler, J. M. & Vallone, P. M. 2011. NIST validation studies on the 3500 Genetic Analyzer. *Forensic Science International: Genetics Supplement Series,* 3**,** e184-e185.

Cadenas, A. M., Zhivotovsky, L. A., Cavalli-Sforza, L. L., Underhill, P. A. & Herrera, R. J. 2007. Y-chromosome diversity characterizes the Gulf of Oman. *Eur J Hum Genet,* 16**,** 374-386.

Chamberlain, J. S., Gibbs, R. A., Ranier, J. E., Nguyen, P. N. & Caskey, C. T. 1988. Deletion screening of the Duchenne muscular dystrophy locus via multiplex DNA amplification. *Nucleic Acids Research,* 16**,** 11141-11156.

Charlesworth, B. 2009. Effective population size and patterns of molecular evolution and variation. *Nat Rev Genet,* 10**,** 195-205.

Chen, N., Manabe, T., Terabe, S., Yohda, M. & Endo, I. 1994. High-resolution separation of oligonucleotides and DNA sequencing reaction products by capillary electrophoresis with linear polyacrylamide and laser-induced fluorescence detection. *Journal of Microcolumn Separations,* 6**,** 539-543.

Chung, D. T., Drábek, J., Opel, K. L., Butler, J. M. & McCord, B. R. 2004. A study on the effects of degradation and template concentration on the amplification efficiency of the STR Miniplex primer sets. *Journal of Forensic Sciences,* 49**,** 733-740.

Cloete, K., Ehrenreich, L., D'Amato, M. E., Leat, N., Davison, S. & Benjeddou, M. 2010. Analysis of seventeen Y-chromosome STR loci in the Cape Muslim population of South Africa. *Legal Medicine,* 12**,** 42-45.

Coble, M. D., Loreille, O. M., Wadhams, M. J., Edson, S. M., Maynard, K., Meyer, C. E., Niederstter, H., Berger, C., Berger, B., Falsetti, A. B., Gill, P., Parson, W. & Finelli, L. N. 2009. Mystery solved: The identification of the two missing Romanov children using DNA analysis. *PLoS ONE,* 4.

Cooper, D., Smith, B., Cooke, H., Niemann, S. & Schmidtke, J. 1985. An estimate of unique DNA sequence heterozygosity in the human genome. *Human Genetics,* 69**,** 201-205.

Crouse, C. A. & Schumm, J. 1995. Investigation of species specificity using nine PCR-based human STR systems. *Journal of Forensic Sciences,* 40**,** 952-956.

D'Amato, M. E., Benjeddou, M. & Davison, S. 2009. Evaluation of 21 Y-STRs for population and forensic studies. *Forensic Science International: Genetics Supplement Series,* 2**,** 446-447.

D'Amato, M. E., Ehrenreich, L., Cloete, K., Benjeddou, M. & Davison, S. 2010. Characterization of the highly discriminatory loci DYS449, DYS481, DYS518, DYS612, DYS626, DYS644 and DYS710. *Forensic Science International: Genetics,* 4**,** 104-110.

Dahm, R. 2005. Friedrich Miescher and the discovery of DNA. *Developmental Biology,* 278**,** 274-288.

Daniels, D. L., Hall, A. M. & Ballantyne, J. 2004. SWGDAM developmental validation of a 19-locus Y-STR system for forensic casework. *Journal of Forensic Sciences,* 49**,** 668-683.

Decker, A. E., Kline, M. C., Vallone, P. M. & Butler, J. M. 2007. The impact of additional Y-STR loci on resolving common haplotypes and closely related individuals. *Forensic Science International: Genetics,* 1**,** 215-217.

Dekairelle, A. F. & Hoste, B. 2001. Application of a Y-STR-pentaplex PCR (DYS19, DYS389I and II, DYS390 and DYS393) to sexual assault cases. *Forensic Science International,* 118**,** 122-125.

Demographic and Social Statistics [Online]. 2012. Available: http://www.uaestatistics.gov.ae [Accessed 12/19/2011 2011].

Derenko, M., Malyarchuk, B., Denisova, G., Wozniak, M., Grzybowski, T., Dambueva, I. & Zakharov, I. 2007. Y-chromosome haplogroup N dispersals from south Siberia to Europe. *J Hum Genet,* 52**,** 763-770.

Detwiler, M. M., Hamp, T. J. & Kazim, A. L. 2004. DNA sequencing using the liquid polymer POP-7™ on an ABI PRISM® 3100 Genetic Analyzer. *BioTechniques,* 36**,** 932-933.

Devaney, J. M., Girard, J. E. & Marino, M. A. 2000. DNA microsatellite analysis using ion-pair reversed-phase high- performance liquid chromatography. *Analytical Chemistry,* 72**,** 858-864.

Douadi, M. I., Gatti, S., Levrero, F., Duhamel, G., Bermejo, M., Vallet, D., Menard, N. & Petit, E. J. 2007. Sex-biased dispersal in western lowland gorillas (Gorilla gorilla gorilla). *Molecular Ecology,* 16**,** 2247-2259.

Edwards, A., Civitello, A., Hammond, H. A. & Thomas Caskey, C. 1991. DNA typing and genetic mapping with trimeric and tetrameric tandem repeats. *American Journal of Human Genetics,* 49**,** 746-756.

Edwards, A., Hammond, H. A., Jin, L., Caskey, C. T. & Chakraborty, R. 1992. Genetic variation at five trimeric and tetrameric tandem repeat loci in four human population groups. *Genomics,* 12**,** 241-253.

Erler, A., Stoneking, M. & Kayser, M. 2004. Development of Y-chromosomal microsatellite markers for nonhuman primates. *Molecular Ecology,* 13**,** 2921-2930.

Excoffier, L. & Lischer, H. E. L. 2010. Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources,* 10**,** 564-567.

Fan, H. & Chu, J. Y. 2007. A Brief Review of Short Tandem Repeat Mutation. *Genomics, Proteomics & Bioinformatics,* 5**,** 7-14.

Fernandes, S., Paracchini, S., Meyer, L. H., Floridia, G., Tyler-Smith, C. & Vogt, P. H. 2004. A Large AZFc Deletion Removes DAZ3/DAZ4 and Nearby Genes from Men in Y Haplogroup N. *American journal of human genetics,* 74**,** 180-187.

Flores, S., Sun, J., King, J. & Budowle, B. 2014. Internal validation of the GlobalFiler™ Express PCR Amplification Kit for the direct amplification of reference DNA samples on a high-throughput automated workflow. *Forensic Science International: Genetics,* 10**,** 33-39.

Frank, R. & Köster, H. 1979. DNA chain length markers and the influence of base composition on electrophoretic mobility of oligodeoxyribonucleotides in plyacrylamide-gels. *Nucleic Acids Research,* 6**,** 2069-2087.

Frégeau, C. J., Bowen, K. L. & Fourney, R. M. 1999. Validation of highly polymorphic fluorescent multiplex short tandem repeat systems using two generations of DNA sequencers. *Journal of Forensic Sciences,* 44**,** 133-166.

French, J. 2011. *Implementation of the PowerPlex® 18D System in a Databasing Laboratory: Transforming Operations to Improve Efficiency* [Online]. Available: http://www.promega.com/resources/articles/profiles-in-dna/2011/implementation-of-the-powerplex-18d-system-in-a-databasing-laboratory/.

Fujii, K., Iwashima, Y., Kitayama, T., Nakahara, H., Mizuno, N. & Sekiguchi, K. 2014. Allele frequencies for 22 autosomal short tandem repeat loci obtained by PowerPlex Fusion in a sample of 1501 individuals from the Japanese population. *Legal Medicine,* 16**,** 234-237.

Gilder, J. R., Doom, T. E., Inman, K. & Krane, D. E. 2007. Run-Specific Limits of Detection and Quantitation for STR-based DNA Testing. *Journal of Forensic Sciences,* 52**,** 97-101.

Gill, P. & Evett, I. 1995. Population genetics of short tandem repeat (STR) loci. *Genetica,* 96**,** 69-87.

Gill, P., Sparkes, R. & Kimpton, C. 1997. Development of guidelines to designate alleles using an STR multiplex system. *Forensic Science International,* 89**,** 185-197.

Gill, P., Urquhart, A., Millican, E., Oldroyd, N., Watson, S., Sparkes, R. & Kimpton, C. P. 1996. A new method of STR interpretation using inferential logic - Development of a criminal intelligence database. *International Journal of Legal Medicine,* 109**,** 14-22.

Gjertson, D. W., Brenner, C. H., Baur, M. P., Carracedo, A., Guidet, F., Luque, J. A., Lessig, R., Mayr, W. R., Pascali, V. L., Prinz, M., Schneider, P. M. & Morling, N. 2007. ISFG: Recommendations on biostatistics in paternity testing. *Forensic Science International: Genetics,* 1**,** 223-231.

Goedbloed, M., Vermeulen, M., Fang, R. N., Lembring, M., Wollstein, A., Ballantyne, K., Lao, O., Brauer, S., Kr++ger, C., Roewer, L., Lessig, R., Ploski, R., Dobosz, T., Henke, L., Henke, J., Furtado, M. R. & Kayser, M. 2009. Comprehensive mutation analysis of 17 Y-chromosomal short tandem repeat polymorphisms included in the AmpFSTR-« Yfiler-« PCR amplification kit. *International Journal of Legal Medicine,* 123**,** 471-482.

Goes, A., Silva, D., Domingues, C. S., Sobrinho, J. M. & Carvalho, E. F. 2002. Identification of a criminal by DNA typing in a rape case in Rio de Janeiro, Brazil. *Sao Paulo Medical Journal,* 120**,** 77-80.

Goodwin, W., Linacre, A. & Hadi, S. 2011. *An Introduction to Forensic Genetics,* John Wiley & Sons.

Gross, A. M., Liberty, A. A., Ulland, M. M. & Kuriger, J. K. 2008. Internal Validation of the AmpFlSTR Yfiler™ Amplification Kit for Use in Forensic Casework. *Journal of Forensic Sciences,* 53**,** 125-134.

Gusmão, L., Butler, J. M., Carracedo, A., Gill, P., Kayser, M., Mayr, W. R., Morling, N., Prinz, M., Roewer, L., Tyler-Smith, C. & Schneider, P. M. 2006. DNA Commission of the International Society of Forensic Genetics (ISFG): An update of the recommendations on the use of Y-STRs in forensic analysis. *Forensic Science International,* 157**,** 187-197.

Gusmaúo, L., Butler, J. M., Carracedo, A., Gill, P., Kayser, M., Mayr, W. R., Morling, N., Prinz, M., Roewer, L., Tyler-Smith, C. & Schneider, P. M. 2006. DNA Commission of the International Society of Forensic Genetics (ISFG): An update of the recommendations on the use of Y-STRs in forensic analysis. *Forensic Science International,* 157**,** 187-197.

Gusmaúo, L., Gonz+ílez-Neira, A., Pestoni, C., Bri+¦n, M., Lareu, M. V. & Carracedo, A. 1999. Robustness of the Y STRs DYS19, DYS389 I and II, DYS390 and DYS393: optimization of a PCR pentaplex. *Forensic Science International,* 106**,** 163-172.

Gusmaúo, L., S+ínchez-Diz, P., Alves, C., Beleza, S., Lopes, A., Carracedo, A. & Amorim, A. 2003. Grouping of Y-STR haplotypes discloses European geographic clines. *Forensic Science International,* 134**,** 172-179.

Hall, A. & Ballantyne, J. 2003. Novel Y-STR typing strategies reveal the genetic profile of the semen donor in extended interval post-coital cervicovaginal samples. *Forensic Science International,* 136**,** 58-72.

Hammer, M. F., Karafet, T. M., Redd, A. J., Jarjanazi, H., Santachiara-Benerecetti, S., Soodyall, H. & Zegura, S. L. 2001. Hierarchical patterns of global human Y-chromosome diversity. *Molecular Biology and Evolution,* 18**,** 1189-1203.

Hanson, E. K. & Ballantyne, J. 2007a. Population data for 48 'Non-Core' Y chromosome STR loci. *Legal Medicine,* 9**,** 221-231.

Hanson, E. K. & Ballantyne, J. 2007b. An ultra-high discrimination Y chromosome short tandem repeat multiplex DNA typing system. *PLoS ONE,* 2.

Harfe, B. D. & Jinks-Robertson, S. 2000. DNA MISMATCH REPAIR AND GENETIC INSTABILITY. *Annual Review of Genetics,* 34**,** 359-399.

Hattori, M. 2005. Finishing the euchromatic sequence of the human genome. *Tanpakushitsu kakusan koso. Protein, nucleic acid, enzyme,* 50**,** 162-168.

Hedman, M., Neuvonen, A. M., Sajantila, A. & Palo, J. U. 2011. Dissecting the Finnish male uniformity: The value of additional Y-STR loci. *Forensic Science International: Genetics,* 5**,** 199-201.

Hedman, M., Pimenoff, V., Lukka, M., Sistonen, P. & Sajantila, A. 2004. Analysis of 16 Y STR loci in the Finnish population reveals a local reduction in the diversity of male lineages. *Forensic Science International,* 142**,** 37-43.

Henegariu, O., Heerema, N. A., Dlouhy, S. R., Vance, G. H. & Vogt, P. H. 1997. Multiplex PCR: Critical parameters and step-by-step protocol. *BioTechniques,* 23**,** 504-511.

Henegariu, O., Hirschmann, P., Kilian, K., Kirsch, S., Lengauer, C., Maiwald, R., Mielke, K. & Vogt, P. 1994. Rapid screening of the Y chromosome in idiopathic sterile men, diagnostic for deletions in AZF, a genetic Y factor expressed during spermatogenesis. *Andrologia,* 26**,** 97-106.

Henson, G. 2005. DYF399S1: A Unique Three-Copy Short Tandem Repeat on the Human Y Chromosome. *Journal of Genetic Genealogy,* 1**,** 8-11.

Holsinger, K. E. & Weir, B. S. 2009. Genetics in geographically structured populations: defining, estimating and interpreting FST. *Nat Rev Genet,* 10**,** 639-650.

Honda, K., Roewer, L. & De Knijff, P. 1999. Malo DNA typing from 25-year-old vaginal swabs using Y chromosomal STR polymorphisms in a retrial request case. *Journal of Forensic Sciences,* 44**,** 868-872.

Inman, K. & Rudin, N. 2001. *Individualisation. Principles and Practices of Criminalistics: the Profession of Forensic Science.* Florida: CRC Press.

Ioulianos, A., Sismani, C., Fourouclas, N., Patroclou, T., Sergiou, C. & Patsalis, P. C. 2002. A nation-based population screening for azoospermia factor deletions in Greek-Cypriot patients with severe spermatogenic failure and normal fertile controls, using a specific study and experimental design. *International Journal of Andrology,* 25**,** 153-158.

Jacobs, M., Janssen, L., Vanderheyden, N., Bekaert, B., Van de Voorde, W. & Decorte, R. 2009. Development and evaluation of multiplex Y-STR assays for application in molecular genealogy. *Forensic Science International: Genetics Supplement Series,* 2**,** 57-59.

Jeffreys, A. J., Wilson, V. & Thein, S. L. 1985. Individual-specific /`fingerprints/' of human DNA. *Nature,* 316**,** 76-79.

Jobling, M. A., Pandya, A. & TylerSmith, C. 1997. The Y chromosome in forensic analysis and paternity testing. *International Journal of Legal Medicine,* 110**,** 118-124.

Jobling, M. A. & Tyler-Smith, C. 2003. The human Y chromosome: an evolutionary marker comes of age. *Nat Rev Genet,* 4**,** 598-612.

Karlin, S. & Altschul, S. F. 1993. Applications and statistics for multiple high-scoring segments in molecular sequences. *Proceedings of the National Academy of Sciences,* 90**,** 5873-5877.

Kashyap, V. K., Sitalaximi, T., Chattopadhyay, P. & Trivedi, R. 2004. DNA profiling technologies in forensic analysis. *International Journal of Human Genetics,* 4**,** 11-30.

Kayser, M., Brauer, S., Sch+ñdlich, H., Prinz, M., Batzer, M. A., Zimmerman, P. A., Boatin, B. A. & Stoneking, M. 2003. Y Chromosome STR Haplotypes and the Genetic Structure of U.S. Populations of African, European, and Hispanic Ancestry. *Genome Research,* 13**,** 624-634.

Kayser, M., Cagli+á, A., Corach, D., Fretwell, N., Gehrig, C., Graziosi, G., Heidorn, F., Herrmann, S., Herzog, B., Hidding, M., Honda, K., Jobling, M., Krawczak, M., Leim, K., Meuser, S., Meyer, E., Oesterreich, W., Pandya, A., Parson, W., Penacino, G., Perez-Lezaun, A., Piccinini, A., Prinz, M., Schmitt, C., Schneider, P. M., Szibor, R., Teifel-Greding, J., Weichhold, G., de Knijff, P. & Roewer, L. 1997a. Evaluation of Y-chromosomal STRs: A multicenter study. *International Journal of Legal Medicine.*

Kayser, M., Caglia, A., Corach, D., Fretwell, N., Gehrig, C., Graziosi, G., Heidorn, F., Herrmann, S., Herzog, B., Hidding, M., Honda, K., Jobling, M., Krawczak, M., Leim, K., Meuser, S., Meyer, E., Oesterreich, W., Pandya, A., Parson, W., Penacino, G., Perez-Lezaun, A., Piccinini, A., Prinz, M., Schmitt, C., Schneider, P., Szibor, R., Teifel-Greding, J., Weichhold, G., de Knijff, P. & Roewer, L. 1997b. Evaluation of Y-Chromosomal STRs: a multicenter study. *Int J Legal Med,* 110**,** 125 - 133.

Kayser, M., Kittler, R., Erler, A., Hedman, M., Lee, A. C., Mohyuddin, A., Mehdi, S. Q., Rosser, Z., Stoneking, M., Jobling, M. A., Sajantila, A. & Tyler-Smith, C. 2004. A Comprehensive Survey of Human Y-Chromosomal Microsatellites. *American journal of human genetics,* 74**,** 1183-1197.

Kayser, M., Lao, O., Anslinger, K., Augustin, C., Bargel, G., Edelmann, J., Elias, S., Heinrich, M., Henke, J., Henke, L., Hohoff, C., Illing, A., Jonkisz, A., Kuzniar, P., Lebioda, A., Lessig, R., Lewicki, S., Maciejewska, A., Monies, D., Paw+éowski, R., Poetsch, M., Schmid, D., Schmidt, U., Schneider, P., Stradmann-Bellinghausen, B., Szibor, R., Wegener, R., Wozniak, M., Zoledziewska, M., Roewer, L., Dobosz, T. & Ploski, R. 2005. Significant genetic differentiation between Poland and Germany follows present-day political borders, as revealed by Y-chromosome analysis. *Human Genetics,* 117**,** 428-443.

Kayser, M. & Sajantila, A. 2001. Mutations at Y-STR loci: implications for paternity testing and forensic analysis. *Forensic Science International Proceedings of the Second Forensic Y Chromosome User Workshop,* 118**,** 116-121.

Kimpton, C., Walton, A. & Gill, P. 1992. A further tetranucleotide repeat polymorphism in the vWF gene. *Human Molecular Genetics,* 1**,** 287.

Kimpton, C. P., Oldroyd, N. J., Watson, S. K., Frazier, R. R. E., Johnson, P. E., Millican, E. S., Urquhart, A., Sparkes, B. L. & Gill, P. 1996. Validation of highly discriminating multiplex short tandem repeat amplification systems for individual identification. *Electrophoresis,* 17**,** 1283-1293.

King, T. E., Bosch, E., Adams, S. M., Parkin, E. J., Rosser, Z. H. & Jobling, M. A. 2005. Inadvertent diagnosis of male infertility through genealogical DNA testing. *Journal of Medical Genetics,* 42**,** 366-368.

Kline, M. C., Duewer, D. L., Redman, J. W. & Butler, J. M. 2005. Results from the NIST 2004 DNA quantitation study. *Journal of Forensic Sciences,* 50**,** 571-578.

Klintschar, M. & Wiegand, P. 2003. Polymerase slippage in relation to the uniformity of tetrameric repeat stretches. *Forensic Science International,* 135**,** 163-166.

Korbie, D. J. & Mattick, J. S. 2008. Touchdown PCR for increased specificity and sensitivity in PCR amplification. *Nat.Protocols,* 3**,** 1452-1456.

Krawczak, M. 2001. Forensic evaluation of Y-STR haplotype matches: a comment. *Forensic Science International,* 118**,** 114-115.

Krenke, B. E., Viculis, L., Richard, M. L., Prinz, M., Milne, S. C., Ladd, C., Gross, A. M., Gornall, T., Frappier, J. R. H., Eisenberg, A. J., Barna, C., Aranda, X. G., Adamowicz, M. S. & Budowle, B. 2005. Validation of a male-specific, 12-locus fluorescent short tandem repeat (STR) multiplex. *Forensic Science International,* 148**,** 1-14.

Kroutil, L. C., Register, K., Bebenek, K. & Kunkel, T. A. 1996. Exonucleolytic Proofreading during Replication of Repetitive DNA. *Biochemistry,* 35**,** 1046-1053.

Kruglyak, S., Durrett, R. T., Schug, M. D. & Aquadro, C. F. 1998. Equilibrium distributions of microsatellite repeat length resulting from a balance between slippage events and point mutations. *Proceedings of the National Academy of Sciences,* 95**,** 10774-10778.

Kuhner, F., Morfill, J., Neher, R. A., Blank, K. & Gaub, H. E. 2007. Force-induced DNA slippage. *Biophysical Journal,* 92**,** 2491-2497.

Kuhr, W. G. 1990. Capillary electrophoresis. *Analytical Chemistry,* 62**,** 403R-414R.

Landers, J. P., Oda, R. P., Spelsberg, T. C., Nolan, J. A. & Ulfelder, K. J. 1993. Capillary electrophoresis: A powerful microanalytical technique for biologically active molecules. *BioTechniques,* 14**,** 98-109+111.

Landsteiner, K. 1900. Zur kenntnis der antifermentativen, lyteschen und agglutinierenden wirkungen des blustrums und der lymphe. *Zentralbl Bakteriol,* 27**,** 357-362.

Landsteiner, K. & Wiener, A. S. 1940. An Agglutinable Factor in Human Blood Recognized by Immune Sera for Rhesus Blood. *Experimental Biology and Medicine,* 43**,** 223.

Lazaruk, K., Wallin, J., Holt, C., Nguyen, T. & Walsh, P. S. 2001. Sequence variation in humans and other primates at six short tandem repeat loci used in forensic identity testing. *Forensic Science International,* 119**,** 1-10.

Lee, H. C., Ladd, C., Bourke, M. T., Pagliaro, E. & Tirnady, F. 1994. DNA Typing in Forensic Science: I. Theory and Background. *The American Journal of Forensic Medicine and Pathology,* 15.

Lewontin, R. C. & Hartl, D. L. 1991. Population genetics in forensic DNA typing. *Science,* 254**,** 1745-1750.

Li, G.-M. 2008. Mechanisms and functions of DNA mismatch repair. *Cell Res,* 18**,** 85-98.

Li, S., Ma, L., Li, L., Vang, S., Hu, Y., Bolund, L. & Wang, J. 2007. Snap: An integrated SNP annotation platform. *Nucleic Acids Research,* 35**,** D707-D710.

Litt, M. & Luty, J. A. 1989. A hypervariable microsatellite revealed by in vitro amplification of a dinucleotide repeat within the cardiac muscle actin gene. *American Journal of Human Genetics,* 44**,** 397-401.

Lygo, J. E., Johnson, P. E., Holdaway, D. J., Woodroffe, S., Whitaker, J. P., Clayton, T. M., Kimpton, C. P. & Gill, P. 1994. The validation of short tandem repeat (STR) loci for use in forensic casework. *International Journal of Legal Medicine,* 107**,** 77-89.

Malécot, G. 1948. *Les Mathématiques de l'hérédité,* Paris, France, Masson et Cie.

Mayntz-Press, K. A. & Ballantyne, J. 2007. Performance Characteristics of Commercial Y-STR Multiplex Systems*. *Journal of Forensic Sciences,* 52**,** 1025-1034.

McCord, B. R., Jung, J. M. & Holleran, E. A. 1993. High resolution capillary electrophoresis of forensic DNA using a non-gel sieving buffer. *Journal of Liquid Chromatography,* 16**,** 1963-1981.

McElreavey, K., Ravel, C., Chantot-Bastaraud, S. & Siffroi, J.-P. 2006. Y chromosome variants and male reproductive function. *International Journal of Andrology,* 29**,** 298-303.

Miles, S. B. 2005. *The countries and tribes of the Persian Gulf,* Kegan Paul.

Miyake, K., Yang, C., Minakuchi, Y., Ohori, K., Soutome, M., Hirasawa, T., Kazuki, Y., Adachi, N., Suzuki, S., Itoh, M., Goto, Y. I., Andoh, T., Kurosawa, H., Oshimura, M., Sasaki, M., Toyoda, A. & Kubota, T. 2013. Comparison of Genomic and Epigenomic Expression in Monozygotic Twins Discordant for Rett Syndrome. *PLoS ONE,* 8.

Mohammad, T., Xue, Y., Evison, M. & Tyler-Smith, C. 2009. Genetic structure of nomadic Bedouin from Kuwait. *Heredity,* 103**,** 425-433.

Mohammed, A. A. A., Linacre, A. M. T., Vanezis, P. & Goodwin, W. 2001. STR data for the GenePrint (TM) PowerPlex (TM) 1.2 system loci from three United Arab Emirates populations. *Forensic Science International,* 119**,** 328-329.

Moller, A., Meyer, E. & Brinkmann, B. 1994. Different types of structural variation in STRs: HumFES/FPS, HumVWA and HumD21S11. *International Journal of Legal Medicine,* 106**,** 319-323.

Moretti, T. R., Baumstark, A. L., Defenbaugh, D. A., Keys, K. M., Smerick, J. B. & Budowle, B. 2001. Validation of short tandem repeats (STRs) for forensic usage: Performance testing of fluorescent multiplex STR systems and analysis of authentic and simulated forensic samples. *Journal of Forensic Sciences,* 46**,** 647-660.

Mulero, J., Ballantyne, J., Ballantyne, K., Budowle, B., Coble, M., Gusmão, L., Roewer, L. & Kayser, M. 2014. Nomenclature update and allele repeat structure for the markers DYS518 and DYS449. *Forensic Science International: Genetics*.

Mulero, J. J., Chang, C. W., Calandro, L. M., Green, R. L., Li, Y., Johnson, C. L. & Hennessy, L. K. 2006. Development and Validation of the AmpFLSTR Yfiler(R) PCR Amplification Kit: A Male Specific, Single Amplification 17 Y-STR Multiplex System*. *Journal of Forensic Sciences,* 51**,** 64-75.

Mullis, K., Faloona, F., Scharf, S., Saiki, R., Horn, G. & Erlich, H. 1986. Specific enzymatic amplification of DNA in vitro: The polymerase chain reaction. *Cold Spring Harbor Symposia on Quantitative Biology,* 51**,** 263-273.

Nakamura, Y., Leppert, M., O'Connell, P., Lathrop, M., Lalouel, J. M. & White, R. 1989. A genetic linkage map of markers for human chromosome 20. *Genomics,* 5**,** 945-947.

NIST 2012. SWGDAM Validation Guidlines. *Validation Guidlines for DNA Analysis Methods.* USA: National Institute of Standard Technology.

Oota, H., Settheetham-Ishida, W., Tiwawech, D., Ishida, T. & Stoneking, M. 2001. Human mtDNA and Y-chromosome variation is correlated with matrilocal versus patrilocal residence. *Nat Genet,* 29**,** 20-21.

Pascali, V. L., Dobosz, M. & Brinkmann, B. 1998. Coordinating Y-chromosomal STR research for the Courts. *International Journal of Legal Medicine,* 112**,** 1-1.

Patzelt, D. 2004. History of forensic serology and molecular genetics in the sphere of activity of the German Society for Forensic Medicine. *Forensic Science International 100th Anniversary of the German Society of Legal Medicine,* 144**,** 185-191.

Pearson, C. E. & Sinden, R. R. 1998. Trinucleotide repeat DNA structures: Dynamic mutations from dynamic DNA. *Current Opinion in Structural Biology,* 8**,** 321-330.

Pena, S. D. J., Prado, V. F. & Epplen, J. T. 1995. DNA diagnosis of human genetic individuality. *Journal of Molecular Medicine,* 73**,** 555-564.

Polymeropoulos, M. H., Rath, D. S., Xiao, H. & Merril, C. R. 1991a. Tetranucleotide repeat polymorphism at the human c-fes/fps proto-oncogene (FES). *Nucleic Acids Research,* 19**,** 4018.

Polymeropoulos, M. H., Xiao, H., Rath, D. S. & Merril, C. R. 1991b. Tetranucleotide repeat polymorphism at the human tyrosine hydroxylase gene (TH). *Nucleic Acids Research,* 19**,** 3753.

Poznik, G. D., Henn, B. M., Yee, M. C., Sliwerska, E., Euskirchen, G. M., Lin, A. A., Snyder, M., Quintana-Murci, L., Kidd, J. M., Underhill, P. A. & Bustamante, C. D. 2013. Sequencing Y Chromosomes Resolves Discrepancy in Time to Common Ancestor of Males Versus Females. *Science,* 341**,** 562-565.

Prinz, M., Boll, K., Baum, H. & Shaler, B. 1997. Multiplexing of Y chromosome specific STRs and performance for mixed samples. *Forensic Science International,* 85**,** 209-218.

Prinz, M. & Sansone, M. 2001. Y chromosome-specific short tandem repeats in forensic casework. *Croatian Medical Journal,* 42**,** 288-291.

Puers, C., Hammond, H. A., Jin, L., Caskey, C. T. & Schumm, J. W. 1993. Identification of repeat sequence heterogeneity at the polymorphic short tandem repeat locus HUMTH01[AATG]n and reassignment of alleles in population analysis by using a locus-specific allelic ladder. *American Journal of Human Genetics*, 53**,** 953-958.

Purps, J., Siegert, S., Willuweit, S., Nagy, M., Alves, C., Salazar, R., Angustia, S. M. T., Santos, L. H., Anslinger, K., Bayer, B., Ayub, Q., Wei, W., Xue, Y., Tyler-Smith, C., Bafalluy, M. B., Martínez-Jarreta, B., Egyed, B., Balitzki, B., Tschumi, S., Ballard, D., Court, D. S., Barrantes, X., Bäßler, G., Wiest, T., Berger, B., Niederstätter, H., Parson, W., Davis, C., Budowle, B., Burri, H., Borer, U., Koller, C., Carvalho, E. F., Domingues, P. M., Chamoun, W. T., Coble, M. D., Hill, C. R., Corach, D., Caputo, M., D'Amato, M. E., Davison, S., Decorte, R., Larmuseau, M. H. D., Ottoni, C., Rickards, O., Lu, D., Jiang, C., Dobosz, T., Jonkisz, A., Frank, W. E., Furac, I., Gehrig, C., Castella, V., Grskovic, B., Haas, C., Wobst, J., Hadzic, G., Drobnic, K., Honda, K., Hou, Y., Zhou, D., Li, Y., Hu, S., Chen, S., Immel, U.-D., Lessig, R., Jakovski, Z., Ilievska, T., Klann, A. E., García, C. C., de Knijff, P., Kraaijenbrink, T., Kondili, A., Miniati, P., Vouropoulou, M., Kovacevic, L., Marjanovic, D., Lindner, I., Mansour, I., Al-Azem, M., Andari, A. E., Marino, M., Furfuro, S., Locarno, L., Martín, P., Luque, G. M., Alonso, A., Miranda, L. S., Moreira, H., Mizuno, N., Iwashima, Y., Neto, R. S. M., Nogueira, T. L. S., Silva, R., Nastainczyk-Wulf, M., Edelmann, J., Kohl, M., Nie, S., Wang, X., Cheng, B., et al. 2014. A global analysis of Y-chromosomal haplotype diversity for 23 STR loci. *Forensic Science International: Genetics*, 12**,** 12-23.

Quintana-Murci, L. & Fellous, M. 2001. The human Y chromosome: The biological role of a "functional wasteland". *Journal of Biomedicine and Biotechnology*, 2001**,** 18-24.

Quintana-Murci, L., Krausz, C. & McElreavey, K. 2001. The human Y chromosome: Function, evolution and disease. *Forensic Science International*, 118**,** 169-181.

Redd, A. J., Agellon, A. B., Kearney, V. A., Contreras, V. A., Karafet, T., Park, H., de Knijff, P., Butler, J. M. & Hammer, M. F. 2002. Forensic value of 14 novel STRs on the human Y chromosome. *Forensic Science International*, 130**,** 97-111.

Redd, A. J., Clifford, S. L. & Stoneking, M. 1997. Multiplex DNA typing of short-tandem-repeat loci on the Y chromosome. *Biological Chemistry*, 378**,** 923-927.

Report on the Current Activities of the Scientific Working Group on DNA Analysis Methods Y-STR Subcommittee. 2004. Unite state of America.

Repping, S., van Daalen, S. K. M., Korver, C. M., Brown, L. G., Marszalek, J. D., Gianotten, J., Oates, R. D., Silber, S., van der Veen, F., Page, D. C. & Rozen, S. 2004. A family of human Y chromosomes has dispersed throughout northern Eurasia despite a 1.8-Mb deletion in the azoospermia factor c region. *Genomics*, 83**,** 1046-1052.

Robert, K. & Neil, V. 2014. *Map Arabian Peninsula* [Online]. Gulf Art Guide. Available: http://gulfartguide.com/essay/contemporary-art-in-the-gulf-a-novel-phenomenon-2/1-4-map-arabian-peninsula-2/ [Accessed 7/6/2014 2014].

Rockenbauer, E., Hansen, S., Mikkelsen, M., Børsting, C. & Morling, N. 2014. Characterization of mutations and sequence variants in the D21S11 locus by next generation sequencing. *Forensic Science International: Genetics*, 8**,** 68-72.

Rodig, H., Roewer, L., Gross, A., Richter, T., de Knijff, P., Kayser, M. & Brabetz, W. 2008. Evaluation of haplotype discrimination capacity of 35 Y-chromosomal short tandem repeat loci. *Forensic Science International,* 174**,** 182-188.

Roewer, L., Arnemann, J., Spurr, N. K., Grzeschik, K. H. & Epplen, J. T. 1992. Simple repeat sequences on the human Y chromosome are equally polymorphic as their autosomal counterparts. *Human Genetics,* 89**,** 389-394.

Roewer, L., Croucher, P., Willuweit, S., Lu, T., Kayser, M., Lessig, R., de Knijff, P., Jobling, M., Tyler-Smith, C. & Krawczak, M. 2005. Signature of recent historical events in the European Y-chromosomal STR haplotype distribution. *Human Genetics,* 116**,** 279-291.

Rolf, B., Schürenkamp, M., Junge, A. & Brinkmann, B. 1997. Sequence polymorphism at the tetranucleotide repeat of the human beta- actin related pseudogene H-beta-Ac-psi-2 (ACTBP2) locus. *International Journal of Legal Medicine,* 110**,** 69-72.

Sachidanandam, R., Weissman, D., Schmidt, S. C., Kakol, J. M., Stein, L. D., Marth, G., Sherry, S., Mullikin, J. C., Mortimore, B. J., Willey, D. L., Hunt, S. E., Cole, C. G., Coggill, P. C., Rice, C. M., Ning, Z., Rogers, J., Bentley, D. R., Kwok, P. Y., Mardis, E. R., Yeh, R. T., Schultz, B., Cook, L., Davenport, R., Dante, M., Fulton, L., Hillier, L., Waterston, R. H., McPherson, J. D., Gilman, B., Schaffner, S., Van Etten, W. J., Reich, D., Higgins, J., Daly, M. J., Blumenstiel, B., Baldwin, J., Stange-Thomann, N., Zody, M. C., Linton, L., Lander, E. S. & Altshuler, D. 2001. A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. *Nature,* 409**,** 928-933.

Saiki, R. K., Gelfand, D. H., Stoffel, S., Scharf, S. J., Higuchi, R., Horn, G. T., Mullis, K. B. & Erlich, H. A. 1988. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science,* 239**,** 487-491.

Saiki, R. K., Scharf, S., Faloona, F., Mullis, K. B., Horn, G. T., Erlich, H. A. & Arnheim, N. 1985. Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science,* 230**,** 1350-1354.

Saks, M. J. & Koehler, J. J. 2008. *The Individualization Fallacy in Forensic Science Evidence*. *Vanderbilt Law Review,* 61**,** 199-219.

Sanger, F., Nicklen, S. & Coulson, A. R. 1977. DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences,* 74**,** 5463-5467.

Schlotterer, C. & Tautz, D. 1992. Slippage synthesis of simple sequence DNA. *Nucleic Acids Research,* 20**,** 211-215.

Schneider, P. M., Balogh, K., Naveran, N., Bogus, M., Bender, K., Lareu, M. & Carracedo, A. 2004. Whole genome amplification-the solution for a common problem in forensic casework? *Progress in Forensic Genetics 10,* 1261**,** 24-26.

Schneider, P. M., Meuser, S., Waiyawuth, W., Seo, Y. & Rittner, C. 1998. Tandem repeat structure of the duplicated Y-chromosomal STR locus DYS385 and frequency studies in the German and three Asian populations. *Forensic Science International,* 97**,** 61-70.

Schoske, R. 2003. *The design, optimisation and testing of Y chromosome short tandem repeat megaplexes.* PhD, American University.

Schoske, R., Vallone, P., Ruitberg, C. & Butler, J. 2003. Multiplex PCR design strategy used for the simultaneous amplification of 10-áY chromosome short tandem repeat (STR) loci. *Analytical and Bioanalytical Chemistry,* 375**,** 333-343.

Schoske, R., Vallone, P. M., Kline, M. C., Redman, J. W. & Butler, J. M. 2004. High-throughput Y-STR typing of U.S. populations with 27 regions of the Y chromosome using two multiplex PCR assays. *Forensic Science International,* 139**,** 107-121.

Schubbert, R. & Rittler, S. 2010. MULTIPLEX MICROSATELLITE ANALYSIS WITH 5 • f DYE-LABELLEDMARKER SETS USING FAM, YAKIMA YELLOW, ATTO 550, ATTO 565, AND ATTO 633 ON AN ABI 3130XL GENETIC ANALYZER. Applied Genetics, Anzinger Str. 7a, 85560 Ebersberg.

Schumm, J. W. 1997. Why use a size marker and allelic ladders in STR analysis. *Profiles in DNA,* 1**,** 3.

Serai, H. 1997. *History of old Arabian Peninsula,* Duabi, Al-Khaleej centre for books.

Seyfert, A. L., Cristescu, M. E. A., Frisse, L., Schaack, S., Thomas, W. K. & Lynch, M. 2008. The rate and spectrum of microsatellite mutation in Caenorhabditis elegans and Daphnia pulex. *Genetics,* 178**,** 2113-2121.

Sharma, V. & Litt, M. 1992. Tetranucleotide repeat polymorphism at the D21S11 locus. *Human Molecular Genetics,* 1**,** 67-67.

Shewale, J. G., Sikka, S. C., Schneida, E. & Sinha, S. K. 2003. DNA profiling of azoospermic semen samples from vasectomized males by using Y-PLEX™6 amplification kit. *Journal of Forensic Sciences,* 48**,** 127-129.

Shi, W., Ayub, Q., Vermeulen, M., Shao, R. g., Zuniga, S., van der Gaag, K., de Knijff, P., Kayser, M., Xue, Y. & Tyler-Smith, C. 2010. A Worldwide Survey of Human Male Demographic History Based on Y-SNP and Y-STR Data from the HGDPGÇôCEPH Populations. *Molecular Biology and Evolution,* 27**,** 385-393.

Shinde, D., Lai, Y., Sun, F. & Arnheim, N. 2003. Taq DNA polymerase slippage mutation rates measured by PCR and quasi-likelihood analysis: (CA/GT)n and (A/T)n microsatellites. *Nucleic Acids Research,* 31**,** 974-980.

Short Tandem Repeats Internet Database [Online]. 2014. Available: http://www.cstl.nist.gov/strbase/mutation.htm [Accessed 7/4/2014 2014].

Shriver, M. D., Jin, L., Chakraborty, R. & Boerwinkle, E. 1993. VNTR allele frequency distributions under the stepwise mutation model: A computer simulation approach. *Genetics,* 134**,** 983-993.

Shuber, A. P., Grondin, V. J. & Klinger, K. W. 1995. A simplified procedure for developing multiplex PCRs. *Genome Research,* 5**,** 488-493.

Shuber, A. P., Skoletsky, J., Sternl, R. & Handelin, B. L. 1993. Efficient 12-mutation testing in the CFTR gene: A general model for complex mutation analysis. *Human Molecular Genetics,* 2**,** 153-158.

Sia, E. A., Jinks-Robertson, S. & Petes, T. D. 1997. Genetic control of microsatellite stability. *Mutation Research/DNA Repair,* 383**,** 61-70.

Sibille, I., Duverneuil, C., Lorin de la Grandmaison, G., Guerrouache, K., Teissière, F., Durigon, M. & de Mazancourt, P. 2002. Y-STR DNA amplification as biological evidence in sexually assaulted female victims with no cytological detection of spermatozoa. *Forensic Science International,* 125**,** 212-216.

Sinha, S. K., Budowle, B., Arcot, S. S., Richey, S. L., Chakraborty, R., Jones, M. D., Wojtkiewicz, P. W., Schoenbauer, D. A., Gross, A. M., Sinha, S. K. & Shewale, J. G. 2003. Development and validation of a multiplexed Y-chromosome STR genotyping system, Y-PLEX™6, for forensic casework. *Journal of Forensic Sciences,* 48**,** 93-103.

Slatkin, M. 1995. A measure of population subdivision based on microsatellite allele frequencies. *Genetics,* 139**,** 457-462.

Smithies, O. 1955. Grouped Variations in the Occurrence of New Protein Components in Normal Human Serum. *Nature,* 175**,** 307-308.

Steinlechner, M., Berger, B., Niederstatter, H. & Parson, W. 2002. Rare failures in the amelogenin sex test. *Int J Legal Med,* 116**,** 117 - 120.

Strachan, T. 2004. *Human Molecular Genetic,* New York, USA, Grand Science /Taylo and Francis Group.

Stumpf, M. P. H. & Goldstein, D. B. 2001. Genealogical and Evolutionary Inference with the Human Y Chromosome. *Science,* 291**,** 1738-1742.

Sup Lee, J., Hanford, M. G., Genova, J. L. & Farber, R. 1999. Relative Stabilities of Dinucleotide and Tetranucleotide Repeats in Cultured Mammalian Cells. *Human Molecular Genetics,* 8**,** 2567-2572.

Thangaraj, K., Reddy, A. & Singh, L. 2002. Is the amelogenin gene reliable for gender identification in forensic casework and prenatal diagnosis? *Int J Legal Med,* 116**,** 121 - 123.

Thomas, M. G., Bradman, N. & Flinn, H. M. 1999. High throughput analysis of 10 microsatellite and 11 diallelic polymorphisms on the human Y-chromosome. *Human Genetics,* 105**,** 577-581.

Thompson, J. M., Ewing, M. M., Frank, W. E., Pogemiller, J. J., Nolde, C. A., Koehler, D. J., Shaffer, A. M., Rabbach, D. R., Fulmer, P. M., Sprecher, C. J. & Storts, D. R. 2013. Developmental validation of the PowerPlex-« Y23 System: A single multiplex Y-STR analysis system for casework and database samples. *Forensic Science International: Genetics,* 7**,** 240-250.

Thompson, R., Zoppis, S. & McCord, B. 2012. *An Overview of DNA Typing Methods for Human Identification: Past, Present, and Future*, HUMANA PRESS INC, 999 RIVERVIEW DR, STE 208, TOTOWA, NJ 07512-1165 USA.

Tiraby, J. G. & Fox, M. S. 1973. Marker discrimination in transformation and mutation of pneumococcus. *Proceedings of the National Academy of Sciences of the United States of America,* 70.

Triki-Fendri, S., Alfadhli, S., Ayadi, I., Kharrat, N., Ayadi, H. & Rebai, A. 2010. Genetic structure of Kuwaiti population revealed by Y-STR diversity. *Annals of human biology,* 37**,** 827-835.

Ullrich, A., Dull, T. J., Gray, A., Brosius, J. & Sures, I. 1980. Genetic variation in the human insulin gene. *Science,* 209**,** 612-615.

Underhill, P. A., Shen, P., Lin, A. A., Jin, L., Passarino, G., Yang, W. H., Kauffman, E., Bonne-Tamir, B., Bertranpetit, J., Francalacci, P., Ibrahim, M., Jenkins, T., Kidd, J. R., Mehdi, S. Q., Seielstad, M. T., Wells, R. S., Piazza, A., Davis, R. W., Feldman, M. W., Cavalli-Sforza, L. L. & Oefner, P. 2000. Y chromosome sequence variation and the history of human populations. *Nat Genet,* 26**,** 358-361.

Urquhart, A., Kimpton, C. P., Downes, T. J. & Gill, P. 1994. Variation in Short Tandem Repeat sequences GÇöa survey of twelve microsatellite loci for use as forensic identification markers. *International Journal of Legal Medicine,* 107**,** 13-20.

Van Oven, M., Huñmmerle, J. M., van Schoor, M., Kushnick, G., Pennekamp, P., Zega, I., Lao, O., Brown, L., Kennerknecht, I. & Kayser, M. 2010. Unexpected island effects at an extreme: reduced Y-chromosome and mitochondrial DNA diversity in Nias. *Molecular Biology and Evolution*.

Van Oven, M., Van Geystelen, A., Kayser, M., Decorte, R. & Larmuseau, M. H. 2014. Seeing the Wood for the Trees: A Minimal Reference Phylogeny for the Human Y Chromosome. *Human Mutation,* 35**,** 187-191.

Venter, J. C., Adams, M. D., Myers, E. W., Li, P. W., Mural, R. J., Sutton, G. G., Smith, H. O., Yandell, M., Evans, C. A., Holt, R. A., Gocayne, J. D., Amanatides, P., Ballew, R. M., Huson, D. H., Wortman, J. R., Zhang, Q., Kodira, C. D., Zheng, X. H., Chen, L., Skupski, M., Subramanian, G., Thomas, P. D., Zhang, J., Gabor Miklos, G. L., Nelson, C., Broder, S., Clark, A. G., Nadeau, J., McKusick, V. A., Zinder, N., Levine, A. J., Roberts, R. J., Simon, M., Slayman, C., Hunkapiller, M., Bolanos, R., Delcher, A., Dew, I., Fasulo, D., Flanigan, M., Florea, L., Halpern, A., Hannenhalli, S., Kravitz, S., Levy, S., Mobarry, C., Reinert, K., Remington, K., Abu-Threideh, J., Beasley, E., Biddick, K., Bonazzi, V., Brandon, R., Cargill, M., Chandramouliswaran, I., Charlab, R., Chaturvedi, K., Deng, Z., Francesco, V. D., Dunn, P., Eilbeck, K., Evangelista, C., Gabrielian, A. E., Gan, W., Ge, W., Gong, F., Gu, Z., Guan, P., Heiman, T. J., Higgins, M. E., Ji, R. R., Ke, Z., Ketchum, K. A., Lai, Z., Lei, Y., Li, Z., Li, J., Liang, Y., Lin, X., Lu, F., Merkulov, G. V., Milshina, N., Moore, H. M., Naik, A. K., Narayan, V. A., Neelam, B., Nusskern, D., Rusch, D. B., Salzberg, S., Shao, W., Shue, B., Sun, J., Wang, Z. Y., Wang, A., Wang, X., Wang, J., Wei, M. H., Wides, R., Xiao, C., Yan, C., et al. 2001. The Sequence of the Human Genome. *Science,* 291**,** 1304-1351.

