U.PORTO

2.º CICLO DE ESTUDOS
CIÊNCIAS FORENSES

Population Data from Cape Verde: Allele Frequencies of the 26 Most Commonly Used Y-STR Loci for Forensic and Medico-Legal Purposes

Rita Gabriela Martins Gonçalves Costa

M

2025

SEDE ADMINISTRATIVA

FACULDADE DE **MEDICINA**

FACULDADE DE **CIÊNCIAS**

FACULDADE DE **DIREITO**

FACULDADE DE **FARMÁCIA**

FACULDADE DE **MEDICINA DENTÁRIA**

FACULDADE DE **PSICOLOGIA E DE CIÊNCIAS DA EDUCAÇÃO**

INSTITUTO DE **CIÊNCIAS BIOMÉDICAS ABEL SALAZAR**

Population Data from Cape Verde: Allele Frequencies of the 26 Most Commonly Used Y-STR Loci for Forensic and Medico-Legal Purposes

Rita Gabriela Martins Gonçalves Costa

Dissertação de Mestrado em Ciências Forenses
Ano: 2024/2025
Orientador/a: Prof. Doutora Laura Cainé
Co-Orientador/a: Doutora Jennifer Fadoni

Agradecimentos

Gostaria de expressar a minha profunda gratidão à Professora Doutora Laura Cainé, pela orientação dedicada, pelas sugestões e por todo o apoio ao longo do desenvolvimento deste trabalho. À Doutora Jennifer Fadoni, pela incansável dedicação, pela disponibilidade constante e pela paciência com que me acompanhou em cada etapa, oferecendo sempre palavras de incentivo e contributos fundamentais para a concretização deste trabalho.

À Faculdade de Medicina da Universidade do Porto, pela oportunidade de frequentar este mestrado. Ao Serviço de Genética e Biologia Forenses do Instituto Nacional de Medicina Legal e Ciências Forenses – Delegação Norte, deixo o meu mais profundo reconhecimento pela oportunidade concedida para desenvolver este trabalho, pela disponibilização dos recursos necessários e pelo acolhimento num ambiente ciêntifico de excelência, Agradeço igualmente a todas as pessoas do SGBF-N, pelo apoio, pela disponibilidade e pela partilha de conhecimentos, que tanto contribuíram para a realização deste percurso.

Às minhas amigas do mestrado, Catarina, Juliana e Maria, agradeço profundamente pela amizade genuína, pela partilha de experiências e por todos os momentos de apoio, alegria e motivação que tornaram este percurso único e inesquecível. À Bárbara, que esteve ao meu lado diariamente no laboratório, deixo um agradecimento muito especial pela companhia, pela paciência e pela entreajuda constante, que fizeram toda a diferença nos desafios do dia a dia e transformaram este caminho numa experiência enriquecedora e gratificante.

Aos meus pais, agradeço com todo o meu coração por todo o amor, apoio e confiança incondicional. Vocês foram e serão sempre o meu maior exemplo de força e dedicação. Esta conquista também é vossa. Sem vocês nunca teria sido possível.

À minha irmã, obrigada pelo teu amor genuíno e pelo coração bonito que tanto me inspira. Obrigada por cada abraço apertado e por cada palavra bonita repetida vezes sem conta. Mesmo sem o perceberes, foste o meu maior apoio e a minha maior força.

Aos meus avós, obrigada por todo o carinho, pela presença constante e pelo amor com que sempre me acompanharam. Cada gesto vosso foi um apoio inestimável e uma inspiração ao longo deste caminho.

À minha avó Lina, à minha estrelinha, que mesmo lá de cima estará sempre a olhar por mim e a celebrar eternamente o meu sucesso.

Aos meus amigos e amigas que, desde o início, me acompanharam neste percurso, agradeço pelo apoio, pela motivação e por estarem sempre presentes, mesmo nos momentos em que a distância ou o tempo podiam ser obstáculos.

Por fim, agradeço a todos os que, de forma direta ou indireta fizeram parte desta etapa. Cada gesto, palavra de incentivo e sorriso tiveram um valor imensurável e permanecerão para sempre guardados no meu coração.

A todos, expresso o meu mais profundo e sincero obrigada!

Original Publication

In the course of this dissertation, two original scientific articles focused on forensic genetics were developed, both published in the international journal *Genes*, by MDPI. This journal is indexed in the main scientific databases and is classified in the second quartile (Q2) of the *Genetics & Heredity* category, according to the Journal Citation Reports. The published works result directly from the research carried out and are incorporated into this dissertation, contributing to the advancement of scientific knowledge and to the application of genetics in forensic investigation.

- ◊ Costa, R., Fadoni, J., Amorim, A., & Cainé, L. (2025). Y-STR Databases—Application in Sexual Crimes. *Genes*, *16*(5), 484. https://doi.org/10.3390/genes16050484
- ◊ Costa, R., Fadoni, J., Amorim, A., & Cainé, L. (2025). The Genetic Structure of Cape Verdean Population Revealed by Y-Chromosome STRs. *Genes*, *16*(9), 999. https://doi.org/10.3390/genes16090999

Resumo

Os marcadores Y-STR são amplamente utilizados na genética forense devido à sua especificidade para o cromossoma Y, sendo particularmente úteis na identificação de perfis genéticos masculinos em misturas biológicas complexas, como em casos de agressão sexual. Apesar da sua utilidade, a aplicação eficaz dos Y-STR em contextos forenses depende da representatividade populacional nas bases de dados genéticas de referência.

Neste trabalho, antes da aplicação do kit Investigator® Argus Y-28 QS (QIAGEN, Hilden, Alemanha) às amostras populacionais, foi realizada uma validação interna no Serviço de Genética e Biologia Forense – Delegação Norte (INMLCF, I.P.). A comparação direta com o kit Yfiler® Plus demonstrou concordância total dos haplótipos obtidos, enquanto a calibração com material de referência certificado confirmou a correta atribuição alélica. Foram ainda definidos limiares analíticos robustos, estabelecidos através da análise de controlos negativos. Os ensaios de sensibilidade revelaram que é possível obter perfis completos e fiáveis a partir de 0,5 ng de DNA, enquanto concentrações mais elevadas conduziram ao aparecimento de artefactos. Estudos adicionais confirmaram a precisão, reprodutibilidade e tolerância do kit a condições adversas, incluindo a exposição a radiação UV, reforçando a sua aplicabilidade em contexto forense.

A presente dissertação integra dois artigos científicos publicados na revista Genes (MDPI), inseridos no contexto da genética forense e das ciências forenses. O primeiro artigo apresenta uma revisão abrangente sobre a aplicação dos Y-STR em crimes sexuais, destacando as vantagens, limitações e a relevância dos marcadores de mutação rápida (RM Y-STR) para a discriminação entre indivíduos aparentados. Sublinha ainda a necessidade de aumentar a representatividade populacional em bases como a YHRD, com enfoque em grupos sub-representados, como a população Cabo-Verdiana.

O segundo artigo consiste num estudo populacional que caracteriza a diversidade genética da população cabo-verdiana residente em Portugal com base em 26 loci Y-STR, incluindo RM Y-STR). Os resultados evidenciam uma elevada diversidade haplotípica e refletem a composição genética mista africana e europeia desta população. A inclusão dos dados na base YHRD contribui para a melhoria da robustez estatística em estudos de parentesco e identificação forense.

Em conjunto, os dois trabalhos reforçam a importância da integração de dados populacionais sub-representados nas bases de dados genéticas e da aplicação crítica dos Y-STR em contextos forenses sensíveis, como os crimes sexuais, promovendo práticas mais equitativas, rigorosas e cientificamente sustentadas na genética forense.

Abstract

Y-chromosomal short tandem repeats (Y-STRs) are widely used in forensic genetics due to their specificity to the Y chromosome, proving particularly useful for the identification of male genetic profiles in complex biological mixtures, such as those encountered in sexual assault cases. Despite their utility, the effective forensic application of Y-STRs depends on the population representativeness within genetic reference databases.

Before applying the Investigator® Argus Y-28 QS kit (QIAGEN, Hilden, Germany) to the population samples, an internal validation was performed at the Forensic Genetics and Biology Services – North Delegation (INMLCF, I.P.). Direct comparison with the Yfiler® Plus kit demonstrated full concordance of haplotypes, while calibration with certified reference material ensured accurate allele assignment. Robust analytical thresholds were established from negative controls, and sensitivity tests showed that complete and reliable profiles could be obtained from as little as 0.5 ng of DNA, whereas higher concentrations led to artefacts. Additional studies confirmed the precision, reproducibility, and tolerance of the kit to challenging conditions, including UV exposure, supporting its forensic applicability.

This dissertation includes two scientific articles published in the journal *Genes* (MDPI), within the scope of forensic and population genetics. The first article provides a comprehensive review of the application of Y-STRs in sexual crimes, highlighting their advantages, limitations, and the relevance of rapidly mutating Y-STRs (RM Y-STRs) for differentiating closely related male individuals. It also emphasizes the need to increase population representation in databases such as the Y Chromosome Haplotype Reference Database (YHRD), with particular focus on underrepresented groups, such as the Cape Verdean population.

The second article presents a population study that characterizes the genetic diversity of the Cape Verdean population residing in Portugal, based on 26 Y-STR loci, including RM Y-STRs. The results reveal a high level of haplotype diversity and reflect the admixed African and European genetic background of this population. The inclusion of these data in the YHRD enhances the statistical robustness of kinship analyses and forensic identification.

Together, these two studies reinforce the importance of integrating underrepresented population data into genetic databases and critically applying Y-STR markers in sensitive forensic contexts such as sexual crimes, thus promoting more equitable, rigorous, and scientifically grounded practices in forensic genetics.

Índice

Agradecimentos 6
Original Publication 8
Resumo 9
Abstract 10
List of Abreviations 12
CHAPTER I - Introduction 13
CHAPTER II – Material and Methods 15
Study Design 15
Review Methodology 15
Sample Collection 16
DNA Extraction 16
Y-STR Amplification and Electrophoresis 16
Statistical Analysis 17
CHAPTER III – Validation Study of the Investigator® Argus Y-28 QS Kit 18
Introduction 18
Materials and Methods 18
Results and Discussion 19
Conclusions 28
CHAPTER IV – Original Research 29
Article I – Y-STR Databases – Application in Sexual Crimes. 29
Article II – The Genetic Structure of Cape Verdean Population Revealed by Y-Chromosome STRs 42
CHAPTER V – Discussion 55
CHAPTER VI – Future Perspectives 57
CHAPTER VII – Conclusions 58

List of Abreviations

CRM	Certified Reference Material
DC	Discrimination Capacity
DNA	Deoxyribonucleic Acid
GD	Gene Diversity
HD	Haplotype Diversity
HMP	Haplotype Match Probability
INMLCF, I.P.	Instituto Nacional de Medicina Legal, Instituto Público
ISFG	International Society for Forensic Genetics
IUPAC	International Union of Pure and Applied Chemistry
MDS	Multidimensional Scaling
NGS	Next-Generation Sequencing
PCR	Polimerase Chain Reaction
QS	Quality Sensor
PD	Power of Discrimination
PIC	Polymorphism Information Content
RFU	Relative Fluorescence Units
RM Y-STRs	Rapidly Mutating Y-chromosome Short Tandem Repeats
SGBF-N	Serviço de Genética e Biologia Forenses – Delegação Norte
STR	Short Tandem Repeat
SWGDAM	Scientific Working Group on DNA Analysis Methods
UV	Utraviolet Radiation
YHRD	Y-Chromosome Haplotype Reference Database
Y-STRs	Y-chromosome Short Tandem Repeats

CHAPTER I - Introduction

Forensic genetics has become an essential tool in human identification and criminal investigations, playing a central role in the analysis of biological evidence, particularly in cases of sexual assault. Among the most commonly used markers, short tandem repeats located on the Y chromosome (Y-STRs) are key to distinguishing male genetic lineages due to their patrilineal inheritance and lack of recombination, which enable the stable transmission of haplotypes across generations [1-4].

However, this stability also limits the ability to discriminate between paternally related individuals, posing a significant challenge in forensic analysis. To address this issue, rapidly mutating Y-STRs (RM Y-STRs) have been developed, which increase the discriminatory power within paternal lineages, although they do not always allow the distinction of very closely related individuals such as consanguineous brothers. These markers exhibit mutation rates above 10^{-2} per generation, making it possible to differentiate between closely related male individuals [5, 6].

The analysis of Y-STR profiles requires comparison with robust population databases, such as the Y-Chromosome Haplotype Reference Database (YHRD), which allows for the assessment of haplotype rarity and the assignment of statistical weight to forensic findings [7]. The utility of these databases depends on their population representativeness. The underrepresentation of specific groups, such as African and Afro-descendant populations, can compromise the reliability of statistical inference and limit the forensic applicability of Y-STR results [8, 9].

The Cape Verdean population is particularly relevant in this context, with a historical background shaped by Portuguese colonization and the transatlantic slave trade, the archipelago presents an admixed genetic profile characterized by predominantly European paternal lineages and African maternal lineages [10-12].

Despite the forensic relevance of the Cape Verdean population, particularly in countries with large Cape Verdean communities such as Portugal, there remains a significant lack of population-specific genetic data in international databases. This limitation may reduce the robustness of forensic analyses in contexts such as individual identification or the resolution of sexual assault cases, where statistical reliability is crucial [8, 13].

Collecting genetic data from underrepresented populations, together with the use of extended Y-STR panels that include RM loci, is essential to improving discriminatory power and ensuring more equitable and representative forensic practices [8,14]. Increasing the diversity of genetic databases contributes to fairer interpretation of forensic results and supports a more robust scientific and social justice framework [9, 14, 15].

In this light, the present work aligns with broader scientific efforts to strengthen the accuracy, inclusiveness, and ethical integrity of forensic genetic practices. By addressing gaps in population

data and applying high-resolution Y-chromosome markers, this work aims to contribute to a more equitable and scientifically grounded application of forensic genetic.

Beyond the importance of population representativeness, forensic applications also require that the laboratory methods themselves are reliable and robust. International organizations such as the International Society for Forensic Genetics (ISFG) and Scientific Working Group on DNA Analysis Methods (SWGDAM) have established that newly introduced kits must undergo internal validation to demonstrate their accuracy, reproducibility, and suitability for forensic use. In this context, the present work also addresses the validation of the Investigator® Argus Y-28 QS kit under the specific conditions of the Forensic Genetics and Biology Services – North Delegation (SGBF-N) of the National Institute of Legal Medicine and Forensic Sciences (INMLCF, I.P.), thereby ensuring that the population data presented herein were generated using a method that had been thoroughly tested and confirmed to be forensically suitable [14, 16].

Objectives of the Dissertation

This dissertation aims to contribute to the advancement of forensic genetics through three interrelated objectives. The first is to undertake a critical review of the forensic use of Y-STRs, with particular emphasis on their application in the investigation of sexual crimes, highlighting their discriminatory potential as well as their limitations. The second objective involves generating original population data for Cape Verdean immigrant males living in Portugal who share Cape Verdean paternal ancestry, through the analysis of 26 Y-STR loci, including rapidly mutating markers.

This work seeks to improve the representativeness of international reference databases such as the YHRD and to strengthen the statistical reliability of forensic analyses involving underrepresented populations. The third objective is to perform the internal validation of the Investigator® Argus Y-28 QS kit under the specific conditions of the SGBF-N laboratory of the INMLCF, I.P., in accordance with international guidelines (ISFG and SWGDAM). The validation was conducted to confirm that the methodology applied to the Cape Verdean samples followed robust, reproducible, and forensically reliable procedures, supporting the generation of scientifically robust population data.

CHAPTER II – Material and Methods

Study Design

This dissertation is structured as a thesis by publication and comprises two scientific articles that address complementary aspects of forensic genetics. The first article is a systematic review focused on the application of Y-chromosome STR markers in the investigation of sexual crimes, based on the analysis and integration of published literature. The second article is an original research study that presents new population data on Cape Verdean males using Y-STR analysis, including sample collection, laboratory procedures, and statistical interpretation. Despite the differences in methodology, both studies contribute to the advancement of forensic genetics by highlighting the value and limitations of Y-STRs in forensic contexts. The methods described in the following sections refer specifically to the original research article.

Review Methodology

The first scientific article included in this dissertation is a narrative literature review on the forensic applications of Y-chromosome short tandem repeats (Y-STRs), particularly in sexual crimes. This review aimed to synthesize current knowledge on Y-STR utility, limitations, and population database coverage, with a special focus on underrepresented populations such as the Cape Verdean community.

A comprehensive literature search was conducted in November 2024 using the PubMed database (U.S. National Library of Medicine), without restrictions on publication date. The following keywords were used: *"Y-STR", "Y chromosome", "forensic genetics", "DNA mixtures", "sexual crimes", "rapidly mutating markers"*, and *"Y-STR databases"*. Inclusion criteria comprised peer-reviewed articles that addressed forensic applications of Y-STRs, particularly in sexual crime contexts, with emphasis on DNA mixture analysis, population databases, and recent technological advancements (e.g., rapidly mutating Y-STRs). Articles addressing ethical, legal, and methodological issues were also included.

Exclusion criteria involved duplicates, articles lacking full-text access, or those not directly related to forensic applications (e.g., evolutionary or clinical genetics).

Additional data were obtained from official reports and statistics from the National Institute of Statistics of Cape Verde (INE), EUROSTAT, UNODC, and the Annual Internal Security Report (RASI).

Sample Collection

A total of 143 DNA samples were obtained from unrelated male individuals of self-declared Cape Verdean paternal ancestry. These samples were collected from paternity cases at the Instituto Nacional de Medicina Legal e Ciências Forenses, Instituto Público (INMLCF, I.P.). Among them, 135 were dried blood spots on FTA® cards, and 8 were buccal swabs collected using sterile swabbing procedures. The inclusion criteria ensured the absence of known paternal relationships between participants. All samples were anonymized prior to analysis, with only the information regarding paternal ancestry being retained.
The study received ethical approval from the Ethics Committee of the INMLCF, I.P. (CE-25/2024).

