MSc

2.º CICLO

FCUP
2016

U.PORTO

Genetic analysis of haplotypic data for 17 Y-chromosome short tandem repeat loci in the population of São Paulo, Brazil.

Jennifer Nascimento Fadoni

FC

U.PORTO

FACULDADE DE CIÊNCIAS
UNIVERSIDADE DO PORTO

Genetic analysis of haplotypic data for 17 Y-chromosome short tandem repeat loci in the population of São Paulo, Brazil.

Jennifer Nascimento Fadoni
Master in Forensic Genetics
Department of Biology

2016

Supervisor
Doctor Luis Álvarez Fernández, Institute for Investigation and Innovation in Health (i3S), Institute of Molecular Pathology and Immunology of the University of Porto (IPATIMUP)

Co-Supervisors
Doctor Verónica Gomes, Researcher, i3S/IPATIMUP
Doctor Nádia Pinto, Researcher, i3S/IPATIMUP, Centre of Mathematics of the University of Porto (CMUP)

Todas as correções determinadas pelo júri, e só essas, foram efetuadas.

O Presidente do Júri,

Porto, _____/_____/_____

Dissertação de candidatura ao grau de mestre em Genética Forense submetida à Faculdade de Ciências da Universidade do Porto.

O presente trabalho foi desenvolvido no Instituto de Investigação e Inovação em Saúde sob orientação dos Doutores Luis Álvarez Fernández, Verónica Gomes e Nádia Pinto.

Dissertation for applying to a Master's degree in Forensic Genetics, submitted to the Faculty of Sciences of the University of Porto.

The present work was developed at the Institute for Research and Innovation in Health under the scientific supervision of Doctors Luis Álvarez Fernández, Verónica Gomes and Nádia Pinto.

AGRADECIMENTOS

Começo por dedicar essa dissertação aos meus pais, por me ensinarem que o conhecimento é o bem mais valioso que podemos ter, pois nunca pode nos ser retirado. É graças a eles que hoje estou aqui, esse sem seu apoio incondicional, amor, paciência e suporte financeiro, este trabalho não teria sido realizado.

Agradeço ao meu ao orientador Luís, e às minhas co-orientadoras Verónica e Nádia pela partilha de conhecimentos, por todas as sugestões e por estarem sempre disponíveis e prontos para ajudar. Levo-vos como exemplos de pessoas e pesquisadores; expresso aqui a minha admiração pelo vosso comprometimento com o trabalho e conduta profissional.

Ao professor Amorim, pela oportunidade de realizar o meu sonho de frequentar este Mestrado. Agradeço-lhe por ter tornado possível este crescimento pessoal e científico. Obrigada por estar sempre disposto a ajudar.

Obrigada à equipa do IMESC, por toda ajuda que me deram nestes últimos anos e por disponibilizarem as amostras para a realização deste trabalho.

Obrigada aos meus amigos e companheiros de curso por tornar esta jornada ainda mais enriquecedora e por em ajudarem todas as vezes que precisei.

A todos os professores que ao longo do meu percurso académico, contribuíram para que eu chegasse onde estou hoje e que me permitirão continuar em frente.

Eu não teria sido capaz de concluir este trabalho sem todas essas pessoas. Muito obrigada pelo apoio, energia, conhecimento e sabedoria.

ABSTRACT

Brazil is the world's fifth-largest country by both area and population. Its territory is divided in five geographical regions, which in turn are divided into 26 states plus Federal District. Brazil harbor a highly diverse population that resulted from different degrees and modes of admixture between Native Americans, Europeans and Africans.

The human Y-chromosome is male specific and genetic markers locating on the non-recombining region of this chromosome have a patrilineal inheritance mode. They are transmitted unchanged (except for the mutations), what makes Y-chromosome a high informative tool for forensic studies also for tracing human migration and evolution through male lineages. While different kind of Y-polymorphisms were proved to be useful in routine forensic casework, at present, Y-STRs are the most used not just in forensic cases but also in Population Genetics and Human Evolution studies due to their high levels of diversity when compared to other polymorphisms.

In this study, were investigated the haplotype composition of Y-chromosome in a sample of 417 genetically unrelated individuals residing in São Paulo state, Southeast of Brazil. Using the Yfiler™ kit (Applied Biosystems), 17 Y-STR were evaluated (DYS19, DYS385a/b, DYS389I, DYS389II, DYS390, DYS391, DYS392, DYS393, DYS437, DYS438, DYS439, DYS448, DYS456, DYS458, DYS635 and Y GATA H4). A total of 410 different haplotypes were found, among which 7 were shared between pair of individuals. The probability of two random individuals showing identical haplotype was 0.25% and the overall Haplotype Diversity was 0.9999. The average Gene Diversity at the 17 Y-STR loci was 0.66295; DYS458 showed the greatest degree of diversity (0.79523) and DYS389I was the locus characterized by the lowest value (0.54636). These results did not differ from previous research results about the São Paulo state population. In addition, the haplogroups for all the samples were inferred throughout the Haplogroup Predictor program, being European the major genetic contribution for São Paulo population (46.52% belong to haplogroup R), followed by the African (20.86% belong to haplogroup E).

We analyzed the genetic relationship between the data from this study and published data from other Brazilian regions. The population of Pernambuco showed the highest value (6.67%) of shared haplotypes with the samples of the present study. Moreover, pairwise genetic distances were estimated (based on R_{ST}) between the haplotype distributions observed for the 17 Y-STRs in our samples and the 20 Brazilian admixed population samples.

The highest genetic distance was found between our samples and the samples from Minas Gerais, Amazonas and Santa Catarina.

Finally, population comparisons with European (Portuguese, Italian, Spanish and German) and African populations (Mozambican and Angolan) were undertaken through pairwise genetic distances (based on R_{ST}), and illustrated through MDS (based on F_{ST}) values. São Paulo populations seems to have a Portuguese and Angolan background as no significant genetic distances were found between them.

Keywords: Y-STR, Y-Chromosome, Yfiler, São Paulo, Population Genetics

RESUMO

Brasil é o quinto maior país do mundo, tanto em área quanto em população. Seu território é dividido geograficamente em cinco regiões, que por sua vez são divididas em 26 estados, além do Distrito Federal. A população do Brasil é grandemente diversa, resultado de diferentes graus e modos de miscigenação entre Nativo-Americanos, Europeus e Africanos.

O cromossoma Y humano é exclusivo aos homens e marcadores genéticos localizados em sua região não recombinante têm um modo de herança patrilínea. Eles são transmitidos inalteradamente (exceto quando há uma mutações), o que faz do cromossoma Y uma ferramenta informativa de análise, através das linhagens masculinas, em casos forenses, em estudos de migração humana e evolução. Apesar de diferentes tipos de polimorfismos do cromossoma Y serem comprovadamente úteis nos casos forenses, atualmente, os Y-STRs são os mais utilizados não apenas nesses campo mas também em Genética Populacional e Evolução Humana, devido às suas elevadas taxas de diversidade quando comparados a outros polimorfismos.

No presente estudo, foi investigada a composição haplotípica do cromossoma Y em uma amostra de 417 indivíduos não relacionados geneticamente, residentes no Estado de São Paulo, sudeste do Brasil. Através do kit Yfiler™ kit (Applied Biosystems), 17 Y-STR foram analisados (DYS19, DYS385a/b, DYS389I, DYS389II, DYS390, DYS391, DYS392, DYS393, DYS437, DYS438, DYS439, DYS448, DYS456, DYS458, DYS635 and Y GATA H4). Foram encontrados 410 haplótipos distintos, dentro dos quais 07 são partilhados entre pares de indivíduos. A probabilidade de dois indivíduos ao acaso apresentarem o mesmo haplótipo foi de 0,25% e a média da Diversidade Haplotípica foi 0.9999. A média da Diversidade Genética para os 17 Y-STR loci analisados foi 0.66295; DYS458 apresentou o maior valor de diversidade (0.79523) e DYS389I foi o locus que apresentou o valor mais baixo (0.54636). Esses resultados não diferem dos resultados de estudos sobre a população de São Paulo realizados anteriormente. Através do Haplogroup Predictor, foram inferidos os haplogrupos de todas as amostras e a maior contribuição genética para a população de São Paulo mostrou ser a Europeia, seguida da Africana, visto que 46,52% dos haplótipos pertencem ao haplogrupo R, um haplogrupo Europeu, e 20,86% das amostras pertencem ao haplogrupo E, um haplogrupo Africano.

Foram analisadas as relações genéticas entre os dados do presente estudo e dados publicados sobre estudos de populações de outras regiões do Brasil. A população de

Pernambuco mostrou a percentagem mais alta (6,67%) de haplótipos compartilhados com as amostras do presente estudo. As distâncias genéticas “Pairwise” (com base em R_{ST}) foram estimadas entre a distribuição haplotípica observada para os 17 Y-STRs das amostras do presente estudo e entre amostras de 20 populações Brasileiras e as amostras da população de São Paulo apresentou maior distância genética com as populações de Minas Gerais, Amazonas e Santa Catarina.

Foram feitas comparações com populações Europeias (Portuguesa, Italiana, Espanhola e Alemã) e Africanas (Moçambicana e Angolana), através das distâncias genéticas “pairwise” (baseadas em R_{ST}) e ilustradas através de uma representação MDS (baseada nos valores F_{ST}). A população do estado de São Paulo mostrou ter uma herança genética Portuguesa e Angolana visto que não foram encontradas distâncias genéticas significativas com essas populações.

Palavras-chave: Y-STR, Cromossoma Y, Yfiler, São Paulo, Genética Populacional

Contents

Tables Index

Figures Index

Abbreviations and symbols

bp – Base Pair

DNA – Deoxyribonucleic Acid

GDB – Genome Database

HLA – Human Leukocyte Antigen

HD – Haplotype Diversity

Indels – Insertions and delections

ISFG – International Society for Forensic Genetics

MDS – Multidimensinal Scaling

MSY – Male-specific Region

mtDNA – Mitochondrial Deoxyribonucleic Acid

ng – Nanogram

NRY – Non Recombining Y

PAR – PseudoAutosomal Region

PCR – Polymerase Chain Reaction

pg – Picogram

RFLP – Restriction Fragment Length Polymorphism

SNP – Single Nucleotide Polymorphism

STR – Short Tandem Repeat

SWGDAM – Scientific Working Group on DNA Analysis Methods

VNTRs – Variable Number of Tandem Repeats

YHRD – Y-STR Haplotype Reference Database

Y-STR – Short Tandem Repeat on the Y Chromosome

µL – Microliter

1. INTRODUCTION

1.1. Population Genetics and Phylogeography

The theoretical basis of Population genetics was developed by Ronald A. Fisher [1], J.B.S. Haldane [2] and Sewall Wright [3] in the 1920s and 1930s after the widespread acceptance of Mendelian genetics. The primary goals of population genetics are to understand the factors determining evolutionary change and stasis, and the amount and pattern of genetic variation within and between populations [4]. This variation is generated by mutations and shaped by natural selection, genetic drift and gene flow [5], which are highly influenced by historical or demographic circumstances [6]. These four evolutionary forces are reflected in patterns of diversity, measured by the numbers of different alleles at a gene locus, the frequencies of each allele, and the interrelatedness of each allele to the others present at the same time. One of the main goals of population genetics is to understand which are the driving forces responsible for variation over time, and to interpret how occurred the accumulation of allelic diversity or, by other words, which were the events responsible for those patterns [5, 7]. The study of human population genetics (also known as anthropological genetics) allows explaining the causes of human diversity in the world today and the evolutionary history that has generated this diversity. Studies of anthropological genetics include efforts to describe population structure, to reconstruct population history, and to understand adaptation to local environments [5]. Studies of population genetics in human populations are important to understand human evolution, the spread of modern humans, migration patterns in certain geographic areas and differentiation in single populations.

Phylogeography is defined as the study of processes that may be responsible for the contemporary geographic distributions of genetic lineages. Its terminology was first described in 1987 when Avise and colleagues [8], with the aim to unite evolutionary biologists in the dissimilar fields of Phylogenetics and Population Genetics, proposing the integration of these areas for investigating the connection between micro- and macroevolutionary pheca [9].

1.2. Forensic Genetics

Forensic genetics can be defined as the application of genetics (in the sense of a science with the purpose of studying inherited characteristics for the analysis if inter- and intraspecific variations in populations) to the resolution of legal conflicts by analysis of genetic variation [10]. The foremost applications of DNA analysis in forensic genetics include criminal investigation, personal identification and paternity testing. Since 1985, when Alec Jeffreys and Colleagues first applied DNA analysis to solve forensic problems [11] and in the following

years, numerous medico-legal cases have been solved based on this method (see, for example Gill and colleaguesor and/or Jeffreys and colleagues [12, 13]). Nowadays, DNA undoubtedly has an role in forensic investigations, for example, through the “DNA Innocence project”, launched in the United States to acquit wrongfully convicted people, 344 persons have been exonerated by DNA testing, (September, 2016) including several individuals who were sentenced to death and 148 real perpetrators have been found [14].

