T.C.

CUKUROVA UNIVERSITY

INSTITUTE OF ADDICTION

AND FORENSICSCIENCES

DEPARTMENT OF FORENSIC SCIENCES

FORENSIC ANALYSIS OF THE Y CHROMOSOME SNPS IN THE ELGON REGION POPULATION OF EASTERN UGANDA

AMOSI WAZODI

FORENSIC SCIENCE MASTER'S THESIS

SUPERVISOR

Prof. Dr. M. Bertan YILMAZ

ADANA 2023

ACCEPTANCE AND APPROVAL

I WAZODI AMOSI declare that this research thesis is my original write up and has never been presented for the award of a degree in any University.

Signature……Date:
…….

This thesis has been approved for submission by the academic Supervisor.

Prof. Dr. M. Bertan YILMAZ

Medical Biology Department

Cukurova University, Turkey

Signature……Date……
…….

TABLE OF CONTNENTS

ACCEPTANCE AND APPROVAL i
TABLE OF CONTNENTS ii
LIST OF FIGURES iv
DEDICATION v
ACKNOWLEDGEMENTS vi
ABBREVIATIONS vii
ÖZET viii
SUMMARY ix
CHAPTER I 1
1.0 Introduction 1
1.1 Background of the Study 1
1.2 Problem Statement 3
1.3 General Objective 4
1.3.1 Specific Objectives 4
1.4 Hypothesis 4
1.5 Significance 4
1.6 Justification 5
1.7 Scope 6
CHAPTER II 7
2.0 Literature Review 7
2.1 The Y Chromosome 7
2.2 The Next-generation Sequencing (NGS) Technology 8
2.3 The Elgon Population in Eastern Uganda 9
2.4. The Bagisu tribe 10
2.5. The Sabiny tribe 11
CHAPTER III 13
3.0 Materials and Methods 13
3.1 Research design 13
3.2 Description of the geographical area 13
3.3 Sampling strategy 13
3.4 Sample collection 13
3.5 DNA Lysis 14
3.6 DNA extraction 15

3.7 DNA Quantification 16
3.8 Library Preparation 16
3.9 Template preparation 18
3.10 Sequencing 18
3.11. Data analysis 19
CHAPTER IV 20
4.0 Results and Discussion 20
4.1 Results 20
4.2 Discussion of the results 24
4.3 Limitations and Challenges of the Study 28
4.4 Conclusion and Recommendations. 29
5. REFERENCES 31
RESUME 35
APPENDIX 36
Appendix I. The Ethical Clearance Letter 36
Appendix II. Table 1 Showing the Y Chromosome SNPs Haplogroups diversity of the Bagisu and the Sabiny of Elgon region in Eastern Uganda. 37
Appendix III. Coverage Analysis report of the samples. 40

LIST OF FIGURES

Figure 1 The map showing the location of the Elgon region in Eastern Uganda. 10
Figure 2 Nucleic cards used for sample collection. 14
Figure 3 Filter column assemblage and sample lysis on Thermo-shakers.....14
Figure 4 Hamilton Bio-robots used for Extraction. 15
Figure 5 Quant Studio5 Instrument used for the DNA quantification of the samples.....16
Figure 6 The Ion Chef Instrument used for Library and Template preparations.....17
Figure 7 The Gene Studio and the 530 Chips used for Sequencing oif the samples. 18
Figure 8 The Y Chromosome SNPs Haplogroups of the Bagisu and the Sabiny of the Elgon region in Eastern Uganda.....20
Figure 9 The Haplogroups Under A Major Haplogroup (A1b And A-M13). 21
Figure 10 The haplogroups under B major haplogroup (B-M5844, B-Y275213 and B-Y12200) 21
Figure 11 The haplogroups under E major haplogroup (E-M41, E-Z830, E-V22, E-U175, E-L514, E-M4254, E-U174 and E-U290). 23
Figure 12 The frequency distribution of the Y Chromosome SNPs Haplogroups of the Bagisu and the Sabiny of the Elgon Population in Eastern Uganda.....24

DEDICATION

I dedicate this thesis to my family and friends who have been a constant source of encouragement, support, and love throughout my academic journey. Your unwavering belief in my potential and your sacrifices made this achievement possible.

I also dedicate this thesis to the resilient and welcoming people of the Elgon region population in Eastern Uganda, whose DNA samples and participation in this research made it possible. Your contribution has shed light on the genetic makeup of the population and deepened our understanding of the complex history and diversity of human populations.

Finally, I dedicate this thesis to all the scholars and researchers who have dedicated their lives to advancing the field of forensic genetics. Your work has paved the way for future generations to continue exploring, discovering, and using genetic data for the betterment of humanity.

ACKNOWLEDGEMENTS

I would like to express my gratitude and appreciation to the following institutions and individuals who made this thesis a success:

Makerere University and Cukurova University for the Financial and Academic support they have provided me and for the opportunity to attain this knowledge. May God bless Prof. William Bazeyo, Prof. Ayse Serin and Prof.Winston T. Ireeta for being there for me.

My supervisor, Prof. Dr. M. Bertan YILMAZ. May God bless you for the guidance, support, and motivation you have provided me throughout the writing of this thesis.

The Directorate of Forensic Services, Uganda Police Force for providing me with the necessary resources, including access to the laboratory facilities and equipment. Special thanks to the Director Forensic Services ACP. Andrew K. Mubiru for providing me with the academic support I needed to complete this research. Your expertise and knowledge in the field of forensic genetics were invaluable.

The Elgon region population in Eastern Uganda for their willingness to participate in this study, and for sharing their time, DNA samples, and personal information.

My family and friends for their unwavering support, encouragement, and understanding during the ups and downs of the research journey. I thank our team leader Mutesi Lilian Doris for being such a good coordinator and you have all it takes to lead.

Lastly, I would like to acknowledge contribution of all the scholars whose work I have reviewed, cited, and referenced in this thesis. Your research has shaped my understanding of the subject matter and inspired me to think critically and creatively.

Thank you all for your support and for making this thesis a reality.

ABBREVIATIONS

DNA:	Deoxyribose Nucleic Acid
ISOGG:	International Society of Genetic Genealogy
MPS:	Massively Parallel Sequencing
MSY:	Male Specific region of the Y chromosome
MtDNA:	Mitochondria Deoxyribose Nucleic Acid
NGS:	Next Generation Sequencing
PCR:	Polymerase Chain Reaction
SBE:	Single Base Extension
SNPs:	Single Nucleotide Polymorphisms
STRs:	Short Tandem Repeats.
YCC:	Y Chromosome Consortium

ÖZET

Y kromozomu SNP'lerinin adli analizi, çeşitli popülasyonların genetik çeşitliliğini anlamada yararlı bir araç haline gelmiştir. Y kromozomu SNP'lerinin tanımlanması, cinsel saldırı, babalık testi veya kayıp şahıs soruşturmaları gibi erkek DNA profilinin gerekli olduğu vakaların adli soruşturmalarında yararlıdır. Bu çalışmada, Doğu Uganda'daki Elgon popülasyonundan aralarında akrabalık ilişkisi bulunmayan ve rastgele seçilen 32 erkekten alınan yanak içi sürüntü örneklerinde Y kromozomu SNP'lerinin adli analizi için Yeni Nesil Dizileme Teknolojisi (NGS) ve Ion AmpliSeq HID Y-SNP Paneli v1 primer havuzu kullanıldı. Haplogruplar için geliştirilen Yleaf bilgisayar analiz programı kullanarak elde ettiğimiz sonuçlar, Elgon popülasyonunun birçok Afrika popülasyonunun özelliği olan %40,6 gibi yüksek bir genetik çeşitlilik sergilediğini gösterdi. Bu çalışmada, adli analiz ve popülasyon genetiği çalışmaları için faydalı genetik belirteçler olarak hizmet edebilecek on üç (13) benzersiz Y kromozomu SNP'i (Haplogrup) tanımlanmıştır. Haplogruplar, Afrika populasyonuna özgü başlıca üç haplogrup A, B ve E'den kaynaklanır ve Haplogrup E %63,63 ile en yüksek frekansı gösterdi, ardından %21,21 sıklıkla haplogrup A ve %15,15 sıklıkla haplogrup B izledi. Haplogruplar A1b, A-M13, B-M5844, B-Y12200, B-Y275213, E-M41, E-Z830, E-V22, E-U175, E-M4254, E-U174, E-L514 ve E-U290 idi. Haplogrup E-U174 en yüksek sıklıkta %24,24 ile yalnızca Bağisu'da; B-Y275213, E-Z830 ve E-U290 haplogrupları yalnızca Sabiny'de gözlendi. Diğer haplogruplar her iki grupta da paylaşıldı. Genel olarak, bu araştırma, Elgon popülasyonunun genetik yapısına ilişkin değerli bilgiler sağlar ve Y kromozomu SNP'lerinin adli analiz ve popülasyon genetiği araştırmalarındaki öneminin altını çizer. Adli bilimciler, Y kromozomundaki SNP'leri analiz ederek, bir DNA örneğine katkıda bulunan erkekleri belirlemek için kullanılabilecek bir genetik profil geliştirebilirler. Bununla birlikte, ülkenin genetik çeşitliliğini tam olarak anlamaya yardımcı olacağından, Uganda'daki diğer popülasyonların genetik çeşitliliğini araştırmak için büyük bir örneklem büyüklüğüne sahip Y kromozomu SNP'leri kullanılarak daha fazla araştırma yapılmalıdır.

Anahtar kelimeler: YSNP, haplogrup, Uganda, adli bilimler

SUMMARY

The forensic analysis of Y chromosome SNPs has become a useful tool in understanding the genetic diversity of various populations. The identification of unique Y chromosome SNPs is useful in Forensic investigations of cases where a male DNA profile is needed, such as in cases of sexual assault, paternity testing, or missing persons investigations. In this research, a forensic analysis of the Y chromosome SNPs from 32 unrelated male individuals of the Elgon population in Eastern Uganda using the Next Generation Sequencing Technology (NGS) and the Ion AmpliSeq HID Y-SNP Research Panel v1 primer pool. Using the Y-leaf analysis tool, our results showed that the Elgon population exhibited a high degree of genetic diversity of 40.6% which is characteristic of many African populations. Thirteen (13) unique Y chromosome SNPs (Haplogroups) that could serve as useful genetic markers for forensic analysis and population genetics studies were identified in this study. The haplogroups originate from the major three Haplogroups A, B and E typical of the African Population and Haplogroup E showed the highest frequency of 63.63%) followed by haplogroup A with the frequency of 21.21% and haplogroup B with a frequency of 15.15%. The haplogroups were A1b, A-M13, B-M5844, B-Y12200, B-Y275213, E-M41, E-Z830, E-V22, E-U175, E-M4254, E-U174, E-L514 and E-U290. Haplogroup E-U174 was observed in highest frequency 24.24% in Bagisu only and haplogroups B-Y275213, E-Z830 and E-U290 were observed in Sabiny only. Other haplogroups were shared in both. Overall, this research provides valuable insights into the genetic makeup of the Elgon population and underscores the importance of Y chromosome SNPs in forensic analysis and population genetics research. By analyzing SNPs on the Y chromosome, forensic scientists can develop a genetic profile that can be used to identify the male contributors to a DNA sample. However, further studies should be conducted to investigate the genetic diversity of other populations in Uganda using Y chromosome SNPs with a large sample size as this will help to fully understand the genetic diversity of the country.

Keyword: YSNP, haplogrup, Uganda, Forensic Sciences

CHAPTER I

1.0 Introduction

1.1 Background of the Study

The analysis of single nucleotide polymorphisms (SNPs) on the Y chromosome has become an important tool in forensic genetics due to its ability to provide information about a person's paternal lineage. In recent years, studies have been conducted to analyze the Y chromosome SNPs in various populations around the world, providing valuable data for forensic investigations (Jobling & Tyler-Smith, 2017; Kayser, 2017).

Eastern Africa is a region with a large amount of ethnic and linguistic diversity. Populations in this region speak languages belonging to all four of the major language families in Africa: Afroasiatic, Nilo-Saharan, Niger-Kordofanian and Khoisan (Lema *et al.*, 2009).

To better understand and trace the geographic origin of ancestral populations, human genetic studies have been utilized basing on the availability of various genetic markers that can be analysed (Salas *et al*., 2002). The Y-chromosome marker has been used because its highly suitable to provide information about the geographic region a person’s paternal ancestors originated from (Kayser, 2017). The sequential accumulation of genetic diversity along lines of common descent makes Y-chromosome a much more powerful tool for phylogeographic analysis than autosomes (Hassan, *et al.*, 2008).

Uganda is one of the East African countries bordered by Kenya in the East, Sudan in the North, DRC in the west, Tanzania in the south and Rwanda in the southwest. İt has a complex and diverse language situation. It’s a multi-ethnic, multi-cultural and multi-lingual society (Namyalo & Nakayiza, 2015).

Elgon region is an Economic region deriving its name from the Mt. Elgon on whose slopes the region is situated and is comprised of two sub-regions namely the Bugisu sub-region and the Sebei sub-region. The Bugisu sub-region is inhabited by mainly the Gisu people also known as Bamasaba and majorly occupy the districts of Mbale, Sironko,

Manafwa, Bududa, Namisindwa and Bulambuli and the Sebei sub-region is inhabited mainly by the Sabiny who majorly occupy the districts of Kapchorwa, Kween and Bukwo. The region is in Eastern Uganda with Mbale the principal town located about 245 km from Kampala. The region borders Kenya which can be accessed through Suam border post in Bukwo district and Lwakhaka border post in Manafwa. It neighbors the Karamoja, Teso and Bukedi regions internally.

Elgon region has a unique genetic history and cultural heritage. The Elgon population is believed to have migrated from the highlands of Ethiopia over 5,000 years ago and has since undergone genetic admixture with neighboring tribes. Previous studies have shown that the Elgon population has a high frequency of the E1b1a haplogroup, which is common in sub-Saharan Africa (Ugwumba et al., 2018).

Some of the population genetic data from Ugandan groups are 2 complete mtDNA sequences from Ugandans of unknown ethnicity (Cruciani *et al., 2002*), complete mtDNA genome of 13 Baganda, 13 Acholi and 14 Lugbara, Ugandan individuals (Isabirye, 2010), several mtDNA control region sequences (Cruciani *et al., 2002*), and Y chromosome microsatellite profiles from 188 Karamojong (Gomes *et al.,* 2010). It is worth noting that the Ugandan mtDNA sequences in (Horai, *et al.,* 1995; Salas *et al.,* 2002) were selected because of their deep divergence from other known sequences.

This study aims to analyze the Y chromosome SNPs of the Bagisu and the Sabiny in the Elgon population to provide insights into their genetic ancestry and paternal lineage. This could have significant implications for forensic investigations in the region, as well as contributing to our understanding of human migration and population history. Through this analysis, we hope to shed more light on the genetic diversity and unique characteristics of the Elgon population. Studying Y chromosome single nucleotide polymorphisms (SNPs) can be highly relevant to forensics. SNPs are variations in the genetic code that are used to identify differences between individuals. By analyzing SNPs on the Y chromosome, forensic scientists can develop a genetic profile that can be used to identify the male contributors to a DNA sample.

Since the Y chromosome is only present in males and is passed down from fathers to their sons, studying Y chromosome SNPs can be particularly useful in cases where a male DNA profile is needed, such as in cases of sexual assault, paternity testing, or

missing persons investigations. In general, studying Y chromosome SNPs is a powerful tool in forensic investigations that can be used to identify male individuals, establish patrilineal relationships, and potentially provide geographic information about the individual.

1.2 Problem Statement

The Elgon region is a mountainous area and is highly susceptible to massive landslides, especially during heavy rains. The previous studies have analyzed the genetics of some of the Ugandan populations, there is a lack of research on the Y chromosome SNPs of this unique population of the Bagisu and the Sabiny of the Elgon region in Eastern Uganda. The Y chromosome is important in forensic investigations due to its ability to provide information about a person's paternal lineage (Kayser, 2017). Strict paternal inheritance of these markers makes them useful for paternity and kinship studies as they provide a lineage marker in form of a single haplotype transferred directly from father to son (Köksal *et al*., 2022). Studying the Y chromosome of the Bagisu and the Sabiny of the Elgon region could provide insights into their genetic diversity, thus filling a gap in the current genetic knowledge of this population.

Single nucleotide polymorphisms (SNPs) have become popular in forensic genetics as an alternative to short tandem repeats (STRs) due to low mutation rates and small amplicon sizes and the ability to handle degraded DNA (Joo, 2022). High resolution analysis of Y chromosome haplogroups using large Y-SNP panels allow improved paternal lineage identification and biogeographic ancestry inference (Ralf *et al*., 2019).

Whereas there were previous genetic studies on the Ugandan population such as 2 complete mtDNA sequences from Ugandans of unknown ethnicity (Cruciani *et al., 2002*), complete mtDNA genome of 13 Baganda, 13 Acholi and 14 Lugbara, Ugandan individuals (Isabirye, 2010) and Y chromosome microsatellite profiles from 188 Karamojong (Gomes *et al.,* 2010), less is known about the Elgon population Y haplogroups.

This study aims to address this research problem by analyzing the Y chromosome SNPs in the Bagisu and the Sabiny of the Elgon population of Eastern Uganda. By doing

so, we hope to provide valuable genetic data that can enable forensic scientists develop a genetic profile that can be used to identify the male contributors to a DNA sample.