DNA Extraction

Buccal swab samples were extracted using the PrepFiler Express™ Forensic DNA Extraction Kit (Thermo Fisher Scientific, MA, USA), following the manufacturer's instructions. For blood samples, direct amplification was carried out from the FTA® cards, bypassing the DNA extraction step. Both sample types were subsequently prepared for amplification and STR analysis.

Y-STR Amplification and Electrophoresis

The Investigator® Argus Y-28 QS PCR kit (QIAGEN, Hilden, Germany) was used to amplify 26 Y-STR loci, including 20 standard and 6 rapidly mutating markers. Amplifications were performed on the GeneAmp® PCR System 9700 thermal cycler (Thermo Fisher Scientific, MA, USA) with modified conditions to reduce reagent usage and avoid overamplification: reaction volumes were halved (12.5 µL), and the cycle number was reduced to 28. Each PCR run included a positive control (DNA 9948) and a negative control (ddH_2O).

PCR products were prepared for capillary electrophoresis by mixing 1 µL of PCR product with 12 µL of Hi-Di™ Formamide and 0.5 µL of DNA Size Standard 24plex (BTO). The mixture was denatured at 95 °C for 3 minutes and cooled to 4 °C before analysis. Fragment separation was performed on an ABI 3500 Genetic Analyzer. Allele calling and data interpretation were conducted using GeneMapper ID-X software (version 1.6), with reference to the manufacturer's allelic ladder.

Statistical Analysis

Allele and haplotype frequencies were calculated by direct counting. Haplotype match probability (HMP), discrimination capacity (DC), and haplotype diversity (HD) were computed using Microsoft Excel and the STRAF software (version 2.2.2). Gene diversity (GD), power of discrimination (PD), and polymorphism information content (PIC) were estimated for each individual locus to assess their informativeness and forensic utility. To evaluate genetic differentiation, pairwise Rst distances between the Cape Verdean sample and selected reference populations were calculated using Arlequin software (version 3.5.2.2). These distances were then visualized through multidimensional scaling (MDS) using R software (version 4.0), providing a graphical representation of interpopulation relationships. For these analyses, DYS389II values were adjusted by subtracting DYS389I to ensure proper alignment with published datasets. Additionally, Y-chromosome haplogroups were predicted based on STR haplotypes using the NevGen Y-DNA Haplogroup Predictor. Only assignments with a prediction probability greater than 50% were considered. The haplogroup distribution was analyzed to infer the predominant paternal lineages within the sample and to explore historical and population-level patterns.

CHAPTER III – Validation Study of the Investigator® Argus Y-28 QS Kit

Introduction

Validation is a fundamental requirement in forensic genetics to ensure that any analytical method performs reliably and consistently under the specific conditions of the laboratory in which it is implemented. International guidelines, particularly those from the ISFG and the GHEP-ISFG, emphasize that newly introduced kits must undergo internal validation before being applied to forensic casework or population studies [14, 16]. This process is essential to verify the robustness, accuracy, and reproducibility of the method, as well as its suitability for the intended applications.

In this dissertation, the Investigator® Argus Y-28 QS kit (QIAGEN, Hilden, Germany), designed to amplify 26 Y-chromosomal short tandem repeat (Y-STR) loci, was subjected to a comprehensive internal validation at the Forensic Genetics and Biology Services, North Delegation (SGBF-N), National Institute of Legal Medicine and Forensic Sciences (INMLCF, I.P.). This validation was performed prior to the analysis of the Cape Verdean samples presented in Article II, thereby ensuring that the population data generated in this study were obtained with a method previously demonstrated to be reliable and forensically suitable.

Materials and Methods

The validation of the Investigator® Argus Y-28 QS kit was performed in two stages: direct evaluation, followed by indirect evaluation. Direct evaluation was carried out first, since it involved the comparison of the kit with previously validated methods and calibration with certified reference materials (CRM). This step also included the optimization of reaction conditions, namely by reducing the PCR reaction volume to half of that recommended by the manufacturer and by lowering the number of amplification cycles from 30 to 28. These adjustments were adopted as the standard protocol for the subsequent indirect evaluation. The direct evaluation aimed to verify the reliability of the kit by comparing its performance with the Yfiler® Plus multiplex, a well-established commercial kit for Y-STR analysis. In addition, calibration experiments were conducted using known standards and CRM to confirm the accuracy of allele calling and fragment sizing. The outcomes of this phase ensured that the kit showed full concordance with internationally accepted forensic DNA typing practices and that the optimized protocol could be reliably applied to further testing.

The indirect evaluation focused on experimental parameters tested under laboratory conditions, using the optimized protocol defined during the direct validation. In accordance with the requirements of the INMLCF Quality Management System (SGQ), each parameter was selected because it assesses a specific performance characteristic essential for forensic DNA analysis. Specificity, which verifies that the kit amplifies only human male DNA, was assessed by testing DNA from several non-human species (horse, pig, dog, and cat), female samples, and standard controls (9947A and 9948) to confirm selective amplification. Analytical thresholds, defined to distinguish true allelic peaks from background noise, were established by analyzing negative controls to determine the minimum fluorescence level (RFU) required for reliable interpretation. Repeatability, which evaluates consistency under identical experimental conditions, was tested by amplifying male DNA samples in triplicate, while reproducibility, assessing performance under varying operators, thermal cycler, and amplification days, was examined by repeating the experiments under different laboratory conditions. Sensitivity was investigated through a dilution series of male DNA, ranging from 0.5 ng to 0.01 ng, as well as with higher DNA inputs up to 30 ng, to evaluate performance with both low-template and high-template DNA, assessing the lowest and highest DNA quantities at which the kit can still produce reliable and interpretable profiles. Additional experiments included the evaluation of precision, contamination and mixtures, and the effect of inhibitors. Precision was examined using allelic ladders to confirm accurate fragment sizing across electrophoretic runs. Mixtures of male DNA with varying ratios of contributors were analyzed to test the ability of the kit to detect minor contributors. The robustness of the kit under suboptimal conditions was assessed by adding known inhibitory substances to PCR reactions.

Results and Discussion

- **Direct Evaluation**

Comparison with Yfiler® Plus

The Investigator® Argus Y-28 QS kit was compared with Yfiler® Plus, a widely used and previously validated commercial multiplex for Y-STR analysis. A set of 28 male DNA samples was amplified with both systems, and complete concordance was observed between the haplotypes generated. The identical results confirmed that the Argus Y-28 QS kit provides profiles consistent with those obtained using an established methodology, supporting its reliability with existing forensic databases.

During this stage, the amplification protocol was also optimized by reducing the PCR reaction volume to half of the manufacturer's recommendation and lowering the number of

amplification cycles from 30 to 28. These optimized conditions maintained high-quality profiles and were subsequently adopted as the standard protocol for the indirect evaluation. This complete concordance highlights the robustness of the Argus Y-28 QS kit and demonstrates that it performs in line with previously validated methodologies. The agreement with Yfiler® Plus, which has been widely used in forensic laboratories worldwide, confirms that haplotypes generated with Argus Y-28 QS can be reliably compared with those in existing databases. Furthermore, the successful optimization of the PCR protocol not only improved cost-effectiveness and efficiency but also showed that the kit maintains high data quality under modified conditions, an important feature for its implementation in routine forensic practice.

Calibration with certified reference materials (CRM)

Calibration was performed using the positive control supplied with the Investigator® Argus Y-28 QS kit, which corresponds to a certified reference material with a genetic profile described in the literature. In all amplification runs, the expected haplotype was consistently obtained, demonstrating accurate allele assignment and confirming the reliability of the method. In line with the laboratory's routine practice, each amplification includes a positive control sample with a previously known genetic profile, ensuring the continuous validation of results during casework and research applications. These findings emphasize the importance of incorporating certified reference materials in validation and routine workflows. By providing a benchmark for allele assignment, CRMs ensure that every amplification can be verified against a reliable standard, thereby reinforcing quality control and guaranteeing the reliability of results across different runs and applications.

- **Indirect Evaluation**

Specificity

The specificity of the Investigator® Argus Y-28 QS kit was assessed using DNA from female human samples, non-human species (horse, pig, dog, and cat), and the standard positive and negative controls.

As expected, no Y-STR profiles were obtained from female or animal DNA, while complete haplotypes were consistently generated from male DNA and the positive control 9948. These findings confirm that the kit selectively amplifies human male DNA, as intended for its forensic application. Demonstrating strict specificity is essential in forensic genetics, as it ensures that the profiles obtained truly represent male genetic material and are not artefacts from female DNA or from other biological sources. This provides confidence that results generated with the kit are reliable and suitable for both casework and population studies.

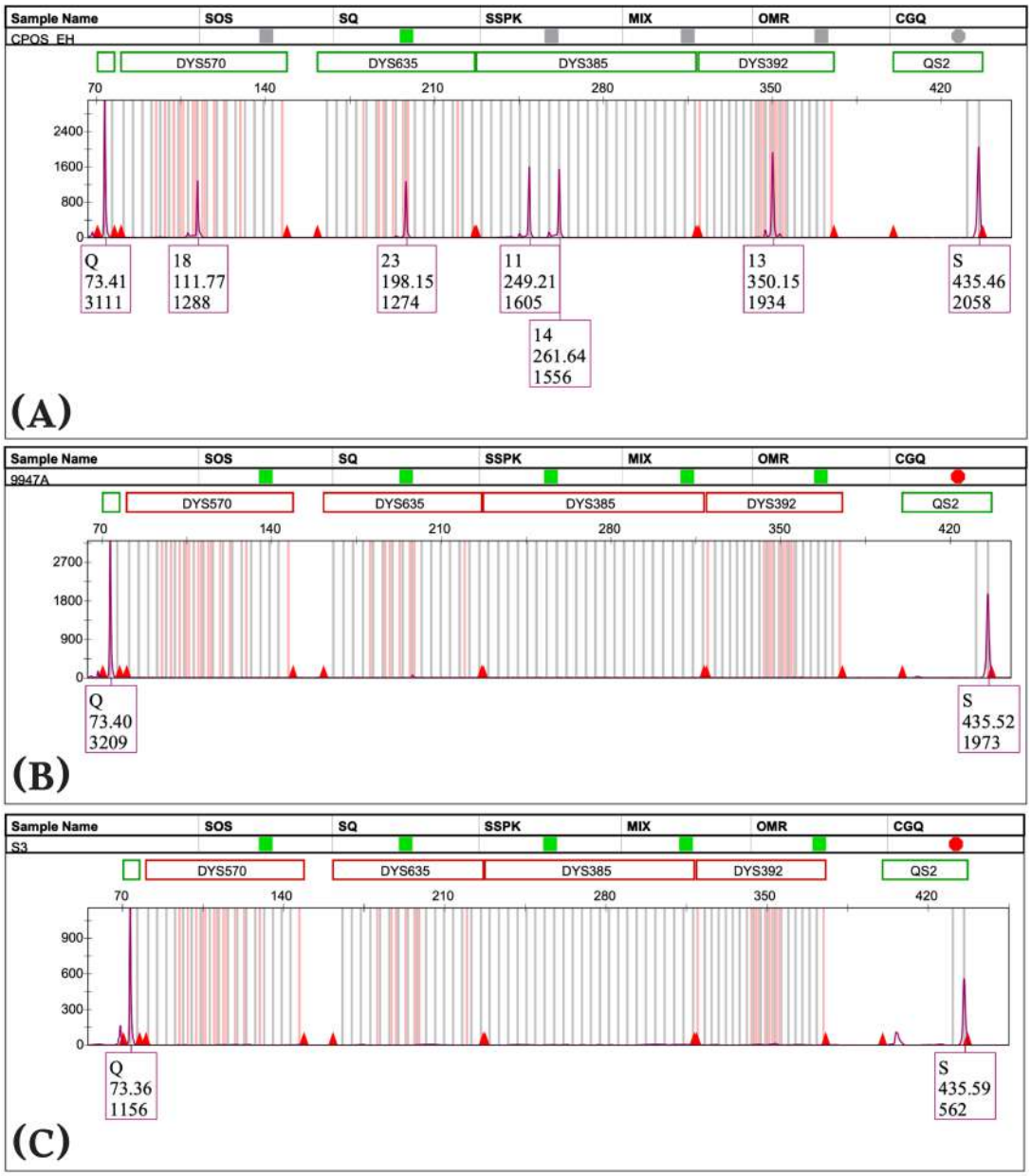

Figure 1. Specificity testing of the Investigator® Argus Y-28 QS kit. (A) Male DNA control(CPOS_EH) showing a complete Y-STR profile; (B) female DNA control (9947A) with no Y-STR alleles detected, only internal quality sensors; (C) non-human DNA (*Sus scrofa*, S3) also showing absence of amplification, with quality sensors confirming successful PCR.

Analytical thresholds

Analytical thresholds were calculated from the analysis of negative amplification controls. The baseline noise detected in each dye channel was measured, and the highest peak values were used as a reference to define the threshold. The analytical thresholds were set above these background signals to ensure that only peaks exceeding the defined RFU value could be

interpreted as true allelic signals. By deriving the thresholds directly from experimental data, the method guarantees that allelic calls are based on genuine signals and not artefacts. This approach strengthens the reliability of haplotype interpretation and ensures that the method performs reliably across different analyses, which is essential in both forensic casework and population studies. The analytical threshold was determined following the method described by Kaiser in 1970, as recommended by IUPAC [17]. Although other formulas exist for threshold calculation, this approach was selected because it is widely used in the validation of other kits implemented at SGBF, INMLCF, I.P. According to this method, the analytical threshold (AT) is defined as:

$$AT = M + 3 * SD$$

Where M corresponds to the mean of the highest peak values observed for each sample, and SD is the associated standard deviation.

The RFU values obtained for each dye channel are summarized in Table I.

Argus Y-28 QS	Blue 6-FAM	Green BTG	Yellow BTY	Red BTR2	Purple BTP
LA (IUPAC + Kaiser) (RFU)	149.88	25.02	32.55	106.52	24.05

Table I - Analytical thresholds calculated for each dye channel using the IUPAC + Kaiser (1970) method.

Repeatability and reproducibility

Repeatability was assessed by amplifying male DNA samples in triplicate under identical conditions. In all cases, the haplotypes obtained were identical across replicates, demonstrating complete concordance. Reproducibility was then tested by repeating the analysis under variable conditions, including different operators, amplification days, and thermal cyclers. The results again showed full concordance, confirming that the Investigator® Argus Y-28 QS kit provides robust and consistent profiles regardless of minor laboratory variations. These findings indicate that the kit performs reliably both within and between runs, ensuring that results are not affected by technical variation. Such robustness is essential in forensic

practice, where analyses are frequently carried out by different operators and under varying laboratory conditions.

Sensitivity

Sensitivity was evaluated using the positive control 9948 and a set of samples tested at different DNA concentrations.

Complete and valid haplotypes were consistently obtained at 0.5 ng/µL. At 0.2 ng/µL, most samples produced full profiles, although allele dropout was observed in three cases. At 0.1 ng/µL, reliable profiles were rare, with only a minority of samples yielding complete haplotypes. At 0.05 and 0.01 ng/µL, all profiles showed significant allele loss, making them unsuitable for interpretation. At higher DNA inputs, valid profiles were obtained at 5 ng/µL. However, at 10 ng/µL and above, artefacts such as additional alleles (drop-in) and peak splitting were observed, likely caused by DNA overload, which compromised the interpretation. Overall, the kit generated reliable haplotypes within the range of 0.5–5 ng/µL. Concentrations below 0.2 ng/µL increased the risk of allele dropout, while concentrations above 5 ng/µL led to artefacts, limiting their forensic applicability.

Table II - Summary of sensitivity testing results obtained with the Investigator® Argus Y-28 QS kit.

DNA Concentration	Profile Quality
0.01 ng/µL	No valid profiles
0.05 ng/µL	Highly incomplete profiles
0.1 ng/µL	Predominantly partial profiles; frequent allele dropouts
0.2 ng/µL	Mostly complete profiles; occasional allele dropouts
0.5 ng/µL	Consistently complete profiles in all samples
5 ng/µL	Complete and robust profiles
≥ 10 ng/µL	Profiles affected by artefacts (drop-in, peak splitting)
Optimal DNA input range = 0.5–5 ng/µL	

Precision

Precision was assessed by analysing 12 allelic ladders from independent electrophoretic runs. For each allele, the mean fragment size and standard deviation were calculated to evaluate consistency in allele sizing.

The largest standard deviation observed was 0.2804312 bp for allele 15 at locus YGATAH4, while the smallest was 0.02384848 bp for allele 21 at locus DYS456. These results confirm that the Investigator® Argus Y-28 QS kit provides high allelic resolution, allowing unequivocal allele assignment across loci. The very low variation in fragment sizing demonstrates the method's suitability for reliable forensic and population genetic applications.

