6-2021

FORENSIC EVALUATION OF QIAGEN INVESTIGATOR ARGUS Y-28 WITH CASE WORK SAMPLES

Afnan Mohammed Alshamsi

United Arab Emirates University

College of Science

Department of Biology

FORENSIC EVALUATION OF QIAGEN INVESTIGATOR ARGUS Y-28 WITH CASE WORK SAMPLES

Afnan Mohammed Alshamsi

This thesis is submitted in partial fulfilment of the requirements for the degree of Master of Science in Molecular Biology and Biotechnology

Under the Supervision of Professor Rabah Iratni

June 2021

Declaration of Original Work

I, Afnan Mohammed Alshamsi, the undersigned, a graduate student at the United Arab Emirates University (UAEU), and the author of this thesis entitled "*Forensic Evaluation of Qiagen Investigator Argus Y-28 with Case Work Samples*", hereby, solemnly declare that this thesis is my own original research work that has been done and prepared by me under the supervision of Professor Rabah Iratni, in the College of Science at UAEU. This work has not previously formed the basis for the award of any academic degree, diploma or a similar title at this or any other university. Any materials borrowed from other sources (whether published or unpublished) and relied upon or included in my thesis have been properly cited and acknowledged in accordance with appropriate academic conventions. I further declare that there is no potential conflict of interest with respect to the research, data collection, authorship, presentation and/or publication of this thesis.

Student's Signature: Afnan Date: 16/6/2021

Advisory Committee

1) Advisor: Rabah Iratni

Title: Professor

Department of Biology

College of Science

2) Co-advisor: Rashed Alghafri

Title: Associate Professor

Department of Biology

College of Science

Approval of the Master Thesis

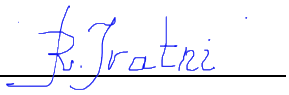
This Master Thesis is approved by the following Examining Committee Members:

1) Advisor (Committee Chair): Rabah Iratni

   Title: Professor

   Department of Biology

   College of Science

   Signature  Date 10/6/2021

2) Member: Rashed Alghafri

   Title: Associate Professor

   Department of Biology

   College of Science

   Signature  Date 13/6/2021

3) Member: Ranjit Vijayan

   Title: Associate Professor

   Department of Biology

   College of Science

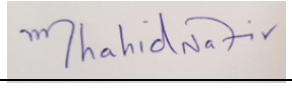
   Signature  Date 13/6/2021

4) Member (External Examiner): Muhammad Shahid Nazir

   Title: Associate Professor

   College of Biotechnology

   Institution: University of Modern Sciences, Dubai, UAE

   Signature  Date 13/6/2021

This Master Thesis is accepted by:

Dean of the College of Science: Professor Maamar Benkraouda

Signature maamar Benkraouda Date July 26, 2021

Dean of the College of Graduate Studies: Professor Ali Al-Marzouqi

Signature Ali Hassan Date July 26, 2021

Copy ____ of ____

Abstract

Short Tandem Repeats polymorphisms on the Y-chromosome (Y-STRs) play a significant role in history of human hereditary. The recently introduced 6-dyes Qiagen investigator Argus Y-28 Kit which include 28Y-STR loci (DYS389I, DYS391, DYS389II, DYS533, DYS390, DYS627, DYS458, DYS393, DYS19, DYS437, DYS449, DYS460, DYS576, YGATAH4, DYS481, DYS448, DYS518, DYS439, DYS549, DYS438, DYS456, DYS643, QS1, DYS570, DYS635, DYS385, DYS392, QS2) has been used for internal validation study used in the forensic laboratories which followed sensitivity study, mixture studies, PCR parameter limits, stability study, and performance on mock casework samples. The results were reported by statistical analysis using ANOVA program. The results validate the kit's robustness, reliability, and suitability as an assay for human identification with casework DNA samples. This kit is expected to improve male relative differentiation in future and human hereditary applications.

Keywords: Y- chromosome, STR, internal validation, casework samples, forensic, PCR, qiagen kit.

Title and Abstract (in Arabic)

تقييم موقع الكروموسوم الذكري في الأدلة الجنائية باستخدام Qiagen Argus Y-28 مع عينات القضايا

الملخص

التكرارات المترادفه القصيرة في موقع الكروموسوم الذكري (STR-Y) يلعب دورًا مهمًا في تاريخ الوراثة البشرية. تم طرح 6 أصباغ من Qiagen investigator Argus Y-28 Kit والتي تشمل 28Y-STR loci

(DYS389I, DYS391, DYS389II, DYS533, DYS390, DYS627, DYS458, DYS393, DYS19, DYS437, DYS449, DYS460, DYS576, YGATAH4, DYS481, DYS448, DYS518, DYS439, DYS549, DYS438, DYS456, DYS643, QS1, DYS570, DYS635, DYS385, DYS392, QS2)

لدراسة التحقق الداخلي المستخدمة في مختبرات الطب الشرعي والتي اشتملت على دراسة الحساسية ودراسات الخليط، وحدود معامل PCR، دراسة الاستقرار، والأداء على عينات وهمية. تم الحصول على النتائج عن طريق التحليل الإحصائي باستخدام برنامج ANOVA. هذه النتائج حققت من خلال قوة المجموعة وموثوقيتها ومدى ملاءمتها كمقايسة لتحديد هوية الإنسان باستخدام عينات الحمض النووي الخاصة بعمل الحالات. من المتوقع أن تعمل هذه المجموعة على تحسين التمايز النسبي للذكور في التطبيقات الوراثية البشرية في المستقبل.

مفاهيم البحث الرئيسية: Y- كروموسوم، STR، التحقق الداخلي، عينات الحالات، الطب الشرعي، PCR، مجموعة qiagen.

Acknowledgment

Above all else I'm incredibly thankful to my supervisors, Prof. Rabah Iratni and Dr. Rashed Alghafri for their important and high support, consistent help, and persistence during my thesis. Their massive information and ample experience have empowered me in all the hour of my scholarly examination and everyday life. I might likewise want to express gratitude toward Tareq Zeyad Ahmed, Reem Lotah Mayda Tracy Bakri, Afra Alowais Shamma Almutawa, and Shamsa Alfalasi for their specialized help on my investigation. I might want to thank every one of the individuals in the Dubai Police Department for providing me with facilities during my work in the thesis research. It is their thoughtful assistance and backing that have made my examination and life in the UAE a great time. Special thanks to major general Mohammed Khalfan Al Romaithi, the commander-in-chief of Abu Dhabi Police for supporting and encouraging me through master's study.

My gratefulness to express my deepest appreciation to my sister Ahlam Alshamsi and Ibtesam Saeed for their comments, suggestions and constant support.

At last, I might want to offer my thanks to my friend Amina Sameh, and my family. Without their huge arrangement and support in the previous few years, it would be inconceivable for me to finish my examination. I express love, appreciation and indebtedness to them.

Dedication

To my beloved parents and family

Table of Contents

List of Tables

List of Figures

List of Abbreviations

AMOVA	Analysis of Molecular Variation
AZF	Azoospermia Factor
CE	Capillary Electrophoresis
CE	Capillary Electrophoresis
MRCA	Most Recent Common Ancestor
PCR	Polymerase Chain Reaction
RFLP	Restriction Fragment Length Polymorphism
SD	Standard Deviation
SRR	Simple Sequence Repeats
STR	Short Tandem Repeat

Chapter 1: Introduction

1.1 Overview

Short Tandem Repeat (STRs) or Simple Sequence Repeats (SRRs) are sequences repeat mostly found surrounding the chromosomal centromere in all chromosomes. DNA regions with repeat units that are 2 to 7 bp in length are mainly used in forensic applications. These repeats are called microsatellites (Tautz, 1993), shaping arrangement with lengths of up to 100 nucleotides (nt). STRs are widely found in prokaryotes and eukaryotes. They are distributed evenly throughout the human genome, representing about 3% of the whole genome. Nonetheless, their distribution within chromosomes is not exactly uniform—they appear less often in sub -telomeric loci (Koreth et al., 1996). Most of the STRs are found in the noncoding regions, while only about 8% located in the coding regions (Ellegren, 2000). Moreover, their densities differ marginally among chromosomes. In humans, chromosome 19 has the most noteworthy density of STRs (Subramanian et al., 2003). By and large, one STR happens per 2,000 bp in the human genome. The most well-known STRs in people are A-rich units: A, AC, AAAN, AAN and AG (Nadir et al., 1996). The STR locus is named as, for example, D3S1266, where D addresses DNA, 3 denotes chromosome number 3 on which STR locus is observed, S stands for STR, and 1266 is the special identifier.

Y human chromosome represents about 2% of the total human genome and contain 78 genes which code for proteins. Male relative separation utilizing Y-chromosome markers is attainable by utilizing Y-STRs. For just about twenty years of Y-STR examination and applications, STR loci better usage in human identification in

forensic laboratories because DNA in such samples may mix or degraded (Butler, 2009).

1.2 Relevant Literature

1.2.1 History of DNA

In 1869, Friedrich Miescher isolated a substance called nuclein in the nuclei of white blood cells. This substance composed of proteins is what known today as DNA (deoxyribonucleic acid) (Inman & Rudin, 2001). Phoebus Levene, an American biochemist, identified the components of DNA-the bases-Adenine (A), Thymine (T), Cytosine (C) and Guanine (G) in addition to sugar molecule (deoxyribose) and phosphate group (Levene, 1940). In 1944, DNA was defined by Oswald Avery. He eliminated all component of the bacterium that causes pneumonia. With the exception of DNA, he found that DNA was capable of transformation a harmless from of the bacterium into a lethal one demonstrating that DNA carries hereditary information. Rosalind franklin developed X-ray of the crystal structure of DNA molecule. The images were captured using this diffraction technique that showed that DNA molecules have helical structure (Maddox, 2014). After referring to the images captured by Rosalind in 1951, Watson and Crick developed a chemical model of the DNA molecule, suggesting that DNA has a double helix structure (Maddox, 2014).

DNA fingerprinting or DNA profiling was first described in 1985 by an English geneticist named Alec Jeffreys. Dr Jeffreys found that certain region of DNA contained DNA sequence that were repeated over and again next to each other. He also discovered that the number repeated sections present in sample could differ from individual to individual. By developing a technique to examine the length variation of these DNA repeat sequences, Dr Jeffreys created the ability to perform human identity

tests. These DNA repeat regions became known as VNTRs, which stands for variable number of tandem repeats. The technique used to examine the VNTR was Restriction Fragment Length Polymorphism (RFLP) because it involved the use of a restriction enzyme to cut the regions of DNA surrounding the VNTRs. This RFLP methods was first used to help in an English immigration case and shortly thereafter to solve a double homicide case (Aartsma-Rus, 2012).

Polymerase Chain Reaction (PCR) is a new tool in molecular biology, which was invented by Kary Mullis in the year 1985. PCR is so sensitive that single DNA amplified they used to make more copy of DNA. The method requires DNA polymerase enzyme that makes new strands of DNA, using existing strands as templates. The DNA polymerase typically used in PCR is called Taq polymerase, after the heat-tolerant bacterium from which it was isolated (Thermus aquaticus) (Butler, 2009).

1.2.2 Human Y Chromosome

In the late 20th century, Theophilus painter studied the role the Y-chromosome plays in the determination and development of the male embryo. His research assists other researcher to determine sex chromosome (Butler, 2009). Y chromosome studies lends itself to several mutating alleles on Y chromosome that can be used to group human Y chromosomes into vertical paternal lineages called Y chromosome haplotype. The early Y chromosome informational sets of data contained not many haplotypes (Dorit et al., 1995; Mallet, 1995; Whitfield et al., 1995), limiting most examinations to a straightforward segment model: consistent populace size. Later informational data sets (Underhill et al., 1997) and microsatellite informational collections, for example, those considered here contain bigger quantities of familial

transformations and accordingly present the chance to research more overwhelming models.

In this investigation, a model was considered of human demography, which takes into consideration populace development beginning sooner or later to before and proceeding to the present. In spite of the fact that is followed the majority of the past Y chromosome concentrates in assessing the opportunity to the Most Recent Common Ancestor (MRCA), additionally assessed a progression of segment boundaries under the expected to be model. Be that as it may, the MRCA time is simply by implication associated with populace history and is famously hard to appraise with high accuracy. Therefore, more can be found out about the historical backdrop of populaces from assessments of segment boundaries. The significant finding of this investigation is that a model of steady populace size delivers a helpless fit to the information for the vast majority of the populaces considered. The model of dramatic development perfor ms obviously better, and for the overall example, a point scale for the populace development pace of 0.008 per age, beginning 18,000 years prior were observed. The predictable example of populace development shows that hereditary distance measures and factual tests that accept consistent populace size are unseemly for the human Y chromosome. It will be essential to assess which distance measures and tree-building strategies are powerful within the sight of populace development. All means considered, the degrees of variety ought to be more comparative (Slatkin & Hudson 1991).

Human cells contain 23 pairs of chromosomes, out of which, 22 pairs of autosomal chromosomes and a pair of sex chromosome (Figure 1) (Butler, 2009).

Females contain two X (XX) chromosomes while male contain a single copy of the X chromosome and single y chromosome (XY) (Aartsma-Rus, 2012).