It is beyond a doubt that DNA analysis has become “a new form of scientific evidence” that is being constantly evaluated by both the public and professionals. More and more courts around the world accept the results of DNA analysis, and nowadays this technology is almost universally accepted in most legal systems [15].

1.3. Genetic Markers

Genetic marker is an observable trait that can be used to trace the presence of genes determining (or linked with) its variable forms [16]. The first human genetic marker was characterized in 1900 when Karl Landsteiner described the ABO blood grouping system and observed that individuals could be placed into different groups based on their blood type [17]. In 1924, Felix Bernstein established that three alleles of one gene accounted for the ABO blood groups [18] were transmitted according to rules of Mendelian inheritance. In the following years numerous other blood groups, the complex HLA system of white blood cells and several polymorphisms in serum proteins or erythrocyte enzymes, were characterized and could be analyzed in combination to produce highly discriminatory profiles. The use of serum protein genetic markers was a huge breakthrough in the field of human identification however the use of these markers were limited in forensic cases due to the amount of biological material required and to the rapid storage protein degradation. In the 1960s and 1970s, developments in molecular biology, including restriction enzymes, Sanger sequencing [19] and Southern blotting [20], enabled scientists to examine DNA sequences and in 1980 the analysis of the first highly polymorphic locus was reported [21]. In 1984, Jeffreys et al. found that certain genomic regions contained DNA sequences that were repeated next to each other over and over again [22]. Jeffreys also discovered that the number of repeats present in a sample could differ from individual to individual [23]. By developing a technique to examine the length variation of these DNA repeat sequences, he increased the ability to perform human identity tests. These repeat regions became known as VNTRs (Variable Number of Tandem Repeats) and the technique used by Jeffreys to examine the length

variation of these DNA repeat sequences was based on RFLP (Restriction Fragment Length Polymorphism) analysis, which involved the use of restriction enzymes to cut regions of DNA surrounding the VNTRs [21]. The RFLP was the first molecular method used in human identification and population genetics [22]. A critical development came in 1986 when Kary Mullis developed a technique that allows exponential amplification of specific regions of DNA, the polymerase chain reaction (PCR) technique [24].

The advent of PCR highly increased the sensitivity of DNA analysis, potentiating the discovery of new genetic markers, as well as the use of small amounts of DNA. Nowadays wide ranges of DNA polymorphisms are available and well characterized. Being the most used markers: Microsatellites also known as Short Tandem Repeats (STRs; classified according to their number of repeat motifs); Single Nucleotide Polymorphisms (SNPs; structurally characterized by differences in a single nucleotide base); and Insertions and deletions (Indels; insertion or the deletion of bases in the DNA).

Two types of genetic markers can be defined according to their mode of transmission: Subjected to recombination, like those from the nuclear DNA located at the autosomes and the sex X-chromosome in females; and not-subjected to recombination, including those contained in the mitochondrial DNA (mtDNA) and in the Male Specific region of Y-chromosome (MSY). Their different characteristics endow the markers with distinct advantages and limitations that need to be taken into account when being analyzed [25].

1.3.1. Uniparental Genetic Markers

Uniparental genetic markers, also called as genetic lineage markers, comprise polymorphisms that are present on the maternally inherited mitochondrial genome or the paternally inherited Y-chromosome [11]. They are transferred directly from generation to generation (**Figure 1**) without variation, unless a mutation occurs, either from mother to child in the case of mtDNA, or from father to son in the case of the Y-chromosome [26]. Uniparental genetic markers are important because they can provide a clear-cut pattern of historical events that is clouded by recombination factors (see, for example, [27]). Due to their haploid condition, these systems present a low effective population size, a quarter when compare to autosomes [9], being therefore more susceptible to demographic events such as genetic drift [25, 28]. Therefore, they are the markers of choice in the great majority of population studies, specially the phylogeographic ones.

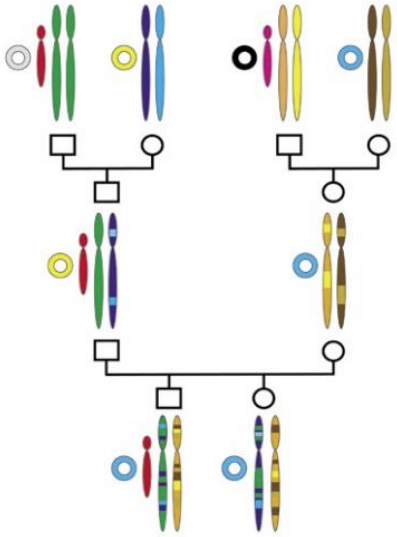

Figure 1 – Schematic representation of human inheritance patterns of autosomal, Y-chromosomal and mtDNA sequences. Autosomal sequences (large pair of chromosomes) recombine each generation, while Y-specific (small chromosome) and mtDNA sequences (circle) do not. Consequently, an individual (bottom) can trace his/her autosomal sequences back to multiple ancestors (top - each is a different colour), but the Y chromosome (excluding the pseudoautosomal regions) and mtDNA have only a single ancestor [29]

1.4. Y-Chromosome

The human Y-chromosome stands out from all other chromosomes because it is male-specific [30]. As result of evolutionary process, exchange between X and Y-chromosomes is limited to two small regions of the X-Y pair and, consequently, to a great extent, the Y-chromosome is paternally inherited and haploid. Along generations, this haploid region is transmitted from father to son unchanged, unless a mutational event takes place [25, 29, 31] . For scientists, more than a simple male determinant, it is a piece of knowledge, which can be used in an abundant range of genetic areas and for several different purposes as forensic evidence examination, paternity testing, historical investigations, studying human migration patterns throughout history, and genealogical research [25] (see, for example, [32, 33]).

1.4.1. Structure

The Y-chromosome, at an average 60 Mb of length, is one of the smallest chromosomes of the human genome [34, 35]. The tips of the Y-chromosome, which are called the pseudoautosomal regions (PAR), recombine with the homologous regions on the tips of X-chromosome. PAR1 located at the tip of the short arm (Yp) of the Y-chromosome is approximately 2.5Mb in length while PAR2 at the tip of the long arm (Yq) is less than 1Mb in size [36]. Since these short regions are homologous of X-chromosome sequences and are

located at the distal portions of the short and long arms of the chromosome they are responsible for correct pairing between the two sex chromosomes during male meiosis [30] (**Figure 2**). The remainder of the Y-chromosome (~95%) was known as the non-recombining portion of the Y-chromosome or NRY [30, 37] and remains the same from father to son unless a mutation occurs [25, 29, 31].

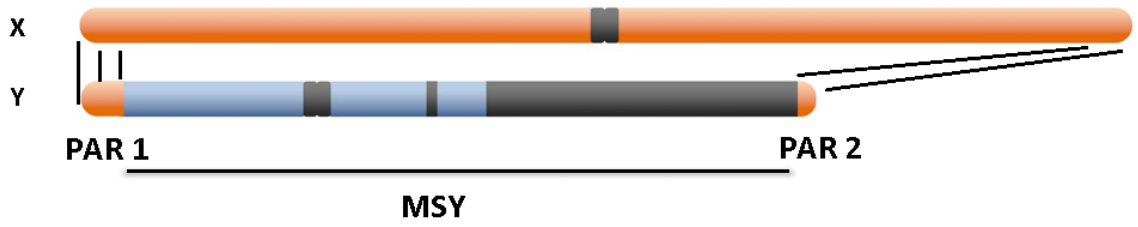

Figure 2 – Schematics of the sex chromosomes. The tips of Y-chromosome (PAR1 and PAR2) recombine with the tips of the X-chromosome. The remaining 95% is the so-called MSY region [38]

Nowadays, the designation of NRY is consider obsolete, due to the evidence of frequent gene conversion or intra-chromosomal recombination; and for this reason, the male-specific region (MSY) designation is commonly accepted [34].

1.4.2. Origin

Homology between the chromosomes X and Y in the PAR regions and the preponderance of shared genes support the proposal that the mammalian X and Y originated from a pair of autosomes [39] (**Figure 3**). Differentiation of the proto-Y began when it acquired a male determining locus. After this, there was accumulation of other male-advantage genes in a region across which recombination was suppressed [40-42]. A chromosome region that is maintained permanently heterozygous without exchange with its homolog will accumulate deleterious recessive mutations, because their increase under mutation pressure will not be resisted by selection [42] and genetic drift becomes important because of the inability to recombine un-mutated regions [43]. Therefore, lack of recombination resulted in progressive Y-chromosome deletions because selection no longer acted upon a single gene, but rather the entire MSY. Under these conditions the Y degraded because of high variation, drift and inefficient selection. Of the ~1100 genes on the ancestral Y (now represented by the X-chromosome), only a total of 45 survive.

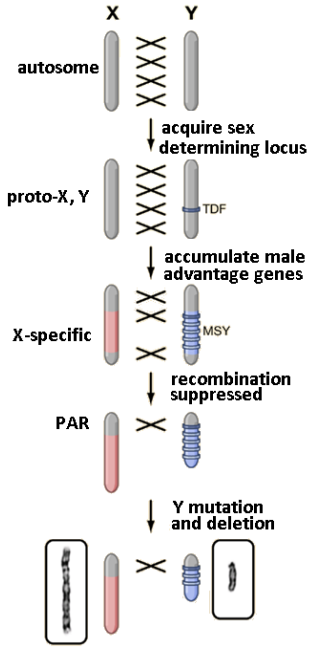

Figure 3 – Mammalian sex chromosomes origin and evolution from an autosomal pair [44]

The Y-chromosome seems to be more subject to mutation, deletion and insertion than the rest of the genome [44]. This bias is because the Y-chromosome must spend every generation in the testis (an oxidative environment and lack repair enzymes). Meanwhile, the autosomes and X-chromosome cycle through the testis is only half or a third as often. Moreover, it takes many more cell divisions to make a spermatozoon than an ovum, providing additional opportunities for damage and in addition, the repetitive structure of the Y-chromosome makes deletions very frequent [39, 40, 44-46].

1.4.3. Genetic Markers in the Y-chromosome

Two broad categories of DNA markers have been using to examine Y-chromosome diversity: Y-SNPs (bi-allelic loci), and Y-STRs (multi-allelic loci). Results from typing of SNP markers are classified into haplogroups while STR results are characterized as haplotypes [25, 47].

1.4.3.1. Y-SNPs

The SNPs are the most abundant class of polymorphisms in the human genome [48], and are estimated to occur at 1 out of every 1,000 bases in the human genome [49, 50]. SNPs are structurally characterized by differences in a single nucleotide base (**Figure 4**).

Normally these markers exhibit only two alleles, reason why they are called bi-allelic polymorphisms, although some rare exceptions for tri-allelic SNPs [51]. When only two alleles are present, one is the ancestral form, from which the other, the derived allele, arose due to mutation. Since the frequency of this type of mutation is very low, approximately 10^{-8} per base per generation [52, 53], it is usually presumed that the derived allele results from a unique event. The first Y bi-allelic marker discovered [54] was the Alu element insertion YAP, being followed by several new detections. Currently a massive amount of these polymorphisms have been identified [55].

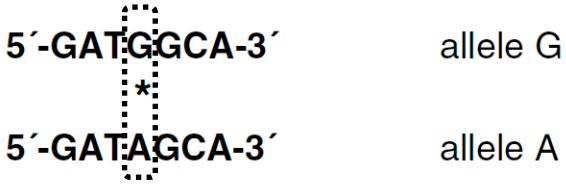

Figure 4 – A single nucleotide polymorphism. Two alleles differ at one position indicated by the star [21]

The SNPs of the Y-chromosome can be used to define haplogroups. The Y-chromosome haplogroup is a group or family of Y-chromosomes related by a common descent that shared the same SNP mutations. The binary polymorphisms associated with the MSY preserve the paternal genetic legacy of our species that has persisted to the present, permitting inference of human evolution, population affinity and demographic history [56-59]. The distribution of Y-haplogroups across populations is not random, but characterized by high levels of geographic specificity (**Figure 5**) [60, 61].

Haplogroups can be named by the major haplogroup information (i.e., H) followed by the name of the terminal mutation that defines a given haplogroup (M36).