1.3 General Objective

To establish the Y-SNPs haplogroups diversity of the Bagisu and the Sabiny of Eastern Uganda.

1.3.1 Specific Objectives

1.2.1.1 To analyze the Y Chromosome SNPs of the Bagisu and the Sabiny of the Elgon region in Eastern Uganda.

1.2.1.2 To determine the frequency and distribution Y Chromosome haplogroups of the Bagisu and the Sabiny of the Elgon region in Eastern Uganda.

1.2.1.3 To establish the relationship between Y Chromosome haplogroups of the Bagisu and the Sabiny of the Elgon region in Eastern Uganda and the rest of the world.

By achieving these research objectives, we hope to significantly contribute to the field of forensic genetics while also providing insights into the genetic history of the Elgon population.

1.4 Hypothesis

Null Hypothesis: The distribution of Y SNPs haplogroups among the Bagisu and the Sabiny is correlated with geography and language.

Alternate Hypothesis: The distribution of Y SNPs haplogroups among the Bagisu and the Sabiny is not correlated with geography and language.

1.5 Significance

This study will establish the Y Chromosome haplogroups diversity and distribution in the Bagisu and the Sabiny of the Elgon region population. The analysis of Y chromosome SNPs is essential in forensic investigations, and this study will provide

valuable genetic data for forensic investigations in the region. By identifying the frequency and distribution of Y chromosome haplogroups in the Bagisu and the Sabiny of the Elgon population, this research could aid in identifying suspects and solving cases in the region (Kayser, 2017). Generally, this research is significant due to its potential to provide valuable genetic data for forensic investigations while also contributing to our understanding of human migration and ancestry (Jobling & Tyler-Smith, 2017). Since the Y chromosome is only present in males and is passed down from fathers to their sons, studying Y chromosome SNPs can be particularly useful in cases where a male DNA profile is needed, such as in cases of sexual assault, paternity testing, or missing persons investigations. In general, studying Y chromosome SNPs is a powerful tool in forensic investigations that can be used to identify male individuals, establish patrilineal relationships, and potentially provide geographic information about the individual.

1.6 Justification

Y haplogroups, which are defined by Y-SNPs, harbor crucial information for evolutionary, population and forensic genetic studies. Strict paternal inheritance of these markers makes them useful for paternity and kinship studies as they provide a lineage marker in form of a single haplotype transferred directly from father to son (Köksal *et al*., 2022). The analysis of Y chromosome SNPs has become an important tool in forensic genetics due to its ability to provide information about a person's paternal lineage. This information can be crucial in forensic investigations, especially in cases where the identity of the perpetrator is unknown. Moreover, analyzing the Y chromosome SNPs of a population can provide valuable insights into their genetic history, ancestry, and migration patterns (Kayser, 2017).

As the Elgon population of Eastern Uganda has a unique genetic history and cultural heritage, analyzing their Y chromosome SNPs could have significant implications for forensic investigations in the region, where little data on the Y chromosome SNPs of the population currently exists. Additionally, this research could significantly contribute to our understanding of human migration and population history. Generally, this research is important due to its potential to shed more light on the genetic

diversity and unique characteristics of the Elgon population while also providing valuable genetic data for forensic investigations.

1.7 Scope

The study will focus on analyzing the Y chromosome single nucleotide polymorphisms (SNPs) to determine the genetic relationships between individuals and groups within the population. The research will involve the collection of DNA samples from 32 unrelated male participants within the Bagisu and the Sabiny of the 9 selected districts namely Bukwo, Kwen, Kapchorwa, Bulambuli, Sironko, Mbale, Manafa, Namisindwa and Bududa of Elgon region in Eastern Uganda, followed by genotyping and sequencing of Y chromosome SNPs. The findings of this research could be useful in forensic investigations, such as in the identification of missing persons or victims of crime.

CHAPTER II

2.0 Literature Review

2.1 The Y Chromosome

Y chromosome is the sex determining chromosome in humans as its presence indicates biological male and its absence indicates biological female. The Y chromosome is passed only from the father to the son and contains sex-determining region which triggers male development. The short arm of the Y chromosome is very important as most of the vital genes for male sex determination, spermatogenesis, and other male-related functions are found on it (Imam & Rana, 2018).

The male specific region of the Y chromosome (MSY) makes up 95% of its length. Unlike autosome chromosomes, the MSY does not undergo recombination with a partner chromosome during meiosis. The Y chromosome is particularly informative for reconstructing human evolution due to uni-parental inheritance via males and the absence of recombination (Lema *et al.,* 2009; Maan *et al.,* 20017). The Y chromosome contains the largest nonrecombining block in the human genome and can be considered one of the most informative haplotyping systems, with applications in evolutionary population studies, forensics, medical genetics, and genealogical reconstruction (Underhill & Kivisild, 2007). Using a combination of genetic markers, it is possible to classify each individual Y chromosome into one of numerous haplogroups (Maan *et al.,* 2017). A standard nomenclature was established by the Y Chromosome Consortium (YCC), 2002, which resolved the global pattern of Y chromosome variation into 18 major haplogroups that were classified using capital letters A through to R. This was revised by (Karafet *et al.,* 2008) to a Y chromosome haplogroup phylogeny that contained 311 branches delineated by approximately 600 markers (primarily bi-allelic) and included an additional two haplogroups (S and T), increasing the major haplogroup number to 20. Currently, the Y chromosome phylogeny is kept up to date by the International Society of Genetic Genealogy (ISOGG), as new publications and the discovery of more variation improve its resolution.

Of over 20 major haplogroups, only 3 exhibit substantial presence throughout Africa namely A, B and E. While haplogroups A and B represent the oldest splits in the Y chromosome phylogeny, it is haplogroup E that has attained the highest frequencies, with some clades spreading out of Africa into the middle east and Europe (Naidoo, 2014). As of 2021, there are over 2,000 known Y chromosome haplogroups identified by the scientific community. These haplogroups are defined by specific variations, or mutations, on the Y chromosome and are used to trace patrilineal ancestry and genetic relationships among human populations.

The Y chromosome has a high mutation rate, making it a useful tool for detecting recent population movements and establishing relationships between individuals or populations unlike mtDNA has a relatively low mutation rate means that it may not provide enough resolution to distinguish between closely related individuals or populations. This can make it difficult to differentiate between certain populations or identify finer-scale patterns of migration and genetic variation (Soares *et al.,* 2009; Jobling & Tyler-Smith, 2013).

2.2 The Next-generation Sequencing (NGS) Technology

With the introduction of single nucleotide polymorphism (SNP) analysis on NGS platforms, HID laboratories can now routinely ascertain biogeographic ancestry, paternal lineage and phenotypic traits using informative genetic marker sets to assist investigators with new types of forensic investigative leads. In addition to the technical advantages of MPS over single base extension (SBE) methods which limit the number of markers that can be included in a reaction targeted Y-SNP MPS assays offer a deeper understanding of paternal lineages to associate potentially very distant relatives in a single assay (Ralf *et al*., 2019). By contrast, Y-STR methods, typically used to make paternal associations between close relatives, do not have the same resolution of large Y-SNP assays. Single nucleotide polymorphisms (SNPs) have been studied as a forensic marker that overcomes the limitations of short tandem repeats (STRs). SNPs are useful for highly degraded samples as they can be amplified to a smaller size than that of STRs. In addition, due to their stable inheritance and lower mutation rate (10^{-8}) than STRs (10^{-3}), SNPs are suitable for paternity testing and complex relationship identification (Butler *et al.,* 2007).

However, due to their bi-allelic nature, many SNPs should be multiplexed at once to achieve discrimination power equivalent to that of the STR typing. Next-generation sequencing (NGS) technology has enabled the simultaneous sequencing of large-scale markers from multiple samples, using a sample specific barcode system (Ballard *et al*., 2020).

In general, NGS technology is a powerful tool for the forensic analysis of Y chromosome SNPs in the Elgon population of Eastern Uganda. It enables the identification of novel SNPs and rare haplotypes and provides accurate and reliable results. Its high-throughput capabilities and precision make it an essential tool in the forensic toolkit.

2.3 The Elgon Population in Eastern Uganda

Uganda is one of the East African countries bordered by Kenya in the East, Sudan in the North, Democratic Republic of Congo in the west, Tanzania in the south and Rwanda in the southwest. İt has a complex and diverse language situation. It's a multi-ethnic, multi-cultural and multi-lingual society (Namyalo & Nakayiza, 2015).

The Elgon region is an Economic region deriving its name from the Mt. Elgon on whose slopes the region is situated and is comprised of two sub-regions namely the Bugisu sub-region and the Sebei sub-region (Fig.1). The Bugisu sub-region is inhabited by mainly the Gisu people also known as Bamasaba and majorly occupy the districts of Mbale, Sironko, Manafwa, Bududa, Namisindwa (Curved from Manafa District) and Bulambuli and the Sebei sub-region is inhabited mainly by the Sabiny who majorly occupy the districts of Kapchorwa, Kween and Bukwo. The region is in Eastern Uganda with Mbale the principal town located about 245 km from Kampala. The region borders Kenya which can be accessed through Suam border post in Bukwo district and Lwakhaka border post in Manafwa. It neighbors the Karamoja, Teso and Bukedi regions internally (Uganda Investment Authority, 2018).

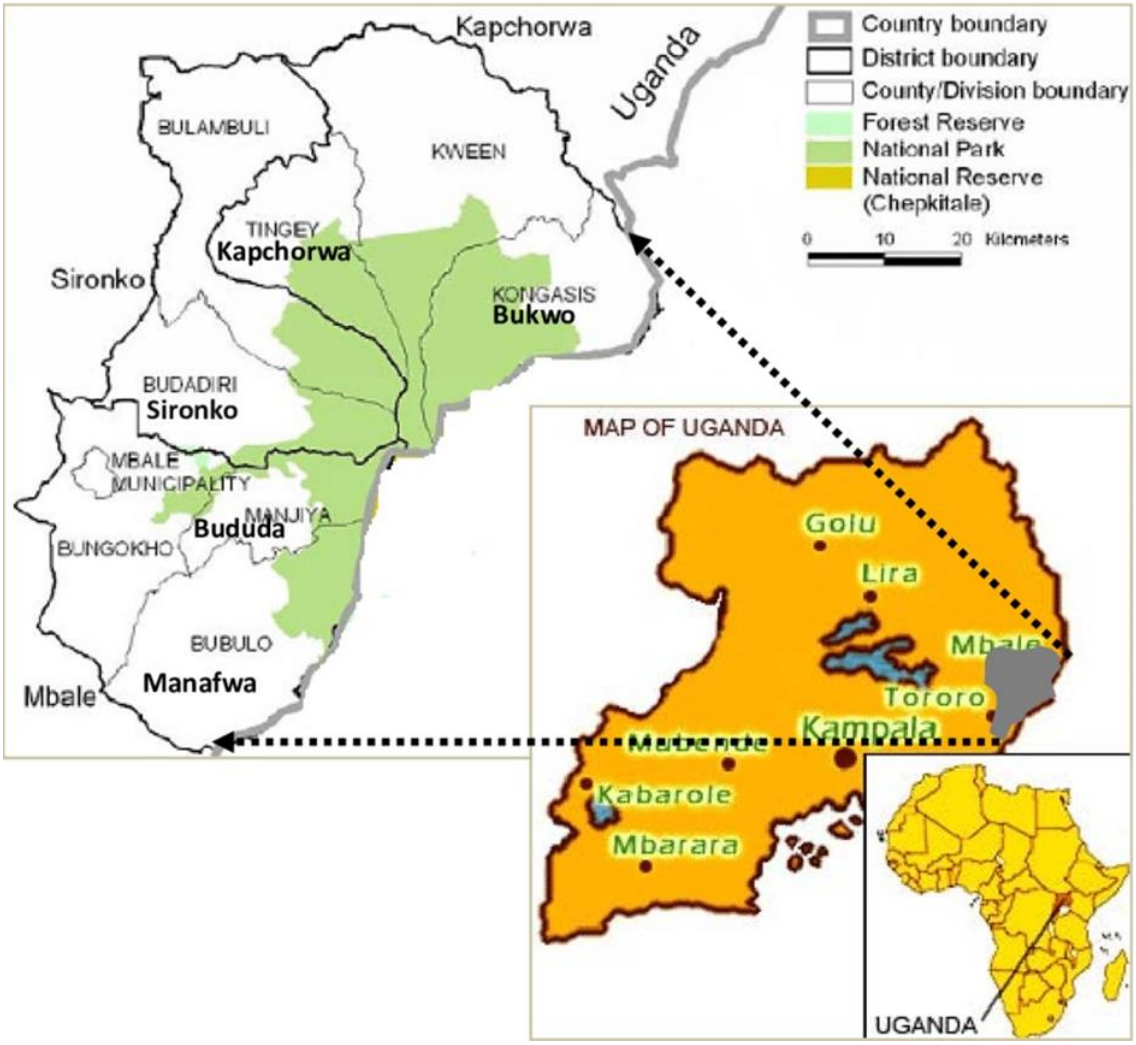


Figure 1 The map showing the location of the Elgon region in Eastern Uganda.

2.4. The Bagisu tribe

The Bagisu tribe is an ethnic group found in eastern Uganda, particularly in the districts of Mbale, Sironko, Manafwa, Bududa, Namisindwa and Bulambuli. The Bagisu are closely related to other Bantu-speaking people in the region and are believed to have migrated into the area several centuries ago.

Research on the origin of the Bagisu is limited, but some studies have suggested that the Bagisu and other Bantu-speaking peoples in this region can trace their ancestry back to the eastern African Great Lakes region. One study analyzed genome-wide data from several ethnic groups in Uganda, including the Bagisu, and found evidence of

genetic admixture between Bantu-speaking groups and other nearby populations such as the Nilotic-speaking Karamojong. The study also identified several genetic variants that were particularly common among the Bagisu and other Bantu-speaking groups, suggesting a shared genetic history (Henn *et al*., 2012).

Cultural traditions among the Bagisu include male circumcision, which is performed during elaborate initiation ceremonies known as "imbalu". These ceremonies are believed to have been practiced among the Bagisu and other ethnic groups in the region for centuries and are thought to be associated with fertility and the transition to adulthood.

Generally, the Bagisu tribe represents an important part of the cultural and genetic diversity of eastern Africa, with a history that is likely closely intertwined with that of other Bantu-speaking populations in the region.

2.5. The Sabiny tribe

The Sabiny tribe, also known as the Sebei, is an ethnic group found primarily in the eastern regions of Uganda near Mount Elgon occupying mainly the districts of Kapchorwa, Kween and Bukwo. The Sabiny speak a language called Sebei, which is part of the Nilotic language family.

The origins of the Sabiny are uncertain, but they are believed to be related to other Nilotic-speaking peoples in eastern Africa. Some anthropologists have suggested that the Sabiny may have migrated into their current territory from the south, possibly from the Lake Victoria area, while others propose that they may be descendants of the Twa people of the Great Lakes region. Cultural traditions among the Sabiny include male circumcision, which is performed during elaborate initiation ceremonies called "soratia". These ceremonies are an important rite of passage for Sabiny boys and are believed to be associated with fertility and the transition to adulthood (Mukhala, 2015; Mbaka, 2017).

The previous population genetic studies conducted on the Ugandan population are 2 complete mtDNA sequences from Ugandans of unknown ethnicity (Cruciani *et al., 2002*), complete mtDNA genome of 13 Baganda, 13 Acholi and 14 Lugbara, Ugandan individuals (Isabirye, 2010), several mtDNA control region sequences (Cruciani *et al.*,

2002), and Y chromosome microsatellite profiles from 188 Karamojong (Gomes *et al.,* 2010). It is worth noting that the Ugandan mtDNA sequences in (Horai, *et al.,* 1995; Salas *et al.,* 2002) were selected because of their deep divergence from other known sequences. Less is known about the Ugandan population Y Haplotypes and this study intends to disclose the Elgon population Y Haplotypes.

The Y chromosome SNPs in the Bantu-speaking populations of Uganda and Tanzania were revealed to have high levels of genetic diversity and within these populations unique Y chromosome SNPs were found in both Ugandan and Tanzanian populations (Quiros *et al*., 2017). A study focused on the genetic structure of African populations by analyzing the Y chromosome SNPs of 2,432 individuals from 113 African populations, including Ugandans showed that the Ugandan population had high levels of genetic diversity, which is characteristic of African populations (Henn *et al*., 2008). These studies also identified unique Y chromosome SNPs in the Ugandan population, which highlights the importance of conducting further research on the genetic diversity of African populations.

In a more recent study, on the Y chromosome SNPs of the Karamojong population of Northeastern Uganda. It was revealed that the Karamojong population had high levels of genetic diversity, which was like other African populations. The study also identified several unique Y chromosome SNPs in the Karamojong populations that could be used as genetic markers for forensic analysis and population genetics studies (Pinhasi *et al.,* 2020).

Therefore, this study on Forensic Analysis of the Y Chromosome SNPs in the Bagisu and the Sabiny of the Elgon Population in Eastern Uganda builds on the previous research on Y chromosome SNPs in Ugandan populations and adds to the growing body of literature on the genetic diversity of African populations. This enables the forensic investigators to easily associate the unknown male DNA.

CHAPTER III

3.0 Materials and Methods

3.1 Research design

This was a descriptive study involving quantitative approaches. It involved determination of the Y-SNPs haplogroup distribution of the Elgon region population in Eastern Uganda through analysis of Y-SNPs haplogroups of the individuals belonging to the nine selected districts namely Bukwo, Kwen, Kapchorwa, Bulambuli, Sironko, Mbale, Manafa and Bududa in the Elgon region.