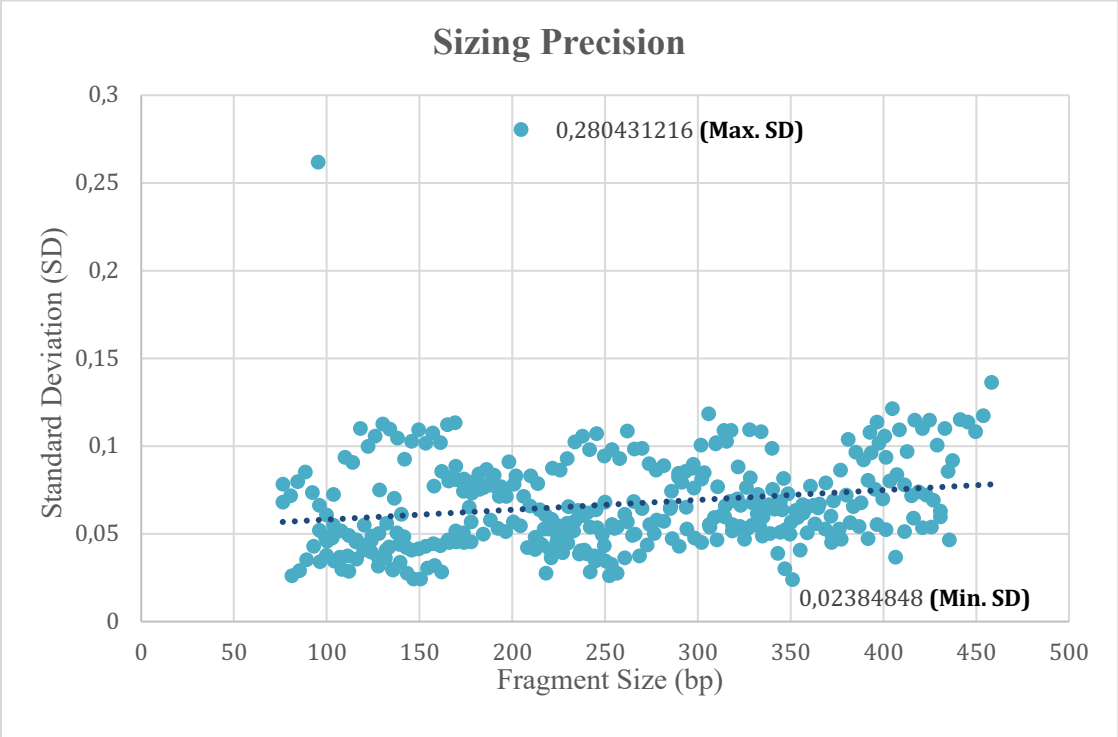

Figure 2. Sizing precision determined from 12 allelic ladders analyzed with the Investigator® Argus Y-28 QS kit. The standard deviation ranged from 0.0238 bp (DYS456) to 0.2804 bp (YGATAH4), demonstrating consistent allele sizing across all loci.

Mixtures and Contamination

The ability of the Investigator® Argus Y-28 QS kit to detect contamination and analyse DNA mixtures was tested using both reference samples and artificially prepared mixtures.

Negative amplification controls consistently showed no detectable DNA, confirming the absence of exogenous contamination during the procedures.

In male/female mixtures, complete male haplotypes were obtained at all tested ratios (1:1, 1:10, and 10:1), with the exception of one 1:10 mixture where three alleles were not detected, likely due to the low quantity of male DNA.

In male/male mixtures, balanced profiles were observed at the 1:1 ratio. At 1:10 and 10:1, the profiles were dominated by the major contributor, and interpretation of the minor contributor's profile became more difficult, particularly when alleles overlapped with stutter positions. Nevertheless, no complete loss of the minor haplotype was observed, although allele dropouts occurred in some loci. These results demonstrate that the Argus Y-28 QS kit can reliably exclude contamination and detect contributors in mixed samples.

However, the detection of minor contributors becomes increasingly limited as their proportion decreases. This behaviour is consistent with other Y-STR kits and highlights the importance of cautious interpretation of mixture profiles in forensic applications.

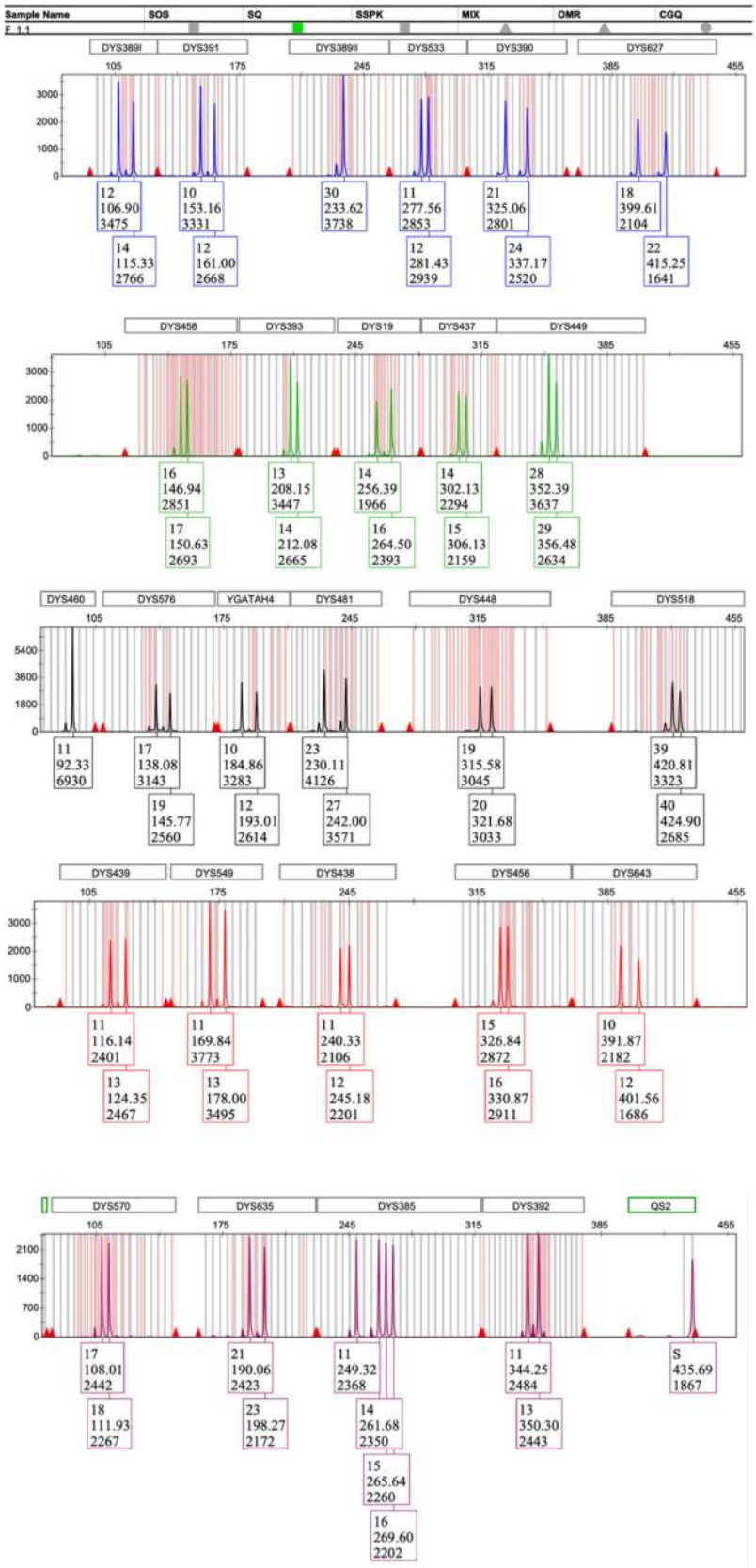

Figure 3. Electropherogram of a 1:1 male DNA mixture (bloodstain and buccal swab) analyzed with the Investigator® Argus Y-28 QS kit. The profile shows the presence of allelic peaks from both contributors, illustrating the kit's performance in detecting and interpreting mixed male DNA samples.

Inhibitors

The effect of physical inhibitors was evaluated by exposing DNA samples to ultraviolet (UV) radiation prior to amplification. Samples exposed to short periods of UV still produced complete haplotypes, although with reduced signal intensity. With longer exposures, progressive allele dropout was observed, and at the highest levels no valid profiles could be recovered.

These findings show that the Investigator® Argus Y-28 QS kit is moderately tolerant to UV-induced DNA damage, but excessive exposure compromises the quality of the genetic profile. This highlights the importance of proper collection, handling, and storage of forensic samples to minimize the risk of degradation caused by environmental factors.

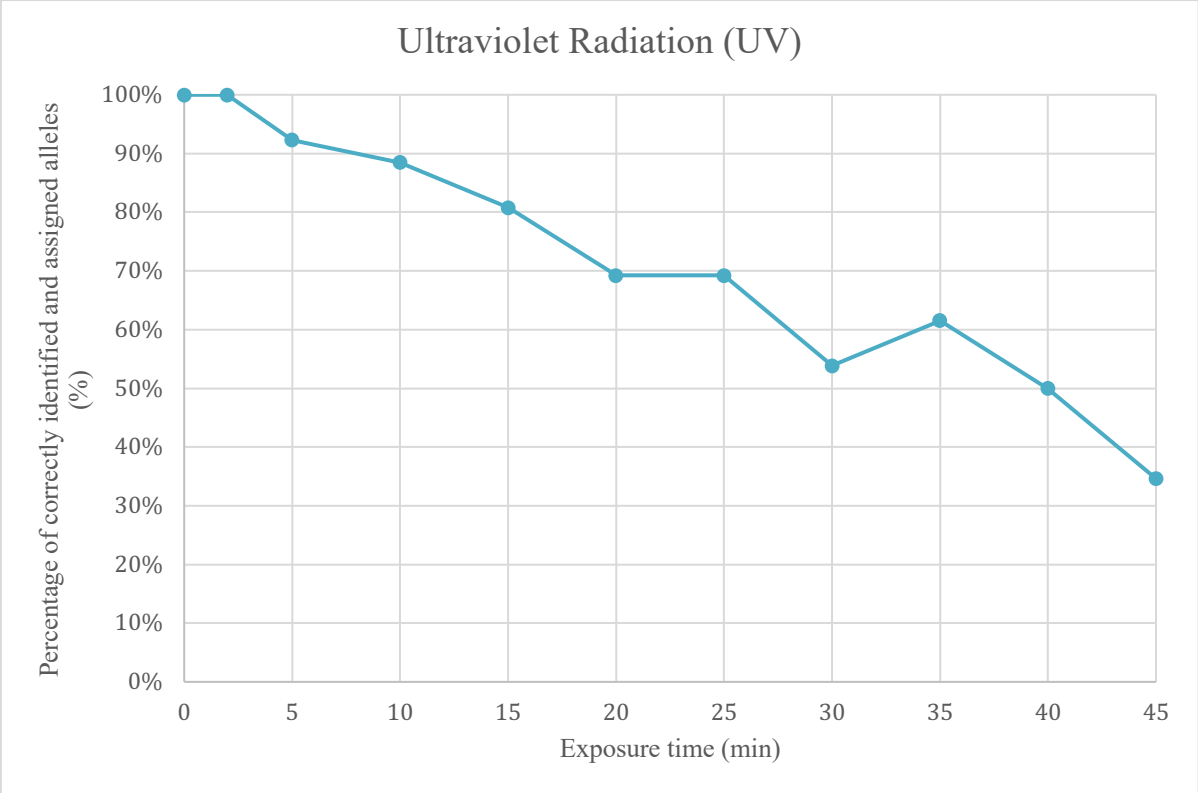

Figure 4. Percentage of alleles correctly identified and assigned as a function of exposure time to ultraviolet (UV) radiation. A decreasing trend was observed in the efficiency of allele identification as exposure time increased, highlighting the negative impact of UV radiation on DNA integrity.

Conclusions

The internal validation of the Investigator® Argus Y-28 QS kit demonstrated that this system provides reliable and reproducible results under the specific conditions of the SGBF-N laboratory. Direct comparison with the Yfiler® Plus kit confirmed full concordance of haplotypes, while calibration with certified reference material ensured accurate allele assignment. Optimized amplification conditions (reduced PCR volume and cycle number) maintained high-quality profiles and were adopted as the standard protocol.

Indirect validation further supported the robustness of the kit. The analytical thresholds defined from negative controls allowed the clear distinction between true allelic signals and background noise. Sensitivity studies indicated that valid and complete profiles can be consistently obtained from as little as 0.5 ng of DNA, whereas higher DNA concentrations (>5 ng) led to artefacts such as allele imbalance and peak splitting. Precision tests confirmed the system's ability to discriminate alleles differing by one or two base pairs. The evaluation of mixtures showed that minor contributors could still be detected even at highly unbalanced ratios, and no artefactual results were obtained in negative controls. Finally, inhibition assays demonstrated that the kit is moderately tolerant to UV exposure, although prolonged radiation significantly compromised allele recovery.

Taken together, these results confirm that the Argus Y-28 QS kit is a robust, accurate, and forensically reliable tool for Y-STR analysis. Its implementation in the laboratory reinforces the reliability of both forensic casework and population studies, including the generation of Cape Verdean reference data presented in Article II of this dissertation.

CHAPTER IV – Original Research

Article I – Y-STR Databases – Application in Sexual Crimes.

genes

MDPI

Review

Y-STR Databases—Application in Sexual Crimes

Rita Costa [1], Jennifer Fadoni [2,3], António Amorim [2,3,4,5] and Laura Cainé [1,3,4,*]

[1] Faculty of Medicine, Porto University, 4200-319 Porto, Portugal; up201906725@fc.up.pt
[2] National Institute of Legal Medicine and Forensic Sciences, I.P., North Branch, 4050-202 Porto, Portugal; jennifer.n.fadoni@inmlcf.mj.pt (J.F.); antonio.j.amorim@inmlcf.mj.pt (A.A.)
[3] LAQV&REQUIMTE, Laboratory of Applied Chemistry, Department of Chemical Sciences, Faculty of Pharmacy, University of Porto, 4050-313 Porto, Portugal
[4] National Institute of Legal Medicine and Forensic Sciences, I.P., Centre Branch, 3000-548 Coimbra, Portugal
[5] Faculty of Sciences, Lisbon University, 1749-016 Lisboa, Portugal
* Correspondence: lauracaine@med.up.pt

Abstract: Background/Objectives: The Y chromosome is a crucial tool in forensic genetics due to its unique characteristics, such as its haploid inheritance and lack of recombination. Y-STRs (short tandem repeats on the Y chromosome) are widely used for identifying male genetic profiles in DNA mixtures, especially in sexual assault cases where high levels of female DNA hinder autosomal analysis. This study evaluates the applicability of Y-STRs in forensic investigations, addressing their limitations and the impact of advanced technologies, such as rapidly mutating Y-STRs (RM Y-STRs). Methods: A comprehensive literature review was conducted to analyze existing knowledge on the application of Y-STRs in sexual crimes. The study also examines the role of population databases, such as YHRD, in estimating haplotype frequencies and enhancing forensic reliability. Results: Y-STR analysis proves essential for male DNA identification in complex mixtures, with RM Y-STRs enhancing discriminatory power. However, limitations persist, particularly in cases involving closely related male lineages. The population database coverage remains insufficient in regions like Cape Verde, affecting forensic reliability. Case studies demonstrate Y-STR effectiveness in solving cold cases and sexual crimes, reinforcing the need for expanded databases and methodological advancements. Conclusions: Y-STRs play a fundamental role in forensic genetics, particularly in sexual assault investigations. Their integration with advanced sequencing technologies and expanded databases is critical for improving forensic accuracy. Ethical considerations regarding genetic data privacy and potential discrimination must be addressed through clear regulations and forensic best practices.

Keywords: Y chromosome; Y-STR; microsatellites; haplotypes; population genetics; genetic diversity; forensic sciences; sexual crimes

Academic Editor: Emiliano Giardina

Received: 1 April 2025
Revised: 23 April 2025
Accepted: 24 April 2025
Published: 25 April 2025

Citation: Costa, R.; Fadoni, J.; Amorim, A.; Cainé, L. Y-STR Databases—Application in Sexual Crimes. *Genes* **2025**, *16*, 484. https://doi.org/10.3390/genes16050484

1. Introduction

The analysis of Y-chromosome STRs (Y-STRs) is widely recognized as an essential tool in forensic genetics, particularly in sexual crime cases. These markers are valuable in situations where male DNA is present in minimal quantities or overwhelmed by a large proportion of female DNA. Studies have demonstrated that Y-STRs enable the precise identification of male genetic profiles, even in complex samples, offering a robust approach for forensic investigations [1].

Portugal and Cape Verde were selected as the focus of this study due to their complementary characteristics in terms of population genetics and forensic infrastructure. Portugal has a relatively well-established Y-STR population database; however, despite

several completed population studies, the available data remain limited in both geographic coverage and the number of loci analyzed, preventing a comprehensive understanding of the country's genetic diversity [2–6]. In contrast, Cape Verde faces significant challenges, including limited forensic infrastructure and a small number of haplotypes recorded in international databases such as YHRD [7]. This comparison between two countries allows for an exploration of both opportunities and limitations in distinct forensic contexts, offering insights that may apply to other nations with similar characteristics.

This article examines the forensic applications and implications of Y-STRs, discussing their advantages and limitations, and provides recommendations to maximize their impact on the criminal justice system, with particular emphasis on Cape Verde and Portugal.

2. Methods

In November 2024, a comprehensive search was conducted in the PubMed database (U.S. National Library of Medicine) without restrictions on publication date or language to identify the relevant literature on the application of Y-STRs in forensic genetics. The keywords used included "Y-STR", "Y chromosome", "forensic genetics", "DNA mixtures", "sexual crimes", "rapidly mutating markers", and "Y-STR databases." The review focused on recent advances in Y-STRs applications, their limitations, and their forensic impact in solving complex cases, particularly those involving sexual crimes.

The inclusion criteria comprised peer-reviewed scientific articles that addressed the application of Y-STRs in forensic genetics, particularly in sexual crimes, with emphasis on DNA mixture analysis, population databases, and technological developments such as RM Y-STR markers. Articles discussing relevant ethical, legal, and methodological aspects were also included. Exclusion criteria involved duplicate records, articles not directly related to the forensic application of Y-STRs (e.g., studies with a purely evolutionary or clinical focus), papers without full-text access, and systematic reviews lacking original data or discussion.