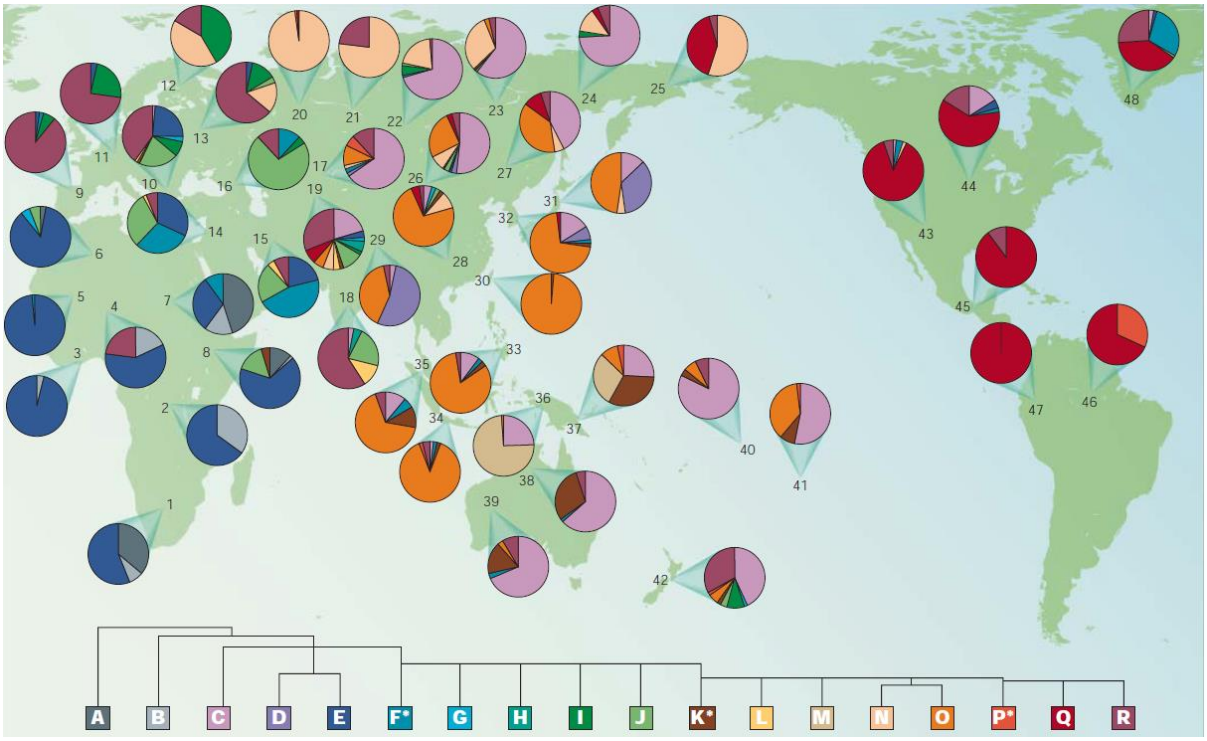

Figure 5 – Global distributions of Y Haplogroups [61]. Each circle represents a population sample with the frequency of the 18 main Y haplogroups identified by the Y Chromosome Consortium (YCC) indicated by the coloured sectors. Populations are numbered as follows: 1, Kung; 2, Biaka Pygmies;3, Bamileke; 4, Fali; 5, Senegalese; 6, Berbers; 7, Ethiopians; 8, Sudanese; 9, Basques; 10, Greeks; 11,Polish; 12, Saami; 13, Russians; 14, Lebanese; 15, Iranians; 16, Kazbegi (Georgia); 17, Kazaks; 18, Punjabis;19, Uzbeks; 20, Forest Nentsi; 21, Khants; 22, Eastern Evenks; 23, Buryats; 24, Evens; 25, Eskimos;26, Mongolians; 27, Evenks; 28, Northern Han; 29, Tibetans; 30, Taiwanese; 31, Japanese; 32, Koreans; 33, Filipinos; 34, Javanese; 35, Malaysians; 36, West New Guineans (highlands); 37, Papua New Guineans (coast);38, Australians (Arnhem); 39, Australians (Sandy Desert); 40, Cook Islanders; 41, Tahitians; 42, Maori; 43, Navajos;
44, Cheyenne; 45, Mixtecs; 46, Makiritare; 47, Cayapa; 48, Greenland Inuit.

Summarily, the geographical distribution of the haplogroups is [62]:

● Haplogroup A (M91) is the older Y-chromosome lineage, from which all Y-branches are descended. Haplogroup A is restricted to Africa, where it is present in several populations at low frequency but is most commonly found in Khoisan populations of Southern Africa and Nilotic people (members of several east-central African groups living in South Sudan, northern Uganda, and western Kenya). Early sub-branches of A have been found in central Africa.

● Haplogroup B (M60) is one of the oldest Y-chromosome lineages in humans. It is found almost exclusively in Africa. This lineage was likely the first to disperse around Africa approximately 90-130 thousand years ago. Haplogroup B appears at low frequency all around Africa, with highest frequency in Pygmy populations.

● Haplogroup C (M130) is found throughout mainland Asia, the south Pacific, New Guinea, Australia, and at low frequencies in Native American populations.

● Haplogroup D (M174) evolved in Asia. This Haplogroup was later displaced from much of Asia by other colonizing groups but is still present at intermediate frequencies in the aboriginal Japanese and on the Tibetan plateau. It is also found at low frequencies in Mongolian populations and the Altais people of central Asia.

● Haplogroup E (M89) is an African lineage. It is currently believe that this haplogroup was dispersed through the Bantu agricultural expansion. This migratory movement started in Central Western Africa, near modern-day Nigeria and Cameroon and dispersed this haplogroup throughout the south. It is also the most common lineage among African Americans. This is a diverse haplogroup with many branches and nowadays is distribute throughout Africa and present at a very low frequency in North Africa, Middle East and Southern Europe.

● Haplogroup F (M89) lineages are extremely rare and are distributed in Europe, the Middle East, and Asia.

● Haplogroup G (M201) was the first branch of Haplogroup F outside of Africa. This haplogroup found mostly in the north central Middle East and the Caucasus, with smaller numbers around the Mediterranean and eastward. Haplogroup G represents one of the first peoples in Europe.

● Haplogroup H (M69) is nearly completely restricted to India, Indonesia, Sri Lanka, and Pakistan.

● Haplogroup I (M170) is found throughout Europe, although some branches may be present in low frequencies in Northeast Africa, Central Siberia, the Near East, and the Caucasus regions. As haplogroup G, haplogroup I represents one of the first peoples in Europe.

● Haplogroup J (M12f2.1) is found at highest frequencies in the Middle East, west of the Zagros Mountains in Iran to the Mediterranean Sea, and encompassing the entire Arabian

Peninsula. It is also found in North African populations where it has been carried by Middle Eastern traders into Europe, central Asia, India, and Pakistan.

● Haplogroup K (M9) is present at low frequencies in Africa, Asia, and in the south Pacific. One descendent lineage of this haplogroup is restricted to aboriginal Australians while another is found at low frequency in southern Europe, Northern Africa, and the Middle East.

● Haplogroup L (M20) is found primarily in India and Sri Lanka and has also spread into several Middle Eastern populations (Turks, Saudis, and Pakistanis). It is also present at very low frequencies in Europe.

● Haplogroup M (M4) is completely confined to the South Pacific. It most probably originated in Melanesia and then spread into Indonesia, Micronesia, and New Guinea.

● Haplogroup N (M231) is distribute throughout Northern Eurasia and Siberia. It is the most common Y-chromosome type in Uralic speakers (Finns and Native Siberians). It is also found in Mongolia.

● Haplogroup O (M175) is a branch of the mega-haplogroup K. O originated about 35,000 years ago in Asia. Its branches have spread into Central and East Asia.

● Haplogroup P (M45) is an extremely rare haplogroup. It is the ancestral line of haplogroups Q and T. It is present at low frequency in India, Pakistan, and central Asia.

● Haplogroup Q (M242) is the lineage that links Asia and the Americas. This lineage is present in North and Central Asian populations as well as Native Americans. Among European populations, haplogroup Q is most frequently distribute in Eastern Europe and Scandinavia. This lineage had originated in Central Asia and migrated through the Altai/Baikal region of northern Eurasia into the Americas. Haplogroup Q-M3 is the only lineage strictly associated with Native American populations. This haplogroup is define by the presence of the M3 mutation that occurred on the Q lineage around 8-12 thousand years ago, as the migration into the Americas was underway.

● Haplogroup R (M207) had origin in Central Asia. Most descendants belong to one of two major lineages. They are present at low frequencies across Central Asia, South Asia, and Europe. Haplogroup R-M173 possibly originated in Eastern Europe and then migrated eastward into Asia.

● Haplogroup S (M230) is an Oceanic lineage and is present primarily in populations from Papau New Guinea with lower frequencies in Melanesia and Indonesia.

● Haplogroup T (M70) is presently found in southern Europe, Northern Africa, and the Middle East.

The process of determining the haplogroup by direct testing of SNPs can sometimes be a lengthy process. Therefore, there is considerable interest in predicting the haplogroup from a set of STR markers [59]. Studies have showed that Y-STR variability is structured in a haplogroup background, thus the haplotypes may also be used to predict haplogroup status [59, 63, 64]. Several approaches have been developing for predicting the Y-chromosome haplogroup from a set of Y-STR markers, among which is that implemented in Haplogroup Predictor [65]. This program has a Bayesian approach, based on the allele frequencies for each haplogroup and how well a given test haplotype fits the pattern of alleles in each haplogroup, which were calculated from collections of haplotypes extracted from published articles and databases [59]. It is has been implemented on a web site since October, 2004 [65], where allows a set of 37 Y-STRs markers to be entered and returns a “goodness of fit score” for 20 haplogroups (E1b1a, E1b1b, G2a, G2c, H, I1, I2a (xI2b1), I2a1, I2b (xI2b1), J1, J2a1b, J2a1h, J2a1 x J2a1-bh, J2b, L, N, Q, R1a, R1b and T). The accuracy of this prediction software have been estimating and the results showed that the Y-haplogroups were correctly predicted in 98.80% samples using only 12 Y-STRs [66].

1.4.3.2. Y-STRs

STRs or microsatellites are the most commonly multi-allelic markers used in Population and Forensic Genetics because: 1) they are robust, since they lead to successful analysis of a wide range of biological material even in non-optimal conditions; 2) the results generated in different laboratories are easily comparable; 3) they are highly discriminatory, especially when analyzing a large number of loci simultaneously; 4) they are very sensitive, requiring only a few cells for a successful analysis; and 5) it is relatively cheap and easy to

generate STR profiles [21]. They are classified according to their number of repeat motifs, which can comprise, in average, 2 to 6 base pairs (bp) of length (**Figure 6**).

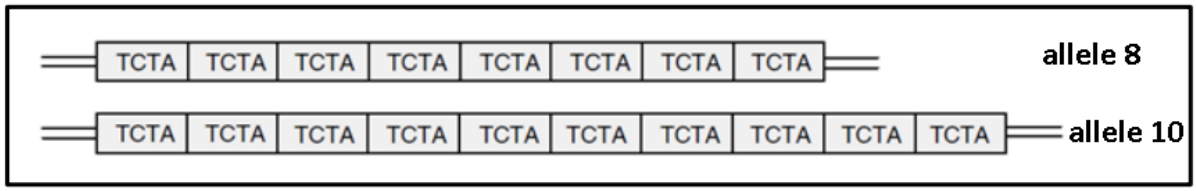

Figure 6 – A short tandem repeat. The aleles are named according to the number of repeats they contain. Extracted from Goodwin, 2011 [21]

Since Y-STRs change more rapidly (mutation rate of 2.1 X 10^{-3} per generation) [47, 67, 68] compared to Y-SNPs (mutation rate of 10^{-8} per base per generation [52, 53], the Y-STR results exhibit more variability and thus have greater use in forensic applications. Furthermore, it is possible to amplify multiple STR loci in a single combined multiplex reaction and direct detection of amplified products through capillary electrophoresis across the introduction of fluorescent dye-labeled primer technology.

The first Y-STR described was Y27H39 known at present as DYS19 [69]. Subsequently, an increasing number of Y-STRs were described and more than 400 Y-STRs [70, 71] have been discovered and deposited in the Genome Database (GDB). Although this database is no longer available online, information regarding Y-STRs, such as nucleotide sequence, can be found on GenBank Database, available at http://www.ncbi.nlm.nih.gov/. In view of the huge diversity of Y-STR, efforts have been necessary to obtain a reliable database for the practical use of these markers [72], so a core set of Y-STR markers (DYS19, DYS389I, DYS389II, DYS390, DYS391, DYS392, DYS393, DYS385), compatible with the International Society for Forensic Genetics (ISFG) guidelines for forensic STR analysis was settled as the “minimum haplotype”. Whilst seven of these markers are unilocal, marker DYS385 detects variation at two loci simultaneously, thereby bringing the total number of loci tested to nine. This core represent a means of safe, sensitive and efficient genotyping of the human Y-chromosome and provides a high discrimination potential [73-75]. In 2003, the Scientific Working Group on DNA Analysis Methods (SWGDAM) recommended the addition of the loci DYS438 and DYS439 to the minimal haplotype; however, nowadays, the SWGDAM haplotype is no longer used so often. These loci are included in the commonly used Y-STR kits .