3.2 Description of the geographical area

This study involved unrelated male individuals belonging to the eight selected districts namely Bukwo, Kwen, Kapchorwa, Bulambuli, Sironko, Mbale, Manafa and Bududa in the Elgon region.

3.3 Sampling strategy

A total of thirty-two (32) unrelated adult male individuals from the districts Bukwo, Kwen, Kapchorwa, Bulambuli, Sironko, Mbale, Manafa and Bududa in Eastern Uganda were utilised in this study after obtaining informed consent. Only health adult males were used in this study as they carry their father's paternal genetic lineage (in their Y chromosome) and not more than one member of any immediate family (siblings, parents, grandparents, first cousins) was recruited into the study.

3.4 Sample collection

Buccal swabs were collected from the participants who have consented using the Nucleic cards (Fig.2). The participants were swabbed in the mouth and the swab transferred and pressed gently on the card for about 15 sec to transfer the sample onto the card. The card is then air dried, packed in a paper envelope, labeled, and then taken to the lab for processing.



Figure 2 Nucleic cards used for sample collection.

3.5 DNA Lysis

Lysis was done following the Prepfiler Automated Forensic DNA Extraction kit and user guide. A Prepfiler lysis buffer and 1.0 M DTT mixture was prepared following the user guide at the ratio of 300 µL Prepfiler lysis buffer to 3 µL 1.0 M DTT for each sample in a 50mL tube. A piece from each sample card of approximately 25mm^2 was cut into separate filter column fitted into a 1.5ml Eppendorf tube (Fig. 3). Using a pippete, 300 µL of the Lysis solution prepared above was added into each assemblage containing the sample and sealed. Samples together with the Lysis solution were then loaded on the thermo-shakers and incubated at the temperature of 70^{0}C and 250 rpm for 60 minutes (Fig.3). After the incubation, the tubes were centrifuged and the substrate in the filter columns were discarded. The sample lysate in the Eppendorf tube was capped properly and proceeded to Extraction.



Figure 3 Filter column assemblage and sample lysis on Thermo-shakers.

3.6 DNA extraction

Genomic DNA was extracted from the sample lysate above using the PrepFiler Automate DNA extraction kit Using Hamilton Biorobot system (Fig.4) following the manufacturer's recommendations. The worklist for the samples was prepared on the Computer of the Biorobot and the corresponding samples loaded respectively according to the worklist. Tips for pipetting, corresponding elusion tubes (1.5mL Eppendorf tubes), magnetic particles, Isopropanol, wash buffer 1, wash buffer 2, wash buffer 3 and elusion buffer including the processing plates were all loaded onto the Biorobot following the manufacture's recommendations. After all the consumables were loaded for extraction, the Bio-robot system was set to start the Extraction process. After the extraction process, the eluted samples (100 μL) in the elusion tubes were sealed and proceeded to DNA quantification.



Figure 4 Hamilton Bio-robots used for Extraction.

3.7 DNA Quantification

DNA concentrations were quantified using the quantstudio5 (Fig.5) and Quantifiler Duo DNA quantification kit and user guide (protocol). 8 standards of the concentrations (ng/ μL) 50.000, 16.700, 5.560, 1.850, 0.620, 0.210, 0.068 and 0.023 were prepared following the guidelines outlined in the protocol. The PCR mixture of Quantifiler Duo Primer Mix and Quantifiler Duo PCR Reaction Mix was prepared following the manufacture's recommendation at the ratio of 10.5 μL Quantifiler Duo Primer Mix and 12.5 μL Quantifiler Duo PCR Reaction Mix for each sample in a 1.5mL Eppendorf tube and vortexed briefly. Using a pippet 23 μL of the PCR mixture was transferred to 41 wells of the 96 wells optical PCR plate. Thereafter, using a fresh tip for each sample, 2 μL from each sample and the standards were added into separate wells containing a PCR mixture. The PCR plate was then sealed with an optical film and then loaded onto the Quantstudio5 instrument. The sample sets were labeled according to the arrangement on the 96 optical well PCR plate. After the run, the samples were analyzed, and the quantities of the total human and total male DNA were established for each sample.



Figure 5 Quant Studio5 Instrument used for the DNA quantification of the samples.

3.8 Library Preparation

All the DNA samples were diluted to a concentration of (0.067ng/μL) using Low-TE buffer following the optimum sample concentration recommendation from the custom SNP panel protocol outlined in the Precision ID SNP Panels with the HID Ion S5/HID Ion

GeneStudio S5 System Application Guide (MAN0017767). Library preparation was done using the Ion chef (Fig.6) by adding 150µLof the Ion AmpliSeq HID Y-SNP Research Panel v1 to Positions A and B tubes of the DL8 reagents cartridge following the custom SNP panel protocol outlined in the Precision ID SNP Panels with the HID Ion S5/HID Ion GeneStudio S5 System Application Guide (MAN0017767). A DNA input amount of 1 ng in 15µL was used for library preparation sample set-up (0.067ng/µL). Cycling parameters for library preparation on the Ion Chef were 20 cycles and 4 minutes anneal and extension time. For the first library, 8 samples were pippeted each into a separate well of 8 wells of column 1 (A1 to H1) of the 96 wells Ion coded plate ranging from Ion code 001 to Ion code 008. The Ion chef was loaded with all the supplies following the loading instructions on the instrument. After the run, the tube D with the combined library barcoded samples was capped and kept under 4^0C. the Ion Chef is then Cleaned for the second library.

The same procedures were repeated for library 2 using the second Ion coded plate ranging from Ion code 009 to Ion code 016, Library 3 using the third Ion coded plate ranging from Ion code 017 to Ion code 024 and Library 4 using the fourth Ion coded plate ranging from Ion code 025 to Ion code 032. After library preparation, the libraries were prepared for template preparation.



Figure 6 The Ion Chef Instrument used for Library and Template preparations.

3.9 Template preparation

Templating was done using the Ion Chef (Fig.6) and plans were created following the protocol outlined in the 'Create a Planned Run for the Custom Ion AmpliSeq™ SNP Panel' section of the Precision ID SNP Panels with the HID Ion S5/HID Ion GeneStudio S5 System Application Guide (MAN0017767) using a 530 chip. Corresponding target y-tree 20191001 designed region BED file and the reference hg19 (DNA) was used to setup templating runs.

Libraries 1 and 2 (16 samples) were pooled together by pippeting 25 µL from each library and combining into a 1.5mL Eppendorf tube, vortexed briefly to mix and then transferred 25 µLof the pooled libraries into Position A of the reagents cartridge. The same was also done with Libraries 3 and 4 and loaded into position B for template preparation. After the chips were loaded, sequencing was done immediately beginning with Chip 1 (Library 1 and 2). Chip 2 was stored under 4^0C.

3.10 Sequencing

Sequencing was done using Ion Gene Studio S5 System (Fig.7) following the procedures in the Precision ID SNP Panels with the HID Ion S5/HID Ion Gene Studio S5 System Application Guide (MAN0017767). The reagent cartridge was thawed 45 minutes to sequencing. The instrument was Initialized first and then the first chip was loaded. The post run cleaning was disabled to enable the loading of the second chip.



Figure 7. The Gene Studio and the 530 Chips used for Sequencing oif the samples.

After the first run, the second chip was loaded, and post run cleaning enabled. After the completion of the runs, the run parameters for each sample were analyzed and the Bam files were downloaded and saved locally for further analysis.

3.11. Data analysis

After the run, the BAM files were saved locally and the haplogroups were evaluated using Y-leaf Software v3.1. using Ubuntu operating system.

CHAPTER IV

4.0 Results and Discussion

4.1 Results

In this study, 13 unique Y chromosome SNPs haplogroups were identified from 32 unrelated male individuals representing 40.6% diversity. Out of over 20 major haplogroups, these haplogroups originate from three major haplogroups namely A, B and E as shown in figures 8, 9, 10 and 11 (a), (b)& (c) that are substantially present throughout Africa from the previous studies (Naidoo, 2014).

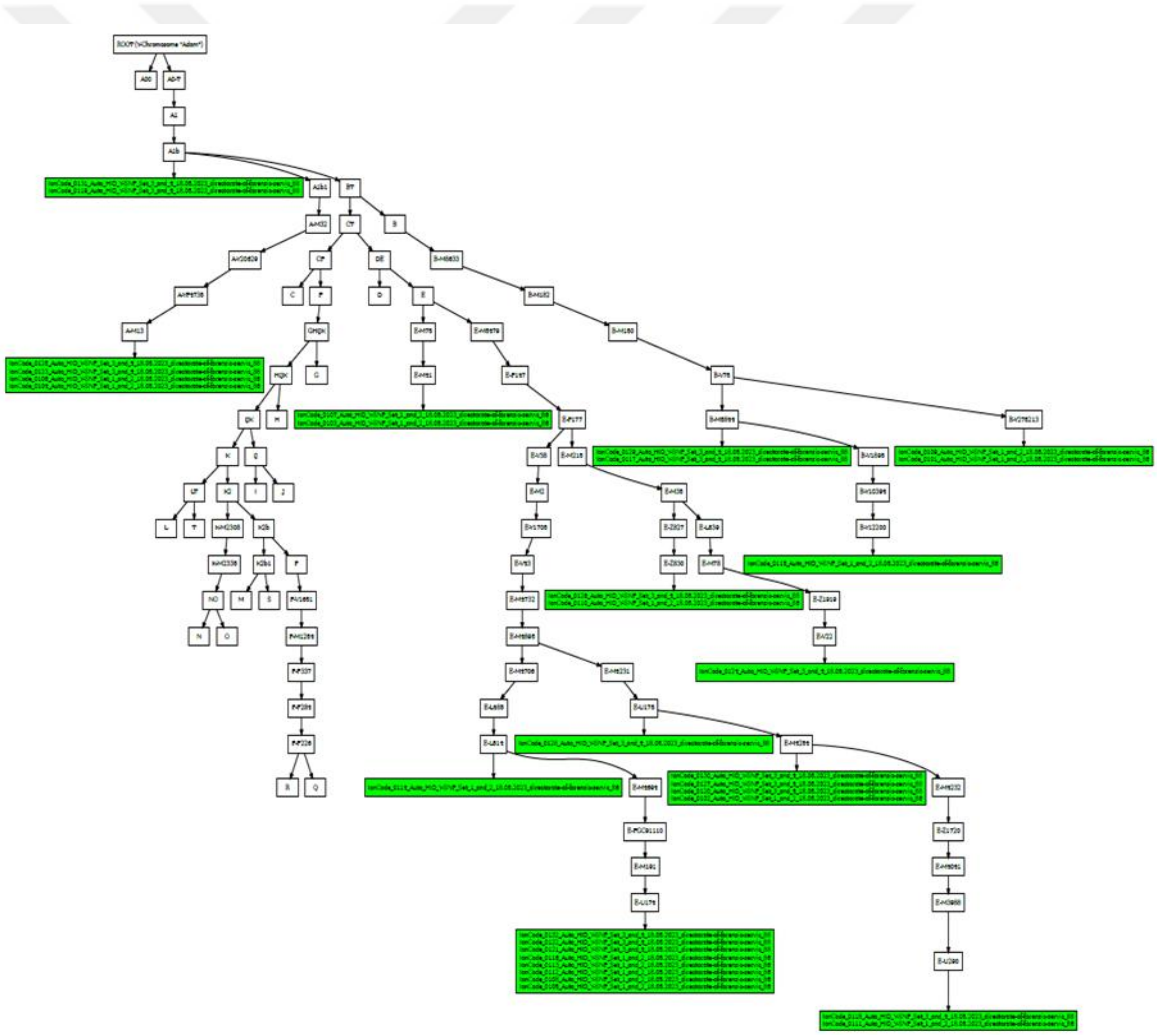


Figure 8 The Y Chromosome SNPs Haplogroups of the Bagisu and the Sabiny of the Elgon region in Eastern Uganda.

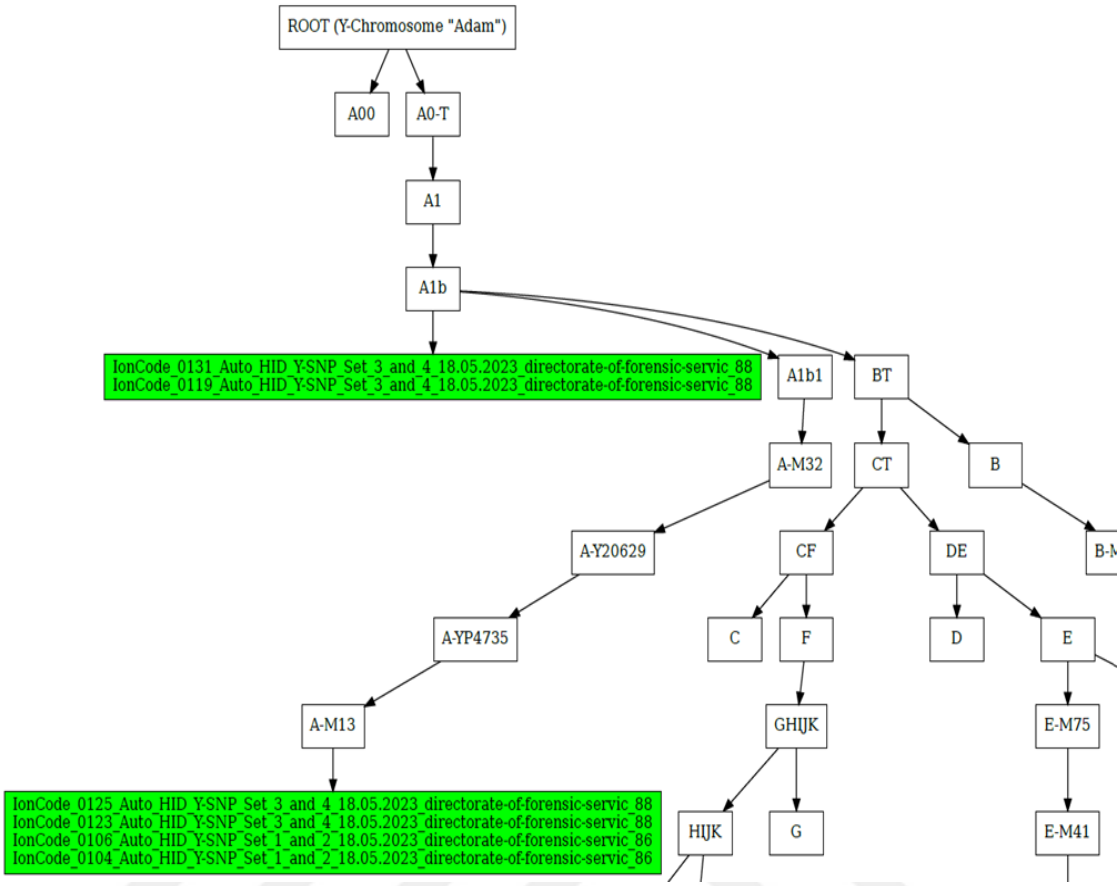


Figure 9 The haplogroups under A major haplogroup (A1b and A-M13).