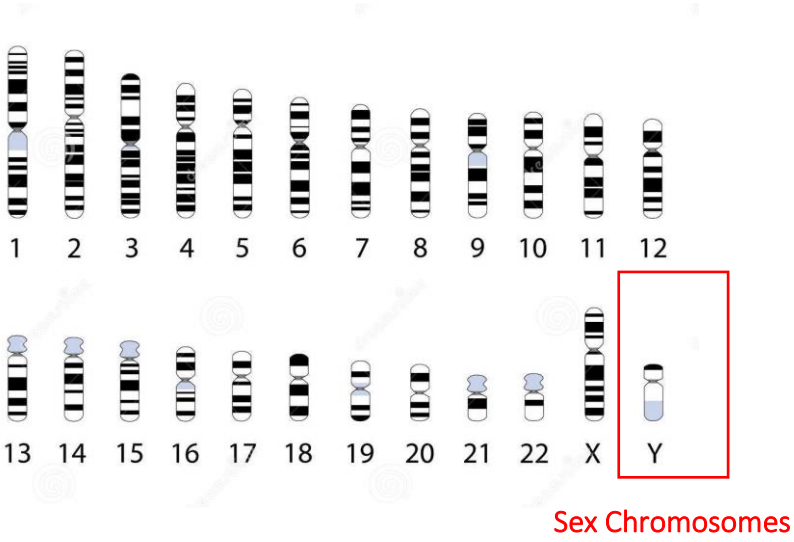


Figure 1: Male chromosome
Source: (Butler, 2009)

There are many applications of Y chromosome analysis including paternity testing, forensic casework on sexual assault, human migration, historical and genealogical. Another significant application of Y chromosome testing is finding missing persons. The value of the Y chromosome in forensic DNA testing is its occurrence only in males. The SRY (Sex-Determining Region of the Y) gene determines maleness. A majority of the Y chromosome is inherited directly from father to son (Figure 2) (Butler, 2009) without recombination to shuffle its genes and provide greater variety to future generations (Zeyad, 2017).

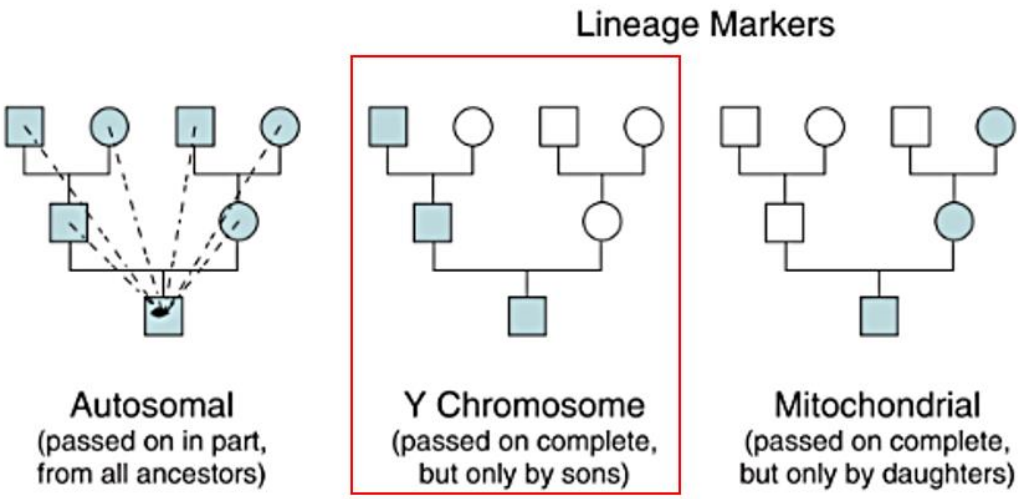


Figure 2: Illustrate of inheritance pattern and the value of Y chromomere to human identity
Source: (Butler, 2009)

- Structure of the Human Y Chromosome

There are two small portions Pseudo Autosomal Regions [PAR1 and PAR2] located at the terminal ends of the Y chromosome. PAR1 located on the green boxes show the genes encoded in these regions (at the tip of Yp), Yp is the short arm of the Y chromosome and the genes (Zeyad, 2017). The long arm, Yq, is composed of both euchromatin and the genetically inactive heterochromatin regions (Colaco, 2018). The pink box shows the genes in the Azoospermia Factor (AZFa) region refers to one of several proteins or their genes, which are coded from the AZF region on the human male Y chromosome. Deletions in this region are associated with inability to produce sperm (Colaco, 2018). The heterochromatin is not known to harbor any known genes. The region beyond the PAR is termed as Male Specific Region on Y (MSY) (Figure 3) (Colaco, 2018).

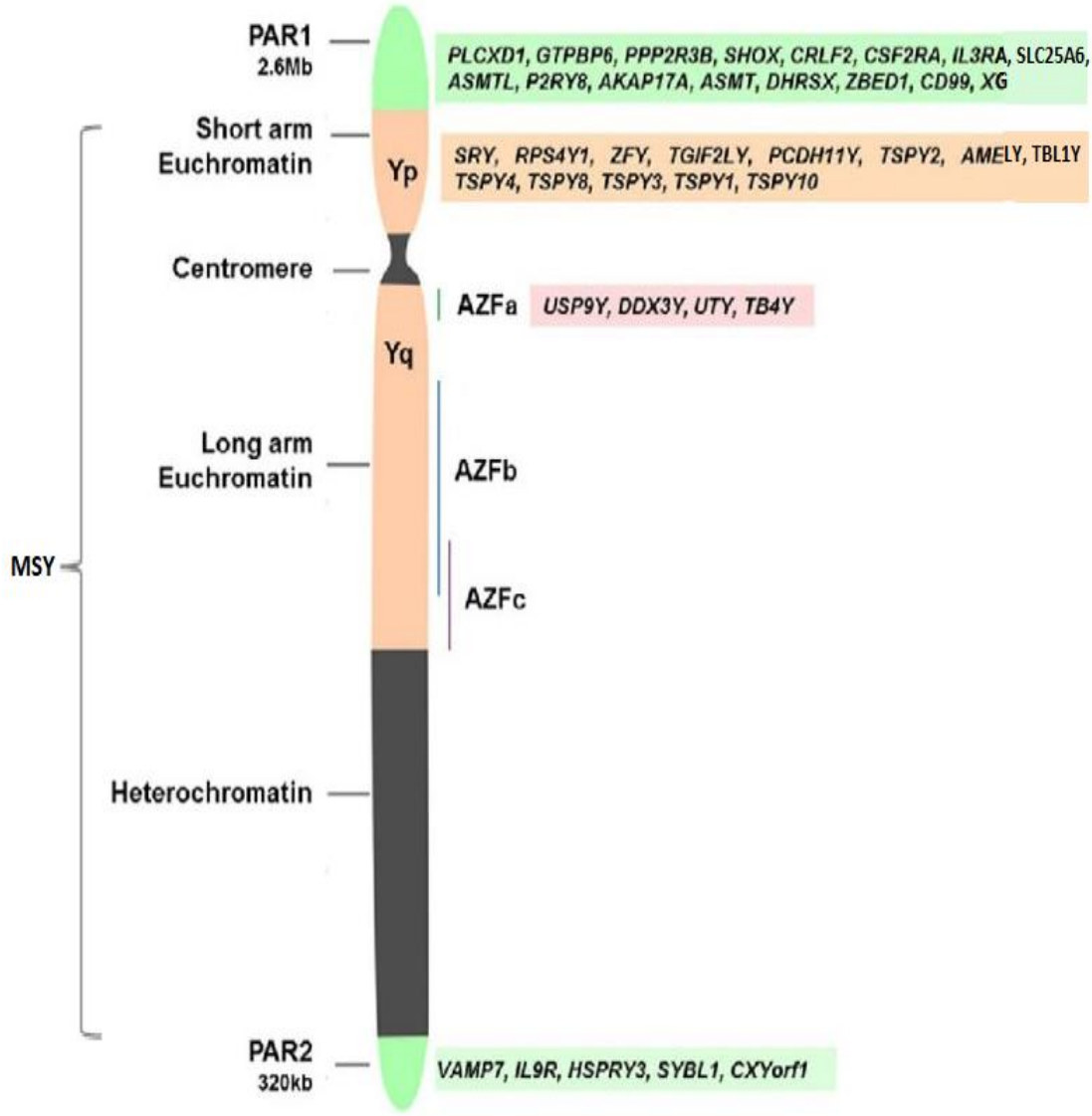


Figure 3: Structure of the human Y chromosome
Source: (Colaco, 2018)

1.2.3 Short Tandem Repeat

Short Tandem Repeat (STR) investigation has developed throughout the most recent 25 years to become and stay the best quality level for human individual identification purposes in measurable forensic genetics (Fregeau & Fourney, 1993), while they are also utilized in other human Y zones. Other than autosomal STRs, the human genome of male people additionally contains many STRs situated on the male-specific part of the human Y-chromosome (Y-STRs).

In studies involving forensics, Y-STRs are particularly valuable for addressing rape cases with DNA combinations normally containing an excess of DNA from the female casualty's epithelial cells contrasted and DNA of the male culprit's sperm cells (Roewer, 2019). In view of such imbalanced male–female DNA blends, it regularly is basically difficult to distinguish the male donor dependent on autosomal STR profiling, even after differential lysis prompting advancement of sperm DNA was applied (Vuichard et al., 2011).

Interestingly, a Y-STR profile (haplotype) of the male supporter can regularly be gotten from such blended material, which permits deciding the fatherly ancestry to which the male crime location follow contributor has a place (Kayser, 2017). In view of the absence of recombination and the relatively low mutations rate (~10^{-3} transformations per marker per meiosis) of the Y-STRS ordinarily utilized in measurable Y-chromosome examination, a Y-STR haplotype features the male culprit along with a significant number of his in a fatherly way related male family members (Claerhout et al., 2020).

1.2.4 Types of STR Loci

STR that are usually used have either 4 bp or 5 bp core repeat and often divided into several categories based on the repeat pattern, those are simple repeats, compound repeats and complex repeats.

Simple repeats: STR are those, which contain one unit of identical length and sequence (eg. AGATA AGATA AGATA). Example of this type are CSFIPO, TH01, TPOX, D5S818, D7S820, D13S317 and D22S1045.

Compound repeats: comprise two or more adjacent simple repeats (e.g. TCTATCTG TCTATCTG). Example of this type of repeats are FGA, VWA and D6S1043.

Complex repeats may contain several repeat blocks of variable unit length as well as variable intervening sequences. Example of this type of repeats is D21S1

1.2.5 Application of STR

Non-Recombination Y Chromosome (NRY) condition and genetic diversity of the most part of Y chromosome NRY are very useful in population history study paternal determination, familial searching and identifying when sample are mixed with male and female DNA. Another important application of Y-STR in forensic is the detecting number of unrelated male contributors in mixed DNA samples of two or more unrelated male individuals. Other uses are paternity or familial testing where DNA samples of alleged father are not available, for missing persons and identification of male victims (Li et al., 2020). Y-STR haplotype data is also used to get the information about a particular Y haplogroup, and therefore Y haplogroups are more informative for biogeographic ancestry inferences. In forensic, this is very helpful in tracing unknown suspects, as it can be possible to infer individual's paternal biogeographic ancestry from Y-chromosome analysis. There is a limitation of Y-STR haplotype, which shows a single locus rather than set of independent markers and also Y-STR haplotype is much more polymorphic than single STR loci. Also, it requires larger databases for analysis. This limitation arises as a result of low mutation rate and had problems in differentiating fathers and brothers and distant paternal relationships (Arwin et al., 2019).

Parental lineage identification is one of the widely explored applications of Y-STR and around 13 RM Y-STRs are identified to be very useful for parental lineage identification and differentiation. In a global study of parental lineage identification, 12,272 unrelated males from 111 global populations, a high accurate, 99% i.e, 12, 156 were differentiated through unique RM-STRs system. Male individual identification is another potential application of Y-STR in forensics. The availability of sufficient numbers of RM-YSTR markers makes the separation of unrelated and distantly related men in the study population becomes easier due to the observed mutations. Also, YSTR can be explored for the assessment of parental bio-geographic ancestry and useful in missing person cases such as disaster victim identification. The candidature of Y-STR is more suitable for inferring bio-geographical identification as the repeated sequence escape from recombination (Kayser, 2017). A number of research has already been carried out in the identification and analysis of YSTR in varying applications. However, the research findings that utilized and validated the efficacy of the YSTR28 QS kit. Thus, the current investigation is intended to validate the newly developed Investigator Argus Y28 QS kit using samples from case studies.

1.3 Hypothesis and Objectives

The current research was aimed to validate the efficacy of the recently developed Qiagen Investigator Argus Y-28 using samples from case studies. Further, the possibilities of the success rates were analyzed through varied subjects and samples, which would validate the kit as an efficient tool in forensic investigations.

The objectives of the study were:

1. To measure the sensitivity and stability of the Qiagen Investigator Argus Y-STR28 assay using lowest sample concentrations and Taqman probes respectively.

2. To validate PCR limits of the kit by altering the parameters such as $MgCl_2$ and EDTA.

3. To analyze mixture sensitivity and inhibition resistance of the subjective kit by mixing and varying the ratio of male and female DNA samples.

Chapter 2: Materials and Methods

2.1 Sample Collection

The samples were collected randomly from unrelated individuals. Ethical approval has been obtained from Dubai Police Scientific research ethic committee in Dubai Health Authority (DHA) for the use of biological samples in this research.

2.2 Quality Control

The main experiments were carried out at the training center of Dubai police and Forensic Science and which is an accredited laboratory (ISO 17025), in accordance with quality control measures. All the methods were carried out in accordance with the approved guidelines of Forensic Science.

2.3 General Considerations

Prior to starting internal validation study, the Investigator Argus Y-28 kit was received on dry ice and stored immediately at −30°C to −15°C in a constant-temperature freezer upon receipt. The primer mix and allelic ladder were stored protected from light. DNA samples and post-PCR reagents (allelic ladder and DNA size standard) were stored separately from PCR reagents as indiacted in the kit. The results given in the guidelines were considered as a standard for the comparison with the current investigation. Long-term storage, storage conditions conditions of buffer compounds and storage of genomic DNA were avoided, which might influence the capacity to create FORENSICS-STR profiles. Only the assigned work regions and pipettes for pre-and present amplification steps were utilized to avoid the potential cross-contamination of among the samples. Gloves were worn and changed regularly considering the high priority DNA samples.