Nowadays, in the Y-Chromosome Haplotype Reference Database (YHRD – yhrd.org) [76], a reference anonymous database developed mainly for the forensic community, the haplotypes available range from 7 to 26 Y-STRs. It comprises the minimal haplotype recommended, but also the maximum number of loci amplified in actual multi-allelic loci commercial kits [77]. These multi-allelic loci can be used to differentiate Y-chromosome haplotypes with fairly high resolution due to their higher mutation rates [48, 70, 78-80]. Information regarding Y-STRs, such as nucleotide sequence, can be found on GenBank Database, available at http://www.ncbi.nlm.nih.gov/ [81].

1.5. Commercial kits

Different multiplexes were developed [82] and some of them successfully employed in forensic cases [73, 83]. Nowadays, forensic routine rely heavily on commercially available kits to perform DNA testing. Most laboratories do not have time or resources to design primers, optimize PCR multiplexes, and control the quality of primer synthesis. The convenience of using ready-made kits is also augmented by the fact that widely used primer sets and conditions allow improved opportunities for sharing data between laboratories without fear of failing the detection of silent alleles (**Figure 7**) [77] as kits from different companies have different primer binding sites.

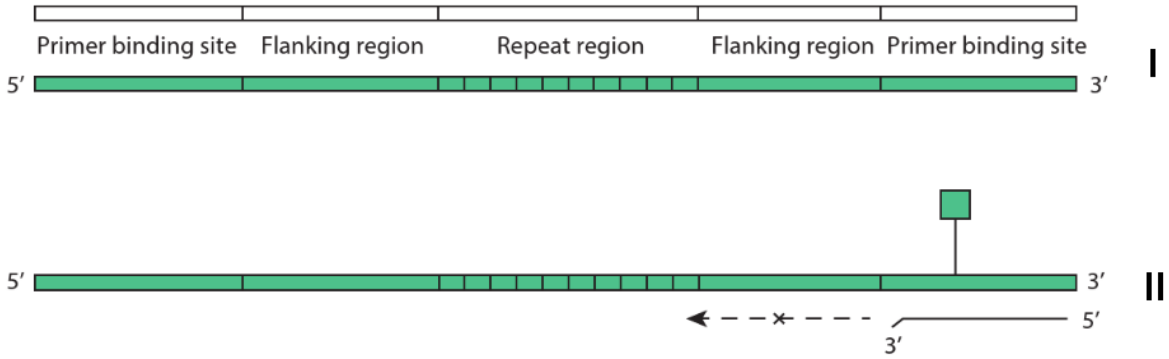

Figure 7– I. Idealized structure of a short tandem repeat locus; II. Primer binding site mutation. Due to no amplification, the allele is silent under the conditions of this Y-STR Kit [84]

Several Y-STRs commercial kits have been released since 2001 (**Table 1**).

Table 1 – Commercial Y-STR kits. Adapted from Butler, 2005 [25, 48, 85, 86]

Kit name (Source)	Release Date	Number of loci amplified	Loci Amplified
Y-PlexTM 6 (ReliaGene Technologies)	2001	6	DYS393, DYS19, DYS389II, DYS390, DYS391, DYS385
Y-PlexTM 5 (ReliaGene Technologies)	2002	5	DYS389I, DYS389II, DYS439, DYS438, DYS392
genRES® DYSplex-1 (Serac)	2002	7	DYS390, DYS39I, DYS385, Amelogenin, DYS5389I/II
genRES® DYSplex-2 (Serac)	2002	5	DYS392, DYS393, DYS19, DYS389I/II
Y-PlexTM 12 (ReliaGene Technologies)	2003	12	DYS392, DYS390, DYS385, DYS393, DYS389I, DYS391, DYS389II, Amelogenin, DYS19, DYS439, DYS438
PowerPlex® Y (Promega Corporation)	2003	12	DYS391, DYS389I, DYS439, DYS389II, DYS438, DYS437, DYS19, DYS392, DYS393, DYS390, DYS385
MenPlex® Argus Y-MU (Biotype)	2004	9	DYS393, DYS390, DYS385, DYS391, DYS19, DYS389I, DYS392, DYS389II
Yfiler™ (Applied Biosystems)	2004	17	DYS456, DYS389I, DYS390, DYS389II, DYS458, DYS19, DYS385, DYS393, DYS391, DYS439, DYS635, DYS392, H4, DYS437, DYS438, DYS448
Investigator® Argus Y-12 QS (Qiagen)	2010	12	DYS19, DYS385, DYS389I, DYS389II, DYS390, DYS391, DYS392, DYS393, DYS437, DYS438, DYS439
PowerPlex® Y23 (Promega Corporation)	2012	23	DYS576, DYS389I, DYS448, DYS389II, DYS19, DYS391, DYS481, DYS549, DYS533, DYS438, DYS437, DYS570, DYS635, DYS390, DYS439, DYS392, DYS643, DYS393, DYS458, DYS385, DYS456, H4
Yfiler® Plus	2014	27	DYS576, DYS389I, DYS635, DYS389II, DYS627, DYS460, DYS458, DYS19, YGATAH4, DYS448, DYS391, DYS456, DYS390, DYS438, DYS392, DYS518 DYS570 DYS437, DYS385, DYS449, DYS393, DYS439, DYS481, DYF387S1, DYS533

1.5.1. Yfiler™

The AmpFℓSTR® Yfiler™ PCR amplification kit (Applied Biosystems) simultaneously amplifies 17 Y-STR loci including the loci in the "European minimal haplotype" (DYS19, DYS385a/b, DYS389I, DYS389II, DYS390, DYS391, DYS392, and DYS393), SWGDAM recommended Y-STR loci (DYS438 and DYS439), and the highly polymorphic loci DYS437, DYS448, DYS456, DYS458, Y GATA H4, and DYS635 (formerly known as Y GATA C4). This kit was validated according to the FBI/National Standards and SWGDAM guidelines. The Yfiler™ showed that full profiles are attainable with low levels of male DNA (below 125pg) and that under optimized conditions, no detectable cross-reactive products were obtained on human female DNA, bacteria, and commonly encountered animal species. Additionally, this kit has the ability to detect male specific profiles in admixed male and female blood samples at a ratio of 1:1000 [87]. Allele calls are assigned on the basis of the Yfiler® kit allelic ladder containing the 137 most common alleles observed for all loci (**Figure 8**).

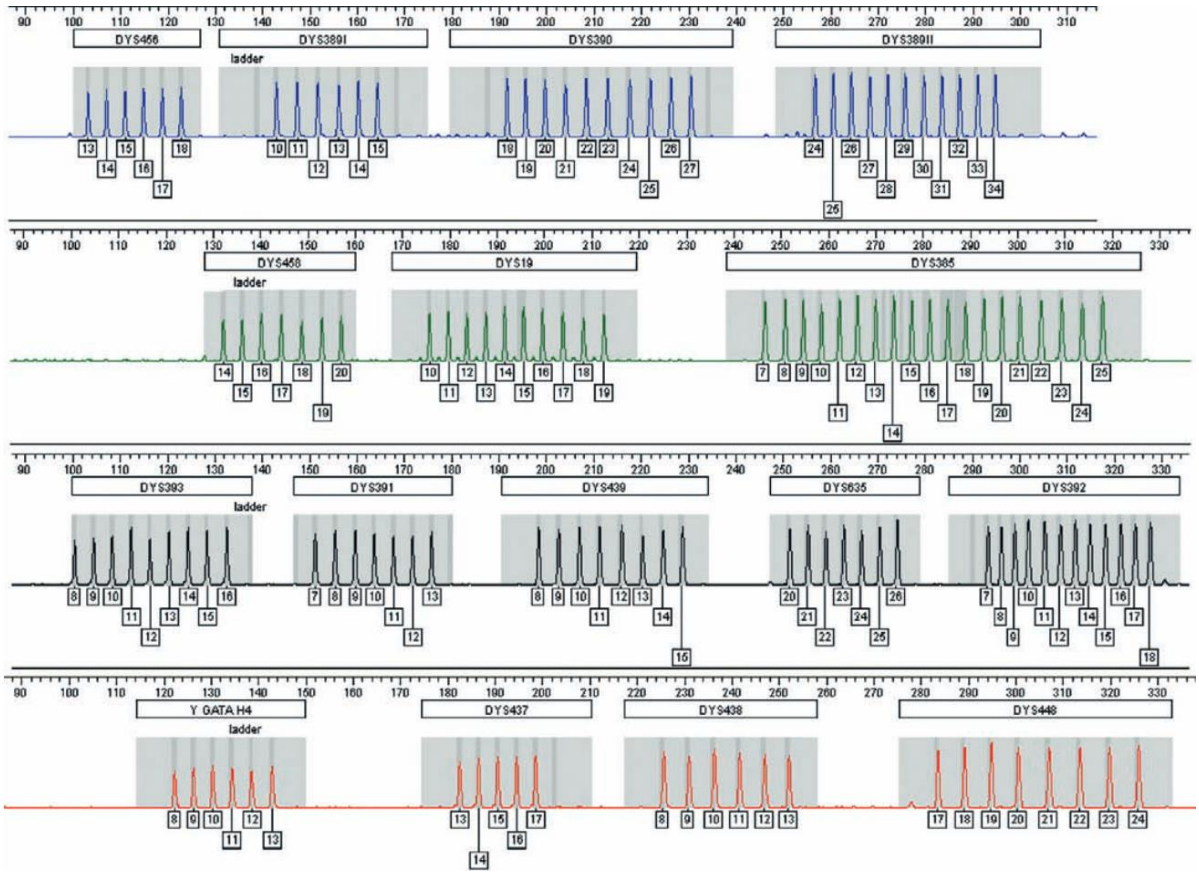

Figure 8 – Y-filer kit allelic ladder [25]

1.6. Brazil

This work focuses in a Brazilian population. Brazil, officially the Federative Republic of Brazil, is the largest country in both South America and Latin America. Its population of nearly 206,000,000 is one-third the population of South America. Brazil occupies 8,515,767,049km^2 (making it the world's fifth-largest country by both area and population) and is divided into five geopolitical regions: North, Northeast, Central-West, Southeast, and South (**Table 2**). It is the largest Lusophone (Portuguese-speaking) and Roman Catholic country in the world, and the only Portuguese speaking country in America [88].

Table 2 - Geopolitical regions and States of Brazil

Northeast	**North**	**Central-West**	**Southeast**	**South**
Maranhão	Amazonas	Mato Grosso	São Paulo	Paraná
Piauí	Pará	Mato Grosso do Sul	Minas Gerais	Santa Catarina
Ceará	Roraima	Goiás	Rio de Janeiro	Rio Grande do Sul
Rio Grande do Norte	Rondônia	Federal District	Espírito Santo	
Paraíba	Amapá			
Pernambuco	Tocantins			
Alagoas	Acre			
Sergipe				
Bahia				

1.7. Ancestry of the Brazilian population

Brazil has a large territory harboring highly diverse populations that resulted from different degrees and modes of admixture. From the genetic point of view, Brazil is known as one of the most heterogeneous population in the world, with an important genetic contribution from three main continental groups: Europeans, Africans and Native Americans. The first people arriving in Brazil were Europeans, coming mainly from Portugal, who arrived in 1500 to a territory that was already inhabited by the Native Americans for at least 11,000 Years [89-92]. Initially, admixture was mainly between native females and Portuguese male navigators [91]. During the slave trade period, which officially lasted from 1538 to 1850, a

huge number of African people were forced to move to Brazil. During that period, approximately 3.6 million slaves are estimated to have entered the country [93, 94]. After the abolition of slavery in Brazil in 1888, a new important migration wave took place, extending the admixed process to new immigrants not just from Europe but also from Asia and Middle East. The number of new incomers was approximately 6 million, coming from Portugal, Italy, Spain, Germany, Syria, Lebanon and Japan [91, 95]. At the same time that people from diverse countries and continents were arriving to different regions in Brazil, important movements were taking place inside the territory, mainly due to economic interests. These internal movements gained a new impetus after the First World War, between 1914 and 1918, mainly from the northeast to the north and southeast regions of the country [96].