Vermeulen, M., Wollstein, A., van der Gaag, K., Lao, O., Xue, Y., Wang, Q., Roewer, L., Knoblauch, H., Tyler-Smith, C., de Knijff, P. & Kayser, M. 2009. Improving global and regional resolution of male lineage differentiation by simple single-copy Y-chromosomal short tandem repeat polymorphisms. *Forensic Science International: Genetics,* 3**,** 205-213.

Victoria, K. 2010. *Genotyping of worldwide panel of rapidly mutating Y-STRs.* MSc Forensic Science MSc Dissertation, University Van Amsterdam.

Viguera, E., Canceill, D. & Ehrlich, S. D. 2001. In vitro replication slippage by DNA polymerases from thermophilic organisms. *Journal of Molecular Biology,* 312**,** 323-333.

Vuichard, S., Borer, U., Bottinelli, M., Cossu, C., Malik, N., Meier, V., Gehrig, C., Sulzer, A., Morerod, M. L. & Castella, V. 2011. Differential DNA extraction of challenging simulated sexual-assault samples: A Swiss collaborative study. *Investigative Genetics,* 2.

Wang, L., Zhao, X.-C., Ye, J., Liu, J.-J., Chen, T., Bai, X., Zhang, J., Ou, Y., Hu, L., Jiang, B.-W. & Wang, F. 2014. Construction of a library of cloned short tandem repeat (STR) alleles as universal templates for allelic ladder preparation. *Forensic Science International: Genetics,* 12**,** 136-143.

Watson, J. D. & Crick, F. H. C. 1953. Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid. *Nature,* 171**,** 737-738.

Weber-Lehmann, J., Schilling, E., Gradl, G., Richter, D. C., Wiehler, J. & Rolf, B. 2014. Finding the needle in the haystack: Differentiating "identical" twins in paternity testing and forensics by ultra-deep next generation sequencing. *Forensic Science International: Genetics,* 9**,** 42-46.

Weber, J. L. & Wong, C. 1993. Mutation of human short tandem repeats. *Human Molecular Genetics,* 2**,** 1123-1128.

Westring, C. G., Kristinsson, R., Gilbert, D. M. & Danielson, P. B. 2007. Validation of Reduced-Scale Reactions for the QuantifilerGäó Human DNA Kit. *Journal of Forensic Sciences,* 52**,** 1035-1043.

Willott, G. M. & Allard, J. E. 1982. Spermatozoa — their persistence after sexual intercourse. *Forensic Science International,* 19**,** 135-154.

Willuweit, S., Caliebe, A., Andersen, M. M. & Roewer, L. 2011. Y-STR Frequency Surveying Method: A critical reappraisal. *Forensic Science International: Genetics,* 5**,** 84-90.

Wright, S. 1949. THE GENETICAL STRUCTURE OF POPULATIONS. *Annals of Eugenics,* 15**,** 323-354.

Wright, S. 1978. *Evolution and the Genetics of Populations,* Chicago, University of Chicago Press.

Wu, J.-Y., Jiang, X.-T., Jiang, Y.-X., Lu, S.-Y., Zou, F. & Zhou, H.-W. 2010. Effects of polymerase, template dilution and cycle number on PCR based 16 S rRNA diversity analysis using the deep sequencing method. *BMC Microbiology,* 10**,** 255.

Xue, Y., Wang, Q., Long, Q., Ng, B. L., Swerdlow, H., Burton, J., Skuce, C., Taylor, R., Abdellah, Z., Zhao, Y., MacArthur, D. G., Quail, M. A., Carter, N. P., Yang, H. & Tyler-Smith, C. 2009. Human Y Chromosome Base-Substitution Mutation Rate Measured by Direct Sequencing in a Deep-Rooting Pedigree. *Current Biology,* 19**,** 1453-1457.

Zerjal, T., Dashnyam, B., Pandya, A., Kayser, M., Roewer, L., Santos, F. R., Schiefenhövel, W., Fretwell, N., Jobling, M. A., Harihara, S., Shimizu, K., Semjidmaa, D., Sajantila, A., Salo, P., Crawford, M. H., Ginter, E. K., Evgrafov, O. V. & Tyler-Smith, C. 1997. Genetic relationships of Asians and northern Europeans, revealed by Y- chromosomal DNA analysis. *American journal of human genetics,* 60**,** 1174-1183.

Zhang, Z., Kermekchiev, M. B. & Barnes, W. M. 2010. Direct DNA Amplification from Crude Clinical Samples Using a PCR Enhancer Cocktail and Novel Mutants of Taq. *The Journal of Molecular Diagnostics,* 12**,** 152-161.

Appendix 1

A1.1 DNA Sequences for 13 RM Y-STR Loci

DYF387S1 Locus

Forward Primer 5'-3': GCCTGGGTGACAGAGCTAGA
Reverse Primer 5'-3': GCCACAGTGTGAGAAGTGTGA
Variable motif: AAAG (13)
Repeat type: Tetra, complex, Multi (2)

Product A location: chrY:25931440-25931696
Product A size: 257bp
Product sequence:

GCCTGGGTGACAGAGCTAGATTCCATTTTACCCCTAACAAGAAAAAAAGAA
GAAGAAGAAAGAGAAAAAAAGAAAGAAAGGTAGGAAGGAAGGAAGGAAG
AAAGAAAGGAAGAAAGAAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGG
AAGGAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA
AGAAAGAAAGAAAATAAAAAAAACTGTGGTAGCACTTCTCACACTTCTCAC
ACTGTGGC

Product B location: chrY:28030679-28030939
Product B size: 261bp
Product sequence:

GCCTGGGTGACAGAGCTAGATTCCATTTTACCCCTAACAAGAAAAAAAGAA
GAAGAAGAAAGAGAAAAAAAGAAAGAAAGGTAGGAAGGAAGGAAGGAAG
AAAGAAAGGAAGAAAGAAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGG
AAGGAAGGAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA
AGAAAGAAAGAAAGAAAATAAAAAAAACTGTGGTAGCACTTCTCACACTTC
TCACACTGTGGC

DYF399S1 Locus

Forward Primer 5'-3': GGGTTTTCACCAGTTTGCAT
Reverse Primer 5'-3': CCATGTTTTGGGACATTCCT
Variable motif: GAAA (10-23)
Repeat type: Tetra, complex, Multi (3)

Product A location: chrY:26730126-26730414
Product A size: 289bp
Product sequence:

GGGTTTTCACCAGTTTGCATAGGTAGAGGGAGGCCAAAAGCCCAACAGGAA
AAAATAAAAGAAAGAAAGAAAAAGAAAAGAAAGAAAGAAAGAAAGAAAG
AAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGA
ACTTTTACCCTTTTGACAGCATATGAGACTTCTGGGTTCTTTTCTCCTGGGTC
CAATCCTAAGCTGTCCAGTTTAATGTTTGGGAAATTAACTCTTCCAAACTTG
GAGGATGCATTGAAGAGGAATGTCCCAAAACATGG

Product B location: chrY:25096369-25096670
Product B size: 302bp
Product sequence:

GGGTTTTCACCAGTTTGCATAGGTAGAGGGAGGCCAAAAGCCCAACAGGAA
AAAATAAAAGAAAGAAAGAAAAAGAAAAGGAAAGAAAGAAAGAAAGAAA
GAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG
AAAGAAAGAAAAAACTTTTACCCTTTTGACAGCATATGAGACTTCTGGGTTC
TTTTCTCCTGGGTCCAATCCTAAGCTGTCCAGTTTAATGTTTGGGAAATTAA
CTCTTCCAAACTTGGAGGATGCATTGAAGAGGAATGTCCCAAAACATGG

Product C location: chrY:27231970-27232262
Product C size: 293bp
Product sequence:

GGGTTTTCACCAGTTTGCATAGGTAGAGGGAGGCCAAAAGCCCAACAGGAA
AAAATAAAAGAAAGAAAGAAAAAGAAAAGAAAGAAAGAAAGAAAGAAAG
AAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGA
AAAAACTTTTACCCTTTTGACGGCATATGAGACTTCTGGGTTCTTTTCTCCTG
GGTCCAATCCTAAGCTGTCCAGTTTAATGTTTGGGAAATTAACTCTTCCAAA
CTTGGAGGATGCATTGAAGAGGAATGTCCCAAAACATGG

DYF403S1a/b Locus

Forward Primer 5'-3': CAAAATTCATGTGGATAATGAG
Reverse Primer 5'-3': ACAGAGCAGGATTCCATCTA
Variable motif: TTCT
Repeat type: Tetra, complex, Multi (3 + 1)

Product A location: chrY:9519608-9519948
Product A size: 341bp
Product sequence:

TGAAATTCATGTGGATAATGAGCAAAATCATAAGTTTTTATTTTTATTCATTTCCTTT
TGGTTTTATGCCTTTCATTCTCTTTCTCCTTCCTTCCCTTCTTCCCTCCCTCCCTTCTTC
CTTCTCTCCCTCCCTTCTCCCTGTCTTTCTTTCTTCCCTCCTTCCCTTCTCCCTGTCTGT
CTTTCTTTCGTTCTCTTTCTTTTTCTTTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTT
TCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTTTCTTTCTTTCTTTCCTCTAACTC
TGTATCTCTGTTTAATATTTTTAGATGGAATCCTGCTCTGT

Product B location: chrY:6225842-6226153
Product B size: 312bp
Product sequence:

CAAAATTCATGTGGATAATGAGTAAAATCATGTTTTTATTTTTATTCATTTCCTTTTG
GTTTCATGCCTTTCATTCTCTTTTTCTCCCTCCCTTCTTCCCTCCCTCCCTTCTCCCTG
TCTTTCTTTCCTTCTTTCCTTCTTTCTTTCCTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT
CTTTCTTTCTTTCTTTCTATCTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCT
TTCTTTCTTTCTTTCTTCCTTTCCTCTATCTCTGTATCTCTGTTTAATATTTTTAGATG
GAATCCTGCTCTGT

Product C location: chrY:9654225-9654540
Product C size: 316bp
Product sequence:

CAAAATTCATGTGGATAATGAGTAAAATCATGTTTTTATTTTTATTCATTTCCTTTTG
GTTTCATGCCTTTCATTCTCTTTTTCTCCCTCCCTTCTTCCCTCCCTCCCTTCTCCCTG
TCTTTCTTTCCTTCTTTCCTTCTTTCTTTCCTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT
CTTTCTTTCTATCTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCT
TTCTTTCTTTCTTTCTTTCTTCCTTTCCTCTATCTCTGTATCTCTGTTTAATATTTTTAG
ATGGAATCCTGCTCTGT

Product D location: chrY:6347622-6348058
Product D size: 437bp
Product sequence:

TGAAATTCATGTGGATAATGAGCAAAATCGTAAGTTTTTATTTTTATTCATTTCCTTT
TGGTTTTATGCCTTTCATTCTCTTTCTCCTTCCTTCCCTTCTTCCCTCCCTCCCTTCTTC
CTTCTCTCCCTCCCTTCTCCCTGTCTTTCTTTCTTCCCTCCTTCCCTTCTCCCTGTCTTT
CTTTCGTTCTCTTTCTTTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTC
TTTCTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCCTTCCTTCCTTCCTTCCTT
CCTTCCTTCCTTCCTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTC
TTTCTTTCTTTTTCTTTCTTTCTTTCCTCTAACTCTGTATCTCTGTTTAATATTTTTAGA
TGGAATCCTGCTCTGT

DYF404S1 Locus

Forward Primer 5'-3': GGCTTAAGAAATTTCAACGCATA
Reverse Primer 5'-3': CCATGATGGAACAATTGCAG
Variable motif: TTTC
Repeat type: Tetra, complex, Multi (2)

Product A location: chrY:25954055-25954251
Product A size: 197bp
Product sequence:

GGCTTAAGAAATTTCAACGCATATTTTTTCTTTCTTTCTTTCTTTCTTTCTTTC
TTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTCTCTTTCACTCTTTCTCTT
TCTCTCTTTCTCTCTTTCTTCCTTTCTTTCTTTCTTGACAGTGTCTCATTTTTTT
GTCAAAGATGGACTGCAATTGTTCCATCATGG

Product B location: chrY:28008132-28008320
Product B size: 189bp
Product sequence:

GGCTTAAGAAATTTCAACGCATATTTTTTCTTTCTTTCTTTCTTTCTTTCTTTC
TTTCTTTCTTTCTTTCTTTCTTTCTTTCTCTCTTTCACTCTTTCTCTTTCTCTCTT
TCTCTCTTTCTTCCTTTCTTTCTTTCTTGACAGTGTCTCATTTTTTTGTCAAAG
ATGGACTGCAATTGTTCCATCATGG

DYS449 Locus

Forward Primer 5'-3': TGGAGTCTCTCAAGCCTGTTC
Reverse Primer 5'-3': CCATTGCACTCTAGGTTGGAC
Variable motif: TTCT
Repeat type: Tetra, complex

Product location: chrY:8217908-8218227
Product size: 320bp
Product sequence:

TGGAGTCTCTCAAGCCTGTTCTATGAATATTTTCCCTTAACTTGTGTGATTTT
TTGTTTAAAAAGTTCCCCCATGTTTTTTTCTTGCTCTTTTTCTTTTCTCTCTTTT
TCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTT
TCTCTCTCTCCTCCTCTTTCTTTCCTTCTTTCTTTCTTTTCCTCTTTCCTTCTTT
CTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTCT
CTTTGTTTCTGTCTTACTCTTGTTGTCCAACCTAGAGTGCAATGG

DYS518 Locus

Forward Primer 5'-3': TGGAGTCTCTCAAGCCTGTTC
Reverse Primer 5'-3': CCATTGCACTCTAGGTTGGAC
Variable motif: AAAG
Repeat type: Tetra, complex

Product location: chrY:17319839-17320117
Product size: 279bp
Product sequence:

GGCAACACAAGTGAAACTGCTTCTCGAAAAAAGAAAGAAAGGAAGAAAGA
AAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAA
AGAAAGAAAGGGAGAAAGAAAGAAAGAAAGGAAGAGAAAGAAAGAAAGA
AAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGGAAGAGGGAG
GGAAGAAGAAAAGAAAGAAGGAAGGAAGGAAGGAGGGAAAAGAAGGAAA
GAAAAGAAAGAAATCACCCATGGTAAGAGCTGA

DYS526a/b Locus

Forward Primer 5'-3': TCTGGTGAACTGATCCAAACC
Reverse Primer 5'-3': GGGTTACTTCGCCAGAAGGT
Variable motif: CCTT
Repeat type: Tetra, complex, multi (1+1)

Product location: chrY:3640396+3640765
Product size: 370bp
Product sequence:

TCTGGTGAACTGATCCAAACCTTTACTTATTTTCTTCCTTTTTTCCTCTCTCCT
TCCCTCCCTCCCTCTTTCTTCCTTTCTTTCTTCCTTTCTTTCTTTCTTTCTTTCT
TTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCCTTCCTTCCTTCCTTCCTTCCTTCC
TTCCTTCCTTCTTTCCTTCTTTCCTTCATTTCTTCTTTCTTTCTCTTTACCTTCT
GGGAAACTGATCCAAACCTTTACTTACTTTCTATTTTTCTCTATTTTTCTTTCT
CTATTTTTTTCTCTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTC
CTTCCTTCCTTCCTTCTCTTTACCTTCTGGCGAAGTAACCC

DYS547 Locus

Forward Primer 5'-3': TCCATGTTACTGCAAAATACAC
Reverse Primer 5'-3': TGACAGAGCATAAACGTGTC
Variable motif: CCTT
Repeat type: Tetra, complex

Product location: chrY:18871923+18872360
Product size: 438bp
Product sequence:

TCCATGTTACTGCAAAATACACAATCTTATTTTTTTTACAGCTGAATAGTATT
TTATTTTGTATCTTTATGACATTTTTTCCTTCATCCCTTCCTTTTCCTTCCTTCT
TTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTTCTTC
CTTCCTTCCTTCCTTCTTTCCTCCCTCCCTTGCTCCCTCTCTCCCTTCCTCCTTC
CCTCCCTCCTTTCTTTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCT
TTCTTTCTTTCTTTCTTTCTTTCCTTATTTCTTCCTTCCTTCCTTCCTTTCTCTTT
CTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTTCTTTCTCTCTT
TTCTTTCTTTCTTTTCTTTCTTTTTTTTTGAGACACGTTTATGCTCTGTCA

DYS570 Locus

Forward Primer 5'-3': GAACTGTCTACAATGGCTCACG
Reverse Primer 5'-3': TCAGCATAGTCAAGAAACCAGACA
Variable motif: TTTC
Repeat type: Tetra, simple

Product location: chrY:6861115-6861370
Product size: 256bp
Product sequence:

GAACTGTCTACAATGGCTCACGTTACCATTAAACTTCAGTTATGTGACATCA
AGGTTATGAAACGTAAAATGAATGATGACTAGGTAGAAATCCTGGCTGTGT
CCTCCAAGTTCCTTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT
TCTTTCTTTCTTTCTTTCTTTCTTTCTTTTTGTAGATAGGGAATATCTGCATTT
CAGCTTAGGTTGCCACCAGTTGTCTGGTTTCTTGACTATGCTGA

DYS576 Locus

Forward Primer 5'-3': TTGGGCTGAGGAGTTCAATC
Reverse Primer 5'-3': GGCAGTCTCATTTCCTGGAG
Variable motif: AAAG
Repeat type: Tetra, simple

Product location: chrY:7053302-7053492
Product size: 191bp
Product sequence:

TTGGGCTGAGGAGTTCAATCTCAGCCAAGCAACATAGCAAGACCTCATCTC
TGAATAAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAA
GAAAGAAAGAAAGAAAGAAAGAAAGAAAAAGCCAAGACAAATACGCTTAT
TACTCCCATCTCCTCCTTCATCTCCAGGAAATGAGACTGCC

DYS612 Locus

Forward Primer 5'-3': CCCCCATGCCAGTAAGAATA
Reverse Primer 5'-3': TGAGGGAAGGCAAAAGAAAA
Variable motif: TCT
Repeat type: tri, complex

Product location: chrY:15752549-15752752
Product size: 204bp
Product sequence:

CCCCCATGCCAGTAAGAATAAAATTACAGCATGAAGTTTCACACAGGTTCA
GAGGTTTGCCTCCTCCTCCTCCTCTTTCTTCTTCTTCTCCTTCTTCTTCTTCTTC
TTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT
CTTCTGTCACTTTTCCAAATTATTTTCTTTTGCCTTCCCTCA

DYS626 Locus

Forward Primer 5'-3': GCAAGACCCCATAGCAAAAG
Reverse Primer 5'-3': AAGAAGAATTTTGGGACATGTTT
Variable motif: GAAA
Repeat type: tetra, complex

Product location: chrY:24416964-24417216
Product size: 253bp
Product sequence:

GCAAGACCCCATAGCAAAAGAAAAAAAAAGCAGAAAGGAAGAAAAAGAA
AGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAA
GAAAGAAAGAAAGAAAGAGAAAGAAAGAGGAAGGAAGGAAGAAAGAAAG
AAAGAGACAGAAAGAAAGAAAGAAAGAAAAAAGAAAGAAAGAAGGAAAG
AAAGAAAGGGAAGAAAGAAAGAGAAAGAAAGAACAAACATGTCCCAAAA
TTCTTCTT

DYS627 Locus

Forward Primer 5'-3': CTAGGTGACAGCGCAGGATT
Reverse Primer 5'-3': GGATAATGAGCAAATGGCAAG
Variable motif: AAAG
Repeat type: tetra, complex

Product location: chrY:8649930+8650266
Product size: 337bp
Product sequence:

CTAGGTGACAGCGCAGGATTCCATCTAAAAACAAAAATCAAAGAGAGACA
CAGGGAGAAAGAAAGAAAGGAAAAAAGAAAGAAAGAGAGAGAGAGAAAG
AAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGA
AAGAAAGAAAGAAAGAAAGAAAAAGAGAAAGAGAGAGGGAGGGAGGAGG
GAGGGAGGAAGTAAAGAAGGTAAGAAGGAAGGATGGAAGTAAGGAAGGA
AAGAAGGAAGGAAGGAACGAAGGAGAGAATGAAAGGCATAAAACCAAAG
GGAAATGAATAAAAATAAAAACTTGCCATTTGCTCATTATCC

A1.2 Allelic Ladder Construction

Table A1.1: Samples selected in developing allelic ladder presenting repeat numbers, mean of allele sizes in base pairs (bp) and standard deviation as a result out of precision study at each locus within blue channel.

Locus	Allele	Size Mean (bp)	Standard Deviation	Sample ID
DYS526a	11	131.4	0.029872	M381
DYS526a	12	135.28	0.02705	M345
DYS526a	13	139.45	0.02705	M194
DYS526a	14	143.53	0.043081	M268
DYS526a	15	147.70	0.025389	M287
DYS526a	16	151.68	0.062517	M226
DYS526a	17	155.77	0.030359	M164
DYS526a	18	159.61	0.082694	M329
DYS526b	30	339.40	0.072548	M411
DYS526b	31	343.38	0.02705	M381
DYS526b	32	347.20	0.044004	R120
DYS526b	33	351.44	0.026553	M345
DYS526b	34	355.50	0.02705	M194
DYS526b	35	359.57	0.031278	M268
DYS526b	36	363.54	0.02705	M204
DYS526b	37	367.57	0.039385	M307
DYS526b	38	371.56	0.037761	M287
DYS526b	39	375.56	0.087381	M226
DYS526b	40	379.51	0.037761	M164
DYS526b	41	383.50	0.047364	M329
DYS526b	42	387.48	0.061004	R131
DYS547	43	412.02	0.030749	M081
DYS547	44	416.03	0.037761	M071
DYS547	45	420.17	0.026379	M044
DYS547	46	424.14	0.021212	M075
DYS547	47	428.20	0.037761	M063
DYS547	48	432.19	0.037761	M032
DYS547	49	436.18	0.055007	M034
DYS547	50	440.13	0.084875	M031
DYS547	51	444.31	0.037761	M098
DYS547	52	448.37	0.062117	M201
DYS547	53	452.25	0.082001	G065
DYS547	54	456.30	0.082565	M316
DYS612	31	184.55	0.038264	M275
DYS612	32	187.68	0.075351	M385
DYS612	33	190.76	0.077031	M371
DYS612	34	193.68	0.042276	M031
DYS612	35	196.86	0.022188	M032
DYS612	36	199.89	0.02705	M087
DYS612	37	202.89	0.090077	M040
DYS612	38	205.95	0.026937	M069

Table A1.1 (Continued)

Locus	Allele	Size Mean (bp)	Standard Deviation	Sample ID
DYS612	39	208.98	0.028585	M066
DYS612	40	211.78	0.066626	M149
DYS612	41	214.80	0.054322	M180
DYF399S1	15	255.80	0.036349	B096
DYF399S1	16	259.84	0.05365	M062
DYF399S1	16.2	261.81	0.071525	M200
DYF399S1	17	263.86	0.081668	G020
DYF399S1	17.2	265.96	0.023534	F154
DYF399S1	18	267.94	0.089929	F143
DYF399S1	19	271.75	0.05383	G020
DYF399S1	19.1	272.85	0.027759	F154
DYF399S1	20	275.70	0.073758	F158
DYF399S1	20.1	276.72	0.02086	M040
DYF399S1	21	279.63	0.027528	G020
DYF399S1	21.1	280.59	0.027759	G020
DYF399S1	22	283.48	0.033407	F143
DYF399S1	22.1	284.46	0.022735	F143
DYF399S1	23	287.36	0.071986	M062
DYF399S1	23.1	288.31	0.041623	M062
DYF399S1	24	291.24	0.074379	B096
DYF399S1	24.1	292.17	0.047216	F158
DYF399S1	25	295.09	0.067985	F158
DYF399S1	25.1	296.04	0.047216	M081
DYF399S1	26	298.95	0.067831	M200
DYF399S1	26.1	299.86	0.055456	M040
DYF399S1	27	302.79	0.083066	B096
DYF399S1	27.1	303.64	0.027759	M294
DYF399S1	28	306.76	0.07456	M411
DYF399S1	28.1	307.66	0.03257	M411
DYF399S1	31	318.55	0.076272	M411

Table A1.2: Samples selected in developing allelic ladder presenting repeat numbers, mean of allele sizes in base pairs (bp) and standard deviation as a result out of precision.

Locus	Allele	Size Mean (bp)	Standard Deviation	Sample ID
DYS626	24	219.80	0.059689	M293
DYS626	25	223.99	0.072009	M081
DYS626	26	227.79	0.038264	M220
DYS626	27	231.80	0.038775	M169
DYS626	28	235.68	0.071823	M208
DYS626	29	239.58	0.02705	M170
DYS626	30	243.56	0.062993	M201
DYS626	31	247.60	0.038264	M194
DYS626	32	251.63	0.040666	M176
DYS626	33	255.41	0.040312	M175
DYS626	34	259.26	0.027631	M173
DYS626	35	263.10	0.063915	M048
DYF404S1	10	159.78	0.028434	M310
DYF404S1	11	173.80	0.051323	M271
DYF404S1	12	177.81	0.087785	M201
DYF404S1	12.2	179.85	0.0329	M287
DYF404S1	13	181.87	0.093132	M044
DYF404S1	13.2	183.99	0.027759	M287
DYF404S1	14	186.02	0.036885	M338
DYF404S1	14.2	188.16	0.02899	M201
DYF404S1	15	190.16	0.027759	M271
DYF404S1	15.2	192.30	0.075334	M288
DYF404S1	16	194.26	0.051979	M044
DYF404S1	16.2	196.30	0.074222	M310
DYF404S1	17	198.36	0.036512	M288
DYF404S1	18	202.28	0.0221	M023
DYF404S1	18.2	204.24	0.0487	M338
DYF403S1a	4	282.01	0.074458	M251
DYF403S1a	7	294.10	0.027759	M251
DYF403S1a	8	298.18	0.064295	M260
DYF403S1a	9	302.07	0.077847	M260
DYF403S1a	10	306.02	0.057707	M314
DYF403S1a	11	310.03	0.037761	Q007
DYF403S1a	12	314.18	0.043125	M347
DYF403S1a	13	318.46	0.027759	M314
DYF403S1a	14	322.74	0.049986	Q007
DYF403S1a	15	326.80	0.084499	M251
DYF403S1a	16	331.08	0.050607	M260
DYF403S1a	17	335.21	0.037761	M314
DYF403S1a	18	339.20	0.040585	Q007
DYF403S1a	19	343.16	0.045671	M347
DYF403S1a	20	347.08	0.035513	M347

study at each locus within green channel.

Table A1.2 (Continued)

Locus	Allele	Size Mean (bp)	Standard Deviation	Sample ID
DYF403S1b	43	419.76	0.024864	M201
DYF403S1b	45	427.74	0.042816	M251
DYF403S1b	46	431.88	0.027759	Q069
DYF403S1b	47	435.85	0.06609	M347
DYF403S1b	48	439.84	0.02854	M301
DYF403S1b	49	443.87	0.032859	M347
DYF403S1b	50	447.89	0.095541	M260
DYF403S1b	51	451.95	0.027759	M211
DYF403S1b	52	456.02	0.027759	M314
DYF403S1b	53	460.08	0.032088	Q007
DYF403S1b	54	464.05	0.043452	M054
DYF403S1b	55	468.06	0.046205	Q003
DYF403S1b	56	472.02	0.086111	G099

Table A1.3: Samples selected in developing allelic ladder presenting repeat numbers, mean of allele sizes in base pairs (bp) and standard deviation as a result out of precision study at each locus within yellow channel.

Locus	Allele	Size Mean (bp)	Standard Deviation	Sample ID
DYS518	33	254.86	0.068098	F161
DYS518	34	258.73	0.068604	M082
DYS518	35	262.54	0.04182	M096
DYS518	36	266.37	0.091975	M119
DYS518	37	270.21	0.073463	M092
DYS518	38	274.12	0.087941	M085
DYS518	39	278.02	0.073545	M093
DYS518	40	281.92	0.038264	M109
DYS518	41	285.76	0.077509	M089
DYS518	42	289.63	0.025841	M169
DYS518	43	293.46	0.043027	M176
DYS518	44	297.24	0.058098	M281
DYS518	45	301.16	0.045604	M355
DYS627	15	362.71	0.031475	M141
DYS627	15.2	364.62	0.052754	B058
DYS627	16	366.58	0.052754	M159
DYS627	16.2	368.49	0.029589	M280
DYS627	17	370.38	0.02705	M158
DYS627	18	374.23	0.021499	M143
DYS627	19	378.09	0.052754	M255
DYS627	20	381.96	0.057911	M270
DYS627	21	385.88	0.024657	M266
DYS627	21.2	387.81	0.038533	G134
DYS627	22	389.79	0.053351	M252
DYS627	22.3	392.73	0.035457	Q014
DYS627	23	393.69	0.056052	M240
DYS627	24	397.51	0.053351	M098
DYS627	25	401.48	0.053351	M372
DYS576	12	169.90	0.047216	B111
DYS576	13	173.83	0.05547	M173
DYS576	14	177.89	0.05547	M082
DYS576	15	181.59	0.05547	M092
DYS576	16	185.51	0.045162	M121
DYS576	17	189.45	0.05547	M085
DYS576	18	193.39	0.05547	M084
DYS576	19	197.33	0.05547	M111
DYS576	20	201.18	0.080616	M109
DYS576	21	205.09	0.026008	M252
DYS576	22	208.87	0.029975	M081

Table A1.4: Samples selected in developing allelic ladder presenting repeat numbers, mean of allele sizes in base pairs (bp) and standard deviation as a result out of precision study at each locus within red channel.

Locus	Allele	Size Mean (bp)	Standard Deviation	Sample ID
DYF387S1	34	240.00	0.059268	M243
DYF387S1	35	243.99	0.069678	B041
DYF387S1	36	248.02	0.027759	B041
DYF387S1	37	251.96	0.037761	M118
DYF387S1	38	255.81	0.027759	M194
DYF387S1	39	259.63	0.080904	B041
DYF387S1	40	263.55	0.021711	B041
DYF387S1	41	267.46	0.087507	M323
DYF387S1	42	271.32	0.044481	M118
DYS449	25	301.05	0.057711	M339
DYS449	26	305.03	0.064108	M329
DYS449	27	309.10	0.03212	M085
DYS449	28	313.07	0.076509	M140
DYS449	29	317.26	0.085749	M143
DYS449	30	321.53	0.027759	M096
DYS449	31	325.51	0.073869	M098
DYS449	32	329.63	0.027759	M092
DYS449	33	333.74	0.047216	M093
DYS449	34	337.74	0.081742	M111
DYS449	35	341.88	0.037761	M207
DYS449	36	345.84	0.027759	M169
DYS570	11	109.42	0.040148	G147
DYS570	12	113.40	0.060783	M190
DYS570	13	117.38	0.037761	E038
DYS570	14	121.33	0.086873	M067
DYS570	15	125.24	0.025159	M066
DYS570	16	129.11	0.08574	M065
DYS570	17	133.12	0.038264	M022
DYS570	18	137.14	0.033469	M023
DYS570	19	141.22	0.08581	M026
DYS570	20	145.30	0.042974	M075
DYS570	21	149.31	0.038264	M170
DYS570	22	153.62	0.048148	M293

A1.3 External Proficiency Testing

Results of 8 quality control received from Prof. Manfred Kayser, Erasmus MC, Rotterdam, Netherland as part of an exercise, and male DNA control 007 are presented in this section.

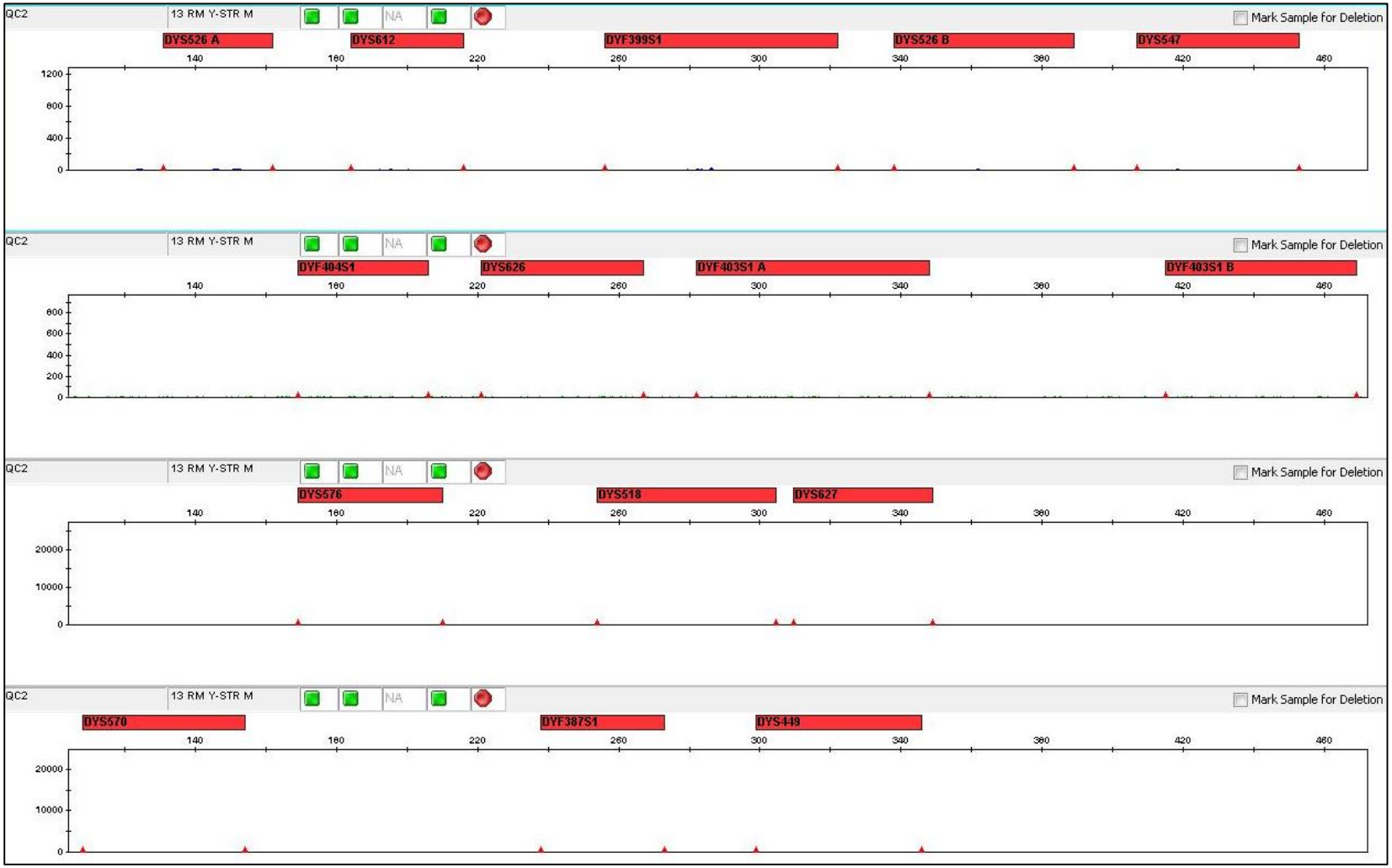

Figure A1.1: Electropherogram of quality control sample number 1.

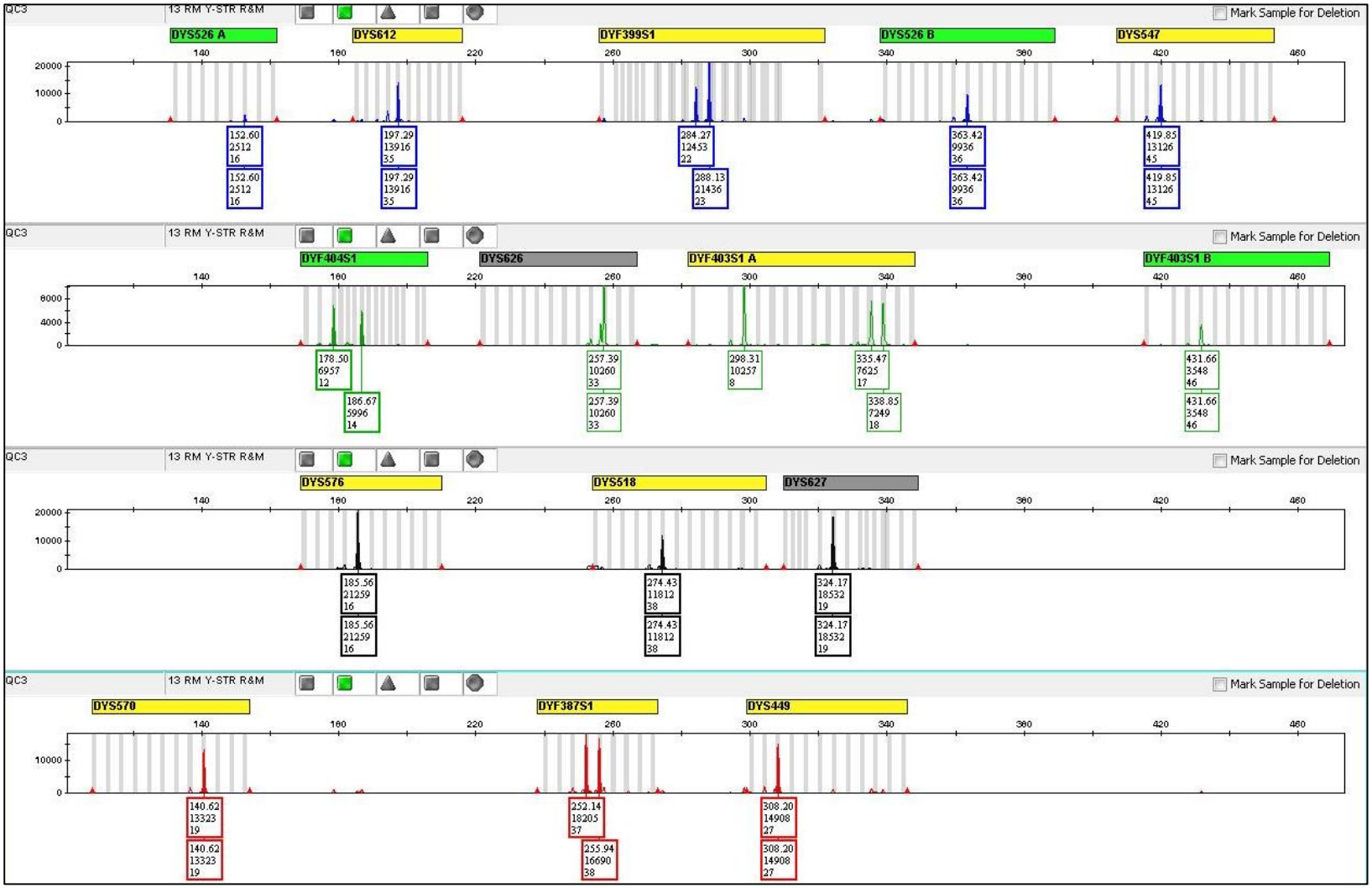

Figure A1.2: Electropherogram of quality control sample number 2.

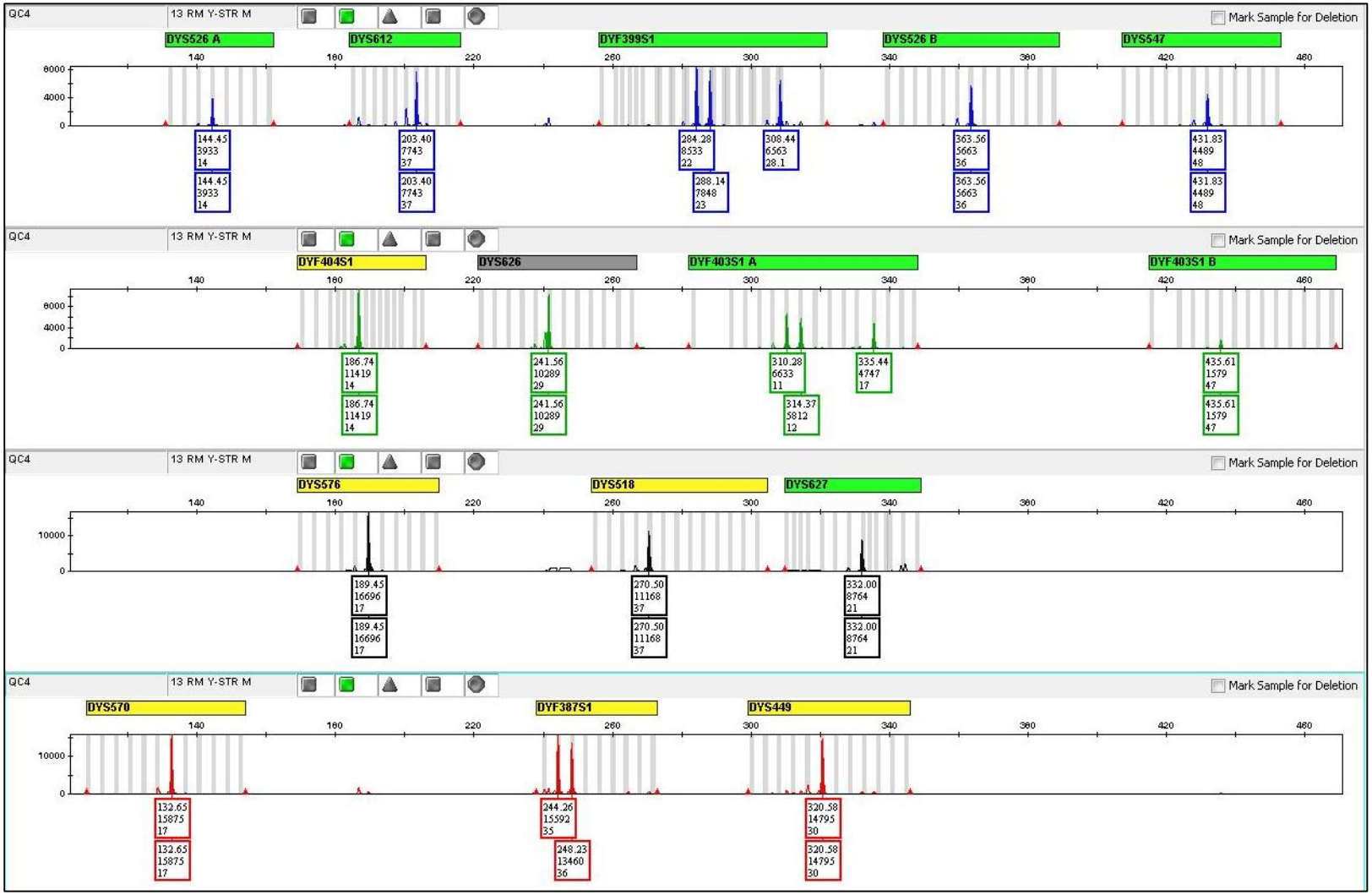

Figure A1.3: Electropherogram of quality control sample number 3.

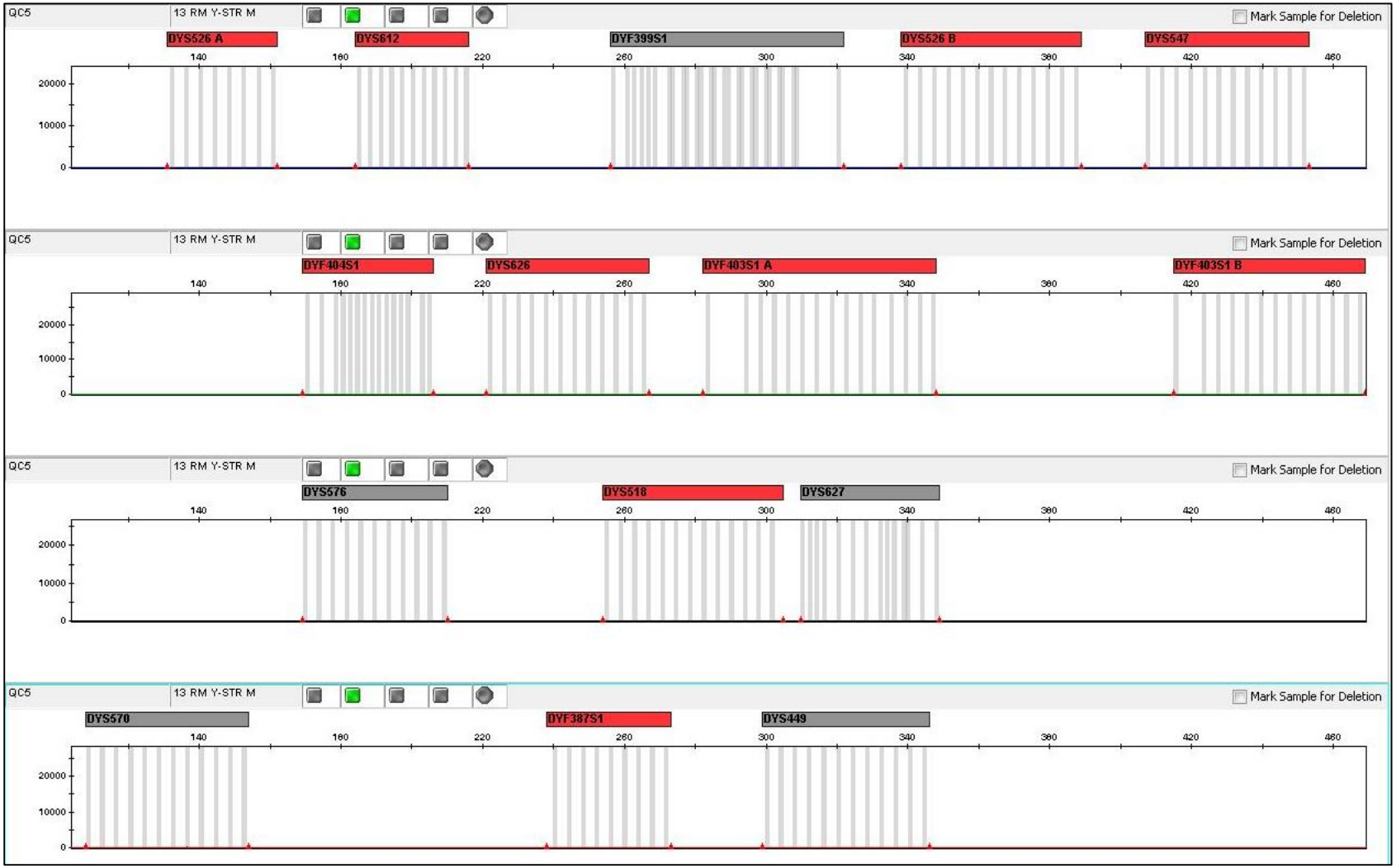

Figure A1.4: Electropherogram of quality control sample number 4.