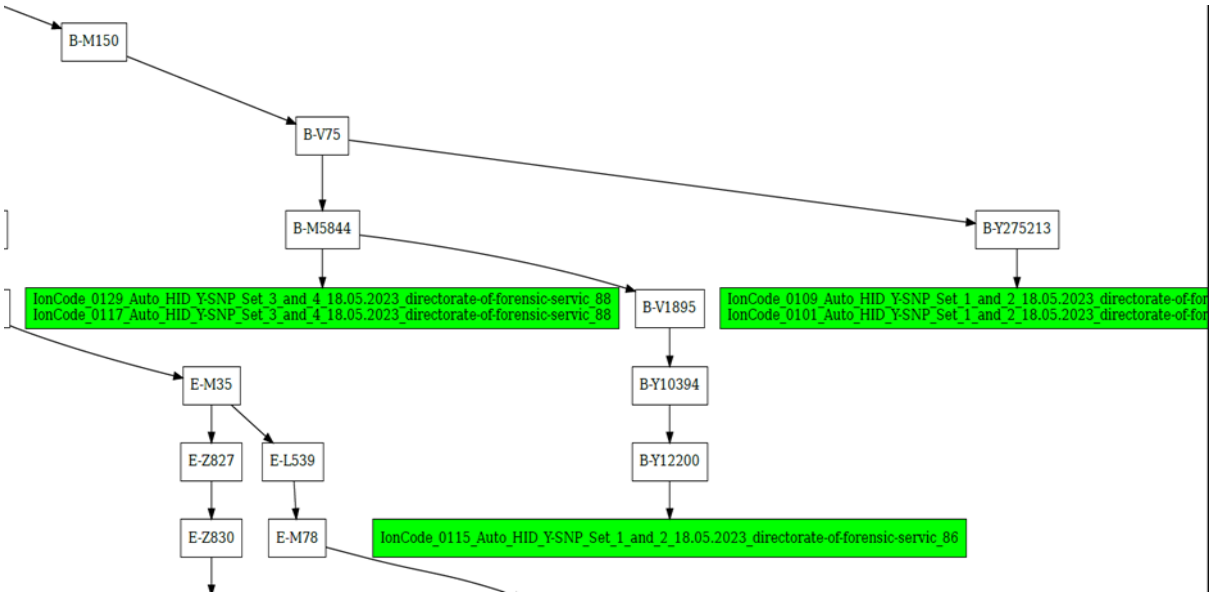
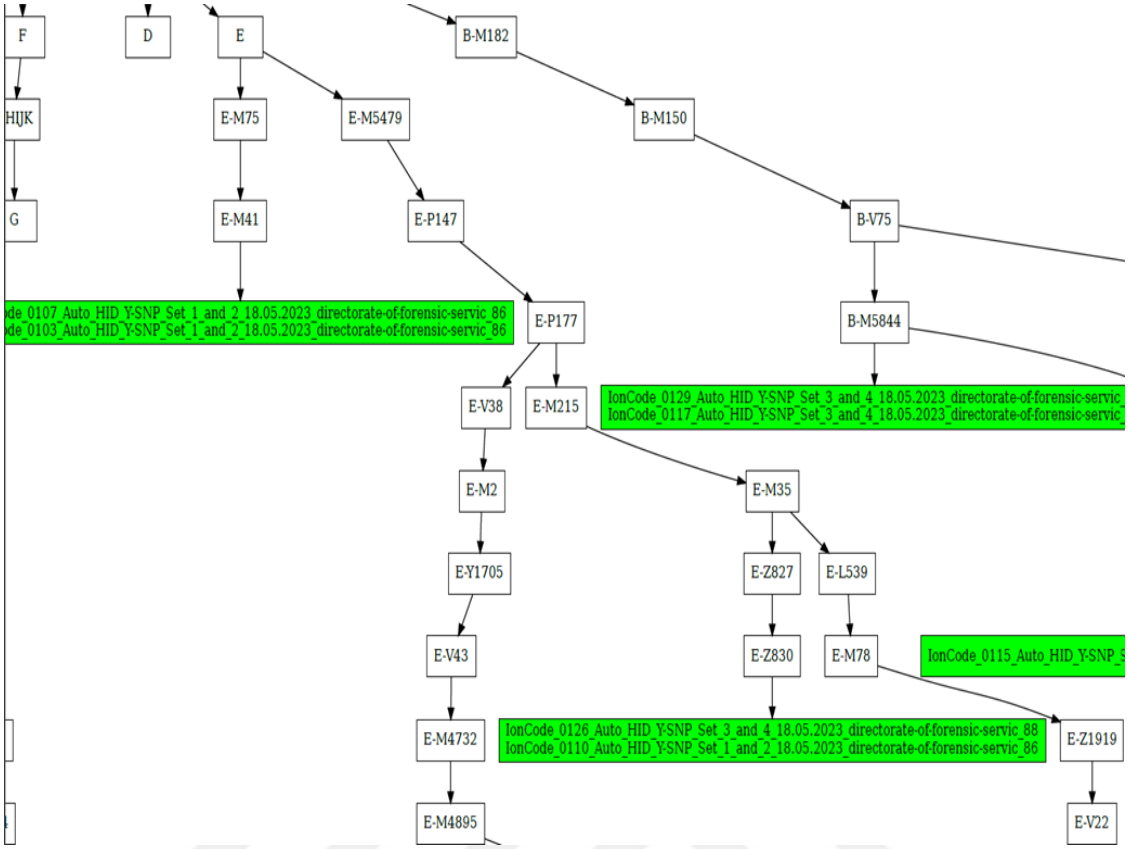
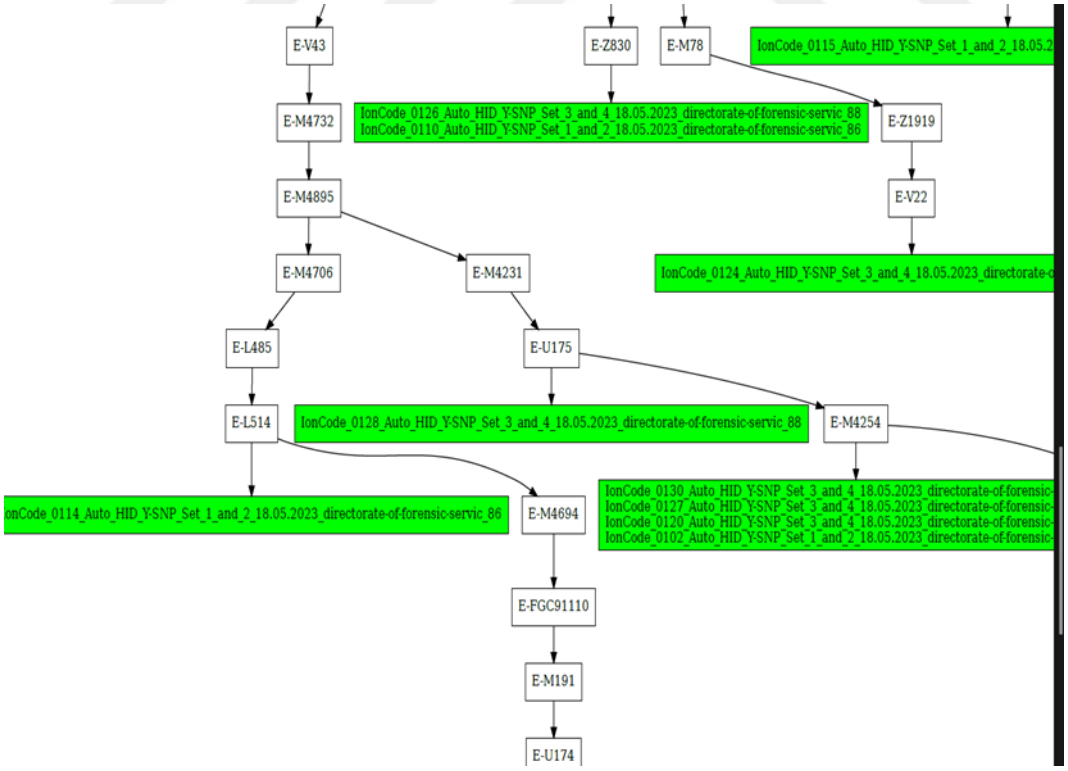


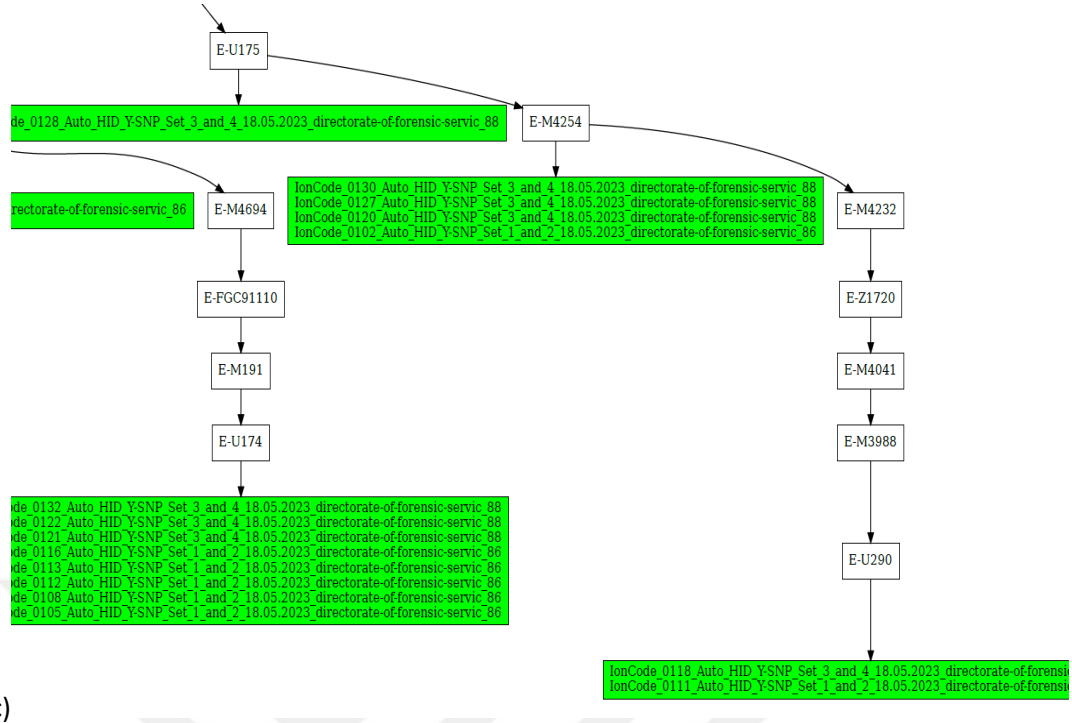
Figure 10. The haplogroups under B major haplogroup (B-M5844, B-Y275213 and B-Y12200)


F
D
E
B-M182
HIJK
E-M75
E-M5479
B-M150
G
E-M41
E-P147
B-V75
E-P177
B-M5844
E-V38
E-M215
E-M2
E-M35
E-Y1705
E-Z827
E-L539
E-V43
E-Z830
E-M78
E-M4732
E-Z1919
E-M4895
E-V22

a)


E-V43
E-Z830
E-M78
E-M4732
E-Z1919
E-M4895
E-V22
E-M4706
E-M4231
E-L485
E-U175
E-L514
E-M4254
E-M4694
E-FGC91110
E-M191
E-U174

b)



c)

Figure 11 The haplogroups under E major haplogroup (E-M41, E-Z830, E-V22, E-U175, E-L514, E-M4254, E-U174 and E-U290).

The frequencies of the observed haplogroups were presented in the form of percentages as shown in the figure below.

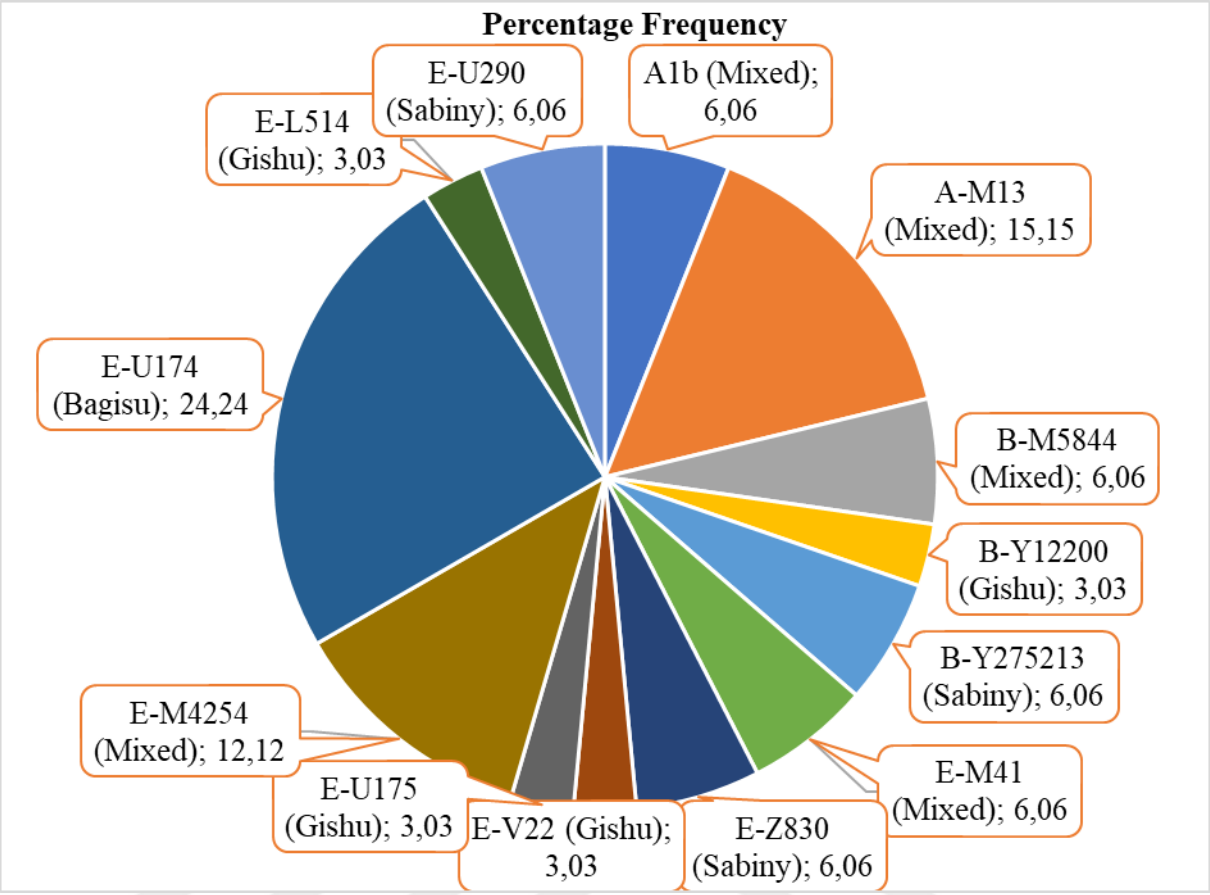


Figure 12 The frequency distribution of the Y Chromosome SNPs Haplogroups of the Bagisu and the Sabiny of the Elgon Population in Eastern Uganda.

4.2 Discussion of the results

Haplogroup E showed the highest frequency of 63.63% in the Elgon population just like the previous studies on the African populations (Naidoo, 2014) followed by haplogroup A with the frequency of 21.21% and haplogroup B with a frequency of 15.15%.

Haplogroup E-U174 (Fig.11c) was observed to be the highest frequency of 24.24% (Fig. 12). and only among the Bagisu population This genetic marker is seen to be found in Africa, particularly among populations in East and Northeast Africa. It is associated with the migration of people to the Horn of Africa and the Arabian Peninsula (Trombetta *et al*., 2015).

Research has shown that haplogroup E-U174 is most frequently found in Ethiopia, particularly among the Oromo ethnic group. It is also found in smaller frequencies among ethnic groups in Sudan, Djibouti, and Somalia. The distribution of haplogroup E-U174 in Africa is particularly interesting because it overlaps with the migration patterns of ancient populations in the region. This suggests that this genetic marker may be linked to the movement of people in this area throughout history (Trombetta *et al*., 2015).

Haplogroups B-Y275213 (Fig.10), E-Z830 (Fig.11a) and E-U290 (Fig.11c) were observed among the Sabiny population only (Fig. 11). These genetic lineages are predominantly found in Africa. Haplogroup B-Y275213 is a rare genetic lineage that is found in Africa. It is a subclade of the larger haplogroup B-M60, which is primarily found in East and Southeast Asia, as well as parts of Oceania and the Americas (Quintana-Murci *et al*., 2004). B-Y275213 is believed to have originated in Africa and is associated with the Bantu-speaking populations, who have a history of migration and cultural expansion across the continent (Scheinfeldt *et al*., 2010). This proves the possibility of inter-marriages between the Bagisu and the Sabiny (Verdu *et al.*, 2013).

E-Z830 is a subclade of the larger haplogroup E, which is believed to have originated in East Africa and subsequently spread to other parts of the continent coinciding with its presence in the Sabiny population. E-Z830 is most common in West and Central Africa but has been found in other parts of the continent as well (Owolodun *et al*., 2018).

One study published in the American Journal of Physical Anthropology analyzed the Y-chromosome data of several ethnic groups in Cameroon and found a small frequency of E-Z830 among the Bantu-speaking populations. The researchers suggested that this haplogroup may have been introduced to Central Africa through ancient migrations and cultural exchanges (Manni *et al*., 2013).

E-U290 lineage is defined by the SNP mutation U290, which was first identified in a Nigerian population. Research into E-U290 has revealed important insights into the genetic diversity and history of African populations. For example, studies have suggested that this lineage may have originated in West Africa around 7,000-10,000 years ago, and subsequently spread across the region and into Central Africa. E-U290 has been found at particularly high frequencies among certain ethnic groups such as the Fulani, Wolof, and

Mandenka peoples, and has been linked to the spread of pastoralism and other cultural practices in the region. One study examined the genetic diversity of E-U290 in various populations across West and Central Africa and suggested that the lineage may have undergone population expansions and contractions over time due to factors such as migration, admixture, and cultural diffusion (Wood *et al*., 2005).

Haplogroups B-Y12200 (Fig.10), E-V22 (Fig.11b), E-U175 (Fig.11b) and E-L514 (Fig.11b) were observed with low frequencies (Fig.11) among the Bagisu population. Haplogroup B-Y12200 is a genetic marker found in, particularly among populations in Central and West Africa. It is a subclade of haplogroup B, which is one of the oldest Y chromosome lineages in humans. Research has shown that haplogroup B-Y12200 is most frequently found in Nigeria, particularly among the Yoruba ethnic group. It is also found in smaller frequencies among ethnic groups in Cameroon, Ghana, and Benin. The distribution of haplogroup B-Y12200 in Africa is particularly interesting because it is associated with the migration of ancient populations in the region (Henn *et al*., 2012).

Haplogroup E-V22 is a genetic lineage that is predominantly found in Africa. It is a subclade of the larger haplogroup E, which is believed to have originated in East Africa and subsequently spread to other parts of the continent. E-V22 is most common in West and Central Africa but has been found in other parts of the continent as well (Batini *et al*., 2011; Diakite *et al*., 2016). Despite its relatively low frequency in some parts of Africa, E-V22 has important implications for the study of genetic diversity and migration patterns in the continent.