2.4 Sensitivity Studies

A set of six different male DNA samples as mentioned (Taqman, 9948, 2800, 007, R1 and R2) were used to test the sensitivity. Serial dilutions of these samples were prepared. Each sample was analyzed using the optimal PCR cycling conditions (as per the manufacturer protocol). Capillary Electrophoresis (CE) instrument utilizing the infusion boundaries were considered for this investigation. This information can be utilized to decide the powerful reach and affectability for the Y-STR framework in blend with the infusion conditions for the CE instrument. Each example was dissected utilizing the ideal PCR cycling conditions (According to the production convention). Various quantities of PCR cycles (mostly 3) were utilized to survey the sensitivity of the assay.

2.5 Mixture Studies

A mixed male/female DNA sample with known ratios Male / female mixture: Male samples were constant while female samples in variable mixture ratio 1:250, 1:500, 1:750, 1:1000 for a total of 0.5 ng of DNA were prepared for the mixture study.

Male/Male mixture: The reason for the male/male combinations is to decide at which proportion a major significant profile and minor profile can be resolved. This examination additionally helps in building up the proportion for reproducible perception of the minor profile. Utilizing three combination sets in the accompanying proportions: 19:1, 9:1, 3:1, 1:1, 1:3, 1:9 and 1:19. The aggregate sum of DNA per response ought to be the ideal layout sum decided during the affectability study.

2.6 Stability Study

The impact of basic inhibitors including haematin, humic corrosive, and tannic corrosive were focused over a scope of various centralizations of each. Stock arrangements were set up by dissolving the fitting measure of inhibitor to create three distinct focuses including 1500 ng/µl, 3000 ng/µl and 4500 ng/µl for each. A concentration of 1 µl from each stock were included the advanced PCR response to arrive at a last required fixation; 100 ng/µl, 200 ng/µl and 300 ng/µl. Three male DNA controls, 007, Taqman and 9948 were examined in triplicate.

2.7 PCR Parameters Limits

To evaluate the effect of Magnesium Chloride ($MgCl_2$) and EDTA on the multiplex PCR reaction, titration experiment was conducted by amplifying 1 ng of 007, 9948 and Taqman male positive controls DNA separately with $MgCl_2$ and EDTA being added to the mix at different input amount. An additional concentration of each additive to PCR reaction were achieved by using stock solution of each as described in Table 1. In each experiment a positive control, containing same amount of DNA template without the addition of $MgCl_2$ and EDTA were prepared and analyzed parallelly with serial dilutions reactions (Table 1).

Table 1: Volumes of $MgCl_2$ and EDTA added for preparing a serial dilution of each

Required concentration of $MgCl_2$	Volume added to 25 ul PCR reaction using 50 mM $MgCl_2$ stock	Required concentration of EDTA	Volume added to 25 ul PCR reaction using 10 mM EDTA stock
2.5 mM	0.75 ul	0.1 mM	0.15 ul
5 mM	1.5 ul	0.25 mM	0.37 ul
7.5 mm	2.25 ul	0.5 mM	0.75 ul
10 mm	3.0 ul	1.0 mM	1.5 ul

2.8 Investigator® Argus Y-28

The Investigator Argus Y-28 Qiagen kit provided from the Thermofisher Scientific Company. In accordance with QIAGEN's ISO-certified Quality Management System, each lot of the Investigator Argus Y-28 QS Kit were tested against pre-determined specifications to ensure consistent product quality. Investigator Argus Y-28 QS Kits meet ISO 18385 requirements. It might be of multiplex application for Y-chromosomal Short Tandem Repeat (STR) loci analysis. The Investigator Argus Y-28 QS Kit Primer Mix contains two innovative internal PCR controls (Quality Sensor QS1 and QS2) to supply helpful information about the efficiency of the PCR and therefore the presence of PCR inhibitors. The standard sensors are amplified simultaneously with the polymorphic STR markers. The Investigator Argus Y-28 QS Kit was developed specifically for fast and reliable generation of male DNA profiles from mixtures of male and female DNA in order that separation of sperm from female cells or differential lysis is not required. The kit utilizes QIAGEN's fast-cycling PCR technology (Table 2). It provides highly robust results with inhibitor-resistant chemistry. The primers were fluorescence-labeled with these dyes:

- 6-FAM™: QS1, DYS389-I, DYS391, DYS389-II, DYS533, DYS390, DYS643, QS2.
- BTG: DYS458, DYS393, DYS19, DYS437, DYS449.
- BTY: DYS460, DYS576, YGATAH4, DYS481, DYS448, DYS518.
- BTR2: DYS439, DYS549, DYS438, DYS456.
- BTP: DYS570, DYS635, DYS385, DYS392, DYS627.

- Spectral Calibration/Matrix Generation

Before conducting DNA fragment size analysis, a spectral calibration with the 6 fluorescent labels 6-FAM, BTG, BTY, BTR2, BTP, and BTO were performed for each analyzer (Table 5). The calibration procedure creates a matrix, which was used to correct the overlapping of the fluorescence emission spectra of the dyes.

Spectral calibration was performed for each new capillary array, which comprises the following steps:

- Preparation of the instrument.
- Preparation of the standard calibration plate.
- Plate assembly and loading into the instrument.
- Software setup of dye set BT6.
- Performing a spectral calibration run.
- Checking the matrix.

Before the spectral calibration process, the spatial calibration was performed as described in the Applied Biosystems 3500/3500xL Genetic Analyzers User Guide.

Table 2: Standard cycling protocol recommended for all DNA samples

Temperature	Time	Number of cycles
96°C	12 min	–
96°C 60°C 72°C	10 s 1 min 25 s 5 s	30 cycles
68°C	5 min	–
60°C	5 min	–
10°C	∞	–

The following protocol was used for PCR amplification of STR loci from forensic samples using the Investigator Argus Y-28 QS Kit. All the reaction mixtures were set up in an area separate from that used for DNA isolation and PCR product analysis (post-PCR analysis). Disposable tips containing hydrophobic filters were used to minimize cross-contamination risks. Recommended amount of DNA under standard conditions is 0.5 ng were used as mentioned in the product guidelines. The tubes containing PCR components were vortexed and then centrifuge briefly to collect the contents at the bottom of the tubes.

Varying PCR Volumes for The Investigator Argus Y-28 QS Kit could be run with half reaction mix volumes (Fast Reaction Mix + Primer Mix).

PCR amplification from buccal swab lysates. This protocol was used for direct PCR amplification of STR loci from crude lysates of buccal swabs using the Investigator Argus Y-28 QS Kit.

The initial STR profile for the DNA positive control 9948 illustrates that the alleles located on each maker could be used as reference during doing mixture study or other tests (Qiagen, 2020).

Table 3: Allele assignment of the investigator Argus Y-28 QS Kit

Locus	DNA 9948
DYS19	14
DYS385	11/14
DYS389-I	13
DYS389-II	31
DYS390	24
DYS391	10
DYS392	13
DYS393	13
DYS437	15
DYS438	11
DYS439	12
DYS448	19
DYS456	17
DYS458	18
DYS449	30
DYS460	11
DYS481	24
DYS518	38
DYS533	12
DYS549	13
DYS570	18
DYS576	16
DYS627	22
DYS635	23
DYS643	11
YGATAH4	12

In general, the electropherogram shows an individual peaks for every Y-STR locus. In some case, locus DYS385 produces two peaks of various lengths or of a similar length (Table 3). These two parts start from copied and inversed duplicates of one Y-chromosomal locus. The initial data from the preliminary works given in the Specialist Argus Y-28 QS Unit could also co-enhance the two homologous loci. In the event that more than one peak in the electropherogram for one or a few markers, this

does not really propose blended examples. Duplications or triplications of STR markers additionally bring about such an impact and have effectively been noticed for DYS385 and DYS19. Infrequently, single frameworks can likewise fall flat as a result of Y-chromosomal erasures as known in azoospermic patients, which has effectively been depicted for DYS385 and DYS392.

Appropriate user guides for the GeneMapper ID-X Software were referred for the automatic sample analysis. Finding the exact lengths of the amplified products depends on the device type, the conditions of electrophoresis, and the DNA size standard used. Due to the complexity of some loci, size determination was based on evenly distributed references. DNA Size Standard 24plex (BTO) was used with the following fragment lengths:

DNA Size Standard 24plex (BTO): 60, 80, 90, 100, 120, 140, 160, 180, 200, 220, 240, 250, 260, 280, 300, 320, 340, 360, 380, 400, 425, 450, 475, 500, 525, and 550 bp (Figure 4).

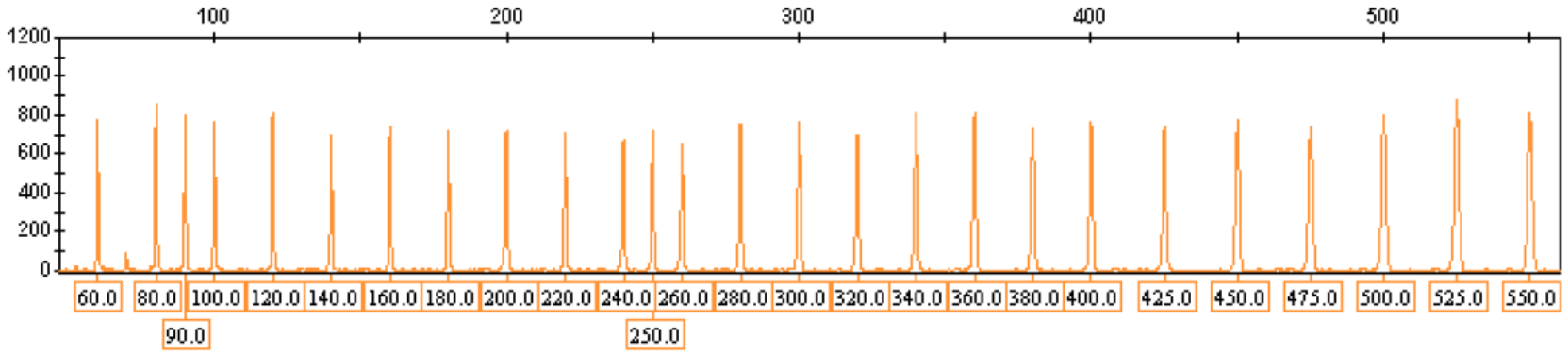


Figure 4: Electropherogram of the DNA size standard 24plex (BTO)
Note: Fragment's lengths in bp

The allelic ladder contains 2 alleles for each quality sensor (QS1 and QS2) (Qiagen, 2020). This allows automated calling of the QS peaks for sample analysis. The Specialist Argus Y-28 QS Pack contains two inward PCR controls (quality sensors QS1 and QS2), which give supportive data about the PCR intensification proficiency by and large and about the presence of PCR inhibitors. The inside quality sensors are encased in the preliminary blend and were enhanced all the while with the polymorphic STR markers. The quality sensors were marked with BTP and show up at part sizes of 74 bp (QS1) and 435 bp (QS2).

To address the issue of redundancy in similarities and to limit the chance of ambiguity, a manufactured internal control DNA layout was executed utilizing an arbitrary calculation. The layout succession varies from all realized DNA groupings, and specifically bears no closeness to human DNA. The possibility of vague restricting with regards to a multiplex PCR intensification response is in this way extremely low.

All the data were expressed in mean ± SD and analysed through one-way analysis of variance (ANOVA) using IBM-SPSS statistical analysis tool. The level of significance was set at $p<0.05$.

Chapter 3: Results

3.1 Statistical Analysis

Data analysis were carried out utilizing GeneMapper software. (Thermo Fisher Scientifics, USA). It dissects information produced from the narrow electrophoresis stages, for example, 3500 hereditary analyses in this venture. The GeneMapper software was utilized to break down the outcomes and create crude information for all samples. The preliminary informational document produced from the GeneMapper was investigated utilizing Microsoft Excel and IBM SPSS to analyze statistical significance through one-way ANOVA to track down the huge contrasts between tests.

Table 4: Sensitivity test of (1 ng) DNA for positive male controls (2800, 9948, R1, R2, Taqman)

Descriptive								
	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
2800 1	28	1476.43	585.981	110.740	1249.21	1703.65	169	2626
9948 1	28	1157.04	346.956	65.569	1022.50	1291.57	502	1994
R1 1	28	4784.07	1811.378	342.318	4081.69	5486.45	496	7904
R2 1	28	868.32	313.834	59.309	746.63	990.01	210	1472
TM 1	28	2123.04	820.741	155.105	1804.79	2441.29	0	3397
Total	140	2081.78	1700.415	143.711	1797.64	2365.92	0	7904

Table 4 shows positive controls (1 ng) which addresses the complete of DNA focus. The main position goes for (R1 1), where the mean is 4784.07 with standard deviation 1811.378 and standard error 342.318, the lower and upper bound of the 95%

certainty stretch for mean are 4081.69 and 5486.45 separately. Followed by (TM 1), where the mean is 2123.04 with standard deviation 820.741 and standard error 155.105, the lower and upper bound of the 95% certainty stretch for mean are 1804.79 and 2441.29 individually. The least significant goes for (R2 1), where the mean is 868.32 with standard deviation 313.834 and standard error 59.309, the lower and upper bound of the 95% certainty stretch for mean are 746.63 and 990.01 individually. The ANOVA shows that there is a huge contrast between certain controls, where Sig. is under 0.05 as shown in Table 5.