Additionally, documents from the National Institute of Statistics of Cape Verde (INE), the Office for Strategy and Studies (GEE), EUROSTAT (the Statistical Office of the European Union), and the United Nations Office on Drugs and Crime (UNODC), as well as the Annual Internal Security Report (RASI), were consulted for complementary data on this topic.

The literature synthesis was performed through a narrative approach, organized around the following thematic axes: (1) characteristics and limitations of Y-STRs; (2) forensic case studies involving Y-STRs; (3) the role of population databases; (4) methodological and technological developments; and (5) ethical implications and public policy recommendations.

3. The Y Chromosome and Its Application in Forensic Genetics

The human Y chromosome is a unique structure in the genome, composed of approximately 60 Mb and transmitted exclusively through the paternal lineage. This chromosome, present only in male individuals, is crucial for sex determination, as it is responsible for the development of the testes during embryogenesis. With minimal changes across generations, except for occasional mutations, the Y chromosome has distinct characteristics compared to other human chromosomes, such as haploid inheritance and the absence of recombination (except in two specific regions that pair with the X chromosome). These features have made the Y chromosome an essential tool in forensic genetics, genetic genealogy, and evolutionary research [8–10].

4. Y-STRs: Exclusive and Highly Informative Markers

Short tandem repeats on the Y chromosome (Y-STRs) are male-specific (XY) markers widely used in forensic analysis due to their specificity and high informativeness. These

markers are fundamental for identifying individuals through the paternal lineage, conducting paternity tests, differentiating male relatives, inferring geographical origins, and performing molecular anthropology research [11–13].

Due to the absence of recombination on the Y chromosome, Y-STR loci are physically linked, forming haplotypes that are directly transmitted from father to son. This inheritance pattern creates an organized structure of alleles and haplotypes that correlates with geographical and ethnolinguistic patterns, allowing inferences about the population origin of a genetic profile. However, the haploid nature of the Y chromosome presents challenges, as paternally related males often share similar haplotypes, making individualization difficult in forensic investigations [11,14].

Although Y-STR typing allows for the precise identification of male DNA, it presents significant challenges, particularly in the interpretation of both male–male and male–female DNA mixtures, which is especially critical in sexual assault cases with multiple male contributors or samples containing an overwhelming excess of female DNA. The Y chromosome is inherited in a haploid manner and transmitted from father to son, enabling the tracing of paternal lineage through Y-STR analysis. However, the low evidential value of a Y-STR profile, due to the lack of recombination, can be an obstacle in forensic cases, especially in samples with an excess of female DNA. To mitigate this limitation, some studies suggest expanding the set of Y-STR loci with more informative markers. Additionally, while matching a single male Y-STR profile is straightforward, interpreting mixtures requires caution, as coincidental inclusions can occur, as evidenced by a wrongful conviction case based on a coincidental Y-STR inclusion [15].

The use of more sensitive systems and the inclusion of additional markers have proven crucial in increasing accuracy and discrimination in forensic investigations, leading to improved case resolution. Guidelines with recommendations on Y-STR result interpretation in forensic analysis have been established by the DNA Commission of the International Society of Forensic Genetics (ISFG) [14].

5. Real Case Studies of Y-STRs in Forensic Investigation

Y-STRs have proven to be essential in various forensic investigations, particularly in cases involving sexual crimes where male DNA is a crucial part of the evidence. Below are two significant cases in which Y-STRs played a central role in solving crimes and identifying suspects.

5.1. The Boston Strangler Case

The case of Mary Sullivan, one of the victims of the Boston Strangler, was solved in 2013, nearly 50 years after the crime. Sullivan was murdered in 1964, and the initial investigation failed to resolve the case. After years of speculation, a connection was finally made between the crime and Albert DeSalvo, the man who had confessed to the murders but was never convicted because he had already passed way. The key to solving the case was Y-STR analysis performed on DNA collected from DeSalvo's nephew. The test indicated a DNA match between the nephew and the sample collected at the crime scene, implicating DeSalvo as Sullivan's murderer. The exhumation of DeSalvo's body and subsequent DNA testing confirmed the genetic relationship, establishing his identity as the true perpetrator with a precision of 1 in 220 billion, ruling out other possibilities [16].

5.2. The Vaatstra Case in The Netherlands

Another significant case involving Y-STR application occurred in the Netherlands with the case of Marianne Vaatstra, who was brutally raped and murdered in 1999. Despite years

of investigation, the case remained unsolved until 2012, when local police implemented an innovative mass DNA collection strategy known as a "DNA dragnet".

More than 6600 men from the region were invited to provide saliva samples for DNA analysis. Y-STRs were used to identify potential relatives of the perpetrator, as the attacker was unlikely to voluntarily provide a DNA sample. The results revealed a rare Y-STR match in two men from the region, both of whom were ruled out as suspects after autosomal DNA marker analysis. However, the use of Y-STRs allowed investigators to trace the paternal lineage of the perpetrator, ultimately leading to the identification of the culprit. This case demonstrated the value of Y-STRs not only in identifying suspects directly, but also in locating relatives, which enabled the resolution of a long-unsolved case [17,18].

These examples highlight how Y-STRs are essential in investigating cold or archived cases, particularly when DNA mixtures are involved or when the offender's profile is difficult to access due to the predominant presence of female DNA. The application of Y-STRs not only aids in identifying perpetrators but also helps exclude suspects and strengthen links to the crime scene, offering a powerful tool for investigators. The integration of Y-STRs with other forensic technologies, such as investigative genealogy, has become increasingly useful in resolving old and complex cases [19,20].

6. Population Databases: A Critical Necessity

The analysis of Y-STRs has proven to be an essential tool in forensic genetics, particularly in complex cases such as sexual crimes. Despite its limitations, including a reduced exclusion power and the absence of Y-STR profiles in national databases, recent technological advancements and growing scientific recognition have strengthened its use. Several commercially developed kits now offer a greater sensitivity and discriminatory power, making Y-STRs even more effective. However, to optimize their impact on criminal investigations, it is crucial to expand population databases and conduct more extensive studies with real samples, ensuring a more consistent and reliable application in diverse forensic contexts [21].

To maximize the potential of Y-STRs, it is essential to have broad and representative population databases, such as the YHRD (Y-Chromosome Haplotype Reference Database) [22]. These databases allow for the estimation of haplotype frequencies and the quantitative analysis of the probative weight of a Y-STR match. They also consider population substructure, which reflects genetic variations associated with regional, cultural, and ethnolinguistic factors. Statistical models, such as Laplace parameters, are often applied to interpret Y-STR profiles in relation to the populations of origin, thus facilitating the statistical interpretation of DNA evidence [14,23–25].

The use of these databases is indispensable, especially in judicial investigations, where the robustness and accuracy of genetic analyses can significantly impact case outcomes [12].

Although countries like the United States, China, and South Korea have thousands of recorded haplotypes for 27 Y-STR loci in YHRD, Cape Verde currently has only 117 haplotypes registered for 17 loci and none for 26 or more loci (https://yhrd.org (accessed on 20 November 2024)), greatly limiting the discriminatory power of forensic analysis and complicating the exclusion or accurate identification of suspects. In comparison, Portugal has haplotypes recorded for 23 Y-STR loci (248 haplotypes), which, although representing progress, remain relatively insufficient when considering the genetic diversity of the Portuguese population. This highlights the urgent need to expand and update these population databases, particularly in underrepresented regions, to strengthen the reliability and probative value of Y-STR analyses in forensic casework [22].

7. Forensic Applications and the Investigation of Sexual Crimes

The success of solving forensic cases is directly related to the efficient recovery of DNA at the crime scene and the application of a precise and discriminatory method for genetic analysis [26].

Y-STRs are particularly effective tools in the investigation of sexual crimes due to their male specificity, allowing the exclusion of female DNA contribution and facilitating the analysis of male DNA even when present in very low proportions (<5%) [17,27–30]. This approach is especially advantageous for interpreting mixed DNA profiles, providing clearer and more accurate conclusions compared to autosomal profiles that include DNA from both sexes [29,31]. Additionally, the use of Y-STRs can be useful for determining the number of contributors in mixtures containing multiple male donors [21].

With advances in forensic DNA analysis, which increasingly employs more sensitive and effective methods, there has been a significant rise in the collection and processing of trace DNA in investigations. These traces are characterized by their unknown origin in bodily fluids or tissues and have become widely analyzed due to improvements in available techniques [32].

Recent studies highlight the need to understand the patterns of male DNA prevalence and transfer, such as in female underwear, to differentiate background traces from crime-related events. Several studies have confirmed that male DNA levels in female underwear are low unless the woman cohabits with a man. Additionally, a consistent trend has been observed in which the amount of male DNA recovered from women's clothing decreases as the intensity and frequency of physical contact with men decline [32–34].

Studies have shown that even 24 to 46 h after a sexual assault, viable Y-STR profiles can still be obtained, as spermatozoa may persist in the vaginal canal for up to three days [35]. Nevertheless, the detection of male DNA can be hindered by the preferential amplification of the female component, requiring complementary methods, such as Y-STR typing, to identify the perpetrator [29,30].

Y-STRs have been widely used in various forensic scenarios, including the exclusion of male suspects, the identification of the paternal lineage of perpetrators, the determination of multiple male contributors in complex biological mixtures, and the inference of population origins of unknown suspects [17].

Most forensic samples obtained in cases of sexual assault are intimate swabs collected from the genital tract of a female victim to identify the DNA of a male perpetrator [36]. However, despite advances in semen detection and recovery, the majority of alleged sexual crimes involving vaginal penetration do not present traces of semen [37].

In situations where there is no presence of semen or spermatozoa (such as in cases of digital penetration, penile penetration without ejaculation, or azoospermic perpetrators, including infertile or vasectomized men), it is known that male epithelial cells from fingers or the penis can be transferred to the vagina and subsequently identified [36–39].

However, due to the large amount of DNA from female epithelial cells in these samples, identifying the perpetrator's DNA using autosomal analysis methods becomes extremely challenging. In this context, the analysis of Y-chromosome short tandem repeats (Y-STRs), which are specific to male DNA, has become an essential tool in forensic investigations of sexual assault cases. This method allows the recovery of male DNA profiles even when female DNA is present in amounts exceeding 1000 times [36,40,41].

8. Recent Advances: Rapidly Mutating Y-STRs (RM Y-STRs)

The development of rapidly mutating Y-STRs (RM Y-STRs) has marked a significant milestone in forensic genetics. These markers, identified by Ballantyne et al. [42], exhibit mutation rates higher than 1×10^{-2} per locus per generation, which allows for a greater

discrimination between related individuals, including first-degree relatives. This advancement has significantly enhanced forensic analysis in cases where traditional Y-STR profiles are not sufficiently discriminatory. RM Y-STRs have also facilitated the differentiation of profiles in complex biological mixtures and improved efficiency in solving criminal cases, emerging as one of the most promising innovations in the field [43,44].

9. Other Relevant Applications

Beyond their forensic applications, Y-STRs are widely used in historical and genealogical studies. These markers enable the tracing of ancestral human lineages and the inference of migratory routes, providing insights into the genetic history of human populations. They are also valuable in the identification of missing persons and disaster victims, particularly in cases where autosomal DNA does not provide sufficient information [8,17].

10. Statistics

10.1. Statistics on Cape Verde in Portugal

In 2021, there were 34,093 Cape Verdeans residing in Portugal, representing 4.88% of the immigrant population in the country. Currently, the Cape Verdean population is the third-largest foreign community in Portugal, with the highest concentrations in Lisbon (21,141), Setúbal (6373), and Faro (1896) [45].

10.2. Statistics on Cape Verde

According to the National Institute of Statistics of Cape Verde, sexual crimes are among the least reported types of crime. In 2015, 104 cases of sexual abuse of children and minors were recorded, representing 0.4% of all incidents that year, marking an 8.3% increase compared to the previous year (2014). Between 2010 and 2015, there was an average annual increase of 9.9% in the number of reported cases of sexual abuse of children and minors. In this category of crime, female children and minors are the primary victims, accounting for 92.5% of the cases. On the other hand, male individuals aged between 22 and 30 years are the main perpetrators of these crimes. In 2015, 106 cases of sexual assault were also recorded. Compared to the previous year, there was a 15.2% increase in sexual assault incidents. Similarly to child sexual abuse cases, female individuals represent the majority of victims (90.2%), while male individuals are the primary perpetrators of this type of crime [46].

10.3. Statistics on Portugal

According to the 2023 Annual Internal Security Report (RASI), the victims of crimes against sexual freedom and self-determination are predominantly female (77.3%), while the defendants are primarily male (94.3%). Among the investigations initiated, the most common type of crime was the sexual abuse of children (39.5%), followed by rape and child pornography, with 20.2% and 12.8%, respectively. More specifically, in rape cases, it was found that 100% of the defendants are male, and 90.7% of the victims are female [47].

10.4. Statistics of the European Union

According to the UNODC (United Nations Office on Drugs and Crime) and EUROSTAT (European Union Statistical Office), between 2008 and 2015, rape crimes showed a 47% increase, with the majority of victims being women (85.8%). On the other hand, the suspects (96.5%) and defendants (98.3% of the convicted) are predominantly male [48]. In the EU, according to Eurostat, 231,456 sexual violence offenses were recorded by the police in 2022, representing a 10.3% increase compared to the previous year. This trend has been rising since 2015, except in 2020 [49].

11. Implications and Public Policy Recommendations

Although Y-STRs offer powerful tools for forensic investigations, their application raises ethical and legal challenges that cannot be overlooked. The creation and use of genetic databases require a careful balance between investigative benefits and the protection of fundamental rights, such as privacy and non-discrimination. Studies have already shown that the use of these databases can generate concerns about misuse, especially in legal systems where regulatory frameworks are weak or outdated [14,17,50]. Recent studies highlight advancements in technologies, such as the use of next-generation sequencing (NGS), which has increased the accuracy of genetic analyses but also amplified concerns about privacy, particularly in countries where specific guidelines for handling genetic information are still lacking [51]. Ethical issues include, for example, the privacy of genetic data, as the inclusion of profiles in databases, especially of individuals who have not been convicted, may violate basic rights [17,52]. Furthermore, the analysis of Y-STRs, being associated with paternal lineage, may reinforce stereotypes or lead investigations unfairly against specific ethnic or cultural groups, especially in minority populations [23,53]. Another relevant ethical concern is informed consent, as the collection of genetic samples is not always accompanied by clear explanations of their use and implications [40,54]. To maximize the benefits and minimize the risks associated with the use of Y-STRs, it is essential for governments and institutions to adopt comprehensive and specific policies. Firstly, the creation of legislation to regulate the use of genetic databases is needed, defining clear criteria for profile inclusion, privacy protection, and measures against discrimination [17,50]. At the same time, expanding population databases is crucial, especially in countries like Cape Verde, which still have a limited number of registered profiles. This effort can be enhanced through international partnerships with platforms such as the Y-Chromosome Haplotype Reference Database (YHRD) [23,55]. Additionally, specialized training for forensic and judicial professionals is essential to ensure the correct interpretation of genetic data, especially in cases involving complex mixtures or where genetic evidence is the primary proof [14,26,56]. Comparing the use of Y-STRs in Cape Verde and Portugal with countries that have more advanced systems, such as South Korea and China, can provide valuable insights. These nations have developed robust and well-structured databases, covering more than 27 loci, which enables a more precise analysis and greater discriminatory power in investigations. These examples demonstrate how the combination of technological advancements and stringent policies can maximize the positive impact of genetic tools [44,57]. Based on this evidence, it is possible to conclude that Y-STRs have the potential to transform forensic investigations, particularly in cases of sexual crimes. However, for their use to be fully effective, continuous investment in technology, the expansion of population databases, and the adaptation of legal systems to handle genetic evidence in an ethical and reliable manner are necessary. Moreover, collaborative initiatives between countries can ensure a more equitable representation of populations and the continuous scientific validation of the methodologies used. These efforts will not only strengthen justice systems in Cape Verde and Portugal but also serve as a model for other nations facing similar challenges in the use of genetic technologies in criminal investigations.

12. Discussion

This study presented highlights the importance of the Y chromosome and Y-STRs in forensic genetics, especially in complex cases of sexual crimes. The ability to identify male DNA in mixtures containing large quantities of female genetic material represents a significant scientific advance. However, despite the well-established applications of this method, it is essential to critically assess the technical and contextual limitations that may affect the reliability of the results.

Among the technical challenges, allelic dropout is a prominent issue, especially in degraded or low-template samples. This phenomenon compromises the completeness of genetic profiles, directly impacting the capacity for comparison and individual identification. Furthermore, DNA degradation (frequently observed in samples exposed to adverse environmental conditions), may impair the uniform amplification of all loci, resulting in partial profiles and limiting the scope of forensic interpretation.

The high sensitivity of amplification protocols employed in Y-STR typing, while enabling the analysis of extremely small quantities of genetic material, also heightens the risk of contamination. This necessitates rigorous laboratory protocols and stringent quality control measures during both the collection and analysis phases.

Another critical consideration involves the representativeness of population databases, such as the Y Chromosome Haplotype Reference Database (YHRD), which are fundamental for estimating the probative value of a genetic profile. Underrepresented populations (such as Cape Verde) may introduce sampling bias and compromise the statistical robustness of the analyses, ultimately limiting the evidential value in forensic casework and increasing the risk of misinterpretation, particularly when inferring geographic origin.

Beyond these technical limitations, it is also essential to consider the ethical and legal implications associated with the use and expansion of genetic databases. In Portugal, Law No. 5/2008, of 12 February establishes clear guidelines for the collection, storage, use, and deletion of DNA profiles, aiming to balance the needs of criminal investigations with the protection of fundamental rights, particularly regarding privacy and informed consent.