Haplogroup E-U175 is a genetic marker found in Africa, predominantly among populations in East and Northeast Africa. This haplogroup is associated with the Bantu expansion, a significant migration of people speaking Bantu languages across the African continent between approximately 3000 and 1500 years ago. Research has shown that haplogroup E-U175 is most frequently found in Tanzania, particularly among the Chaga and Pare ethnic groups. It is also found in smaller frequencies among ethnic groups in Ethiopia, Sudan, Kenya, and Uganda. The distribution of haplogroup E-U175 in Africa is particularly interesting because it overlaps with the migration patterns of the Bantu peoples (Macholdt *et al*., 2020). This suggests that this genetic marker may be linked to the spread of Bantu languages across the continent.

Haplogroup E-L514 is a genetic marker found in Africa, predominantly among populations in West and Central Africa. This haplogroup is associated with the Bantu expansion, a significant migration of people speaking Bantu languages across the African continent between approximately 3000 and 1500 years ago. Research has shown that haplogroup E-L514 is most frequently found in Cameroon, particularly among the Bamileke and Bantu ethnic groups. It is also found in smaller frequencies among ethnic groups in Nigeria, Gabon, Equatorial Guinea, and Angola. The distribution of haplogroup E-L514 in Africa is particularly interesting because it overlaps with the migration patterns of the Bantu peoples (Henn *et al*., 2012; Pickrell *et al*., 2012 and Oliveira *et al*., 2015). This suggests that this genetic marker may be linked to the spread of Bantu languages across the continent.

However, some haplogroups were observed in both populations and these included A1b (Fig.9), A-M13 (Fig.9), B-M5844 (Fig.10), E-M41 (Fig.11a) and E-M4254 (Fig.11b). Haplogroup A1b is a subclade of the larger A1 haplogroup, which is found at low frequencies in some populations in Central and West Africa. A study published in the American Journal of Physical Anthropology in 2011 found that haplogroup A1b was present in about 6% of individuals from the Bantu-speaking Luba and Songye populations in the Democratic Republic of Congo (Verdu *et al*., 2011).

Haplogroup A-M13 is a subclade of the larger haplogroup A, which is one of the oldest and most widely distributed Y-chromosome haplogroups in Africa. Haplogroup A-M13 is found in low frequencies in many populations across the African continent, as well as in some individuals of African descent in other parts of the world. A study published in the journal Molecular Biology and Evolution in 2014 analyzed the frequency and distribution of haplogroup A-M13 in populations from Ethiopia, Sudan, and Kenya. The authors found that the highest frequencies of A-M13 were observed in the Gumuz and Nilo-Saharan groups in Ethiopia, as well as in the Nubian and Beja populations in Sudan. In addition, the authors noted that the presence of A-M13 in coastal Kenyan populations may be the result of historical migrations (Trombetta *et al*., 2015). However, more research is needed to fully understand the distribution and genetic history of this haplogroup.

Haplogroup B-M5844 is a genetic marker found in Africa, particularly among populations in Central and West Africa. It is a subclade of haplogroup B, which is one of the oldest Y chromosome lineages in humans. Research has shown that haplogroup B-M5844 is most frequently found in Cameroon, particularly among the Bantu ethnic groups in the region (Henn *et al*., 2011). It is also found in smaller frequencies among ethnic groups in Nigeria, Gabon, and Equatorial Guinea (Batini *et al*., 2011). This suggests that this genetic marker may be linked to the expansion of Bantu languages and cultures across Africa.

Haplogroup E-M41 also known as E1b1, is a genetic lineage that is predominantly found in Africa. It is a subclade of the larger haplogroup E, which is believed to have originated in East Africa and subsequently spread to other parts of the continent. E-M41 is most common in West and Central Africa but has been found in other parts of the continent as well (Cruciani *et al*., 2007). One study published in the American Journal of Physical Anthropology analyzed the genetic data of several populations in Africa and identified E-M41 as the most common haplogroup in Bantu-speaking populations (Salas *et al*., 2002).

Haplogroup E-M4254 is a genetic marker found in Africa, primarily among populations in East and Northeast Africa. This haplogroup is thought to have originated in Ethiopia and is associated with the migration of people to the Horn of Africa and the Arabian Peninsula. Research has shown that haplogroup E-M4254 is most frequently found in Ethiopia, particularly among the Amhara and Tigray ethnic groups (Pagan *et al*., 2015). It is also found in smaller frequencies among ethnic groups in Somalia, Djibouti, and Yemen (Trombetta *et al.*, 2015). The distribution of haplogroup E-M4254 in Africa is particularly interesting because it overlaps with the migration patterns of ancient populations in the region. This suggests that this genetic marker may be linked to the movement of people in this area throughout history.

4.3 Limitations and Challenges of the Study

However, the study had some limitations and challenges which included the following.

The Y chromosome is present in males only, so it can be used to study the genetic diversity of males in a population but not females.

Y chromosome SNPs haplogroups could help investigators determine where the suspect is from, which could be valuable information in identifying the suspect but cannot be used to single out an individual independently.

It is very expensive and costly to run.

4.4 Conclusion and Recommendations.

The results of this study are consistent with previous research on Y chromosome SNPs in African populations, which have shown high levels of genetic diversity (Henn et al., 2008; Quiros et al., 2017; Pinhasi et al., 2020). The 13 unique Y chromosome SNPs identified in this study could be used to differentiate the Elgon population from other populations in Uganda and are important for forensic DNA analysis.

In addition, the high degree of genetic diversity in the Elgon population suggests that there has been little genetic drift, which is consistent with the historical migration patterns and cultural practices of many African populations. This phenomenon been studied in previous research on African populations, which have shown that traditional cultural practices such as polygamy can increase genetic diversity (Henn et al., 2008).

In conclusion, the research on the Forensic Analysis of the Y Chromosome SNPs in the Bagisu and the Sabiny of Elgon Population in Eastern Uganda revealed a high degree of genetic diversity in the Elgon population, which is characteristic of many African populations. The identification of unique Y chromosome SNPs could be useful in cases where a male DNA profile is needed, such as in cases of sexual assault, paternity testing, or missing persons investigations. The results of this study add to the growing body of literature on the genetic diversity of African populations.

Based on the findings of the research on the Forensic Analysis of the Y Chromosome SNPs in the Elgon Population of Eastern Uganda, the following recommendations should be considered:

Further studies should be conducted to investigate the genetic diversity of other populations in Uganda using Y chromosome SNPs with a large sample size as this will help to fully understand the genetic diversity of the country.

The unique Y chromosome SNPs identified in this research should be considered for use in forensic analysis and population genetics studies, as they could be used to differentiate the Elgon population from other populations in Uganda.

The results of this study could be used to improve the accuracy of forensic DNA analysis and could be helpful in solving crimes, especially in cases where there is limited information about the suspect.

5. REFERENCES

1. **Ballard D, Winkler-Galicki J, Wesoły J.** Massive parallel sequencing in forensics: advantages, issues, technicalities, and prospects. *Int J Legal Med*. **2020;**134(4):1291-1303. doi:10.1007/s00414-020-02294-0.

2. **Batini C, Ferri G, Destro-Bisol G, et al.** Signatures of the pre-agricultural peopling processes in sub-Saharan Africa as revealed by the phylogeography of early Y chromosome lineages. Mol Biol Evol. **2011**;28(9):2603-2613.

3. **Butler JM, Coble MD, Vallone PM.** STRs Vs SNPs: thoughts on the future of forensic DNA testing. *Forensic Sci Med.* **2007;** 3:200–205. https://doi.org/10.1007/s12024-007-0018-1

4. **Court DS.** The y chromosome and its use in forensic dna analysis. *Emerg Top Life Sci*. **2021**;5(3):427-441. doi:10.1042/ETLS20200339

5. **Cruciani, F., Santolamazza, P., Shen, P., Macaulay, V., Moral, P., Olckers, A., … Underhill, P. A. A** back migration from Asia to sub-Saharan Africa is supported by high-resolution analysis of human Y-chromosome haplotypes. *American Journal of Human Genetics*. **2002**;*70*(5), 1197–1214. https://doi.org/10.1086/340257.

6. **Cruciani, F., La Fratta, R., Trombetta, B., Santolamazza, P., Sellitto, D., Colomb, E. B., & Scozzari, R.** (2007). Tracing past human male movements in northern/eastern Africa and western Eurasia new clues from Y-chromosomal haplogroups E-M78 and J-M12. Molecular biology and evolution, 24(6), 1300-1311

7. **Diakite, M., Fagny, M., & Galvani, A. P.** (2016). Fine-scale study of local and global genetic diversity of Ghanaians and Nigerians using genome-wide SNP data. BMC genetics, 17(1), 1-13.

8. **Gomes, V., Sánchez-Diz, P., Amorim, A., Carracedo, Á., & Gusmão, L.** Digging deeper into East African human y chromosome lineages. *Human Genetics*. **2010;***127*(5), 603–613. https://doi.org/10.1007/s00439-010-0808-5

9. **Hassan, H. Y., Underhill, P. A., Cavalli-Sforza, L. L., & Ibrahim, M. E.** (2008). Y-chromosome variation among sudanese: Restricted gene flow, concordance with language, geography, and history. *American Journal of Physical Anthropology*, *137*(3), 316–323. https://doi.org/10.1002/ajpa.20876

10. **Henn BM, Gignoux CR, Feldman MW, et al.** (2008). "Y-chromosomal evidence of a pastoralist migration through Tanzania to southern Africa." Proceedings of the National Academy of Sciences, Vol. 105, no. 31, pp. 10693-10698.

11. **Henn BM, Gignoux CR, Jobin M, et al.** Hunter-gatherer genomic diversity suggests a southern African origin for modern humans. Proc Natl Acad Sci USA. **2011**;108(13):5154-5162.

12. **Henn BM, Botigué LR, Gravel S, et al.** Gen ancestry of North Africans supports back-to-Africa migrations. PLoS Genet. **2012**;8(1): e1002397.

13. **Henn BM, Cavalli-Sforza LL, Feldman MW.** (2012). "The great human expansion". Proceedings of the National Academy of Sciences, 109(44), pp. 17758-17764.

14. **Horai, S., Hayasaka, K., Kondo, R., Tsugane, K., & Takahata, N.** Recent African origin of modern humans revealed by complete sequences of hominoid mitochondrial DNAs. *Proceedings of the National Academy of Sciences of the United States of America*. **1995**;*92*(2), 532–536. https://doi.org/10.1073/pnas.92.2.532

15. **Imam J, Rana AK, Forensic C.** DNA Fingerprinting: Advancements and Future Endeavors. *DNA*

Fingerprinting Adv Futur Endeavors. **2018**;(July 2019). doi:10.1007/978-981-13-1583-1

16. **Isabirye, D.** *The mitochondrial DNA heritage of the Baganda , Lugbara and Acholi from Uganda The mitochondrial DNA heritage of the Baganda , Lugbara and Acholi from Uganda*. **2010**.

17. **Jobling MA, Tyler-Smith C.** The human Y chromosome: an evolutionary marker comes of age. Nat Rev Genet. **2003**;4(8):598-612. doi:10.1038/nrg1124.

18. **Jobling, M. A., & Tyler-Smith,** *C.* (2017). The human Y chromosome: an evolutionary marker comes of age. Nature Reviews Genetics, 18(2), 80-92.

19. **Joo SM.** Investigation of SNPs in the Precision ID Identity Panel using next-generation sequencing in a Myanmar population. Published online **2022**:1-12.

20. **Karafet, T. M., Mendez, F. L., Meilerman, M. B., Underhill, P. A., Zegura, S. L., & Hammer, M. F.** New binary polymorphisms reshape and increase resolution of the human Y chromosomal haplogroup tree. *Genome Research*. **2008**;*18*(5), 830–838. https://doi.org/10.1101/gr.7172008

21. **Kayser M.** Forensic use of Y-chromosome DNA: a general overview. *Hum Genet*. **2017**;136(5):621-635. doi:10.1007/s00439-017-1776-9

22. **Köksal Z, Burgos G, Carvalho E, et al.** Testing the Ion AmpliSeq™ HID Y-SNP Research Panel v1 for performance and resolution in admixed South Americans of haplogroup Q. *Forensic Sci Int Genet*. **2022**;59(January). doi: 10.1016/j.fsigen.2022.102708

23. **Lema, G., Juma, A. T., Moore, J. H., Hirbo, J. B., Froment, A., Weber, J. L., … Mortensen, H.** The Genetic Structure and History of Africans and African Americans. *Science*. **2009;***324*(5930), 1035–1044. https://doi.org/10.1126/science.1172257

24. **Maan, A. A., Eales, J., Akbarov, A., Rowland, J., Xu, X., Jobling, M. A., … Tomaszewski, M.** The y chromosome: A blueprint for men's health? *European Journal of Human Genetics*. **2017**;*25*(11), 1181–1188. https://doi.org/10.1038/ejhg.2017.128.

25. **Macholdt E, Slatkin M, Pakendorf B, Stoneking M.** New insights into the history of the Bantu languages and Bantu speaker populations of Africa from high-density genetic data. Am J Hum Genet. **2020**;106(6):789-804.

26. **Manni, F., Guérard, E., Heyer, E., & Chaix, R.** (2013). Fine-scale genetic structure and differential male and female migration rates in Baoulé and Djimini populations of Ivory Coast. American Journal of Physical Anthropology, 152(4), 495-504.

27. **Mbaka JN**. (2017). "Biodiversity of cosmology and indigenous knowledge systems among the Sabiny of Uganda". Journal of Ethnobiology and Ethnomedicine, 13(1), p. 11.

28. **Mukhala E.** (2015). "Sebei Indigenous Cattle Breeds: Development Prospects and Conservation Challenges in the Context of a Changing Global Environment". CABI.

29. **Naidoo, F. T. *P***. *Hylogeography of Y Chromosome Haplogroups a & B in a Frica*. **2014.**

30. **Namyalo, S., & Nakayiza, J.** Dilemmas in implementing language rights in multilingual Uganda. *Current Issues in Language Planning*. **2015**;*16*(4), 409–424. https://doi.org/10.1080/14664208.2014.987425.

31. **Oliveira S, LeslieJ, Okafor FC, et al.** Cristóvão Colombo, African slaves, and the new world: a genetic legacy. Ann Hum Genet. **2015**;79(6):391-401

32. **Owolodun, O. A., Ogunbiyi, O. J., & Oyedeji, S. I.** (2018). Phylogenetic Analysis of Y Chromosome Haplogroup E-M183 and E-M215 in Three Ethnic Groups in Southwest Nigeria. Journal of Medical Sciences, 38(2), 52-59.

33. **Quintana-Murci, L., Chaix, R., Wells, R. S., Behar, D. M., Sayar, H., Scozzari, R., ... & Semino, O.** (2004). Where west meets east: the complex mtDNA landscape of the southwest and Central Asian corridor. The American Journal of Human Genetics, 74(5), 827-845.

34. **Quiros P, Mendoza-Revilla J, Leal S, et al.** (2017). "Y-chromosome variation in South African Bantu-speakers: A reflection of a click-speaking pastoralist ancestry." PLoS ONE, Vol. 12, no. 6, e0179433.

35. **Pickrell JK, Patterson N, Loh PR, et al.** Ancient west African foragers in the context of African population history. Nature. **2014**;506(7481):353-357.

36. **Pinhasi R, Fernandes D, Sirak K, et al.** (2020). "Genetic history of the Karamojong population from northeastern Uganda." Human Genetics, Vol. 139, no.2, pp. 211-223.

37. **Ralf, A., Oven, M., Gonzalez, O., de Kniff, P., der Beek, K., Wootton, S., Lagace, R., Kayser, M.** 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 Sci Int Genet. **2019**;41(July), 93-106.

38. **Salas, A., Richards, M., De la Fe, T., Lareu, M. V., Sobrino, B., Sánchez-Diz, P., … Carracedo, Á.** The making of the African mtDNA landscape. *American Journal of Human Genetics*. **2002**;*71*(5), 1082–1111. https://doi.org/10.1086/344348.

39. **Scheinfeldt, L. B., Tishkoff, S. A., & Wolpoff, M. H.** (2010). The structure of Bantu genetic diversity. Evolutionary Anthropology: Issues, News, and Reviews, 19(6), 259-273.

40. **Soares P, Ermini L, Thomson N, et al.** Correcting for purifying selection: an improved human mitochondrial molecular clock. Am J Hum Genet. **2009**;84(6):740-759. doi: 10.1016/j.ajhg.2009.05.001.

41. **Trombetta B, D'Atanasio E, Massaia A, et al.** Phylogeographic refinement and large-scale genotyping of human Y chromosome haplogroup E provide new insights into the dispersal of early pastoralists in the African continent. Genome Biol Evol. **2015**;7(7):1940-1950.

42. **Uganda Investment Authority. Elgon investment profile** 2018. Published online 2019. https://www.ugandainvest.go.ug/wp-content/uploads/2019/06/Elgon-Investment-Profile.pdf

43. **Ugwumba, A. O., Okeke, U. O., Ejike, C. E., Ezechukwu, C. C., & Okoli, C. C.** (2018). Y-Chromosome DNA Polymorphism among Igbo Populations of Nigeria. Advances in Anthropology, 8(02), 98-110.

44. **Underhill, P. A., & Kivisild, T.** Use of Y Chromosome and Mitochondrial DNA Population Structure in Tracing Human Migrations. *Annual Review of Genetics*, *41*(1), 539–564. https://doi.org/10.1146/annurev.genet. **2007**;41.110306.130407.