Table 5: ANOVA test for sensitivity test of (1 ng) DNA positive male controls

ANOVA					
	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	279948621.386	4	69987155.346	77.472	0.000
Within Groups	121957700.750	135	903390.376		
Total	401906322.136	139			

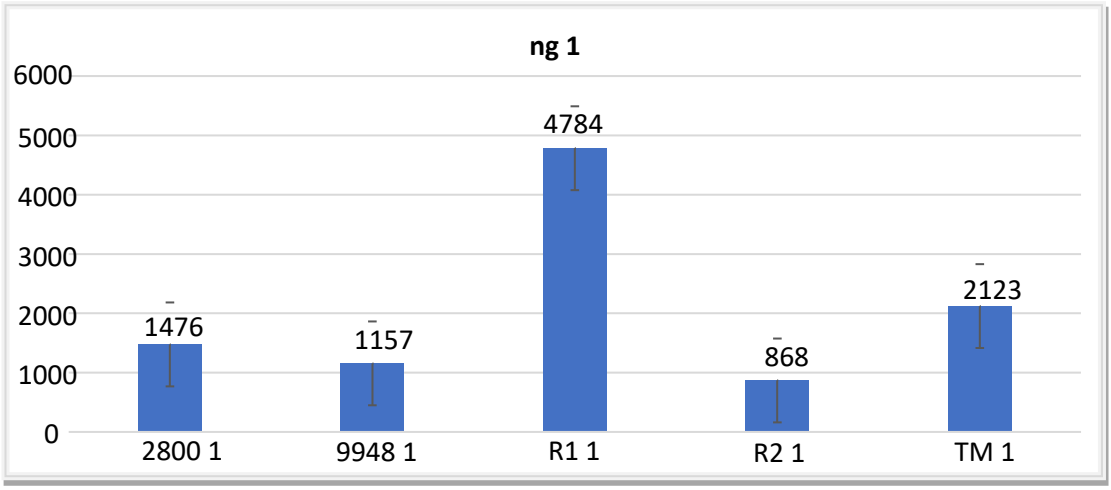


Figure 5: The average peaks height of (1 ng) DNA for positive male controls (2800, 9948, R1, R2, Taqman)

As known, the error bars pass on the variety in the information and the precision of the mean measurement and long error bar would show that the qualities are more fanned out and less dependable. The error bars at the bar diagram above are short which implies that the qualities are concentrated, it affirmed that plotted arrived at the midpoint of significant worth is more probable. Moreover, (R2 1) is less exact than others as demonstrated in Figure 5.

Table 6: Sensitivity test of (0.500 ng) positive male controls (2800, 9948, R1, R2 and Taqman)

Descriptive								
Height 2								
	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
2800 2	17	572.41	387.576	94.001	373.14	771.69	188	1485
9948 2	12	513.00	134.611	38.859	427.47	598.53	279	702
R1 2	24	7056.33	8105.127	1654.452	3633.84	10478.83	125	25119
R2 2	24	1521.21	386.571	78.908	1357.97	1684.44	707	2282
TM 2	25	1940.96	829.810	165.962	1598.43	2283.49	249	3248
Total	102	2649.73	4636.076	459.040	1739.11	3560.34	125	25119

Table 6 addresses peak 2 for positive controls (0.500 ng) which illustrates the complete of DNA focus. The principal rank goes for (R1 2), where the mean is 7056.33 with standard deviation 8105.127 and standard error 1654.452, the lower and upper bound of the 95% certainty span for mean are 3633.84 and 10478.83 individually, while most minimal position goes for (9948 2), where the mean is 513 with standard deviation 134.611 and standard error 38.859, the lower and upper bound of the 95% certainty stretch for mean are 42747 and 598.53 separately. The analysis done in

ANOVA shows that there is a significant difference between positive controls, where significant level was $p<0.05$ as shown in Table 7.

Table 7: ANOVA test for sensitivity test of (0.500 ng) positive male controls

ANOVA					
Peak 2					
	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	637306659.944	4	159326664.986	10.078	0.000
Within Groups	1533506692.369	97	15809347.344		
Total	2170813352.314	101			

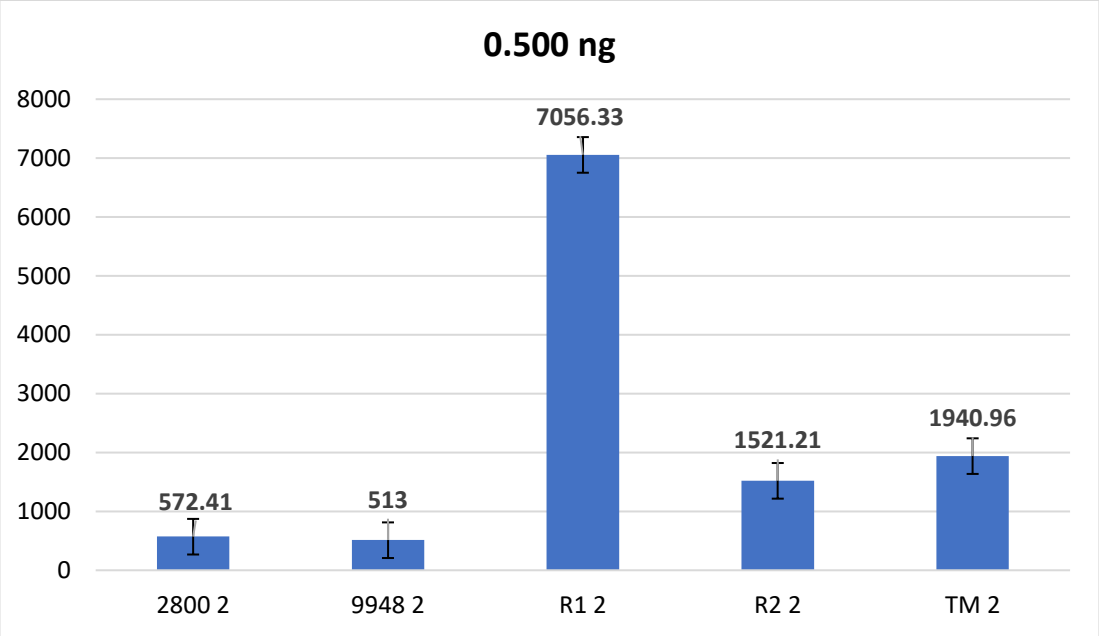


Figure 6: The average peaks height of (0.500 ng) DNA for positive male controls (2800, 9948, R1, R2, Taqman)

Figure 6 shows data for 0.500 ng DNA for positive male controls. The error bars idicates that (9800 2), (9948 2) and (R1 2) are more precise than others.

Table 8: Sensitivity test of (0.250 ng) DNA for positive male controls (2800, 9948, R1, R2, Taqman)

Descriptive								
	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
2800 3	11	552.91	129.230	38.964	466.09	639.73	308	709
9948 3	1	85.00	.	.	.	.	85	85
R1 3	21	1281.10	763.796	166.674	933.42	1628.77	494	4254
R2 3	19	757.32	332.814	76.353	596.90	917.73	93	1433
TM 3	16	794.94	242.930	60.733	665.49	924.39	440	1283
Total	68	884.97	550.456	66.753	751.73	1018.21	85	4254

Table 8 shows positive controls adding (0.250 ng) which addresses DNA fixation. Statistically, the most significant was observed to be (R1 3), where the mean is 1281.1 with standard deviation 763.796 and standard error 166.674, the lower and upper bound of the 95% certainty stretch for mean are 933.42 and 1628.77 individually, while for the least significant was (2800 3), where the mean is 552.91 with standard deviation 129.23 and standard error 38.964, the lower and upper bound of the 95% certainty span for mean are 466.09 and 639.73 separately. The ANOVA shows that there is a contradictory finding between control test DNA samples, where significance was less than 0.05 as shown in Table 9.

Table 9: ANOVA test for sensitivity test of (0.250 ng) DNA positive male controls

ANOVA					
	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	5587390.180	4	1396847.545	5.981	0.000
Within Groups	14713701.761	63	233550.822		
Total	20301091.941	67			

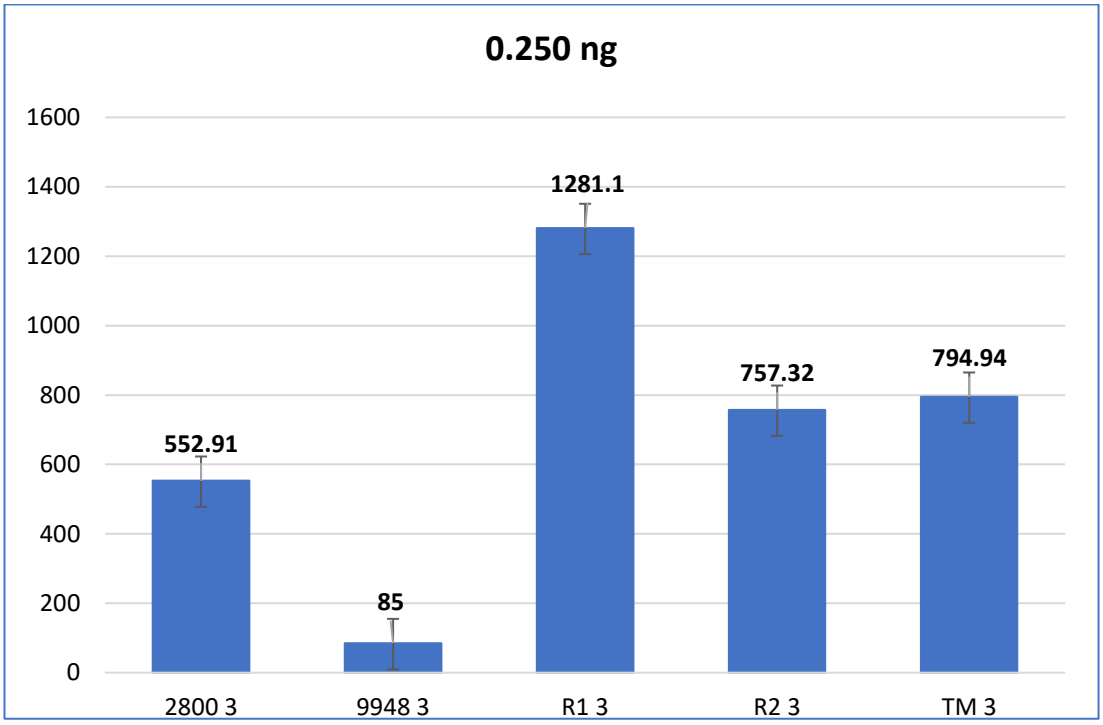


Figure 7: The average peaks height of (0.250 ng) DNA for positive male controls (2800, 9948, R1, R2, TaqMman)

Figure 7 shows data for 0.250 ng DNA for positive male controls. The error bars are short for all groups, which means that the values are concentrated. Moreover, the error bars most often represent the Standard Deviation (SD) of a data set for positive male controls 9948.

Table 10: Mean of the pure samples after addition of Magnesium Chloride ($MgCl_2$) for 007, 9948 and TM

Mg				Positive Mean	Height 1 Mean
Sample Name	007	Condition	0.75 ul $MgCl_2$	2870.82	2594
			1.5 ul $MgCl_2$		2512
			2.25 ul $MgCl_2$		900
			3.0 ul $MgCl_2$		207
	9948	Condition	0.75 ul $MgCl_2$	3196.61	2742
			1.5 ul $MgCl_2$		1421
			2.25 ul $MgCl_2$		1793
			3.0 ul $MgCl_2$		711
	TM	Condition	0.75 ul $MgCl_2$	.	1484
			1.5 ul $MgCl_2$	.	187
			2.25 ul $MgCl_2$	2138.04	360
			3.0 ul $MgCl_2$		894

Table10 shows the mean of the pure samples after addition of $MgCl_2$, it can be seen that more statistical significance was observed in the the sample (9948) with 3196.61 especially for 0.75 ul $MgCl_2$ with mean 2742, followed by other samples as evidenced in the table. Table 11 shows interesting observations after the addition of $MgCl_2$ in µl.

Table 11: Descriptive statistics of different concentrations for $MgCl_2$

Descriptive statistics

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
0.75 ul MgCl2	84	2273.26	970.148	105.852	2062.73	2483.80	350	4852
1.5 ul MgCl2	81	1417.09	1152.004	128.000	1162.36	1671.82	58	4377
2.25 ul MgCl2	61	1265.64	804.710	103.033	1059.54	1471.74	72	4662
3.0 ul MgCl2	82	607.43	530.834	58.621	490.79	724.06	61	2576
Total	308	1405.04	1086.863	61.930	1283.18	1526.90	58	4852

Table 12: ANOVA analysis different concentrations of $MgCl_2$

ANOVA Analysis

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	116684433.847	3	38894811.282	48.072	.000
Within Groups	245965636.760	304	809097.489		
Total	362650070.607	307			

The ANOVA analysis as shown in Table 12, it can be observed that there are significant differences between samples, where calculated F is 48.072 with a level of statistical significance less than 0.05.

Table 13: Descriptive statistics positive controls (007, 9948 and TM)

	N	Mean	Standard Deviation	Standard Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
007	111	1565.45	1251.293	118.767	1330.08	1800.82	61	4377
9948	111	1675.20	993.694	94.317	1488.28	1862.11	156	4852
TM	86	849.29	721.787	77.832	694.54	1004.04	58	3013
Total	308	1405.04	1086.863	61.930	1283.18	1526.90	58	4852

Table 13 illustrates descriptive statistics of the three samples of positive male control 007, 9948 and TM. According to 007, the mean is 1565.45 with standard deviation 1251.293 and standard error 118.767. For 9948, the mean is 1675.2 with standard deviation 993.694 and standard error 94.317. For TM, the mean is 849.29 with standard deviation 721.787 and standard error 77.832.

Table 14: ANOVA analysis the three samples of positive male control

ANOVA Analysis					
	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	37519309.757	2	18759654.879	17.598	0.000
Within Groups	325130760.850	305	1066002.495		
Total	362650070.607	307			

By reference to the consequences of the ANOVA analysis (Table 14), it tends to be seen that there are huge contrasts between tests, where determined F is 17.598 with a degree of factual importance under 0.05.