In Cape Verde, personal data protection is regulated by Decree-Law No. 13/2007, of 26 February, which classifies genetic data as "sensitive data" and requires explicit consent for its processing. However, to date, no specific regulation has been enacted to address the unique challenges inherent to the forensic use of genetic profiles, leaving significant gaps in the protection of citizens' rights and in the establishment of best practices for forensic applications.

13. Future Perspectives

The progress in the identification of RM Y-STRs represents a significant advancement in the field of forensic genetics, enabling a more precise discrimination between individuals belonging to the same paternal lineage and substantially expanding the potential for solving complex cases. However, the successful implementation of these advances requires not only modernized laboratory infrastructure and validated methodologies but also the continuous training and qualification of forensic experts to ensure the accurate interpretation and application of genetic data.

Simultaneously, the expansion of genetic databases and the use of mass screening strategies, such as those employed in large-scale investigations (e.g., DNA dragnets), must be approached with extreme caution, particularly regarding informed consent, the risk of unauthorized secondary use of genetic data, and the potential stigmatization of minority groups.

The increasing adoption of Y-STR analysis and the enhancement of genetic databases should not be regarded as a standalone solution, but rather as a complementary tool, achieving its full potential only when integrated into a judicial system that is ethical, technically proficient, and socially responsible.

In this context, international cooperation plays a vital role, facilitating not only the exchange of best practices in both laboratory and judicial domains but also the harmonization of data protection standards and ethical procedures, with special consideration given to the national contexts of Portugal and Cape Verde.

The future of forensic genetics and of Y-STR analysis will inevitably depend on maintaining a careful balance between technological innovation, scientific rigor, and the protection of fundamental rights. This balance is essential not only to ensure technical reliability but also to strengthen public trust and uphold the legitimacy of using genetic data in judicial processes.

14. Conclusions

The expansion and diversification of Y-STR databases, combined with investments in technology and professional training, are essential to maximize the potential of these tools. At the same time, implementing clear and transparent ethical regulations will be indispensable to ensure public trust in the use of these technologies. By combining these efforts, it will be possible to strengthen forensic investigations and offer more robust and fair judicial responses, both in Cape Verde and Portugal.

The Y chromosome and Y-STRs emerge as indispensable tools in forensic genetics, offering precision in identifying male individuals and inferring paternal lineages. Their application is especially relevant in sexual crime cases, where the unique characteristics of Y-STRs enable effective analysis even in situations involving complex DNA mixtures. Additionally, recent advancements, such as RM Y-STRs, have enhanced their discriminatory capacity, enabling differentiation between individuals with genetic connections and significantly increasing the scope of forensic investigations.

The statistics from Cape Verde and Portugal clearly illustrate the urgency of effective strategies to address sexual crimes. In Cape Verde, the steady increase in reported cases, especially child sexual abuse, reflects a severe social issue, with most victims being young women and the perpetrators adult men. In Portugal, the prevalence of sexual crimes such as child abuse and rape, with nearly all suspects being male, reinforces the need for robust genetic identification methods. These data, along with European trends showing a rise in sexual violence, underline the importance of tools such as Y-STRs in supporting criminal justice.

Therefore, the use of Y-STRs, combined with extensive and representative population databases, emerges as a viable and effective solution to assist forensic investigations, especially in contexts where conventional methods face limitations. By integrating advanced technologies with a detailed understanding of regional realities, as demonstrated by the cases in Cape Verde and Portugal, it is possible to significantly improve the response to sexual crimes and strengthen the justice system.

While this article has focused on the specific realities of Cape Verde and Portugal, its findings can be extended to other countries that share similar forensic and infrastructural challenges. These insights may contribute to the development of forensic strategies tailored to local contexts while ensuring the protection of individual rights and ethical standards in the handling of genetic data.

Furthermore, investigations should be accompanied by clear regulations on the ethical use of genetic information, ensuring privacy and non-discrimination. The future of forensic investigations will depend on the combination of these advanced technologies with the continuous education of justice professionals and the implementation of robust public policies that ensure efficiency and equity across legal systems.

Author Contributions: Conceptualization, R.C.; methodology, R.C. and J.F.; writing—original draft preparation, R.C.; writing—review and editing, J.F., L.C. and A.A.; supervision, J.F. and L.C.; funding acquisition, L.C. and A.A. All authors have read and agreed to the published version of the manuscript.

Funding: The APC was funded by Faculty of Medicine, Porto University, Portugal.

Data Availability Statement: No new data were created or analyzed in this study.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

DNA	Deoxyribonucleic Acid
EUROSTAT	Statistical Office of the European Union
INE	National Institute of Statistics
ISFG	International Society of Forensic Genetics
NGS	Next-Generation Sequencing
RASI	Annual Internal Security Report
RM Y-STR	Rapid Mutation Y-Chromosome Short Tandem Repeats
UNODC	United Nations Office on Drugs and Crime
Y-STR	Short Tandem Repeats of the Y Chromosome

References

1. Sibille, I.; Duverneuil, C.; Lorin de la Grandmaison, G.; Guerrouache, K.; Teissière, F.; Durigon, M.; de Mazancourt, P. Y-STR DNA amplification as biological evidence in sexually assaulted female victims with no cytological detection of spermatozoa. *Forensic Sci. Int.* **2002**, *125*, 212–216. [CrossRef]
2. Bento, A.M.; Carvalho, M.; Lopes, V.; Serra, A.; Costa, H.A.; Andrade, L.; Balsa, F.; Oliveira, C.; Batista, L.; Gamero, J.; et al. Distribution of Y-chromosomal haplotypes in the Central Portuguese population using 17-STRs. *Forensic Sci. Int. Genet.* **2009**, *4*, e35–e36. [CrossRef]
3. Pontes, M.L.; Cainé, L.; Abrantes, D.; Lima, G.; Pinheiro, M.F. Allele frequencies and population data for 17 Y-STR loci (AmpFlSTR Y-filer) in a Northern Portuguese population sample. *Forensic Sci. Int.* **2007**, *170*, 62–67. [CrossRef]
4. Fernandes, A.; Brehm, A. Y-chromosome STR haplotypes in the Açores Archipelago (Portugal). *Forensic Sci. Int.* **2003**, *135*, 239–242. [CrossRef] [PubMed]
5. Carvalho, M.; Anjos, M.J.; Andrade, L.; Lopes, V.; Santos, M.V.; Gamero, J.J.; Corte Real, F.; Vide, M.C. Y-chromosome STR haplotypes in two population samples: Azores Islands and Central Portugal. *Forensic Sci. Int.* **2003**, *134*, 29–35. [CrossRef]
6. Fernandes, A.T.; Brehm, A.; Gusmão, L.; Amorim, A. Y-chromosome STR haplotypes in the Madeira archipelago population. *Forensic Sci. Int.* **2001**, *122*, 178–180. [CrossRef] [PubMed]
7. Fernandes, A.T.; Velosa, R.; Jesus, J.; Carracedo, A.; Brehm, A. Genetic differentiation of the Cabo Verde archipelago population analysed by STR polymorphisms. *Ann. Hum. Genet.* **2003**, *67*, 340–347. [CrossRef] [PubMed]
8. Hodișan, R.; Zaha, D.C.; Jurca, C.M.; Irina, S.; Bembea, M. Genetic Diversity Based on the Analysis of 27 Y- Short Tandem Repetition (STR) Loci in Two Populations in the Apuseni Mountains, Romania. *Cureus* **2024**, *16*, e62505. [CrossRef]
9. Jobling, M.A.; Pandya, A.; Tyler-Smith, C. The Y chromosome in forensic analysis and paternity testing. *Int. J. Leg. Med.* **1997**, *110*, 118–124. [CrossRef]
10. Noordam, M.J.; Repping, S. The human Y chromosome: A masculine chromosome. *Curr. Opin. Genet. Dev.* **2006**, *16*, 225–232. [CrossRef]
11. Roewer, L. Y chromosome STR typing in crime casework. *Forensic Sci. Med. Pathol.* **2009**, *5*, 77–84. [CrossRef]
12. Parkin, E.J.; Kraayenbrink, T.; van Driem, G.L.; Tshering Of Gaselô, K.; de Knijff, P.; Jobling, M.A. 26-Locus Y-STR typing in a Bhutanese population sample. *Forensic Sci. Int.* **2006**, *161*, 1–7. [CrossRef] [PubMed]
13. Haarkötter, C.; Isabel Medina-Lozano, M.; Vinueza-Espinosa, D.C.; Saiz, M.; Gálvez, X.; Carlos Álvarez, J.; Antonio Lorente, J. Evaluating the efficacy of three Y-STRs commercial kits in degraded skeletal remains. *Sci. Justice* **2024**, *64*, 543–548. [CrossRef]
14. Roewer, L.; Andersen, M.M.; Ballantyne, J.; Butler, J.M.; Caliebe, A.; Corach, D.; D'Amato, M.E.; Gusmão, L.; Hou, Y.; de Knijff, P.; et al. DNA commission of the International Society of Forensic Genetics (ISFG): Recommendations on the interpretation of Y-STR results in forensic analysis. *Forensic Sci. Int. Genet.* **2020**, *48*, 102308. [CrossRef] [PubMed]
15. Hampikian, G.; Peri, G.; Lo, S.S.; Chin, M.H.; Liu, K.L. Case report: Coincidental inclusion in a 17-locus Y-STR mixture, wrongful conviction and exoneration. *Forensic Sci. Int. Genet.* **2017**, *31*, 1–4. [CrossRef] [PubMed]
16. Bulman, P. Solving Cold Cases with DNA: The Boston Strangler Case. Available online: https://nij.ojp.gov/topics/articles/solving-cold-cases-dna-boston-strangler-case (accessed on 26 November 2024).
17. Kayser, M. Forensic use of Y-chromosome DNA: A general overview. *Hum. Genet.* **2017**, *136*, 621–635. [CrossRef]
18. Syndercombe Court, D. The Y chromosome and its use in forensic DNA analysis. *Emerg. Top. Life Sci.* **2021**, *5*, 427–441. [CrossRef]

19. Ge, J.; Crysup, B.; Peters, D.; Franco, R.; Liu, M.; Wang, X.; Huang, M.; Budowle, B. MPKin-YSTR: Interpretation of Y chromosome STR haplotypes for missing persons cases. *Electrophoresis* **2023**, *44*, 1080–1087. [CrossRef]
20. Zhang, J.; Li, X.; Li, L.; Chen, A.; Zhang, S. Y-STR analysis of highly degraded DNA from skeletal remains over 70 years old. *Forensic Sci. Res.* **2024**, *9*, owae020. [CrossRef]
21. Purps, J.; Geppert, M.; Nagy, M.; Roewer, L. Validation of a combined autosomal/Y-chromosomal STR approach for analyzing typical biological stains in sexual-assault cases. *Forensic Sci. Int. Genet.* **2015**, *19*, 238–242. [CrossRef]
22. Y-Chromosome Haplotype Reference Database (YHRD). Available online: https://yhrd.org (accessed on 20 November 2024).
23. Roewer, L. The Y-Short Tandem Repeat Haplotype Reference Database (YHRD) and Male Population Stratification in Europe—Impact on Forensic Genetics. *Forensic Sci. Rev.* **2003**, *15*, 165–172. [PubMed]
24. Al-Shoba, K.; Al-Hamadi, N.; Almohammed, E.K.; Hadi, S.; Goodwin, W.; Lazim, H. The Yemeni genetic structure revealed by the Y chromosome STRs. *Forensic Sci. Med. Pathol.* **2025**, 1–7. [CrossRef]
25. Purps, J.; Siegert, S.; Willuweit, S.; Nagy, M.; Alves, C.; Salazar, R.; Angustia, S.M.; Santos, L.H.; Anslinger, K.; Bayer, B.; et al. A global analysis of Y-chromosomal haplotype diversity for 23 STR loci. *Forensic Sci. Int. Genet.* **2014**, *12*, 12–23. [CrossRef] [PubMed]
26. D'Amato, M.E.; Bajic, V.B.; Davison, S. Design and validation of a highly discriminatory 10-locus Y-chromosome STR multiplex system. *Forensic Sci. Int. Genet.* **2011**, *5*, 122–125. [CrossRef]
27. Ralf, A.; Zieger, M.; Kayser, M. Considerations on expanding criminal offender DNA databases with Y-STR profiles. *J. Law. Biosci.* **2024**, *11*, lsae017. [CrossRef] [PubMed]
28. Ashirbekov, Y.; Zhunussova, A.; Abaildayev, A.; Bukayeva, A.; Sabitov, Z.; Zhabagin, M. Genetic polymorphism of 27 Y-STR loci in Kazakh populations from Central Kazakhstan. *Ann. Hum. Biol.* **2024**, *51*, 2377571. [CrossRef]
29. Hall, A.; Ballantyne, J. Novel Y-STR typing strategies reveal the genetic profile of the semen donor in extended interval post-coital cervicovaginal samples. *Forensic Sci. Int.* **2003**, *136*, 58–72. [CrossRef]
30. Davies, A.; Wilson, E. The persistence of seminal constituents in the human vagina. *Forensic Sci.* **1974**, *3*, 45–55. [CrossRef]
31. Dooley, K.B.; Thabang Madisha, M.; Strümpher, S.; Ehlers, K. Forensic genetic value of 27 Y-STR loci (Y-Filer® Plus) in the South African population. *Sci. Justice* **2022**, *62*, 358–364. [CrossRef]
32. Roersma, R.; Storm, J.; Joling, J.; Kokshoorn, B. Prevalence of male DNA on female worn undergarments. How small datasets may support robust opinions in activity level evaluative reporting. *Forensic Sci. Int.* **2024**, *361*, 112097. [CrossRef]
33. Doole, S.; Hargreaves, C.; Davidson, G.; Alexander, K.; Boyce, M.; Fraser, I.; Murphy, C.; Lee-Gorman, M.; Rogers, C. Examining the transfer of male DNA onto female underwear in simulated non sexual social interaction versus digital penetration to assist in the evaluation of results in casework scenarios. *Sci. Justice* **2023**, *63*, 90–108. [CrossRef] [PubMed]
34. Murphy, C.; Kenna, J.; Flanagan, L.; Lee Gorman, M.; Boland, C.; Ryan, J. A Study of the Background Levels of Male DNA on Underpants Worn by Females. *J. Forensic Sci.* **2020**, *65*, 399–405. [CrossRef] [PubMed]
35. Cerri, N.; Ricci, U.; Sani, I.; Verzeletti, A.; De Ferrari, F. Mixed stains from sexual assault cases: Autosomal or Y-chromosome short tandem repeats? *Croat. Med. J.* **2003**, *44*, 289–292.
36. Henry, J.; Scandrett, L. Assessment of the Yfiler® Plus PCR amplification kit for the detection of male DNA in semen-negative sexual assault cases. *Sci. Justice* **2019**, *59*, 480–485. [CrossRef] [PubMed]
37. Owers, R.; McDonald, A.; Montgomerie, H.; Morse, C. A casework study comparing success rates and expectations of detecting male DNA using two different Y-STR multiplexes on vaginal swabs in sexual assault investigations where no semen has been detected. *Forensic Sci. Int. Genet.* **2018**, *37*, 1–5. [CrossRef]
38. Dziegelewski, M.; Simich, J.P.; Rittenhouse-Olson, K. Use of a Y chromosome probe as an aid in the forensic proof of sexual assault. *J. Forensic Sci.* **2002**, *47*, 601–604. [CrossRef]
39. McDonald, A.; Jones, E.; Lewis, J.; O'Rourke, P. Y-STR analysis of digital and/or penile penetration cases with no detected spermatozoa. *Forensic Sci. Int. Genet.* **2015**, *15*, 84–89. [CrossRef]
40. Prinz, M.; Boll, K.; Baum, H.; Shaler, B. Multiplexing of Y chromosome specific STRs and performance for mixed samples. *Forensic Sci. Int.* **1997**, *85*, 209–218. [CrossRef] [PubMed]
41. Mulero, J.J.; Chang, C.W.; Calandro, L.M.; Green, R.L.; Li, Y.; Johnson, C.L.; Hennessy, L.K. Development and validation of the AmpFlSTR Yfiler PCR amplification kit: A male specific, single amplification 17 Y-STR multiplex system. *J. Forensic Sci.* **2006**, *51*, 64–75. [CrossRef]
42. Ballantyne, K.N.; Keerl, V.; Wollstein, A.; Choi, Y.; Zuniga, S.B.; Ralf, A.; Vermeulen, M.; de Knijff, P.; Kayser, M. A new future of forensic Y-chromosome analysis: Rapidly mutating Y-STRs for differentiating male relatives and paternal lineages. *Forensic Sci. Int. Genet.* **2012**, *6*, 208–218. [CrossRef]
43. Semikhodskii, A.; Krassotkin, Y.; Makarova, T.; Zavarin, V.; Ilina, V.; Sutyagina, D. Population genetic data and forensic parameters of the 27 Y-STR panel Yfiler(®) Plus in Russian population. *Int. J. Leg. Med.* **2021**, *135*, 1785–1787. [CrossRef] [PubMed]
44. Alghafri, R.; Zupanič Pajnič, I.; Zupanc, T.; Balažic, J.; Shrivastava, P. Rapidly mutating Y-STR analyses of compromised forensic samples. *Int. J. Leg. Med.* **2018**, *132*, 397–403. [CrossRef] [PubMed]

45. Gestão de Dados e Análise Estatística de Estudos. População Estrangeira Com Estatuto Legal de Residente em Portugal—Cabo Verde. Available online: https://www.gee.gov.pt/pt/lista-publicacoes/estatisticas-de-imigrantes-em-portugal-por-nacionalidade/paises/Cabo%20Verde/3947-populacao-estrangeira-com-estatuto-legal-de-residente-em-portugal-cabo-verde/file (accessed on 20 November 2024).
46. Instituto Nacional de Estatística. Cabo Verde, Justiça e Segurança em Números—2015. Available online: https://ine.cv/wp-content/uploads/2017/03/justica-e-seguranca_rev01.pdf (accessed on 20 November 2024).
47. Secretário-Geral, G.d. Relatório Anual de Segurança Interna 2023. Available online: https://www.portugal.gov.pt/pt/gc24/comunicacao/documento?i=relatorio-anual-de-seguranca-interna-2023 (accessed on 20 November 2024).
48. Eurostat. Estatísticas Sobre a Criminalidade: Crimes de Violência Sexual. Available online: https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Archive:Estat%C3%ADsticas_sobre_a_criminalidade&oldid=364651#Crimes_de_viol.C3.AAncia_sexual (accessed on 22 November 2024).
49. Eurostat. Crime Statistics: 231 456 Sexual Violence Crimes in the EU in 2022, a Continued Rise. Available online: https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Crime_statistics#231_456_sexual_violence_crimes_in_the_EU_in_2022.2C_a_continued_rise (accessed on 22 November 2024).
50. Heffernan, L. Genetic Policing: The Use of DNA in Criminal Investigations. By Robin Williams and Paul Johnson (Willan Publishing, 2008). *Br. J. Criminol.* **2008**, *48*, 699–701. [CrossRef]
51. Alvarez-Cubero, M.J.; Saiz, M.; Martínez-García, B.; Sayalero, S.M.; Entrala, C.; Lorente, J.A.; Martinez-Gonzalez, L.J. Next generation sequencing: An application in forensic sciences? *Ann. Hum. Biol.* **2017**, *44*, 581–592. [CrossRef]
52. Murphy, E. Relative Doubt: Familial Searches of DNA Databases. *Mich. L. Rev.* **2009**, *109*, 291.
53. Montague, S. Genetic Suspects: Global Governance of Forensic DNA Profiling and Databasing. *Yale J. Biol. Med.* **2011**, *84*, 326.
54. Laurie, G.; Hepple, B.; Cooke, G.; Holm, S.; Parry, B.; Read, A.; Williams, R. *The Forensic Use of Bioinformation: Ethical Issues*; Nuffield Council on Bioethics: London, UK, 2007.
55. Jobling, M.A.; Gill, P. Encoded evidence: DNA in forensic analysis. *Nat. Rev. Genet.* **2004**, *5*, 739–751. [CrossRef]
56. Schneider, P.M. Scientific standards for studies in forensic genetics. *Forensic Sci. Int.* **2007**, *165*, 238–243. [CrossRef]
57. Lee, H.C.; Ladd, C. Preservation and collection of biological evidence. *Croat. Med. J.* **2001**, *42*, 225–228.