45. **Verdu P, Becker NS, Froment A, Georges M, Grugni V, Quintana-Murci L, et al**. (2011) Sociocultural behavior, sex-biased admixture, and effective population sizes in Central African Pygmies and non-Pygmies. Am J Phys Anthropol 145: 11–22.

46. **Verdu, P., Becker, N. S., Froment, A., Georges, M., Grugni, V., Quintana-Murci, L., ... & Heyer, E.** (2013). Sociocultural behavior, sex-biased admixture, and effective population sizes in Central African Pygmies and non-Pygmies. Molecular biology and evolution, 30(4), 918-937.

47. **Wood ET, Stover DA, Ehret C, et al.** (2005). "Contrasting patterns of Y chromosome and mtDNA variation in Africa: evidence for sex-biased demographic processes". European Journal of Human Genetics, 13(7), pp. 867-876.

48. **YCC**. A nomenclature system for the tree of Human Y-Chromosomal binary haplogroups. *Genome Res.* **2002**;*12* (2), 339-48

RESUME

Wazodi Amosi was born on 23rd April 1988 in Bulambuli District, Uganda. I completed O' level education in the year 2006 at Nabbongo Secondary School and A' level (High School) in the year 2008 at Mbale Secondary School. I graduated with a bachelor's degree in Fisheries and Aquaculture from Makerere University, College of Natural Sciences in the year 2012 with research on the performance of the African Catfish Larvae fed on *Artemia salina* and Zoo planktons. In the year 2014, I joined the Uganda Police Force and in the year 2018, I joined the Directorate of Forensic Services, DNA analysis Section where am currently working up to date. I have received trainings in DNA profiling, Collection, Preservation and Presentation of Forensic evidence, Precision ID NGS System for Human Identification from Africa Biosystems, and Hamilton STARline Operators from Hamilton. I have experience in Basic Human Identification workflow using both Genetic Analyzer and Next Generation Sequencing Technology.

Referees

1. **Andrew K. Mubiru**
   Director Forensic Services
   Uganda Police Force
   +256781402888, +256779621751

2. **Dr. Eric Sande**
   Head, Department of Zoology, Entomology and Fisheries Sciences
   Makerere University College of Natural Sciences
   eric.sande@mak.ac.ug, ericsandephd@gmail.com
   +256772 688 552, +256701688552, +256703368630.

4. **Dr. Inosters W. Nzuki (PhD)**

   Africa Biosystems Limited

   +254724502004

Appendix II. Table 1 Showing the Y Chromosome SNPs Haplogroups diversity of the Bagisu and the Sabiny of Elgon region in Eastern Uganda.

Sample_name	Haplogroups Diversity
IonCode_0119 (KWNY23S003)	A1b*(xA1b1, B, CF, D, E-P147, E-M41, E-M90)
IonCode_0131 (NAMY23S003)	A1b*(xA1b1, B, CF, D, E-P147, E-M41, E-M98)
IonCode_0104 (MBLY23S001)	A-M13*(xA-V5071, A-Y95250)
IonCode_0106 (BULY23S001)	A-M13*(xA-V5071, A-Y95250)
IonCode_0123 (NAMY23S002)	A-M13*(xA-V5071, A-Y95250)
IonCode_0125 (KAPY23S004)	A-M13*(xA-V5071, A-Y95250)
IonCode_0117 (KAPY23S003)	B-M5844*(xB-A9832, B-Y12200, B-Y33569)
IonCode_0129 (SRKY23S003)	B-M5844*(xB-A9832, B-Y10394, B-Y33569)
IonCode_0115 (NAMY23S001)	B-Y12200
IonCode_0109 (KAPY23S002)	B-Y275213
IonCode_0101 (KAPY23S001)	B-Y275213
IonCode_0103 (KWNY23S001)	E-M41
IonCode_0107 (MANY23S001)	E-M41
IonCode_0110 (BUKY23S002)	E-Z830*(xE-M123, E-Y90711, E-FT18871, E-M293, E-Y30000, E-V42, E-V6)

IonCode_0126 (BUKY23S004)	E-Z830*(xE-M123, E-Y90711, E-PH1646, E-FT18871, E-M293, E-Y30000, E-V42, E-V6)
IonCode_0124 (BUDY23S003)	E-V22*(xE-BY7566, E-FT185103)
IonCode_0128 (MBLY23S004)	E-U175*(xE-M4254)
IonCode_0102 (BUKY23S001)	E-M4254*(xE-U290)
IonCode_0120 (MBLY23S003)	E-M4254*(xE-U290)
IonCode_0127 (KWNY23S004)	E-M4254*(xE-U290)
IonCode_0130 (MANY23S004)	E-M4254*(xE-U290)
IonCode_0114 (BULY23S002)	E-L514*(xE-M191)
IonCode_0105 (SRKY23S001)	E-U174*(xE-FT212537, E-Y96183)
IonCode_0108 (BUDY23S001)	E-U174*(xE-FT212537, E-Y96183)
IonCode_0112 (MBLY23S002)	E-U174*(xE-FT212537, E-Y96183)
IonCode_0113 (SRKY23S002)	E-U174*(xE-FT212537, E-Y96183)
IonCode_0116 (BUDY23S002)	E-U174*(xE-FT212537, E-Y96183)
IonCode_0121 (BULY23S003)	E-U174*(xE-FT212537, E-Y96183)
IonCode_0122 (MANY23S003)	E-U174*(xE-FT212537, E-Y96183)
IonCode_0132 (BULY23S004)	E-U174*(xE-FT212537, E-Y96183)

IonCode_0111 (KWNY23S002)	E-U290*(xE-U181, E-L649)
IonCode_0118 (BUKY23S003)	E-U290*(xE-U181, E-L649)

Appendix III. Coverage Analysis report of the samples.

Coverage Analysis Report

Sample Name: KAPY23S001

IonCode_0101_Auto_HID_Y-SNP_Set_1_and_2_18.05.2023_directorate-of-forensic-servic_86

Library type: AmpliSeq DNA
Reference: hg19 (DNA)
Target regions: ytree.20191001.designed
Read filters: Sample tracking



Number of mapped reads	1,213,633
Percent reads on target	71.50%
Percent sample tracking reads	0.00%
Average base coverage depth	1,433
Uniformity of base coverage	97.84%

Amplicon Read Coverage		Target Base Coverage	
Number of amplicons	602	Bases in target regions	47,438
Percent assigned amplicon reads	71.50%	Percent base reads on target	58.78%
Average reads per amplicon	1,441	Average base coverage depth	1,433
Uniformity of amplicon coverage	99.17%	Uniformity of base coverage	97.84%
Amplicons with at least 1 read	100.00%	Target base coverage at 1x	100.00%
Amplicons with at least 20 reads	100.00%	Target base coverage at 20x	99.83%
Amplicons with at least 100 reads	100.00%	Target base coverage at 100x	99.50%
Amplicons with at least 500 reads	93.85%	Target base coverage at 500x	91.80%
Amplicons with no strand bias	91.53%	Target bases with no strand bias	87.78%
Amplicons reading end-to-end	92.52%	Percent end-to-end reads	90.96%
Amplicon base composition bias	0.064		

Sample Name: BUKY23S001

IonCode_0102_Auto_HID_Y-SNP_Set_1_and_2_18.05.2023_directorate-of-forensic-servic_86

Library type: AmpliSeq DNA
Reference: hg19 (DNA)
Target regions: ytree.20191001.designed
Read filters: Sample tracking



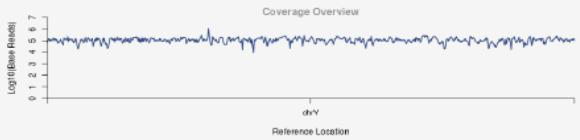
Number of mapped reads	1,248,510
Percent reads on target	70.65%
Percent sample tracking reads	0.00%
Average base coverage depth	1,458
Uniformity of base coverage	97.68%

Amplicon Read Coverage		Target Base Coverage	
Number of amplicons	602	Bases in target regions	47,438
Percent assigned amplicon reads	70.65%	Percent base reads on target	57.95%
Average reads per amplicon	1,465	Average base coverage depth	1,458
Uniformity of amplicon coverage	98.67%	Uniformity of base coverage	97.68%
Amplicons with at least 1 read	100.00%	Target base coverage at 1x	100.00%
Amplicons with at least 20 reads	100.00%	Target base coverage at 20x	99.86%
Amplicons with at least 100 reads	99.83%	Target base coverage at 100x	99.73%
Amplicons with at least 500 reads	94.85%	Target base coverage at 500x	92.42%
Amplicons with no strand bias	91.20%	Target bases with no strand bias	87.95%
Amplicons reading end-to-end	93.19%	Percent end-to-end reads	91.96%
Amplicon base composition bias	0.250		

Sample Name: KWNY23S001

IonCode_0103_Auto_HID_Y-SNP_Set_1_and_2_18.05.2023_directorate-of-forensic-servic_86

Library type: AmpliSeq DNA
Reference: hg19 (DNA)
Target regions: ytree.20191001.designed
Read filters: Sample tracking



Number of mapped reads	1,273,211
Percent reads on target	71.83%
Percent sample tracking reads	0.00%
Average base coverage depth	1,506
Uniformity of base coverage	96.72%

Amplicon Read Coverage		Target Base Coverage	
Number of amplicons	602	Bases in target regions	47,438
Percent assigned amplicon reads	71.83%	Percent base reads on target	58.98%
Average reads per amplicon	1,519	Average base coverage depth	1,506
Uniformity of amplicon coverage	98.50%	Uniformity of base coverage	96.72%
Amplicons with at least 1 read	100.00%	Target base coverage at 1x	99.84%
Amplicons with at least 20 reads	99.83%	Target base coverage at 20x	99.66%
Amplicons with at least 100 reads	99.67%	Target base coverage at 100x	99.46%
Amplicons with at least 500 reads	94.68%	Target base coverage at 500x	92.23%
Amplicons with no strand bias	91.69%	Target bases with no strand bias	88.18%
Amplicons reading end-to-end	93.36%	Percent end-to-end reads	92.46%
Amplicon base composition bias	0.167		

Sample Name: MBLY23S001

IonCode_0104_Auto_HID_Y-SNP_Set_1_and_2_18.05.2023_directorate-of-forensic-servic_86

Library type: AmpliSeq DNA
Reference: hg19 (DNA)
Target regions: ytree.20191001.designed
Read filters: Sample tracking



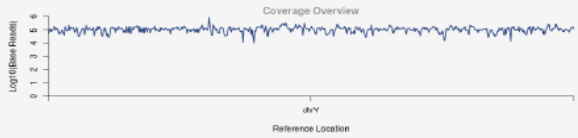
Number of mapped reads	1,236,124
Percent reads on target	71.68%
Percent sample tracking reads	0.00%
Average base coverage depth	1,484
Uniformity of base coverage	97.13%

Amplicon Read Coverage		Target Base Coverage	
Number of amplicons	602	Bases in target regions	47,438
Percent assigned amplicon reads	71.68%	Percent base reads on target	59.36%
Average reads per amplicon	1,472	Average base coverage depth	1,484
Uniformity of amplicon coverage	97.84%	Uniformity of base coverage	97.13%
Amplicons with at least 1 read	100.00%	Target base coverage at 1x	100.00%
Amplicons with at least 20 reads	99.83%	Target base coverage at 20x	99.57%
Amplicons with at least 100 reads	99.83%	Target base coverage at 100x	99.51%
Amplicons with at least 500 reads	93.02%	Target base coverage at 500x	91.07%
Amplicons with no strand bias	91.20%	Target bases with no strand bias	88.11%
Amplicons reading end-to-end	93.69%	Percent end-to-end reads	92.28%
Amplicon base composition bias	0.245		

Sample Name: SRKY23S001

IonCode_0105_Auto_HID_Y-SNP_Set_1_and_2_18.05.2023_directorate-of-forensic-servic_86

Library type: AmpliSeq DNA
Reference: hg19 (DNA)
Target regions: ytree.20191001.designed
Read filters: Sample tracking



Number of mapped reads	1,133,174
Percent reads on target	74.67%
Percent sample tracking reads	0.00%
Average base coverage depth	1,427
Uniformity of base coverage	98.16%

Amplicon Read Coverage		Target Base Coverage	
Number of amplicons	602	Bases in target regions	47,438
Percent assigned amplicon reads	74.67%	Percent base reads on target	62.53%
Average reads per amplicon	1,405	Average base coverage depth	1,427
Uniformity of amplicon coverage	99.00%	Uniformity of base coverage	98.16%
Amplicons with at least 1 read	100.00%	Target base coverage at 1x	100.00%
Amplicons with at least 20 reads	100.00%	Target base coverage at 20x	99.82%
Amplicons with at least 100 reads	99.83%	Target base coverage at 100x	99.75%
Amplicons with at least 500 reads	93.36%	Target base coverage at 500x	90.90%
Amplicons with no strand bias	90.53%	Target bases with no strand bias	87.08%
Amplicons reading end-to-end	93.85%	Percent end-to-end reads	92.68%
Amplicon base composition bias	0.264		

Sample Name: BULY23S001

IonCode_0106_Auto_HID_Y-SNP_Set_1_and_2_18.05.2023_directorate-of-forensic-servic_86

Library type: AmpliSeq DNA
Reference: hg19 (DNA)
Target regions: ytree.20191001.designed
Read filters: Sample tracking



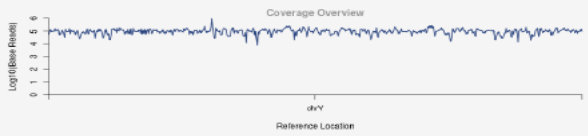
Number of mapped reads	732,080
Percent reads on target	74.59%
Percent sample tracking reads	0.00%
Average base coverage depth	878.2
Uniformity of base coverage	96.18%

Amplicon Read Coverage		Target Base Coverage	
Number of amplicons	602	Bases in target regions	47,438
Percent assigned amplicon reads	74.59%	Percent base reads on target	62.30%
Average reads per amplicon	907.1	Average base coverage depth	878.2
Uniformity of amplicon coverage	97.84%	Uniformity of base coverage	96.18%
Amplicons with at least 1 read	100.00%	Target base coverage at 1x	100.00%
Amplicons with at least 20 reads	99.83%	Target base coverage at 20x	99.53%
Amplicons with at least 100 reads	99.34%	Target base coverage at 100x	98.71%
Amplicons with at least 500 reads	79.90%	Target base coverage at 500x	78.06%
Amplicons with no strand bias	92.03%	Target bases with no strand bias	88.22%
Amplicons reading end-to-end	92.36%	Percent end-to-end reads	90.02%
Amplicon base composition bias	0.195		

Sample Name: MANY23S001

IonCode_0107_Auto_HID_Y-SNP_Set_1_and_2_18.05.2023_directorate-of-forensic-servic_86

Library type: AmpliSeq DNA
Reference: hg19 (DNA)
Target regions: ytree.20191001.designed
Read filters: Sample tracking



Number of mapped reads	1,062,639
Percent reads on target	74.39%
Percent sample tracking reads	0.00%
Average base coverage depth	1,281
Uniformity of base coverage	97.44%

Amplicon Read Coverage		Target Base Coverage	
Number of amplicons	602	Bases in target regions	47,438
Percent assigned amplicon reads	74.39%	Percent base reads on target	62.06%
Average reads per amplicon	1,313	Average base coverage depth	1,281
Uniformity of amplicon coverage	98.50%	Uniformity of base coverage	97.44%
Amplicons with at least 1 read	99.83%	Target base coverage at 1x	99.84%
Amplicons with at least 20 reads	99.83%	Target base coverage at 20x	99.66%
Amplicons with at least 100 reads	99.67%	Target base coverage at 100x	99.15%
Amplicons with at least 500 reads	93.02%	Target base coverage at 500x	89.86%
Amplicons with no strand bias	92.86%	Target bases with no strand bias	89.34%
Amplicons reading end-to-end	93.36%	Percent end-to-end reads	92.41%
Amplicon base composition bias	0.169		

Sample Name: BUDY23S001

IonCode_0108_Auto_HID_Y-SNP_Set_1_and_2_18.05.2023_directorate-of-forensic-servic_86

Library type: AmpliSeq DNA
Reference: hg19 (DNA)
Target regions: ytree.20191001.designed
Read filters: Sample tracking



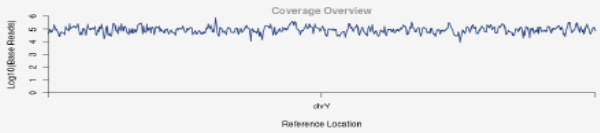
Number of mapped reads	1,087,443
Percent reads on target	73.49%
Percent sample tracking reads	0.00%
Average base coverage depth	1,233
Uniformity of base coverage	95.46%

Amplicon Read Coverage		Target Base Coverage	
Number of amplicons	602	Bases in target regions	47,438
Percent assigned amplicon reads	73.49%	Percent base reads on target	61.67%
Average reads per amplicon	1,328	Average base coverage depth	1,233
Uniformity of amplicon coverage	97.51%	Uniformity of base coverage	95.46%
Amplicons with at least 1 read	100.00%	Target base coverage at 1x	100.00%
Amplicons with at least 20 reads	100.00%	Target base coverage at 20x	99.81%
Amplicons with at least 100 reads	99.67%	Target base coverage at 100x	99.45%
Amplicons with at least 500 reads	93.36%	Target base coverage at 500x	88.94%
Amplicons with no strand bias	90.86%	Target bases with no strand bias	87.88%
Amplicons reading end-to-end	92.86%	Percent end-to-end reads	92.53%
Amplicon base composition bias	0.385		