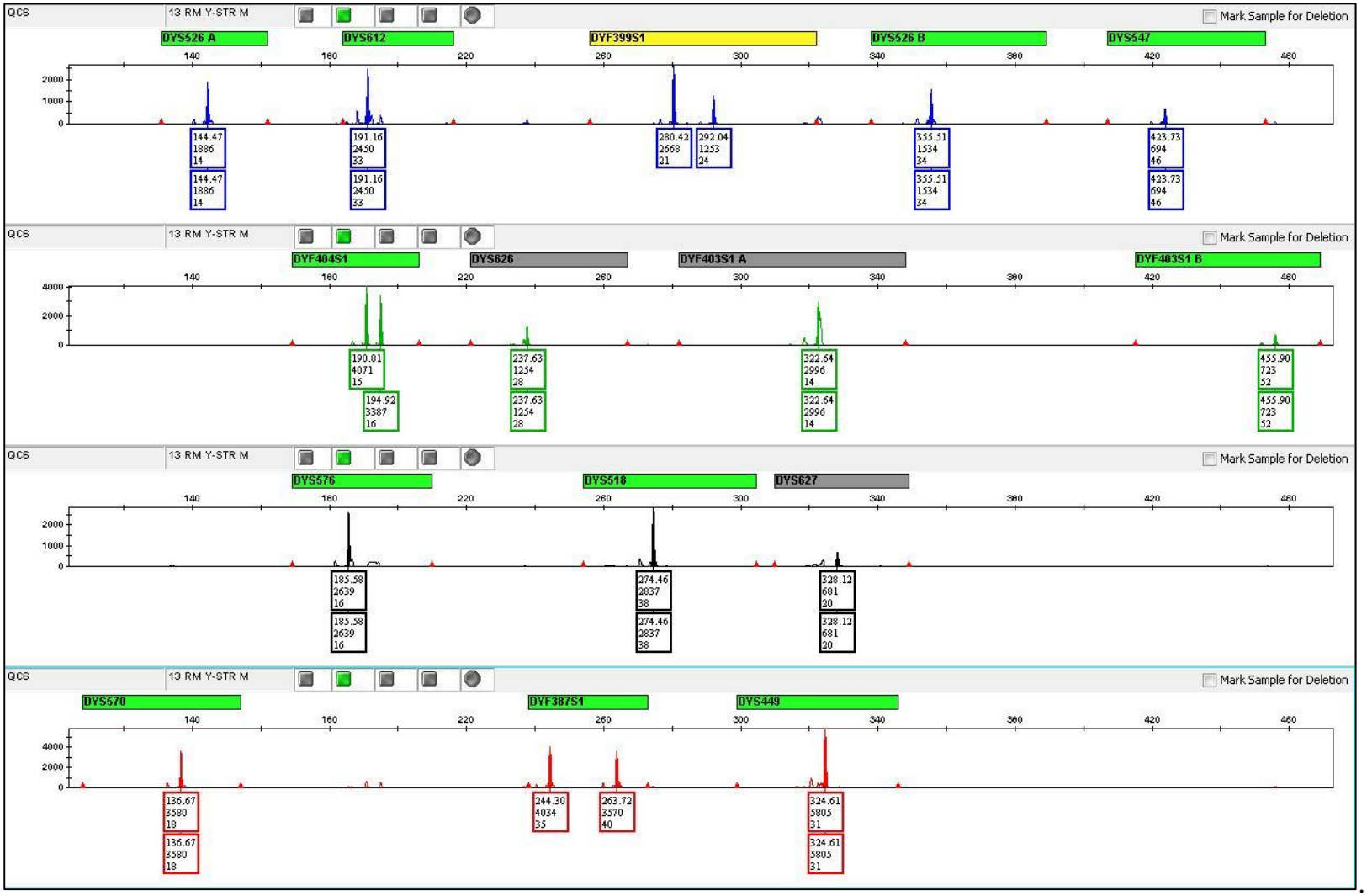

Figure A1.5: Electropherogram of quality control sample number 5.

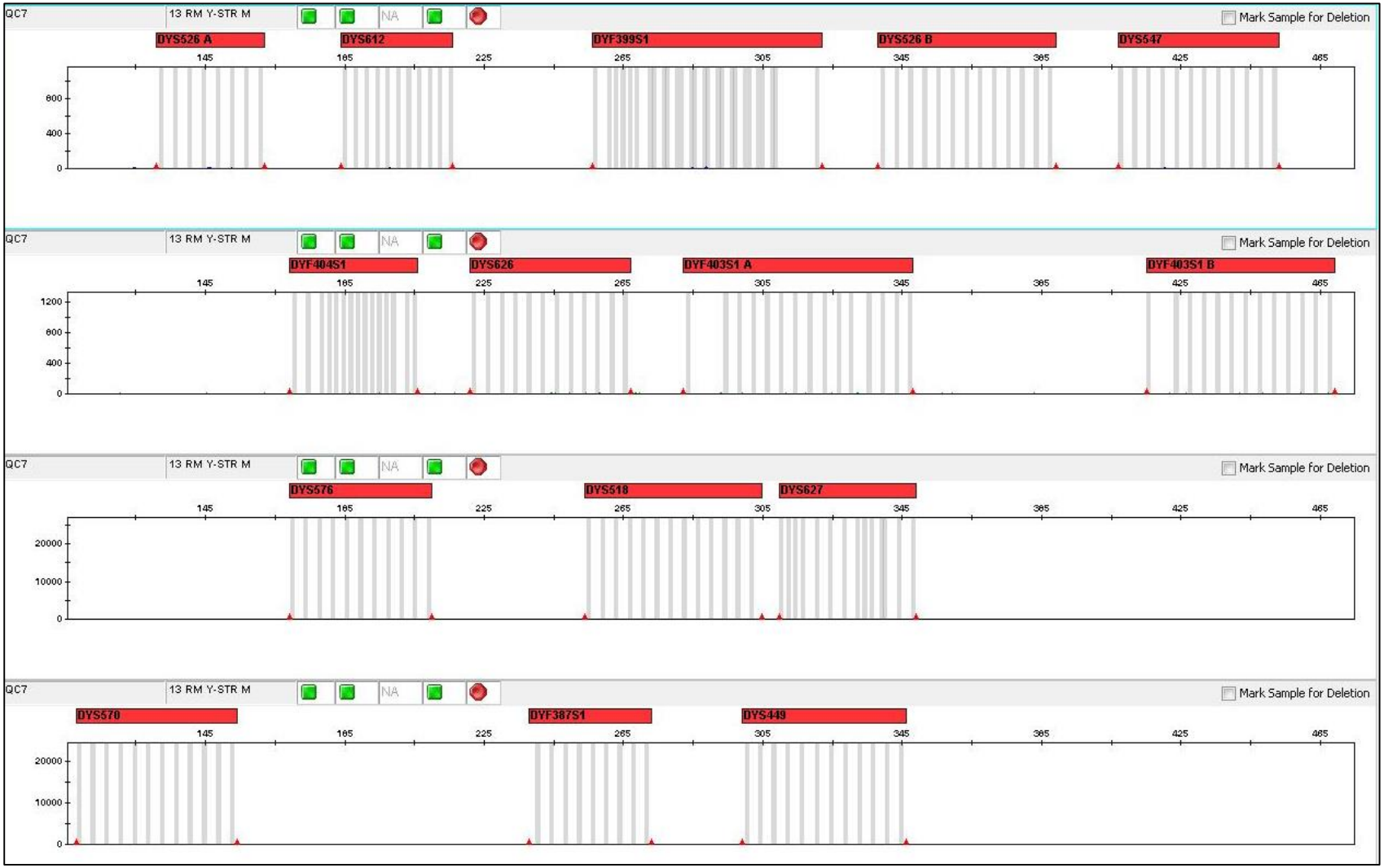

Figure A1.6: Electropherogram of quality control sample number 6.

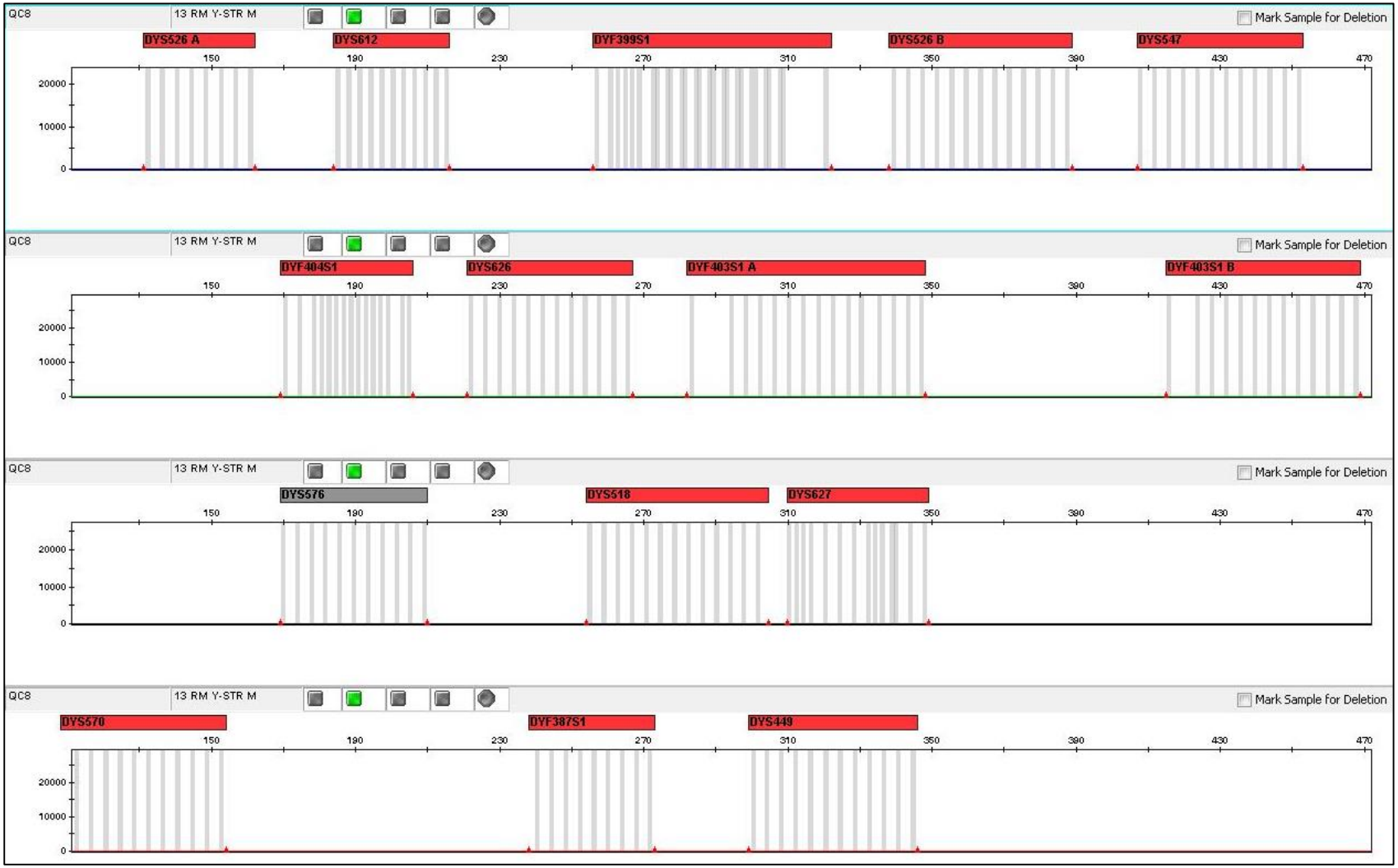

Figure A1.7: Electropherogram of quality control sample number 7.

Figure A1.8: Electropherogram of quality control sample number 8.

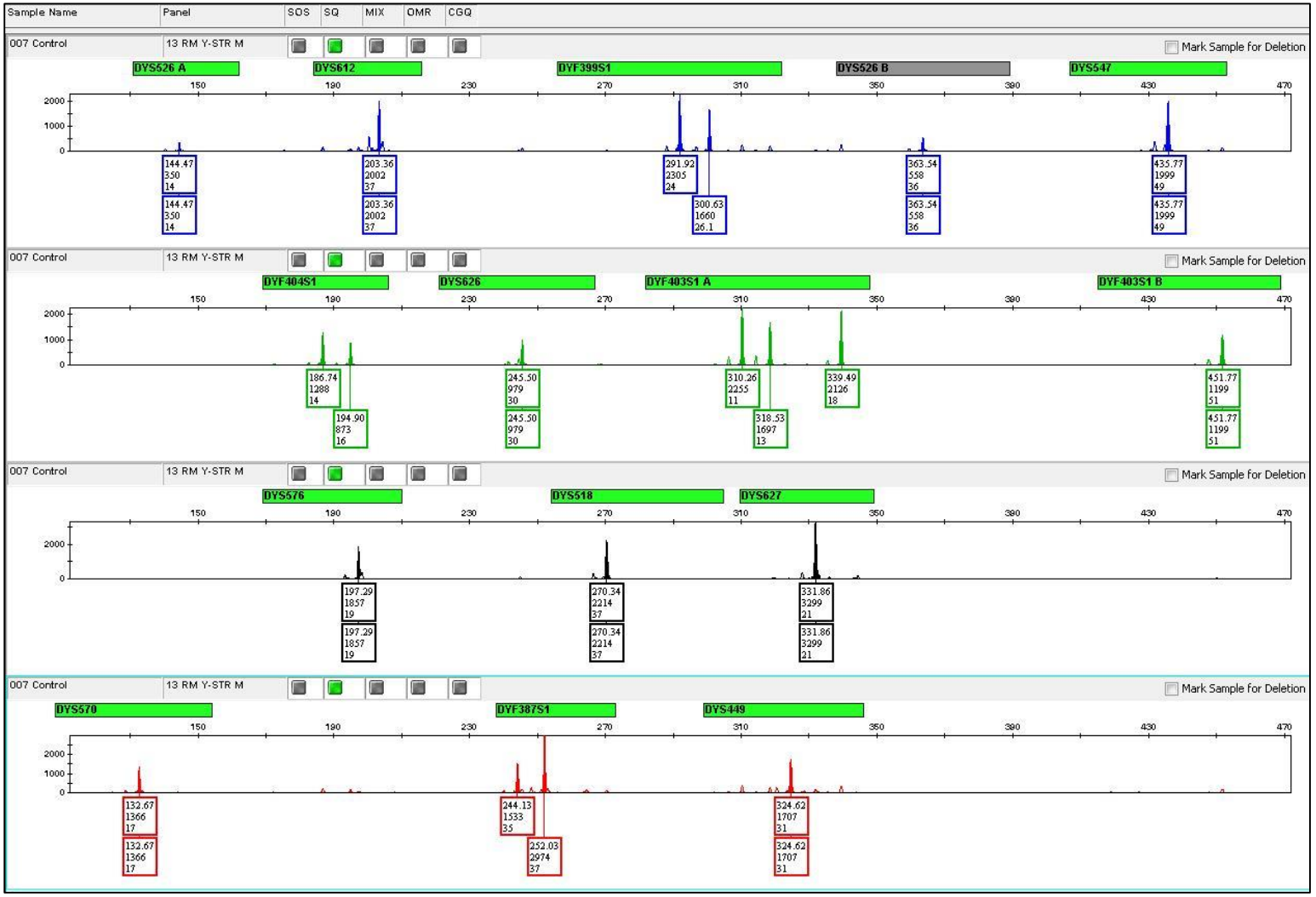

Figure A1.9: Electropherogram of positive male DNA control 007.

Appendix 2

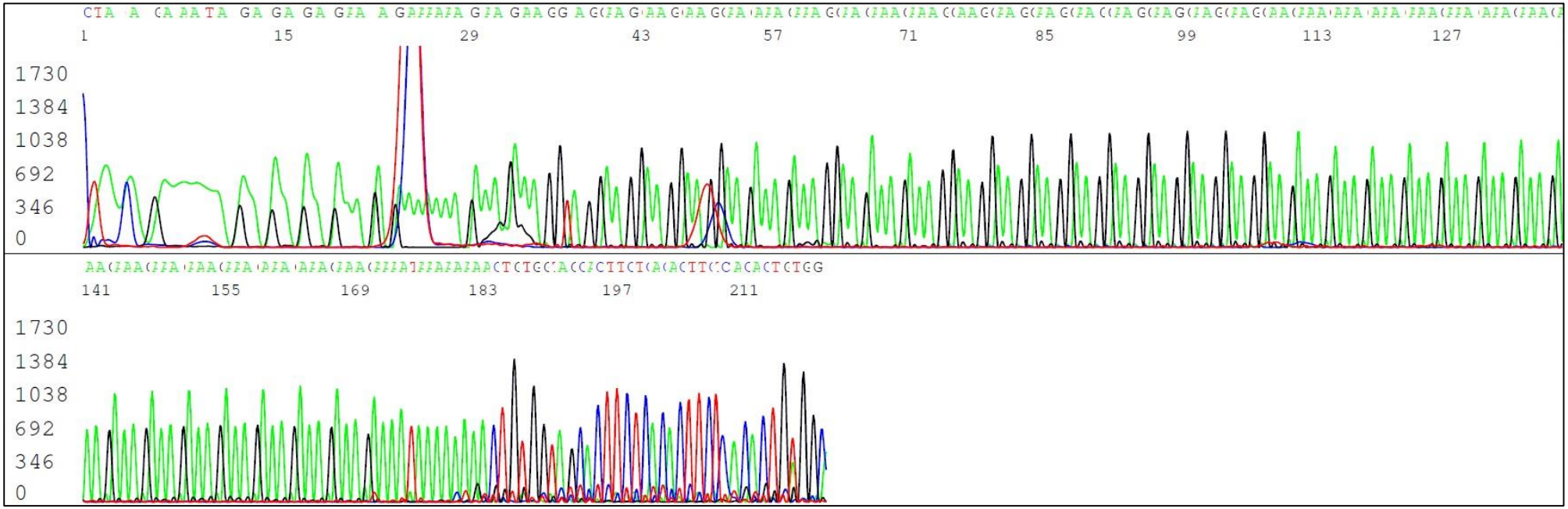

Figure A2.1: Electriopherogram presenting a sequence of Allele 16 at DYF387S1 locus with Repeat motif [GAAA].

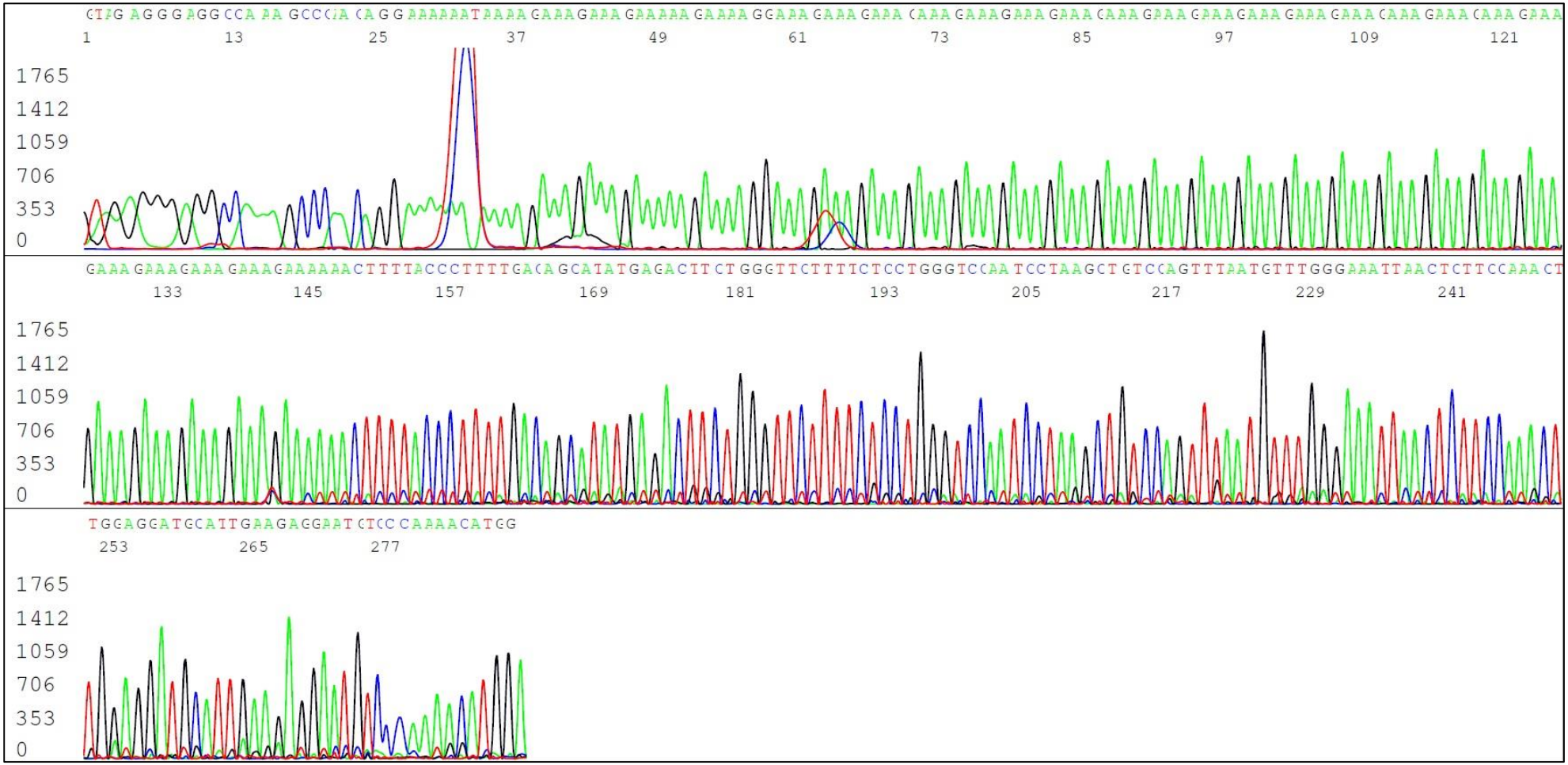

Figure A2.2: Electropherogram presenting a sequence of allele 22 at DYF399S1 locus with repeat motif [GAAA]

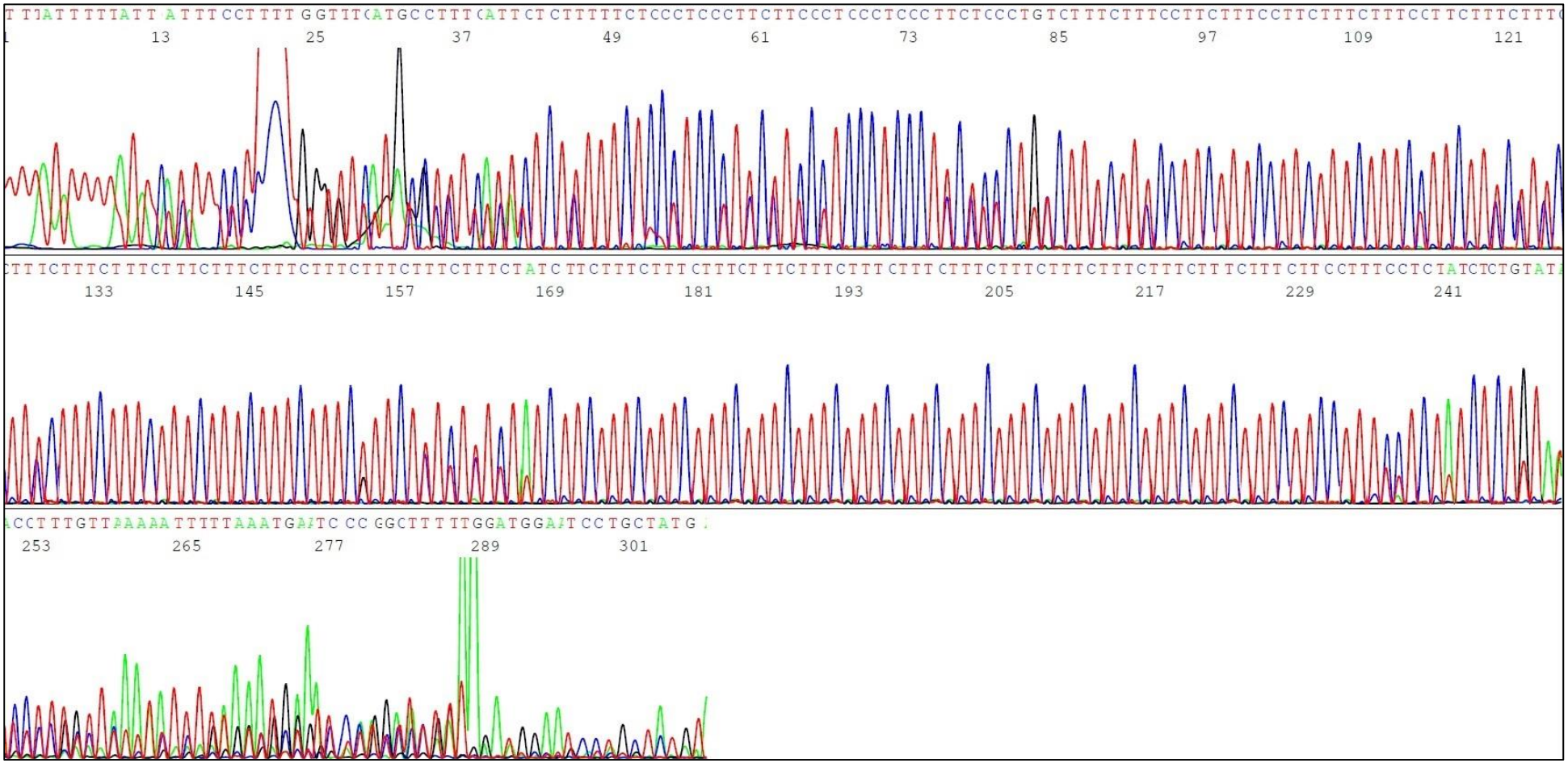

Figure A2.3: Electropherogram presenting a sequence of allele 15 at DYF403S1a locus with repeat motif [TTCT]

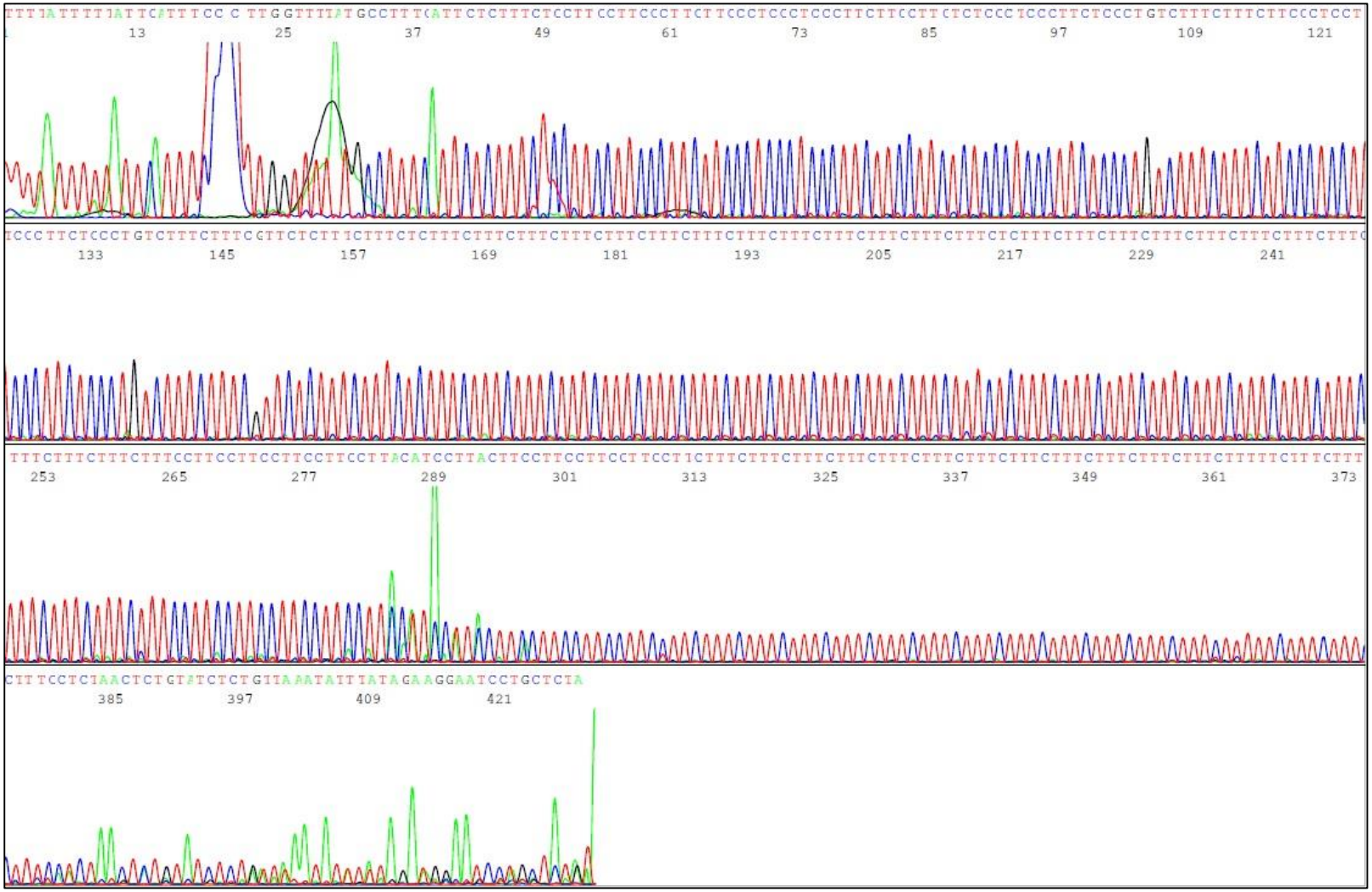

Figure A2.4: Electropherogram presenting a sequence of allele 49 at DYF403S1b locus with repeat motif [TTCT] and [TTCC]

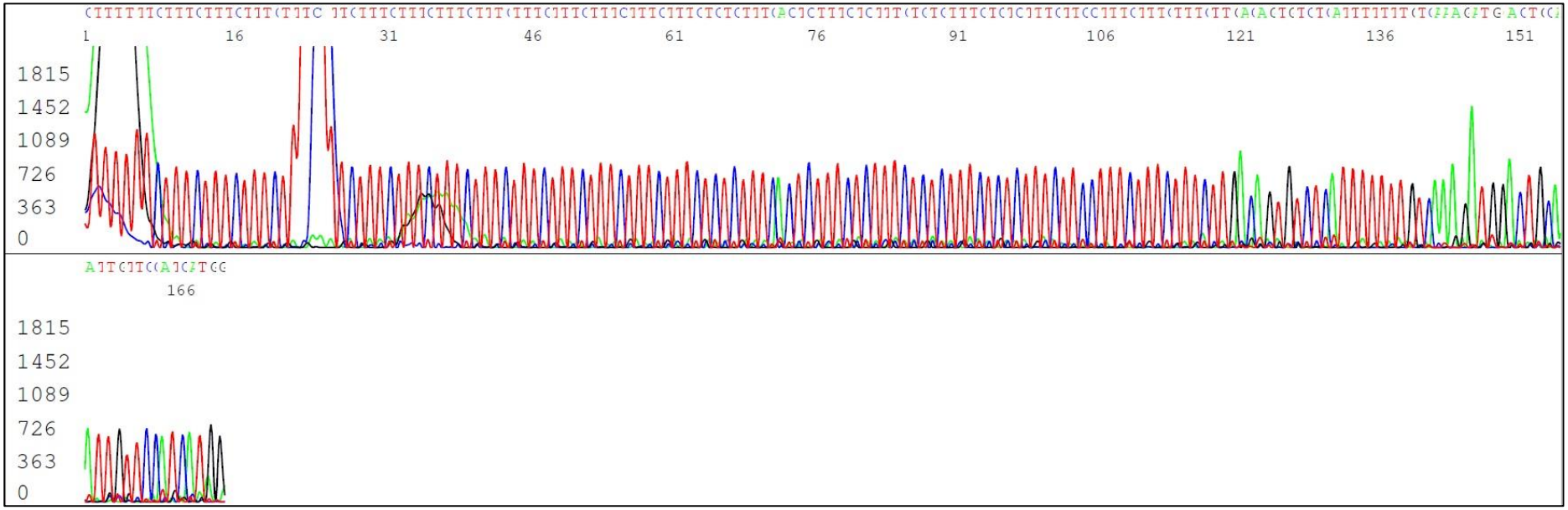

Figure A2.5: Electropherogram presenting a sequence of allele 15 at DYF404S1 locus with repeat motif [TTTC]

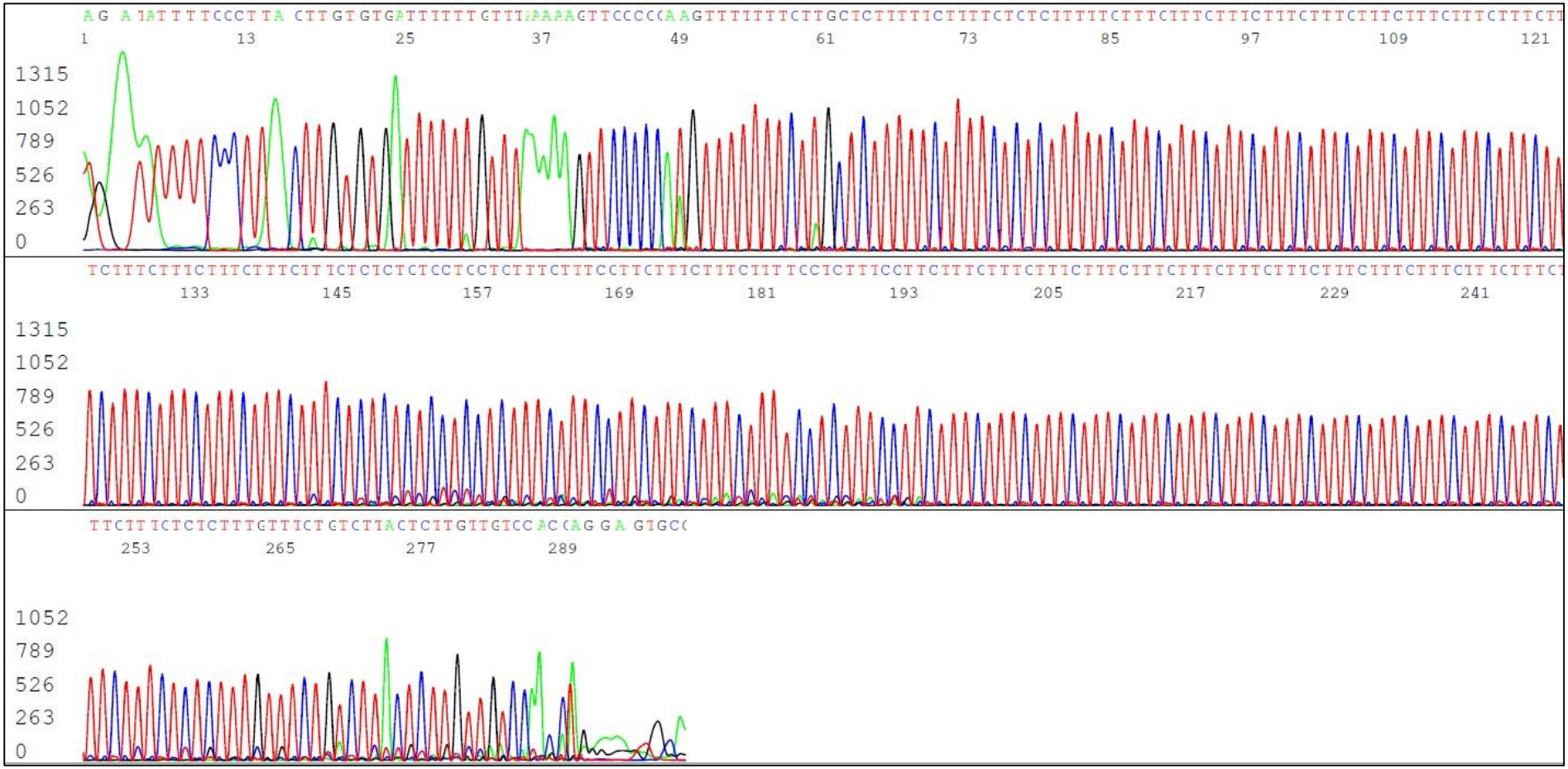

Figure A2.6: Electropherogram presenting a sequence of allele 31 at DYS449 locus with repeat motif [TTTC]

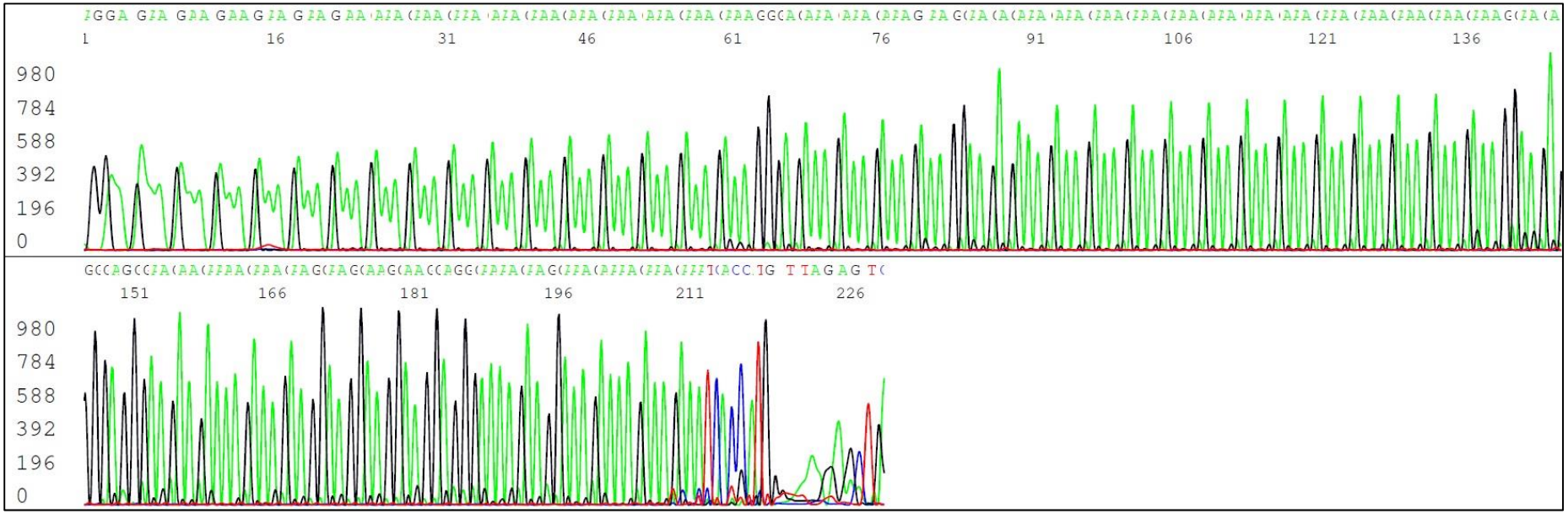

Figure A2.7: Electropherogram presenting a sequence of allele 29 at DYS518 locus with repeat motif [AAAG]

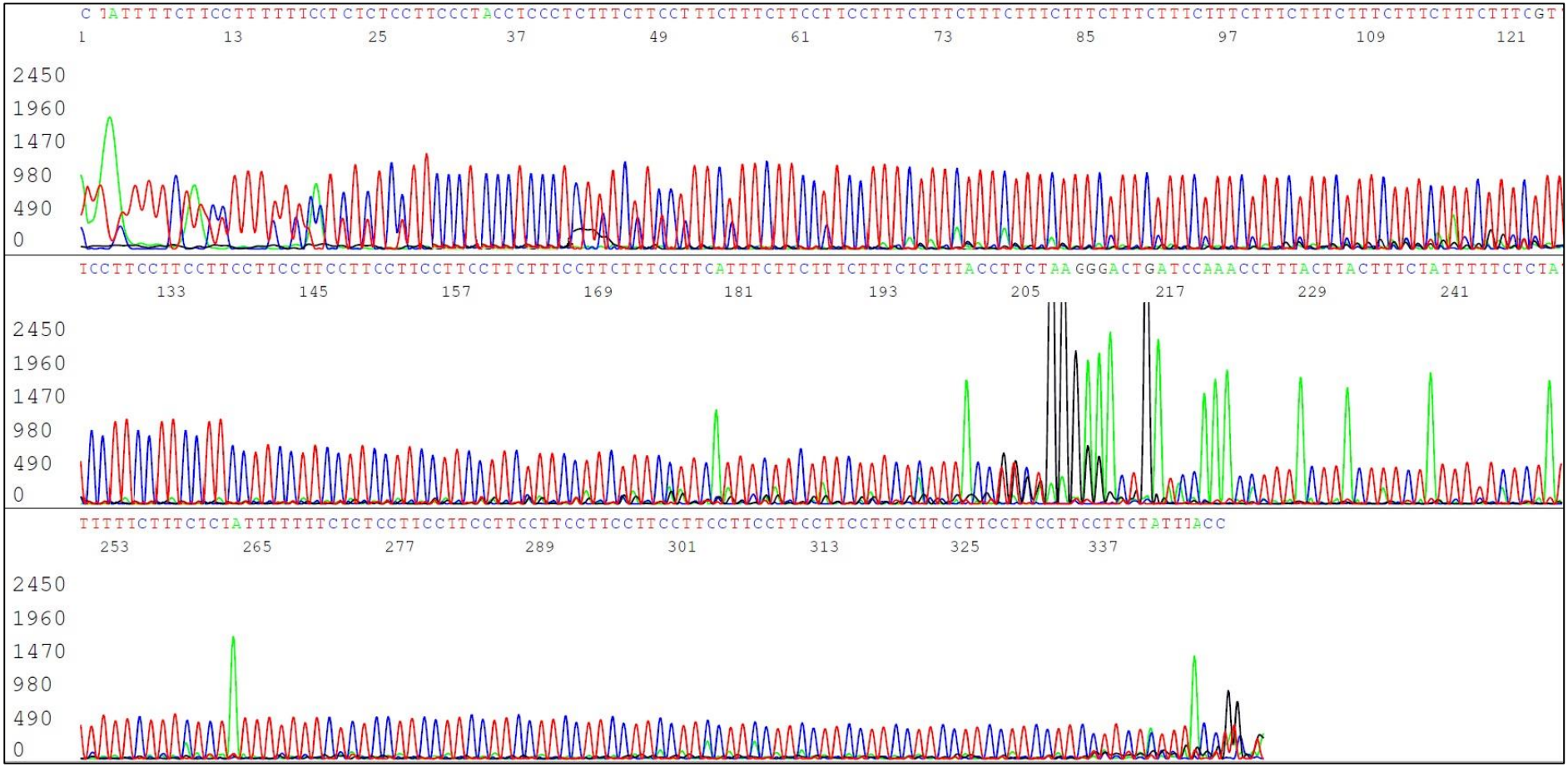

Figure A2.8: Electropherogram presenting a sequence of alleles a:16 and b:40 at DYS526ab locus with repeat motif [CTTT] and [CCTT].

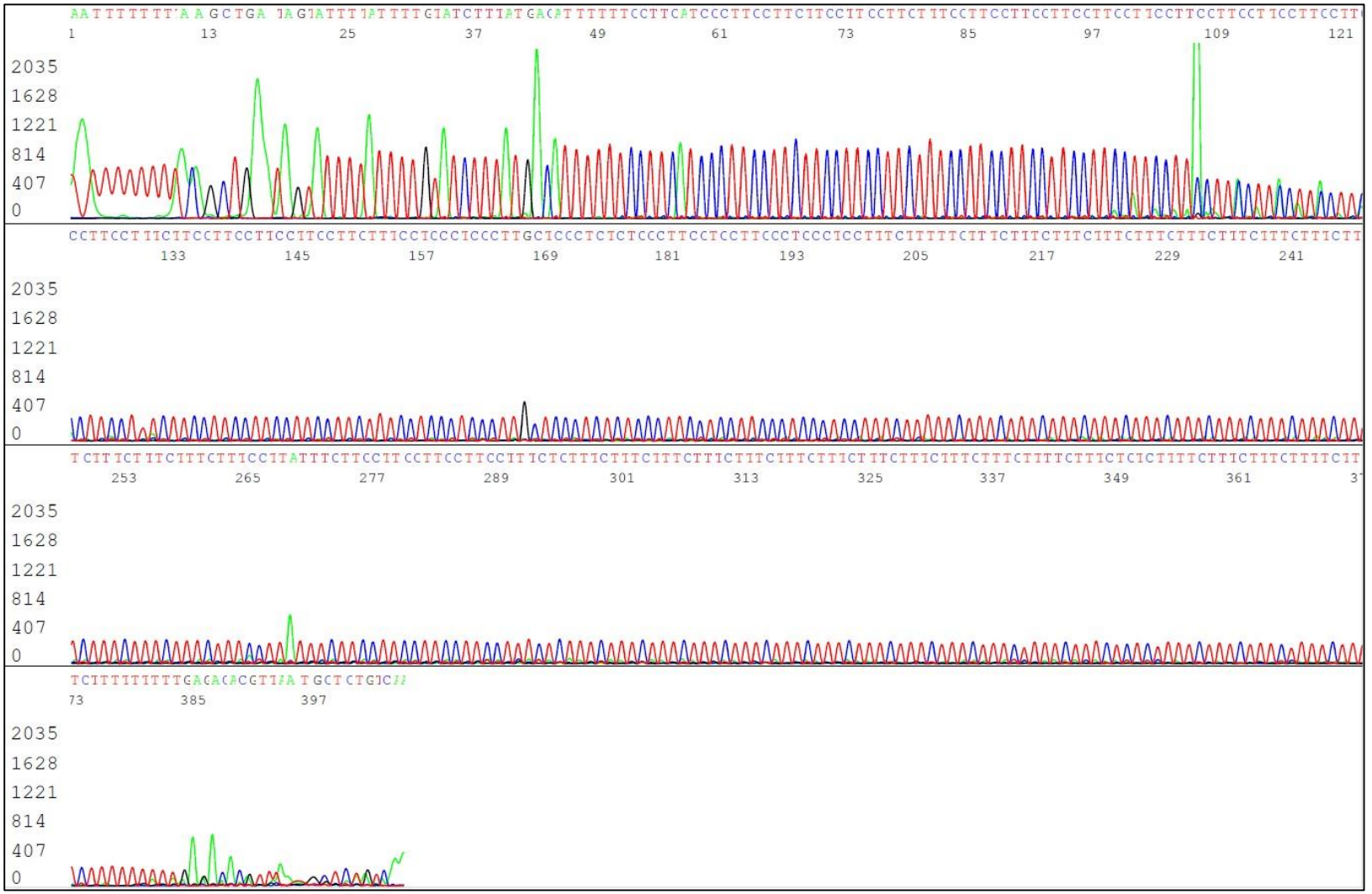

Figure A2.9: Electropherogram presenting a sequence of allele 43 at DYS547 locus with repeat motif [CCTT], [CTTC] and [TTTC]

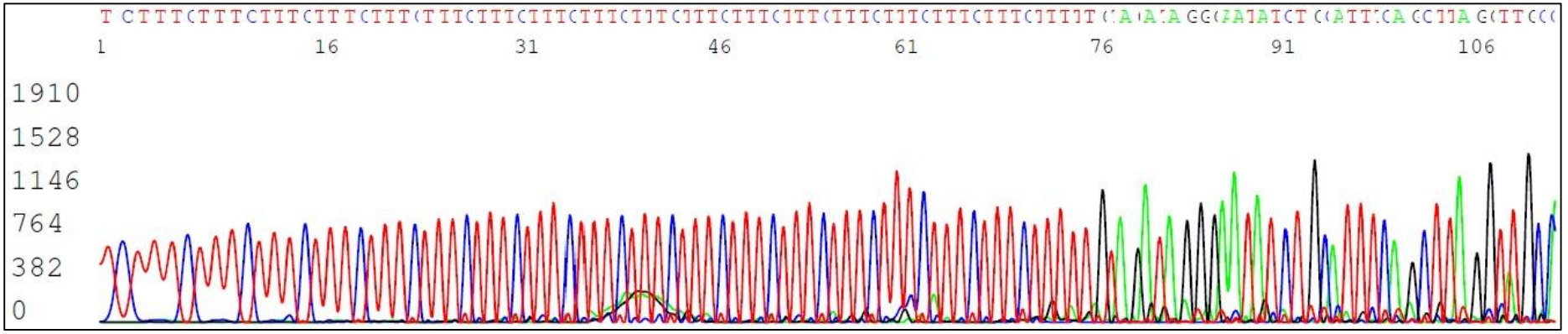

Figure A2.10: Electropherogram presenting a sequence of allele 17 at DYS570 locus with repeat motif [TTTC].