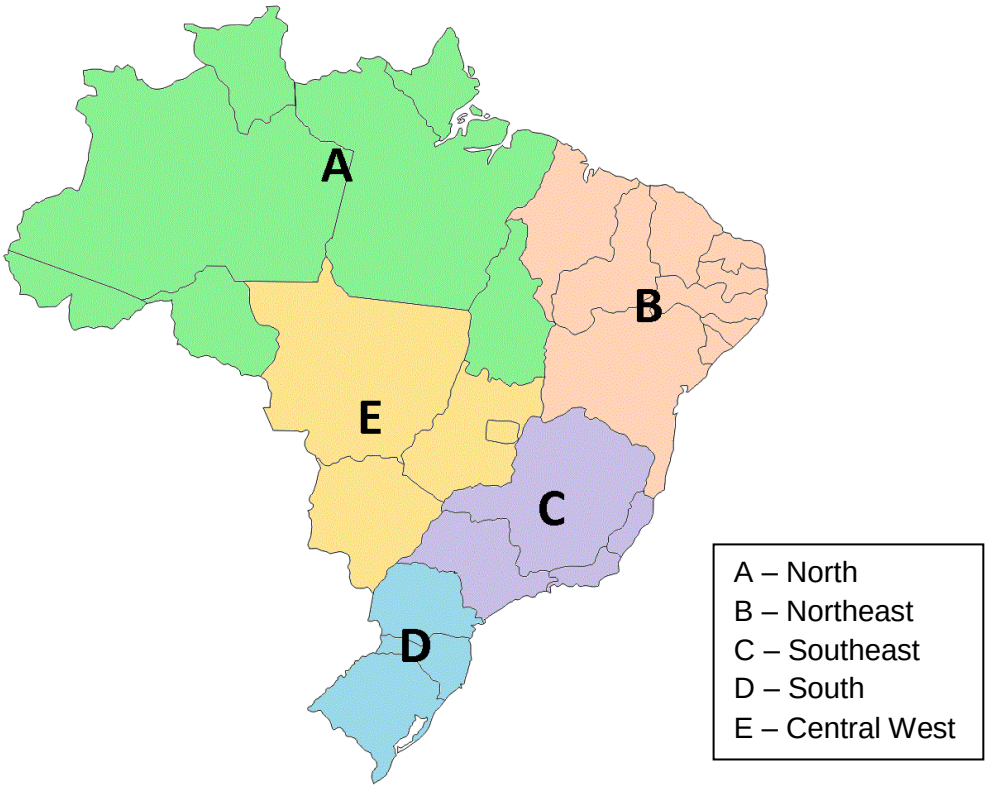

Figure 9 - Map of Brazil showing the 5 geopolitical regions. Adapted from d-maps.com [97]

Consequently, the modern Brazilian population is genetically very diverse and considered heterogeneous when considering the 5 main geopolitical regions of the country (**Figure 9**): (A) the northern region hosts the people with the largest Native American ancestry; (B) northeast region has the highest African contribution; (C) the southeast and (D) the south are the regions where the European contribution is more important, and; (E) the central west was the last colonized region by the influx of people coming from all the other Brazilian regions, mainly from the northeast and southeast [91, 95, 98]. A previous study [91] showed no

significant differentiation for 17 populations, distributed among the 5 geopolitical regions of Brazil, which was attributed to the high frequency of European male lineages in all populations [92].

1.7.1. The State of São Paulo

São Paulo is one of the 27 states of Brazil and it is located at southeast region of the country (**Figure 10**). It is the most populous state in Brazil, with approximately 44.771.000 inhabitants of that 49.24% are male and 50.76% are female [99].

Figure 10 – Map of Brazil showing the 5 geopolitical regions and the state of São Paulo. Adapted from d-maps.com [97]

Nowadays the state of São Paulo is administratively divided into 15 mesoregions (**Figure 11**) and its population represents a typical sample of Brazilian mixed ethnicity.

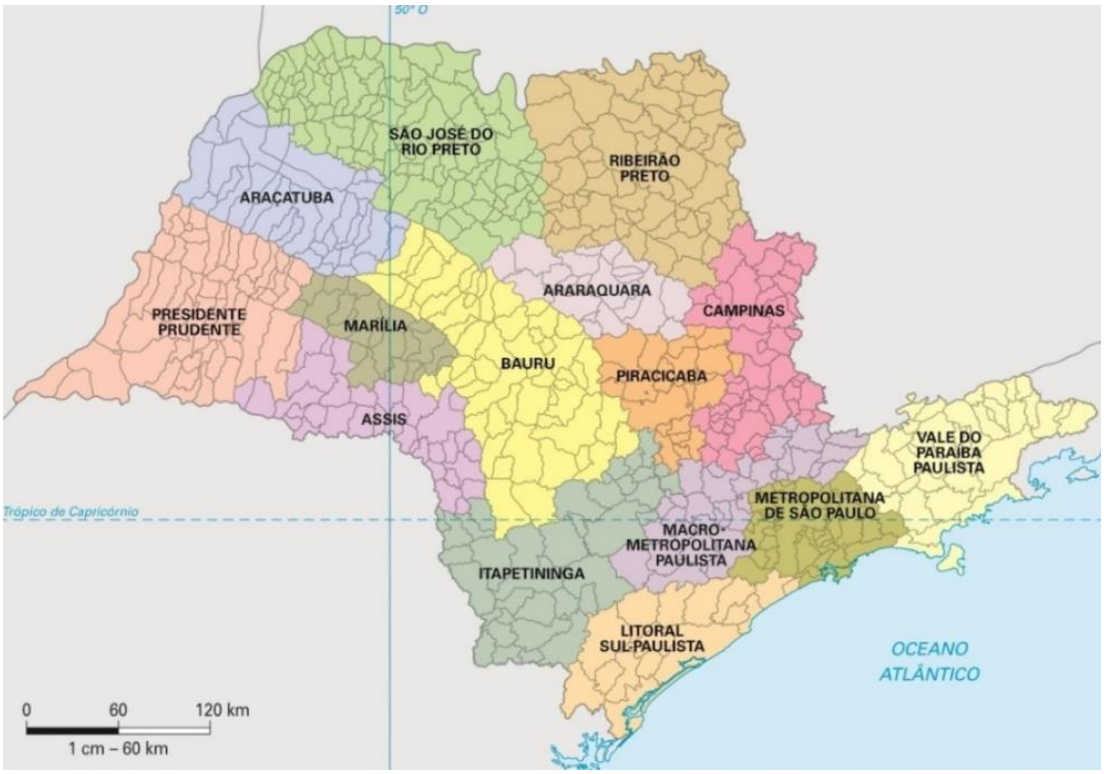

Figure 11– Mesoregions of São Paulo [100]

The colonial settlement of São Paulo state began in 1532, by an expedition led by the Portuguese Martim Afonso de Sousa. Initially, admixture was mainly between native Tupi females and Portuguese male navigators [92]. Tupi was the first Amerindian group to have contact with Portuguese community. Soon, a process of miscegenation between Portuguese settlers and indigenous women started. The colonists rarely brought women, making the Indian women the breeding matrix of the Brazilian people. Polygyny, a common practice among South American Indians, was quickly adopted by European settler and, this way, a single European man could have dozens of Indian wives [101].

Settlers were heavily dependent on indigenous labor. During the initial phases of settlement, natives were often captured by expeditions called "*bandeiras*" originated in São Paulo. However, the cultivation of sugarcane increased the demand of workforce and the use of hand labor of African slaves spread across the country and throughout the second half of the sixteenth century, the African hand labor gradually replaced the native labor. The transatlantic slave trade promoted the settlement of Brazil and of São Paulo by people coming from different regions of Africa. The Tietê-Paranapanema river system (which winds into the interior of São Paulo), made it an ideal base for enslaving expeditions and during the 16th , 17th and mid-18th centuries, São Paulo state received slaves mainly from West-Central

Africa (in particular the territory currently occupied by Angola) and later, from the African East Cost, particularly from Mozambique. Africans did not stop coming until the middle-19th century, when the traffic was definitely abolished. Subsequent migratory waves of Spanish, Italian, German and Japanese contributed to an ethnic variety of this population [94, 102].

2. AIMS

In attempting to increase the genetic information relatively to the Brazilian state of São Paulo, a sample of unrelated males from São Paulo was characterized for a set of Y-chromosome STRs included in YFiler™ kit, aiming to achieve the following goals:

- Calculate several diversity parameters with relevance for the forensic and population genetics fields;

- Compare the results from this study with those from previous ones trying to:

    - Apprehend if the present study corroborates results from published data;

    - Search for possible patterns of genetic relatedness between São Paulo and populations from other Brazilian regions;

    - Explore the European and African populations that had contributed to the present genetic background of São Paulo, searching for possible relatedness.

3. MATERIAL AND METHODS

3.1. Sampling

We analyzed 417 unrelated males from paternity caseworks available from the Institute of Social Medicine and Criminology of São Paulo. The samples include individuals living in the 15 administrative regions of the state. DNA was extracted and purified from blood stains collected on FTA® cards (Whatman®,BioScienses) using FTA® Purification Reagent (Whatman®) following the manufacturer's protocol [103].

3.2. Genotyping

3.2.1. STRs

Target DNA (0.5 – 2.0 ng/µL) was amplified by PCR using the AmpFlSTR® YFiler™ PCR Amplification Kit (Applied Biosystems), at the following loci: DYS19, DYS385, DYS389I/II, DYS390, DYS391, DYS392, DYS393, DYS437, DYS438, DYS439, DYS448, DYS456, DYS458, DYS635, and Y-GATA-H4 (**Table 3**) – A positive control was used for evaluating the efficiency of the amplification and genotyping.

Table 3 – Characteristics of the 17 Y-STR loci amplified with the Yfiler® System [66, 80] . The mutation rate for each loci is available in the YHRD database [76]

Locus designation	**Repeat structure**	**Alleles included in AmpFlSTR® Yfiler® Allelic Ladder**	**Mutation rate (95%CI)**
DYS456	$(AGAT)_n$	13, 14, 15, 16, 17, 18	4.29×10^{-3}
DYS389 I	$(TCTG)_3(TCTA)_n$	10, 11, 12, 13, 14, 15	2.45×10^{-3}
DYS390	$(tcta)_2(TCTG)_n(TCTA)_n(TCTG)_n(TCTA)_ntca(tcta)_2$	18, 19, 20, 21, 22, 23, 24, 25, 26, 27	2.11×10^{-3}
DYS389 II	$(TCTG)_n(TCTA)_nN_{28}(TCTG)_3(TCTA)_n$	24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34	4.12×10^{-3}
DYS458	$(GAAA)_n$	14, 15, 16, 17, 18, 19, 20	6.36×10^{-3}
DYS19	$(TAGA)_3tagg(TAGA)_n$	10, 11, 12, 13, 14, 15, 16, 17, 18, 19	2.24×10^{-3}
DYS385 a/b	$(aagg)_{6-7}(GAAA)_n$	7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25	2.45×10^{-3}
DYS393	$(AGAT)_n$	8, 9, 10, 11, 12, 13, 14, 15, 16	1.05×10^{-3}

Table 3 – continued

DYS391	$(tctg)_3(TCTA)_n$	7, 8, 9, 10, 11, 12, 13	2.45×10^{-3}
DYS439	$(GATA)_n$	8, 9, 10, 11, 12, 13, 14, 15	5.45×10^{-3}
DYS635	$(TCTA)_4(TGTA)_2(TCTA)_2(TGTA)_2(TCTA)_2(TGTA)_{0,2}(TCTA)_n$	20, 21, 22, 23, 24, 25, 26	4.33×10^{-3}
DYS392	$(TAT)_n$	7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18	5.19×10^{-4}
Y GATA H4	$(AGAT)_4CTAT(AGAT)_2(AGGT)3(AGAT)_nN24(ATAG)_4(ATAC)_1(ATAG)_2$	8, 9, 10, 11, 12, 13	3.03×10^{-3}
DYS437	$(TCTA)_n(TCTG)_{1-3}(TCTA)4$	13, 14, 15, 16, 17	1.22×10^{-3}
DYS438	$(TTTTC)_1(TTTTA)_{0-1}(TTTTC)_n$	8, 9, 10, 11, 12, 13	3.75×10^{-4}
DYS448	$(AGAGAT)_nN42(AGAGAT)_n$	17, 18, 19, 20, 21, 22, 23, 24	1.52×10^{-3}

Amplification was performed according to the instructions provided by the manufacturer [81] although adaptations were done according to the laboratory's procedures (adjusted for a final volume of 12.5 µL and under the conditions described in **Table 4)**. One FTA punch was used per polymerase chain reaction (PCR), and an unused FTA punch was prepared as a negative control.

Table 4 – PCR reaction conditions for STR amplification with Yfiler® System

PCR Reaction	
Reagents	**Volume per reaction (µL)**
Y filer Kit PCR Reaction Mix	4.6
Y filer Kit Primer Set	2.5
AmpliTaq Gold® DNA Polymerase	0.4
H_20 (RNase/DNase/Protease Free)	3.2
DNA	1.8
Total	12.5

PCR was performed using a GeneAmp PCR System 9700 (Applied Biosystems) under the conditions described in **Table 5**.

Table 5 – PCR program conditions for STR amplification Yfiler® System

		PCR Program	
		Temperature (ºC)	Time (min)
Initial		95	11
25 cycles	Denaturation	94	1
	Annealing	59	1
	Extension	72	1
Final Extension		60	60
Hold		4	∞

Electrophoresis of the PCR products, together with GeneScan-500 Internal Lane Size Standard (LIZ-500) to determine the base-pair sizes, was performed on the ABI PRISM®3130xl Genetic Analyzer (Applied Biosystems). GeneMapper® ID-X v.1.4. Software version 4.0 was used in order to determine the allelic repeats by applying YFiler Allelic Ladder. Alleles were named as suggested by the manufacturer's nomenclature.