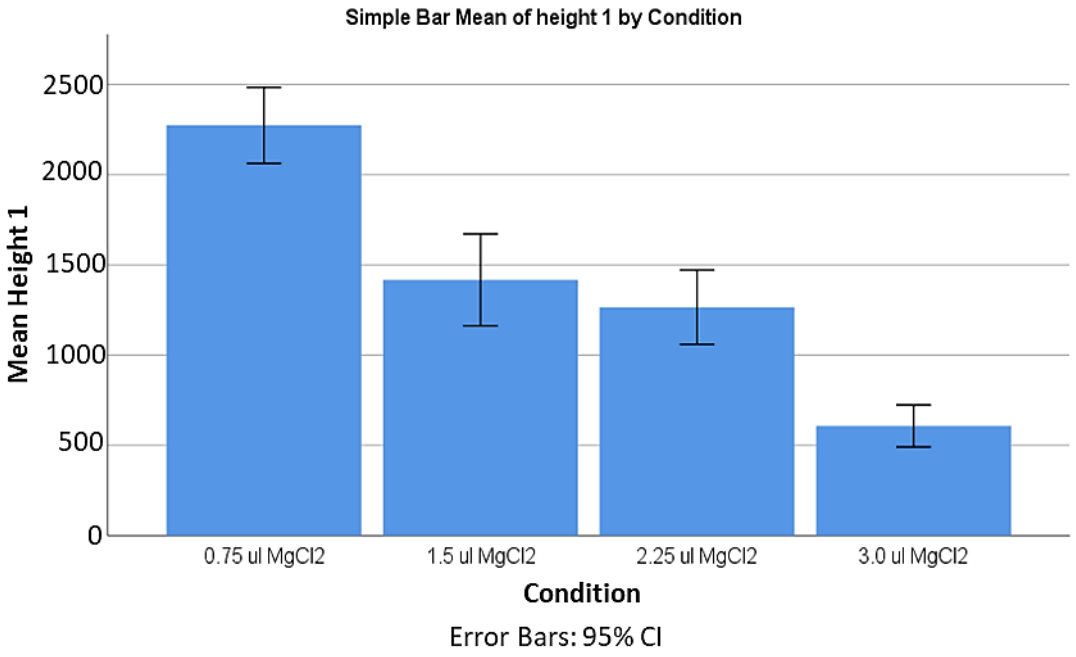


Figure 8: Heights of the four conditions of $MgCl_2$ and their errors

Figure 8 shows the mean peaks of the four conditions and their errors. Statistically significant peaks were observed in the 0.75 ul $MgCl_2$ with mean 2273.26 followed by 1.5 ul $MgCl_2$ with mean 1417.09. the third rank goes to 2.25 ul $MgCl_2$

with mean 1265.64. The least significance was observed in 3.0 ul $MgCl_2$ with mean 607. 43. In this case, all conditions are somewhat same in spreading around the mean value except for 3.0 ul $MgCl_2$, it appears that measurements in that group are more precise than others.

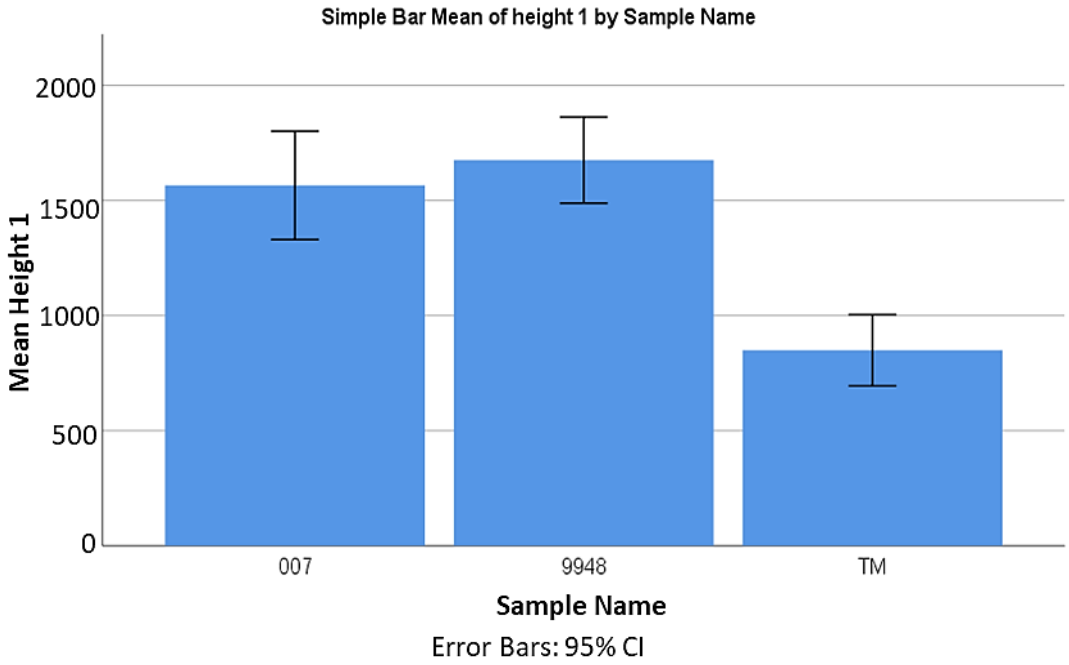


Figure 9: The mean heights of the of three sample (007, 9948 and Taqman)

Table 15: The mean of the pure samples after addition of EDTA for (007, 9948 and TM)

EDTA				Positive Mean	Height 1 Mean
Sample Name	007	Condition	0.15 ul EDTA	2870.82	3373
			0.37 ul EDTA		5102
			0.75 ul EDTA		5157
			1.5 ul EDTA		2157
	9948	Condition	0.15 ul EDTA	3196.61	2763
			0.37 ul EDTA		2560
			0.75 ul EDTA		2547
			1.5 ul EDTA		2346
	TM	Condition	0.15 ul EDTA	2138.04 .	1055
			0.37 ul EDTA		1695
			0.75 ul EDTA		.
			1.5 ul EDTA		.

Figure 9 shows the mean heights of the of three sample (007, 9948 and Taqman). Table 15 shows the mean of the pure samples after addition of EDTA, it could be seen that high statistical significance was observed in the sample (9948) with mean 3196.61 especially after adding 0.15 µl EDTA with mean 2763, followed by other samples. Table 16 shows descriptive statistics of EDTA peaks measured in microliter including 0.15 ul EDTA, 0.37 ul EDTA, 0.75 ul EDTA and 1.5 ul EDTA.

Table 16: Descriptive statistics of EDTA height including positive controls (007, 9948 and TM)

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
0.15 ul EDTA	84	2397.31	1322.216	144.266	2110.37	2684.25	235	5510
0.37 ul EDTA	84	3118.69	1801.288	196.537	2727.79	3509.59	426	8143
0.75 ul EDTA	56	3852.04	2037.744	272.305	3306.32	4397.75	61	8160
1.5 ul EDTA	56	2251.57	886.046	118.403	2014.29	2488.86	462	4843
Total	280	2875.52	1677.759	100.265	2678.15	3072.89	61	8160

Table 17: ANOVA analysis different concentrations of EDTA

ANOVA Analysis					
	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	99378746.324	3	33126248.775	13.328	0.005
Within Groups	685971405.548	276	2485403.643		
Total	785350151.871	279			

According to the results of the ANOVA analysis (Table 17), it can be observed that there are significant differences between samples, where calculated F is 13.328 with a level of statistical significance less than 0.05.

Table 18: Descriptive statistics positive controls (007, 9948 and TM)

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
007	112	3947.26	1907.270	180.220	3590.14	4304.38	61	8160
9948	112	2553.97	901.563	85.190	2385.16	2722.78	462	4894
TM	56	1375.14	674.958	90.195	1194.39	1555.90	235	2786
Total	280	2875.52	1677.759	100.265	2678.15	3072.89	61	8160

Table 18 illustrates descriptive statistics of the positive control samples (007), (9948) and (TM). According to (007), the mean is 3947.26 with standard deviation 1907.270 and standard error 180.220. For (9948), the mean is 2553.97 with standard deviation 901.563 and standard error 85.190. For TM, the mean is 2875.52 with standard deviation 1677.759 and standard error 100.265.

Table 19: ANOVA test for EDTA of positive controls (007, 9948 and TM)

ANOVA Analysis					
	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	266289234.604	2	133144617.302	71.053	0.000
Within Groups	519060917.268	277	1873866.127		
Total	785350151.871	279			

Regarding to the results of the ANOVA analysis (Table 19), it can be seen that there are significant differences between samples, where calculated F is 71.053 with a level of statistical significance less than 0.05.

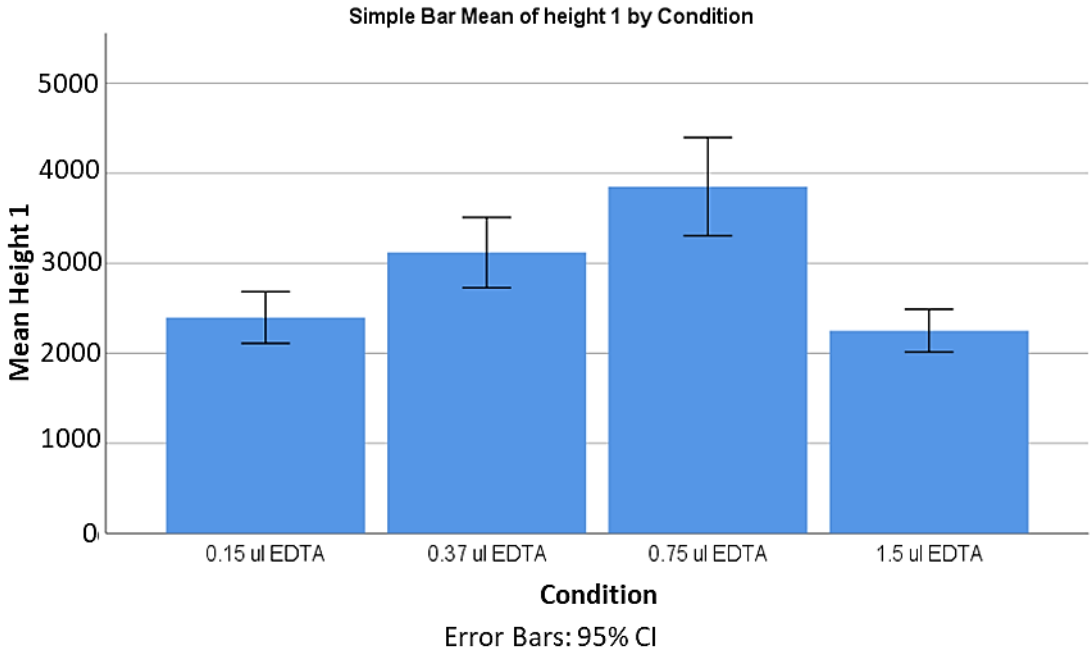


Figure 10: Height of the four conditions of EDTA and their errors

The bar chart in Figure 10 shows the mean height of the four conditions and their errors.

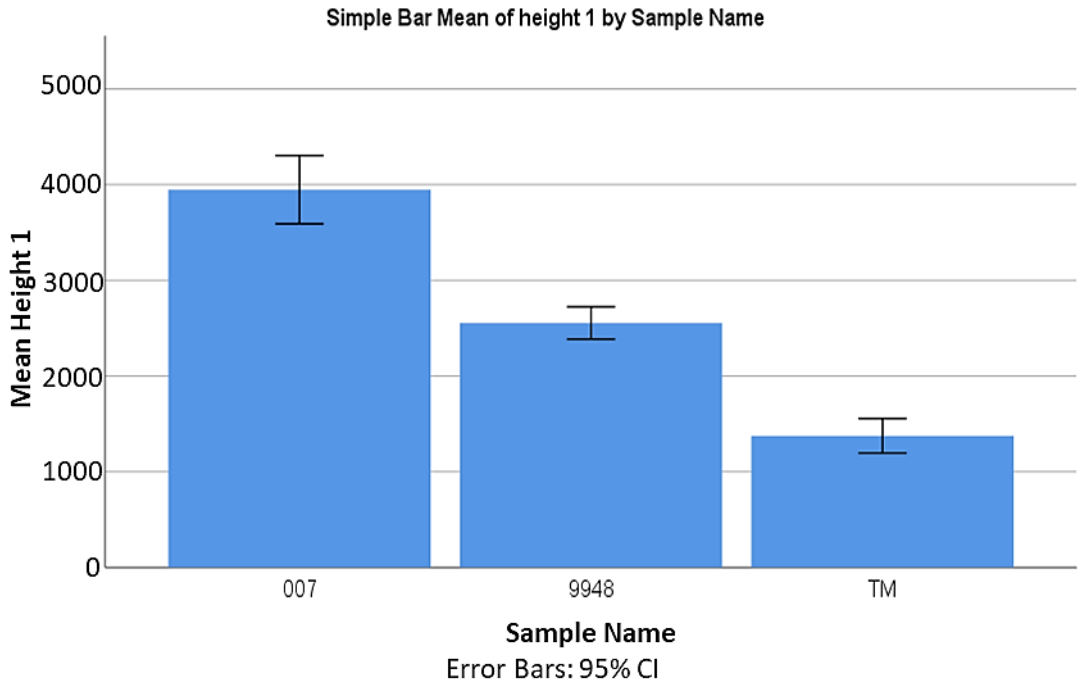


Figure 11: The mean heights of the of three sample (007,9948 and Taqman)

The bar chart in Figure 11 shows the mean height the three samples and their errors.

3.2 Profiling Sample

All the test samples were amplified using PCR with the addition and modifications of various parameters including DNA sample size, EDTA and $MgCl_2$. Since YSTR repeats are unique to the individuals, the obtained results were subjected for Capillary electrophoresis, which express the presence or the absence of the specific YSTR in terms of observable peaks. Numer of the peaks give the information about the repeated occurrence and the height of the peaks give information about genetic dominance, recessive and co-dominance. The peak observed for various parameters through capillary electrophoresis analysis are shown in the Figure 12 (positive control); Figure 13 (mixture study 007 and Taqman); Figure 14 (off ladder alleles for male positive controls) and Figure 15 (Stutter peaks).