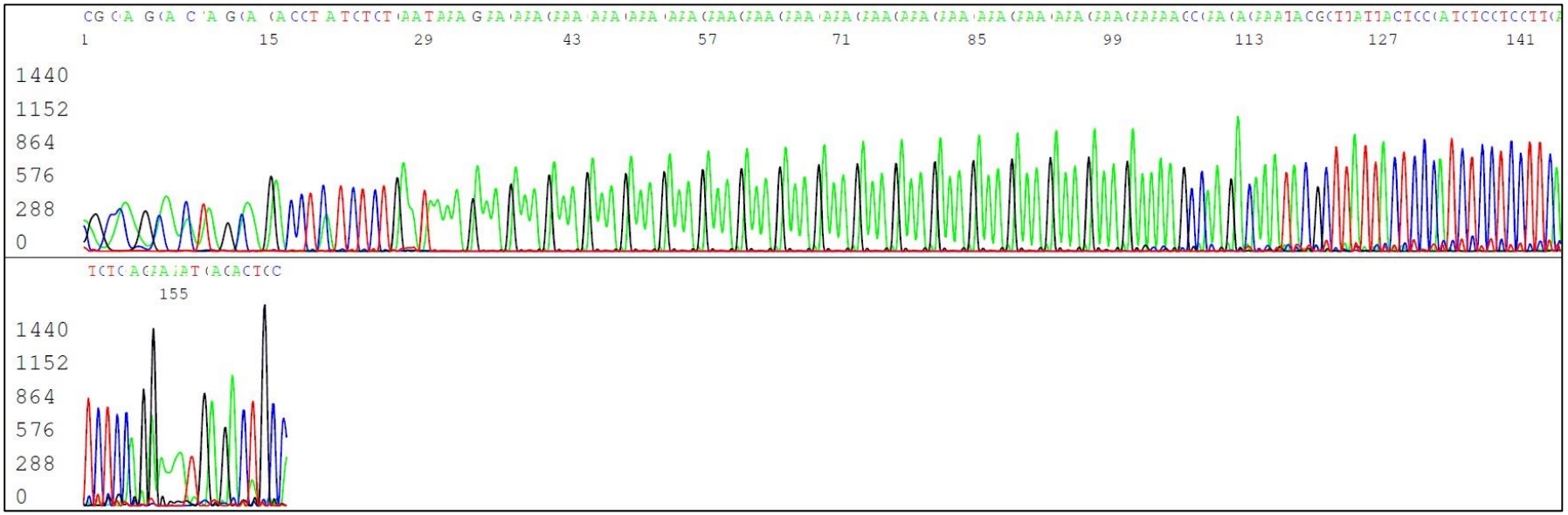

Figure A2.11: Electropherogram presenting a sequence of allele 18 at DYS576 locus with repeat motif [AAAG].

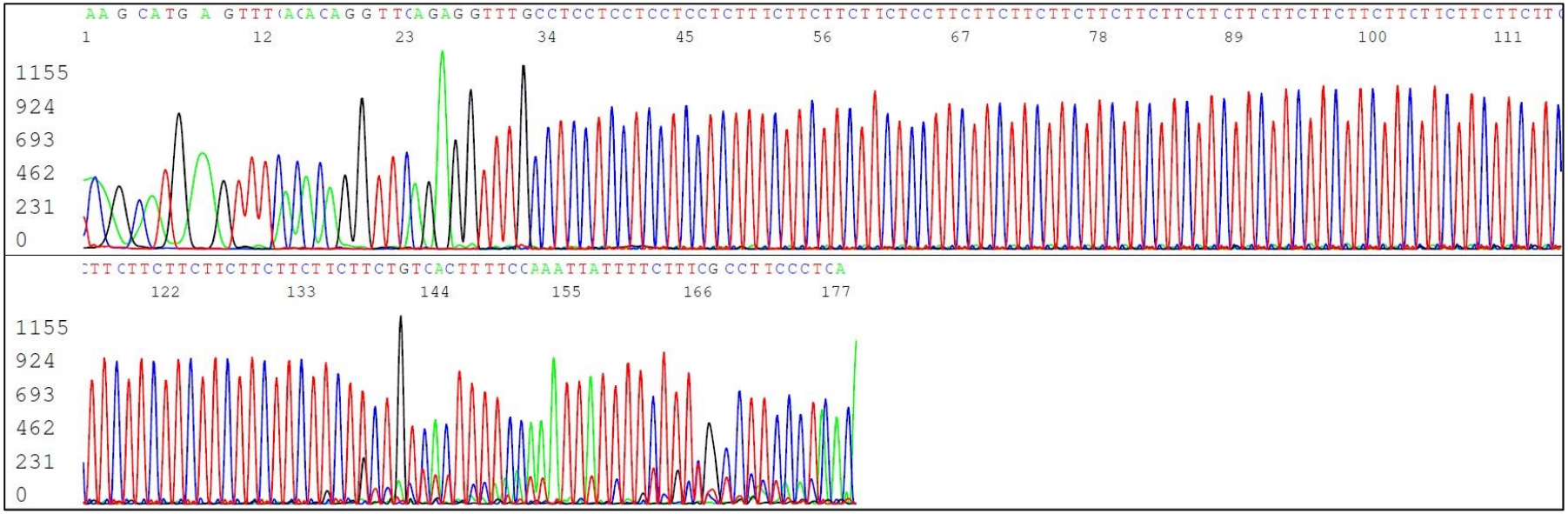

Figure A2.12: Electropherogram presenting a sequence of allele 25 at DYS612 locus with repeat motif [TCT].

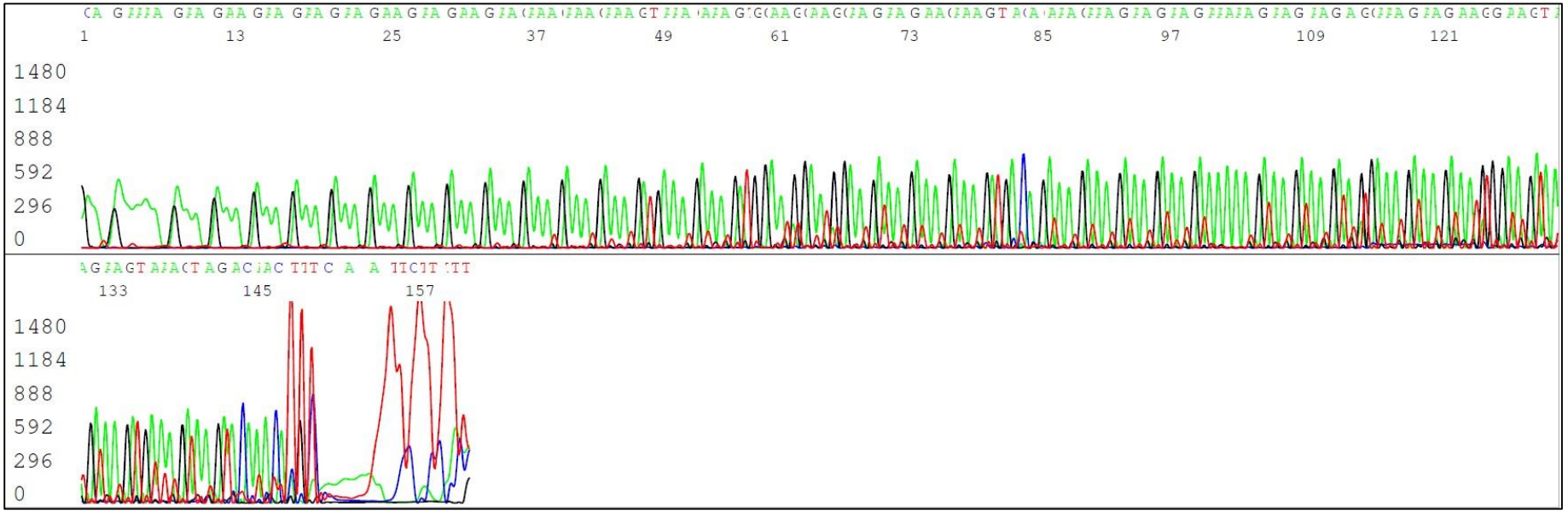

Figure A2.13: Electropherogram presenting a sequence of allele 27 at DYS626 locus with repeat motif [GAAA].

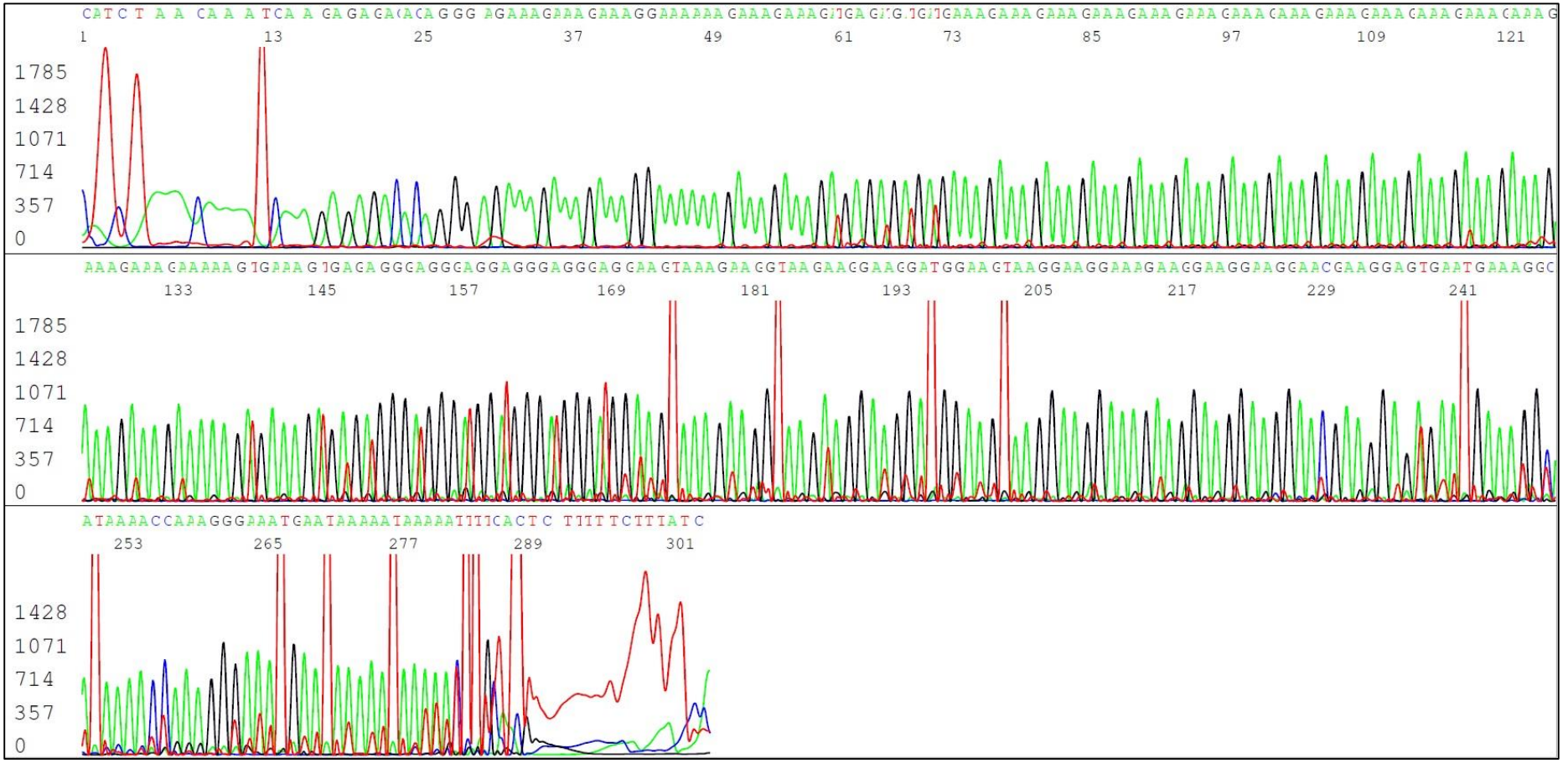

Figure A2.14: Electropherogram presenting a sequence of allele 18 at DYS627 locus with repeat motif [AAAG].

Appendix 3

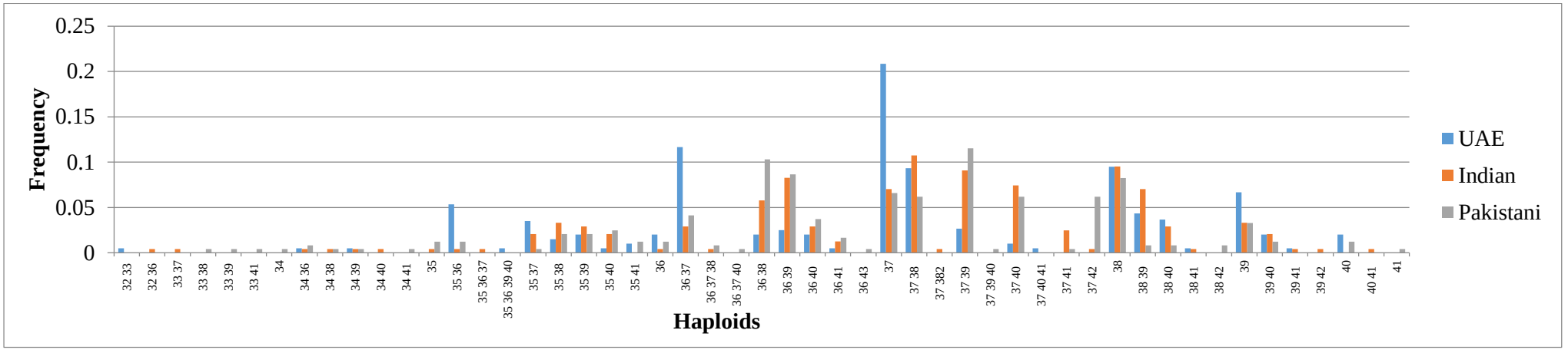

Figure A3.1: DYF387S1 haploids frequency for UAE Arabs, Indian and Pakistani populations.

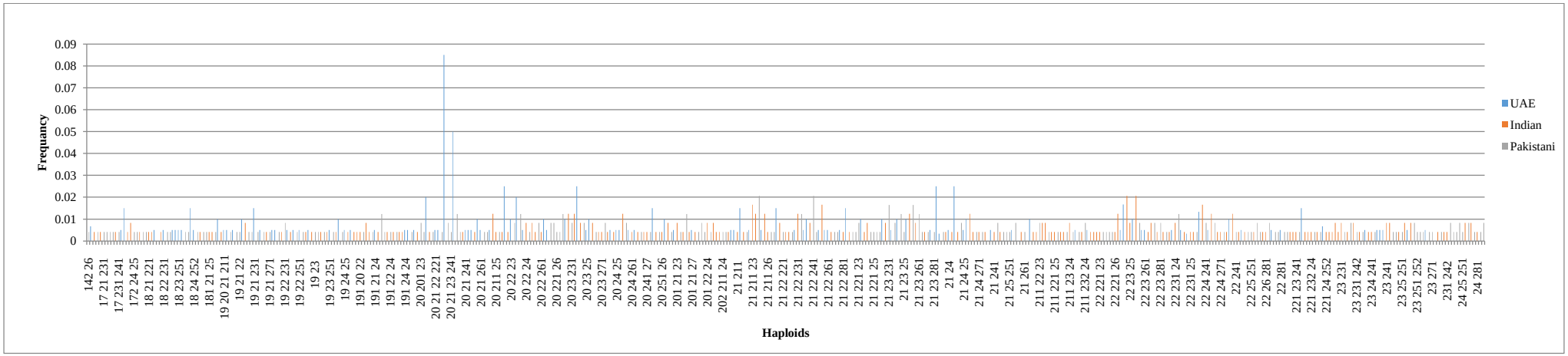

Figure A3.2: DYF399S1 haploids frequency for UAE Arabs, Indian and Pakistani populations.

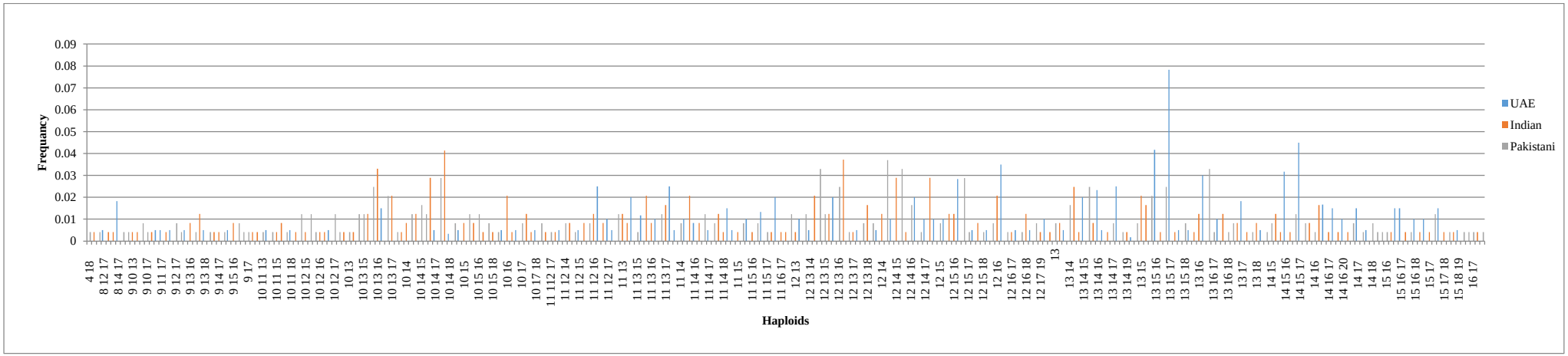

Figure A3.3: DYF403S1a haploids frequency for UAE Arabs, Indian and Pakistani populations.

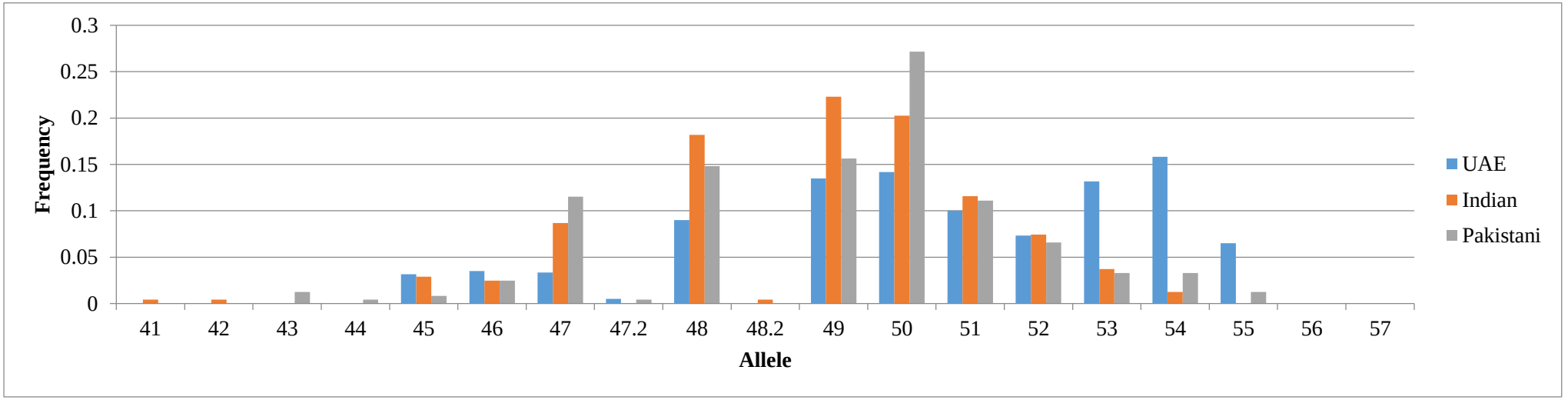

Figure A3.4: DYF403S1b alleles frequency for UAE Arabs, Indian and Pakistani populations.

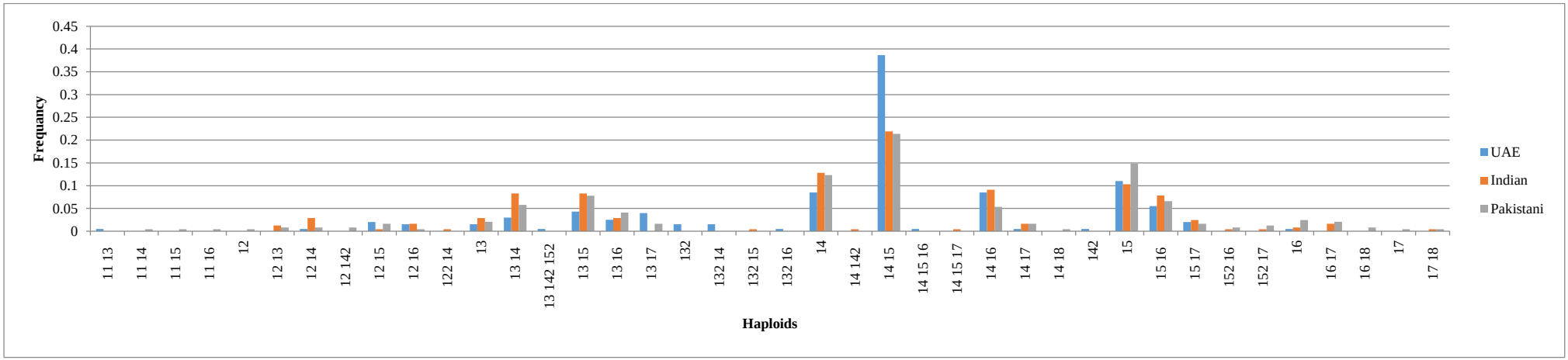

Figure A3.5: DYF404S1 Haploids frequency for UAE Arabs, Indian and Pakistani populations.

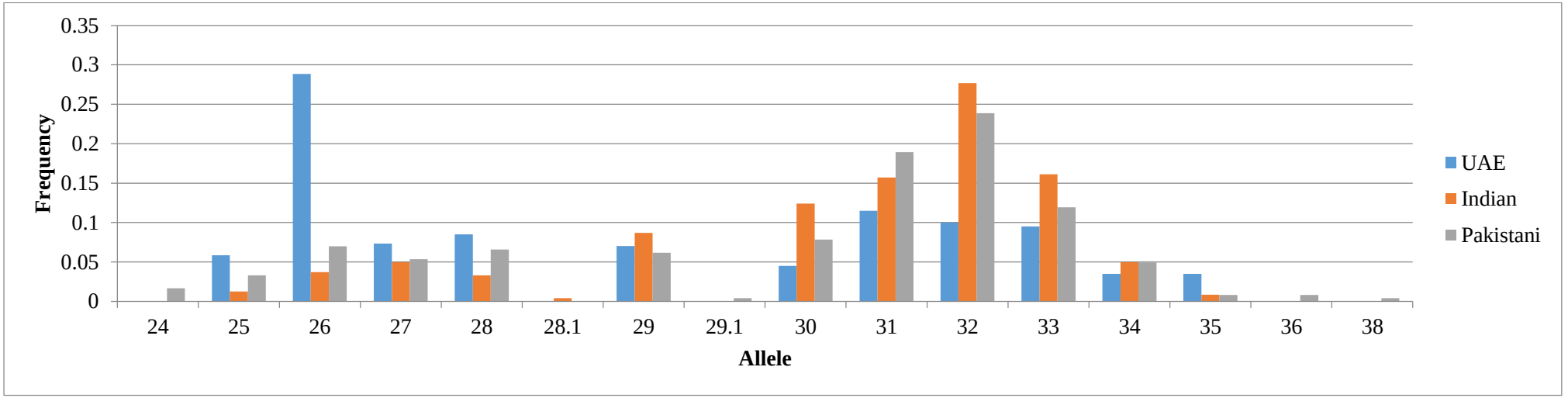

Figure A3.6: DYS449 alleles frequency for UAE Arabs, Indian and Pakistani populations.

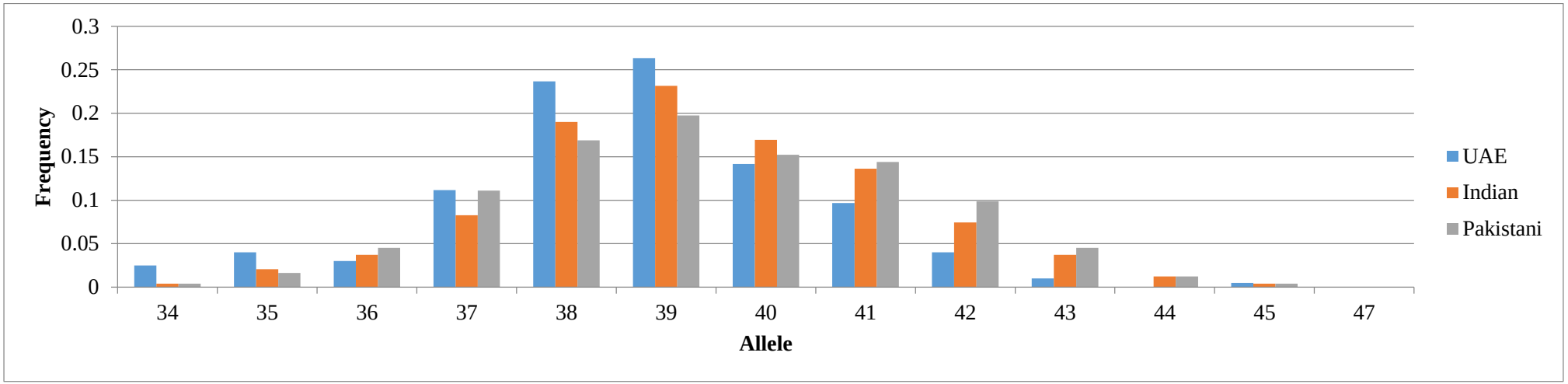

Figure A3.7: DYS518 alleles frequency for UAE Arabs, Indian and Pakistani populations.

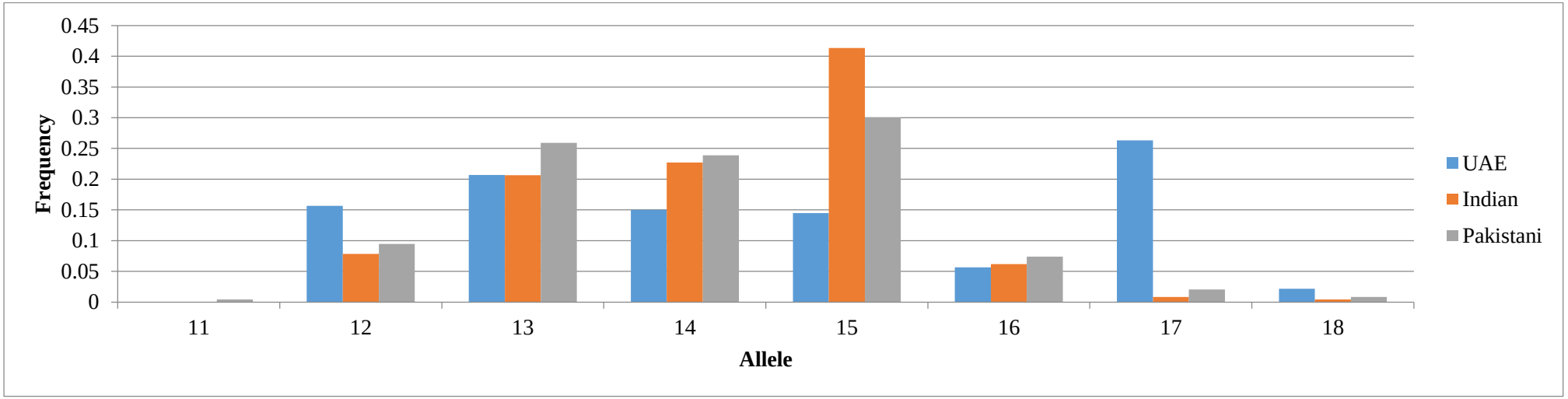

Figure A3.8: DYS526a alleles frequency for UAE Arabs, Indian and Pakistani populations.

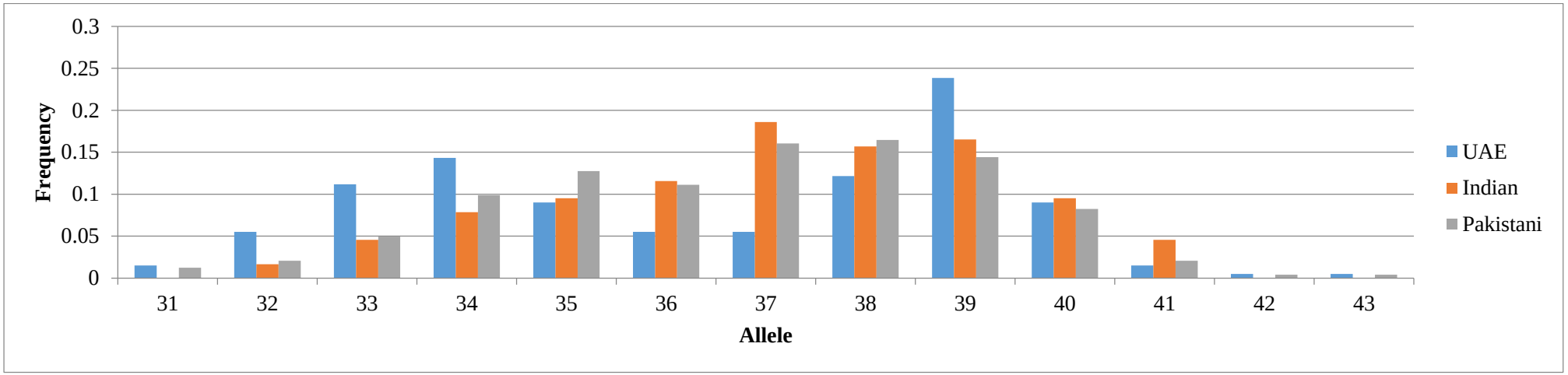

Figure A3.9: DYS526b alleles frequency for UAE Arabs, Indian and Pakistani populations.

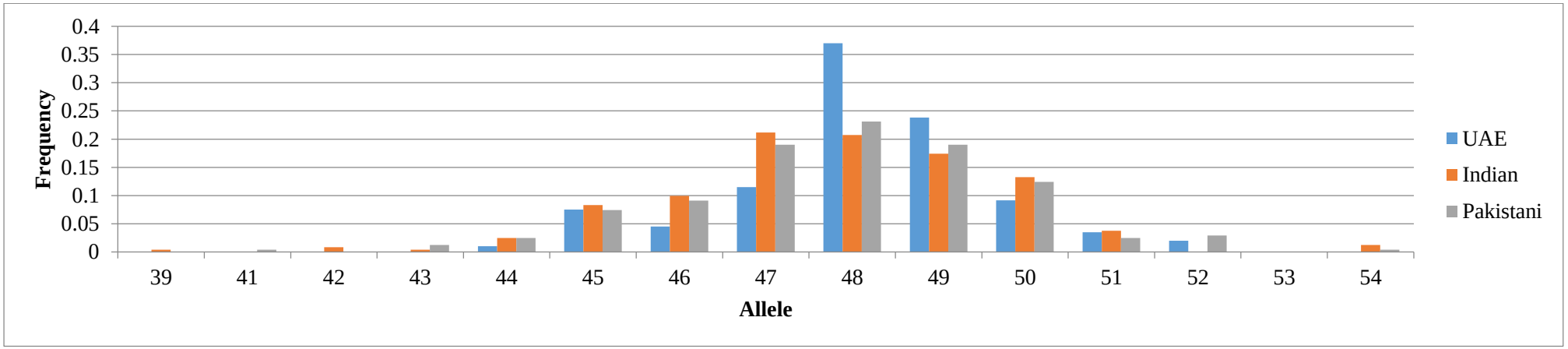

Figure A3.10: DYS547 alleles frequency for UAE Arabs, Indian and Pakistani populations.

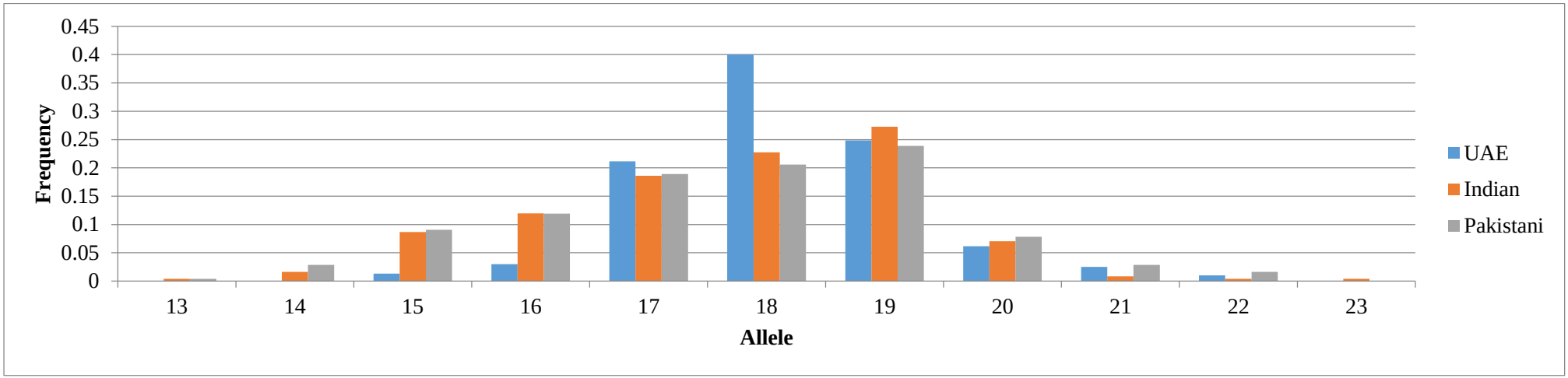

Figure A3.11: DYS570 alleles frequency for UAE Arabs, Indian and Pakistani populations.

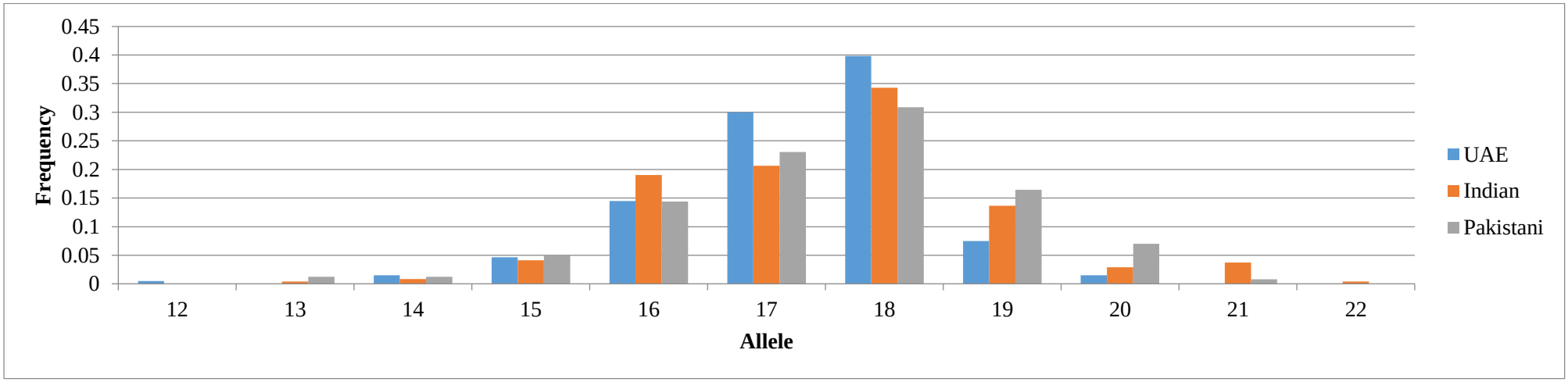

Figure A3.12: DYS576 alleles frequency for UAE Arabs, Indian and Pakistani populations.

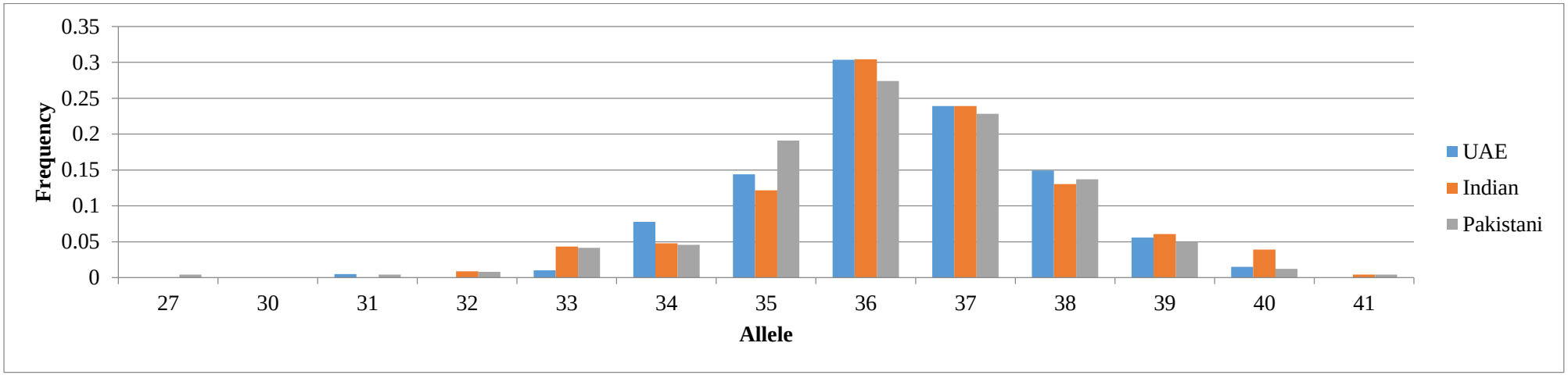

Figure A3.13: DYS612 alleles frequency for UAE Arabs, Indian and Pakistani populations.

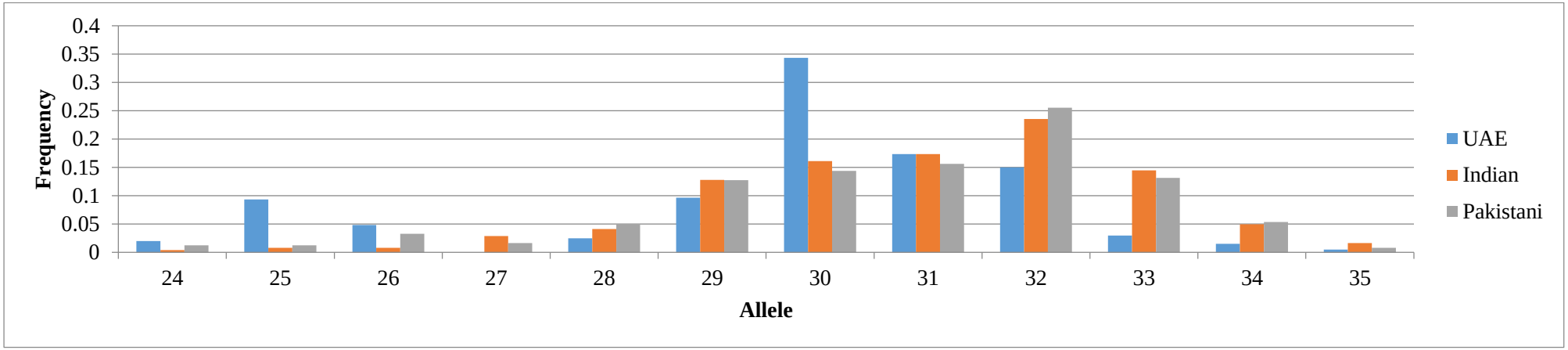

Figure A3.14: DYS626 alleles frequency for UAE Arabs, Indian and Pakistani populations.

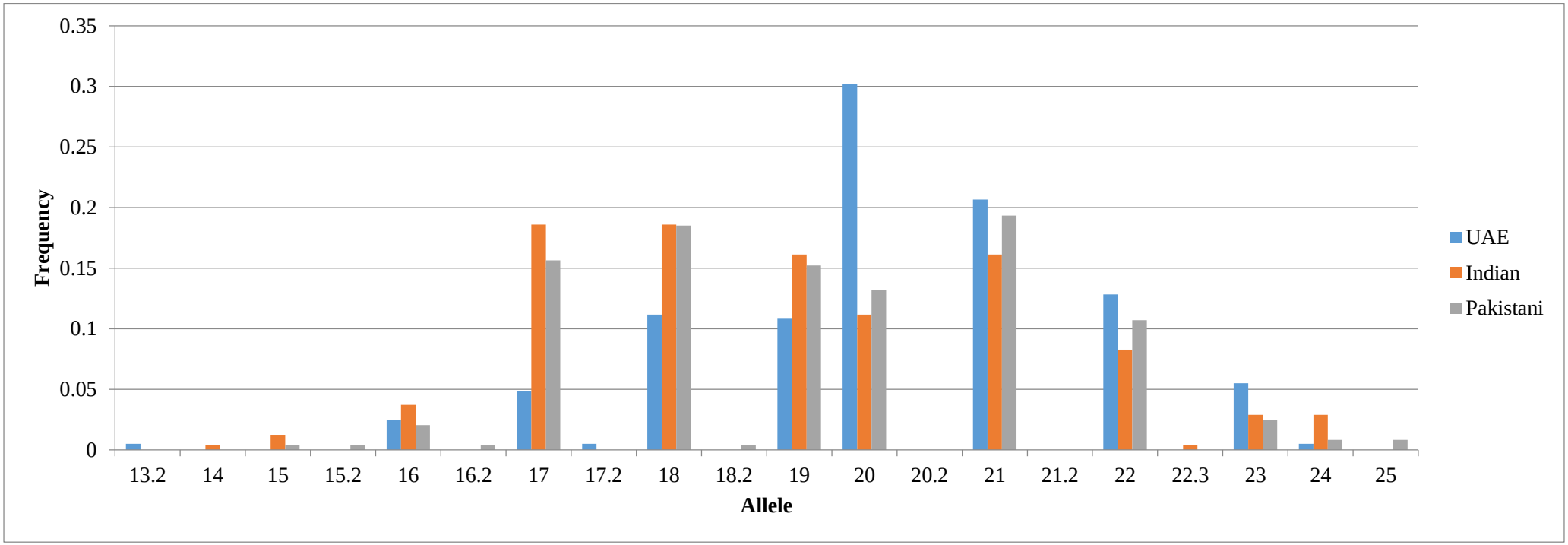

Figure A3.15: DYS627 alleles frequency for UAE Arabs, Indian and Pakistani populations.

Table A3.1: Database for 13 RM Y-STR haplotypes generated using RM-plex for 8 populations.