Article II – The Genetic Structure of Cape Verdean Population Revealed by Y-Chromosome STRs

genes

MDPI

Article

The Genetic Structure of Cape Verdean Population Revealed by Y-Chromosome STRs

Rita Costa [1], Jennifer Fadoni [2,3], António Amorim [2,3,4,5] and Laura Cainé [1,3,4,*]

[1] Faculty of Medicine, Porto University, 4200-319 Porto, Portugal; up201906725@fc.up.pt
[2] National Institute of Legal Medicine and Forensic Sciences, I.P., North Branch, 4050-202 Porto, Portugal; jennifer.n.fadoni@inmlcf.mj.pt (J.F.); antonio.j.amorim@inmlcf.mj.pt (A.A.)
[3] LAQV&REQUIMTE, Laboratory of Applied Chemistry, Department of Chemical Sciences, Faculty of Pharmacy, University of Porto, 4050-313 Porto, Portugal
[4] National Institute of Legal Medicine and Forensic Sciences, I.P., Centre Branch, 3000-548 Coimbra, Portugal
[5] Faculty of Sciences, Lisbon University, 1749-016 Lisboa, Portugal
* Correspondence: laura.m.caine@inmlcf.mj.pt

Abstract

Background/Objectives: Y-chromosomal short tandem repeats (Y-STR) are genetic markers widely used in forensic and population genetics. However, despite their importance, many populations remain under-represented in published studies and genetic databases. One such population is the Cape Verdean, which, despite its unique history of admixture between European and sub-Saharan African populations, continues to be under-represented in global Y-STR reference databases. This study aims to characterize the Y-STR haplotype diversity and paternal lineage composition of the Cape Verdean population using a high-resolution STR panel. **Methods:** A total of 143 unrelated Cape Verdean men were analyzed using a set of 26 Y-STR loci, including rapidly mutating markers. Allele and haplotype frequencies were calculated, along with standard forensic parameters such as gene and haplotype diversity. Paternal lineages were inferred, and genetic relationships with other populations were evaluated using distance-based and graphical methods. **Results:** A total of 135 haplotypes were detected, with 88.8% being unique, yielding a haplotype diversity of 0.999. The most common haplogroups reflected both West African and European ancestry. Genetic distance analysis positioned the Cape Verdean population between African and European groups, supporting its intermediate and admixed genetic background. **Conclusions:** This study provides the first high-resolution Y-STR dataset for Cape Verdeans, contributing valuable reference data for forensic casework and population genetic studies. The results highlight the utility of extended Y-STR panels in admixed populations and underscore the need to enhance the representation of admixed populations in international forensic reference databases.

Keywords: population genetics; haplotype; Y chromosome; Cape Verde; Argus Y28; forensic genetics

Academic Editor: Mitchell Mark Holland

Received: 3 August 2025
Revised: 21 August 2025
Accepted: 23 August 2025
Published: 25 August 2025

Citation: Costa, R.; Fadoni, J.; Amorim, A.; Cainé, L. The Genetic Structure of Cape Verdean Population Revealed by Y-Chromosome STRs. *Genes* **2025**, *16*, 999. https://doi.org/10.3390/genes16090999

1. Introduction

Cape Verde is an archipelago located off the west coast of Africa in the Atlantic Ocean, discovered by Portuguese navigators in the 15th century. Permanent settlement began around 1460, driven by the arrival of a small number of European men, mainly Portuguese, and a much larger number of enslaved sub-Saharan Africans, primarily from the Senegambia region. Historical accounts also report the presence of Guanche individuals from the

Canary Islands among the first slaves. This founding population laid the groundwork for the unique demographic and genetic landscape observed in the archipelago today [1–5] (Figure 1).

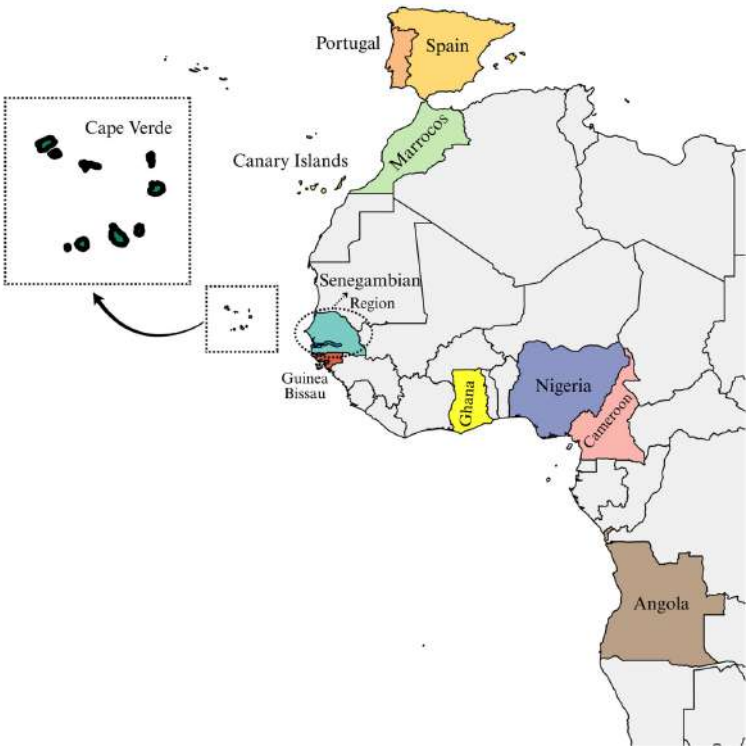

Figure 1. Geographic location of Cape Verde and the populations that contributed to the genetic landscape of the archipelago. Only countries with available Y-STR haplotype data are represented. The map was created using MapChart (https://www.mapchart.net, accessed on 2 July 2025).

Over the centuries, Cape Verde has experienced intense migratory flows, mainly involving sub-Saharan Africans and Europeans. This admixture has led to an asymmetric genetic pattern, with a predominance of African maternal lineages and European paternal lineages, as documented by previous genetic studies [6,7].

Beyond its unique demographic history, the Cape Verdean population maintains a strong historical and cultural connection with Portugal, reflected in significant and sustained migration [8,9]. According to the 2023 Migration and Asylum Report, Cape Verdeans are the third-largest foreign community legally residing in Portugal, with 48,885 residents, accounting for 4.7% of the total foreign population. However, this figure does not fully capture the true demographic weight of Cape Verdeans due to the increasing acquisition of Portuguese nationality and ongoing migration to other countries [10,11].

These migration patterns and historical admixture have important implications in forensic genetics. Genetic studies indicate that the Cape Verdean population exhibits substructure and diverges significantly from other sub-Saharan African groups. Notably, short tandem repeat (STR) marker analysis has revealed differences between the northern (Barlavento) and southern (Sotavento) islands, likely due to genetic drift and limited gene flow. Such differentiation underscores the importance of establishing population-specific reference data, particularly for forensic applications [2,6].

Y-chromosomal STRs (Y-STRs) are commonly used in forensic casework involving male individuals, such as sexual assault investigations and paternity testing [12]. Among them, rapidly mutating Y-STRs (RM Y-STRs), which have elevated mutation rates (greater than 10^{-2}), are especially valuable for distinguishing closely related men, as they provide higher resolution in pedigree analysis and in complex forensic cases [13,14].

The interpretation of Y-STR profiles relies on the availability of comprehensive population data. For this purpose, international databases like the Y Chromosome Haplotype Reference Database (YHRD) are essential tools that enable global comparison of haplotypes and support statistical evaluation of matches in forensic casework [15,16]. However, the Cape Verdean population remains under-represented in the YHRD, especially with respect to data generated from expanded multiplex kits such as the Argus Y-28, which includes both conventional and rapidly mutating loci, with only 117 haplotypes currently registered for 17 loci and none for 23 or more loci (https://yhrd.org (accessed on 20 August 2025)) [15,17–19].

Although previous studies have described the genetic characteristics of the Cape Verdean population, none have provided high-resolution data on allele and haplotype frequencies using expanded Y-STR panels, limiting their forensic applicability [2,6].

The aim of this study is to determine allele frequencies, haplotype diversity, and forensic parameters for 26 Y-STR loci in a sample of men residing in continental Portugal with Cape Verdean paternal ancestry, using an Argus Y-28 QS kit (QIAGEN, Hilden, Germany), and to contribute these data to the YHRD. This will enhance the representation of the Cape Verdean population in international forensic reference databases and support reliable Y-STR-based identification in medico-legal contexts.

2. Materials and Methods

2.1. Sample Collection

A total of 143 samples (135 dried blood spots and 8 buccal swabs) from unrelated men of Cape Verdean paternal ancestry were selected and analyzed. Paternal ancestry was assessed based on the individual's identification documents, which confirmed their place of birth. Information regarding paternal lineage was further obtained from the individual's self-reported account of their father's origin. The samples, collected between 2012 and 2022 from paternity cases available at INMLCF, I.P., were fully anonymized prior to analysis, with no personal information of the individuals being accessible, except for their paternal ancestry. They remained stored until their use in this study (2024–2025). Buccal swabs were stored at room temperature in paper envelopes in a dry and dark environment until DNA extraction.

This study received ethical approval from the Ethics Committee of the National Institute of Legal Medicine and Forensic Sciences. The approval was granted on 25 July 2024, under the reference code CE-25/2024. The use of these samples complied with Portuguese legislation (Law No. 12/2005, of January 26) and with INMLCF regulations, which permit the use of anonymized samples stored for more than two years after casework completion.

2.2. DNA Extraction and Y-STR Fragment Analysis

The saliva samples collected using buccal swabs were extracted with a PrepFiler Express™ Forensic DNA Extraction Kit (Thermo Fisher Scientific, Waltham, MA, USA), following the manufacturer's protocol. The blood samples on FTA® cards were processed by direct amplification, without any prior DNA extraction procedure. Amplification of 26 Y-STR loci was performed using an Investigator® Argus Y-28 QS PCR kit (QIAGEN, Hilden, Germany) on Applied Biosystems GeneAmp® PCR System 9700 thermal cyclers (Thermo Fisher Scientific, Waltham, MA, USA). Reactions followed the manufacturer's

protocol, with two modifications to optimize reagent use and prevent overamplification: the reaction volume was halved to 12.5 µL, and the number of PCR cycles was reduced to 28.

For more details, the PCR reaction was performed using the following program: pre-denaturation at 96 °C for 12 min; denaturation at 96 °C for 10 s, annealing at 61.5 °C for 85 s, and extension at 72 °C for 5 s, for 28 cycles; and final extension at 68 °C for 5 min and 60 °C for 5 min, with maintenance at 10 °C. Control DNA 9948 (QIAGEN, Hilden, Germany) was used as the positive control and ddH_2O as a negative control (without DNA) for each batch of Y-STR fragment analysis. The Investigator® Argus Y-28 QS PCR kit included 20 standard Y-STR loci (DYS389I, DYS391, DYS389II, DYS533, DYS390, DYS458, DYS393, DYS19, DYS437, DYS460, YGATAH4, DYS448, DYS439, DYS549, DYS438, DYS456, DYS643, DYS635, DYS385, DYS392), 6 rapidly mutating markers (DYS449, DYS481, DYS570, DYS576, DYS518, DYS627), and quality sensors QS1 and QS2. PCR products were prepared for capillary electrophoresis (CE) by adding 1 µL of each PCR product to a mixture of 12 µL Hi-Di™ Formamide (Thermo Fisher Scientific, Waltham, MA, USA) and 0.5 µL of DNA Size Standard 24plex (BTO) (QIAGEN, Hilden, Germany). The reaction plate was then heat denatured at 95 °C for 3 min and subsequently cooled using a thermal cycler set to 4 °C. Capillary electrophoresis was carried out on an ABI 3500 Genetic Analyzer (Thermo Fisher Scientific, Waltham, MA, USA). DNA typing was carried out following the manufacturer's protocol, using the provided locus panel, allele bins, and allele designations based on the supplied allelic ladder. The results were processed and analyzed using GeneMapper ID-X version 1.6 software (Thermo Fisher Scientific, Waltham, MA, USA), following the reference allelic ladder for genotype assignment. Reagent blanks were included alongside each batch of samples to verify the accuracy and reliability of the analyses. Additionally, samples that showed apparent null alleles, off-ladder (OL) peaks, or mutations in specific markers were reamplified to confirm the observed results.

2.3. Statistical Analyses

The number of distinct haplotypes, the frequency of unique haplotypes, discrimination capacity (DC), haplotype match probability (HMP), and haplotype diversity (HD) were calculated using a direct counting method in Microsoft Office Excel (version 16.94). The haplotype diversity (HD) was calculated applying the formula $HD = n(1 - \sum pi^2)/(n - 1)$, where n is the sample size, and pi is the frequency of the i-th haplotype [20]. The discrimination capacity (DC) was calculated as the ratio between the number of different haplotypes and the total number of individuals in the sample. The frequency of unique haplotypes was determined as the ratio between the number of unique haplotypes and the total number of individuals analyzed in the sample. The haplotype match probability (HMP) was calculated by summing the squared relative frequencies of all observed haplotypes ($HMP = \sum pi^2$), where pi represents the frequency of the i-th haplotype. STRAF software (STR Analysis for Forensics) (version 2.2.2) was used to calculate forensic parameters, including allele frequencies at each locus, gene diversity (GD), polymorphism information content (PIC), and power of discrimination (PD) [21]. Pairwise genetic distances (R_{ST}), estimated in Arlequin v3.5.2.2 software, were used to measure genetic distances between the samples from this study and the other compiled populations. For inter-population comparisons, the DYS389II allele length was calculated by subtracting the repeat count at DYS389I from that at DYS389II. For this analysis, only 17 Y-STR markers common to all population datasets were used. Using R statistical software version 4.0, a multidimensional scaling (MDS) plot was generated to visualize the relationships between populations based on the calculated matrix of pairwise genetic distances. Haplogroup prediction was performed using NevGen Y-DNA Haplogroup Predictor (https://www.nevgen.org, accessed on 16 June

2025) [22], assigning the haplogroup with the highest membership probability based on Y-STR haplotype.

3. Results

3.1. Analysis of Haplotype and Allele Diversity (and Forensic Parameters)

A total of 135 different haplotypes were obtained from 143 Cape Verdean samples, of which 127 (88.8%) were unique. The complete haplotype data for all 26 Y-STR loci are detailed in Supplementary Table S1. The allelic diversity observed across the loci ranged from 4 distinct alleles at DYS391 and DYS460 to 13 distinct alleles at DYS449, as shown in Supplementary Figure S1. Allele frequency data for all loci are available in Supplementary Table S2. The haplotype match probability (HMP) was calculated to be 0.0078, and the haplotype diversity (HD) was found to be 0.999. The corresponding allelic frequencies varied from 0.0070 to 0.6643. The lowest GD value was 0.497 in locus DYS391 and the highest was 0.859 in locus DYS627. Among the male Cape Verdean samples, biallelic patterns were observed. These included the DYS389II locus (alleles 29/30 and 28/29), DYS437 (alleles 14/15), DYS448 (alleles 19/20, 20/21, and 19/22), and DYS439 (alleles 10/11), suggesting possible duplication events or structural variations on the Y chromosome. Additionally, microvariant alleles such as 17.2 and 16.2 were detected at the DYS458 locus, further highlighting the genetic diversity and mutational dynamics within the studied population. No meaningful differences in profile completeness were observed between saliva and blood samples. Nonetheless, blood stains occasionally exhibited slightly higher signal intensities and a slightly elevated level of background peaks. These variations did not affect the overall quality or the interpretation of the profiles.