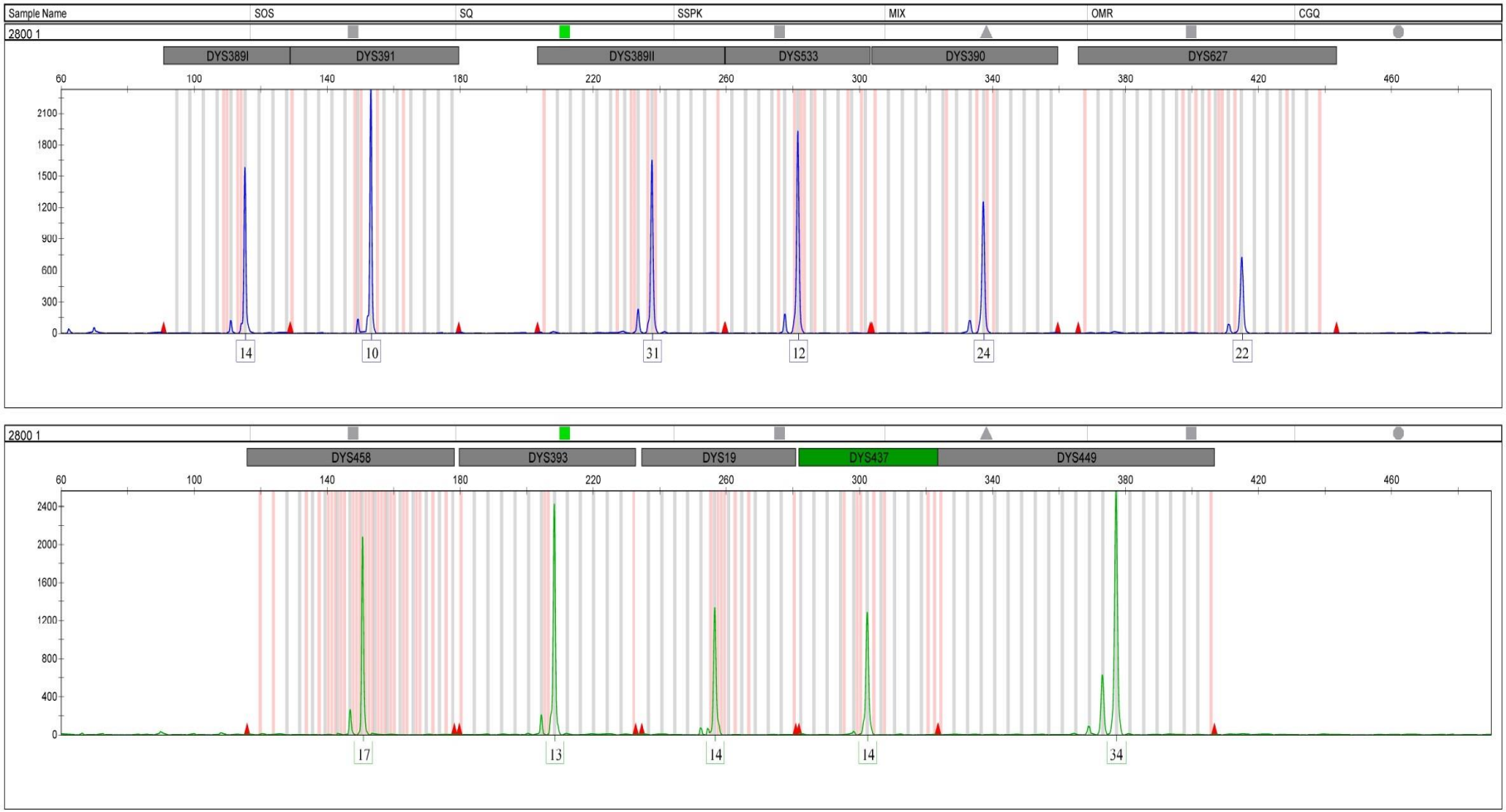


Figure 12: Electropherogram of the positive control DNA 2800 analyzed on an Applied Biosystems 3500xL Genetic Analyzer
Source: (Qiagen, 2020)

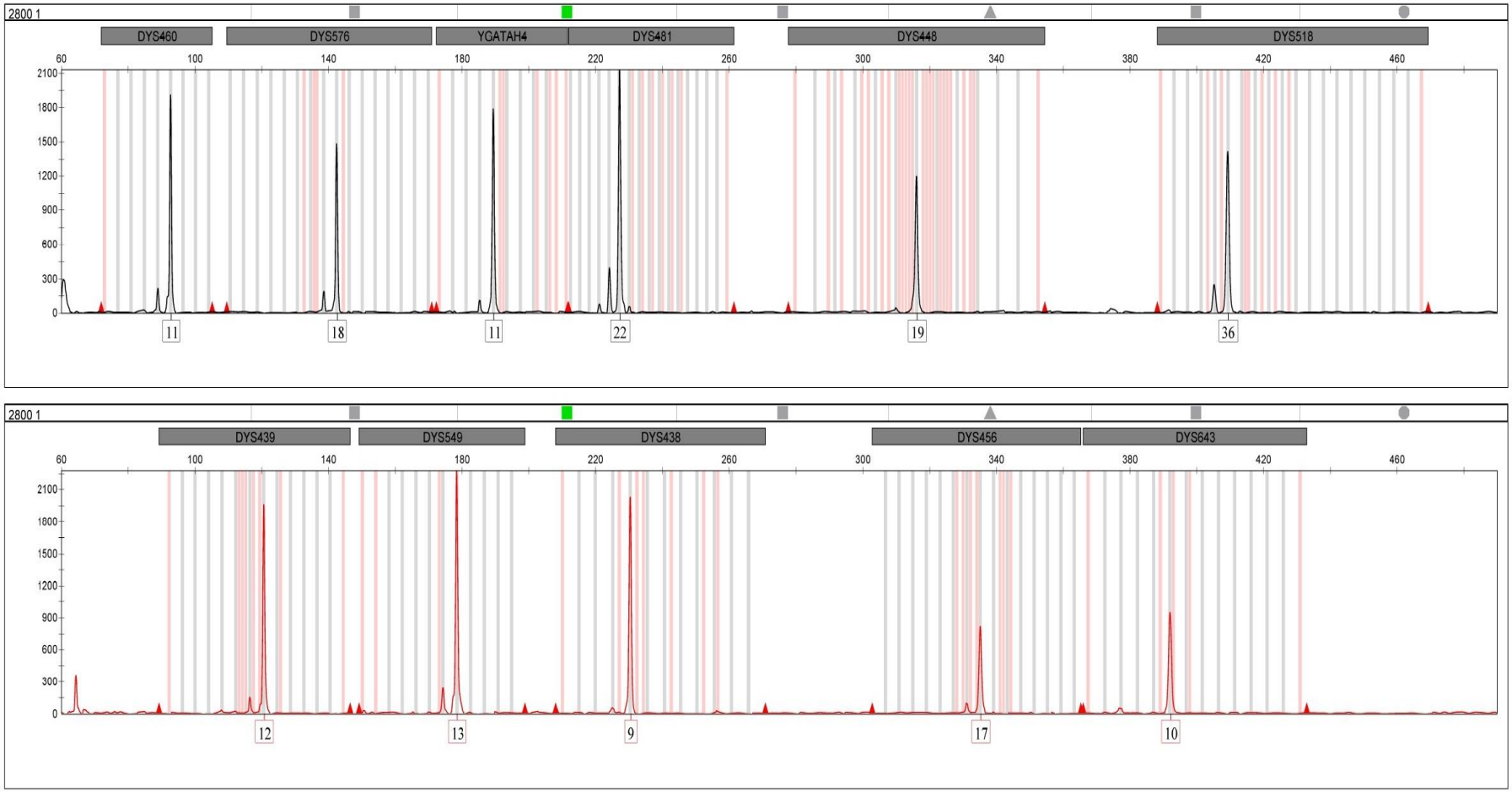


Figure 12: Electropherogram of the positive control DNA 2800 analyzed on an Applied Biosystems 3500xL Genetic Analyzer (continued)
Source: (Qiagen, 2020)

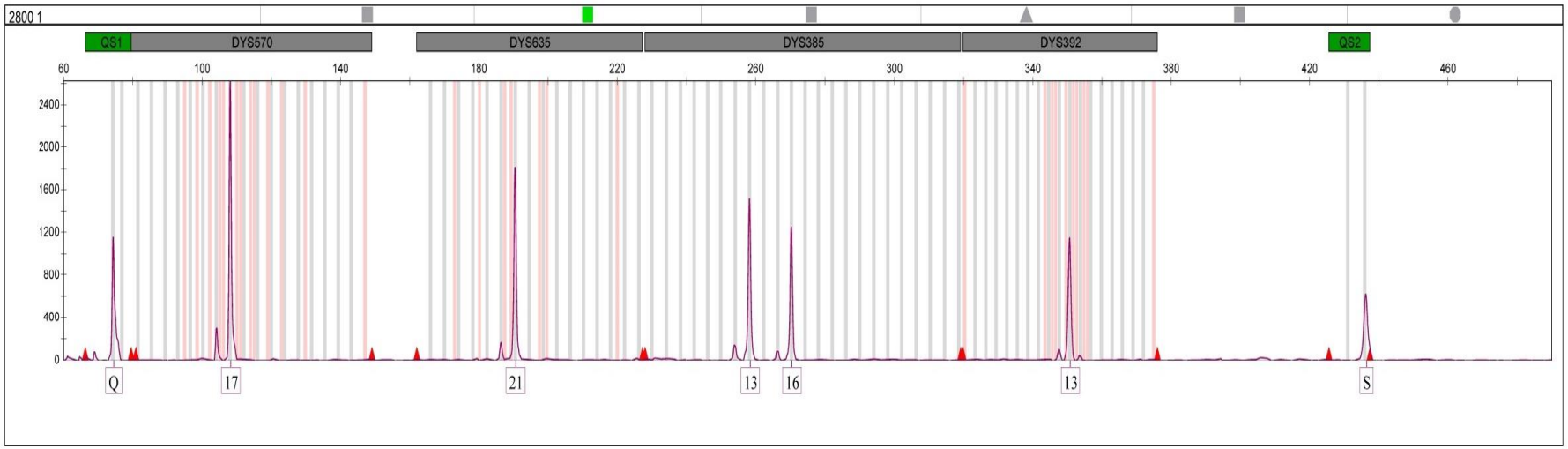


Figure 12: Electropherogram of the positive control DNA 2800 analyzed on an Applied Biosystems 3500xL Genetic Analyzer (continued)
Source: (Qiagen, 2020)

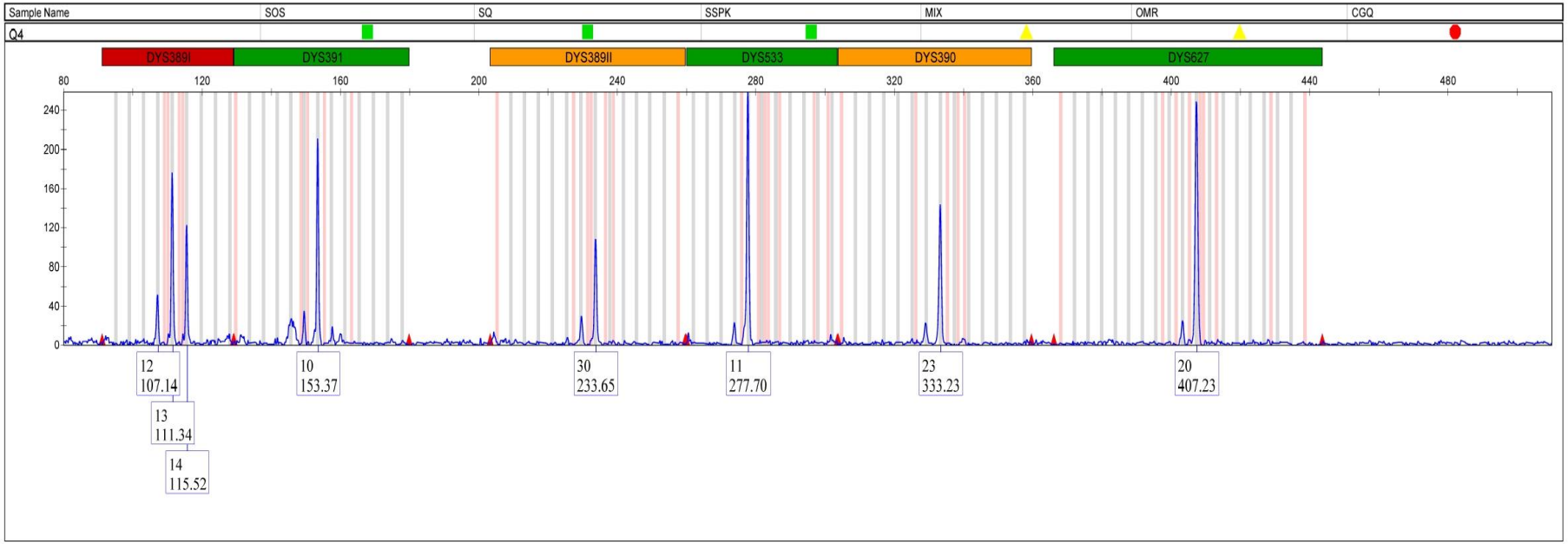


Figure 13: Result of mixture study electropherogram for 007 and Taqman
Source: (Qiagen, 2020)