ID	Population	DYF387S1				DYF399S1					DYF403S1				DYF403S1b	DYF404S1			DYS449	DYS518	DYS526a	DYS526b	DYS547	DYS570	DYS576	DYS612	DYS626		DYS627
B001	UAE_B	38	39	-	-	18	24	25.1	-	-	8	14	17	-	53	14	15	-	31	38	13	33	48	18	18	36	25	-	18
B002	UAE_B	39	40	-	-	21	24	25	-	-	12	15	16	-	52	14	-	-	26	41	14	34	48	17	16	35	32	-	21
B003	UAE_B	37	38	-	-	22	23.1	24.1	-	-	11	14	19	-	46	12	15	-	35	34	12	34	45	19	16	37	31	-	18
B004	UAE_B	34	39	-	-	20	23	24.1	-	-	12	17	20	-	48	15	-	-	32	40	13	34	45	18	18	36	24	-	22
B005	UAE_B	36	37	-	-	20	21	23.1	-	-	14	16	-	-	55	14	15	-	26	37	17	40	48	18	18	38	30	-	21
B006	UAE_B	38	-	-	-	20	23	24.1	-	-	12	15	16	-	50	14	15	-	28	40	14	35	48	18	17	37	30	-	22
B007	UAE_B	39	-	-	-	20	23	23.1	-	-	12	14	18	-	51	14	16	-	34	38	13	32	48	18	18	33	26	-	18
B008	UAE_B	37	-	-	-	20	21.1	-	-	-	14	15	17	-	53	14	15	-	26	37	17	39	49	17	18	35	30	-	20
B009	UAE_B	37	38	-	-	19	21	24.1	-	-	15	16	17	-	54	14	-	-	25	39	16	39	49	17	18	36	30	-	23
B010	UAE_B	36	37	-	-	20	21	22.1	-	-	12	16	17	-	53	14	16	-	25	39	17	40	49	18	17	36	30	-	21
B011	UAE_B	37	-	-	-	20	21	24.1	-	-	14	16	17	-	54	14	15	-	26	38	17	39	49	17	18	35	30	-	20
B012	UAE_B	38	39	-	-	19	22	24	-	-	11	12	15	-	52	14	16	-	31	39	14	33	49	17	18	36	31	-	18
B013	UAE_B	39	-	-	-	20	23	25.1	-	-	13	17	-	-	50	15	16	-	31	41	13	33	50	18	17	37	24	-	18
B014	UAE_B	38	39	-	-	21	21.1	22	-	-	13	15	16	-	52	15	-	-	33	36	12	32	51	19	17	37	25	-	20
B015	UAE_B	36	39	-	-	22	23	24.1	-	-	12	15	16	-	50	14	16	-	32	41	15	38	49	18	18	34	32	-	16
B016	UAE_B	39	-	-	-	21	23	28	-	-	12	14	15	-	48	14	15	-	28	34	13	35	47	18	16	37	34	-	20
B017	UAE_B	35	39	-	-	20	21	22.1	-	-	11	12	16	-	51	14	15	-	28	39	13	35	48	18	18	35	30	-	20
B018	UAE_B	35	37	-	-	22	23	24.1	-	-	11	15	16	-	53	15	-	-	31	37	13	33	50	21	17	36	26	-	21
B019	UAE_B	39	41	-	-	21	25	26	-	-	12	15	16	-	49	14	-	-	30	41	14	34	48	17	18	36	31	-	20
B020	UAE_B	36	37	-	-	20	21	22.1	-	-	14	15	17	-	54	14	16	-	26	38	17	39	48	18	18	37	30	-	20
B021	UAE_B	38	41	-	-	20	23	25	-	-	13	17	-	-	50	14	-	-	32	40	14	33	49	17	16	38	31	-	19
B022	UAE_B	37	38	-	-	21	23.1	-	-	-	12	14	16	-	52	14	15	-	35	38	12	34	45	17	18	34	33	-	20
B023	UAE_B	38	-	-	-	22	23	25.1	-	-	15	16	17	-	53	12	16	-	30	42	15	37	49	17	17	36	28	-	21
B024	UAE_B	37	-	-	-	20	21	22.1	-	-	13	16	17	-	54	14	15	-	26	38	17	39	48	19	17	39	30	-	20
B025	UAE_B	36	37	-	-	20	21	22.1	-	-	14	15	17	-	53	14	15	-	26	39	17	39	48	18	18	37	30	-	20
B026	UAE_B	37	-	-	-	20	21	26.1	-	-	13	14	17	-	52	14	15	-	25	40	18	41	50	19	18	35	30	-	23
B027	UAE_B	35	36	-	-	20	21	22	23	-	11	13	-	-	47	13	15	-	30	38	12	35	48	16	16	40	31	-	19
B028	UAE_B	34	36	-	-	18	23	25.1	-	-	15	16	-	-	50	15	16	-	33	40	12	33	49	22	18	40	25	-	20
B029	UAE_B	36	40	-	-	20	21	23.1	-	-	13	15	17	-	53	14	15	-	26	37	18	40	48	18	17	36	30	-	20
B030	UAE_B	38	-	-	-	21	22	22.1	-	-	12	16	-	-	52	15	17	-	32	36	12	32	51	19	17	36	25	-	21
B031	UAE_B	37	-	-	-	20	21	23.1	-	-	13	15	17	-	54	14	15	-	26	39	17	39	48	18	18	36	31	-	20
B032	UAE_B	37	-	-	-	20	21	23.1	-	-	13	15	17	-	54	14	15	-	26	39	17	39	49	18	18	36	31	-	20

ID	Population	DYF387S1				DYF399S1					DYF403S1				DYF403S1b	DYF404S1			DYS449	DYS518	DYS526a	DYS526b	DYS547	DYS570	DYS576	DYS612	DYS626		DYS627
B033	UAE_B	35	37	-	-	20	21	23.1	-	-	14	15	17	-	55	14	16	-	27	38	17	39	48	18	19	39	30	-	20
B034	UAE_B	37	38	-	-	21	24	24.1	-	-	12	15	16	-	49	13	17	-	30	39	14	38	48	19	17	38	31	-	21
B035	UAE_B	40	-	-	-	21	-	-	-	-	14	16	20	-	49	14.2	-	-	28	42	13	36	45	19	17	37	28	-	19
B036	UAE_B	35	36	-	-	19.1	24	24	-	-	11	13	15	-	51	15	-	-	28	37	14	35	50	19	15	31	29	-	22
B037	UAE_B	36	38	-	-	19	21	23.1	-	-	13	14	16	-	51	14	-	-	26	38	15	37	49	18	18	34	30	-	21
B038	UAE_B	36	37	-	-	20	21	23.1	-	-	13	16	-	-	53	12	14	-	27	38	16	37	51	18	20	37	31	-	22
B039	UAE_B	37	38	-	-	21	22.1	23.1	-	-	12	13	-	-	49	15	16	-	26	41	13	34	45	19	17	39	31	-	21
B040	UAE_B	36	40	-	-	20	23	24.1	-	-	13	15	19	-	49	15	17	-	33	39	13	35	47	18	19	34	29	-	21
B041	UAE_B	35	36	39	40	19	22	25.1	-	-	12	14	-	-	48	14	15	-	28	40	14	42	47	15	17	35	32	-	23
B042	UAE_B	37	-	-	-	21	22.1	-	-	-	14	15	16	-	53	14	15	-	26	38	17	39	50	18	19	35	30	-	20
B043	UAE_B	38	-	-	-	21.1	22	-	-	-	14	15	16	-	50	15	16	-	31	37	12	32	50	18	17	35	25	-	20
B044	UAE_B	38	-	-	-	21.1	22	-	-	-	13	15	16	-	52	14	15	-	33	36	12	31	50	19	17	38	25	-	21
B045	UAE_B	37	38	-	-	21	22.1	23.1	-	-	12	13	-	-	49	14	16	-	26	41	13	34	48	21	17	38	31	-	20
B046	UAE_B	37	-	-	-	20	21	23.1	-	-	14	16	17	-	55	14	15	-	26	38	17	39	48	18	17	39	30	-	20
B047	UAE_B	37	38	-	-	20	21	23.1	-	-	13	15	17	-	53	15	-	-	26	38	17	40	49	17	18	33	30	-	20
B048	UAE_B	35	38	-	-	18	21	26.1	-	-	13	14	16	-	53	14	16	-	26	37	16	38	49	18	18	34	30	-	20
B049	UAE_B	35	37	-	-	22	23.1	24	-	-	11	15	17	-	51	14	15	-	31	38	12	33	49	19	17	37	26	-	22
B050	UAE_B	39	40	-	-	22	24	25	-	-	12	16	-	-	52	14	15	-	27	41	14	35	49	17	16	36	32	-	21
B051	UAE_B	35	41	-	-	20	22	23	-	-	9	13	16	-	52	14	15	-	29	39	12	31	48	19	16	34	28	-	20
B052	UAE_B	37	-	-	-	21	22	25.1	-	-	13	15	16	-	49	14	-	-	33	40	14	36	44	18	17	35	31	-	20
B053	UAE_B	37	38	-	-	21	24	24.1	-	-	11	12	16	-	49	13	15	-	33	40	15	38	47	19	17	37	32	-	18
B054	UAE_B	37	-	-	-	20	22	23.1	-	-	11	15	17	-	54	14	16	-	26	39	16	38	48	19	18	36	30	-	22
B055	UAE_B	38	-	-	-	20	24	25.1	-	-	11	14	18	-	49	14	15	-	31	42	15	38	46	20	17	37	34	-	18
B056	UAE_B	35	36	-	-	22	23	24.1	-	-	9	12	16	-	50	15	-	-	30	40	14	38	48	18	16	37	30	-	24
B057	UAE_B	38	40	-	-	20	26	-	-	-	11	13	14	-	48	13	17	-	33	39	13	34	47	19	17	38	32	-	22
B058	UAE_B	37	38	-	-	22.1	23	25.1	-	-	11	13	17	-	45	12	15	-	35	35	12	34	45	17	16	36	31	-	15.2
B059	UAE_B	36	37	-	-	21	24	24.1	-	-	12	15	18	-	51	15	-	-	31	39	12	33	51	19	16	38	26	-	22
B060	UAE_B	37	-	-	-	20	22	22.1	-	-	13	16	-	-	54	14	15	-	26	38	17	39	48	18	18	35	30	-	20
B061	UAE_B	36	-	-	-	20	22.1	-	-	-	14	15	17	-	53	14	15	-	26	39	17	39	48	18	18	38	30	-	20
B062	UAE_B	36	39	-	-	22	23	24.1	-	-	12	15	16	-	50	14	16	-	32	41	15	38	49	18	18	34	32	-	16
B063	UAE_B	40	-	-	-	22	23	23.1	24	24.1	12	17	20	-	49	14	15	-	34	39	13	34	49	18	17	38	24	-	21
B064	UAE_B	38	-	-	-	17.2	21	21.1	-	-	15	17	-	-	46	13.2	-	-	29	37	13	35	46	17	18	36	32	-	23
B065	UAE_B	38	-	-	-	17.2	21	21.1	-	-	15	17	-	-	47	13.2	-	-	29	37	14	35	46	17	18	36	31	-	23
B066	UAE_B	36	37	-	-	21	22	24.1	-	-	11	14	17	-	50	13	14	-	32	43	14	37	47	20	17	40	31	-	17

ID	Population	DYF387S1				DYF399S1					DYF403S1				DYF403S1b	DYF404S1			DYS449	DYS518	DYS526a	DYS526b	DYS547	DYS570	DYS576	DYS612	DYS626		DYS627
B067	UAE_B	36	37	-	-	21	22.1	-	-	-	14	17	-	-	54	15	-	-	26	39	17	39	48	19	18	38	30	-	20
B068	UAE_B	36	37	-	-	20	21	21.1	-	-	13	14	17	-	55	13	15	-	26	37	17	39	48	19	18	35	30	-	20
B069	UAE_B	35	36	-	-	23	24	26.1	-	-	12	14	17	-	47	15	16	-	31	38	15	38	46	17	16	38	29	-	22
B070	UAE_B	38	-	-	-	23	25.1	-	-	-	15	16	19	-	54	15	-	-	27	39	15	37	49	19	18	36	29	-	19
B071	UAE_B	37	-	-	-	21	23.1	-	-	-	13	14	16	-	53	13	14	-	27	37	16	37	49	18	18	37	31	-	22
B072	UAE_B	37	-	-	-	20	21	22.1	-	-	15	16	17	-	54	14	15	-	26	38	17	39	48	18	17	38	30	-	20
B073	UAE_B	35	36	-	-	18	23	24.1	-	-	14	16	-	-	50	15	16	-	32	41	12	34	49	21	18	39	25	-	21
B074	UAE_B	39	-	-	-	20	24	27	-	-	12	13	15	-	49	13	15	-	28	35	13	35	48	18	18	37	33	-	21
B075	UAE_B	37	-	-	-	21	23.1	-	-	-	12	14	16	-	53	14	15	-	25	38	17	36	49	18	18	32	30	-	21
B076	UAE_B	39	-	-	-	20	23	23.1	-	-	11	14	18	-	51	14	17	-	33	37	13	32	49	18	19	36	26	-	19
B077	UAE_B	36	38	-	-	19	21	23.1	-	-	12	14	16	-	54	14	-	-	25	39	16	37	50	17	19	34	32	-	17
B078	UAE_B	37	38	-	-	20	21	22.1	-	-	13	15	17	-	53	14	15	-	26	38	17	39	49	19	17	37	30	-	20
B079	UAE_B	36	37	-	-	20	22.1	-	-	-	14	15	16	-	55	15	-	-	26	38	17	39	48	18	18	37	30	-	20
B080	UAE_B	39	-	-	-	21	23	24	-	-	12	13	15	-	48	13	16	-	30	38	13	35	48	15	16	38	31	-	21
B081	UAE_B	35	38	-	-	20	22	23	-	-	11	12	17	-	51	14	15	-	27	39	14	34	49	18	14	35	30	-	19
B082	UAE_B	38	39	-	-	18	24	25.1	-	-	8	14	17	-	52	14	15	-	31	38	13	33	48	18	18	36	25	-	18
B083	UAE_B	38	40	-	-	18	24	25.2	-	-	8	14	17	-	53	14	15	-	31	38	13	34	48	18	18	36	25	-	18
B084	UAE_B	38	39	-	-	18	24	25.1	-	-	8	14	17	-	54	14	16	-	31	38	13	33	48	18	18	36	25	-	18
B085	UAE_B	40	-	-	-	20	24	25	-	-	12	16	-	-	51	14	15	-	26	41	14	34	48	17	16	35	32	-	22
B086	UAE_B	40	-	-	-	20	25	-	-	-	12	16	-	-	48	14	15	-	26	41	14	34	48	17	16	35	32	-	21
B087	UAE_B	38	39	-	-	21	23	24.1	-	-	12	16	-	-	49	13	17	-	31	39	14	38	48	20	17	37	31	-	21
B088	UAE_B	38	40	-	-	20	21	26	-	-	11	14	-	-	48	13	17	-	32	39	13	36	47	20	16	37	31	-	21
B089	UAE_B	36	39	-	-	21	23	23.1	-	-	13	18	19	-	47.2	14	16	-	29	41	13	36	47	20	17	37	24	-	18
B090	UAE_B	39	-	-	-	21	22	28	-	-	12	15	-	-	48	13	15	-	28	34	13	32	47	18	17	38	32	-	21
B091	UAE_B	36	-	-	-	21	24.1	-	-	-	14	16	20	-	45	16	-	-	35	36	12	38	50	16	14	33	32	-	19
B092	UAE_B	36	37	-	-	20	21	22.1	-	-	13	14	17	-	54	14	15	-	25	38	18	38	49	18	18	36	29	-	22
B093	UAE_B	39	-	-	-	19.1	21	24	-	-	10	12	17	-	50	15	-	-	33	39	15	40	48	19	20	38	32	-	17
B094	UAE_B	35	39	-	-	20	21	25	-	-	11	12	14	-	49	14	15	-	33	37	13	34	49	19	18	37	35	-	20
B095	UAE_B	38	39	-	-	21	21.1	22	-	-	13	15	16	-	52	15	16	-	33	36	12	33	52	19	17	36	25	-	20
B096	UAE_B	37	39	-	-	15	24	27	-	-	13	14	15	-	49	14	15	-	28	40	14	34	50	19	14	34	29	-	18
B097	UAE_B	36	39	-	-	19	24	25	-	-	11	15	-	-	50	14	16	-	27	37	15	34	48	20	15	36	30	-	20
B098	UAE_B	38	40	-	-	18	23	26.1	-	-	8	12	17	-	50	14	15	-	31	38	13	34	49	18	17	37	25	-	18
B099	UAE_B	35	36	-	-	21	23	24.1	-	-	9	11	17	-	49	14	-	-	29	39	14	39	52	17	17	36	31	-	22
B100	UAE_B	37	-	-	-	22.1	23	25.1	-	-	11	12	17	-	45	12	16	-	35	35	12	38	45	17	16	36	31	-	21

ID	Population	DYF387S1				DYF399S1					DYF403S1				DYF403S1b	DYF404S1			DYS449	DYS518	DYS526a	DYS526b	DYS547	DYS570	DYS576	DYS612	DYS626		DYS627
B101	UAE_B	37	38	-	-	21	24	-	-	-	14	17	19	-	49	14	15	-	34	39	13	35	45	17	18	36	32	-	23
B102	UAE_B	38	-	-	-	21	23	23.1	-	-	12	16	-	-	48	15	16	-	33	37	12	32	52	19	17	36	25	-	20
B103	UAE_B	32	33	-	-	18	23	24	-	-	13	14	-	-	48	13	-	-	32	38	15	39	45	18	19	38	31	-	19
B104	UAE_B	36	37	-	-	21	24	24.1	-	-	11	13	17	-	50	13	14	-	34	45	15	40	47	20	15	39	31	-	17
B105	UAE_B	37	40	-	-	19	24	26	-	-	13	14	15	-	48	14	15	-	29	41	14	33	49	18	15	36	30	-	18
B106	UAE_B	36	41	-	-	20	24	24.1	-	-	12	14	16	-	50	13	14	-	27	38	14	39	48	16	16	35	29	-	21
B107	UAE_B	35	36	-	-	22	26.1	28	-	-	11	13	16	-	49	13	17	-	28	38	14	36	46	17	17	37	31	-	23
B108	UAE_B	36	40	-	-	19	22	24.1	-	-	11	12	16	-	49	14	-	-	30	40	13	33	46	17	15	35	30	-	19
B109	UAE_B	37	40	-	-	20	20.1	23	-	-	11	13	17	-	49	14	-	-	31	42	15	36	48	19	18	38	33	-	17
B110	UAE_B	37	39	-	-	19	23.1	25	-	-	10	13	17	-	49	14	16	-	33	41	15	34	48	20	19	36	32	-	17
B111	UAE_B	37	39	-	-	21	24	25.1	-	-	15	16	-	-	50	14	15	-	29	38	14	33	49	17	12	34	25	-	19
B112	UAE_B	37	-	-	-	20	24	26	-	-	12	13	15	-	48	13	15	-	28	35	13	36	48	17	18	37	33	-	21
B113	UAE_B	37	40	41	-	19	20	21	21.1	-	13	14	15	-	47	13	14.2	15.2	28	42	13	39	45	20	18	39	28	-	18
B114	UAE_B	37	-	-	-	20	21	22.1	-	-	13	15	-	-	53	14	15	-	27	38	17	41	49	20	18	36	30	-	21
B115	UAE_B	37	38	-	-	21	21.1	25	-	-	10	11	18	-	48	13	14	-	29	38	12	39	50	17	17	35	29	-	20
B116	UAE_B	37	-	-	-	20	21	22.1	-	-	13	17	-	-	53	14	15	-	25	38	17	41	49	18	18	36	30	-	21
B117	UAE_B	37	38	-	-	20	23	24.1	-	-	13	15	17	-	51	14	15	-	31	40	15	38	46	18	16	35	31	-	17
B118	UAE_B	37	-	-	-	20	23	24.1	-	-	11	12	16	-	50	15	16	-	33	40	15	39	49	19	18	36	33	-	19
B119	UAE_B	38	39	-	-	21	23.1	-	-	-	11	14	16	-	47	11	13	-	29	37	14	35	47	17	17	36	30	-	20
B120	UAE_B	38	-	-	-	21	21.1	22	-	-	13	15	16	-	52	15	16	-	32	37	12	32	49	19	16	36	25	-	21
B121	UAE_B	36	37	-	-	21	24	25.1	-	-	11	13	15	-	49	14	16	-	29	39	14	38	47	17	16	35	29	-	21
B122	UAE_B	37	-	-	-	20	21	22	22.1	-	14	15	16	-	55	14	15	16	26	38	17	39	48	17	17	36	30	-	19
B123	UAE_B	37	39	-	-	19	21	22.1	-	-	14	16	-	-	53	14	15	-	26	39	17	40	48	19	18	36	30	-	18
B124	UAE_B	36	37	-	-	20	21	22.1	-	-	14	15	17	-	53	14	15	-	26	39	17	39	48	17	18	37	30	-	20
B125	UAE_B	37	-	-	-	21	23.1	-	-	-	13	16	17	-	53	14	-	-	27	42	16	38	48	18	19	35	30	-	19
B126	UAE_B	36	37	-	-	20	21	21.1	-	-	14	15	17	-	54	14	15	-	26	39	17	39	48	18	18	38	30	-	20
B127	UAE_B	37	38	-	-	21	23	25.1	-	-	10	11	14	16	49	14	15	-	34	40	15	39	48	19	17	36	31	-	18
B128	UAE_B	37	-	-	-	20	22	22.1	-	-	14	15	17	-	54	14	15	-	26	38	17	39	48	18	17	39	30	-	20
B129	UAE_B	35	40	-	-	20	21	25.1	-	-	12	15	-	-	50	13	16	-	27	38	15	38	48	16	17	37	32	-	22
B130	UAE_B	38	40	-	-	20	26	-	-	-	11	13	14	-	48	13	17	-	33	38	13	34	47	19	16	37	32	-	22
B131	UAE_B	36	-	-	-	19	22.1	-	-	-	13	14	16	-	55	14	-	-	25	39	17	40	48	18	17	36	30	-	21
B132	UAE_B	37	-	-	-	19	-	-	-	-	13	15	17	-	55	15	-	-	25	40	17	40	49	18	19	38	30	-	20
B133	UAE_B	38	-	-	-	23	24	25.1	-	-	9	15	16	-	49	14	-	-	33	42	15	39	48	18	18	34	32	-	16
B134	UAE_B	36	37	-	-	20	21	22.1	-	-	13	14	17	-	54	14	15	-	26	39	17	39	48	19	18	38	30	-	20

ID	Population	DYF387S1				DYF399S1					DYF403S1				DYF403S1b	DYF404S1			DYS449	DYS518	DYS526a	DYS526b	DYS547	DYS570	DYS576	DYS612	DYS626		DYS627
B135	UAE_B	37	-	-	-	19	22	22.1	-	-	12	13	17	-	52	14	15	-	26	40	16	38	50	17	18	37	30	-	22
B136	UAE_B	35	39	-	-	20	21	22.1	-	-	11	12	16	-	51	14	15	-	28	40	13	35	48	18	18	35	30	-	20
B137	UAE_B	38	-	-	-	17.2	21	21.1	-	-	15	17	-	-	46	13.2	-	-	29	37	13	35	46	17	18	36	31	-	23
B138	UAE_B	37	-	-	-	20	21	23.1	-	-	13	15	17	-	55	14	15	-	26	38	17	39	48	18	18	35	30	-	20
B139	UAE_B	39	-	-	-	22	24.1	-	-	-	15	16	18	-	54	15	-	-	27	39	15	37	49	19	18	34	29	-	19
B140	UAE_B	36	37	-	-	20	21	22.1	-	-	13	15	17	-	51	14	15	-	26	39	17	39	49	18	19	38	30	-	20
B141	UAE_B	36	37	-	-	21	22	22.1	-	-	13	14	17	-	55	14	15	-	26	39	18	40	48	19	18	36	29	-	19
B142	UAE_B	39	40	-	-	21	21.1	-	-	-	13	15	17	-	46	13.2	14	-	28	38	13	34	44	17	18	38	29	-	19
B143	UAE_B	36	40	-	-	17	24	25.1	-	-	10	16	-	-	50	14	15	-	30	41	13	33	45	18	15	37	31	-	20
B144	UAE_B	35	37	-	-	22	24	24.1	-	-	11	15	16	-	53	15	-	-	31	37	12	33	50	21	17	36	26	-	21
B145	UAE_B	37	-	-	-	20	22	23.1	-	-	13	15	17	-	54	14	15	-	27	39	16	39	48	17	18	36	30	-	20
B146	UAE_B	37	-	-	-	19	20	21	23.1	-	13	15	17	-	54	15	-	-	26	40	17	40	48	19	18	37	30	-	21
B147	UAE_B	37	-	-	-	19	21.1	22	-	-	13	15	16	-	55	14	15	-	26	40	17	40	49	18	17	38	31	-	21
B148	UAE_B	39	-	-	-	21	22	26	-	-	12	13	15	-	48	13	15	-	28	34	13	36	47	18	17	36	34	-	20
B149	UAE_B	37	38	-	-	18	22	23.1	-	-	10	17	18	-	49	13.2	14	-	29	35	12	35	45	18	16	36	28	-	19
B150	UAE_B	38	-	-	-	20.1	21	23	-	-	10	14	17	-	50	13	15	-	32	41	15	40	47	18	17	36	33	-	17
B151	UAE_B	37	-	-	-	19	23.1	25	-	-	10	13	17	-	49	14	16	-	33	41	15	38	49	20	19	37	32	-	17
B152	UAE_B	36	39	-	-	22	24	24.1	-	-	12	15	17	-	50	14	15	-	32	40	15	38	50	18	18	35	32	-	16
B153	UAE_B	35	36	-	-	22	23.1	25	-	-	11	14	18	-	51	15	16	-	31	38	17	33	49	20	17	37	26	-	20
B154	UAE_B	39	-	-	-	22	24.1	-	-	-	15	16	18	-	54	15	-	-	27	39	15	37	49	19	20	36	29	-	19
B155	UAE_B	37	-	-	-	20	22	22.1	-	-	13	16	-	-	54	14	15	-	26	39	17	39	48	18	17	37	30	-	19
B156	UAE_B	35	41	-	-	19	23	25.1	-	-	12	13	14	-	47	14	15	-	27	39	15	43	47	15	17	35	32	-	22
B157	UAE_B	37	-	-	-	21	22.1	-	-	-	14	15	16	-	54	14	15	-	26	38	17	39	49	18	19	35	30	-	20
B158	UAE_B	38	-	-	-	21.1	25	-	-	-	9	13	18	-	51	14	-	-	32	39	15	38	47	18	16	35	32	-	17
B159	UAE_B	37	-	-	-	21	21.1	23	-	-	13	15	16	-	50	13	16	-	25	39	12	32	49	18	17	34	25	-	20
B160	UAE_B	37	-	-	-	22.1	23	25.1	-	-	11	13	18	-	45	12	16	-	34	35	12	34	45	17	15	36	31	-	22
B161	UAE_B	37	-	-	-	20	22	23.1	-	-	13	15	17	-	53	14	15	-	26	39	17	39	48	18	18	36	30	-	20
B162	UAE_B	36	37	-	-	21	23.1	24	-	-	12	14	18	-	51	15	-	-	31	39	12	33	51	19	16	37	26	-	21
B163	UAE_B	35	37	-	-	22	23	25.1	-	-	13	16	-	-	50	15	-	-	32	37	12	33	47	19	15	36	26	-	20
B164	UAE_B	36	38	-	-	19	20	24.1	-	-	13	15	16	-	52	13	-	-	25	39	17	39	49	19	18	36	32	-	23
B165	UAE_B	35	39	-	-	21	22	24.1	-	-	12	14	15	-	51	13	14	-	26	36	14	37	51	16	16	39	30	-	21
B166	UAE_B	39	40	-	-	23	23.2	25.1	-	-	13	14	16	17	46	13.2	16	-	29	34	13	33	48	18	16	37	29	-	21
B167	UAE_B	37	38	-	-	21	22	22.1	-	-	14	16	-	-	54	14	15	-	26	40	17	40	48	17	18	36	30	-	19
B168	UAE_B	39	-	-	-	22	23	26	-	-	11	13	16	-	47	13	16	-	28	38	13	36	47	19	17	35	32	-	20

ID	Population	DYF387S1				DYF399S1					DYF403S1				DYF403S1b	DYF404S1			DYS449	DYS518	DYS526a	DYS526b	DYS547	DYS570	DYS576	DYS612	DYS626		DYS627
B169	UAE_B	37	-	-	-	19	21	22.1	-	-	13	16	-	-	53	14	15	-	26	39	17	40	48	18	18	36	30	-	18
B170	UAE_B	37	-	-	-	20	22	23.1	-	-	11	15	17	-	54	14	16	-	26	39	17	39	48	19	17	37	31	-	21
B171	UAE_B	38	40	-	-	20	21	26	-	-	11	14	-	-	48	13	17	-	32	39	13	34	47	19	16	37	32	-	20
B172	UAE_B	38	-	-	-	23	24.1	-	-	-	15	16	19	-	54	15	-	-	27	39	15	37	50	19	18	36	29	-	19
B173	UAE_B	37	38	-	-	23	26.1	-	-	-	10	15	-	-	50	13	-	-	31	40	16	40	48	18	19	36	29	-	18
B174	UAE_B	37	-	-	-	20	21	21.1	-	-	14	16	17	-	54	14	15	-	26	38	17	39	49	17	18	36	30	-	20
B175	UAE_B	37	-	-	-	19	21	23.1	-	-	13	15	18	-	53	13	15	-	26	39	17	39	48	19	18	36	30	-	18
B176	UAE_B	36	37	-	-	20	21	21.1	-	-	13	14	16	-	52	14	15	-	26	38	16	39	51	18	19	34	30	-	22
B177	UAE_B	37	39	-	-	21	22	26.1	-	-	10	17	-	-	50	15	-	-	31	43	15	40	50	19	17	37	32	-	16
B178	UAE_B	36	37	-	-	20	22	22.1	-	-	13	15	17	-	55	14	15	-	26	39	17	38	48	18	18	37	30	-	20
B179	UAE_B	37	38	-	-	22.1	24	25.1	-	-	11	13	17	-	45	12	15	-	35	35	12	34	45	16	16	37	31	-	22
B180	UAE_B	38	-	-	-	19.2	21.1	23	-	-	12	17	-	-	46	13.2	14	-	30	40	13	38	45	17	18	38	29	-	21
B181	UAE_B	35	37	-	-	21	24	24.1	-	-	11	15	17	-	51	14	15	-	31	40	12	33	50	19	17	37	26	-	22
B182	UAE_B	38	40	-	-	20	25	-	-	-	11	13	14	-	48	13	17	-	33	39	13	34	47	19	15	38	32	-	21
B183	UAE_B	37	38	-	-	20	21	23.1	-	-	14	15	16	-	54	14	15	-	26	38	17	39	48	18	18	34	30	-	17.2
B184	UAE_B	36	-	-	-	19	-	-	-	-	13	14	15	-	53	14	-	-	26	40	16	38	49	17	18	36	30	-	23
B185	UAE_B	35	37	-	-	20	21	22.1	-	-	13	15	17	-	54	14	15	-	26	38	17	39	48	18	18	37	30	-	20
B186	UAE_B	38	-	-	-	20	23	25.1	-	-	11	15	16	-	50	14	15	-	28	40	14	35	48	18	18	37	30	-	22
B187	UAE_B	35	36	-	-	22	23	26.1	-	-	10	13	17	-	50	14	-	-	31	41	14	35	46	17	17	36	29	-	21
B188	UAE_B	35	38	-	-	22	28.1	-	-	-	15	18	19	-	51	15	17	-	32	42	12	32	49	21	17	37	25	-	21
B189	UAE_B	38	40	-	-	20	25	-	-	-	11	13	14	-	48	13	16	-	33	39	13	34	47	19	17	38	32	-	22
B190	UAE_B	35	36	-	-	21	23	25.1	-	-	15	16	-	-	50	15	-	-	35	40	12	33	48	22	18	38	25	-	22
B191	UAE_B	37	-	-	-	20	22	22.1	-	-	13	15	17	-	54	15	-	-	26	40	17	40	48	19	18	37	30	-	21
B192	UAE_B	37	-	-	-	21.1	23	25.1	-	-	11	13	17	-	45	12	15	-	34	35	12	34	49	17	16	35	31	-	22
B193	UAE_B	39	-	-	-	20	22	24	-	-	11	15	-	-	51	14	15	-	32	37	14	31	48	17	18	35	31	-	19
B194	UAE_B	38	-	-	-	21	22	23.1	-	-	12	16	-	-	51	15	17	-	32	37	12	32	52	18	17	37	25	-	20
B195	UAE_B	36	37	-	-	20	21	22.1	-	-	14	17	-	-	54	14	15	-	26	39	17	39	48	18	18	38	30	-	20
B196	UAE_B	37	-	-	-	20	22.1	23	-	-	14	17	-	-	55	14	15	-	26	38	17	39	48	18	17	39	29	-	20
B197	UAE_B	36	37	-	-	20	21	22.1	-	-	14	15	17	-	53	14	15	-	26	39	17	39	48	18	18	38	30	-	21
B198	UAE_B	35	36	-	-	22	24.1	25	-	-	9	11	16	-	49	14	-	-	29	38	14	36	50	17	17	39	31	-	23
B199	UAE_B	37	39	-	-	20.1	21	27	-	-	10	14	18	-	49	14	15	-	32	41	15	39	47	19	19	37	32	-	18
B200	UAE_B	36	38	-	-	21	22	24	-	-	12	14	17	-	50	14	15	-	32	39	14	34	50	20	15	34	30	-	18
U001	UAE_U	37	39	-	-	21	21.1	22	-	-	9	12	16	-	46	13.2	-	-	28	38	14	39	52	18	17	39	31	-	20
U002	UAE_U	36	39	-	-	15	24	27	-	-	11	13	14	-	55	13	14	-	26	35	15	39	50	19	17	37	30	-	20

ID	Population	DYF387S1				DYF399S1					DYF403S1				DYF403S1b	DYF404S1			DYS449	DYS518	DYS526a	DYS526b	DYS547	DYS570	DYS576	DYS612	DYS626		DYS627
U003	UAE_U	38	40	-	-	19	24	25	-	-	11	13	17	-	54	15	-	-	31	42	16	38	48	16	18	37	30	-	20
U004	UAE_U	35	36	-	-	18	23	26.1	-	-	12	15	18	-	51	13	15	-	33	38	13	35	49	22	18	36	30	-	20
U005	UAE_U	37	-	-	-	21	23	24.1	-	-	13	16	-	-	55	15	16	-	26	38	13	35	52	18	19	38	26	-	22
U006	UAE_U	37	38	-	-	22.1	23	25.1	-	-	14	15	17	-	46	15	-	-	26	38	14	39	45	19	17	34	32	-	21
U007	UAE_U	38	-	-	-	21	24	-	-	-	12	15	16	-	50	13	14	-	26	40	17	37	45	18	17	38	28	-	20
U008	UAE_U	32	33	-	-	21	23	23.1	-	-	12	17	20	-	53	14	15	-	26	40	12	39	52	18	15	37	31	-	20
U009	UAE_U	36	37	-	-	18	23	24	-	-	15	17	-	-	53	15	16	-	31	37	12	40	45	18	18	35	32	-	18
U010	UAE_U	37	40	-	-	21	24	24.1	-	-	15	17	-	-	54	13	15	-	27	37	13	34	47	19	20	36	30	-	22
U011	UAE_U	36	41	-	-	19	24	26	-	-	11	14	17	-	55	14	15	-	29	39	17	33	49	19	17	36	34	-	18
U012	UAE_U	35	36	-	-	20	24	24.1	-	-	14	17	-	-	48	14	17	-	33	38	17	33	48	19	19	38	30	-	24
U013	UAE_U	36	40	-	-	22	26.1	28	-	-	13	14	17	-	49	14	-	-	33	39	16	39	46	18	17	37	32	-	22
U014	UAE_U	37	40	-	-	19	22	24.1	-	-	12	14	17	-	50	14	15	-	26	39	12	40	46	18	19	36	31	-	13.2
U015	UAE_U	37	39	-	-	20	20.1	23	-	-	15	16	19	-	49	15	-	-	31	42	14	40	48	19	17	36	26	-	22
U016	UAE_U	37	39	-	-	19	23.1	25	-	-	13	14	16	-	50	13	16	-	30	39	12	36	48	18	17	37	30	-	20
U017	UAE_U	37	-	-	-	21	24	25.1	-	-	15	16	17	-	51	14	15	-	33	40	14	35	49	15	17	38	30	-	20
U018	UAE_U	37	40	41	-	20	24	26	-	-	14	16	-	-	54	14	15	-	35	38	15	40	48	18	17	38	32	-	16
U019	UAE_U	37	-	-	-	19	20	21	21.1	-	12	13	15	-	54	14	15	-	31	38	16	38	45	18	18	34	24	-	21
U020	UAE_U	37	38	-	-	20	24	25	-	-	12	14	16	-	47	14	16	-	26	38	15	38	49	19	18	37	32	-	23
U021	UAE_U	37	-	-	-	20	25	-	-	-	11	14	18	-	54	14	15	-	26	39	14	33	50	21	17	38	31	-	23
U022	UAE_U	37	38	-	-	21	23	24.1	-	-	12	14	16	-	51	14	15	-	32	40	13	37	49	18	16	37	25	-	21
U023	UAE_U	37	-	-	-	20	21	26	-	-	13	15	17	-	50	13	17	-	34	42	12	39	46	17	16	38	32	-	22
U024	UAE_U	38	39	-	-	21	23	23.1	-	-	14	15	16	-	45	13	17	-	29	39	12	43	49	18	17	34	25	-	22
U025	UAE_U	38	-	-	-	21	22	28	-	-	12	13	15	-	53	14	15	-	29	40	17	39	47	19	17	36	30	-	21
U026	UAE_U	36	37	-	-	21	24.1	-	-	-	11	12	17	-	51	15	16	-	32	40	17	38	49	17	18	37	31	-	22
U027	UAE_U	37	-	-	-	20	21	22.1	-	-	8	14	17	-	50	13	-	-	26	37	15	32	47	19	17	37	31	-	19
U028	UAE_U	37	-	-	-	19.1	21	24	-	-	8	14	17	-	52	13	14	-	26	37	13	34	48	18	17	36	25	-	20
U029	UAE_U	36	37	-	-	20	21	25	-	-	8	14	17	-	51	14	15	-	31	39	13	39	48	19	18	37	30	-	20
U030	UAE_U	37	-	-	-	20	21	21.1	-	-	12	16	-	-	49	13	14	-	27	38	13	33	48	19	17	38	29	-	20
U031	UAE_U	36	37	-	-	23	24	26.1	-	-	12	16	-	-	48	13	17	-	27	38	14	33	48	20	16	38	30	-	21
U032	UAE_U	37	38	-	-	23	25.1	-	-	-	12	16	-	-	47.2	14	-	-	26	38	17	39	48	18	18	37	31	-	23
U033	UAE_U	37	-	-	-	21	23.1	-	-	-	11	14	-	-	48	14	-	-	32	39	17	37	48	17	17	35	32	-	18
U034	UAE_U	37	-	-	-	20	21	22.1	-	-	13	18	19	-	45	14	16	-	28	39	15	33	48	19	19	35	30	-	18
U035	UAE_U	36	37	-	-	18	23	24.1	-	-	12	15	-	-	54	14	15	-	25	42	15	39	48	18	15	37	32	-	16
U036	UAE_U	35	37	-	-	20	24	27	-	-	14	16	20	-	50	13	15	-	33	40	16	39	47	18	15	38	30	-	20

ID	Population	DYF387S1				DYF399S1					DYF403S1				DYF403S1b	DYF404S1			DYS449	DYS518	DYS526a	DYS526b	DYS547	DYS570	DYS576	DYS612	DYS626		DYS627
U037	UAE_U	36	38	-	-	21	23.1	-	-	-	13	14	17	-	49	13	14.2	15.2	25	39	17	34	48	18	16	39	31	-	22
U038	UAE_U	35	39	-	-	20	23	23.1	-	-	10	12	17	-	52	14	15	-	26	35	12	37	49	21	17	38	29	-	21
U039	UAE_U	39	40	-	-	19	21	23.1	-	-	11	12	14	-	49	13	14	-	26	39	13	37	48	18	15	39	28	-	19
U040	UAE_U	37	38	-	-	20	21	22.1	-	-	13	15	16	-	50	14	15	-	30	38	17	39	48	18	18	37	29	-	22
U041	UAE_U	39	-	-	-	20	22.1	-	-	-	13	14	15	-	50	14	15	-	27	39	13	34	50	18	19	34	30	-	21
U042	UAE_U	37	-	-	-	21	23	24	-	-	11	15	-	-	49	15	16	-	31	41	16	35	48	17	12	36	31	-	22
U043	UAE_U	37	-	-	-	20	22	23	-	-	14	16	20	-	45	11	13	-	31	39	17	36	49	19	18	35	31	-	21
U044	UAE_U	38	40	-	-	18	24	25.1	-	-	11	13	15	-	49	15	16	-	26	37	17	32	49	18	18	37	29	-	21
U045	UAE_U	38	-	-	-	18	24	25.2	-	-	13	14	16	-	46	14	16	-	33	37	13	37	50	18	18	36	33	-	20
U046	UAE_U	37	38	-	-	18	24	25.1	-	-	13	16	-	-	53	14	15	16	28	39	14	39	49	18	17	38	28	-	21
U047	UAE_U	37	-	-	-	22.1	24	25.1	-	-	12	13	-	-	52	14	15	-	26	40	12	39	48	17	18	37	30	-	20
U048	UAE_U	37	-	-	-	21	22	22.1	-	-	13	15	19	-	46	14	15	-	31	38	13	35	48	19	16	33	30	-	20
U049	UAE_U	36	37	-	-	22	23	26	-	-	12	14	-	-	48	15	-	-	33	39	17	34	49	18	18	35	30	-	23
U050	UAE_U	37	39	-	-	19	21	22.1	-	-	14	15	16	-	55	14	15	-	26	38	14	33	48	18	17	36	31	-	19
U051	UAE_U	36	37	-	-	20	22	23.1	-	-	14	15	16	-	50	13	15	-	26	43	13	33	48	18	16	36	25	-	20
U052	UAE_U	37	38	-	-	20	21	26	-	-	13	15	16	-	51	13.2	14	-	26	39	17	33	48	18	16	35	30	-	20
U053	UAE_U	38	-	-	-	23	24.1	-	-	-	12	13	-	-	53	13	15	-	26	35	16	34	48	17	17	36	25	-	21
U054	UAE_U	35	37	-	-	23	26.1	-	-	-	14	16	17	-	54	14	16	-	31	40	17	34	48	17	18	37	31	-	20
U055	UAE_U	38	40	-	-	20	21	21.1	-	-	13	15	17	-	53	14	15	-	27	40	17	38	48	18	18	37	31	-	20
U056	UAE_U	37	38	-	-	19	21	23.1	-	-	13	14	16	-	54	15	16	-	29	39	14	36	47	17	18	34	30	-	20
U057	UAE_U	36	39	-	-	20	21	21.1	-	-	11	15	17	-	52	15	-	-	31	38	13	36	47	17	18	37	31	-	21
U058	UAE_U	39	-	-	-	21	22	26.1	-	-	12	16	-	-	50	14	15	-	26	40	12	32	47	18	18	35	32	-	21
U059	UAE_U	36	-	-	-	20	22	22.1	-	-	9	13	16	-	52	14	15	-	35	38	15	38	50	19	18	36	31	-	21
U060	UAE_U	36	37	-	-	22.1	24	25.1	-	-	13	15	16	-	50	14	15	-	32	40	13	38	48	18	18	35	31	-	21
U061	UAE_U	39	-	-	-	19.2	21.1	23	-	-	11	12	16	-	48	14	-	-	26	41	13	39	49	18	19	37	24	-	18
U062	UAE_U	35	39	-	-	21	24	24.1	-	-	11	15	17	-	51	13	16	-	28	42	12	38	48	18	18	36	32	-	21
U063	UAE_U	38	39	-	-	20	25	-	-	-	11	14	18	-	53	12	16	-	34	39	14	37	48	21	18	38	32	-	19
U064	UAE_U	36	-	-	-	20	21	23.1	-	-	9	12	16	-	49	14	15	-	26	40	17	40	48	17	15	37	32	-	23
U065	UAE_U	35	37	-	-	19	-	-	-	-	11	13	14	-	54	15	-	-	25	40	14	39	48	18	17	33	30	-	20
U066	UAE_U	38	-	-	-	20	21	22.1	-	-	11	13	17	-	50	15	-	-	25	35	12	39	48	17	18	35	25	-	20
U067	UAE_U	35	36	-	-	20	23	25.1	-	-	12	15	18	-	52	13	-	-	26	37	15	39	48	17	18	36	25	-	21
U068	UAE_U	35	38	-	-	22	23	26.1	-	-	13	16	-	-	53	13	14	-	31	37	17	40	47	17	17	36	30	-	20
U069	UAE_U	38	40	-	-	22	28.1	-	-	-	14	15	17	-	54	13.2	16	-	31	39	17	38	47	19	17	35	29	-	22
U070	UAE_U	35	36	-	-	20	25	-	-	-	12	15	16	-	53	14	15	-	33	38	18	34	47	17	16	36	29	-	19

ID	Population	DYF387S1				DYF399S1					DYF403S1				DYF403S1b	DYF404S1			DYS449	DYS518	DYS526a	DYS526b	DYS547	DYS570	DYS576	DYS612	DYS626		DYS627
U071	UAE_U	37	-	-	-	21	23	25.1	-	-	12	17	20	-	52	13	16	-	32	43	12	38	50	18	17	37	31	-	22
U072	UAE_U	37	-	-	-	20	22	22.1	-	-	15	17	-	-	47	14	15	-	28	39	12	33	49	17	19	37	30	-	20
U073	UAE_U	39	-	-	-	21.1	23	25.1	-	-	15	17	-	-	50	14	-	-	28	37	18	34	48	18	16	34	25	-	21
U074	UAE_U	38	-	-	-	20	22	24	-	-	11	14	17	-	53	12	15	-	31	38	12	39	49	18	16	37	33	-	21
U075	UAE_U	36	37	-	-	21	22	23.1	-	-	14	17	-	-	46	15	-	-	30	41	17	38	48	19	18	35	30	-	21
U076	UAE_U	37	-	-	-	20	21	22.1	-	-	13	14	17	-	54	14	15	-	26	34	17	39	46	18	18	36	26	-	19
U077	UAE_U	36	37	-	-	20	22.1	23	-	-	12	14	17	-	47	14	15	-	32	40	17	35	48	16	17	36	32	-	17
U078	UAE_U	35	36	-	-	20	21	22.1	-	-	15	16	19	-	53	14	16	-	35	37	14	35	49	19	18	37	30	-	20
U079	UAE_U	37	39	-	-	22	24.1	25	-	-	13	14	16	-	54	14	15	-	30	40	13	32	49	18	18	38	30	-	20
U080	UAE_U	36	38	-	-	20.1	21	27	-	-	15	16	17	-	48	14	-	-	26	38	17	34	48	18	18	34	31	-	21
U081	UAE_U	38	39	-	-	21	22	24	-	-	14	16	-	-	54	14	16	-	26	37	17	33	44	18	17	36	30	-	19
U082	UAE_U	39	40	-	-	18	24	25.1	-	-	12	13	15	-	50	14	15	-	25	39	15	40	45	19	18	39	25	-	18
U083	UAE_U	37	38	-	-	21	24	25	-	-	12	14	16	-	54	14	16	-	30	39	17	34	50	17	18	37	25	-	18
U084	UAE_U	34	39	-	-	22	23.1	24.1	-	-	11	14	18	-	53	15	16	-	33	38	16	31	48	18	17	35	25	-	18
U085	UAE_U	36	37	-	-	20	23	24.1	-	-	12	14	16	-	52	15	-	-	26	37	13	32	48	17	17	40	32	-	22
U086	UAE_U	38	-	-	-	20	21	23.1	-	-	13	15	17	-	50	14	16	-	32	40	13	39	49	18	18	40	32	-	17
U087	UAE_U	39	-	-	-	20	23	24.1	-	-	14	15	16	-	55	14	15	-	26	38	17	39	47	19	16	36	30	-	20
U088	UAE_U	37	-	-	-	21.1	22	-	-	-	12	13	15	-	45	14	15	-	26	37	15	39	45	18	18	36	30	-	20
U089	UAE_U	37	38	-	-	21.1	22	-	-	-	11	12	17	-	46	15	-	-	27	39	17	38	47	18	17	36	31	-	21
U090	UAE_U	36	37	-	-	21	22.1	23.1	-	-	8	14	17	-	51	14	-	-	30	40	18	34	46	20	18	36	29	-	18
U091	UAE_U	37	-	-	-	20	21	23.1	-	-	8	14	17	-	48	14	16	-	28	38	13	35	47	18	18	39	30	-	20
U092	UAE_U	38	39	-	-	20	21	23.1	-	-	8	14	17	-	54	14	-	-	28	37	13	35	48	20	16	38	29	-	22
U093	UAE_U	39	-	-	-	18	21	26.1	-	-	12	16	-	-	55	14	15	-	26	36	12	37	48	19	18	37	32	-	17
U094	UAE_U	38	39	-	-	22	23.1	24	-	-	12	16	-	-	47	15	-	-	27	41	17	33	46	18	17	31	35	-	20
U095	UAE_U	36	39	-	-	22	24	25	-	-	12	16	-	-	54	12	15	-	26	38	17	39	49	15	17	34	25	-	20
U096	UAE_U	39	-	-	-	20	22	23	-	-	11	14	-	-	53	14	15	-	33	38	17	37	49	18	18	37	29	-	18
U097	UAE_U	35	39	-	-	21	22	25.1	-	-	13	18	19	-	54	15	17	-	31	37	12	33	48	18	18	39	30	-	20
U098	UAE_U	38	-	-	-	21	24	24.1	-	-	12	15	-	-	50	14	15	-	32	38	15	40	49	18	16	34	25	-	18
U099	UAE_U	37	-	-	-	20	22	23.1	-	-	14	16	20	-	49	14	15	-	33	41	14	36	48	17	18	35	31	-	22
U100	UAE_U	36	39	-	-	20	24	25.1	-	-	13	14	17	-	53	14	15	-	35	39	14	40	49	18	17	36	31	-	21
U101	UAE_U	35	36	-	-	22	23	24.1	-	-	10	12	17	-	51	14	-	-	26	38	14	39	49	19	18	36	32	-	23
U102	UAE_U	39	-	-	-	20	26	-	-	-	11	12	14	-	54	14	15	-	34	41	18	34	50	19	16	36	25	-	20
U103	UAE_U	37	-	-	-	22.1	23	25.1	-	-	13	15	16	-	53	14	15	-	32	35	13	37	49	19	14	35	31	-	19
U104	UAE_U	35	41	-	-	21	24	24.1	-	-	13	14	15	-	55	14	15	-	32	38	13	40	48	16	18	36	31	-	17