Sample Name: KAPY23S002

IonCode_0109_Auto_HID_Y-SNP_Set_1_and_2_18.05.2023_directorate-of-forensic-servic_86

Library type: AmpliSeq DNA
Reference: hg19 (DNA)
Target regions: ytree.20191001.designed
Read filters: Sample tracking



Number of mapped reads	977,373
Percent reads on target	70.76%
Percent sample tracking reads	0.00%
Average base coverage depth	1,279
Uniformity of base coverage	97.60%

Amplicon Read Coverage	
Number of amplicons	602
Percent assigned amplicon reads	70.76%
Average reads per amplicon	1,149
Uniformity of amplicon coverage	97.67%
Amplicons with at least 1 read	100.00%
Amplicons with at least 20 reads	100.00%
Amplicons with at least 100 reads	100.00%
Amplicons with at least 500 reads	77.41%
Amplicons with no strand bias	91.36%
Amplicons reading end-to-end	92.86%
Amplicon base composition bias	0.168

Target Base Coverage	
Bases in target regions	47,438
Percent base reads on target	58.47%
Average base coverage depth	1,279
Uniformity of base coverage	97.60%
Target base coverage at 1x	100.00%
Target base coverage at 20x	99.82%
Target base coverage at 100x	99.74%
Target base coverage at 500x	82.09%
Target bases with no strand bias	87.22%
Percent end-to-end reads	90.24%

Sample Name: BUKY23S002

IonCode_0110_Auto_HID_Y-SNP_Set_1_and_2_18.05.2023_directorate-of-forensic-servic_86

Library type: AmpliSeq DNA
Reference: hg19 (DNA)
Target regions: ytree.20191001.designed
Read filters: Sample tracking



Number of mapped reads	1,133,479
Percent reads on target	71.00%
Percent sample tracking reads	0.00%
Average base coverage depth	1,377
Uniformity of base coverage	97.07%

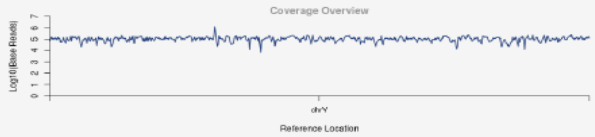
Amplicon Read Coverage	
Number of amplicons	602
Percent assigned amplicon reads	71.00%
Average reads per amplicon	1,337
Uniformity of amplicon coverage	98.17%
Amplicons with at least 1 read	99.83%
Amplicons with at least 20 reads	99.67%
Amplicons with at least 100 reads	99.67%
Amplicons with at least 500 reads	91.20%
Amplicons with no strand bias	91.53%
Amplicons reading end-to-end	93.19%
Amplicon base composition bias	0.137

Target Base Coverage	
Bases in target regions	47,438
Percent base reads on target	58.53%
Average base coverage depth	1,377
Uniformity of base coverage	97.07%
Target base coverage at 1x	99.87%
Target base coverage at 20x	99.28%
Target base coverage at 100x	98.96%
Target base coverage at 500x	90.69%
Target bases with no strand bias	87.62%
Percent end-to-end reads	92.17%

Sample Name: KWNY23S002

IonCode_0111_Auto_HID_Y-SNP_Set_1_and_2_18.05.2023_directorate-of-forensic-servic_86

Library type: AmpliSeq DNA
Reference: hg19 (DNA)
Target regions: ytree.20191001.designed
Read filters: Sample tracking



Number of mapped reads	1,212,469
Percent reads on target	71.89%
Percent sample tracking reads	0.00%
Average base coverage depth	1,406
Uniformity of base coverage	97.55%

Amplicon Read Coverage		Target Base Coverage	
Number of amplicons	602	Bases in target regions	47,438
Percent assigned amplicon reads	71.89%	Percent base reads on target	59.52%
Average reads per amplicon	1,448	Average base coverage depth	1,406
Uniformity of amplicon coverage	98.84%	Uniformity of base coverage	97.55%
Amplicons with at least 1 read	100.00%	Target base coverage at 1x	100.00%
Amplicons with at least 20 reads	100.00%	Target base coverage at 20x	99.82%
Amplicons with at least 100 reads	99.83%	Target base coverage at 100x	99.48%
Amplicons with at least 500 reads	94.52%	Target base coverage at 500x	91.60%
Amplicons with no strand bias	91.53%	Target bases with no strand bias	88.10%
Amplicons reading end-to-end	93.02%	Percent end-to-end reads	92.75%
Amplicon base composition bias	0.194		

Sample Name: MBLY23S002

IonCode_0112_Auto_HID_Y-SNP_Set_1_and_2_18.05.2023_directorate-of-forensic-servic_86

Library type: AmpliSeq DNA
Reference: hg19 (DNA)
Target regions: ytree.20191001.designed
Read filters: Sample tracking



Number of mapped reads	1,174,816
Percent reads on target	71.53%
Percent sample tracking reads	0.00%
Average base coverage depth	1,473
Uniformity of base coverage	98.34%

Amplicon Read Coverage		Target Base Coverage	
Number of amplicons	602	Bases in target regions	47,438
Percent assigned amplicon reads	71.53%	Percent base reads on target	59.22%
Average reads per amplicon	1,396	Average base coverage depth	1,473
Uniformity of amplicon coverage	99.00%	Uniformity of base coverage	98.34%
Amplicons with at least 1 read	100.00%	Target base coverage at 1x	100.00%
Amplicons with at least 20 reads	100.00%	Target base coverage at 20x	99.86%
Amplicons with at least 100 reads	100.00%	Target base coverage at 100x	99.76%
Amplicons with at least 500 reads	89.70%	Target base coverage at 500x	90.32%
Amplicons with no strand bias	92.19%	Target bases with no strand bias	88.57%
Amplicons reading end-to-end	93.02%	Percent end-to-end reads	92.01%
Amplicon base composition bias	0.078		

Sample Name: SRKY23S002

IonCode_0113_Auto_HID_Y-SNP_Set_1_and_2_18.05.2023_directorate-of-forensic-servic_86

Library type: AmpliSeq DNA
Reference: hg19 (DNA)
Target regions: ytree.20191001.designed
Read filters: Sample tracking



Number of mapped reads	1,331,891
Percent reads on target	69.88%
Percent sample tracking reads	0.00%
Average base coverage depth	1,578
Uniformity of base coverage	97.97%

Amplicon Read Coverage		Target Base Coverage	
Number of amplicons	602	Bases in target regions	47,438
Percent assigned amplicon reads	69.88%	Percent base reads on target	57.28%
Average reads per amplicon	1,546	Average base coverage depth	1,578
Uniformity of amplicon coverage	99.00%	Uniformity of base coverage	97.97%
Amplicons with at least 1 read	100.00%	Target base coverage at 1x	100.00%
Amplicons with at least 20 reads	100.00%	Target base coverage at 20x	99.86%
Amplicons with at least 100 reads	100.00%	Target base coverage at 100x	99.66%
Amplicons with at least 500 reads	95.02%	Target base coverage at 500x	93.41%
Amplicons with no strand bias	91.53%	Target bases with no strand bias	87.60%
Amplicons reading end-to-end	93.85%	Percent end-to-end reads	92.67%
Amplicon base composition bias	0.136		

Sample Name: BULY23S002

IonCode_0114_Auto_HID_Y-SNP_Set_1_and_2_18.05.2023_directorate-of-forensic-servic_86

Library type: AmpliSeq DNA
Reference: hg19 (DNA)
Target regions: ytree.20191001.designed
Read filters: Sample tracking



Number of mapped reads	1,153,996
Percent reads on target	74.61%
Percent sample tracking reads	0.00%
Average base coverage depth	1,482
Uniformity of base coverage	98.12%

Amplicon Read Coverage		Target Base Coverage	
Number of amplicons	602	Bases in target regions	47,438
Percent assigned amplicon reads	74.61%	Percent base reads on target	62.60%
Average reads per amplicon	1,430	Average base coverage depth	1,482
Uniformity of amplicon coverage	99.00%	Uniformity of base coverage	98.12%
Amplicons with at least 1 read	100.00%	Target base coverage at 1x	100.00%
Amplicons with at least 20 reads	100.00%	Target base coverage at 20x	99.82%
Amplicons with at least 100 reads	99.83%	Target base coverage at 100x	99.76%
Amplicons with at least 500 reads	93.52%	Target base coverage at 500x	91.37%
Amplicons with no strand bias	91.36%	Target bases with no strand bias	87.77%
Amplicons reading end-to-end	93.19%	Percent end-to-end reads	93.12%
Amplicon base composition bias	0.068		

Sample Name: NAMY23S001

IonCode_0115_Auto_HID_Y-SNP_Set_1_and_2_18.05.2023_directorate-of-forensic-servic_86

Library type: AmpliSeq DNA
Reference: hg19 (DNA)
Target regions: ytree.20191001.designed
Read filters: Sample tracking



Number of mapped reads	1,022,847
Percent reads on target	75.34%
Percent sample tracking reads	0.00%
Average base coverage depth	1,303
Uniformity of base coverage	97.02%

Amplicon Read Coverage		Target Base Coverage	
Number of amplicons	602	Bases in target regions	47,438
Percent assigned amplicon reads	75.34%	Percent base reads on target	63.68%
Average reads per amplicon	1,280	Average base coverage depth	1,303
Uniformity of amplicon coverage	98.84%	Uniformity of base coverage	97.02%
Amplicons with at least 1 read	100.00%	Target base coverage at 1x	100.00%
Amplicons with at least 20 reads	100.00%	Target base coverage at 20x	99.78%
Amplicons with at least 100 reads	99.67%	Target base coverage at 100x	99.27%
Amplicons with at least 500 reads	92.52%	Target base coverage at 500x	90.56%
Amplicons with no strand bias	91.53%	Target bases with no strand bias	87.98%
Amplicons reading end-to-end	93.85%	Percent end-to-end reads	93.52%
Amplicon base composition bias	0.167		

Sample Name: BUDY23S002

IonCode_0116_Auto_HID_Y-SNP_Set_1_and_2_18.05.2023_directorate-of-forensic-servic_86

Library type: AmpliSeq DNA
Reference: hg19 (DNA)
Target regions: ytree.20191001.designed
Read filters: Sample tracking



Number of mapped reads	885,554
Percent reads on target	73.27%
Percent sample tracking reads	0.00%
Average base coverage depth	1,081
Uniformity of base coverage	96.69%

Amplicon Read Coverage		Target Base Coverage	
Number of amplicons	602	Bases in target regions	47,438
Percent assigned amplicon reads	73.27%	Percent base reads on target	61.40%
Average reads per amplicon	1,078	Average base coverage depth	1,081
Uniformity of amplicon coverage	98.17%	Uniformity of base coverage	96.69%
Amplicons with at least 1 read	100.00%	Target base coverage at 1x	100.00%
Amplicons with at least 20 reads	100.00%	Target base coverage at 20x	99.82%
Amplicons with at least 100 reads	99.67%	Target base coverage at 100x	99.45%
Amplicons with at least 500 reads	84.88%	Target base coverage at 500x	83.34%
Amplicons with no strand bias	91.86%	Target bases with no strand bias	88.17%
Amplicons reading end-to-end	93.36%	Percent end-to-end reads	92.36%
Amplicon base composition bias	0.195		

Sample Name: KAPY23S003

IonCode_0117_Auto_HID_Y-SNP_Set_3_and_4_18.05.2023_directorate-of-forensic-servic_88

Library type: AmpliSeq DNA
Reference: hg19 (DNA)
Target regions: ytree.20191001.designed
Read filters: Sample tracking



Number of mapped reads	1,185,429
Percent reads on target	71.54%
Percent sample tracking reads	0.00%
Average base coverage depth	1,416
Uniformity of base coverage	95.17%

Amplicon Read Coverage		Target Base Coverage	
Number of amplicons	602	Bases in target regions	47,438
Percent assigned amplicon reads	71.54%	Percent base reads on target	58.96%
Average reads per amplicon	1,409	Average base coverage depth	1,416
Uniformity of amplicon coverage	97.51%	Uniformity of base coverage	95.17%
Amplicons with at least 1 read	100.00%	Target base coverage at 1x	99.93%
Amplicons with at least 20 reads	99.67%	Target base coverage at 20x	99.29%
Amplicons with at least 100 reads	99.34%	Target base coverage at 100x	98.60%
Amplicons with at least 500 reads	94.19%	Target base coverage at 500x	91.37%
Amplicons with no strand bias	91.86%	Target bases with no strand bias	88.35%
Amplicons reading end-to-end	92.19%	Percent end-to-end reads	90.89%
Amplicon base composition bias	0.159		

Sample Name: BUKY23S003

IonCode_0118_Auto_HID_Y-SNP_Set_3_and_4_18.05.2023_directorate-of-forensic-servic_88

Library type: AmpliSeq DNA
Reference: hg19 (DNA)
Target regions: ytree.20191001.designed
Read filters: Sample tracking



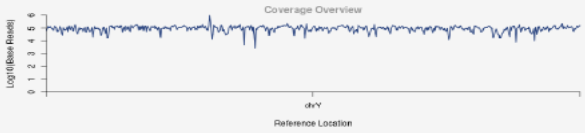
Number of mapped reads	896,905
Percent reads on target	73.55%
Percent sample tracking reads	0.00%
Average base coverage depth	1,028
Uniformity of base coverage	92.41%

Amplicon Read Coverage		Target Base Coverage	
Number of amplicons	602	Bases in target regions	47,438
Percent assigned amplicon reads	73.55%	Percent base reads on target	61.20%
Average reads per amplicon	1,096	Average base coverage depth	1,028
Uniformity of amplicon coverage	95.65%	Uniformity of base coverage	92.41%
Amplicons with at least 1 read	100.00%	Target base coverage at 1x	100.00%
Amplicons with at least 20 reads	100.00%	Target base coverage at 20x	99.74%
Amplicons with at least 100 reads	99.34%	Target base coverage at 100x	97.95%
Amplicons with at least 500 reads	88.04%	Target base coverage at 500x	82.72%
Amplicons with no strand bias	72.43%	Target bases with no strand bias	68.00%
Amplicons reading end-to-end	91.03%	Percent end-to-end reads	90.26%
Amplicon base composition bias	0.415		

Sample Name: KWNY23S003

IonCode_0119_Auto_HID_Y-SNP_Set_3_and_4_18.05.2023_directorate-of-forensic-servic_88

Library type: AmpliSeq DNA
Reference: hg19 (DNA)
Target regions: ytree.20191001.designed
Read filters: Sample tracking



Number of mapped reads	1,082,223
Percent reads on target	73.29%
Percent sample tracking reads	0.00%
Average base coverage depth	1,261
Uniformity of base coverage	95.15%

Amplicon Read Coverage		Target Base Coverage	
Number of amplicons	602	Bases in target regions	47,438
Percent assigned amplicon reads	73.29%	Percent base reads on target	61.01%
Average reads per amplicon	1,318	Average base coverage depth	1,261
Uniformity of amplicon coverage	97.09%	Uniformity of base coverage	95.15%
Amplicons with at least 1 read	99.67%	Target base coverage at 1x	99.73%
Amplicons with at least 20 reads	99.50%	Target base coverage at 20x	99.29%
Amplicons with at least 100 reads	98.84%	Target base coverage at 100x	98.41%
Amplicons with at least 500 reads	92.86%	Target base coverage at 500x	88.71%
Amplicons with no strand bias	91.20%	Target bases with no strand bias	87.09%
Amplicons reading end-to-end	91.86%	Percent end-to-end reads	91.61%
Amplicon base composition bias	0.311		

Sample Name: MBLY23S003

IonCode_0120_Auto_HID_Y-SNP_Set_3_and_4_18.05.2023_directorate-of-forensic-servic_88

Library type: AmpliSeq DNA
Reference: hg19 (DNA)
Target regions: ytree.20191001.designed
Read filters: Sample tracking



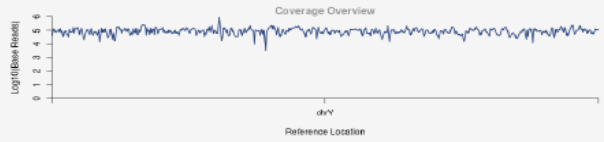
Number of mapped reads	1,010,029
Percent reads on target	73.17%
Percent sample tracking reads	0.00%
Average base coverage depth	1,233
Uniformity of base coverage	96.17%

Amplicon Read Coverage		Target Base Coverage	
Number of amplicons	602	Bases in target regions	47,438
Percent assigned amplicon reads	73.17%	Percent base reads on target	60.88%
Average reads per amplicon	1,228	Average base coverage depth	1,233
Uniformity of amplicon coverage	97.83%	Uniformity of base coverage	96.17%
Amplicons with at least 1 read	100.00%	Target base coverage at 1x	100.00%
Amplicons with at least 20 reads	100.00%	Target base coverage at 20x	99.77%
Amplicons with at least 100 reads	99.34%	Target base coverage at 100x	99.02%
Amplicons with at least 500 reads	91.03%	Target base coverage at 500x	88.55%
Amplicons with no strand bias	91.20%	Target bases with no strand bias	87.83%
Amplicons reading end-to-end	92.52%	Percent end-to-end reads	91.61%
Amplicon base composition bias	0.292		

Sample Name: BULY23S003

IonCode_0121_Auto_HID_Y-SNP_Set_3_and_4_18.05.2023_directorate-of-forensic-servic_88