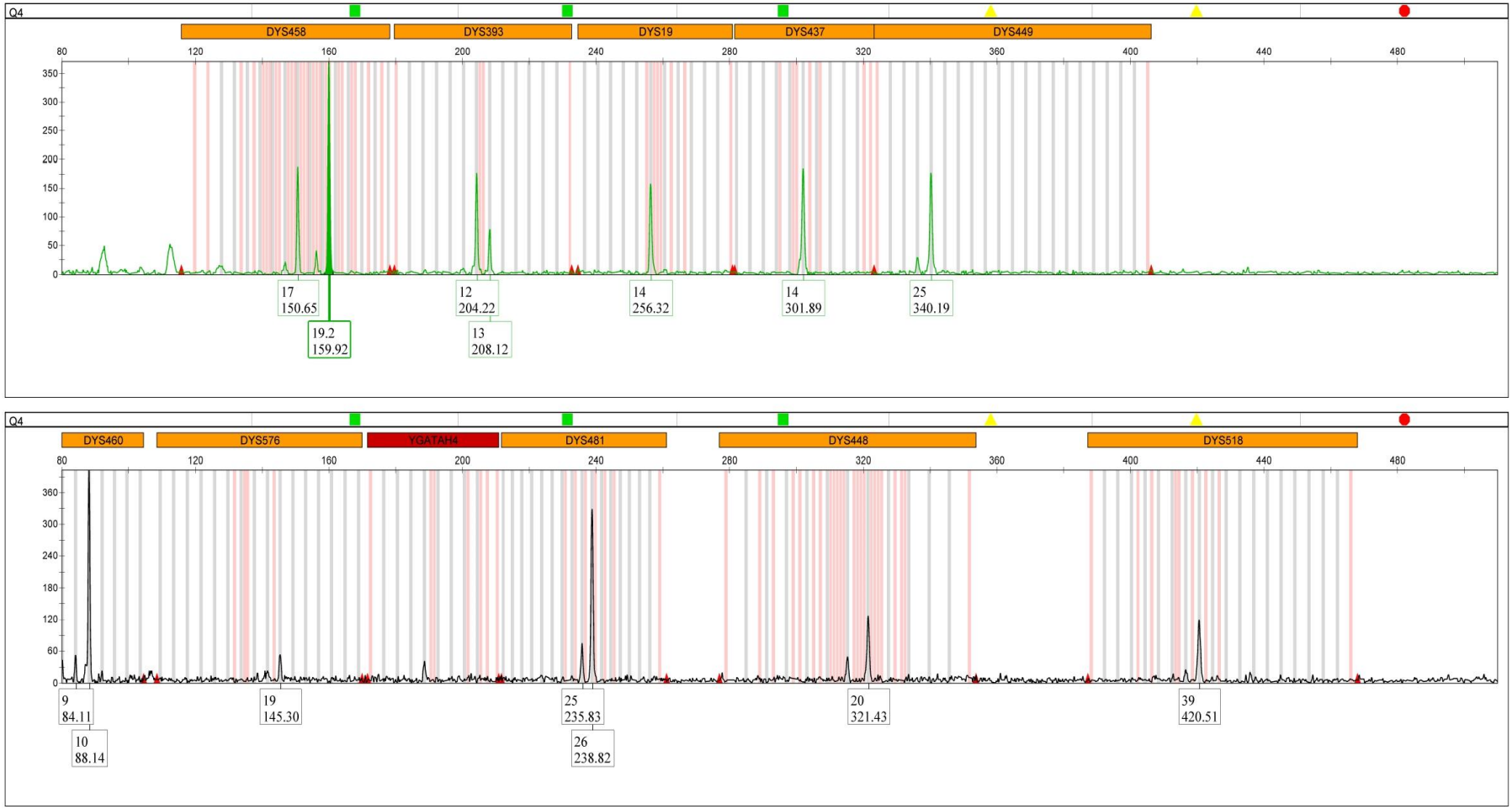


Figure 13: Result of mixture study electropherogram for 007 and Taqman (continued)
Source: (Qiagen, 2020)

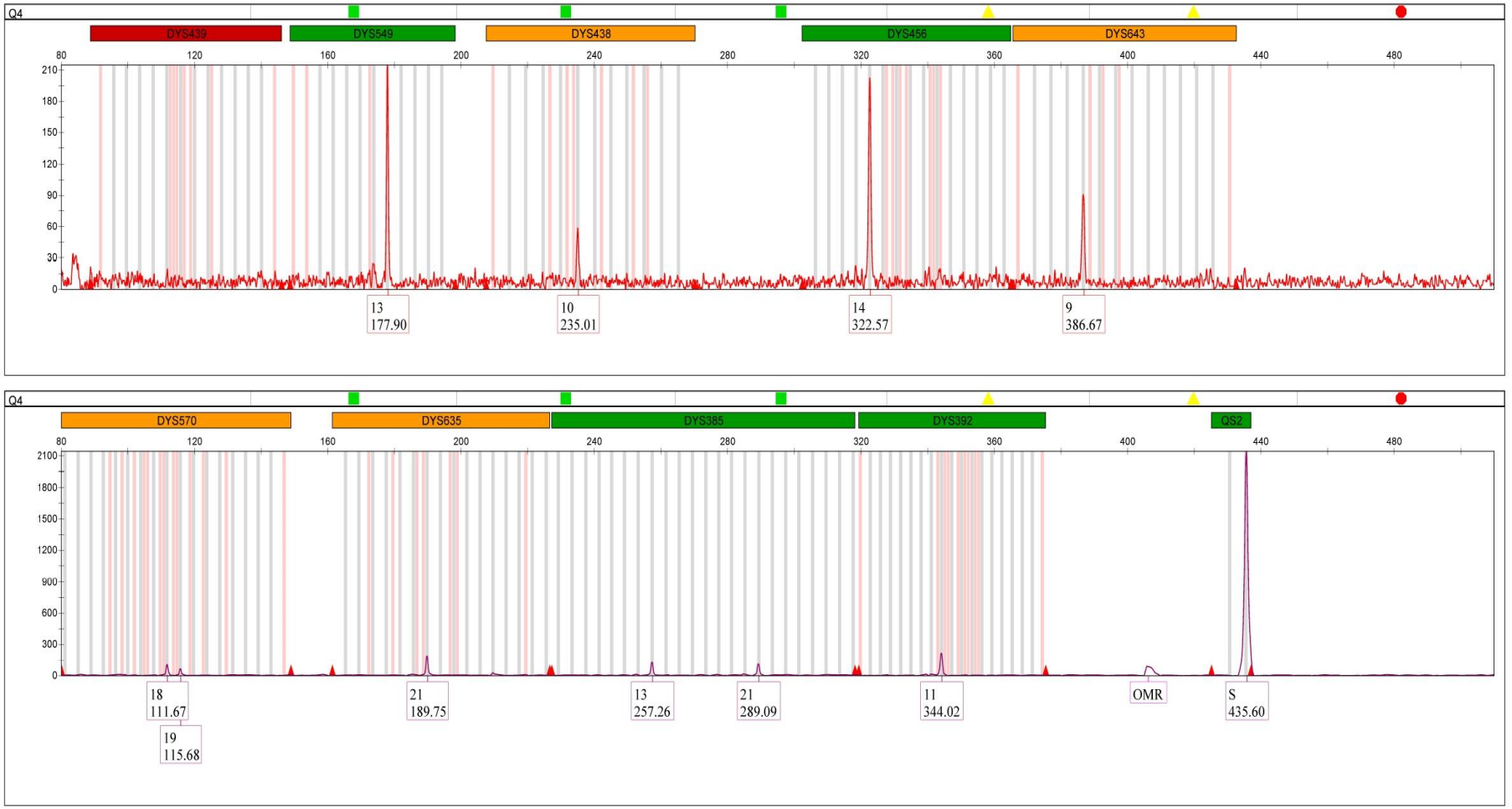


Figure 13: Result of mixture study electropherogram for 007 and Taqman (continued)
Source: (Qiagen, 2020)

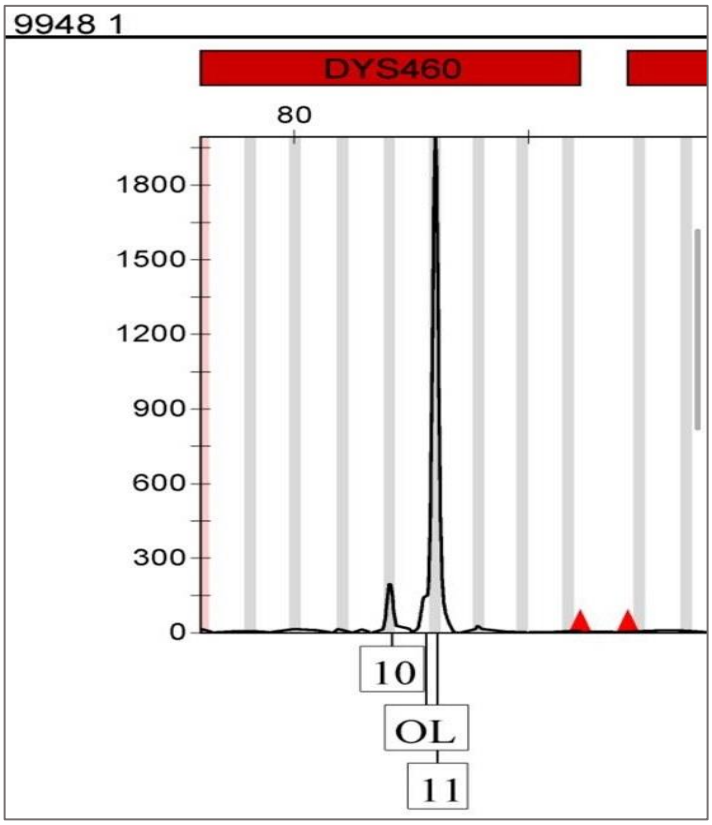


Figure 14: Off-ladder alleles for male positive control 9948

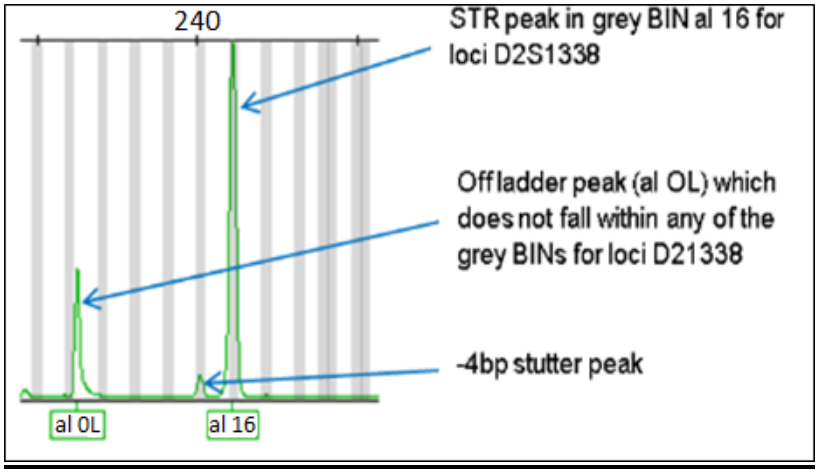


Figure 15: An example of the stutter peaks

3.3 Mock Test

Mock casework samples DNA from mock casework samples were extracted and amplified with Qiagen Kit. DNA inputs differed in the case of highly dilute samples (<0.1 ng/μL), as the kit have different maximum sample input volumes (25 μL). Electropherogram profiles were analyzed and total alleles recovered for each sample. In this study, saliva, blood, cell, sperm and epithelial cell with different concentrations were used for the analysis. Table 20 shows sperm is highest average peak heights than other samples the lowest was saliva 365.22, this is due to the concentration of sample. Figure 16 shows the average peak height of the mock casework samples (sperm, blood, cell, sperm+epithelial, epithelial and saliva).

Table 20: Mock casework samples quantification values for the extracted samples presented, along with DNA input amounts for the STR amplification reactions

Sample	Name	Sample DNA (ng/μL)	DNA input (ng/reaction) 0.5 μL	Average peak height
Q-E2	sperm	1.1751	1.00	22380.44828
Q-B	blood	0.5517	1.00	17440.27586
Q-C	Cells: swabs from phone and keyboard	0.2034	1.00	1998.04878
Q-E1+2	sperm+epithelial	0.1585	1.00	1498.296296
Q-E1	epithelial	0.0272	1.00	452.3953488
Q-S	saliva	0.0087	1.00	365.2285714

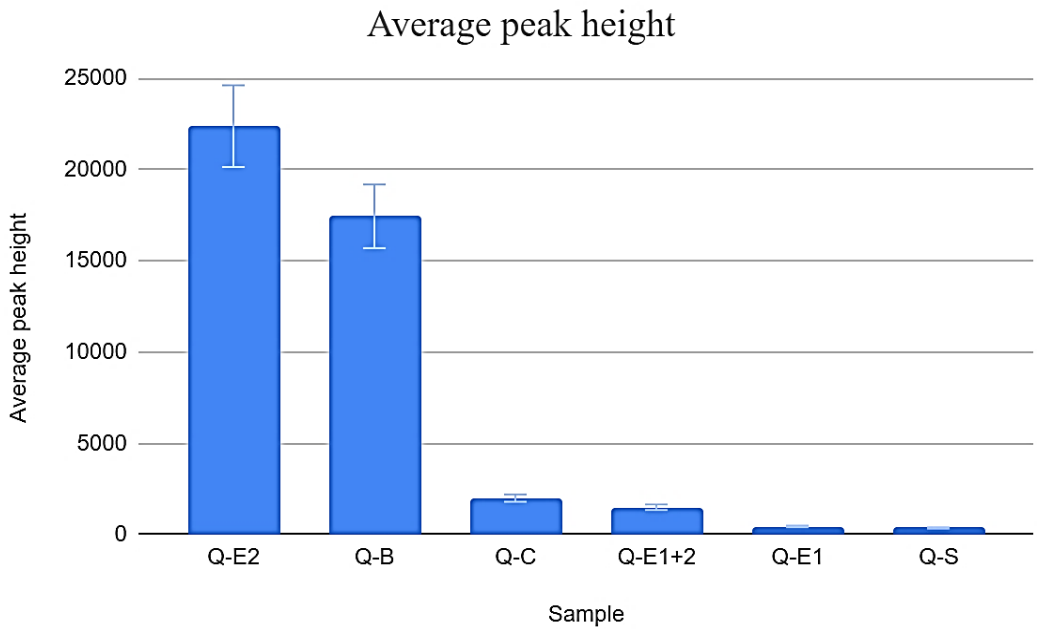


Figure 16: The average peak height of the mock casework samples (sperm, blood, cell, sperm+epithelial, epithelial and saliva)

Chapter 4: Discussion

The investigator Argus Y-28 kit had optimal amount of DNA under standard conditions was 0.5 ng and the internal validations demonstrated reliable results with <0.1 ng DNA. The experiments reveal significant results with <0.1 ng DNA. The sensitivity test for the positive controls (9948, Taqman, R1, R2, 2800) showed that Qiagen kit gave results with DNA in between (1 ng, 0.500 ng, 0.250 ng). As representation of 2800 STR profile for positive control 2800 used 1 ng DNA shows the extraordinary results (Figure 12). The drop out started to showed up in 0.125 ng, 0.06 ng and 0.03 ng.