3.2.2. Haplogroups

The haplotypes obtained in São Paulo sample studied were submitted to the Haplogroup Predictor web based software [65] with equal priors, obtaining probabilities for the most probable correspondent haplogroup. This software gives as a result the probability of the Y-STR haplotypes belonging to each group of the 20 haplogroups (E1b1a, E1b1b, G2a, G2c, H, I1, I2a (xI2b1), I2a1, I2b (xI2b1), J1, J2a1b, J2a1h, J2a1 x J2a1-bh, J2b, L, N, Q, R1a, R1b and T). The haplogroup assigned was the one that scored the highest probability.

3.3. Data Analysis

3.3.1. Population Data

3.3.1.1. Brazilian Populations

Comparisons among the 417 samples here analyzed and Brazilian samples from previous studies were made in order to evaluate differences between the present data and 20 populations belonging to the 5 geopolitical regions of Brazil: North, (n=1092) [91, 104-106], Northeast (n=279) [91], Central West (n=344)[91, 107], Southeast (n=1256) [90, 91, 104, 108] and South (n=670) [91, 109, 110]. We also compared our data with samples belonging to São Paulo state, analyzed in previous studies [91, 104] (**Figure 12**).

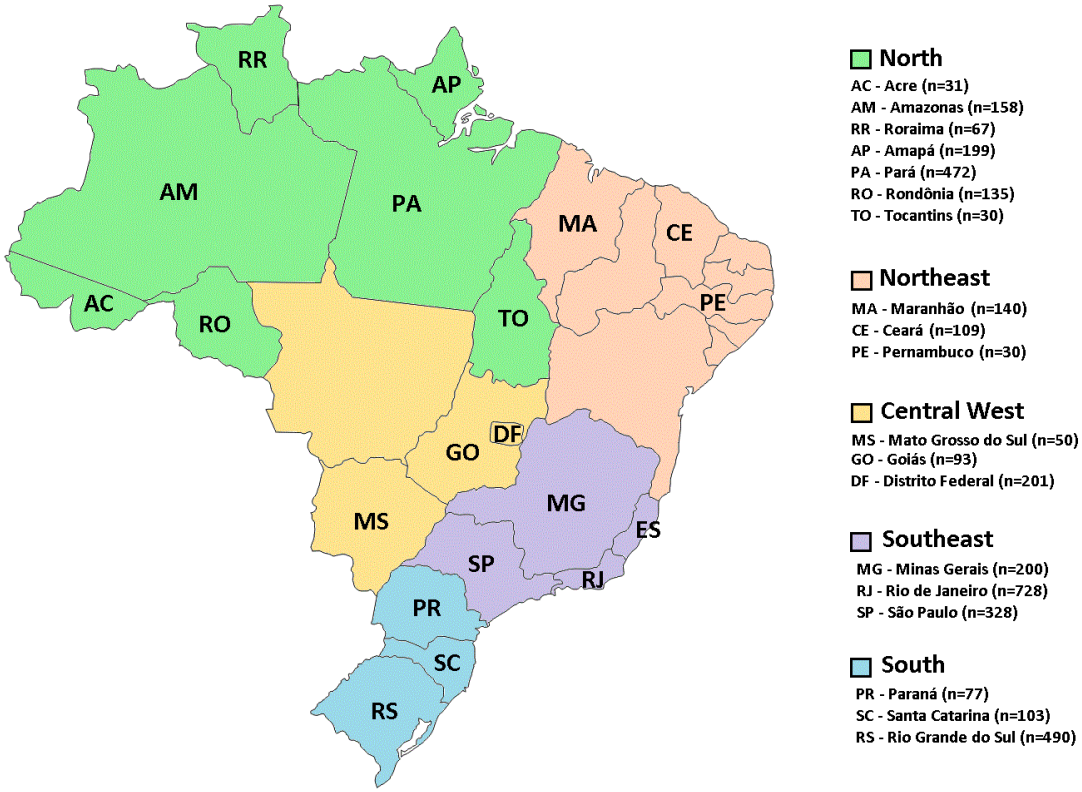

Figure 12 – Map of Brazil indicating the geographical location of the 20 different populations used for comparison located in 5 geopolitical regions. Adapted from d-maps.com [101]

3.3.1.2. European and African Populations

In order to search the main ancestral sources of modern São Paulo population, we also did comparisons between our sample and four reference samples from European and two African countries that have been documented in anthropological and historical terms as the main countries that have contributed to the present genetic background of São Paulo [95] (**Table 6**).

Table 6 – Country of origin and number of European and African samples from previous studies

		Number of samples
Europe		
	Germany [105]	1718
	Italy [104, 105]	1462
	Portugal[104, 105]	298
	Spain [105]	706
Africa		
	Angola [91]	61
	Mozambique [91]	35

The haplogroup composition of these Europeans and Africans populations were also estimated through their STR haplotypes using Haplogroup Predictor Software [65]. From each European country, just the chromosmes belonging to lineages identified as the main European haplogroups - R, I, J or G - were taken into account for the subsequent analysis; and only chromosomes belonging to haplogroup E1b1a (identified as the main sub-Saharan Africa haplogroup) were taken from the African samples for the subsequent analysis.

3.3.2. Statistical Analysis

The haplogroup and haplotype frequencies were calculated through direct counting in the characterized samples. Diversity estimators as the gene diversity and the haplotype diversity, the amount of shared haplotypes between our samples and the compiled populations , and the value of unique haplotypes were calculated using Arlequin v3.5.1.2 software [111]. Moreover, pairwise genetic distances (R_{ST}), also calculated in Arlequin, were used to measure genetic distances between the samples from this study and the other compiled populations. In inter-population comparisons, the DYS389II allele length was obtained by subtracting the number of repeats in DYS389I from the allele at DYS389II. To illustrate the relationship between populations based on the calculated pairwise genetic distances matrix, an MDS plot was created by using Statistica, version 5 software (Statsoft) [112]. Finally, genetic relationships between haplotypes inside of specific haplogroups were analyzed using NETWORK 5.0 software (available online, http://www.fluxus-engineering.com), using frequency>1 criterion and applying the reduced median and median-joining method. STR weighting was applied in accordance with Qamar et al. (2002) [113]. For network construction, the duplicated locus DYS385 was not used since the constituent loci are not distinguished in this assay, and the number of repeats at DYS389II was calculated after subtracting the number of repeats at DYS389I.

4. RESULTS AND DISCUSSION

4.1. Frequency and diversity measurements

4.1.1. Haplotype data

In the set of the 417 samples from unrelated males from São Paulo, considering the total 17 Y-STR markers typed, a total of 410 different haplotypes were found, among which 403 (96.64%) haplotypes were unique and 7 haplotypes were shared each of them between a pair of individuals. All the haplotypes found are described in Supplementary **Table S1**. The probability of two random individuals showing identical haplotype was 0.25% and a high overall Y-STR haplotype diversity was observed (0.9999±0.0002) in the pooled sample representing São Paulo state. The average gene diversity at the 17 Y-STR loci was 0.66295±0.07783. The Y-STR gene diversity for each locus for São Paulo population is shown in **Table 7**. These results show a high level of haplotype diversity in a population of São Paulo, indicating applicability of these markers for individual identification and forensic studies.

Table 7 – Gene diversity for each locus

Locus	Gene diversity
DYS392	0.60282
DYS391	0.57895
DYS390	0.71910
DYS393	0.56282
DYS389I	0.54636
DYS389II	0.59123
DYS19	0.67002
DYS458	0.79523
DYS439	0.67657
GATA H4	0.57707
DYS456	0.71524
DYS437	0.61339

Table 7 – continued

DYS635	0.69637
DYS448	0.68925
DYS438	0.71009
Mean	**0.66295±0.07783**

Of the 17 analyzed markers, DYS458 showed the greatest degree of diversity, 0.79523, being this locus considered as a rapidly mutating one [80]. In the lower range of levels of diversity, DYS389I was the locus characterized by the lowest value (0.54636). Despite the fact that the locus with the highest value of gene diversity also has the highest value of mutation rate, the relation between both parameters is far from being linear as is shown in **Figure 13**. This explains why the loci showing the lower level of gene diversity is not the one with the lowest mutation rate.

Figure 13 – Genetic diversity and Mutation rate locus by locus, log10 scale. The mutation rate for each loci is available in the YHRD database [76]

Comparing this results with those from a previous study [114], there are no significant differences regarding the results (**Table 8**).

Table 8 – Comparison among the data of São Paulo Population from the present study and a previous one [115]

	This study	Previous Study
Unique Haplotypes (%)	96.64	97.4
Haplotype found in two individuals (%)	3.36	2.4
Haplotype found in three individuals (%)	0	0.2
Haplotype Diversity	0.9999±0.0002	0.9999±0.0001
Mean of Gene Diversity	0.663	0.664
Probability of two random individuals showing identical haplotypes (%)	0.25	0.21

4.1.2. Haplogroup data

The 17 Y-STR haplotypes considered in this study allowed classifying the pooled set of 417 Y-chromosomes from São Paulo state population in 10 different haplogroups (**Table 9**).

Table 9 – Haplogroup frequencies in São Paulo state. N represents the number of samples

General haplogroup	Sub-clade	N	Frequency
R	R1a	7	0.016786
	R1b	187	0.448441
E	E1b1a	34	0.081535
	E1b1b	53	0.127098
J	J1	11	0.026379
	J2	37	0.088729
I	I1	20	0.047962
	I2	24	0.057554
G	G2a	21	0.05036
Q		9	0.021583
T		6	0.014388
L		5	0.011990
N		2	0.004796
H		1	0.002398

4.1.2.1. Haplogroup R (M207)

Chromosomes analyzed in the present study have a major European contribution since the haplogroup R1 (M173) was present in 194 samples, representing 46.52% of the total. All

the samples belonging to haplogroup R (M207) were associated to the haplogroup R1. Typically, more than 50% of men in Europe are affiliated within haplogroup R, essentially to its sub-clade R1 (M173) [115], which is mainly represented by two lineages – R1a (M420) and R1b (M343) [116, 117]. In the present study, 7 samples (3.61%), belong to sub-clade R1a that is common in Eurasia, and 187 haplotypes (44.8% of the total) belongs to sub-clade R1b. Men belonging to haplogroup R1b are believed to be the descendants of the first modern humans who entered Europe, where is now the most common European Y-haplogroup [118].

4.1.2.2. Haplogroup E (M96)

The African contribution was the second largest to the gene pool of São Paulo since the second haplogroup more frequent was the E1b1 (E-P2). This haplogroup may have arisen in East Africa [28] and is restricted to African populations [119] however, its subclade E1b1b(M215), was widespread in southern Europe. Haplogroup E (M96) is present in 87 haplotypes, representing 20.86% of the samples. Within the samples belonging to haplogroup E1b1–P2, 12.71% of total belong to sub-clade E1b1b (M215) (formerly known as E3b); and 8.15% (of total) belong to sub-clade E1b1a (M2). The high frequency of haplogroup E in our samples is consistent with the fact that many Africans were taken to Brazil during the period of slavery. The network of E1b1a lineages was constructed using the STRs haplotypes (**Figure 14).** From this analysis it is possible to assert that chromosomes from São Paulo belonging to E1b1a sub-haplogroup did not tend to be located in isolated clusters but in the same clusters than those from the African haplotypes. These results seemed to reinforce the contribution of these African countries to the São Paulo population. In the network, it was possible to observe shared haplotypes between São Paulo and Angola (H250, H380 and H400) São Paulo and Mozambique (H290 and H334) what was not observed when including the loci DYS385a/b in the comparisons. Despite of this exclusion in the construction of the network, Angola and Mozambique did not share any haplotype.