3.2. Haplogroup Distribution

Out of the 143 Y-STR profiles analyzed, haplogroups were predicted for 120 individuals with a probability greater than 50% using the NevGen Y-DNA Haplogroup Predictor (https://www.nevgen.org, accessed on 16 June 2025) [22]. Within this subset, a total of 13 distinct haplogroups were identified: E1a, E1b1a, E1b1b, E3a, G2a2b1, I2a2a, I2a2b, J1a2a1a2, J2b2a, L1a, L1b, R1b, and T. The most frequently observed haplogroups were R1b and E1b1a (Table 1). Interestingly, haplogroups R and E showed similar overall frequencies in the sample, with R1b representing 44% and the combined E sub-haplogroups accounting for 45%. This suggests a balanced genetic contribution from both European (R) and African (E) paternal lineages. The remaining haplogroups appeared in lower frequencies, reflecting additional genetic diversity from Middle Eastern and Asian origins.

Table 1. Y-chromosome haplogroup and ancestry frequencies in Cape Verde individuals, with predicted haplogroups (n = 120).

Haplogroup (H)	Sub-Haplogroup (SH)	Frequency SH (%)	Frequency H (%)	Ancestry
E	E1a	3%	45%	African
	E1b1a	28%		
	E1b1b	14%		
	E3a	1%		
L	L1a	1%	2%	Asian
	L1b	1%		

Table 1. *Cont.*

Haplogroup (H)	Sub-Haplogroup (SH)	Frequency SH (%)	Frequency H (%)	Ancestry
I	I2a2a	1%	4%	European
	I2a2b	3%		
R	R1b	44%	44%	
G	G2a2b1	1%	1%	Middle Eastern
J	J1a2a1a2	2%	4%	
	J2b2a	3%		
T	-	1%	1%	

3.3. Genetic Distance Analysis Based on R_{ST} Distances

To explore the genetic relationships between the Cape Verdean population and other African and European groups, genetic distances were estimated based on the sum of squared size differences (R_{ST}) between the haplotype distributions observed for 17 Y-STR loci. Comparative analysis could not be performed for the entire set of loci due to the absence of corresponding population data. However, for 17 loci in common (DYS389I, DYS391, DYS389II, DYS390, DYS458, DYS393, DYS19, DYS437, YGATAH4, DYS448, DYS439, DYS438, DYS456, DYS635, DYS392, DYS385a, DYS385b), comparisons were possible using published data from Guinea Bissau [23], Ghana [24], Nigeria [25], Cameroon [26], Angola [27], Spain [25], Portugal [28], and Morocco [29]. The pairwise R_{ST} values and corresponding *p*-values between the Cape Verdean population and the eight reference populations are displayed in Supplementary Table S3. The resulting distances were visualized through a multidimensional scaling (MDS) plot (Figure 2). The lowest genetic distances were observed with Cameroon (R_{ST} = 0.01230), Morocco (R_{ST} = 0.01747), Portugal (R_{ST} = 0.02098), and Angola (R_{ST} = 0.02100), indicating greater genetic affinity of Cape Verde with these African and Iberian populations. Although Spain showed a slightly negative RST value (–0.00347), this difference was not statistically significant (p = 0.88288). The highest genetic distances were recorded with Ghana (R_{ST} = 0.04314), Guinea-Bissau (R_{ST} = 0.03635), and Nigeria (R_{ST} = 0.03307). Among the pairwise comparisons, only some R_{ST} distances involving the Cape Verdean population were statistically significant ($p < 0.05$), as presented in Supplementary Table S3. The statistical analysis of R_{ST} distances revealed significant genetic differentiation between the Cape Verdean population and six of the eight populations compared. *p*-values below the conventional threshold of 0.05 were observed in the comparisons with Ghana (p = 0.00000), Nigeria (p = 0.02703), Cameroon (p = 0.01802), Angola (p = 0.01802), Portugal (p = 0.00000), and Morocco (p = 0.03604), indicating genetic divergence from these groups. In contrast, the comparisons with Guinea-Bissau (p = 0.0901) and Spain (p = 0.8829) did not reach statistical significance, suggesting no strong evidence of genetic differentiation from these populations. This may be consistent with a relatively closer genetic relationship, particularly with Iberian and some West African populations. The two-dimensional MDS plot provided an initial visualization of genetic relationships. In this representation, a cluster can be observed in which Cape Verde is positioned near Spain, Portugal, Ghana, and Guinea-Bissau, suggesting relative genetic proximity among these populations. However, the proportion of variance explained (PVE) was relatively low, with the first dimension accounting for 49.4% and the second for only 7.1%. These low cumulative values indicate that the 2D configuration does not accurately preserve the original genetic distances. Therefore, a three-dimensional MDS plot was generated (Supplementary Figure S2) to enhance the spatial resolution and improve the interpretation

of population structure. Together, these results highlight the importance of combining genetic distances with statistical support and appropriate dimensional representations when interpreting population affinities. This integrative approach contributes to a more accurate understanding of the paternal genetic structure and the intermediate position of the Cape Verdean population between African and European lineages.

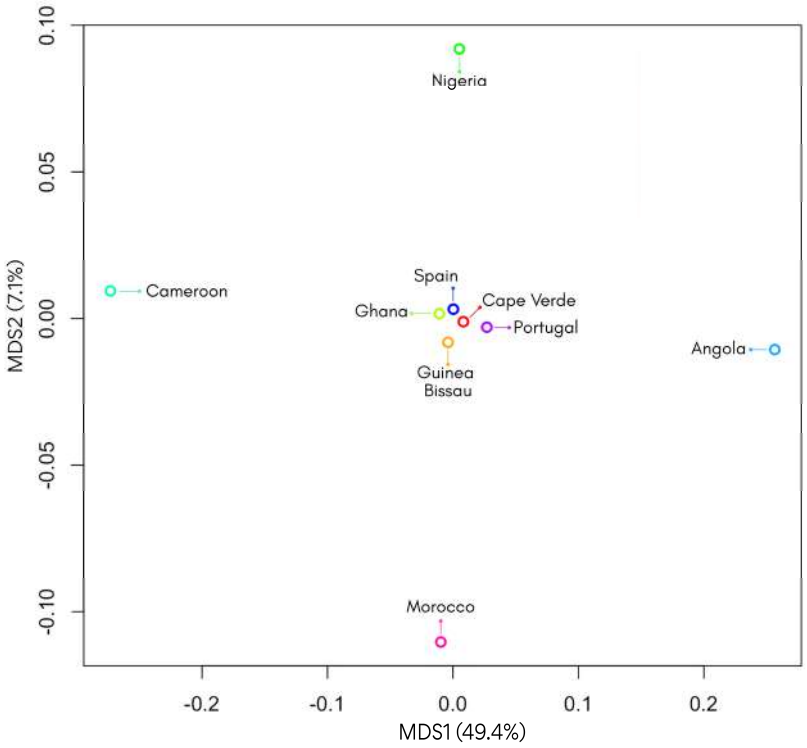

Figure 2. Multidimensional scaling (MDS) plot based on R_{ST} genetic distances among Cape Verde and reference populations. The analysis was performed using 17 shared Y-STR loci. The plot illustrates the genetic affinities between Cape Verde and other European and African populations, showing Cape Verde clustering closely with Guinea-Bissau, Spain, Ghana, and Portugal, while more distant populations include Angola, Cameroon, Nigeria, and Morocco. The percentages on the axes correspond to the proportion of variance explained (PVE) by each dimension.

4. Discussion

The genetic analysis of 143 unrelated Cape Verdean men using 26 Y-STR loci from an Argus Y-28 QS kit revealed a high level of haplotype diversity (HD = 0.999), with 135 different haplotypes and 88.8% being unique. These values reflect a high discrimination capacity, confirming the forensic utility of this multiplex system in admixed populations such as Cape Verde. The presence of only a few (n = 8) repeated haplotypes among unrelated individuals supports the system's ability to differentiate between male individuals, including those with potential close kinship, particularly due to the inclusion of rapidly mutating Y-STRs that enhance resolution. The practical value of rapidly mutating Y-STRs lies in their elevated mutation rates, which generate novel allelic differences even between close paternal relatives, such as brothers or father–son pairs. This mutational dynamic reduces the probability of observing identical haplotypes among relatives and therefore increases the power of Y-STR analysis to exclude or distinguish individuals in forensic contexts. Such resolution is particularly important in cases involving sexual assault with multiple male

contributors or in kinship testing scenarios where conventional Y-STR panels may yield indistinguishable profiles [14].

The detection of microvariant alleles and biallelic patterns in this study can be attributed to well-known mutational mechanisms affecting Y-STRs. Microvariants are most often generated by slippage events during DNA replication, where partial repeat units are inserted or deleted, producing intermediate allele sizes [13]. In contrast, duplicated alleles at certain loci may reflect structural rearrangements of the Y chromosome, such as segmental duplications or gene conversion events, which have been described in other populations [16,30]. Although relatively rare, these phenomena are relevant in forensic practice as they may influence allele designation and the interpretation of Y-STR profiles in casework.

From a forensic perspective, the availability of population-specific Y-STR reference data is essential for accurate statistical interpretation in casework. This need is particularly critical for admixed populations such as Cape Verdeans, whose paternal genetic profiles result from asymmetric historical admixture between European (predominantly R1b) and West African (mainly E1b1a and E1b1b) lineages [6,7]. In addition to the admixture events associated with the transatlantic slave trade, the observed genetic affinities with Iberian populations may also reflect more recent migratory flows during the 20th century, when thousands of Cape Verdeans settled in Portugal and other parts of Europe for economic and political reasons [10,11]. While the primary source of enslaved individuals brought to the archipelago during the transatlantic slave trade is historically associated with the Senegambian region [5], haplotypes from that area are under-represented or even absent in current Y-STR databases, such as YHRD [15,16]. This limits both comparative analyses and forensic resolution, especially when evaluating the statistical weight of evidence in investigations involving biogeographic ancestry inference. This absence of available reference data prevented the direct inclusion of populations from Senegal and Gambia in the RST-based comparisons, despite their well-documented historical relevance for the genetic makeup of Cape Verdeans. Additionally, many existing studies on African and admixed populations rely on panels with a reduced number of Y-STR loci or use datasets with limited haplotypic resolution, which restricts the robustness of comparative population analyses [31]. This limitation is particularly problematic for admixed populations or those with recent common paternal ancestors, where finer resolution is needed to distinguish close male relatives. In contrast, the use of 26 Y-STR loci in this study, including rapidly mutating markers, enabled the detection of rare allelic patterns such as microvariants and biallelic loci [13,14]. These markers are particularly informative in complex forensic scenarios, including sexual assault or kinship testing, where male individuals of the same paternal lineage must be distinguished. The genetic distinctiveness of Cape Verdeans also has practical implications in forensic casework. Y-STR profiles from individuals with Cape Verdean ancestry may be misclassified if interpreted against reference populations that do not adequately reflect their genetic background. This risk has been highlighted in studies emphasizing the forensic implications of misinterpretation in cases involving under-represented populations, particularly when Y-STR evidence is central to the investigation [32]. It should be noted that haplogroup prediction based on Y-STR profiles using NevGen is inherently probabilistic and particularly prone to misclassification in admixed populations such as Cape Verdeans. While SNP-based genotyping would provide more reliable assignments, the present results should be regarded as probabilistic inferences rather than definitive classifications. The RST-based comparisons performed in this study provide further insight into the genetic structure of the Cape Verdean population within a broader African–European context. The multidimensional scaling (MDS) analysis provided a visual approximation of genetic relationships among the studied populations. However,

due to the limited amount of genetic variance represented in the two-dimensional configuration, the spatial distribution observed should be interpreted with caution. To mitigate this limitation, a three-dimensional MDS plot was generated, offering a more accurate depiction of the underlying genetic structure. In this higher-resolution model, the Cape Verdean population appears to occupy an intermediate position between Iberian and West/Central African groups. While this distribution does not reflect strict geographic proximity, it is consistent with the historical and demographic processes that shaped the paternal genetic landscape of Cape Verde, including founder effects, asymmetric admixture, and successive migratory events. Although this study focuses on Cape Verdean individuals residing in Portugal, the findings may not be directly generalizable to the population living in the Cape Verde islands due to possible migration effects, integration processes, and genetic drift within the diaspora. Moreover, the sample size (n = 143) reflects the limited availability of Cape Verdean paternal lineage samples accessible within the studied diaspora context, which inevitably constrains the statistical power of some analyses. These findings contribute valuable high-resolution data to the global Y-STR reference framework, particularly for under-represented admixed populations, and underscore the need to broaden the geographic and demographic scope of future studies [25,30,33]. Future studies should aim to include a broader sampling of individuals from specific islands, as well as from Cape Verdean communities in other regions of the diaspora, to build a more comprehensive and representative genetic dataset for forensic and population genetic purposes [8,9].

5. Conclusions

This study presents the first comprehensive analysis of 26 Y-STR loci in a sample of Cape Verdean men residing in Portugal, revealing high haplotype diversity and excellent discriminatory power. These findings underscore the effectiveness of the Argus Y-28 QS kit in forensic applications involving admixed populations. The detection of unique haplotypes, including rare allelic variants such as microvariants and biallelic patterns, highlights the added value of using extended Y-STR panels, especially those including rapidly mutating loci, for resolving cases involving close paternal relatives. Given the genetic distinctiveness and under-representation of Cape Verdean populations in international forensic databases, this dataset strengthens the basis for more reliable statistical evaluations in forensic casework and ancestry inference involving individuals of Cape Verdean descent. To ensure fair and accurate interpretation of Y-STR evidence, it is essential to expand population-specific reference databases, particularly for admixed and minority populations. Future research should prioritize the inclusion of samples from different Cape Verdean islands and other diaspora communities to refine our understanding of their genetic structure and improve global forensic practices. In addition, our findings highlight the importance of expanding forensic reference databases to include samples from diaspora communities. When island-based sampling is limited, diaspora populations can provide valuable insights into the genetic structure of under-represented groups, thereby strengthening the reliability of Y-STR data for both forensic and population genetic applications. Incorporating such data will ultimately contribute to more equitable forensic investigations and support judicial systems with scientifically robust genetic evidence.

Supplementary Materials: The following supporting information can be downloaded at https://www.mdpi.com/article/10.3390/genes16090999/s1. Table S1. Haplotype database of Cape Verde population. Table S2. Allele frequency in Cape Verde population. Table S3. R_{ST} genetic distances and corresponding p-values between the Cape Verdean population and eight reference populations based on 17 shared Y-STR loci. Figure S1. Allele frequency distributions of 26 Y-STR loci in the Cape Verde population. The x-axis represents allele values per locus; the y-axis indicates the observed allele

frequencies. Figure S2. Interactive 3D MDS plot based on RST genetic distances between the Cape Verdean population and eight reference populations using 17 shared Y-STR loci.

Author Contributions: Conceptualization: L.C. and A.A.; methodology, R.C. and J.F.; writing—original draft preparation, R.C. and J.F.; writing—review and editing, L.C. and A.A.; supervision, J.F., L.C. and A.A.; funding acquisition, L.C. and A.A. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Instituto Nacional de Medicina Legal e Ciências Forenses, Portugal.

Institutional Review Board Statement: This study received ethical approval from the Ethics Committee of the National Institute of Legal Medicine and Forensic Sciences. The approval was granted on 25 July 2024, under the reference code CE-25/2024.

Informed Consent Statement: Informed consent was waived due to the retrospective nature of the study and use of anonymized data.

Data Availability Statement: The original contributions presented in this study are included in this article/Supplementary Material. Further inquiries can be directed to the corresponding author(s).