The Qiagen Argus Y-28 Specialist kit showed extraordinary results in the sensitivity study. The kit exhibited best results at 0.500 ng, with a height around 2,180. A similar height (2,124) was found in instances of 1.0 ng, at any rate these models conveyed a higher number of waver tops than 0.5 ng of DNA. This was especially clear when 1 ng of control test R1 - that has a gathering of 1.627 ng/ul - made a high of 73 stammer tops as the Argus Y-28 pack handbook proposes using 0.5 ng of DNA during PCR. The unit 0.5 ng performed well as a rule for the kit, sorting out some way to recover adequate markers as low as 0.06 ng. Regardless on whole, below 0.06 ng made less marker recovery, for all intents and purposes none. It is at this point possible to convey better results by perhaps increasing the PCR cycles.

Test R1 for 1 ng of DNA this model has higher typical of stammers per marker. This might be R1 has a concnetration of 1.627 ng per smaller than normal. There are drop out in the model (2800 and 9948) in 0.03 ng - 0.06 ng of DNA. Marker DYS627 reliably drop out Locus Information on DYS627 (Change rate 1.51e-02 (55 out of 3650) taking into account (YHRD) in the marker DYS460.

The error bars represented the following information in the one way ANOVA analysis:

- The standard deviation bar is larger spread where the data are more variable from the mean value because amount of DNA is 0.250 ng, which could not be detected through 3500 genetic analyzers.
- It is important to note the reliability as representative number for the data set is less reliable possibly, because Hi-Di Formamide in DNA sample help to stop reannealing of DNA single strands after denaturation at high temperature 95° C evaporate from the sample well during data collection.

4.1 PCR Parameters

In this experiment, 1 ng of 3 male (007, 9948, Taqman) controls was utilized with various measures of $MgCl_2$ and EDTA. The PCR limitations study meant to consider the impact of $MgCl_2$ and EDTA on outcomes. $MgCl_2$ is known to enahnce DNA amplification during PCR, while EDTA is a well known anticoagulant utilized in blood tests.

The Electropherograms (EPGs) showed no huge effect or any restraint of results with one or the other compound. Peaks remained commonly steady and unaffected. The huge drop in top stature happened when the concnetrations of $MgCl_2$ was increased, thus, positive control 1 ng of control 007 was ustilized (Table 21).

Table 21: The drop of sample 007 upon increasing the concentration of $MgCl_2$

Sample	Peak Height avg
+ve 007 1 ng/ul	2,845
007 + 0.75 ul $MgCl_2$	2,532
007 + 1.5 ul $MgCl_2$	2,413
007 + 2.25 ul $MgCl_2$	874
007 + 3.0 ul $MgCl_2$	196

4.2 Combination Study

Combinations of DNA starting from male individual are as often as possible experienced in legal casework tests. Considering the whole, the capacity of any STR multiplex examine to recognize minor and significant amount of DNA is vital. To display a legal combination test and test the presentation of the Qiagen Unit in this specific situation, three control DNAs, Taqman, R2, and 007 DNA were used. These DNAs were mixed in different proportions with reducing relative sums from one (007 DNA). The aggregate sum of consolidated information DNA at every proportion was held steady at 1 ng for all responses. The YSTR peak analysis revealed that the individuals were heterozygous for STRs DYS3801 and DYS 458, which were recognized by the double peaks**.** Rest of the YSTRs namely, DYS301, DYS 38011, DYS533, DYS390, DYS627, DYS393, DYS19, DYS437, DYS449, DYS231, DYS443, DYS518, DYS545 with the single peak denotes that the individuals were homozygous for the specific YSTRs**.** Further, the peaks for few YSTR such as DYS439, DYS575, DYS635, DYS345, and DYS392 were not significantly observed which could be inferred that the STRs were lost due to mutations**.** The longest peaks were observed for few YSTRs such as DYS458, DYS490, and DYS545 revealing longest cluster of repeats**.** An electropherogram with profiles for 007 and Taqman is appeared in Figure 13. In male/female no extra peaks of female sample on the grounds that Y chromosome pack just for male sample. Since this is a Y-explicit unit, thus the assumption is that solitary male profiles will show because of Y chromosome, with no restraint. True to form this is the thing that occurred, and the female examples that were added into the cylinder didn't show up profiles. Henceforth the pack is efficient when performing male/female blended examples and does not bring about any disturbance on DNA amplification.

4.3 Inhibitor Study

Substances known to be inhibitory toward PCR responses are regularly experienced in DNA extracted and separated from legal casework tests impacts noticed comprising of loss of peaks as well as bringing down peaks sizes in profiles produced. These inhibitory chemicals such as haematin from blood, humic corrosive from soil and Tannic corrosive are well known. To demonstrate such situations and test the efficacy of the Qiagen kit in this unique circumstance, an increasing concentration of such inhibitory substances were added straightforwardly to the intensification responses alongside 1 ng of 3 Controls of DNA 007, 9948, Taqman. Execution levels were analyzed dependent on the quantity of alleles accurately brought in profiles got back from these responses. As a result, Qiagen kit returned profiles with a fundamentally higher normal number of alleles calls than both of different units within the sight of 3.0 hematin alleles within the sight of high inhibitory substances.

Off-ladder alleles, also called as microvariants, contain nucleotide changes or incorporations or eradications in the STR repeat region or speedy flanking regions are known to exist and can be settled with a high precision in Capillary Electrophoresis (CE) instrument. Off ladder showed up in electropherogram in some model tests (Figure 14). These ladders are used to produce ‘BINS’. These are locations where the software locates a particular STR amplified product and converts the size of the DNA fragment. The data of this current investigation had many off-ladder in all tests. However, occasionally some sample may give an amplified product which does not fall into any of the known ‘BINS’.

A remarkable feature for the electropherogram was displayed in the marker DYS385. It produces two peaks of different lengths. These two peaks start from

replicated and inversed copies of one Y - chromosomal locus generating two alleles with different peak heights (Table 22).

Table 22: Allele assignment of the investigator Argus Y-28 Kit

Locus	DNA 9948
DYS19	14
DYS385	11/14
DYS389-I	13
DYS389-II	31
DYS390	24
DYS391	10
DYS392	13
DYS393	13
DYS437	15
DYS438	11
DYS439	12
DYS448	19
DYS456	17
DYS458	18
DYS449	30
DYS460	11
DYS481	24
DYS518	38
DYS533	12
DYS549	13
DYS570	18
DYS576	16
DYS627	22
DYS635	23
DYS643	11
YGATAH4	12

Biological artifacts that can complicate the current STR data interpretation include stutter peaks (Figure 15). Peaks that show up primarily one repeat less than the true allele due to strand slippage during DNA synthesis. During the PCR amplification study, the polymerase can lose its place while reproducing a strand of DNA, ordinarily slipping advances or in invert four base matches. The occasion of waver tops depends

upon the gathering of the repetitive development and the amount of alleles. These peaks interpreted using the investigator template files for GeneMapper ID-X Programming is the software for automated genotyping solution for forensic casework and database analysis which grants tops more than 50 stature to be distinguished. The assessment method could be modified to restrict the proportion of stammer tops.

Terminal transferase activity of the Taq DNA Polymerase may cause incomplete adenylation at the 3' end of the amplified DNA fragments. The artifact peak is one base shorter than expected (–1 peaks). All primers included in the Investigator Argus Y-28 QS Kit were designed to minimize these artifacts. The peak height of the artifact correlates with the amount of DNA.

The other artifacts that influence the data of the current investigation was the room temperature, which may influence the performance of PCR products on instruments resulting in shoulder peaks or split peaks. In such cases of shoulder or split peaks, injecting the sample once again is recommended. Futher, the ambient conditions were ensured as recommended by the instrument manufacturer. In addition, the buffers were equilibrated to the ambient conditions.

Chapter 5: Conclusion

The current study was aimed to analyze the internal forensic evaluation of Investigator Argus Y-28 QS using samples such as saliva, blood cells, sperm and epithelial cells from the case studies. The PCR amplified samples followed by capillary electrophoresis revealed remarkable insights on the number of Y-STR repeats and information about the genetic inheritance of every individual. Some of the individuals were homozygous for the particular YSTRs in the given samples. Whereas, most of the individuals were found to be heterozygous for most of the YSTRs (characterized by double peaks). The results further states that there is loss of numerous YSTRs due to mutation over the course of time. The observed findings shows that the Investigator Argus 28 YSTR kit is highly accurate in analyzing the forensic samples with unambiguous insights. Further, the kit can be of greater importance in giving solutions to the parental disputes, parental lineage identifications, criminological case and bio-geographical ancestry analysis through clear results of repeated STR, homozygous and heterozygous nature.

References

Aartsma-Rus, A. (2012). Overview on AON design. Exon Skipping, 5, 117-129.

Arwin, R., Mannis, O., Diego, M., Kees, V., Sharon, W., Robert, L., & Manfred, K. (2019). Forensic Y-SNP analysis beyond SNaPshot: High-resolution Y-chromosomal haplogrouping from low quality and quantity DNA using Ion AmpliSeq and targeted massively parallel sequencing. Forensic Science International: Genetics, 41, 93-106.

Ballantyne, K. N., Ralf, A., Aboukhalid, R., Achakzai, N. M., Anjos, M. J., Ayub, Q., & Berger, B. (2014). Toward male individualization with rapidly mutating Y-chromosomal short tandem repeats. Human Mutation, 35(8), 1021-1032.

Butler, J. (2009). Forensic DNA Typing. Academic Press Amsterdam. Elsevier.

Claerhout, S., Roelens, J., Van der Haegen, M., Verstraete, P., Larmuseau, M. H. D., & Decorte, R. (2020). Ysurnames? The patrilineal Y-chromosome and surname correlation for DNA kinship research. Forensic Science International: Genetics, 44, 102-113.

Colaco, S. (2018) Genetics of the human Y chromosome and its association with male infertility. Reprod Biology Endocrinol, 16, 14-26.

Dorit, R. L., Akashi, D., & Gilbert, W. (1995). Absence of polymorphism at the ZFY locus on the human Y chromosome. Science, 268, 1183-1185.

Ellegren, H. (2000). Heterogeneous mutation processes in human microsatellite DNA sequences. Nature Genetics, 24(4), 400-402.

Fregeau, C. J., & Fourney, R. M. (1993). DNA typing with fluorescently tagged short tandem repeats: A sensitive and accurate approach to human identification. Biotechniques, 15(1), 100-119.

Inman, K., & Rudin, N. (2001). An introduction to forensic DNA analysis. CRC press.

Kayser, M. (2017). Forensic use of Y-chromosome DNA: A general overview. Human Genetics, 136(5), 621-635.

Koreth, J., Oleary, J. F., & Mcgee, J. A. (1996). Microsatellites and PCR genomic analysis. The Journal of Pathology, 178(3), 239-248.

Levene, P. A. (1940). Biographical memoir of Phoebus. Retrieved 3 November 2020 from http://www.nasonline.org/publications/biographical-memoirs/memoir-pdfs/levene-phoebus-a.pdf

Li, M., Zhou, W., Zhang, Y., Huang, L., Wang, X., Wu, J., & Bian, Y. (2020). Development and validation of a novel 29-plex Y-STR typing system for forensic application. Forensic Science International: Genetics, 44, 102-118.

Maddox, B. (2014). Rosalind Franklin: The dark lady of DNA. New York: HarperCollins Press.

Mallet, J. (1995). A species definition for the modern synthesis. Trends in Ecology & Evolution, 10(7), 294-299.

Nadir, E., Margalit, H., Gallily, T., & Ben-Sasson, S. A. (1996). Microsatellite spreading in the human genome: evolutionary mechanisms and structural implications. Proceedings of the National Academy of Sciences, 93(13), 6470-6475.

Qiagen (2020). Investigator® Argus Y-28 QS Handbook. Retrieved 16 January 2020 from https://www.qiagen.com/us/search/products/

Roewer, L. (2019). Y-chromosome short tandem repeats in forensics-Sexing, profiling, and matching male DNA. Wiley Interdisciplinary Reviews: Forensic Science, 1(4), 1336-1349.

Slatkin, M., & Hudson, R. (1991). Pairwise comparisons of mitochondrial DNA sequences in stable and exponentially growing populations. Genetics, 129, 555-562.

Subramanian, S., Mishra, R. K., & Singh, L. (2003). Genome-wide analysis of microsatellite repeats in humans: their abundance and density in specific genomic regions. Genome Biology, 4(2), 1-10.

Tautz, D. (1993). Notes on the definition and nomenclature of tandemly repetitive DNA sequences. DNA Fingerprinting: State of the Science, 56, 21-28.

Underhill, P. A., Jin, L., Lin, A. A., Mehdi, S. Q., Jenkins, T., Vollrath, D., & Oefner, P. J. (1997). Detection of numerous Y chromosome biallelic polymorphisms by denaturing high-performance liquid chromatography. Genome Research, 7(10), 996-1005.

Vuichard, S., Borer, U., Bottinelli, M., Cossu, C., Malik, N., Meier, V., & Castella, V. (2011). Differential DNA extraction of challenging simulated sexual-assault samples: A Swiss collaborative study. Investigative Genetics, 2(1), 11-26.

Whitfield, L. S., Sulston, J. E., & Goodfellow, P. N. (1995). Sequence variation of the human Y chromosome. Nature, 378(55), 379-380.

Zeyad, M. (2017). Study of Y-Chromosome STR Markers in United Arab Emirates Population, Master thesis, UAE University, UAE.