ID	Population	DYF387S1				DYF399S1					DYF403S1				DYF403S1b	DYF404S1			DYS449	DYS518	DYS526a	DYS526b	DYS547	DYS570	DYS576	DYS612	DYS626		DYS627
U105	UAE_U	37	-	-	-	20	22	22.1	-	-	11	15	-	-	48	14	-	-	26	39	13	39	49	18	18	36	30	-	18
U106	UAE_U	38	-	-	-	20	22.1	-	-	-	8	12	17	-	51	12	15	-	26	37	12	39	50	17	18	37	29	-	21
U107	UAE_U	37	-	-	-	22	23	24.1	-	-	9	11	17	-	52	15	-	-	26	36	15	39	49	19	16	35	31	-	23
U108	UAE_U	36	38	-	-	22	23	23.1	24	24.1	11	12	17	-	54	14	15	-	29	39	15	40	49	18	16	38	30	-	19
U109	UAE_U	37	38	-	-	17.2	21	21.1	-	-	14	17	19	-	54	14	15	-	32	39	15	38	48	19	17	36	33	-	17
U110	UAE_U	36	37	-	-	17.2	21	21.1	-	-	12	16	-	-	51	14	16	-	32	38	17	34	47	19	16	35	32	-	17
U111	UAE_U	39	-	-	-	21	22	24.1	-	-	13	14	-	-	48	14	15	-	31	39	15	38	49	19	17	34	25	-	19
U112	UAE_U	35	38	-	-	21	22.1	-	-	-	11	13	17	-	48	14	-	-	26	42	17	33	47	18	17	38	33	-	21
U113	UAE_U	39	-	-	-	18	24	25.1	-	-	13	14	15	-	48	14	16	-	35	37	15	34	49	17	14	35	28	-	18
U114	UAE_U	37	-	-	-	21	24	25	-	-	12	14	16	-	49	14	15	-	32	38	17	39	45	19	18	37	30	-	21
U115	UAE_U	38	39	-	-	22	23.1	24.1	-	-	11	13	16	-	49	14	16	-	26	38	15	38	47	20	20	36	29	-	20
U116	UAE_U	39	40	-	-	20	23	24.1	-	-	11	12	16	-	49	13	15	-	28	41	12	39	49	18	18	37	30	-	21
U117	UAE_U	37	38	-	-	20	21	23.1	-	-	11	13	17	-	49	16	-	-	34	39	12	35	50	16	17	37	31	-	17
U118	UAE_U	34	39	-	-	20	23	24.1	-	-	10	13	17	-	50	14	15	-	26	40	17	35	50	17	14	38	33	-	19
U119	UAE_U	36	37	-	-	20	23	23.1	-	-	15	16	-	-	48	15	-	-	25	38	12	32	48	17	15	36	30	-	20
U120	UAE_U	38	-	-	-	20	21.1	-	-	-	12	13	15	-	47	14	15	-	25	37	12	34	48	19	17	38	25	-	21
U121	UAE_U	39	-	-	-	19	21	24.1	-	-	13	14	15	-	53	15	16	-	26	36	17	33	47	20	18	35	29	-	21
U122	UAE_U	37	-	-	-	20	21	22.1	-	-	13	17	-	-	48	14	15	-	31	41	14	40	50	17	17	38	30	-	19
U123	UAE_U	37	38	-	-	20	21	24.1	-	-	10	11	18	-	53	14	16	-	31	38	13	34	48	17	16	34	30	-	18
U124	UAE_U	36	37	-	-	19	22	24	-	-	13	17	-	-	51	14	15	-	33	38	17	31	45	20	17	38	30	-	20
U125	UAE_U	37	-	-	-	20	23	25.1	-	-	13	15	17	-	50	14	-	-	32	37	13	32	45	18	16	36	30	-	19
U126	UAE_U	38	39	-	-	21	21.1	22	-	-	11	12	16	-	47	12	16	-	28	38	17	39	50	17	16	36	30	-	20
U127	UAE_U	39	-	-	-	22	23	24.1	-	-	11	14	16	-	52	14	15	-	28	41	17	39	47	18	18	40	31	-	18
U128	UAE_U	38	39	-	-	21	23	28	-	-	13	15	16	-	49	14	-	-	31	39	13	39	48	18	18	38	30	-	20
U129	UAE_U	36	39	-	-	20	21	22.1	-	-	11	13	15	-	55	14	15	-	30	40	17	36	49	19	18	35	32	-	22
U130	UAE_U	39	-	-	-	22	24	24.1	-	-	14	15	16	-	53	14	15	-	26	40	12	39	49	17	17	38	32	-	22
U131	UAE_U	35	39	-	-	21	25	26	-	-	13	16	-	-	53	14	15	-	32	39	17	34	49	19	18	36	30	-	21
U132	UAE_U	35	37	-	-	20	21	22.1	-	-	14	15	17	-	53	13	16	-	35	38	15	33	49	17	18	37	30	-	20
U133	UAE_U	39	41	-	-	20	23	25	-	-	13	16	17	-	54	13	17	-	30	41	15	34	48	19	17	38	32	-	16
U134	UAE_U	36	37	-	-	21	23.1	-	-	-	14	15	17	-	49	14	-	-	26	39	14	37	44	17	18	39	30	-	20
U135	UAE_U	38	41	-	-	22	23	25.1	-	-	10	11	14	16	54	15	16	-	26	38	12	39	47	18	18	37	30	-	22
U136	UAE_U	37	38	-	-	20	21	22.1	-	-	14	15	17	-	50	15	17	-	25	41	14	39	48	21	16	38	30	-	20
U137	UAE_U	38	-	-	-	20	21	22.1	-	-	13	15	17	-	48	14	15	-	30	34	17	41	46	18	18	36	31	-	23
U138	UAE_U	37	-	-	-	20	21	26.1	-	-	9	15	16	-	55	14	15	-	33	40	17	35	48	19	18	36	30	-	20

ID	Population	DYF387S1				DYF399S1					DYF403S1				DYF403S1b	DYF404S1			DYS449	DYS518	DYS526a	DYS526b	DYS547	DYS570	DYS576	DYS612	DYS626		DYS627
U139	UAE_U	36	37	-	-	20	21	22	23	-	13	14	17	-	55	15	16	-	26	37	17	33	50	18	17	40	29	-	19
U140	UAE_U	37	-	-	-	18	23	25.1	-	-	12	13	17	-	49	14	15	-	32	40	16	40	48	18	18	38	30	-	20
U141	UAE_U	35	36	-	-	20	21	23.1	-	-	11	12	16	-	54	14	16	-	26	38	17	36	48	18	18	35	29	-	19
U142	UAE_U	34	36	-	-	21	22	22.1	-	-	15	17	-	-	52	14	15	-	26	37	15	40	45	19	18	38	29	-	19
U143	UAE_U	36	40	-	-	20	21	23.1	-	-	13	15	17	-	51	15	-	-	27	39	17	39	45	18	19	36	31	-	20
U144	UAE_U	38	-	-	-	20	21	23.1	-	-	15	16	18	-	50	14	16	-	30	39	15	34	48	17	19	37	26	-	21
U145	UAE_U	38	-	-	-	20	21	23.1	-	-	14	15	17	-	48	14	15	-	28	38	13	40	48	17	18	38	30	-	20
U146	UAE_U	37	-	-	-	21	24	24.1	-	-	15	16	17	-	50	14	15	-	28	39	13	34	48	20	18	39	30	-	21
U147	UAE_U	39	-	-	-	21	-	-	-	-	12	16	17	-	45	14	15	-	26	41	15	33	49	19	19	37	31	-	21
U148	UAE_U	36	37	-	-	19.1	24	24	-	-	14	16	17	-	54	14	-	-	27	36	16	34	49	19	18	37	34	-	20
U149	UAE_U	36	37	-	-	19	21	23.1	-	-	11	12	15	-	51	13	15	-	26	41	17	34	49	17	17	36	28	-	19
U150	UAE_U	39	40	-	-	20	21	23.1	-	-	13	17	-	-	50	14	16	-	28	34	17	34	49	19	17	36	33	-	17
U151	UAE_U	36	40	-	-	21	22.1	23.1	-	-	13	15	16	-	45	14	15	-	29	39	16	39	49	18	17	36	32	-	17
U152	UAE_U	35	37	-	-	20	23	24.1	-	-	12	15	16	-	53	15	-	-	26	37	15	38	50	18	16	36	32	-	16
U153	UAE_U	37	-	-	-	19	22	25.1	-	-	12	14	15	-	51	13	17	-	27	41	17	35	51	18	17	34	26	-	20
U154	UAE_U	37	-	-	-	21	22.1	-	-	-	11	12	16	-	50	12	15	-	26	38	12	32	49	18	19	35	29	-	19
U155	UAE_U	37	-	-	-	20	21	22.1	-	-	11	15	16	-	52	15	-	-	26	40	13	39	47	17	18	35	30	-	19
U156	UAE_U	39	-	-	-	21	21.1	25	-	-	12	15	16	-	51	14	15	-	28	38	12	40	48	17	17	35	32	-	22
U157	UAE_U	37	38	-	-	20	21	22.1	-	-	14	15	17	-	46	14	15	-	30	42	13	33	50	20	18	38	30	-	20
U158	UAE_U	38	-	-	-	20	23	24.1	-	-	13	17	-	-	54	14	16	-	31	38	17	39	48	20	18	38	32	-	17
U159	UAE_U	37	-	-	-	20	23	24.1	-	-	12	14	16	-	47	14	15	-	26	39	16	36	48	20	17	39	25	-	20
U160	UAE_U	36	39	-	-	21	23.1	-	-	-	15	16	17	-	53	13.2	-	-	26	40	17	33	49	18	17	33	31	-	22
U161	UAE_U	35	36	-	-	21	21.1	22	-	-	13	16	17	-	54	12	16	-	26	38	14	36	45	16	18	34	30	-	20
U162	UAE_U	39	-	-	-	21	24	25.1	-	-	14	15	17	-	48	14	15	-	28	40	14	34	49	18	16	37	26	-	21
U163	UAE_U	37	-	-	-	20	21	22	22.1	-	13	14	17	-	54	14	15	-	29	39	12	33	48	19	18	36	26	-	20
U164	UAE_U	35	41	-	-	19	21	22.1	-	-	11	13	-	-	50	14	15	-	32	41	13	36	48	19	17	34	32	-	23
U165	UAE_U	37	-	-	-	20	21	22.1	-	-	15	16	-	-	54	13	15	-	33	36	14	39	50	19	18	37	30	-	21
U166	UAE_U	38	-	-	-	21	23.1	-	-	-	13	15	17	-	53	15	16	-	32	41	13	41	48	19	18	36	29	-	21
U167	UAE_U	37	-	-	-	20	21	21.1	-	-	12	16	-	-	52	14	15	-	31	34	13	39	49	20	16	37	30	-	19
U168	UAE_U	37	-	-	-	21	23	25.1	-	-	13	15	17	-	50	15	17	-	27	39	13	41	48	18	18	37	32	-	20
U169	UAE_U	37	-	-	-	20	22	22.1	-	-	13	15	17	-	55	14	15	-	26	38	14	38	51	17	17	37	30	-	18
U170	UAE_U	36	37	-	-	20	21	25.1	-	-	14	15	17	-	45	14	15	-	27	38	14	39	48	17	17	36	31	-	21
U171	UAE_U	35	37	-	-	20	26	-	-	-	12	15	16	-	46	14	16	-	26	41	14	35	49	17	18	36	32	-	20
U172	UAE_U	36	38	-	-	19	22.1	-	-	-	14	16	20	-	54	13	17	-	32	41	13	32	48	19	18	37	29	-	19

ID	Population	DYF387S1				DYF399S1					DYF403S1				DYF403S1b	DYF404S1			DYS449	DYS518	DYS526a	DYS526b	DYS547	DYS570	DYS576	DYS612	DYS626		DYS627
U173	UAE_U	35	39	-	-	19	-	-	-	-	11	13	15	-	50	14.2	-	-	25	39	13	38	48	18	16	35	29	-	18
U174	UAE_U	39	40	-	-	23	24	25.1	-	-	11	14	16	-	50	15	-	-	34	39	13	39	45	16	18	36	30	-	20
U175	UAE_U	37	38	-	-	20	21	22.1	-	-	13	15	16	-	51	14	-	-	26	41	12	40	50	18	17	36	30	-	18
U176	UAE_U	39	-	-	-	19	22	22.1	-	-	11	13	15	-	48	12	14	-	31	34	18	39	49	17	17	36	30	-	22
U177	UAE_U	37	-	-	-	20	21	22.1	-	-	14	15	16	-	50	14	15	-	32	36	15	38	49	19	18	34	32	-	16
U178	UAE_U	37	-	-	-	17.2	21	21.1	-	-	13	16	-	-	54	14	15	-	25	38	13	39	49	18	18	37	30	-	20
U179	UAE_U	35	37	-	-	20	21	23.1	-	-	11	14	19	-	45	15	16	-	26	39	12	39	50	19	19	37	31	-	22
U180	UAE_U	39	41	-	-	22	24.1	-	-	-	12	17	20	-	51	14	15	-	29	37	14	39	51	19	17	38	29	-	21
U181	UAE_U	36	37	-	-	20	21	22.1	-	-	14	16	-	-	51	14	16	-	33	36	15	38	49	19	17	35	26	-	22
U182	UAE_U	38	41	-	-	21	22	22.1	-	-	12	15	16	-	54	14	15	-	33	40	13	34	47	18	15	36	32	-	22
U183	UAE_U	37	38	-	-	21	21.1	-	-	-	12	14	18	-	55	15	-	-	26	37	14	40	48	17	18	36	30	-	17.2
U184	UAE_U	38	-	-	-	17	24	25.1	-	-	14	15	17	-	53	14	16	-	31	38	12	40	50	19	20	36	30	-	23
U185	UAE_U	37	-	-	-	22	24	24.1	-	-	15	16	17	-	49	14	15	-	30	39	13	32	48	18	17	35	30	-	20
U186	UAE_U	36	37	-	-	20	22	23.1	-	-	12	16	17	-	49	14	15	-	33	35	12	39	48	19	19	35	30	-	22
U187	UAE_U	37	-	-	-	19	20	21	23.1	-	14	16	17	-	50	14	15	-	31	39	15	39	49	18	17	37	29	-	21
U188	UAE_U	35	36	-	-	19	21.1	22	-	-	11	12	15	-	53	15	-	-	26	37	15	39	45	16	19	37	25	-	21
U189	UAE_U	34	36	-	-	21	22	26	-	-	13	17	-	-	52	14	15	-	26	38	14	38	49	17	17	37	32	-	22
U190	UAE_U	36	40	-	-	18	22	23.1	-	-	13	15	16	-	46	12	15	-	31	45	14	36	48	19	17	38	30	-	23
U191	UAE_U	38	-	-	-	20.1	21	23	-	-	12	15	16	-	48	13.2	14	-	32	41	14	35	48	19	17	37	30	-	21
U192	UAE_U	37	-	-	-	19	23.1	25	-	-	12	14	15	-	55	14	15	-	29	38	13	37	50	18	17	36	30	-	20
U193	UAE_U	37	-	-	-	22	24	24.1	-	-	11	12	16	-	50	13	17	-	28	38	15	37	48	17	18	38	31	-	18
U194	UAE_U	35	37	-	-	22	23.1	25	-	-	11	15	16	-	51	14	15	-	35	40	15	34	49	18	18	37	24	-	18
U195	UAE_U	37	38	-	-	22	24.1	-	-	-	12	15	16	-	53	14	-	-	25	42	14	35	48	18	17	36	25	-	20
U196	UAE_U	40	-	-	-	20	22	22.1	-	-	14	15	17	-	54	14	15	-	33	41	13	42	51	17	16	34	32	-	16
U197	UAE_U	35	36	-	-	19	23	25.1	-	-	13	17	-	-	53	14	15	-	33	40	14	39	48	21	16	36	34	-	20
U198	UAE_U	36	38	-	-	21	22.1	-	-	-	12	14	16	-	54	14	-	-	33	40	13	32	49	19	15	37	30	-	20
U199	UAE_U	36	37	-	-	21.1	25	-	-	-	15	16	17	-	52	15	17	-	28	39	14	31	48	22	18	36	26	-	21
U200	UAE_U	37	38	-	-	21	21.1	23	-	-	13	16	17	-	50	13	16	-	27	42	15	36	48	19	16	36	31	-	20
R001	UAE_R	36	40	-	-	22.1	23	25.1	-	-	14	15	17	-	52	15	-	-	31	40	16	39	45	17	15	36	30	-	20
R002	UAE_R	35	36	39	40	20	22	23.1	-	-	13	14	17	-	50	14	15	-	29	39	13	34	50	17	18	36	31	-	19
R003	UAE_R	37	-	-	-	21	23.1	24	-	-	11	13	-	-	48	14	16	-	35	35	13	33	49	18	16	38	33	-	20
R004	UAE_R	38	-	-	-	22	23	25.1	-	-	15	16	-	-	54	13	17	-	34	39	14	34	51	18	16	39	28	-	21
R005	UAE_R	38	-	-	-	19	20	24.1	-	-	13	15	17	-	53	15	-	-	33	38	17	34	45	18	18	36	30	-	20
R006	UAE_R	37	-	-	-	21	22	24.1	-	-	12	16	-	-	50	13	-	-	32	39	12	34	47	18	17	35	30	-	20

ID	Population	DYF387S1				DYF399S1					DYF403S1				DYF403S1b	DYF404S1			DYS449	DYS518	DYS526a	DYS526b	DYS547	DYS570	DYS576	DYS612	DYS626		DYS627
R007	UAE_R	35	40	-	-	23	23.2	25.1	-	-	13	15	17	-	48	14	15	-	34	38	12	40	45	17	18	37	30	-	23
R008	UAE_R	38	40	-	-	20	21	23.1	-	-	13	15	17	-	45	13	15	-	29	39	13	35	51	19	17	35	31	-	19
R009	UAE_R	36	-	-	-	22	24.1	-	-	-	14	15	17	-	51	14	15	-	27	39	17	32	48	20	16	38	25	-	20
R010	UAE_R	37	-	-	-	20	21	22.1	-	-	12	15	16	-	54	14	15	-	28	39	17	39	48	19	20	36	30	-	20
R011	UAE_R	38	-	-	-	21	22	22.1	-	-	12	15	-	-	53	13.2	-	-	30	38	16	39	49	18	17	38	25	-	18
R012	UAE_R	36	37	-	-	21	21.1	-	-	-	11	13	14	-	50	13.2	-	-	31	41	12	40	49	16	17	37	32	-	21
R013	UAE_R	37	-	-	-	17	24	25.1	-	-	13	14	16	-	49	13	14	-	33	37	14	34	46	17	19	39	31	-	18
R014	UAE_R	35	39	-	-	22	24	24.1	-	-	13	15	17	-	46	15	-	-	29	39	12	39	46	19	16	36	24	-	22
R015	UAE_R	38	-	-	-	20	22	23.1	-	-	9	15	16	-	46	13	15	-	28	40	14	34	47	18	17	35	30	-	21
R016	UAE_R	37	-	-	-	19	20	21	23.1	-	13	14	17	-	50	15	16	-	28	40	15	39	48	17	17	36	30	-	22
R017	UAE_R	39	-	-	-	19	21.1	22	-	-	12	13	17	-	54	15	-	-	25	34	16	40	48	19	19	35	26	-	18
R018	UAE_R	36	37	-	-	21	22	26	-	-	11	12	16	-	55	13	14	-	29	35	15	38	46	18	15	35	30	-	20
R019	UAE_R	36	37	-	-	18	22	23.1	-	-	15	17	-	-	47	14	15	-	25	41	14	33	49	18	18	35	30	-	23
R020	UAE_R	39	40	-	-	20.1	21	23	-	-	13	15	17	-	54	15	16	-	31	41	13	35	49	21	16	34	30	-	21
R021	UAE_R	36	40	-	-	19	23.1	25	-	-	15	16	18	-	53	13	15	-	33	40	12	31	48	18	16	36	30	-	20
R022	UAE_R	35	37	-	-	22	24	24.1	-	-	13	15	17	-	54	14	15	-	29	38	12	36	49	18	18	36	31	-	18
R023	UAE_R	37	-	-	-	19	21	23.1	-	-	13	14	17	-	50	14	17	-	32	39	17	38	49	18	18	37	24	-	18
R024	UAE_R	37	-	-	-	20	21	22.1	-	-	13	15	17	-	49	14	-	-	29	39	17	38	51	17	17	36	25	-	20
R025	UAE_R	37	-	-	-	20	22.1	-	-	-	10	16	-	-	53	14	15	-	26	39	15	38	45	19	18	36	32	-	16
R026	UAE_R	39	-	-	-	21	23	24	-	-	12	16	-	-	51	15	-	-	26	38	17	38	47	18	18	39	34	-	20
R027	UAE_R	37	38	-	-	20	22	23	-	-	11	13	17	-	54	13	16	-	26	39	14	34	47	15	18	37	30	-	20
R028	UAE_R	38	39	-	-	18	24	25.1	-	-	13	14	15	-	53	14	15	-	27	39	12	34	49	18	19	36	26	-	21
R029	UAE_R	38	39	-	-	18	24	25.2	-	-	12	14	16	-	55	14	15	-	26	35	15	33	50	18	18	35	31	-	20
R030	UAE_R	40	-	-	-	18	24	25.1	-	-	11	13	16	-	48	14	15	-	34	39	17	39	50	18	18	36	25	-	21
R031	UAE_R	40	-	-	-	20	24	25	-	-	11	12	16	-	51	14	-	-	26	39	17	39	48	18	19	37	31	-	20
R032	UAE_R	38	39	-	-	20	25	-	-	-	11	13	17	-	52	14	15	-	27	37	18	38	48	17	17	36	31	-	20
R033	UAE_R	38	40	-	-	21	23	24.1	-	-	10	13	17	-	54	14	15	-	33	39	12	34	49	17	18	35	30	-	20
R034	UAE_R	38	39	-	-	20	21	26	-	-	15	16	-	-	54	14	15	-	25	36	12	35	49	20	16	37	31	-	21
R035	UAE_R	37	-	-	-	21	23	23.1	-	-	12	13	15	-	51	13	16	-	25	34	18	35	49	20	18	35	28	-	19
R036	UAE_R	37	-	-	-	21	22	28	-	-	13	14	15	-	48	13	17	-	33	40	12	37	49	20	17	38	29	-	22
R037	UAE_R	35	37	-	-	21	24.1	-	-	-	13	17	-	-	49	14	-	-	26	38	17	39	48	18	17	36	30	-	21
R038	UAE_R	37	38	-	-	20	21	22.1	-	-	10	11	18	-	48	15	-	-	35	39	17	39	44	16	18	34	31	-	22
R039	UAE_R	40	-	-	-	19.1	21	24	-	-	13	17	-	-	47.2	14	-	-	31	39	17	33	47	18	18	37	31	-	21
R040	UAE_R	35	36	-	-	20	21	25	-	-	13	15	17	-	48	14	15	-	26	39	13	33	48	19	18	39	29	-	21

ID	Population	DYF387S1				DYF399S1					DYF403S1				DYF403S1b	DYF404S1			DYS449	DYS518	DYS526a	DYS526b	DYS547	DYS570	DYS576	DYS612	DYS626		DYS627
R041	UAE_R	36	38	-	-	21	21.1	22	-	-	11	12	16	-	45	14	15	-	26	37	13	32	46	19	18	36	32	-	23
R042	UAE_R	36	37	-	-	15	24	27	-	-	11	14	16	-	47	14	15	-	32	41	13	38	48	19	17	35	30	-	20
R043	UAE_R	37	38	-	-	19	24	25	-	-	13	15	16	-	52	13.2	-	-	34	38	14	35	47	19	17	36	25	-	20
R044	UAE_R	36	40	-	-	21	23	28	-	-	11	13	15	-	49	14	15	-	29	40	17	35	45	20	17	35	25	-	21
R045	UAE_R	35	36	39	40	20	21	22.1	-	-	14	15	16	-	55	15	-	-	29	38	12	33	51	18	18	36	31	-	20
R046	UAE_R	37	-	-	-	22	24	24.1	-	-	13	16	-	-	53	14	15	-	32	42	17	40	48	17	18	36	30	-	20
R047	UAE_R	38	-	-	-	21	25	26	-	-	14	15	17	-	53	14	15	-	26	38	12	36	48	17	16	36	25	-	22
R048	UAE_R	38	-	-	-	20	21	22.1	-	-	13	16	17	-	53	13.2	14	-	26	39	14	36	49	17	18	35	30	-	21
R049	UAE_R	37	38	-	-	20	23	25	-	-	14	15	17	-	54	14	15	-	31	40	12	32	49	19	17	36	31	-	22
R050	UAE_R	37	-	-	-	21	23.1	-	-	-	10	11	14	16	49	15	-	-	27	38	17	38	46	18	18	36	31	-	19
R051	UAE_R	37	38	-	-	22	23	25.1	-	-	10	13	17	-	54	14	15	-	27	40	17	38	48	20	17	37	25	-	20
R052	UAE_R	35	38	-	-	20	21	22.1	-	-	12	15	17	-	50	13	15	-	26	37	17	40	51	18	18	35	30	-	20
R053	UAE_R	35	37	-	-	20	21	22.1	-	-	11	14	18	-	48	13	14.2	15.2	32	36	14	34	47	16	17	38	29	-	20
R054	UAE_R	39	40	-	-	20	21	26.1	-	-	15	16	18	-	55	14	15	-	28	39	15	33	49	17	16	37	30	-	21
R055	UAE_R	35	41	-	-	20	21	22	23	-	13	16	-	-	55	13	14	-	25	39	14	34	51	17	17	38	31	-	23
R056	UAE_R	37	-	-	-	20	21.1	-	-	-	12	13	14	-	49	14	15	-	32	38	13	34	48	19	17	35	32	-	18
R057	UAE_R	37	38	-	-	19	21	24.1	-	-	14	15	16	-	54	14	15	-	34	39	14	34	48	20	16	36	30	-	18
R058	UAE_R	37	-	-	-	20	21	22.1	-	-	9	13	18	-	52	15	16	-	29	42	12	39	47	17	18	36	25	-	18
R059	UAE_R	38	-	-	-	20	21	24.1	-	-	13	15	16	-	51	11	13	-	27	37	13	38	48	17	17	36	32	-	21
R060	UAE_R	35	36	-	-	19	22	24	-	-	11	15	16	-	46	15	16	-	28	38	17	35	48	20	19	35	31	-	18
R061	UAE_R	38	40	-	-	20	23	25.1	-	-	10	13	17	-	55	14	16	-	30	38	14	32	47	18	15	35	24	-	22
R062	UAE_R	37	38	-	-	21	21.1	22	-	-	15	18	19	-	54	14	15	16	31	41	13	39	50	17	15	37	30	-	21
R063	UAE_R	36	37	-	-	22	23	24.1	-	-	11	13	14	-	51	14	16	-	33	39	17	40	48	18	18	37	30	-	22
R064	UAE_R	37	-	-	-	20	23	23.1	-	-	15	16	-	-	55	14	15	-	29	37	16	33	49	18	17	37	26	-	18
R065	UAE_R	36	-	-	-	22	24.1	-	-	-	13	15	17	-	46	15	-	-	28	38	17	39	48	19	16	34	30	-	20
R066	UAE_R	36	39	-	-	20	22	22.1	-	-	11	13	17	-	50	13	17	-	28	39	17	36	51	17	17	37	26	-	22
R067	UAE_R	40	-	-	-	19	23	25.1	-	-	11	15	-	-	53	12	15	-	25	37	14	39	50	19	17	36	32	-	22
R068	UAE_R	38	-	-	-	21	22.1	-	-	-	12	16	-	-	53	15	-	-	29	38	13	38	48	17	17	36	30	-	17.2
R069	UAE_R	38	-	-	-	21.1	25	-	-	-	14	17	-	-	54	14	15	-	25	41	12	37	45	17	19	36	30	-	23
R070	UAE_R	36	37	-	-	21	21.1	23	-	-	14	17	-	-	55	14	15	-	31	35	15	37	45	18	18	37	30	-	20
R071	UAE_R	36	37	-	-	22.1	23	25.1	-	-	14	15	17	-	48	14	16	-	33	38	13	39	50	17	17	38	30	-	22
R072	UAE_R	36	37	-	-	20	22	23.1	-	-	9	11	16	-	49	14	-	-	29	37	13	40	47	18	20	34	29	-	21
R073	UAE_R	36	37	-	-	15	24	27	-	-	11	14	16	-	47	14	15	-	32	41	13	38	48	19	17	35	30	-	20
R074	UAE_R	38	-	-	-	22	23	25.1	-	-	12	14	17	-	49	15	-	-	29	38	12	39	49	18	17	39	30	-	20

ID	Population	DYF387S1				DYF399S1					DYF403S1				DYF403S1b	DYF404S1			DYS449	DYS518	DYS526a	DYS526b	DYS547	DYS570	DYS576	DYS612	DYS626		DYS627
R075	UAE_R	37	-	-	-	19	20	24.1	-	-	8	14	17	-	50	14	-	-	26	38	15	38	48	17	19	37	30	-	20
R076	UAE_R	37	-	-	-	21	22	24.1	-	-	12	15	16	-	51	14	15	-	26	38	17	35	48	18	16	35	26	-	22
R077	UAE_R	35	36	-	-	23	23.2	25.1	-	-	11	14	19	-	51	14	15	-	26	39	16	35	46	18	18	40	32	-	21
R078	UAE_R	39	-	-	-	21	22	22.1	-	-	12	17	20	-	53	14	15	-	25	38	13	39	49	17	17	40	28	-	20
R079	UAE_R	37	-	-	-	22	23	26	-	-	14	16	-	-	49	13.2	-	-	26	38	13	37	47	21	17	36	31	-	20
R080	UAE_R	39	-	-	-	19	21	22.1	-	-	12	15	16	-	54	14	15	-	26	38	17	39	48	19	17	36	32	-	18
R081	UAE_R	38	40	-	-	20	22	23.1	-	-	12	14	18	-	50	15	-	-	30	41	15	40	48	22	16	36	30	-	22
R082	UAE_R	38	-	-	-	20	21	26	-	-	14	16	-	-	52	14	15	-	27	41	17	34	49	19	17	36	34	-	18
R083	UAE_R	37	38	-	-	20	22	22.1	-	-	11	13	16	-	53	14	15	-	31	39	18	33	48	17	19	39	30	-	24
R084	UAE_R	37	-	-	-	20	21	25.1	-	-	13	16	-	-	54	13.2	14	-	31	39	13	33	52	17	18	38	32	-	22
R085	UAE_R	37	-	-	-	20	26	-	-	-	11	15	17	-	53	14	15	-	31	41	13	39	48	18	18	37	31	-	13.2
R086	UAE_R	36	37	-	-	19	22.1	-	-	-	8	12	17	-	52	15	-	-	26	34	12	40	48	18	18	31	26	-	22
R087	UAE_R	37	39	-	-	19	-	-	-	-	9	11	17	-	47	14	15	-	26	36	17	40	48	18	20	34	30	-	20
R088	UAE_R	36	37	-	-	23	24	25.1	-	-	11	12	17	-	54	14	15	-	31	38	17	36	50	18	18	37	30	-	20
R089	UAE_R	37	38	-	-	20	21	22.1	-	-	14	17	19	-	54	14	-	-	33	39	17	35	47	17	17	36	32	-	16
R090	UAE_R	38	-	-	-	19	22	22.1	-	-	13	14	-	-	55	13	16	-	25	37	13	40	50	19	14	39	24	-	21
R091	UAE_R	35	37	-	-	20	21	22.1	-	-	13	15	17	-	49	12	16	-	33	36	12	38	48	20	15	37	32	-	23
R092	UAE_R	38	40	-	-	17.2	21	21.1	-	-	13	15	17	-	49	14	15	-	26	40	15	38	48	18	17	36	31	-	23
R093	UAE_R	37	38	-	-	21	22	26.1	-	-	13	15	16	-	51	15	-	-	26	37	15	33	45	17	18	35	31	-	17
R094	UAE_R	36	-	-	-	20	22	22.1	-	-	12	13	15	-	51	15	-	-	28	38	15	37	45	19	18	36	30	-	20
R095	UAE_R	35	37	-	-	22.1	24	25.1	-	-	10	17	18	-	53	13	-	-	29	39	17	39	48	18	18	34	30	-	20
R096	UAE_R	38	-	-	-	19.2	21.1	23	-	-	10	14	17	-	49	13	14	-	26	40	15	43	48	18	18	37	29	-	22
R097	UAE_R	35	36	-	-	21	24	24.1	-	-	10	13	17	-	49	13.2	16	-	27	40	17	39	48	18	18	37	29	-	19
R098	UAE_R	35	38	-	-	20	25	-	-	-	12	15	17	-	48	14	15	-	26	34	13	38	49	18	19	37	31	-	22
R099	UAE_R	38	40	-	-	20	21	23.1	-	-	11	14	18	-	54	13	16	-	26	35	14	32	49	17	18	37	30	-	20
R100	UAE_R	35	36	-	-	19	-	-	-	-	15	16	18	-	50	14	15	-	28	41	17	34	48	17	18	36	29	-	21
R101	UAE_R	37	-	-	-	20	21	22.1	-	-	13	16	-	-	52	14	16	-	30	41	15	39	48	18	15	37	30	-	19
R102	UAE_R	37	-	-	-	20	23	25.1	-	-	12	13	14	-	49	13	17	-	31	40	12	34	49	17	17	38	32	-	20
R103	UAE_R	39	-	-	-	22	23	26.1	-	-	14	15	16	-	55	15	-	-	26	39	12	35	47	17	18	38	30	-	18
R104	UAE_R	38	-	-	-	22	28.1	-	-	-	9	13	18	-	53	13	-	-	26	39	17	36	45	18	18	39	31	-	21
R105	UAE_R	36	37	-	-	20	25	-	-	-	13	15	16	-	53	14	15	-	26	35	12	32	47	19	17	34	32	-	20
R106	UAE_R	37	-	-	-	21	23	25.1	-	-	11	13	18	-	51	13	15	-	28	39	12	37	49	18	17	35	29	-	19
R107	UAE_R	36	37	-	-	20	22	22.1	-	-	13	15	17	-	52	14	15	-	29	39	17	39	50	18	15	35	29	-	18
R108	UAE_R	35	36	-	-	21.1	23	25.1	-	-	12	14	18	-	52	15	-	-	32	37	14	39	49	18	18	35	30	-	20

ID	Population	DYF387S1				DYF399S1					DYF403S1				DYF403S1b	DYF404S1			DYS449	DYS518	DYS526a	DYS526b	DYS547	DYS570	DYS576	DYS612	DYS626		DYS627
R109	UAE_R	37	39	-	-	20	22	24	-	-	13	16	-	-	49	14	15	-	33	39	13	35	49	21	16	38	30	-	18
R110	UAE_R	36	38	-	-	21	22	23.1	-	-	13	15	16	-	49	12	15	-	32	36	17	34	48	17	15	38	30	-	22
R111	UAE_R	38	-	-	-	20	21	22.1	-	-	12	14	15	-	54	13.2	14	-	31	34	13	33	47	18	18	39	29	-	19
R112	UAE_R	36	37	-	-	20	22.1	23	-	-	13	14	16	17	49	14	15	-	27	40	17	33	49	17	16	33	30	-	19
R113	UAE_R	37	-	-	-	20	21	22.1	-	-	14	16	-	-	50	13	17	-	26	38	17	33	47	17	16	34	32	-	22
R114	UAE_R	37	-	-	-	22	24.1	25	-	-	11	13	16	-	48	14	15	-	27	39	13	34	49	17	18	37	30	-	20
R115	UAE_R	36	37	-	-	20.1	21	27	-	-	13	16	-	-	45	14	-	-	26	41	15	34	45	19	17	36	33	-	21
R116	UAE_R	37	-	-	-	21	22	24	-	-	11	15	17	-	51	14	15	-	32	39	16	38	48	18	18	34	30	-	21
R117	UAE_R	36	37	-	-	18	23	26.1	-	-	11	14	-	-	54	14	15	-	25	37	17	38	51	19	17	35	26	-	19
R118	UAE_R	37	38	-	-	21	23	24.1	-	-	15	16	19	-	53	14	-	-	32	37	17	39	47	16	16	37	32	-	17
R119	UAE_R	37	-	-	-	22.1	23	25.1	-	-	10	15	-	-	50	15	17	-	29	43	16	35	49	22	18	36	30	-	20
R120	UAE_R	35	40	-	-	21	24	-	-	-	14	16	17	-	49	13	16	-	28	39	13	32	51	18	19	37	30	-	20
R121	UAE_R	38	39	-	-	21	23	23.1	-	-	13	15	18	-	46	15	-	-	35	39	15	38	48	19	18	37	31	-	21
R122	UAE_R	40	-	-	-	18	23	24	-	-	13	14	16	-	46	15	-	-	25	39	15	39	48	18	18	38	30	-	19
R123	UAE_R	40	-	-	-	21	24	24.1	-	-	10	17	-	-	50	12	15	-	33	40	15	39	47	18	19	36	25	-	18
R124	UAE_R	38	39	-	-	19	24	26	-	-	13	15	17	-	54	14	15	-	33	38	15	38	48	18	17	38	25	-	18
R125	UAE_R	38	40	-	-	20	24	24.1	-	-	11	13	17	-	53	13	17	-	33	39	17	36	48	19	19	35	25	-	18
R126	UAE_R	36	39	-	-	22	26.1	28	-	-	12	17	-	-	54	14.2	-	-	28	38	12	35	47	19	16	38	32	-	22
R127	UAE_R	39	-	-	-	19	22	24.1	-	-	11	15	17	-	50	15	-	-	27	43	13	37	50	19	16	34	32	-	21
R128	UAE_R	36	-	-	-	20	20.1	23	-	-	11	13	14	-	50	14	-	-	31	39	12	37	48	18	17	36	31	-	21
R129	UAE_R	36	37	-	-	19	23.1	25	-	-	14	15	16	-	51	12	14	-	29	35	13	34	49	18	16	38	31	-	21
R130	UAE_R	39	-	-	-	21	24	25.1	-	-	13	14	15	-	48	15	16	-	26	40	17	35	48	19	17	37	24	-	18
R131	UAE_R	35	39	-	-	20	24	26	-	-	11	15	17	-	50	15	17	-	26	40	16	42	51	18	17	36	32	-	21
R132	UAE_R	38	39	-	-	19	20	21	21.1	-	11	13	14	-	54	13	-	-	32	39	17	39	49	15	14	34	32	-	19
R133	UAE_R	37	39	-	-	20	21	22.1	-	-	14	15	16	-	45	13	14	-	27	38	14	32	48	18	16	36	29	-	22
R134	UAE_R	36	39	-	-	21	21.1	25	-	-	13	14	15	-	51	14	15	-	31	40	14	31	49	18	16	37	32	-	17
R135	UAE_R	38	40	-	-	20	21	22.1	-	-	13	15	17	-	51	13	14	-	26	38	12	34	52	19	18	36	35	-	20
R136	UAE_R	35	36	-	-	20	23	24.1	-	-	11	15	16	-	54	13	17	-	26	40	13	39	50	21	18	36	25	-	20
R137	UAE_R	37	-	-	-	20	23	24.1	-	-	10	13	17	-	55	14	-	-	26	41	12	40	48	18	18	36	29	-	18
R138	UAE_R	37	38	-	-	23	24.1	-	-	-	15	18	19	-	53	14	-	-	31	42	17	38	49	17	17	36	30	-	20
R139	UAE_R	38	-	-	-	23	26.1	-	-	-	11	13	14	-	49	14	16	-	26	39	12	33	52	18	18	38	25	-	18
R140	UAE_R	32	33	-	-	20	21	21.1	-	-	15	16	-	-	49	14	16	-	35	40	14	35	45	19	18	39	31	-	22
R141	UAE_R	36	37	-	-	19	21	23.1	-	-	13	15	17	-	50	15	16	-	30	40	12	31	45	17	17	35	31	-	21
R142	UAE_R	37	40	-	-	20	21	21.1	-	-	11	13	17	-	51	15	-	-	31	35	17	36	52	19	18	35	32	-	23

ID	Population	DYF387S1				DYF399S1					DYF403S1				DYF403S1b	DYF404S1			DYS449	DYS518	DYS526a	DYS526b	DYS547	DYS570	DYS576	DYS612	DYS626		DYS627
R143	UAE_R	36	41	-	-	18	23	25.1	-	-	11	15	-	-	48	14	16	-	33	37	17	38	45	18	18	37	25	-	20
R144	UAE_R	35	36	-	-	20	21	23.1	-	-	12	16	-	-	53	14	15	-	26	37	17	38	47	19	16	38	31	-	19
R145	UAE_R	36	40	-	-	21	22	22.1	-	-	14	17	-	-	52	14	15	-	26	39	14	38	49	19	18	39	31	-	17
R146	UAE_R	37	40	-	-	20	21	23.1	-	-	14	17	-	-	54	15	-	-	26	38	15	38	48	20	18	38	30	-	18
R147	UAE_R	37	39	-	-	20	21	23.1	-	-	14	15	17	-	54	14	-	-	28	39	14	34	46	18	17	39	29	-	21
R148	UAE_R	37	39	-	-	20	21	23.1	-	-	9	11	16	-	55	14	16	-	31	38	17	34	46	19	18	37	31	-	23
R149	UAE_R	37	-	-	-	21	24	24.1	-	-	10	14	18	-	49	14	-	-	32	41	17	33	48	17	18	34	30	-	19
R150	UAE_R	37	40	41	-	21	-	-	-	-	12	14	17	-	49	14	15	-	33	39	16	39	48	19	18	37	33	-	17
R151	UAE_R	37	-	-	-	19.1	24	24	-	-	11	13	18	-	51	12	16	-	35	39	17	39	49	18	19	35	32	-	17
R152	UAE_R	37	38	-	-	19	21	23.1	-	-	13	15	17	-	51	14	15	-	26	40	15	39	48	18	19	35	25	-	20
R153	UAE_R	37	-	-	-	20	21	23.1	-	-	12	14	18	-	53	14	15	-	34	42	17	40	45	18	17	37	31	-	22
R154	UAE_R	37	38	-	-	21	22.1	23.1	-	-	13	16	-	-	49	14	15	-	32	39	15	39	49	18	18	34	30	-	20
R155	UAE_R	37	-	-	-	20	23	24.1	-	-	13	15	16	-	49	13	15	-	32	40	13	38	49	17	16	36	26	-	21
R156	UAE_R	38	39	-	-	19	22	25.1	-	-	12	14	15	-	48	15	16	-	26	40	17	39	48	17	14	36	26	-	20
R157	UAE_R	38	40	-	-	21	22.1	-	-	-	13	14	16	17	54	14	15	-	26	37	17	39	48	18	18	37	32	-	23
R158	UAE_R	36	-	-	-	21.1	22	-	-	-	13	16	-	-	50	15	17	-	26	38	15	39	46	19	18	36	30	-	21
R159	UAE_R	37	-	-	-	21.1	22	-	-	-	12	13	-	-	52	14	15	-	29	39	14	38	49	19	18	38	25	-	19
R160	UAE_R	38	-	-	-	21	22.1	23.1	-	-	13	15	19	-	49	14	15	-	32	39	13	34	47	18	16	37	33	-	21
R161	UAE_R	36	37	-	-	20	21	23.1	-	-	12	14	-	-	55	14	16	-	32	39	13	40	48	19	18	38	28	-	18
R162	UAE_R	37	-	-	-	20	21	23.1	-	-	14	15	16	-	53	15	17	-	27	38	17	33	48	18	17	34	30	-	21
R163	UAE_R	35	39	-	-	18	21	26.1	-	-	14	15	16	-	53	14	15	-	26	41	12	34	49	16	17	36	29	-	20
R164	UAE_R	37	38	-	-	22	23.1	24	-	-	13	15	16	-	51	14	15	-	34	37	17	34	48	19	18	37	30	-	21
R165	UAE_R	35	38	-	-	22	24	25	-	-	12	13	-	-	52	14	15	-	26	38	15	34	52	18	18	36	31	-	17
R166	UAE_R	35	37	-	-	20	22	23	-	-	14	16	17	-	52	14	-	-	27	39	15	40	48	18	18	37	33	-	19
R167	UAE_R	39	40	-	-	21	22	25.1	-	-	13	15	17	-	49	14	15	-	33	39	14	35	48	18	18	36	30	-	20
R168	UAE_R	35	41	-	-	21	24	24.1	-	-	13	14	16	-	49	14	15	-	25	39	12	32	48	19	17	39	25	-	21
R169	UAE_R	37	-	-	-	20	22	23.1	-	-	11	15	17	-	54	14	15	-	34	39	14	39	50	17	17	37	29	-	21
R170	UAE_R	37	38	-	-	20	24	25.1	-	-	12	16	-	-	49	15	-	-	26	35	17	39	47	18	17	37	30	-	19
R171	UAE_R	37	-	-	-	22	23	24.1	-	-	9	13	16	-	54	14	15	-	31	39	17	40	50	17	18	36	30	-	18
R172	UAE_R	38	-	-	-	20	26	-	-	-	13	15	16	-	50	13	15	-	32	37	15	39	49	18	18	38	30	-	20
R173	UAE_R	35	36	-	-	22.1	23	25.1	-	-	11	12	16	-	49	13.2	14	-	25	38	15	33	46	19	16	34	30	-	19
R174	UAE_R	38	40	-	-	21	24	24.1	-	-	11	15	17	-	52	14	15	-	35	45	16	33	49	18	18	38	30	-	20
R175	UAE_R	37	38	-	-	20	22	22.1	-	-	11	14	18	-	49	14	15	-	34	41	17	32	47	19	17	37	31	-	18
R176	UAE_R	36	37	-	-	20	22.1	-	-	-	13	14	16	-	50	14	16	-	33	38	12	38	49	17	18	35	30	-	20