Library type: AmpliSeq DNA
Reference: hg19 (DNA)
Target regions: ytree.20191001.designed
Read filters: Sample tracking



Number of mapped reads	977,061
Percent reads on target	69.78%
Percent sample tracking reads	0.00%
Average base coverage depth	1,180
Uniformity of base coverage	96.61%

Amplicon Read Coverage		Target Base Coverage	
Number of amplicons	602	Bases in target regions	47,438
Percent assigned amplicon reads	69.78%	Percent base reads on target	56.84%
Average reads per amplicon	1,133	Average base coverage depth	1,180
Uniformity of amplicon coverage	98.17%	Uniformity of base coverage	96.61%
Amplicons with at least 1 read	100.00%	Target base coverage at 1x	100.00%
Amplicons with at least 20 reads	99.83%	Target base coverage at 20x	99.85%
Amplicons with at least 100 reads	99.50%	Target base coverage at 100x	99.32%
Amplicons with at least 500 reads	86.05%	Target base coverage at 500x	85.96%
Amplicons with no strand bias	91.20%	Target bases with no strand bias	87.18%
Amplicons reading end-to-end	92.69%	Percent end-to-end reads	91.32%
Amplicon base composition bias	0.147		

Sample Name: MANY23S003

IonCode_0122_Auto_HID_Y-SNP_Set_3_and_4_18.05.2023_directorate-of-forensic-servic_88

Library type: AmpliSeq DNA
Reference: hg19 (DNA)
Target regions: ytree.20191001.designed
Read filters: Sample tracking



Number of mapped reads	788,638
Percent reads on target	73.24%
Percent sample tracking reads	0.00%
Average base coverage depth	1,018
Uniformity of base coverage	96.69%

Amplicon Read Coverage		Target Base Coverage	
Number of amplicons	602	Bases in target regions	47,438
Percent assigned amplicon reads	73.24%	Percent base reads on target	60.60%
Average reads per amplicon	959.5	Average base coverage depth	1,018
Uniformity of amplicon coverage	98.19%	Uniformity of base coverage	96.69%
Amplicons with at least 1 read	100.00%	Target base coverage at 1x	100.00%
Amplicons with at least 20 reads	100.00%	Target base coverage at 20x	99.77%
Amplicons with at least 100 reads	99.50%	Target base coverage at 100x	99.19%
Amplicons with at least 500 reads	79.07%	Target base coverage at 500x	81.21%
Amplicons with no strand bias	91.69%	Target bases with no strand bias	88.44%
Amplicons reading end-to-end	92.69%	Percent end-to-end reads	91.14%
Amplicon base composition bias	0.151		

Sample Name: NAMY23S002

IonCode_0123_Auto_HID_Y-SNP_Set_3_and_4_18.05.2023_directorate-of-forensic-servic_88

Library type: AmpliSeq DNA
Reference: hg19 (DNA)
Target regions: ytree.20191001.designed
Read filters: Sample tracking



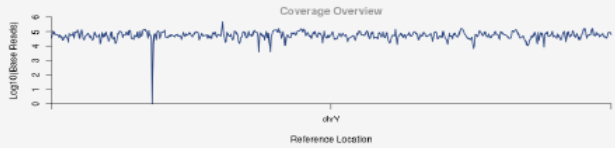
Number of mapped reads	946,327
Percent reads on target	72.36%
Percent sample tracking reads	0.00%
Average base coverage depth	1,087
Uniformity of base coverage	94.45%

Amplicon Read Coverage		Target Base Coverage	
Number of amplicons	602	Bases in target regions	47,438
Percent assigned amplicon reads	72.36%	Percent base reads on target	60.28%
Average reads per amplicon	1,138	Average base coverage depth	1,087
Uniformity of amplicon coverage	96.51%	Uniformity of base coverage	94.45%
Amplicons with at least 1 read	100.00%	Target base coverage at 1x	100.00%
Amplicons with at least 20 reads	99.83%	Target base coverage at 20x	99.52%
Amplicons with at least 100 reads	99.34%	Target base coverage at 100x	98.36%
Amplicons with at least 500 reads	89.04%	Target base coverage at 500x	84.77%
Amplicons with no strand bias	92.03%	Target bases with no strand bias	88.69%
Amplicons reading end-to-end	91.86%	Percent end-to-end reads	91.25%
Amplicon base composition bias	0.383		

Sample Name: BUDY23S003

IonCode_0124_Auto_HID_Y-SNP_Set_3_and_4_18.05.2023_directorate-of-forensic-servic_88

Library type: AmpliSeq DNA
Reference: hg19 (DNA)
Target regions: ytree.20191001.designed
Read filters: Sample tracking



Number of mapped reads	628,873
Percent reads on target	73.39%
Percent sample tracking reads	0.00%
Average base coverage depth	797.1
Uniformity of base coverage	96.60%

Amplicon Read Coverage		Target Base Coverage	
Number of amplicons	602	Bases in target regions	47,438
Percent assigned amplicon reads	73.39%	Percent base reads on target	60.45%
Average reads per amplicon	766.6	Average base coverage depth	797.1
Uniformity of amplicon coverage	98.01%	Uniformity of base coverage	96.60%
Amplicons with at least 1 read	99.83%	Target base coverage at 1x	99.76%
Amplicons with at least 20 reads	99.83%	Target base coverage at 20x	99.58%
Amplicons with at least 100 reads	99.00%	Target base coverage at 100x	98.55%
Amplicons with at least 500 reads	69.93%	Target base coverage at 500x	71.92%
Amplicons with no strand bias	89.70%	Target bases with no strand bias	84.76%
Amplicons reading end-to-end	94.35%	Percent end-to-end reads	91.50%
Amplicon base composition bias	0.343		

Sample Name: KAPY23S004

IonCode_0125_Auto_HID_Y-SNP_Set_3_and_4_18.05.2023_directorate-of-forensic-servic_88

Library type: AmpliSeq DNA
Reference: hg19 (DNA)
Target regions: ytree.20191001.designed
Read filters: Sample tracking



Number of mapped reads	873,724
Percent reads on target	71.34%
Percent sample tracking reads	0.00%
Average base coverage depth	1,031
Uniformity of base coverage	96.44%

Amplicon Read Coverage		Target Base Coverage	
Number of amplicons	602	Bases in target regions	47,438
Percent assigned amplicon reads	71.34%	Percent base reads on target	58.81%
Average reads per amplicon	1,035	Average base coverage depth	1,031
Uniformity of amplicon coverage	97.67%	Uniformity of base coverage	96.44%
Amplicons with at least 1 read	100.00%	Target base coverage at 1x	100.00%
Amplicons with at least 20 reads	99.83%	Target base coverage at 20x	99.56%
Amplicons with at least 100 reads	99.17%	Target base coverage at 100x	98.81%
Amplicons with at least 500 reads	85.88%	Target base coverage at 500x	84.57%
Amplicons with no strand bias	92.86%	Target bases with no strand bias	89.09%
Amplicons reading end-to-end	92.86%	Percent end-to-end reads	90.09%
Amplicon base composition bias	0.229		

Sample Name: BUKY23S004

IonCode_0126_Auto_HID_Y-SNP_Set_3_and_4_18.05.2023_directorate-of-forensic-servic_88

Library type: AmpliSeq DNA
Reference: hg19 (DNA)
Target regions: ytree.20191001.designed
Read filters: Sample tracking



Number of mapped reads	1,037,811
Percent reads on target	71.90%
Percent sample tracking reads	0.00%
Average base coverage depth	1,140
Uniformity of base coverage	95.75%

Amplicon Read Coverage		Target Base Coverage	
Number of amplicons	602	Bases in target regions	47,438
Percent assigned amplicon reads	71.90%	Percent base reads on target	59.34%
Average reads per amplicon	1,239	Average base coverage depth	1,140
Uniformity of amplicon coverage	97.34%	Uniformity of base coverage	95.75%
Amplicons with at least 1 read	99.67%	Target base coverage at 1x	99.46%
Amplicons with at least 20 reads	99.67%	Target base coverage at 20x	99.23%
Amplicons with at least 100 reads	99.00%	Target base coverage at 100x	98.48%
Amplicons with at least 500 reads	93.69%	Target base coverage at 500x	88.12%
Amplicons with no strand bias	91.36%	Target bases with no strand bias	87.23%
Amplicons reading end-to-end	92.52%	Percent end-to-end reads	90.72%
Amplicon base composition bias	0.398		

Sample Name: KWNY23S004

IonCode_0127_Auto_HID_Y-SNP_Set_3_and_4_18.05.2023_directorate-of-forensic-servic_88

Library type: AmpliSeq DNA
Reference: hg19 (DNA)
Target regions: ytree.20191001.designed
Read filters: Sample tracking



Number of mapped reads	1,000,273
Percent reads on target	74.42%
Percent sample tracking reads	0.00%
Average base coverage depth	1,246
Uniformity of base coverage	95.85%

Amplicon Read Coverage		Target Base Coverage	
Number of amplicons	602	Bases in target regions	47,438
Percent assigned amplicon reads	74.42%	Percent base reads on target	62.36%
Average reads per amplicon	1,237	Average base coverage depth	1,246
Uniformity of amplicon coverage	97.84%	Uniformity of base coverage	95.85%
Amplicons with at least 1 read	100.00%	Target base coverage at 1x	100.00%
Amplicons with at least 20 reads	100.00%	Target base coverage at 20x	99.78%
Amplicons with at least 100 reads	99.67%	Target base coverage at 100x	99.45%
Amplicons with at least 500 reads	90.53%	Target base coverage at 500x	87.96%
Amplicons with no strand bias	90.20%	Target bases with no strand bias	85.91%
Amplicons reading end-to-end	93.19%	Percent end-to-end reads	91.54%
Amplicon base composition bias	0.305		

Sample Name: MBLY23S004

IonCode_0128_Auto_HID_Y-SNP_Set_3_and_4_18.05.2023_directorate-of-forensic-servic_88

Library type: AmpliSeq DNA
Reference: hg19 (DNA)
Target regions: ytree.20191001.designed
Read filters: Sample tracking



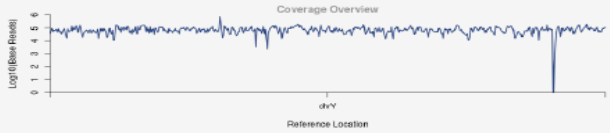
Number of mapped reads	534,736
Percent reads on target	12.32%
Percent sample tracking reads	0.00%
Average base coverage depth	58.99
Uniformity of base coverage	24.37%

Amplicon Read Coverage		Target Base Coverage	
Number of amplicons	602	Bases in target regions	47,438
Percent assigned amplicon reads	12.32%	Percent base reads on target	4.789%
Average reads per amplicon	109.4	Average base coverage depth	58.99
Uniformity of amplicon coverage	23.58%	Uniformity of base coverage	24.37%
Amplicons with at least 1 read	59.97%	Target base coverage at 1x	33.49%
Amplicons with at least 20 reads	26.25%	Target base coverage at 20x	19.61%
Amplicons with at least 100 reads	3.82%	Target base coverage at 100x	2.06%
Amplicons with at least 500 reads	1.83%	Target base coverage at 500x	0.77%
Amplicons with no strand bias	90.53%	Target bases with no strand bias	95.62%
Amplicons reading end-to-end	28.24%	Percent end-to-end reads	62.36%
Amplicon base composition bias	0.036		

Sample Name: SRKY23S003

IonCode_0129_Auto_HID_Y-SNP_Set_3_and_4_18.05.2023_directorate-of-forensic-servic_88

Library type: AmpliSeq DNA
Reference: hg19 (DNA)
Target regions: ytree.20191001.designed
Read filters: Sample tracking



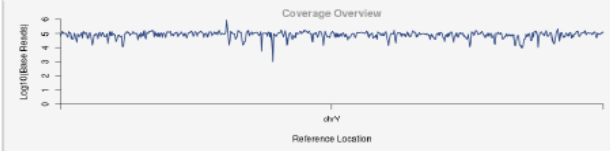
Number of mapped reads	755,518
Percent reads on target	73.75%
Percent sample tracking reads	0.00%
Average base coverage depth	950.3
Uniformity of base coverage	96.17%

Amplicon Read Coverage		Target Base Coverage	
Number of amplicons	602	Bases in target regions	47,438
Percent assigned amplicon reads	73.75%	Percent base reads on target	61.71%
Average reads per amplicon	925.6	Average base coverage depth	950.3
Uniformity of amplicon coverage	97.84%	Uniformity of base coverage	96.17%
Amplicons with at least 1 read	99.83%	Target base coverage at 1x	99.65%
Amplicons with at least 20 reads	99.67%	Target base coverage at 20x	99.27%
Amplicons with at least 100 reads	99.17%	Target base coverage at 100x	98.47%
Amplicons with at least 500 reads	77.74%	Target base coverage at 500x	79.14%
Amplicons with no strand bias	92.19%	Target bases with no strand bias	88.48%
Amplicons reading end-to-end	91.36%	Percent end-to-end reads	90.51%
Amplicon base composition bias	0.142		

Sample Name: MANY23S004

IonCode_0130_Auto_HID_Y-SNP_Set_3_and_4_18.05.2023_directorate-of-forensic-servic_88

Library type: AmpliSeq DNA
Reference: hg19 (DNA)
Target regions: ytree.20191001.designed
Read filters: Sample tracking



Number of mapped reads	990,958
Percent reads on target	71.95%
Percent sample tracking reads	0.00%
Average base coverage depth	1,128
Uniformity of base coverage	94.04%

Amplicon Read Coverage		Target Base Coverage	
Number of amplicons	602	Bases in target regions	47,438
Percent assigned amplicon reads	71.95%	Percent base reads on target	60.55%
Average reads per amplicon	1,184	Average base coverage depth	1,128
Uniformity of amplicon coverage	96.70%	Uniformity of base coverage	94.04%
Amplicons with at least 1 read	100.00%	Target base coverage at 1x	100.00%
Amplicons with at least 20 reads	100.00%	Target base coverage at 20x	99.50%
Amplicons with at least 100 reads	99.50%	Target base coverage at 100x	99.21%
Amplicons with at least 500 reads	91.86%	Target base coverage at 500x	87.44%
Amplicons with no strand bias	92.36%	Target bases with no strand bias	88.59%
Amplicons reading end-to-end	92.69%	Percent end-to-end reads	92.29%
Amplicon base composition bias	0.461		

Sample Name: NAMY23S003

IonCode_0131_Auto_HID_Y-SNP_Set_3_and_4_18.05.2023_directorate-of-forensic-servic_88

Library type: AmpliSeq DNA
Reference: hg19 (DNA)
Target regions: ytree.20191001.designed
Read filters: Sample tracking



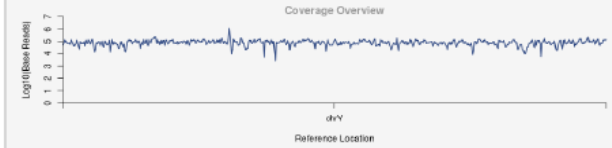
Number of mapped reads	975,109
Percent reads on target	73.16%
Percent sample tracking reads	0.00%
Average base coverage depth	1,264
Uniformity of base coverage	96.66%

Amplicon Read Coverage		Target Base Coverage	
Number of amplicons	602	Bases in target regions	47,438
Percent assigned amplicon reads	73.16%	Percent base reads on target	60.81%
Average reads per amplicon	1,185	Average base coverage depth	1,264
Uniformity of amplicon coverage	98.17%	Uniformity of base coverage	96.66%
Amplicons with at least 1 read	100.00%	Target base coverage at 1x	100.00%
Amplicons with at least 20 reads	100.00%	Target base coverage at 20x	99.96%
Amplicons with at least 100 reads	99.50%	Target base coverage at 100x	99.03%
Amplicons with at least 500 reads	85.71%	Target base coverage at 500x	85.75%
Amplicons with no strand bias	91.53%	Target bases with no strand bias	87.87%
Amplicons reading end-to-end	93.19%	Percent end-to-end reads	91.69%
Amplicon base composition bias	0.096		

Sample Name: BULY23S004

IonCode_0132_Auto_HID_Y-SNP_Set_3_and_4_18.05.2023_directorate-of-forensic-servic_88

Library type: AmpliSeq DNA
Reference: hg19 (DNA)
Target regions: ytree.20191001.designed
Read filters: Sample tracking



Number of mapped reads	924,328
Percent reads on target	72.32%
Percent sample tracking reads	0.00%
Average base coverage depth	1,077
Uniformity of base coverage	95.82%

Amplicon Read Coverage		Target Base Coverage	
Number of amplicons	602	Bases in target regions	47,438
Percent assigned amplicon reads	72.32%	Percent base reads on target	59.97%
Average reads per amplicon	1,110	Average base coverage depth	1,077
Uniformity of amplicon coverage	97.67%	Uniformity of base coverage	95.82%
Amplicons with at least 1 read	100.00%	Target base coverage at 1x	100.00%
Amplicons with at least 20 reads	100.00%	Target base coverage at 20x	99.77%
Amplicons with at least 100 reads	99.34%	Target base coverage at 100x	98.71%
Amplicons with at least 500 reads	89.53%	Target base coverage at 500x	85.39%
Amplicons with no strand bias	91.36%	Target bases with no strand bias	87.74%
Amplicons reading end-to-end	92.86%	Percent end-to-end reads	92.42%
Amplicon base composition bias	0.338		