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

CE	Capillary Electrophoresis
DC	Discrimination Capacity
GD	Gene Diversity
HD	Haplotype Diversity
HMP	Haplotype Match Probability
INMLCF, I.P.	Instituto Nacional de Medicina Legal e Ciências Forenses, Instituto Público
MDS	Multidimensional Scaling
NGS	Next-Generation Sequencing
OL	Off-Ladder
PCR	Polymerase Chain Reaction
PIC	Polymorphism Information Content
PVE	Proportion of Variance Explained
QS	Quality Sensor
RM Y-STRs	Rapidly Mutating Y-Chromosomal Short Tandem Repeats
STR	Short Tandem Repeat
YHRD	Y Chromosome Haplotype Reference Database
Y-STR	Y-Chromosomal Short Tandem Repeat

References

1. Gonçalves, R.; Fernandes, A.T.; Brehm, A. Cabo Verde islands: Different maternal and paternal heritage testifies the nature of its first settlers. *Int. Congr. Ser.* **2004**, *1261*, 372–373. [CrossRef]
2. Fernandes, A.T.; Velosa, R.; Jesus, J.; Carracedo, A.; Brehm, A. Genetic differentiation of the Cabo Verde archipelago population analysed by STR polymorphisms. *Ann. Hum. Genet.* **2003**, *67*, 340–347. [CrossRef] [PubMed]
3. Laurent, R.; Szpiech, Z.A.; da Costa, S.S.; Thouzeau, V.; Fortes-Lima, C.A.; Dessarps-Freichey, F.; Lémée, L.; Utgé, J.; Rosenberg, N.A.; Baptista, M.; et al. A genetic and linguistic analysis of the admixture histories of the islands of Cabo Verde. *eLife* **2023**, *12*, e79827. [CrossRef]
4. Parra, E.J.; Ribeiro, J.C.; Caeiro, J.L.; Riveiro, A. Genetic structure of the population of Cabo Verde (west Africa): Evidence of substantial European admixture. *Am. J. Phys. Anthr.* **1995**, *97*, 381–389. [CrossRef]
5. Brehm, A.; Pereira, L.; Bandelt, H.J.; Prata, M.J.; Amorim, A. Mitochondrial portrait of the Cabo Verde archipelago: The Senegambian outpost of Atlantic slave trade. *Ann. Hum. Genet.* **2002**, *66*, 49–60. [CrossRef]
6. Beleza, S.; Campos, J.; Lopes, J.; Araújo, I.I.; Hoppfer Almada, A.; Correia e Silva, A.; Parra, E.J.; Rocha, J. The admixture structure and genetic variation of the archipelago of Cape Verde and its implications for admixture mapping studies. *PLoS ONE* **2012**, *7*, e51103. [CrossRef]

7. Gonçalves, R.; Rosa, A.; Freitas, A.; Fernandes, A.; Kivisild, T.; Villems, R.; Brehm, A. Y-chromosome lineages in Cabo Verde Islands witness the diverse geographic origin of its first male settlers. *Hum. Genet.* **2003**, *113*, 467–472. [CrossRef]
8. Resende, A.; Amorim, A.; da Silva, C.V.; Ribeiro, T.; Porto, M.J.; Costa Santos, J.; Afonso Costa, H. Study of genetic markers of CODIS and ESS systems in a population of individuals from Cabo Verde living in Lisboa. *Int. J. Leg. Med.* **2017**, *131*, 119–121. [CrossRef] [PubMed]
9. Amorim, A.; Marques-Santos, R.; Vieira-Silva, C.; Afonso-Costa, H.; Espinheira, R.; Ferreira-Gomes, P.; Costa-Santos, J. Genetic portrait of a native population of Cabo Verde living in Lisboa. *Forensic Sci. Int. Genet.* **2012**, *6*, e166–e169. [CrossRef]
10. Gabinete de Estratégia e Estudos. População estrangeira com estatuto legal de residente em Portugal—Cabo Verde. 2023. Available online: https://www.gee.gov.pt/pt/lista-publicacoes/estatisticas-de-imigrantes-em-portugal-por-nacionalidade/paises/Cabo%20Verde/3947-populacao-estrangeira-com-estatuto-legal-de-residente-em-portugal-cabo-verde/file (accessed on 5 May 2025).
11. *AIMA I.P.—DPEE—Direção de Planeamento, Estudos e Estatistica. Relatório de Migrações e Asilo 2023*; AIMA: London, UK, 2024.
12. Kayser, M. Forensic use of Y-chromosome DNA: A general overview. *Hum. Genet.* **2017**, *136*, 621–635. [CrossRef] [PubMed]
13. Ballantyne, K.N.; Goedbloed, M.; Fang, R.; Schaap, O.; Lao, O.; Wollstein, A.; Choi, Y.; van Duijn, K.; Vermeulen, M.; Brauer, S.; et al. Mutability of Y-chromosomal microsatellites: Rates, characteristics, molecular bases, and forensic implications. *Am. J. Hum. Genet.* **2010**, *87*, 341–353. [CrossRef] [PubMed]
14. Ballantyne, K.N.; Keerl, V.; Wollstein, A.; Choi, Y.; Zuniga, S.B.; Ralf, A.; Vermeulen, M.; de Knijff, P.; Kayser, M. A new future of forensic Y-chromosome analysis: Rapidly mutating Y-STRs for differentiating male relatives and paternal lineages. *Forensic Sci. Int. Genet.* **2012**, *6*, 208–218. [CrossRef]
15. Y-Chromosome Haplotype Reference Database (YHRD). Available online: https://yhrd.org (accessed on 5 May 2025).
16. Roewer, L. The Y-Short Tandem Repeat Haplotype Reference Database (YHRD) and Male Population Stratification in Europe—Impact on Forensic Genetics. *Forensic Sci. Rev.* **2003**, *15*, 165–172. [PubMed]
17. QIAGEN. *Investigator® Argus Y-28 QS Handbook*; QIAGEN: Hilden, Germany, 2022. Available online: https://www.qiagen.com/in/resources/download.aspx?id=cf7ae42a-72bc-4d3d-812d-684d0625a179&lang=en (accessed on 5 May 2025).
18. Steffen, C.R.; Huszar, T.I.; Borsuk, L.A.; Vallone, P.M.; Gettings, K.B. A multi-dimensional evaluation of the 'NIST 1032' sample set across four forensic Y-STR multiplexes. *Forensic Sci. Int. Genet.* **2022**, *57*, 102655. [CrossRef]
19. Haarkötter, C.; Isabel Medina-Lozano, M.; Vinueza-Espinosa, D.C.; Saiz, M.; Gálvez, X.; Carlos Álvarez, J.; Antonio Lorente, J. Evaluating the efficacy of three Y-STRs commercial kits in degraded skeletal remains. *Sci. Justice* **2024**, *64*, 543–548. [CrossRef]
20. Nei, M.; Tajima, F. Genetic drift and estimation of effective population size. *Genetics* **1981**, *98*, 625–640. [CrossRef]
21. Gouy, A.; Zieger, M. STRAF—A convenient online tool for STR data evaluation in forensic genetics. *Forensic Sci. Int. Genet.* **2017**, *30*, 148–151. [CrossRef]
22. Cetkovic Gentula, M.; Nevski, A. NEVGEN Y-DNA Haplogroup Predictor. Available online: https://www.nevgen.org (accessed on 16 June 2025).
23. Carvalho, M.; Brito, P.; Bento, A.M.; Gomes, V.; Antunes, H.; Costa, H.A.; Lopes, V.; Serra, A.; Balsa, F.; Andrade, L.; et al. Paternal and maternal lineages in Guinea-Bissau population. *Forensic Sci. Int. Genet.* **2011**, *5*, 114–116. [CrossRef] [PubMed]
24. Kofi, A.E.; Hakim, H.M.; Khan, H.O.; Ismail, S.A.; Ghansah, A.; David, A.A.; Mat, N.F.C.; Chambers, G.K.; Edinur, H.A. Population data of 23 Y chromosome STR loci for the five major human subpopulations of Ghana. *Int. J. Leg. Med.* **2020**, *134*, 1313–1315. [CrossRef]
25. Purps, J.; Siegert, S.; Willuweit, S.; Nagy, M.; Alves, C.; Salazar, R.; Angustia, S.M.; Santos, L.H.; Anslinger, K.; Bayer, B.; et al. A global analysis of Y-chromosomal haplotype diversity for 23 STR loci. *Forensic Sci. Int. Genet.* **2014**, *12*, 12–23. [CrossRef]
26. Della Rocca, C.; Cannone, F.; D'Atanasio, E.; Bonito, M.; Anagnostou, P.; Russo, G.; Barni, F.; Alladio, E.; Destro-Bisol, G.; Trombetta, B.; et al. Ethnic fragmentation and degree of urbanization strongly affect the discrimination power of Y-STR haplotypes in central Sahel. *Forensic Sci. Int. Genet.* **2020**, *49*, 102374. [CrossRef]
27. Palha, T.; Gusmão, L.; Ribeiro-Rodrigues, E.; Guerreiro, J.F.; Ribeiro-Dos-Santos, A.; Santos, S. Disclosing the genetic structure of Brazil through analysis of male lineages with highly discriminating haplotypes. *PLoS ONE* **2012**, *7*, e40007. [CrossRef]
28. Pontes, M.L.; Cainé, L.; Abrantes, D.; Lima, G.; Pinheiro, M.F. Allele frequencies and population data for 17 Y-STR loci (AmpFℓSTR® Y-filer™) in a Northern Portuguese population sample. *Forensic Sci. Int.* **2007**, *170*, 62–67. [CrossRef]
29. Aboukhalid, R.; Bouabdellah, M.; Abbassi, M.; Bentayebi, K.; Elmzibri, M.; Squalli, D.; Amzazi, S. Haplotype frequencies for 17 Y-STR loci (AmpFlSTRY-filer) in a Moroccan population sample. *Forensic Sci. Int. Genet.* **2010**, *4*, e73–e74. [CrossRef]
30. Gusmão, L.; Butler, J.M.; Carracedo, A.; Gill, P.; Kayser, M.; Mayr, W.R.; Morling, N.; Prinz, M.; Roewer, L.; Tyler-Smith, C.; et al. DNA Commission of the International Society of Forensic Genetics (ISFG): An update of the recommendations on the use of Y-STRs in forensic analysis. *Int. J. Leg. Med.* **2006**, *120*, 191–200. [CrossRef]
31. Rosa, A.; Ornelas, C.; Brehm, A.; Villems, R. Population data on 11 Y-chromosome STRs from Guiné-Bissau. *Forensic Sci. Int.* **2006**, *157*, 210–217. [CrossRef]

32. Costa, R.; Fadoni, J.; Amorim, A.; Cainé, L. Y-STR Databases-Application in Sexual Crimes. *Genes* **2025**, *16*, 484. [CrossRef]
33. Excoffier, L.; Laval, G.; Schneider, S. Arlequin (version 3.0): An integrated software package for population genetics data analysis. *Evol. Bioinform. Online* **2007**, *1*, 47–50. [CrossRef]

CHAPTER V – Discussion

This dissertation integrates two complementary contributions to the field of forensic genetics, both centred on the analysis of Y-STR markers. The first article explores the role of Y-STRs in sexual crime investigations, emphasizing their usefulness in identifying male profiles in complex biological mixtures and the importance of rapidly mutating Y-STRs (RM Y-STRs) in differentiating related male individuals. The second article presents original population data on Cape Verdean males residing in Portugal, highlighting the need to include underrepresented populations in reference databases such as the YHRD.

The first study highlights the relevance of Y-STRs in sensitive forensic contexts, particularly in sexual assault cases, where male DNA is often present in low quantities and masked by female DNA. While standard Y-STRs cannot distinguish between paternally related males, the incorporation of RM Y-STRs significantly enhances discrimination power, especially in cases involving related suspects or multiple male contributors [5, 6].

The second, population-based study revealed a high level of haplotype diversity among Cape Verdean males, with 88.8% unique haplotypes and a haplotype diversity (HD) of 0.999. These findings reflect the admixed (African and European) ancestry of this population, consistent with the settlement history of Cape Verde. The balanced presence of haplogroups E1b1a (African origin) and R1b (European origin) supports previous studies describing the asymmetric genetic structure of Cape Verdeans—predominantly European paternal and African maternal lineages [10-12].

The genetic distance analysis, based on 17 shared Y-STR loci, revealed that the Cape Verdean population exhibits affinities with both West and Central African groups as well as Iberian populations, occupying an intermediate position in the three-dimensional multidimensional scaling (MDS) model. While the results support a mixed genetic structure consistent with the historical processes of migration and admixture, several pairwise comparisons yielded statistically significant differences, highlighting the complexity of the population's paternal composition. These findings underscore the importance of incorporating admixed populations, such as Cape Verdeans, into international reference databases like YHRD, particularly in forensic contexts where adequate population representation is essential for the accurate interpretation of Y-STR evidence [7, 8].

Both articles converge on the urgent need to improve the representation of African and Afro-descendant populations in forensic genetic databases. Underrepresentation compromises

the statistical robustness of match interpretations and may negatively impact the reliability of genetic evidence in judicial settings, raising ethical and legal concerns. This challenge has been increasingly acknowledged by the forensic community, which calls for more equitable strategies in forensic DNA interpretation [9].

The integration of methodological validation with population analysis adds an important dimension to this dissertation. By validating the Investigator® Argus Y-28 QS kit under the specific conditions of the SGBF-N laboratory, the study ensured that the subsequent population data were generated using a protocol demonstrated to be robust and reproducible. This methodological step not only strengthens the reliability of the Cape Verdean haplotypic data but also highlights the broader principle, emphasized by international guidelines, that population studies must be grounded on validated laboratory methods. In this sense, the validation presented here bridges the gap between laboratory robustness and population representativeness, reinforcing the forensic applicability of the findings [14, 16].

In summary, this dissertation reinforces the importance of a critical and inclusive approach to the application of Y-STRs in forensic genetics—one that combines technological innovation (e.g., RM Y-STRs) with the ethical imperative of ensuring adequate representation of all populations in forensic reference databases.

CHAPTER VI – Future Perspectives

The findings presented in this dissertation point to several opportunities for further research and practical developments in forensic genetics. One of the most pressing priorities concerns the expansion of population reference databases, particularly for groups that remain underrepresented. The inclusion of Cape Verdean Y-STR haplotypes in the YHRD was a significant contribution, but future efforts should aim to collect additional data from a broader and more diverse sample of individuals across the Cape Verde archipelago. Given the historical and geographic diversity of the islands, such an initiative would help capture potential substructure within the population, thereby increasing the statistical power and reliability of forensic comparisons.

Alongside population data expansion, technological advances in DNA analysis are likely to shape the future of Y-chromosome research. Next-generation sequencing (NGS) platforms, which enable the simultaneous examination of Y-STRs and Y-SNPs, offer enhanced resolution for differentiating male lineages and for inferring ancestral origins. While this approach has not yet become routine in most forensic laboratories, its integration could greatly strengthen the ability to resolve complex cases, particularly those involving related male suspects or mixed DNA samples. Research into the application of NGS in real casework, especially in combination with expanded and representative databases, remains an area of great potential.

Equally important are the broader ethical and societal dimensions of forensic genetics. The persistent underrepresentation of African and Afro-descendant populations in forensic databases raises legitimate concerns about equity and scientific fairness. As forensic genetic tools continue to evolve, it is essential that data collection practices and database policies reflect principles of inclusiveness and transparency. Future studies should engage with these questions more directly, promoting guidelines that support both scientific rigor and social responsibility. The integration of population diversity, technological innovation, and ethical foresight will be critical to advancing a forensic genetic system that is both effective and just.

CHAPTER VII – Conclusions

This dissertation explored the application of Y-chromosome STR markers in forensic genetics through two complementary approaches: a critical review of the literature and a population study involving Cape Verdean males. The review highlighted the relevance of Y-STRs in sexual crime investigations, particularly in the identification of male contributors in complex DNA mixtures, and underscored the added value of rapidly mutating Y-STRs in distinguishing between closely related individuals. It also drew attention to the limitations posed by the underrepresentation of certain populations in global reference databases.

The population study addressed this issue by generating the first comprehensive dataset of 26 Y-STR loci, including rapidly mutating markers, for Cape Verdean individuals residing in Portugal. The results revealed high haplotype diversity and a mixed African–European paternal ancestry, consistent with the population's historical background. The inclusion of these data in the YHRD strengthens the statistical basis for forensic analysis involving this population and contributes to broader efforts toward database inclusivity.

Together, these studies reinforce the importance of combining technological advancement with representative population data to ensure that forensic genetics remains both scientifically rigorous and socially equitable.

Bibliografia

1. Roewer, L., *Y chromosome STR typing in crime casework.* Forensic Sci Med Pathol, 2009. **5**(2): p. 77-84.
2. Kayser, M., *Forensic use of Y-chromosome DNA: a general overview.* Hum Genet, 2017. **136**(5): p. 621-635.
3. Jobling, M.A., A. Pandya, and C. Tyler-Smith, *The Y chromosome in forensic analysis and paternity testing.* Int J Legal Med, 1997. **110**(3): p. 118-24.
4. Syndercombe Court, D., *The Y chromosome and its use in forensic DNA analysis.* Emerg Top Life Sci, 2021. **5**(3): p. 427-441.
5. Alghafri, R., et al., *Rapidly mutating Y-STR analyses of compromised forensic samples.* Int J Legal Med, 2018. **132**(2): p. 397-403.
6. Ballantyne, K.N., et al., *A new future of forensic Y-chromosome analysis: rapidly mutating Y-STRs for differentiating male relatives and paternal lineages.* Forensic Sci Int Genet, 2012. **6**(2): p. 208-18.
7. *Y-Chromosome Haplotype Reference Database (YHRD)*. 20 November 2024]; Available from: https://yhrd.org.
8. Willuweit, S. and L. Roewer, *The new Y Chromosome Haplotype Reference Database.* Forensic Sci Int Genet, 2015. **15**: p. 43-8.
9. Ralf, A., M. Zieger, and M. Kayser, *Considerations on expanding criminal offender DNA databases with Y-STR profiles.* J Law Biosci, 2024. **11**(2): p. lsae017.
10. Parra, E.J., et al., *Genetic structure of the population of Cabo Verde (west Africa): evidence of substantial European admixture.* Am J Phys Anthropol, 1995. **97**(4): p. 381-9.
11. Fernandes, A.T., et al., *Genetic differentiation of the Cabo Verde archipelago population analysed by STR polymorphisms.* Ann Hum Genet, 2003. **67**(Pt 4): p. 340-7.
12. Beleza, S., et al., *The admixture structure and genetic variation of the archipelago of Cape Verde and its implications for admixture mapping studies.* PLoS One, 2012. **7**(11): p. e51103.
13. AIMA I.P. - DPEE – Direção de Planeamento, E.e.E., *Relatório de Migrações e Asilo 2023*. 2024.
14. Roewer, L., et al., *DNA commission of the International Society of Forensic Genetics (ISFG): Recommendations on the interpretation of Y-STR results in forensic analysis.* Forensic Sci Int Genet, 2020. **48**: p. 102308.
15. Costa, R., et al., *Y-STR Databases-Application in Sexual Crimes.* Genes (Basel), 2025. **16**(5).
16. (SWGDAM), S.W.G.o.D.A.M., *SWGDAM Interpretation Guidelines for Y-Chromosome STR Typing by Forensic DNA Laboratories*. 2022.
17. Kaiser, H., *Report for analytical chemists. II. Quantitation in elemental analysis.* Analytical Chemistry, 1970. **42**(4): p. 26A-59a.