ID	Population	DYF387S1				DYF399S1					DYF403S1				DYF403S1b	DYF404S1			DYS449	DYS518	DYS526a	DYS526b	DYS547	DYS570	DYS576	DYS612	DYS626		DYS627
R177	UAE_R	37	-	-	-	22	23	24.1	-	-	13	15	17	-	50	14	15	-	26	38	13	35	47	18	17	36	32	-	22
R178	UAE_R	36	-	-	-	22	23	23.1	24	24.1	10	16	-	-	49	14	15	-	28	40	17	35	48	21	18	35	32	-	22
R179	UAE_R	36	39	-	-	17.2	21	21.1	-	-	11	15	16	-	45	13	17	-	26	42	13	33	48	18	17	34	30	-	21
R180	UAE_R	40	-	-	-	17.2	21	21.1	-	-	13	15	17	-	49	13	17	-	26	41	16	34	48	19	19	38	30	-	20
R181	UAE_R	38	-	-	-	21	22	24.1	-	-	13	15	17	-	48	14	16	-	32	38	17	39	48	18	15	36	32	-	16
R182	UAE_R	38	-	-	-	21	22.1	-	-	-	13	15	16	-	48	13	15	-	27	35	17	33	48	18	16	38	30	-	20
R183	UAE_R	36	37	-	-	20	21	21.1	-	-	12	13	15	-	50	16	-	-	31	42	13	34	48	18	17	37	30	-	22
R184	UAE_R	36	37	-	-	23	24	26.1	-	-	10	17	18	-	48	14	15	-	26	38	14	37	48	18	15	36	30	-	20
R185	UAE_R	36	37	-	-	23	25.1	-	-	-	10	14	17	-	50	15	-	-	26	38	13	39	48	20	18	36	31	-	23
R186	UAE_R	35	36	-	-	21	23.1	-	-	-	13	15	17	-	49	14	15	-	26	38	13	39	47	18	19	37	30	-	20
R187	UAE_R	38	-	-	-	20	21	22.1	-	-	13	14	17	-	49	15	16	-	31	40	13	41	48	20	12	38	29	-	19
R188	UAE_R	37	-	-	-	18	23	24.1	-	-	11	14	-	-	49	14	15	-	26	40	14	35	49	19	18	36	30	-	20
R189	UAE_R	37	-	-	-	20	24	27	-	-	15	16	19	-	49	14	16	-	35	37	14	33	48	18	18	36	29	-	19
R190	UAE_R	35	36	-	-	21	23.1	-	-	-	10	15	-	-	50	14	15	-	30	37	14	40	48	15	18	36	29	-	19
R191	UAE_R	39	-	-	-	20	23	23.1	-	-	14	16	17	-	48	14	-	-	31	39	13	32	50	18	17	37	31	-	20
R192	UAE_R	37	-	-	-	21	23.1	-	-	-	13	15	18	-	47	12	16	-	33	38	13	33	48	18	18	35	26	-	21
R193	UAE_R	39	-	-	-	21	21.1	22	-	-	13	14	16	-	53	14	15	-	26	39	13	36	46	18	16	37	30	-	20
R194	UAE_R	36	38	-	-	21	24	25.1	-	-	10	17	-	-	48	15	16	-	26	39	12	34	48	17	18	36	30	-	21
R195	UAE_R	37	38	-	-	20	21	22	22.1	-	13	15	17	-	53	13	15	-	26	42	18	33	49	18	17	37	31	-	21
R196	UAE_R	36	37	-	-	19	21	22.1	-	-	11	13	17	-	51	14	16	-	28	39	15	36	49	19	16	36	34	-	20
R197	UAE_R	39	-	-	-	20	21	22.1	-	-	12	17	-	-	50	14	15	-	26	40	13	39	48	19	16	36	28	-	19
R198	UAE_R	35	38	-	-	21	23.1	-	-	-	14	15	17	-	54	15	16	-	29	38	12	41	44	19	17	36	33	-	17
R199	UAE_R	38	39	-	-	20	21	21.1	-	-	12	15	-	-	54	15	-	-	26	38	14	39	45	20	18	38	32	-	17
R200	UAE_R	38	39	-	-	21	23	25.1	-	-	11	13	14	-	47	14	15	-	28	38	15	41	50	19	18	37	32	-	16
F130	Indian	38	-	-	-	18	25.1	-	-	-	9	17	-	-	45	14	-	-	30	38	13	37	49	18	16	39	30	-	19
F131	Indian	37	-	-	-	22.1	23.2	-	-	-	12	13	14	-	48	13	-	-	31	38	15	38	49	16	17	37	32	-	21
F132	Indian	37	39	-	-	20.1	22	25	-	-	9	14	17	-	51	14	15	-	31	41	15	41	48	17	19	37	31	-	18
F133	Indian	38	39	-	-	22.2	25	27.1	-	-	13	-	-	-	47	14	15	-	30	38	15	34	45	17	17	35	31	-	21
F134	Indian	36	38	-	-	23	25	-	-	-	14	15	-	-	51	15	16	-	31	40	13	37	44	19	17	39	31	-	22
F135	Indian	39	-	-	-	20.1	23	26	-	-	11	14	16	-	50	14	15	-	33	39	15	33	49	18	17	36	32	-	17
F136	Indian	37	-	-	-	20	22	24	-	-	13	15	16	-	50	13	16	-	32	43	14	35	46	18	14	36	31	-	21
F137	Indian	37	-	-	-	21.1	22.2	25	-	-	12	13	17	-	48	14	-	-	32	39	14	36	47	18	18	36	32	-	18
F138	Indian	36	39	-	-	21.1	22	23	-	-	12	13	16	-	49	14	16	-	32	41	15	39	48	20	19	-	34	-	18
F139	Indian	34	38	-	-	20	25	25.1	-	-	11	13	-	-	54	13	-	-	33	36	15	38	50	19	18	36	28	-	18

ID	Population	DYF387S1				DYF399S1					DYF403S1				DYF403S1b	DYF404S1			DYS449	DYS518	DYS526a	DYS526b	DYS547	DYS570	DYS576	DYS612	DYS626		DYS627
F140	Indian	37	40	-	-	22	23	27.1	-	-	13	15	16	-	48	16	17	-	31	41	13	34	50	16	16	40	33	-	20
F142	Indian	36	39	-	-	21	22	24.1	-	-	13	16	17	-	47	15	16	-	31	40	13	35	47	15	17	38	30	-	18
F144	Indian	36	38	-	-	22	23.1	24.1	-	-	12	15	-	-	48	16	17	-	25	39	13	36	46	19	16	39	31	-	19
F145	Indian	38	39	-	-	21	22	23.1	-	-	11	13	15	-	51	13	-	-	32	38	13	34	50	18	17	36	29	-	22
F148	Indian	38	40	-	-	23	23.1	24.2	-	-	13	15	-	-	49	13	14	-	31	39	15	36	46	20	14	37	33	-	18
F149	Indian	37	38	-	-	22	23.1	25	-	-	14	16	-	-	47	15	16	-	32	40	14	39	44	18	16	38	31	-	19
F150	Indian	39	-	-	-	22.1	23.2	24	-	-	12	14	17	-	48	14	-	-	31	38	15	37	47	18	15	38	32	-	17
F151	Indian	37	40	-	-	23	24.1	24.2	-	-	9	13	16	-	49	13	14	-	32	38	15	39	46	17	17	38	32	-	18
F152	Indian	38	-	-	-	20.1	21.1	-	-	-	11	12	16	-	41	15	16	-	30	40	16	39	46	17	19	41	29	-	19
F153	Indian	37	39	-	-	21	22.1	24	-	-	12	14	17	-	50	14	-	-	30	40	14	37	47	19	18	-	30	-	17
F155	Indian	37	38	-	-	20	24	24.1	-	-	10	13	15	-	52	14	15	-	30	38	15	38	46	17	18	32	28	-	22
F156	Indian	36	41	-	-	21	21.1	24	-	-	11	14	15	-	53	12	13	-	30	38	14	38	49	16	15	35	30	-	18
F157	Indian	36	39	-	-	21	21.1	23	-	-	12	16	18	-	49	14	17	-	30	40	12	34	47	18	19	37	30	-	20
F158	Indian	35	38	-	-	20	24.1	25	-	-	12	14	15	-	51	13	14	-	32	36	15	37	49	19	17	39	28	-	19
F159	Indian	36	38	-	-	21	22	25.1	-	-	10	16	-	-	50	15	-	-	32	39	12	32	51	17	16	36	33	-	19
F160	Indian	36	39	-	-	20	22	-	-	-	12	14	-	-	46	13	16	-	35	41	13	37	47	17	13	37	32	-	24
F161	Indian	37	38	-	-	21	22	24.1	-	-	10	14	16	-	49	15	17	-	33	45	16	39	49	20	18	37	32	-	17
F162	Indian	38	39	-	-	20	23	24	-	-	11	13	17	-	52	14	-	-	31	40	14	33	48	18	18	37	32	-	20
F163	Indian	38	39	-	-	20	23	24	-	-	11	13	17	-	52	14	-	-	31	40	14	38	48	18	18	33	32	-	20
F164	Indian	38	-	-	-	18	21	24	-	-	12	13	14	-	50	15	17	-	31	38	13	38	47	15	18	36	31	-	17
F165	Indian	38	-	-	-	22	24.1	24.2	-	-	12	13	18	-	48	12	16	-	32	37	15	36	47	19	16	37	32	-	21
F166	Indian	37	41	-	-	22	24.1	-	-	-	13	14	-	-	53	12	14	-	27	39	14	38	47	15	17	37	30	-	21
F167	Indian	37	-	-	-	21	23.1	26	-	-	10	13	16	-	50	15	16	-	32	39	15	38	50	19	18	34	32	-	18
F168	Indian	37	39	-	-	22.1	23	24.1	-	-	9	13	16	-	49	15	17	-	26	40	14	37	47	20	17	38	31	-	20
F170	Indian	35	40	-	-	19.1	23.2	26	-	-	13	18	-	-	47	13	15	-	30	39	15	37	45	18	18	37	32	-	19
F171	Indian	36	39	-	-	21.1	22	26	-	-	10	13	16	-	51	14	-	-	33	44	14	38	48	18	18	-	33	-	17
F172	Indian	36	37	-	-	22	23	25.1	-	-	11	12	14	-	51	13	14	-	29	37	15	38	44	20	16	35	33	-	21
F173	Indian	37	38	-	-	23	23.1	24.2	-	-	11	13	-	-	48	14	-	-	32	40	15	38	45	17	17	35	32	-	18
F174	Indian	39	-	-	-	21.2	22.1	-	-	-	13	14	-	-	47	14	-	-	32	39	15	38	46	17	18	37	32	-	19
F175	Indian	35	40	-	-	21	23	27.1	-	-	12	14	-	-	48	14	15	-	27	38	14	34	47	15	16	36	33	-	22
F176	Indian	37	39	-	-	21	21.1	23	-	-	12	14	17	-	50	14	16	-	33	41	15	41	49	20	20	33	35	-	15
F177	Indian	35	36	37	-	22	23	23.1	-	-	11	13	14	-	46	14	15	-	31	39	13	36	48	17	19	36	31	-	20
F178	Indian	37	39	-	-	21.1	22	25	-	-	12	16	18	-	52	14	16	-	32	40	16	40	49	19	18	35	32	-	16
F179	Indian	37	38	-	-	21	22	25.1	-	-	13	16	17	-	53	13	15	-	29	37	14	37	50	18	19	36	31	-	22

ID	Population	DYF387S1				DYF399S1					DYF403S1				DYF403S1b	DYF404S1			DYS449	DYS518	DYS526a	DYS526b	DYS547	DYS570	DYS576	DYS612	DYS626		DYS627
F180	Indian	37	38	-	-	23	23.1	-	-	-	12	13	16	-	51	14	16	-	31	39	15	40	46	20	16	36	32	-	21
F182	Indian	37	40	-	-	21.2	22.1	25	-	-	15	16	17	18	46	12.2	14	-	27	36	13	36	48	19	22	40	29	-	24
F183	Indian	37	-	-	-	22.1	23.2	25	-	-	12	13	16	-	52	13	15	-	33	39	15	37	46	18	17	36	32	-	20
F184	Indian	36	40	-	-	19.1	20	26	-	-	10	13	-	-	51	14	15	-	31	42	15	39	45	19	18	36	33	-	18
F185	Indian	36	38	-	-	21	21.1	-	-	-	10	11	15	-	48	13	14	-	29	38	13	38	50	19	16	36	30	-	19
F186	Indian	37	42	-	-	21	22	25.1	-	-	10	13	15	-	49	12	13	-	29	35	15	39	50	15	15	35	31	-	23
F187	Indian	36	39	-	-	20	22	24.1	-	-	10	11	12	-	53	12	16	-	31	40	14	37	47	19	16	36	30	-	20
F188	Indian	38	39	-	-	17.2	24	25	-	-	12	13	18	-	48	15.2	17	-	28	38	13	34	47	17	16	36	29	-	18
F189	Indian	35	37	-	-	22	23	25.1	-	-	10	15	16	-	49	14	-	-	33	36	12	32	50	19	18	34	26	-	18
F190	Indian	37	38	-	-	20	23	26.1	-	-	11	13	16	-	52	13	15	-	34	41	15	39	48	19	16	36	32	-	16
F191	Indian	39	40	-	-	23	25.1	-	-	-	13	17	18	-	48	13	16	-	30	38	14	37	50	15	19	34	32	-	20
F192	Indian	37	40	-	-	21	24	24.1	-	-	13	15	22	-	52	14	-	-	32	39	14	37	47	19	21	36	29	-	21
F193	Indian	37	40	-	-	24	28.1	-	-	-	13	14	17	-	46	14	15	-	32	39	13	34	49	19	17	37	32	-	19
F194	Indian	35	40	-	-	21	22	28.1	-	-	13	15	-	-	52	14	17	-	33	38	13	35	48	17	17	36	32	-	19
F195	Indian	37	38	-	-	22	24	24.1	-	-	11	14	15	-	51	14	-	-	32	36	12	34	49	15	16	35	31	-	18
F196	Indian	37	39	-	-	24	26.1	-	-	-	12	16	-	-	48	13	15	-	31	39	14	37	45	16	16	38	30	-	21
F197	Indian	36	38	-	-	20	24	25.1	-	-	13	15	17	-	48	16	17	-	33	38	13	37	47	20	18	37	30	-	22
F198	Indian	37	40	-	-	20.1	22	25	-	-	10	14	17	-	50	14	-	-	32	41	16	39	46	19	18	40	33	-	17
F199	Indian	37	38	-	-	20	23	24.1	-	-	10	13	16	-	47	13	14	-	27	37	12	35	49	16	18	37	29	-	20
F200	Indian	38	-	-	-	20	21	24	26.1	-	9	13	17	-	48	14	15	17	30	40	15	40	49	19	17	38	31	-	17
F201	Indian	36	37	-	-	21	21.1	22	-	-	11	15	16	-	50	15	-	-	34	41	15	39	46	19	21	40	32	-	18
F202	Indian	38	-	-	-	24	25.1	-	-	-	12	15	16	-	50	13	-	-	30	40	14	35	49	18	19	37	31	-	19
F203	Indian	38	40	-	-	21	24	26	-	-	12	14	16	-	52	14	-	-	34	38	13	40	49	17	18	38	30	-	18
F204	Indian	35	40	-	-	22	24.1	-	-	-	12	14	15	-	53	12	14	-	28	39	14	39	48	15	16	33	32	-	23
F205	Indian	35	37	-	-	22	24.1	25	-	-	11	13	15	-	52	12	14	-	32	35	15	37	50	18	18	38	28	-	19
F206	Indian	34	39	-	-	20	26	26.1	-	-	14	15	-	-	53	13	-	-	33	34	15	32	47	17	18	36	29	-	18
F207	Indian	36	40	-	-	22	23	25.1	-	-	13	14	-	-	52	12	14	-	28	37	14	38	49	15	17	36	30	-	24
F208	Indian	37	-	-	-	20	24	28.1	-	-	13	14	19	-	48	15	-	-	31	38	13	38	49	19	16	37	30	-	19
F209	Indian	36	39	-	-	21	24.1	25	-	-	12	14	15	-	53	13	-	-	34	39	15	37	46	16	19	38	28	-	18
F210	Indian	36	38	-	-	20.1	22	24	-	-	10	13	17	-	50	14	15	-	32	41	16	36	48	19	19	37	31	-	19
F211	Indian	35	38	-	-	23	23.1	24	-	-	11	13	-	-	48	13	14	-	30	38	14	36	51	13	17	36	32	-	19
F212	Indian	36	38	-	-	23.1	24	-	-	-	15	16	-	-	45	13	14	-	32	37	12	37	44	14	16	39	30	-	23
F213	Indian	36	40	-	-	19	24	25.1	-	-	12	14	17	-	49	15	16	-	32	40	13	41	50	16	17	38	28	-	19
F216	Indian	38	-	-	-	23	24.1	-	-	-	9	11	15	-	53	15	-	-	33	43	15	39	47	19	18	36	32	-	18

ID	Population	DYF387S1				DYF399S1					DYF403S1				DYF403S1b	DYF404S1			DYS449	DYS518	DYS526a	DYS526b	DYS547	DYS570	DYS576	DYS612	DYS626		DYS627
F217	Indian	36	39	-	-	19	27	29.1	-	-	14	15	16	-	50	15	-	-	29	39	13	37	50	18	16	34	31	-	23
F219	Indian	37	38	-	-	21	21.1	25	-	-	11	14	17	-	51	15	-	-	32	42	15	36	48	23	19	36	32	-	16
F220	Indian	38	39	-	-	24	25.1	-	-	-	12	15	16	-	48	15	-	-	31	38	14	37	46	19	17	37	29	-	21
F221	Indian	38	-	-	-	21.1	22	24	-	-	10	14	17	-	50	14	15	-	32	39	15	39	47	19	18	36	33	-	17
F222	Indian	37	38	-	-	21	22	23	23.1	-	12	16	-	-	48	15	-	-	281	39	12	35	50	18	18	35	29	-	23
F223	Indian	37	38	-	-	21	21.1	25	-	-	11	14	17	-	50	14	16	-	32	42	14	37	49	19	18	-	33	-	17
F224	Indian	35	39	-	-	21	24.1	-	-	-	12	13	15	-	50	14	15	-	26	39	14	37	47	15	16	38	32	-	20
F225	Indian	37	-	-	-	22.2	23	23.1	-	-	8	14	-	-	47	15	16	-	32	40	15	39	46	16	15	34	34	-	16
F226	Indian	36	39	-	-	22	24.1	-	-	-	12	16	-	-	45	13	14	-	32	39	13	40	48	19	18	37	27	-	16
F227	Indian	36	-	-	-	20	25.1	26	-	-	10	15	17	-	49	13	15	-	31	39	13	33	47	16	18	40	34	-	17
F228	Indian	37	39	-	-	22	23	24.1	-	-	12	14	17	-	50	14	15	-	30	42	15	41	48	19	18	36	33	-	18
F229	Indian	37	39	-	-	20.1	21	23	-	-	10	14	17	-	49	14	-	-	32	41	15	39	48	19	18	37	33	-	17
F230	Indian	36	39	-	-	19.1	20	22	-	-	9	13	17	-	50	14	15	-	33	40	15	38	49	19	16	37	33	-	17
F231	Indian	38	-	-	-	20	23	25.1	-	-	10	14	16	-	50	14	16	-	32	43	15	40	47	19	18	39	33	-	17
F232	Indian	38	-	-	-	23	24.1	-	-	-	10	11	16	-	51	14	-	-	32	40	15	41	50	18	16	36	34	-	18
F233	Indian	38	-	-	-	21	22	22.1	-	-	11	13	15	-	47	14	15	-	26	37	12	34	49	16	17	33	29	-	19
F235	Indian	36	37	-	-	22	23	27.1	-	-	11	17	-	-	49	15	17	-	33	38	13	35	47	18	18	36	32	-	19
F236	Indian	37	-	-	-	21	22.1	-	-	-	10	13	17	-	50	15	-	-	32	39	15	38	50	17	16	37	33	-	19
F237	Indian	39	-	-	-	22.1	23	-	-	-	11	13	14	-	51	14	15	-	27	38	14	39	48	15	19	37	32	-	22
F238	Indian	39	-	-	-	22	23.1	23.2	-	-	13	14	-	-	49	14	16	-	33	40	15	37	47	17	20	36	33	-	18
F239	Indian	36	39	-	-	22	23	24.1	-	-	13	16	17	-	48	15	-	-	32	41	13	36	47	20	18	37	31	-	23
F240	Indian	38	39	-	-	22	23	27	-	-	10	14	-	-	48	13	15	-	34	40	14	36	50	18	15	37	34	-	19
F241	Indian	36	37	-	-	19	21	22.1	-	-	14	15	-	-	52	13	15	-	25	37	17	40	49	18	17	36	30	-	21
F242	Indian	36	37	-	-	21	24	25.1	-	-	11	13	15	-	48	14	16	-	29	39	14	38	47	17	16	35	29	-	21
F243	Indian	37	39	-	-	22	23	25.1	-	-	13	18	-	-	49	16	17	-	33	39	13	35	49	17	18	36	30	-	20
F244	Indian	38	39	-	-	22.1	23	23.2	-	-	11	13	15	-	47	13	15	-	32	39	15	37	47	17	18	35	33	-	18
F245	Indian	38	-	-	-	23	23.2	24.1	-	-	13	16	-	-	48	12	16	-	32	38	15	36	48	18	16	36	32	-	19
F246	Indian	35	37	-	-	20	22	25.1	-	-	11	14	15	-	49	14	-	-	32	37	12	34	49	22	18	36	24	-	17
F247	Indian	38	40	-	-	19	21	26	-	-	12	13	15	-	49	14	15	-	30	39	13	36	47	16	18	35	32	-	21
F248	Indian	39	40	-	-	21.2	22	22.1	-	-	12	13	18	-	48	13	14	-	32	40	15	37	47	17	18	37	29	-	18
F249	Indian	36	39	-	-	20	23	24	-	-	12	16	-	-	49	15	-	-	33	39	12	33	50	17	17	34	31	-	19
F250	Indian	38	-	-	-	21	24	25.1	-	-	10	14	17	-	50	14	16	-	32	42	15	39	49	19	18	37	32	-	17
M001	Indian	37	38	-	-	18.1	21	25	-	-	12	14	17	-	50	13	14	-	32	41	15	40	49	21	18	37	34	-	17
E035	Indian	37	38	-	-	21	22	22.1	-	-	10	11	16	-	47	13	14	-	29	37	12	35	50	17	16	37	29	-	19

ID	Population	DYF387S1				DYF399S1					DYF403S1				DYF403S1b	DYF404S1			DYS449	DYS518	DYS526a	DYS526b	DYS547	DYS570	DYS576	DYS612	DYS626		DYS627
E036	Indian	39	-	-	-	20	22	24.1	-	-	12	13	15	-	49	15	-	-	29	40	13	33	46	19	19	37	29	-	21
M004	Indian	36	39	-	-	19	22	26.1	-	-	14	15	18	-	49	14	16	-	32	37	13	35	47	18	17	39	30	-	22
M005	Indian	35	39	-	-	19.1	21	24	-	-	10	13	17	-	54	14	15	-	32	42	15	41	48	19	18	36	32	-	17
M006	Indian	35	37	-	-	21	22	26.1	-	-	14	16	17	-	49	15	16	-	32	39	13	35	46	18	17	38	31	-	21
M007	Indian	38	39	-	-	18	19	23	-	-	10	12	14	-	49	14	15	-	30	38	13	35	49	16	17	33	32	-	21
M008	Indian	37	41	-	-	20.2	23	24.1	-	-	14	17	-	-	47	13	15	-	30	37	15	37	46	18	16	38	32	-	18
M009	Indian	36	38	-	-	20.1	21	25	-	-	10	13	18	-	50	15	16	-	31	40	16	41	49	19	18	38	35	-	17
M010	Indian	35	38	-	-	22	22.1	26	-	-	12	14	15	-	53	13	14	-	34	36	15	37	48	18	19	38	28	-	18
M011	Indian	39	-	-	-	21	21.1	23	-	-	11	14	18	-	48	15	-	-	33	44	16	40	47	18	18	-	32	-	17
M012	Indian	38	-	-	-	20.1	22	24	-	-	10	14	17	-	49	14	15	-	32	41	16	40	51	20	18	36	33	-	16
M013	Indian	37	38	-	-	19.1	19.2	22	24	-	15	17	-	-	48	13.2	15	-	28	38	13	35	48	17	17	37	31	-	20
M014	Indian	36	39	-	-	23	24.1	24.2	-	-	12	13	16	-	45	13	15	-	30	38	15	36	46	19	19	36	32	-	18
M015	Indian	37	39	-	-	21	23	23.1	-	-	11	14	17	-	49	15	-	-	31	42	15	38	48	18	18	-	30	-	17
E037	Indian	39	40	-	-	22	23.1	-	-	-	12	16	18	-	50	15	-	-	32	39	15	39	45	18	19	37	32	-	17
M017	Indian	37	40	-	-	22	24	25.1	-	-	11	12	15	-	47	15	-	-	30	40	14	37	50	20	17	37	31	-	20
M018	Indian	36	37	-	-	21.1	22.1	25	-	-	9	12	16	-	42	13	15	-	26	38	15	38	48	19	18	37	30	-	22
M019	Indian	37	38	-	-	22	23	28.1	-	-	13	14	15	-	46	16	-	-	29	38	13	35	50	14	18	34	27	-	21
M022	Indian	38	-	-	-	21	23	25.1	-	-	10	12	15	-	50	13	15	-	30	39	15	37	42	17	17	32	27	-	21
M023	Indian	35	39	-	-	22	24	27.1	-	-	14	16	20	-	49	17	18	-	32	39	14	36	49	18	18	37	28	-	19
M024	Indian	37	38	-	-	23	24	-	-	-	10	12	16	-	50	14	-	-	32	40	15	37	48	18	18	35	31	-	17
M025	Indian	38	39	-	-	17	21	22	-	-	12	13	14	-	49	15	16	-	31	39	13	33	47	15	19	36	31	-	21
M026	Indian	37	40	-	-	21	22.1	23.2	-	-	10	13	15	-	49	14	15	-	31	39	15	37	45	19	17	35	34	-	21
M027	Indian	33	37	-	-	21	21.1	25	-	-	13	16	-	-	50	13	15	-	29	40	14	36	47	20	19	38	30	-	15
M028	Indian	38	39	-	-	20	21.1	23	-	-	9	15	16	-	50	15	-	-	32	42	15	39	49	19	18	36	32	-	17
M029	Indian	37	41	-	-	21.1	23	24	-	-	8	13	17	-	50	13	14	-	30	43	15	39	49	19	18	-	31	-	17
M030	Indian	38	-	-	-	20	23	23.1	-	-	11	12	15	-	49	13	15	-	29	39	16	38	47	17	17	37	27	-	22
M031	Indian	38	-	-	-	22	23	25	-	-	10	16	-	-	48	14	15	-	33	35	12	33	50	19	19	34	30	-	17
M032	Indian	37	-	-	-	22	24	26.1	-	-	9	15	16	-	50	14	15	-	33	42	15	38	48	19	19	35	33	-	17
M033	Indian	37	41	-	-	19	21	22.1	-	-	11	12	13	-	49	15	-	-	31	39	13	34	46	18	16	34	29	-	20
M035	Indian	38	40	-	-	21.1	23	24	-	-	10	13	17	-	49	14	15	-	33	41	15	40	49	17	18	36	32	-	18
M036	Indian	37	39	-	-	22	24	24.1	-	-	13	14	-	-	49	13	16	-	30	40	15	36	45	20	15	36	31	-	20
M038	Indian	38	-	-	-	20.1	22	26	-	-	10	14	17	-	50	14	15	-	34	40	15	40	50	18	18	36	34	-	18
M039	Indian	36	38	-	-	18	21.1	24	-	-	12	15	-	-	48	13	15	-	29	41	12	34	50	15	16	40	29	-	19
M040	Indian	36	38	-	-	20.1	21	26.1	-	-	10	17	-	-	48	13	-	-	28	37	12	33	49	17	16	37	31	-	19

ID	Population	DYF387S1				DYF399S1					DYF403S1				DYF403S1b	DYF404S1			DYS449	DYS518	DYS526a	DYS526b	DYS547	DYS570	DYS576	DYS612	DYS626		DYS627
M041	Indian	35	39	-	-	19	23	27.1	-	-	14	15	18	-	48	14	-	-	34	36	15	38	49	16	18	39	30	-	21
M042	Indian	38	-	-	-	23.1	24.2	-	-	-	12	13	-	-	49	15	-	-	31	40	14	37	46	16	17	37	31	-	19
M043	Indian	37	38	-	-	20	23	25.1	-	-	12	13	16	-	48	14	-	-	30	43	15	39	50	19	18	35	32	-	17
M044	Indian	37	39	-	-	21.1	22.2	24.1	-	-	10	12	20	-	50	13	16	-	32	38	15	38	45	18	20	36	31	-	21
M045	Indian	38	39	-	-	23	23.1	-	-	-	10	14	16	-	48	15	16	-	31	41	14	39	45	19	17	37	30	-	19
M046	Indian	37	39	-	-	21	21.1	26	-	-	11	16	17	-	50	14	-	-	32	41	14	39	48	19	19	-	32	-	17
E038	Indian	37	-	-	-	20	24	25.1	-	-	11	12	14	-	52	13	14	-	29	42	14	39	48	16	21	36	29	-	23
M049	Indian	34	40	-	-	19	23	24.1	-	-	11	14	15	-	50	15	-	-	26	38	14	39	47	15	17	38	33	-	20
M050	Indian	38	40	-	-	22	23	23.1	-	-	11	13	17	-	49	14	15	-	32	43	15	38	48	18	21	33	32	-	17
M051	Indian	36	39	-	-	20	23	24.1	-	-	10	15	18	-	49	14	-	-	34	40	13	35	48	17	18	35	31	-	21
M052	Indian	36	40	-	-	21	22	25.1	-	-	12	13	14	-	51	15	-	-	27	39	14	41	47	15	17	35	33	-	22
M053	Indian	35	36	-	-	24	26.1	-	-	-	12	16	-	-	47	14	-	-	30	37	14	37	48	18	16	37	29	-	22
M054	Indian	36	38	-	-	22	23.1	24	-	-	16	17	-	-	54	15	-	-	33	38	14	39	51	19	17	37	31	-	21
M055	Indian	35	38	-	-	21	24	-	-	-	9	10	14	-	45	13	15	-	32	38	13	35	50	21	18	36	31	-	21
E039	Indian	37	38	-	-	21	24	27.1	-	-	12	15	17	-	49	14	15	-	33	41	15	39	48	18	18	36	33	-	16
M057	Indian	36	40	-	-	18	21	22.1	-	-	10	13	16	-	50	13	14	-	32	39	13	33	45	16	16	35	32	-	19
M058	Indian	37	40	-	-	21	22.1	24	-	-	10	14	16	-	49	14	16	-	32	38	16	40	49	18	20	34	33	-	17
M060	Indian	37	-	-	-	22	24	25.1	-	-	9	14	16	-	49	13	15	-	27	40	14	34	51	16	17	36	30	-	15
M062	Indian	32	36	-	-	16	23	23.1	-	-	12	13	16	-	48.2	14	-	-	29	38	15	40	47	19	18	36	29	-	19
M063	Indian	37	38	-	-	20.1	21	23	-	-	10	13	17	-	49	14	15	-	33	41	15	38	47	19	19	38	32	-	17
M064	Indian	37	40	-	-	20	21.1	25	-	-	10	13	16	-	49	14	15	-	31	43	15	38	48	18	19	36	33	-	17
M065	Indian	36	39	-	-	21	22	23	-	-	13	14	16	-	51	12	15	-	27	38	14	36	50	16	16	38	29	-	21
M066	Indian	37	40	-	-	21	23	26.1	-	-	10	14	16	-	51	14	16	-	33	38	14	37	-	15	17	39	31	-	17
M067	Indian	36	41	-	-	22	24	30	-	-	12	15	17	-	49	14	15	-	34	40	13	35	50	14	16	36	30	-	21
M068	Indian	37	39	-	-	22	24	24.1	-	-	10	16	-	-	50	15	16	-	31	42	15	38	48	19	21	39	33	-	18
M072	Indian	35	39	-	-	22	23	24.1	-	-	14	16	-	-	47	14	15	-	32	39	15	38	44	18	19	39	29	-	19
M084	Indian	38	-	-	-	22	23.2	24.1	-	-	10	14	15	-	48	12	16	-	32	39	15	37	46	20	18	36	33	-	20
M085	Indian	38	41	-	-	20	22	24	-	-	13	15	16	-	49	14	16	-	26	38	14	36	49	15	17	35	29	-	18
M086	Indian	39	41	-	-	20	24	25.1	-	-	11	13	16	-	52	13	14	-	32	39	14	34	47	16	19	33	27	-	18
M087	Indian	38	-	-	-	22	23	23.1	-	-	10	13	16	-	49	14	16	-	32	38	16	40	51	17	18	36	31	-	16
M088	Indian	37	38.2	-	-	18	21	29	-	-	13	14	16	-	52	13	14	-	25	39	17	39	50	20	18	37	30	-	20
M089	Indian	37	39	-	-	20	23	27.1	-	-	13	15	16	-	49	14	15	-	33	41	13	35	47	15	17	35	30	-	19
E040	Indian	35	-	-	-	22	24	25.1	-	-	11	13	17	-	48	14	16	-	31	39	14	37	48	17	18	37	30	-	21
M091	Indian	36	38	-	-	22	22.1	23	-	-	11	12	17	-	-	13	16	-	27	37	14	38	47	19	18	33	30	-	21

ID	Population	DYF387S1				DYF399S1					DYF403S1				DYF403S1b	DYF404S1			DYS449	DYS518	DYS526a	DYS526b	DYS547	DYS570	DYS576	DYS612	DYS626		DYS627
M092	Indian	37	41	-	-	20	23	23.1	-	-	10	15	16	-	49	12	14	-	31	37	15	39	48	16	15	37	30	-	22
M093	Indian	35	39	-	-	23	25	26.1	-	-	12	14	15	-	47	15	-	-	32	39	15	38	45	18	18	39	29	-	21
M094	Indian	37	39	-	-	20	22	22.1	-	-	12	18	19	-	47	15	-	-	30	38	12	35	45	16	18	37	32	-	24
M095	Indian	36	37	38	-	20	21.1	23	-	-	15	18	19	-	47	14	15	-	31	38	12	34	45	18	17	35	29	-	24
M096	Indian	37	41	-	-	20	24.1	27	-	-	10	15	-	-	47	12	14	-	29	35	15	39	50	16	15	39	32	-	21
M097	Indian	36	39	-	-	22	25	30.1	-	-	12	13	16	-	48	14	17	-	32	38	13	37	47	19	18	37	31	-	21
M098	Indian	39	40	-	-	23	25.1	-	-	-	10	18	19	-	50	14	16	-	30	38	15	40	51	17	17	37	30	-	24
M099	Indian	37	40	-	-	22.2	23	23.1	-	-	13	-	-	-	48	14	-	-	31	41	14	37	54	16	18	33	33	-	19
M101	Indian	37	38	-	-	21.1	22	24	-	-	11	15	17	-	48	14	15	-	32	42	15	41	54	17	18	-	34	-	18
M102	Indian	37	-	-	-	17.2	21.1	24	-	-	12	14	15	-	45	14	-	-	30	35	14	37	54	17	21	35	29	-	18
M104	Indian	39	-	-	-	22	25.1	27	-	-	13	16	-	-	52	14	15	-	28	39	15	35	48	18	17	36	31	-	14
M106	Indian	38	-	-	-	19	-	-	-	-	12	15	-	-	52	15	16	-	29	41	13	36	48	15	18	37	31	-	20
M107	Indian	37	38	-	-	20	22	26.1	-	-	11	12	16	-	50	14	14.2	-	30	40	15	39	49	18	18	38	30	-	17
M109	Indian	38	-	-	-	20	20.1	23	-	-	10	14	17	-	49	15	16	-	32	40	15	39	49	19	20	36	34	-	17
M110	Indian	36	37	-	-	19	22	23.1	-	-	15	17	18	-	45	15	16	-	26	39	18	41	50	18	21	38	29	-	20
M111	Indian	37	-	-	-	22	23	24.1	-	-	10	14	15	-	50	14	15	-	33	41	15	39	48	19	19	34	33	-	17
M112	Indian	35	40	-	-	21	22	23.1	-	-	12	14	15	-	50	15	16	-	26	39	14	39	48	14	18	36	33	-	21
M115	Indian	37	39	-	-	19	22.1	-	-	-	10	13	16	-	47	14	16	-	26	37	13	36	45	16	15	35	27	-	21
M116	Indian	37	-	-	-	22.1	24.2	27	-	-	12	13	16	-	47	13	14	-	29	41	15	37	39	17	20	35	32	-	21
M117	Indian	35	38	-	-	22	23.1	24	-	-	14	16	-	-	49	14	15	-	33	41	15	38	43	18	18	38	29	-	21
M118	Indian	37	42	-	-	19	22	24	-	-	10	16	-	-	51	15	17	-	29	38	14	36	48	16	15	36	28	-	24
M119	Indian	38	39	-	-	20	23	23.1	-	-	10	16	-	-	51	13	16	-	33	36	12	34	51	16	18	36	30	-	21
M120	Indian	37	-	-	-	17	22.1	22.1	23	-	10	13	16	-	50	14	15	-	31	40	15	40	49	19	19	36	34	-	16
M121	Indian	37	-	-	-	22.1	24.2	-	-	-	13	15	-	-	47	15	-	-	31	37	15	36	45	18	16	36	33	-	18
M122	Indian	37	40	-	-	21.1	23.2	24	-	-	13	15	-	-	48	12	14	-	33	40	15	37	45	17	17	37	32	-	19
M123	Indian	36	40	-	-	23	24	24.1	-	-	14	16	-	-	49	14	15	-	27	39	14	36	46	19	16	36	25	-	18
M124	Indian	36	40	-	-	21	24	28.1	-	-	13	15	18	-	48	14	15	-	35	42	13	35	48	17	17	38	30	-	20
M125	Indian	38	40	-	-	19.1	23	25	-	-	10	14	17	-	51	14	-	-	31	41	15	41	49	19	17	36	35	-	18
M126	Indian	35	39	-	-	22	23	25.1	-	-	13	14	-	-	50	14	15	-	30	41	15	37	49	16	18	40	32	-	22
M127	Indian	38	40	-	-	22.1	23.1	25	-	-	12	13	16	-	48	15	16	-	32	43	15	38	45	16	18	36	31	-	18
M130	Indian	38	39	-	-	21.1	22	25	-	-	10	14	17	-	50	14	-	-	33	41	16	40	46	19	20	37	33	-	17
M131	Indian	38	39	-	-	17.2	24	25	-	-	12	13	18	-	48	15.2	16	-	28	38	13	34	47	17	16	38	29	-	18
M132	Indian	37	38	-	-	21	21.1	24	-	-	10	17	18	-	50	14	15	-	33	41	14	37	47	19	17	40	33	-	17
M133	Indian	36	41	-	-	21	23	25.1	-	-	13	17	-	-	49	14	15	-	33	40	13	34	49	17	17	38	30	-	22

ID	Population	DYF387S1				DYF399S1					DYF403S1				DYF403S1b	DYF404S1			DYS449	DYS518	DYS526a	DYS526b	DYS547	DYS570	DYS576	DYS612	DYS626		DYS627
M134	Indian	37	-	-	-	20	23.1	26.1	-	-	10	14	16	-	50	14	17	-	32	39	16	37	50	19	19	37	31	-	17
M135	Indian	37	38	-	-	21	22	23.1	-	-	10	17	-	-	50	14	-	-	29	39	12	36	51	17	18	37	28	-	22
M136	Indian	36	39	-	-	22.1	24	25.2	-	-	12	14	17	-	48	13	15	-	32	39	15	39	47	17	17	36	32	-	20
M137	Indian	37	38	-	-	21	23	25.1	-	-	10	15	-	-	51	13	15	-	29	40	15	38	42	18	18	33	27	-	22
M138	Indian	36	38	-	-	21	24	25.1	-	-	10	14	17	-	51	14	16	-	33	37	15	38	48	20	16	35	32	-	18
M139	Indian	35	38	-	-	18.1	24	-	-	-	10	12	17	-	52	14	15	-	29	40	16	39	48	19	18	36	31	-	21
M140	Indian	39	40	-	-	22	23	25	-	-	9	10	13	-	51	14	-	-	27	38	13	34	48	18	19	36	29	-	20
E075	Indian	36	39	-	-	20	21	21.1	-	-	10	16	17	-	49	14	15	-	32	42	15	40	47	17	18	39	30	-	17
E076	Indian	37	40	-	-	22	24	24.1	-	-	13	17	-	-	49	16	-	-	32	40	13	35	48	19	16	36	31	-	20
E077	Indian	37	40	-	-	19.1	21	22	-	-	10	17	-	-	51	14	15	-	32	41	15	39	47	19	17	38	33	-	17
E078	Indian	34	36	-	-	21	26.1	-	-	-	11	14	15	-	50	15	17	-	32	38	13	33	48	15	18	35	26	-	19
E044	Indian	37	39	-	-	21	21.1	23	-	-	7	15	16	-	50	14	16	-	31	44	13	35	47	19	19	-	32	-	18
E045	Indian	37	40	-	-	19.1	22	23	-	-	9	13	17	-	50	14	15	-	33	42	15	39	49	19	18	36	35	-	18
E046	Indian	40	41	-	-	19	23	25	-	-	12	15	16	-	51	13	14	-	30	42	14	33	48	17	18	36	31	-	20
E047	Indian	37	38	-	-	20	24	26.1	-	-	11	15	-	-	51	14	16	-	33	39	15	40	48	16	16	36	32	-	17
E048	Indian	36	39	-	-	22	26	28.1	-	-	12	17	19	-	49	15	16	-	33	38	13	36	47	18	17	35	30	-	18
E049	Indian	36	38	-	-	22	23	24.1	-	-	14	15	17	-	48	15	16	-	34	41	13	39	48	18	16	36	29	-	22
E050	Indian	37	39	-	-	20	21	25.1	-	-	11	12	16	-	51	14	15	-	27	37	14	36	48	17	16	38	30	-	18
E051	Indian	35	38	-	-	20	26	26.1	-	-	10	14	-	-	51	12	13	-	33	36	15	38	47	17	19	37	29	-	19
E052	Indian	35	37	-	-	21	23	26.1	-	-	15	16	19	-	49	14	-	-	33	39	12	32	50	19	16	37	25	-	22
E053	Indian	37	39	-	-	19.1	22	25	-	-	11	12	17	-	51	14	15	-	34	42	15	40	47	19	19	35	33	-	20
E054	Indian	37	40	-	-	21	21.1	24	-	-	10	14	16	-	49	14	15	-	31	42	14	36	49	20	19	-	34	-	18
E055	Indian	38	39	-	-	23.1	24	25.2	-	-	12	13	14	-	49	14	16	-	30	41	15	37	45	17	16	38	33	-	19
E056	Indian	35	38	-	-	21	23	23.1	-	-	12	14	-	-	51	14	16	-	28	40	14	40	47	15	16	40	31	-	22
E057	Indian	37	40	-	-	21.1	22	-	-	-	13	15	-	-	48	14	15	-	31	39	16	40	44	18	21	37	32	-	19
E058	Indian	37	39	-	-	20	21.1	23	-	-	10	14	15	-	49	14	15	-	32	41	15	39	48	18	18	-	32	-	17
E059	Indian	37	39	-	-	19.1	20	26	-	-	11	14	16	-	50	14	15	-	33	43	15	38	45	18	21	36	32	-	17
E060	Indian	38	39	-	-	23	25	26.1	-	-	12	16	17	-	48	14	15	-	31	39	14	38	46	19	16	38	30	-	21
M141	Pakistani	37	40	-	-	19.1	21	24	-	-	10	13	17	-	50	14	15	-	30	39	15	40	46	19	19	36	34	-	15
M142	Pakistani	37	-	-	-	20	22	23.1	-	-	14	17	-	-	52	16	-	-	32	40	12	35	49	17	15	37	30	-	20
M143	Pakistani	38	39	-	-	18.2	25	25	-	-	12	13	19	-	49	15.2	16	-	28	37	14	35	47	17	16	36	29	-	18
M144	Pakistani	36	39	-	-	21.1	23.2	24	-	-	12	13	14	-	47	14	15	-	31	37	15	36	48	16	19	35	32	-	19
M145	Pakistani	38	39	-	-	20	22	23	-	-	11	17	-	-	50	14	15	-	32	38	14	33	48	17	18	36	30	-	19
M146	Pakistani	36	38	-	-	20	21	23.1	-	-	13	15	18	-	52	14	15	-	25	38	16	40	49	18	17	35	30	-	19

ID	Population	DYF387S1				DYF399S1					DYF403S1				DYF403S1b	DYF404S1			DYS449	DYS518	DYS526a	DYS526b	DYS547	DYS570	DYS576	DYS612	DYS626		DYS627
M147	Pakistani	35	39	-	-	20	22.2	27.1	-	-	11	13	-	-	51	15	16	-	30	38	16	37	47	16	19	38	32	-	18
M149	Pakistani	36	40	-	-	22	23.1	-	-	-	10	14	15	-	52	13	16	-	27	36	14	38	50	17	15	40	31	-	20
M151	Pakistani	40	-	-	-	19.2	23	-	-	-	10	14	17	-	50	15.2	16	-	32	40	15	40	48	18	17	38	32	-	17
M153	Pakistani	36	38	-	-	19	22	23.1	-	-	12	16	-	-	50	14	15	-	25	40	14	36	49	17	17	36	30	-	21
M157	Pakistani	36	38	-	-	20	21	23.1	-	-	13	15	18	-	52	14	15	-	25	38	16	40	49	18	17	35	30	-	19
M158	Pakistani	36	39	-	-	20.1	21	25	-	-	12	13	19	-	50	13	15	-	32	39	15	38	50	19	18	36	31	-	17
M159	Pakistani	36	37	-	-	22	23	25.1	-	-	12	14	15	-	47	14	15	-	28	39	13	35	45	18	15	39	34	-	16
M160	Pakistani	34	36	-	-	21	24.1	-	-	-	10	14	15	-	48	13	15	-	30	39	13	35	47	16	19	34	28	-	21
M161	Pakistani	36	41	-	-	20	22	25.1	-	-	13	15	16	-	51	14	-	-	26	40	14	40	48	16	16	37	32	-	21
M162	Pakistani	36	37	-	-	23	23.2	24.1	-	-	14	18	-	-	49	14	-	-	32	38	16	38	49	17	17	37	33	-	19
M163	Pakistani	37	39	-	-	21	24	24.1	-	-	11	13	16	-	50	14	15	-	33	42	15	42	51	18	17	35	32	-	18
M164	Pakistani	36	37	-	-	19	22	22.1	-	-	15	17	-	-	54	14	16	-	26	41	17	40	50	17	18	37	30	-	21
M165	Pakistani	35	38	-	-	22	25	25.1	-	-	13	14	-	-	50	13	14	-	34	35	15	36	48	19	19	38	29	-	18
M166	Pakistani	37	39	-	-	21.1	23	24	-	-	10	18	-	-	51	13	15	-	32	40	15	39	49	19	17	35	32	-	17
M167	Pakistani	36	38	-	-	23	24.1	26	-	-	9	13	14	-	48	14	-	-	28	40	15	39	49	21	18	36	28	-	19
M168	Pakistani	37	39	-	-	21	22.1	24	-	-	11	14	17	-	51	15	-	-	31	44	16	41	47	19	19	35	33	-	17
M169	Pakistani	37	39	-	-	22	23.1	24	-	-	12	13	-	-	51	12	15	-	35	42	13	35	47	18	18	35	27	-	18
M170	Pakistani	35	39	-	-	23	23.1	24	-	-	14	16	-	-	47	15	16	-	32	39	15	37	45	21	18	37	29	-	22
M172	Pakistani	37	40	-	-	20	22.1	24	-	-	10	14	16	-	49	13	14	-	33	42	15	37	47	19	18	35	33	-	17
M173	Pakistani	37	39	-	-	21	23	24	-	-	12	14	15	-	47	13	17	-	38	39	13	35	48	17	13	36	34	-	21
M175	Pakistani	36	39	-	-	21	22.1	23	-	-	12	14	17	-	50	13	15	-	32	42	15	40	47	18	17	-	33	-	17
M176	Pakistani	36	39	-	-	20	23	-	-	-	10	14	-	-	46	13	14	-	32	43	15	36	48	19	17	36	32	-	22
M177	Pakistani	37	40	-	-	21	25	25.1	-	-	13	16	18	-	49	14	15	-	32	41	13	34	49	18	18	35	30	-	22
M178	Pakistani	37	40	-	-	20	23	23.1	-	-	10	14	15	-	51	14	16	-	32	41	14	36	45	20	18	37	35	-	21
E041	Pakistani	35	36	-	-	18.1	19	22	-	-	13	15	-	-	47	13	16	-	31	41	13	35	48	17	18	37	26	-	19
M180	Pakistani	36	43	-	-	24.1	-	-	-	-	10	13	14	-	47	15	-	-	24	40	13	34	44	18	19	41	31	-	19
M181	Pakistani	37	-	-	-	27.1	-	-	-	-	9	10	13	-	47	15	-	-	27	37	13	34	47	19	19	36	30	-	21
M182	Pakistani	37	38	-	-	20	23	25.1	-	-	12	15	16	-	47	16	17	-	33	40	15	38	48	20	17	39	31	-	17
M183	Pakistani	36	37	-	-	20	22	23.1	-	-	12	14	16	-	53	14	15	-	26	40	16	38	50	17	19	35	30	-	20
M185	Pakistani	35	40	-	-	20.1	21	25	-	-	11	15	16	-	50	15	-	-	32	42	15	40	48	18	17	33	34	-	17
M186	Pakistani	37	38	-	-	21	24	27	-	-	10	14	17	-	49	14	15	-	32	39	15	39	49	19	19	36	33	-	16
E042	Pakistani	35	39	-	-	22.1	22.2	23	-	-	11	14	16	-	48	14	16	-	32	40	15	39	51	17	18	35	32	-	18
M188	Pakistani	37	39	-	-	19	22	25.1	-	-	14	15	18	-	48	14	15	-	33	38	13	34	49	16	19	37	34	-	19
E043	Pakistani	37	39	-	-	24	25.1	-	-	-	10	15	16	-	48	14	15	-	31	38	14	37	45	20	16	39	32	-	21