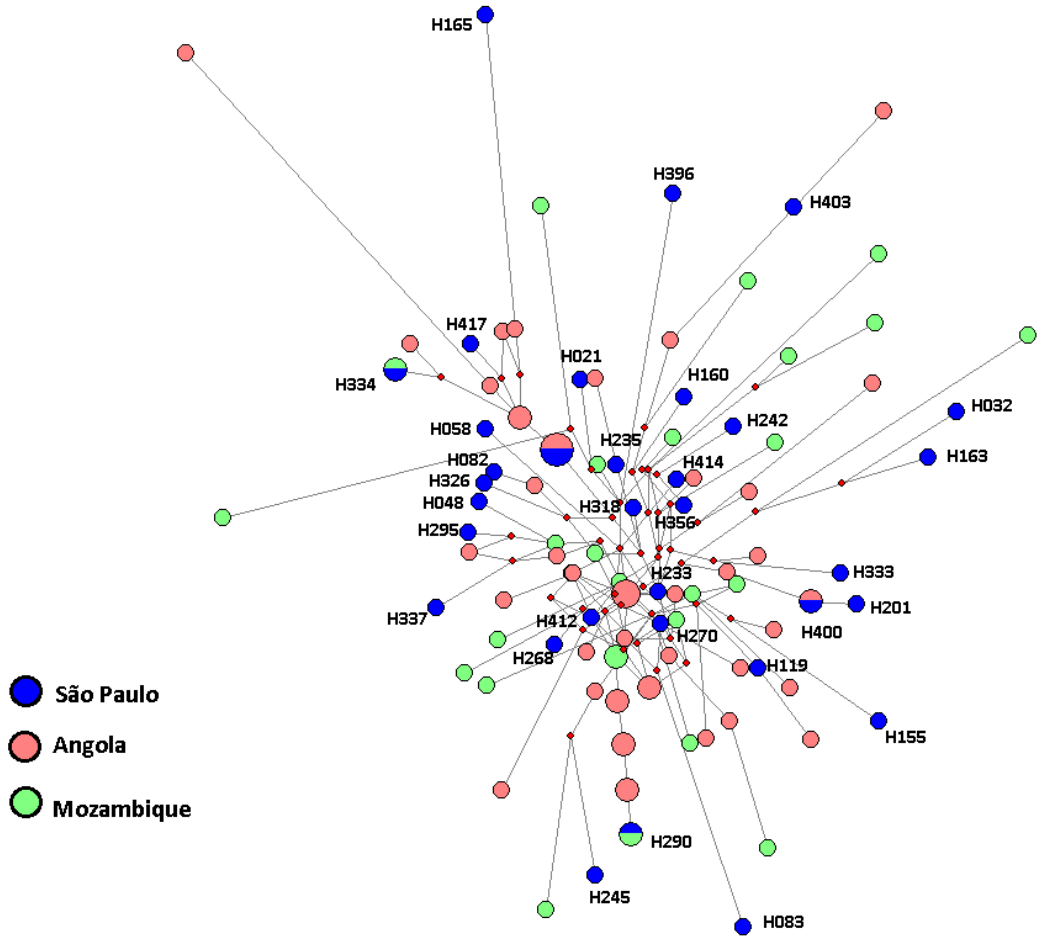

Figure 14 – Median-joining network within sub-haplogroup E1b1a. The network was based in 15 STRs (DYS393, DYS390, DYS19, DYS391, DYS439, DYS389I, DYS389II, DYS392, DYS458, DYS437, DYS448, GATA_H4,DYS456, DYS438 and DYS635). The area of the circles is proportional to the frequency: The smallest circle is equivalent to only one individual/ haplotype. The line length, between each haplotype, is proportional to the number of mutations. Each circle colour represents a different population under the considerations of the figure legend

4.1.2.3. Haplogroup J (M12f2.1)

Belonging to haplogroup J (M12f2.1), were found 48 haplotypes, which represent 11.50% of the samples. Within this Haplogroup, 37 samples (8.87% of total) belong to sub-haplogroup J2 (M172), and 11 samples (2.64% of total) belong to sub-haplogroup J1 (M267). Y-DNA haplogroup J evolved in the ancient Near East and was carried into North Africa, Europe, Central Asia, Pakistan and India [120]. J2 lineages originated in the area known as the Fertile Crescent. The main spread of J2 into the Mediterranean area is thought to have coincided with the agricultural expansion during the Neolithic period. The world's highest frequency of J2 is found among people in the Northeast Caucasus, however one fourth of the Vlach people (isolated communities of Romance language speakers in the Balkans) belong to

J2, considerably more than the average of Macedonia and northern Greece where they live. This, combined to the fact that they speak a language descended from Latin, suggests that they could have a greater part of Roman (or at least Italian) ancestry than other ethnic groups in the Balkans. The Romans probably helped spread haplogroup J2 within their borders, judging from the distribution of J2 within Europe (frequency over 5%), which bears an uncanny resemblance to the borders of the Roman Empire [121, 122]. Since there are no historical records of large migratory waves from Caucasian populations to the state of São Paulo the haplogroup J2 can be assumed as an European contribution.

J1 also represents an European contribution to the males of São Paulo. These lineages may have a more southern origin than the J2 lineages, as they are more often found in the Levant region, other parts of the Near East, and North Africa, with a sparse distribution in the southern Mediterranean flank of Europe, and in Ethiopia. There is a descending gradient in the frequency of occurrence of haplogroup J from the Middle East toward the northwest of Europe, reaching about 3% of the population on the northwest Atlantic coast. The occurrence of J in Europe is undoubtedly due both to the Neolithic expansion and to episodic migrations [55]. A network for Haplogroup J was constructed (**Figure 15**) using

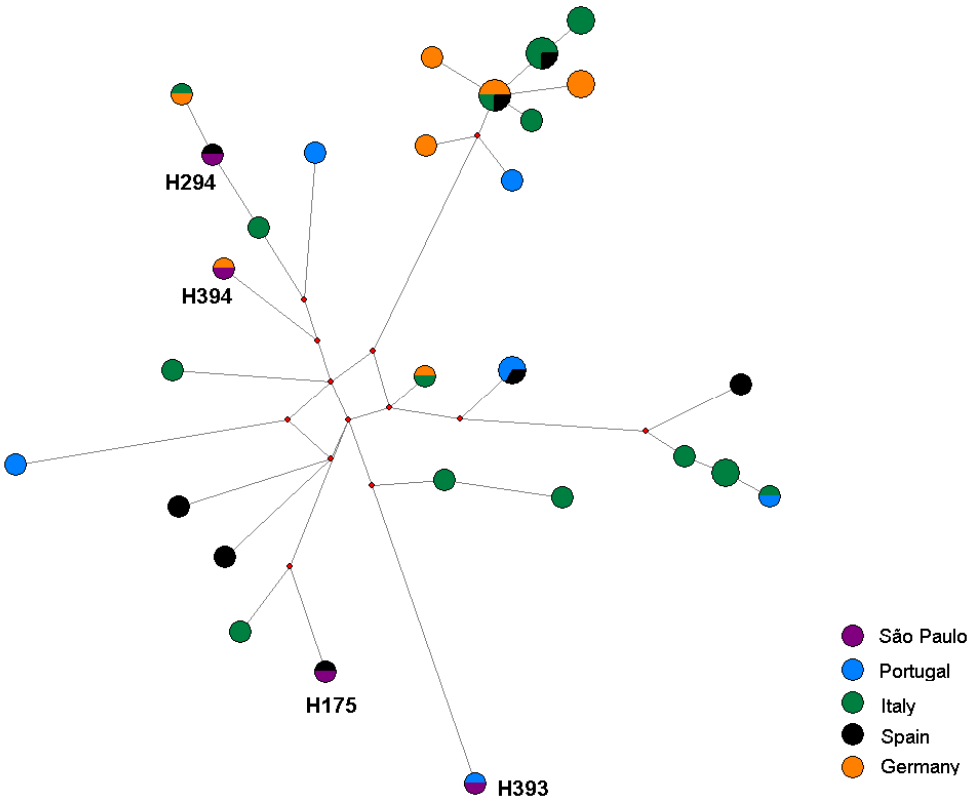

Figure 15 – Median-joining network within haplogroup J. The network was based in the same 15 Y-STR loci referred in Figure 14

STRs haplotypes belonging to clades J1 and J2. From this analysis it is possible to assert that chromosomes from São Paulo belonging to J haplogroup did not tend to be located in isolated clusters but in the same clusters than those from the European haplotypes. These results seemed to reinforce the contribution of these European countries to the São Paulo population. In the network, it was possible to observe shared haplotypes between São Paulo and Spain, Portugal and Germany. These four cases of STR haplotype sharing involved three J2 haplotypes (H175, H393 and H394) and one J1 (H294).

4.1.2.4. Haplogroup I (M170)

The haplogroup I (M170) appears in 10.55% of the samples (44 individuals). This haplogroup is predominantly an European haplogroup and it is considered as the only native European haplogroup, accounting, on average, for 18% of the total paternal lineages in Europe. Its virtual absence elsewhere, including the Near East, suggests that it arose in Europe, likely before the Last Glacial Maximum [121, 123]. The network of I lineages was constructed using the STRs haplotypes (**Figure 16)**. Despite this haplogroup has a central

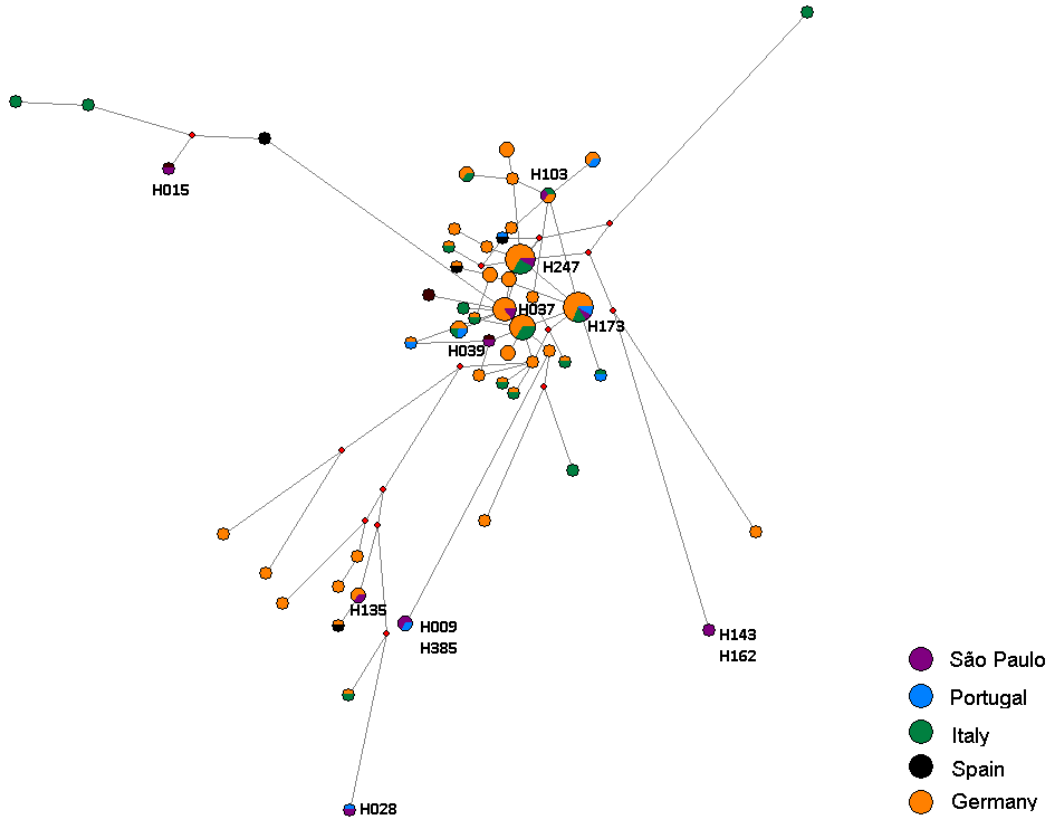

Figure 16 – Median-joining network within haplogroup I. The network was based in the same 15 Y-STR loci referred in Figure 14.

European distribution, represented in the network analysis by the German population, the São Paulo samples appeared in nodes shared with samples from Italy, Portugal and Spain. The only shared haplotype specific with the German population is found in the H135. Therefore an origin from these countries for this haplogroup can not be discarded (mainly considering both countries as the main source of settlers in the region).

4.1.2.5. Haplogroup G (M201)

The haplogroup G (M201), specifically the sub-clade G2a (P15) is present in 21 samples, what represents 5.04% of the total. The presence of haplogroup G was first reported in Europe and Georgia [123]. This haplogroup is common in the region of Caucasus, Mediterranean and Middle East. A strong influence of these groups in the colonization of São Paulo is not recognized but G2a was found in medieval remains in a seventh century tomb in Ergolding, Bavaria (Germany), and the German colonization is recognized in the territory of Brazil. His greatest influence was primarily in the South of Brazil, though its presence in São Paulo can be explained due to migration flows, which took place over the years, within the Brazilian territory. The phylogeographic demarcation zone of haplogroup G is largely restricted to populations of the Caucasus and the Near/Middle East and southern Europe. This haplogroup occurs at frequencies ranging from 5 to 15% in both the rest of Near/Middle East and southern European countries [124]. A network was also constructed for the Haplogroup G **(Figure 17)**.