ID	Population	DYF387S1				DYF399S1					DYF403S1				DYF403S1b	DYF404S1			DYS449	DYS518	DYS526a	DYS526b	DYS547	DYS570	DYS576	DYS612	DYS626		DYS627
M190	Pakistani	38	39	-	-	18	20	24	-	-	11	12	14	-	49	15	-	-	30	39	13	38	48	15	18	-	34	-	23
M191	Pakistani	35	40	-	-	21	24.1	-	-	-	11	12	14	-	50	13	15	-	27	39	13	38	47	15	16	36	32	-	21
M193	Pakistani	36	40	-	-	23.1	24	24.2	-	-	13	14	15	-	48	14	16	-	31	41	15	37	45	17	18	38	32	-	19
M194	Pakistani	38	-	-	-	20.1	22	23.1	-	-	10	12	15	-	49	14	-	-	27	39	13	34	44	18	18	35	31	-	21
M195	Pakistani	37	-	-	-	20	22.1	26	-	-	11	12	16	-	50	15	16	-	32	37	15	37	48	18	20	36	30	-	17
M197	Pakistani	36	40	-	-	22	23	28.1	-	-	13	15	16	-	47	15	-	-	26	37	15	37	46	20	17	-	30	-	20
M198	Pakistani	36	37	-	-	19	22	23.1	-	-	10	13	16	-	47	14	15	-	31	42	16	38	48	19	18	35	33	-	17
M200	Pakistani	40	-	-	-	14.2	14.2	26	-	-	10	14	17	-	49	15	-	-	32	41	15	39	50	18	19	36	32	-	17
M201	Pakistani	37	38	-	-	18	23.1	24.1	-	-	15	18	-	-	43	12	14.2	-	29	39	14	36	52	20	16	36	30	-	20
M202	Pakistani	38	39	-	-	17.2	24	25	-	-	12	15	18	-	47	15.2	17	-	28	38	13	34	47	17	16	-	29	-	18
M203	Pakistani	38	-	-	-	21	22	26.1	-	-	11	15	-	-	47	15	-	-	31	37	16	31	52	17	16	37	29	-	18
M204	Pakistani	35	-	-	-	21	25	26.1	-	-	12	13	16	-	47	13	15	-	29	40	14	36	48	18	19	38	31	-	20
M205	Pakistani	37	38	-	-	20.1	22	23	-	-	10	12	17	-	50	15	16	-	31	41	15	38	47	19	20	36	31	-	18
M206	Pakistani	36	39	-	-	20.2	21.1	25	-	-	13	16	-	-	49	14	15	-	32	42	15	38	45	17	18	-	32	-	18
M207	Pakistani	38	40	-	-	21.1	22	25	-	-	10	14	17	-	51	13	16	-	34	42	16	38	47	20	17	35	34	-	17
M208	Pakistani	39	-	-	-	21	22	24.1	-	-	10	13	14	-	48	14	-	-	30	38	13	36	49	16	17	35	28	-	20
M209	Pakistani	34	38	-	-	21	24.1	26	-	-	13	14	15	-	55	14	15	-	34	39	12	32	50	19	18	-	26	-	19
M210	Pakistani	36	38	-	-	19	23	24.1	-	-	15	16	18	-	43	14	15	-	32	37	13	35	48	17	16	32	31	-	21
M211	Pakistani	38	41	-	-	21.1	23	-	-	-	10	17	-	-	51	13	15	-	33	43	15	40	49	19	19	39	33	-	17
M212	Pakistani	36	39	-	-	23	27.1	-	-	-	13	16	-	-	49	13	14	-	35	39	13	35	49	17	17	35	31	-	20
M213	Pakistani	38	-	-	-	17	23.1	24.1	-	-	14	15	17	-	-	15	-	-	30	37	14	37	46	19	18	-	29	-	20
M214	Pakistani	36	39	-	-	21	22	22.1	-	-	10	13	16	-	51	14	15	-	32	43	15	38	48	19	20	36	32	-	17
M215	Pakistani	38	39	-	-	23	24	24.1	-	-	12	15	16	-	48	15	-	-	32	40	14	37	46	19	16	35	30	-	21
M216	Pakistani	35	41	-	-	22	25.1	-	-	-	12	14	15	-	47	16	18	-	26	38	14	39	49	15	16	36	29	-	22
M217	Pakistani	37	38	-	-	22	24.1	25	-	-	10	16	-	-	50	14	-	-	30	41	15	39	51	18	19	36	33	-	16
M218	Pakistani	35	38	-	-	22	22.1	25.2	-	-	13	17	-	-	51	13	-	-	31	40	15	37	48	16	18	35	32	-	19
M219	Pakistani	36	37	-	-	20	21	22.1	-	-	13	15	17	-	55	14	15	-	25	38	18	43	49	17	19	-	30	-	20
M220	Pakistani	36	38	-	-	20	22	25.1	-	-	10	12	18	-	49	15	16	-	29	39	13	33	47	16	20	35	26	-	21
M221	Pakistani	37	40	-	-	21	21.1	24	-	-	10	14	17	-	52	14	15	-	33	38	15	40	48	19	18	36	32	-	18
M222	Pakistani	37	40	-	-	24	24.1	24.2	-	-	12	15	-	-	47	13	16	-	31	39	15	37	45	17	19	-	33	-	20
M223	Pakistani	38	-	-	-	18	23	-	-	-	10	15	-	-	48	14	16	-	31	41	13	36	46	15	20	34	31	-	22
M224	Pakistani	38	39	-	-	19.1	21	24	-	-	10	13	17	-	49	14	15	-	31	43	15	37	50	19	19	36	33	-	17
M225	Pakistani	37	39	-	-	22.1	22.2	24	-	-	12	13	14	-	47	15	-	-	33	39	15	37	45	15	18	37	32	-	19
M226	Pakistani	37	40	-	-	19.1	22	24	-	-	10	14	17	-	51	15	-	-	32	42	16	39	50	19	19	35	33	-	18

ID	Population	DYF387S1				DYF399S1					DYF403S1				DYF403S1b	DYF404S1			DYS449	DYS518	DYS526a	DYS526b	DYS547	DYS570	DYS576	DYS612	DYS626		DYS627
M227	Pakistani	36	37	-	-	19	21	22.1	-	-	13	16	-	-	54	14	15	-	25	40	17	41	49	19	18	-	30	-	24
M228	Pakistani	37	38	-	-	20	23.1	24	-	-	10	15	16	-	50	15	-	-	32	40	16	38	50	19	20	38	30	-	17
M229	Pakistani	38	-	-	-	18	22	24	-	-	12	13	15	-	48	12	14.2	-	32	40	13	35	49	15	19	35	34	-	23
M231	Pakistani	36	39	-	-	23	25	25.1	-	-	12	14	-	-	50	14	15	-	26	39	14	40	48	15	17	38	32	-	24
M232	Pakistani	37	38	-	-	20	21	23.1	-	-	12	14	15	-	48	14	-	-	27	39	13	38	46	17	17	37	31	-	18
M233	Pakistani	35	40	-	-	20	24	25.1	-	-	12	14	-	-	52	14	15	-	26	38	14	39	49	15	19	37	30	-	21
M235	Pakistani	38	39	-	-	22	22.1	23	-	-	9	15	16	-	50	15	-	-	34	43	14	37	48	19	19	36	33	-	19
M237	Pakistani	37	39	-	-	20.1	21	24	-	-	13	15	16	-	50	14	15	-	31	40	14	37	47	19	18	38	32	-	18
M238	Pakistani	36	37	38	-	20	22.1	23	-	-	13	14	15	-	53	14	15	-	31	40	14	36	48	18	18	37	31	-	22
M240	Pakistani	36	40	-	-	20	21	27.1	-	-	13	14	-	-	48	14	15	-	27	38	14	40	47	15	17	38	32	-	23
M242	Pakistani	38	-	-	-	24	25	25.1	-	-	12	14	15	-	48	15	-	-	30	37	14	38	46	18	17	36	32	-	23
M243	Pakistani	34	-	-	-	22	26.1	-	-	-	12	14	-	-	52	16	18	-	26	39	14	39	48	15	17	36	32	-	21
M245	Pakistani	37	-	-	-	20	23	-	-	-	10	14	-	-	54	13	-	-	32	37	15	36	50	19	20	39	29	-	18
M247	Pakistani	37	40	-	-	20	24.2	25.1	-	-	12	13	15	-	47	14	17	-	33	39	14	36	44	17	19	36	32	-	18
M248	Pakistani	36	38	-	-	17	21.1	25	-	-	10	12	15	-	48	13	14	-	30	40	12	35	48	14	16	37	29	-	20
M249	Pakistani	38	39	-	-	20	25	25.1	-	-	11	13	16	-	52	14	17	-	33	38	13	36	48	18	16	35	25	-	18
M250	Pakistani	37	39	-	-	21	21.1	22	-	-	10	15	17	-	54	13	17	-	33	41	15	39	-	19	19	35	32	-	17
M251	Pakistani	36	-	-	-	24	25	-	-	-	4	7	15	-	47	13	-	-	29	41	13	37	46	17	19	36	29	-	21
M252	Pakistani	39	40	-	-	21	25	26.1	-	-	13	16	18	-	48	13	16	-	24	37	14	38	50	14	21	37	31	-	22
M254	Pakistani	37	40	-	-	17	21	23.1	-	-	11	12	15	-	50	14	15	-	31	38	13	33	45	16	17	35	31	-	21
M255	Pakistani	38	39	-	-	20	22	23	-	-	10	17	-	-	50	14	15	-	32	38	14	33	48	17	18	36	30	-	19
M256	Pakistani	36	39	-	-	21	22	25	-	-	13	15	17	-	49	16	17	-	34	38	13	35	45	19	17	36	29	-	19
M257	Pakistani	37	39	-	-	20	23	23.1	-	-	12	14	-	-	47	13	14	-	29	37	12	34	52	15	16	36	29	-	21
M258	Pakistani	37	39	-	-	22	23	27.1	-	-	13	16	-	-	48	14	15	-	34	42	13	35	49	15	18	36	30	-	19
M259	Pakistani	36	39	-	-	23	24.1	25	-	-	14	18	20	-	50	15	-	-	33	42	13	34	48	19	19	34	32	-	22
M260	Pakistani	36	38	-	-	20.1	21	22	-	-	8	9	16	-	50	14	-	-	33	43	15	39	48	19	18	37	33	-	19
M262	Pakistani	38	-	-	-	21	23	24	-	-	10	15	17	-	51	14	-	-	32	40	16	40	48	19	15	34	34	-	18
M263	Pakistani	35	38	-	-	21	22.1	26	-	-	10	13	15	-	49	13	16	-	31	40	13	34	49	20	17	35	25	-	21
M264	Pakistani	37	-	-	-	19	21	24	-	-	13	15	17	-	54	14	15	-	26	41	17	40	49	18	18	36	30	-	22
M265	Pakistani	36	40	-	-	20	21	27	-	-	11	17	-	-	50	15	17	-	29	37	13	40	48	19	17	35	29	-	20
M266	Pakistani	34	39	-	-	22	27.1	-	-	-	14	17	19	-	50	16	17	-	32	39	13	35	46	16	18	35	29	-	21
M267	Pakistani	36	37	-	-	23	23.1	24	-	-	14	15	17	-	49	14	-	-	34	38	13	33	50	19	18	37	31	-	17
M268	Pakistani	36	38	-	-	20	21	24	-	-	13	15	17	-	50	14	-	-	29	40	14	35	50	19	15	36	31	-	19
M269	Pakistani	38	-	-	-	18	24	27.1	-	-	12	13	16.2	-	50	16	-	-	33	37	13	32	47	18	17	36	26	-	21

ID	Population	DYF387S1				DYF399S1					DYF403S1				DYF403S1b	DYF404S1			DYS449	DYS518	DYS526a	DYS526b	DYS547	DYS570	DYS576	DYS612	DYS626		DYS627
M270	Pakistani	34	41	-	-	22	26.1	-	-	-	12	14	-	-	50	14	-	-	26	37	14	39	47	15	18	38	32	-	20
M271	Pakistani	37	-	-	-	22	24.1	-	-	-	12	13	16	-	51	11	15	-	29	39	14	35	46	20	16	34	31	-	19
M272	Pakistani	37	39	-	-	21	22.1	-	-	-	12	14	-	-	50	13	14	-	27	36	13	36	43	18	18	37	31	-	19
M273	Pakistani	39	-	-	-	20	22	24	-	-	12	15	16	-	50	14	-	-	31	39	14	33	50	17	15	36	32	-	20
M274	Pakistani	37	39	-	-	19.1	23	24	-	-	10	14	16	-	50	14	-	-	31	43	15	39	49	19	18	38	34	-	17
M275	Pakistani	34	36	-	-	21	22	23.1	-	-	10	14	15	-	49	13	16	-	32	41	13	35	49	16	18	31	28	-	22
M276	Pakistani	37	38	-	-	19	21	22	-	-	12	14	-	-	50	16	-	-	30	37	13	36	47	14	18	36	33	-	21
M277	Pakistani	39	-	-	-	22	24	-	-	-	13	15	16	-	49	14	15	-	26	39	14	36	50	16	17	36	29	-	18
M278	Pakistani	38	39	-	-	17.2	24	26	-	-	12	13	17	-	48	15.2	17	-	28	38	13	34	47	17	16	37	29	-	18
M279	Pakistani	37	39	-	-	21.1	23.2	24	-	-	11	13	-	-	48	13	15	-	28	41	15	37	47	17	16	37	32	-	19
M280	Pakistani	36	38	-	-	22	23.1	24	-	-	9	13	17	-	50	15	-	-	34	42	15	39	46	18	17	35	31	-	16.2
M281	Pakistani	37	-	-	-	20	22.1	23	-	-	10	13	16	-	50	14	15	-	32	41	16	38	49	19	19	37	31	-	17
M282	Pakistani	37	41	-	-	21	22.2	23.2	24.1	-	12	13	-	-	47	13	15	-	32	39	14	37	47	16	18	36	32	33	20
M283	Pakistani	36	38	-	-	21	26.1	-	-	-	12	16	17	-	49	15	16	-	32	39	15	39	47	19	18	39	30	-	19
M285	Pakistani	36	39	-	-	19	20	23.1	-	-	14	15	18	-	49	15	-	-	31	40	13	36	50	21	17	36	31	-	19
M286	Pakistani	35	40	-	-	21	23	25.1	-	-	11	14	-	-	48	13	15	-	26	39	14	39	51	15	17	38	31	-	22
M287	Pakistani	37	40	-	-	21	21.1	26	-	-	10	13	18	-	49	12.5	13.5	-	32	42	15	38	49	18	20	36	32	-	17
M288	Pakistani	38	39	-	-	19	23	25	-	-	12	13	17	-	48	15.2	17	-	28	38	13	34	47	17	16	37	28	-	18
M289	Pakistani	36	38	-	-	17	21.1	26	-	-	10	15	-	-	48	14	-	-	28	38	12	36	49	15	16	37	29	-	18
M290	Pakistani	36	38	-	-	21.1	23	27	-	-	9	17	-	-	52	14	-	-	32	40	14	38	47	19	20	-	32	-	17
M291	Pakistani	38	-	-	-	20	23.1	24	-	-	10	15	-	-	52	13	17	-	28	39	15	38	43	17	19	33	27	-	20
M293	Pakistani	36	-	-	-	19	22	25	-	-	12	14	16	-	49	14	-	-	33	41	12	33	49	22	18	35	24	-	16
M294	Pakistani	37	-	-	-	27.1	-	-	-	-	12	13	14	-	51	14	-	-	29	35	13	36	44	18	18	34	33	-	21
M295	Pakistani	36	39	-	-	19.1	21	24	-	-	11	14	17	-	49	15	16	-	32	41	15	37	48	19	18	37	35	-	17
M296	Pakistani	36	38	-	-	19	21	23	-	-	9	10	14	-	52	14	15	-	32	39	13	35	48	19	18	36	32	-	21
M297	Pakistani	36	41	-	-	21	23	24.1	-	-	13	16	-	-	50	11	16	-	31	44	14	35	49	18	18	35	32	-	19
M298	Pakistani	38	39	-	-	23	25.1	25.2	-	-	13	14	-	-	47	13	14	-	32	38	16	37	47	17	18	37	33	-	19
M299	Pakistani	37	40	-	-	21	22	24.1	-	-	13	16	-	-	49	16	-	-	31	40	12	34	46	18	15	35	30	-	20
M300	Pakistani	37	-	-	-	21	22	27.1	-	-	11	13	16	-	49	15	17	-	30	37	13	35	49	15	17	35	32	-	22
M301	Pakistani	35	39	-	-	23	26.1	-	-	-	15	19	20	-	48	16	17	-	33	38	13	35	47	18	16	36	32	-	19
M302	Pakistani	35	40	-	-	23	25.1	26	-	-	13	14	-	-	50	13	16	-	27	41	14	39	49	16	17	39	32	-	22
M303	Pakistani	37	39	-	-	21	21.1	24	-	-	10	14	16	-	50	14	15	-	32	41	15	39	49	19	19	35	34	-	17
M305	Pakistani	37	-	-	-	20.1	22	22	-	-	12	13	14	-	48	13	15	-	28	36	12	36	48	18	20	36	31	-	18.2
M306	Pakistani	35	41	-	-	21	23	27.1	-	-	12	14	15	-	49	15	-	-	26	40	14	39	47	15	16	37	31	-	22

ID	Population	DYF387S1				DYF399S1					DYF403S1				DYF403S1b	DYF404S1			DYS449	DYS518	DYS526a	DYS526b	DYS547	DYS570	DYS576	DYS612	DYS626		DYS627
M307	Pakistani	36	38	-	-	23.1	25	-	-	-	15	16	-	-	45	13	15	-	32	37	12	37	44	14	16	33	30	-	22
M308	Pakistani	38	39	-	-	23	23.1	-	-	-	13	17	-	-	51	15	-	-	32	41	15	38	49	19	18	36	33	-	18
M309	Pakistani	38	39	-	-	20.2	21.1	24	-	-	12	13	16	-	48	14	15	-	33	39	15	36	45	17	16	38	32	-	19
M310	Pakistani	38	-	-	-	23	25.1	-	-	-	12	13	-	-	49	10	16.2	-	27	43	13	31	46	18	17	27	29	-	21
M311	Pakistani	37	39	-	-	22	23.1	24	-	-	11	14	16	-	50	14	17	-	32	42	15	38	50	20	19	-	32	-	17
M312	Pakistani	36	39	-	-	23	23.1	24.2	-	-	13	-	-	-	49	13	15	-	32	39	15	37	47	17	17	36	33	-	22
M313	Pakistani	36	38	-	-	22	24	25.1	-	-	13	14	15	-	47	16	-	-	31	39	13	33	52	14	17	34	27	-	20
M314	Pakistani	35	36	-	-	21	23	25	-	-	10	13	17	-	52	15	17	-	29	39	14	34	47	18	20	35	26	-	20
M315	Pakistani	37	38	-	-	22	24	25.1	-	-	10	12	16	-	50	14	-	-	30	40	15	39	48	18	18	37	33	-	17
M316	Pakistani	36	39	-	-	21.1	22	23	-	-	11	14	-	-	51	12	13	-	29	37	14	38	54	16	14	36	29	-	18
M317	Pakistani	37	39	-	-	20.1	21	25	-	-	10	14	18	-	50	15	-	-	31	43	15	39	46	22	19	37	33	-	17
M318	Pakistani	36	39	-	-	20.1	22	23	-	-	10	13	17	-	50	15	16	-	32	40	15	40	47	18	17	37	32	-	18
M320	Pakistani	38	39	-	-	22	24	24.1	-	-	12	15	16	-	48	15	16	-	31	38	14	37	46	19	16	38	30	-	23
M322	Pakistani	36	38	-	-	22	25.1	26.1	-	-	13	14	18	-	46	17	-	-	31	37	13	34	50	15	20	33	28	-	22
M323	Pakistani	41	-	-	-	23	25.1	-	-	-	12	16	18	-	48	16	-	-	33	40	13	35	45	20	17	36	31	-	20
M324	Pakistani	36	37	38	-	20	21	23	24.1	-	13	18	-	-	44	12	-	-	31	35	14	33	46	21	17	37	32	-	21
M325	Pakistani	40	-	-	-	21	23	27	-	-	12	13	14	-	47	13	17	-	30	38	13	37	50	16	17	37	31	-	18
M326	Pakistani	36	40	-	-	21	23	24.1	-	-	12	15	16	-	47	14	-	-	31	40	13	37	48	17	14	36	30	-	25
M327	Pakistani	36	39	-	-	21	21.1	24	-	-	9	15	17	-	50	15	-	-	32	42	15	39	48	20	18	36	33	-	18
M328	Pakistani	37	39	-	-	21	21.1	24	-	-	11	17	-	-	50	14	15	-	32	43	14	38	48	19	20	37	30	-	18
M329	Pakistani	37	-	-	-	19	21	21.1	-	-	13	15	16	-	54	14	16	-	25	39	18	41	50	19	17	36	30	-	21
M330	Pakistani	36	40	-	-	21	22.1	25	-	-	10	13	16	-	49	14	15	-	33	41	14	39	49	18	21	35	32	-	17
M331	Pakistani	35	-	-	-	21	24	29	-	-	13	16	-	-	50	15	16	-	31	36	13	35	48	18	15	35	26	-	21
M332	Pakistani	38	39	-	-	18	19	23	-	-	10	12	14	-	49	14	15	-	30	38	13	35	49	16	17	33	32	-	21
M333	Pakistani	39	-	-	-	21	24.1	27.1	-	-	16	-	-	-	50	14	16	-	28	38	14	39	50	13	18	37	33	-	18
M334	Pakistani	36	40	-	-	21	22	-	-	-	14	17	-	-	49	14	16	-	33	40	12	35	47	18	18	35	31	-	21
M335	Pakistani	35	37	-	-	19	23	25.1	-	-	16	17	-	-	50	14	16	-	31	42	12	32	50	19	16	38	25	-	22
M337	Pakistani	33	39	-	-	21	21.1	24	-	-	9	10	17	-	51	14	16	-	34	41	16	41	48	18	18	38	32	-	18
M338	Pakistani	36	-	-	-	22	23	25.1	-	-	13	19	-	-	50	14	18	-	30	41	14	37	47	17	16	37	29	-	20
M339	Pakistani	38	40	-	-	19	21	27.1	-	-	14	15	17	-	47.2	13	15	-	24	38	14	34	41	18	19	38	30	-	22
M340	Pakistani	36	38	-	-	17	21.1	24	-	-	12	16	-	-	50	14	-	-	28	39	12	35	47	16	17	37	29	-	18
M341	Pakistani	37	-	-	-	19	21	24.1	-	-	13	15	17	-	54	14	15	-	26	41	17	40	49	18	18	36	30	-	22
M343	Pakistani	36	38	-	-	22	25.1	-	-	-	12	17	18	-	49	15	16	-	33	40	12	34	47	21	17	38	29	-	19
M344	Pakistani	35	36	-	-	21	23	25.1	-	-	10	18	-	-	50	14	-	-	30	39	14	37	50	17	20	37	28	-	22

ID	Population	DYF387S1				DYF399S1					DYF403S1				DYF403S1b	DYF404S1			DYS449	DYS518	DYS526a	DYS526b	DYS547	DYS570	DYS576	DYS612	DYS626		DYS627
M345	Pakistani	37	-	-	-	18	25.1	-	-	-	10	13	15	-	50	15	-	-	32	41	12	33	48	22	18	37	24	-	18
M346	Pakistani	38	-	-	-	21	23	23.1	-	-	9	12	17	-	46	15	-	-	29	36	12	34	49	16	16	39	29	-	20
M347	Pakistani	37	40	-	-	22.1	24	24.2	-	-	12	19	20	-	49	13	15	-	31	42	15	37	48	17	19	39	32	-	20
M348	Pakistani	36	37	-	-	20	21.1	26	-	-	13	16	17	-	52	14	15	-	33	42	12	33	52	18	15	36	26	-	19
M349	Pakistani	37	38	-	-	21.1	22.2	23	-	-	12	13	16	-	48	13	16	-	31	38	15	38	45	18	17	35	33	-	18
M350	Pakistani	36	39	-	-	21	22	24.1	-	-	13	16	-	-	46	15	-	-	291	37	12	34	50	16	16	35	29	-	20
M351	Pakistani	36	37	-	-	18	23	26.1	-	-	11	15	-	-	50	14	-	-	31	42	12	32	50	21	18	36	24	-	18
M352	Pakistani	37	39	-	-	21	23	26.1	-	-	10	13	16	-	50	13	15	-	32	42	15	39	49	20	18	35	32	-	16
M353	Pakistani	36	39	-	-	21	26.1	-	-	-	14	18	19	-	49	15	16	-	34	40	13	33	50	17	14	36	29	-	21
M354	Pakistani	39	40	-	-	21.1	22	23	-	-	12	17	18	-	48	15	-	-	31	40	12	36	48	16	19	36	31	-	21
M355	Pakistani	39	-	-	-	20	21	-	-	-	11	15	16	-	48	15	-	-	31	45	13	36	48	19	13	37	32	-	21
M356	Pakistani	36	39	-	-	22.1	23.2	24	-	-	11	13	-	-	45	13	14	-	30	41	16	38	45	17	18	39	32	-	19
M357	Pakistani	37	39	-	-	21	23.1	24	-	-	9	15	17	-	51	14	16	-	34	41	15	38	47	20	18	37	34	-	17
M358	Pakistani	37	39	-	-	19	23	24	-	-	9	14	15	-	53	13	14	-	32	43	15	35	48	18	17	35	32	-	19
M359	Pakistani	35	38	-	-	22	23	28.1	-	-	13	14	15	-	53	13	-	-	32	36	15	38	47	18	18	34	28	-	19
M360	Pakistani	37	39	-	-	21	21.1	22	-	-	11	15	17	-	50	13	17	-	33	41	13	37	48	19	19	35	32	-	17
M361	Pakistani	39	-	-	-	22	27	-	-	-	10	11	13	-	53	13	15	-	31	39	13	34	48	17	16	38	30	-	18
M362	Pakistani	38	39	-	-	23	26	26.1	-	-	15	17	-	-	48	15	-	-	31	40	14	36	51	17	13	38	31	-	17
M363	Pakistani	37	39	-	-	21	23	25.1	-	-	10	15	16	-	50	14	15	-	31	41	15	39	47	17	19	38	32	-	21
M365	Pakistani	36	38	-	-	20	21	22	-	-	9	10	14	-	49	14	15	-	32	39	13	35	48	20	18	36	31	-	21
M366	Pakistani	38	-	-	-	21	24	27.1	-	-	12	13	14	-	51	12	15	-	28	39	13	37	48	16	18	37	29	-	21
M367	Pakistani	37	38	-	-	23.2	24.1	-	-	-	11	14	16	-	48	17	18	-	33	38	15	37	46	17	17	37	33	-	21
M368	Pakistani	36	41	-	-	20	22	23.1	-	-	12	15	-	-	50	12	14	-	27	37	14	39	47	15	18	37	30	-	23
M369	Pakistani	35	-	-	-	23	26.1	-	-	-	14	15	-	-	50	14	-	-	27	38	14	39	49	16	15	36	31	-	22
M370	Pakistani	35	38	-	-	20	23	24.1	-	-	13	14	15	-	53	12	13	-	32	36	16	38	47	18	19	38	29	-	18
M371	Pakistani	38	39	-	-	19	23	-	-	-	10	12	14	-	49	14	15	-	30	38	13	35	49	16	17	33	32	-	21
M372	Pakistani	33	41	-	-	21	23	28.1	-	-	10	13	14	-	50	14	-	-	26	38	14	39	48	15	17	38	32	-	25
M373	Pakistani	37	39	-	-	21	22.1	23	-	-	10	13	16	-	50	14	-	-	32	39	15	37	47	19	18	33	33	-	18
M374	Pakistani	36	41	-	-	22	24	24.1	-	-	10	13	15	-	53	14	-	-	25	35	14	38	49	17	17	36	30	-	21
M375	Pakistani	38	-	-	-	21	25	29	-	-	17	-	-	-	48	16	17	-	32	39	12	34	47	18	16	37	29	-	21
M376	Pakistani	36	39	-	-	20	20.1	23	-	-	12	13	14	-	47	15	-	-	31	42	15	38	50	19	18	38	32	-	20
M377	Pakistani	37	-	-	-	20	21	22.1	-	-	13	14	17	-	54	15	16	-	25	38	17	41	48	19	18	40	30	-	21
M378	Pakistani	38	42	-	-	24	25.1	-	-	-	12	15	16	-	47	14	15	-	31	38	14	38	46	19	16	38	31	-	21
M379	Pakistani	37	38	-	-	21	22	23.1	-	-	12	13	14	-	48	14	-	-	28	36	13	34	44	18	19	37	33	-	20

ID	Population	DYF387S1				DYF399S1					DYF403S1				DYF403S1b	DYF404S1			DYS449	DYS518	DYS526a	DYS526b	DYS547	DYS570	DYS576	DYS612	DYS626		DYS627
M380	Pakistani	37	40	-	-	17	18	24	-	-	10	12	14	-	50	15	-	-	31	37	14	37	50	16	18	34	32	-	20
M381	Pakistani	37	-	-	-	21	23	25.1	-	-	12	15	17	-	52	14	16	-	33	44	11	31	48	18	18	37	26	-	22
M382	Pakistani	38	-	-	-	21	23	24.1	-	-	11	15	17	-	51	14	-	-	32	40	16	40	48	19	15	34	34	-	18
M383	Pakistani	39	40	-	-	21	23	26.1	-	-	10	11	17	-	48	14	15	-	32	37	13	36	45	14	19	38	31	-	21
M384	Pakistani	37	39	-	-	20.1	22	24	-	-	10	14	17	-	52	14	16	-	32	38	15	40	48	19	19	36	33	-	18
M385	Pakistani	39	-	-	-	20	24	25.1	-	-	10	12	15	-	53	12	16	-	29	39	15	37	43	17	18	32	27	-	22
M386	Pakistani	37	40	-	-	19.1	20	25	-	-	15	17	-	-	50	15	16	-	33	41	15	39	50	19	17	36	32	-	18
M387	Pakistani	37	-	-	-	20	23	-	-	-	13	-	-	-	49	13	-	-	29	41	13	35	48	16	20	38	31	-	21
M388	Pakistani	35	39	-	-	22	23	27.1	-	-	10	14	-	-	51	13	14	-	27	38	14	40	47	16	17	38	33	-	19
M389	Pakistani	35	40	-	-	22	22.1	23.2	-	-	13	15	-	-	50	13	14	-	36	39	15	37	45	18	18	39	32	-	18
M390	Pakistani	35	41	-	-	22	24	27.1	-	-	14	15	-	-	51	14	15	-	26	37	14	39	47	15	18	38	32	-	22
M391	Pakistani	37	38	-	-	21	23	23.1	-	-	9	12	17	-	46	15	-	-	29	36	12	34	49	17	17	40	29	-	20
M393	Pakistani	36	38	-	-	21	23	23.1	-	-	10	12	17	-	46	13	16	-	28	39	12	34	51	16	16	37	29	-	18
M394	Pakistani	38	42	-	-	24	25	-	-	-	12	15	16	-	47	14	15	-	24	38	14	38	46	19	16	38	31	-	21
M395	Pakistani	36	38	-	-	21	23	23.1	-	-	10	13	-	-	51	13	14	-	33	37	13	34	49	20	16	36	29	-	17
M396	Pakistani	36	38	-	-	24	27.1	-	-	-	12	14	15	-	47	15	-	-	30	36	13	35	52	14	18	36	28	-	19
M397	Pakistani	37	38	-	-	21	23	26.1	-	-	11	12	16	-	50	14	15	-	31	40	15	38	49	18	18	36	30	-	17
M398	Pakistani	38	-	-	-	23.1	25	-	-	-	11	13	15	-	48	15	-	-	31	41	14	38	45	20	17	37	31	-	20
M399	Pakistani	33	38	-	-	20	24.1	-	-	-	11	11.2	17	-	55	12	14	-	31	38	14	38	46	19	17	38	33	-	21
E061	Pakistani	36	40	-	-	20	20.1	23	-	-	9	16	18	-	50	13	14	-	33	42	15	39	48	19	18	35	28	-	17
E062	Pakistani	37	40	-	-	19.1	21	23	-	-	10	13	17	-	50	15	-	-	31	39	15	38	46	19	15	36	33	-	18
E063	Pakistani	36	38	-	-	21	22	23.1	-	-	10	12	17	-	50	15	16	-	33	42	15	39	49	21	18	34	32	-	17
E064	Pakistani	38	39	-	-	21	24	24.1	-	-	12	14	16	-	51	14	15	-	27	41	14	36	52	19	16	37	30	-	18
E065	Pakistani	36	38	-	-	19	23.1	25	-	-	12	14	-	-	51	11	14	-	33	36	15	37	47	20	17	37	28	-	18
E066	Pakistani	37	39	-	-	22.1	23	25	-	-	14	18	-	-	51	14	15	-	28	41	15	38	47	20	20	37	32	-	17
E067	Pakistani	37	39	-	-	22	22.1	23.1	24.2	-	12	14	15	-	49	13	15	-	32	42	15	37	45	15	17	36	32	-	20
E068	Pakistani	38	41	-	-	20	21	22	23	-	10	11	15	-	49	14	17	-	34	37	13	36	49	22	16	37	33	-	22
E069	Pakistani	36	37	40	-	21	22	23	24.1	-	12	13	15	-	50	14	-	-	26	38	14	38	48	17	18	39	31	-	21
E070	Pakistani	36	39	-	-	20	23	24.1	-	-	13	15	17	-	50	15	17	-	36	39	12	32	50	16	18	35	31	-	19
E071	Pakistani	37	39	-	-	20	22.1	24	-	-	10	14	18	-	51	15	-	-	31	41	13	37	47	20	20	33	33	-	17
E072	Pakistani	37	39	40	-	21	22.1	22.2	-	-	12	14	-	-	48	13	15	-	32	39	15	37	46	17	17	33	32	33	20
E073	Pakistani	38	39	-	-	19.1	21	25	-	-	12	13	16	-	50	12	15	-	32	39	15	39	48	18	18	36	32	-	17
E074	Pakistani	37	38	-	-	21	22	24.1	-	-	12	13	16	-	52	12	15	-	31	34	15	38	49	18	18	37	28	-	19
Q001	Qatari	37	38	-	-	20	23.1	24	-	-	11	14	16	-	50	14	16	-	32	41	17	41	46	19	20	36	31	-	18

ID	Population	DYF387S1				DYF399S1					DYF403S1				DYF403S1b	DYF404S1			DYS449	DYS518	DYS526a	DYS526b	DYS547	DYS570	DYS576	DYS612	DYS626		DYS627
Q002	Qatari	36	41	-	-	19.1	21	25	-	-	11	18	-	-	50	14	15	-	31	41	15	38	46	20	19	39	34	-	17
Q003	Qatari	36	-	-	-	20	21	23.1	-	-	15	-	-	-	55	14	15	-	26	39	17	42	49	18	17	37	30	-	21
Q004	Qatari	37	38	-	-	20	23	26	-	-	14	15	-	-	50	14	16.2	-	29	38	14	33	50	19	14	35	31	-	18
Q005	Qatari	36	37	-	-	20	21	22.1	-	-	14	15	-	-	55	13	15	-	25	37	17	40	49	19	17	36	30	-	21
Q008	Qatari	37	38	-	-	20.2	23.1	23	-	-	12	14	16	-	49	14.2	15	-	28	40	13	33	51	18	18	36	30	-	21
Q009	Qatari	36	37	-	-	19	21	22.1	-	-	13	15	16	-	54	14	15	-	25	39	17	40	48	17	16	38	30	-	21
Q010	Qatari	36	37	-	-	19	22	24.1	-	-	13	16	17	-	53	14	15	-	27	40	16	38	49	18	18	33	30	-	18
Q011	Qatari	37	38	-	-	20	21	23.1	-	-	13	15	16	-	53	12	14	-	26	38	16	37	50	18	18	37	32	-	22
Q012	Qatari	36	-	-	-	19	20	22.1	-	-	15	16	-	-	55	14	-	-	26	39	16	39	48	18	17	36	30	-	21
Q014	Qatari	37	38	-	-	20	22	22.1	-	-	13	15	17	-	54	14	15	-	27	39	17	39	47	17	17	36	30	-	22.3
Q015	Qatari	36	38	-	-	20	22	23.1	-	-	11	15	16	-	54	14	-	-	24	38	16	38	49	17	18	36	31	-	20
Q017	Qatari	36	37	-	-	20	21	22.1	-	-	13	15	16	-	54	14	15	-	24	40	17	38	49	18	17	36	30	-	21
Q021	Qatari	38	41.2	-	-	25	26	26.1	-	-	14	14.2	17	-	49	11	14	-	34	42	12	33	47	17	17	39	34	-	20
Q024	Qatari	36	37	-	-	19	20	22	22.1	-	13	15	16	-	55	14	15	-	25	39	17	40	48	18	18	36	30	-	21
G003	Qatari	39	-	-	-	21	23	-	-	-	11	13	16	-	47	13	15	-	29	38	13	37	47	18	15	35	33	-	20
G008	Qatari	37	38	-	-	19	22	23.1	-	-	13	15	17	-	54	14	15	-	25	40	17	40	49	18	17	37	30	-	22
G016	Qatari	37	38	-	-	20	22.1	23	-	-	11	13	16	-	51	14	15	-	31	43	15	38	47	19	17	37	30	-	17
G019	Qatari	36	-	-	-	20	-	-	-	-	13	14	16	-	54	14	-	-	26	40	17	37	49	18	18	39	30	-	21
G020	Qatari	37	-	-	-	17	19	22	22.1	-	13	14	16	-	55	13	15	-	25	40	17	40	48	18	17	37	31	-	21
G034	Qatari	37	40	-	-	23	23.1	24.2	-	-	13	14	-	-	49	12	14	-	32	39	16	39	46	20	18	37	32	-	20
G036	Qatari	35	37	-	-	20	24	24.1	-	-	11	14	18	-	52	15	-	-	31	37	12	34	49	21	17	36	26	-	23
G046	Qatari	37	-	-	-	20	21	22.1	-	-	13	14	17	-	54	15	16	-	26	39	16	39	49	18	19	38	30	-	19
G052	Qatari	36	38	-	-	21	24	24.1	-	-	9	14	17	-	49	14	15	-	33	40	15	38	48	18	20	34	30	-	16
G055	Qatari	36	38	-	-	20	22	23.1	-	-	12	14	16	-	54	14	-	-	24	40	16	38	49	17	18	36	32	-	20
G056	Qatari	37	-	-	-	19	22	23.1	-	-	14	16	18	-	54	14	15	-	25	41	17	39	49	18	18	37	30	-	20
G059	Qatari	36	38	-	-	21	24	24.1	-	-	9	14	17	-	49	14	15	-	33	40	15	38	48	18	20	34	30	-	16
G060	Qatari	36	37	-	-	19	21	22.1	-	-	14	15	16	-	55	14	15	-	25	38	17	41	48	19	17	37	30	-	21
G063	Qatari	37	39	-	-	19	23	24	-	-	12	15	-	-	51	14	15	-	28	37	14	34	50	20	15	33	30	-	20
G099	Qatari	36	37	-	-	19	21	22.1	-	-	13	16	-	-	56	14	15	-	25	40	17	38	50	18	18	38	31	-	20
G114	Qatari	36	37	-	-	19	21	22.1	-	-	14	16	-	-	55	14	-	-	25	39	17	39	48	18	17	36	30	-	21
G122	Qatari	37	38	-	-	19	22	22.1	-	-	13	14	17	-	53	14	15	-	25	40	17	40	49	18	17	38	30	-	22
G141	Qatari	36	37	-	-	19	21	22.1	-	-	14	15	-	-	55	13	15	-	25	37	17	40	49	18	17	36	30	-	21
G147	Qatari	37	38	-	-	22.1	23.1	24	-	-	14	15	16	-	45	16	-	-	33	38	12	35	44	13	15	30	29	-	21
B001	Bahraini	37	-	-	-	19	20	23.1	-	-	13	16	17	-	53	14	-	-	25	40	17	40	49	18	18	38	29	-	21

ID	Population	DYF387S1				DYF399S1					DYF403S1				DYF403S1b	DYF404S1			DYS449	DYS518	DYS526a	DYS526b	DYS547	DYS570	DYS576	DYS612	DYS626		DYS627
B002	Bahraini	40	-	-	-	23.1	-	-	-	-	14	16	18	-	51	14.2	-	-	27	38	12	34	46	18	17	35	31	-	22
G006	Bahraini	36	37	-	-	19	21	22.1	-	-	14	15	-	-	57	13	15	-	25	37	17	40	49	18	17	36	30	-	21
G018	Bahraini	37	-	-	-	19	-	-	-	-	13	16	17	-	54	15	-	-	25	40	17	40	49	18	19	36	30	-	20
G053	Bahraini	37	38	-	-	18	19.1	23	-	-	15	18	-	-	46	13.2	14	-	30	37	13	36	47	17	18	34	33	-	19
G067	Bahraini	37	38	-	-	19	22	23.1	-	-	14	16	-	-	53	13	14	-	26	41	15	37	49	18	16	36	31	-	21
G086	Bahraini	35	39	-	-	22	24	24.1	-	-	12	14	16	-	50	13	14	-	29	38	14	36	44	17	17	38	31	-	21
G094	Bahraini	36	37	-	-	19	21.1	22	-	-	13	14	16	-	54	14	16	-	25	40	17	39	49	18	17	38	30	-	21
G133	Bahraini	37	42.2	-	-	23	24	25.1	-	-	13	16	-	-	47	14	-	-	33	47	12	34	47	16	18	37	34	-	21
G139	Bahraini	39	-	-	-	20	22.1	23	-	-	11	14	18	-	52	14	15	-	33	38	13	32	48	18	19	34	26	-	19
G140	Bahraini	37	38	-	-	20	23	23.1	-	-	12	15	16	-	50	14	15	-	32	40	17	41	46	19	19	36	32	-	18
G145	Bahraini	37	38	-	-	20	23	23.1	-	-	12	15	16	-	51	14	16	-	32	40	17	41	46	19	19	36	32	-	18
D053	KSA	37	-	-	-	19	21.1	22	-	-	13	15	17	-	53	14	15	-	25	40	17	39	49	18	18	37	30	-	22
G025	KSA	36	37	-	-	19	20	25.1	-	-	13	14	17	-	52	13	-	-	25	39	17	38	49	18	18	36	30	-	21
G026	KSA	37	-	-	-	20	21	23.1	-	-	13	16	17	-	54	14	15	-	25	40	17	38	50	18	18	37	30	-	20
G035	KSA	39	40	-	-	18	22	24	26.1	-	12	13	16	-	51	14	16	-	29	40	13	35	49	21	18	36	25	-	19
G050	KSA	36	38	-	-	19	23	23.1	-	-	13	14	19	-	54	13	14	-	25	38	17	37	48	18	16	38	31	-	22
G051	KSA	37	-	-	-	20	-	-	-	-	12	16	-	-	54	13	-	-	27	39	16	38	49	17	17	36	31	-	23
G068	KSA	37	39	-	-	20	24	-	-	-	11	13	15	-	49	14	15	-	30	37	13	35	48	16	19	34	32	-	21
G091	KSA	35	36	-	-	19	23	24.1	-	-	15	16	-	-	49	15	-	-	33	40	12	33	49	22	18	36	24	-	21
G120	KSA	38	39	-	-	20	24	-	-	-	11	13	15	-	49	14	15	-	31	37	12	34	48	16	19	34	30	-	21
G128	KSA	36	-	-	-	20	24.1	-	-	-	15	16	-	-	50	12	13	-	25	41	15	38	49	17	18	36	29	-	23
G134	KSA	37	-	-	-	21	24.1	-	-	-	14	17	-	-	52	16	-	-	25	38	17	38	49	18	16	38	30	-	21.2
G142	KSA	36	37	-	-	19	21	25.1	-	-	13	16	17	-	53	13	-	-	25	39	18	40	50	18	18	36	30	-	21
G156	KSA	37	-	-	-	21	22	25.1	27.1	-	12	17	-	-	52	14	-	-	26	38	17	39	48	16	18	34	30	-	21
Q013	Kuwaiti	37	-	-	-	19	21	23.1	-	-	14	15	16	-	55	14	15	-	25	40	17	39	49	18	19	35	30	-	20
Q016	Kuwaiti	36	37	-	-	21.2	22	22.1	-	-	11	15	16	-	44	12.2	13	-	27	37	14	35	46	17	16	38	31	-	21
Q018	Kuwaiti	36	37	-	-	20	23	24.1	-	-	13	14	17	-	50	13	15	-	32	43	15	39	47	19	16	39	31	-	17
Q025	Kuwaiti	37	-	-	-	19	21	23.1	-	-	13	15	17	-	54	14	15	-	25	40	16	38	49	18	18	37	30	-	22
G043	Kuwaiti	37	-	-	-	19	22	22.1	-	-	13	15	17	-	54	14	15	-	26	40	17	40	48	18	18	37	30	-	21
G066	Kuwaiti	36	-	-	-	19	21	23.1	-	-	13	16	17	-	55	14	15	-	26	39	17	38	48	18	18	37	32	-	21
Q019	Yemeni	37	-	-	-	20	23.1	-	-	-	13	14	19	-	52	15	-	-	25	38	17	39	49	18	16	37	31	-	21.2
Q022	Yemeni	36	38	-	-	19	21	24.1	-	-	13	14	17	-	52	14	-	-	26	38	16	39	49	17	17	36	30	-	21
Q023	Yemeni	36	37	-	-	20	21	24.1	-	-	16	17	-	-	54	14	-	-	26	38	17	39	49	18	18	36	30	-	21
G024	Yemeni	36	37	-	-	22	24.1	-	-	-	13	14	17	-	51	14	-	-	27	41	17	39	51	17	19	35	30	-	20

ID	Population	DYF387S1				DYF399S1					DYF403S1				DYF403S1b	DYF404S1			DYS449	DYS518	DYS526a	DYS526b	DYS547	DYS570	DYS576	DYS612	DYS626		DYS627
G065	Yemeni	37	40	-	-	21	22	-	-	-	14	15	-	-	54	15	16	-	28	39	13	35	53	16	18	37	30	-	21
G092	Yemeni	36	37	-	-	22	24.1	-	-	-	13	14	17	-	51	14	-	-	27	41	17	39	51	17	19	35	30	-	20
G123	Yemeni	36	37	-	-	19	23.1	-	-	-	13	16	17	-	53	13	15	-	25	36	17	38	50	17	18	36	30	-	20
Q006	Yemeni	37	-	-	-	19	22	24.1	-	-	13	18	-	-	52	14	-	-	25	40	17	38	49	18	18	36	31	-	20
Q007	Yemeni	37	-	-	-	17	21	26.1	-	-	11	14	18	-	53	14	-	-	24	39	17	39	48	17	17	39	30	-	19
G155	Yemeni	35	38	-	-	18	21	24.1	-	-	14	16	-	-	53	14	17	-	27	37	16	39	49	18	18	34	30	-	21