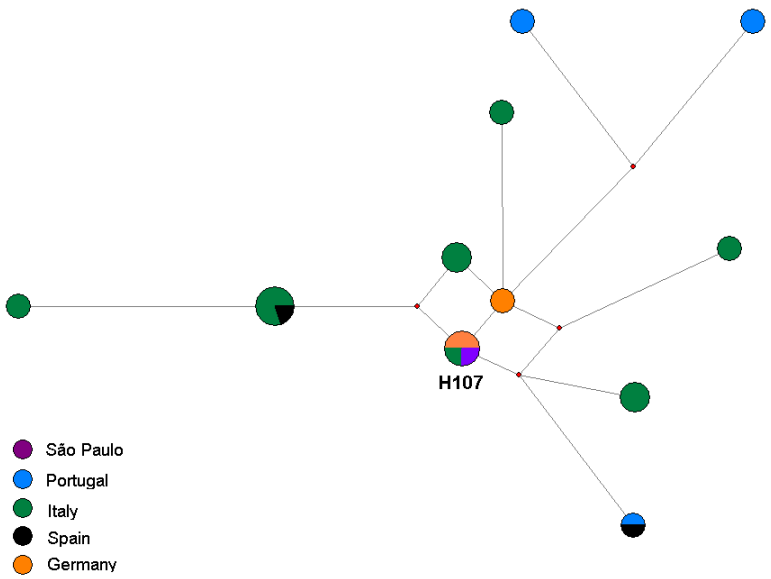

Figure 17 – Median-joining network within haplogroup G. The network was based in the same 15 Y-STR loci referred in Figure 14.

Chromosomes from São Paulo belonging to G haplogroup shared the same haplotype with samples from Italy and Germany. The starlike nature of the network suggests that these lineages have undergone a recent expansion.

4.1.2.6. Haplogroup Q (M242)

Nine of our samples belong to the haplogroup Q (defined by M242), which represent 2.16% of the total. Its presence is a legacy of the first inhabitants of the territory as approximately 90% of pre-Columbian Native Americans belonged to haplogroup Q. Its low frequency, compared to European and African haplogroups shows that nowadays in São Paulo the gene pool has higher contribution from settlers population than from native Americans. At this point, São Paulo differs from the Brazilian Amazon region where, despite most haplogroup belonged to European lineages (89.2%), are more Amerindian (7.2%) than to African haplogroups (3.6%) [106], supporting a diverse demographic history of the populations from these areas.

Haplogroup T (M184)

Belonging to haplogroup T (defined by M184) we find 6 samples that represent 1.44% of the total. The maximal worldwide frequency for haplogroup T is observed in East Africa (Eritrea, Ethiopia, Somalia, Kenya, Tanzania) and in the Middle East (especially the South Caucasus, southern Iraq, south-west Iran, Oman and southern Egypt), where it accounts for approximately 5 to 15% of the male lineages. There are no historical record of strong migrations from these countries to São Paulo, despite its presence can be a contribution of Europeans as the frequency of this haplogroup in our sample is similar to Europe, where it makes up 1% of the population on most of the continent [55, 125] .

4.1.2.7. Haplogroup L (M20)

Belonging to haplogroup L (defined by M20), 5 samples are detected (1.2%). Y-DNA haplogroup L derives from haplogroup T and is found at low frequencies in the Middle East and Europe as sub-haplogroups L1b (M317), L1b1(M349) and L2 (L595), and at significant frequencies in South Asia as sub-haplogroups L1a1 (M27) and L1a2 (M357). The sub-clade L1b1 (M349) is known for Mediterranean branch L for being found, although in low percentages, only in Europe's southern countries from Turkey to Portugal [55] . Using the Haplogroup Predictor, we could not reach these sub-clades, but the presence of samples

belonging to sub-haplogroup L2a could be explained by the colonization of São Paulo by Europeans.

4.1.2.8. Haplogroup N (M231)

Haplogroup N (M231) is found in 2 samples, which represents 0.48% of the total. Haplogroup N is found throughout Northern Eurasia and the highest frequency of this haplogroup occurs among the Finnic and Baltic peoples of northern Europe. Haplogroup N has also been found at moderate concentration in Eastern Europe and at low concentration in Anatolia. N1c-Tat (former N3) represents the western extent of haplogroup N, which is found all over the Far East (China, Korea, and Japan), Mongolia and Siberia, especially among Uralic speakers of northern Siberia. In the beginning of 19th century, migration from Asia (particularly Japanese and Syrian-Lebanese) to Brazil has increased considerably, most of this immigrants was allocated to coffee farms in São Paulo [95] . The presence of sub-haplogroup N in the population of São Paulo could be explained by this Japanese migratory flow since haplogroup N is the fourth most common haplogroup in Japan [122] .

4.1.2.9. Haplogroup H (M2939)

Finally, was found only one sample (0.24%) belonging to haplogroup H (M69). This haplogroup has not been comprehensively studied yet. Nowadays the greatest proportion of members of this haplogroup live in the southern Asia subcontinent area. The Romani (also known as Gypsy) people, who apparently originated in India, are the main source of haplogroup H in Western Europe. Haplogroup H is also found in south of Iberian Pensinsula, as the presence of haplogroup H in the São Paulo population could be explained by the Portuguese and Spanish migration to the Brazilian territory. The sub-haplogroup H2-P96 is usually common in Sardinia (Italy), and also during the 19th century many Italians migrated to Brazil, but using the Haplogroup Predictor [65] we couldn't reach this sub-haplogroup.

4.2. Population comparisons

4.2.1. Brazilian Populations

4.2.1.1. Shared haplotypes

The haplotypes shared between our samples and the other Brazilian populations (from previous studies) are shown in **Table 10**. The 417 samples of male population of São

Paulo state of the present study did not share any haplotype with the Acre, Tocantins, Goiás and Minas Gerais samples. The sample of Pernambuco showed the highest percentage of common haplotypes with our samples (6.67%) followed by Paraná (5.19%), Rio Grande do Sul (4.69%) and Mato Grosso do Sul (4.00%). This result did not show a correlation with the geopolitical divisions as these populations with the highest percentage of sharing haplotypes with our samples respectively belong to Northeast, South and Central West regions. A sample (328 haplotypes) collected in São Paulo in previous studies [91, 104] shared with the present study 3.35% of its haplotypes.

Table 10 – Number and frequencies of shared haplotypes between our 417 samples from São Paulo and Brazilian Populations from previous studies [90, 91, 104-107, 109, 110]

Population	n of shared haplotypes	n of samples	% of common haplotypes with the present study
Acre	0	31	0
Tocantins	0	30	0.00
Goiás	0	93	0.00
Minas Gerais	0	200	0.00
Pará	2	472	0.42
Amazonas	1	158	0.63
Ceará	1	109	0.92
Santa Catarina	1	103	0.97
Maranhão	2	140	1.43
Roraima	1	67	1.49
Rio de Janeiro	13	728	1.79
Amapá	4	199	2.01
Rondónia	3	135	2.22
Espírito Santo	6	253	2.37
Distrito Federal	5	201	2.49
São Paulo	11	328	3.35
Mato Grosso do Sul	2	50	4.00
Rio Grande do Sul	23	490	4.69
Paraná	4	77	5.19
Pernambuco	2	30	6.67

4.2.1.2. Genetic distances

Pairwise genetic distances were estimated based on the sum of squared size differences (R_{ST}) between the haplotype distributions observed for the 17 Y-STRs in our samples and the 20 Brazilian admixed population samples mentioned before. The obtained matrix is displayed on Supplementary Table S2. As in previous studies [91], the samples showed significant genetic distances within their geopolitical regions and therefore we analyzed them separately. Significant genetic distances were found between our samples and Amazonas [91, 104, 105], Amapá [91, 106], Pará [91], Roraima [91], Distrito Federal [107], Goiás [91], Minas Gerais [91], Santa Catarina [110] and Rio Grande do Sul [91, 109].

4.2.1.3. MDS representation

A multidimensional scaling (MDS) was performed by using Statistica, version 5 software (Statsoft) [112] with the R_{ST} distances to illustrate the relationship between the brazilian populations based on pairwise genetic distances matrix (**Figure 18**). It shows that the São Paulo population is closely related to Espírito Santo and Rio de Janeiro. These populations belong to the same geopolitical region (southeast); however, Minas Gerais also belongs to southeast region and shows an outlier position. Following, Minas Gerais, Amazonas and Santa Catarina (from North and South, respectively) show an outlier position.

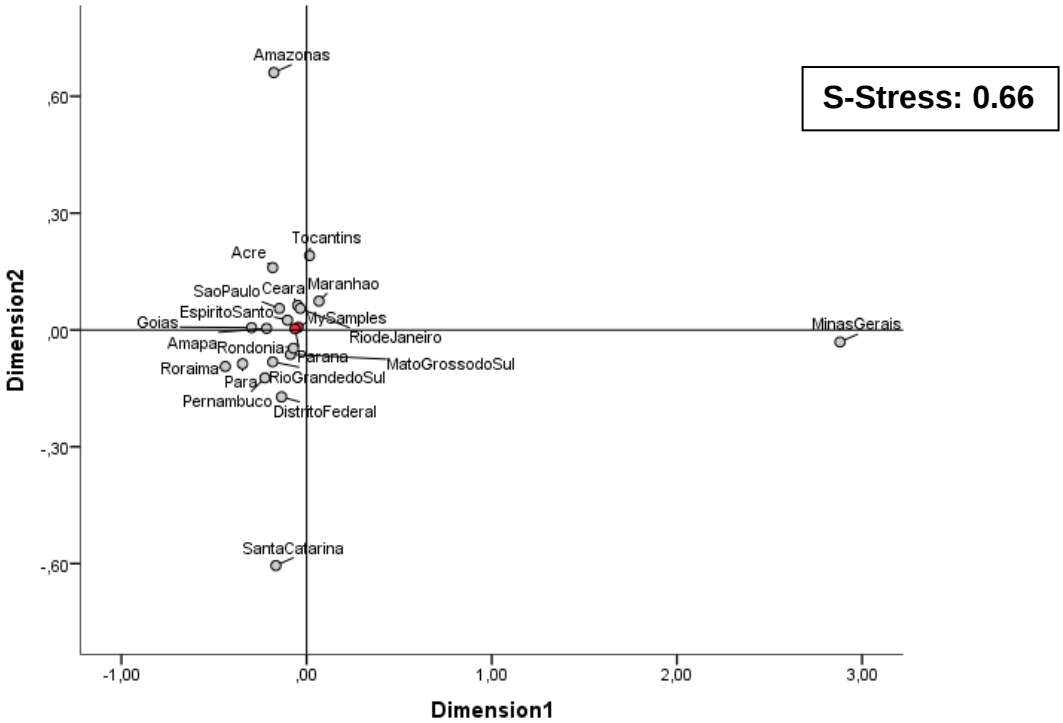

Figure 18 – MDS plot based on R_{ST} values for 17 Y-STR based haplotypes showing relationships among brazilian populations. This plot was based on the matrix data of Table S1.

4.2.2. European and African Populations

From each European country, just the haplotypes belonging to haplogroups identified as the main European haplogroups (R, I, J and G) were taken into account for the comparisons; and only haplotypes belonging to the main sub-Saharan haplogroup (E1b1a-M2) were taken from the African samples. The number of European haplotypes belonging to the main European haplogroups and number of the African haplotypes belonging to main African haplogroups are shown in **Table 11**.

Table 11 – Country of origin and number of European [104, 105] and African [91] samples used for comparisons

Country	Total n of samples	n of samples belonging to the respective continent main haplogroup	% of samples belonging to the respective continent main haplogroup
Germany	1718	1549	90.16
Italy	1462	1184	80.98
Portugal	298	246	82.55
Spain	706	618	87.54
Present study	417	307	73.62
Angola	61	46	75.41
Mozambique	35	26	74.29
Present study	417	34	8.15

4.2.2.1. Shared haplotypes

The haplotypes shared between our samples and the European and African populations (from previous studies and belonging to their respective continents main haplogroups) used in the comparisons, are shown in **Table 12**.

Table 12 – Number and frequencies of shared haplotypes between our samples from São Paulo and European [104, 105] and African [91] Populations

Population	n of shared haplotypes with our samples	n of samples used for the comparison	% of haplotypes present in our samples
Germany	12	1549	0.77
Italy	6	1184	0.51
Portugal	8	246	3.25
Spain	9	618	1.46
Angola	0	46	0.00
Mozambique	0	26	0.00

4.2.2.2. Genetic distances

Pairwise genetic distances were estimated based on R_{ST} between the haplotype distributions observed for the 17 Y-STRs in our samples and the European and African samples mentioned before. The obtained matrix for the European and African samples are displayed on Supplementary Table S3 and Table S4, respectively. Within the European countries, Portugal seems to be the country that gave the biggest contribution to the present genetic background of São Paulo as significant genetic distances were found between our samples and Germany [105], Italy [104, 105] and Spain [105] but not with Portugal [104, 105]. And for the African countries, our sample from São Paulo state shows to be closely related to Angola as no significant genetic distance was found between our samples and samples from this country.

4.2.2.3. MDS representation

A multidimensional scaling (MDS) performed by using Statistica, version 5 software (Statsoft) [112] using the F_{ST} distances clearly shows that our sample is closely related to samples from Portugal [105] (**Figure 19**).

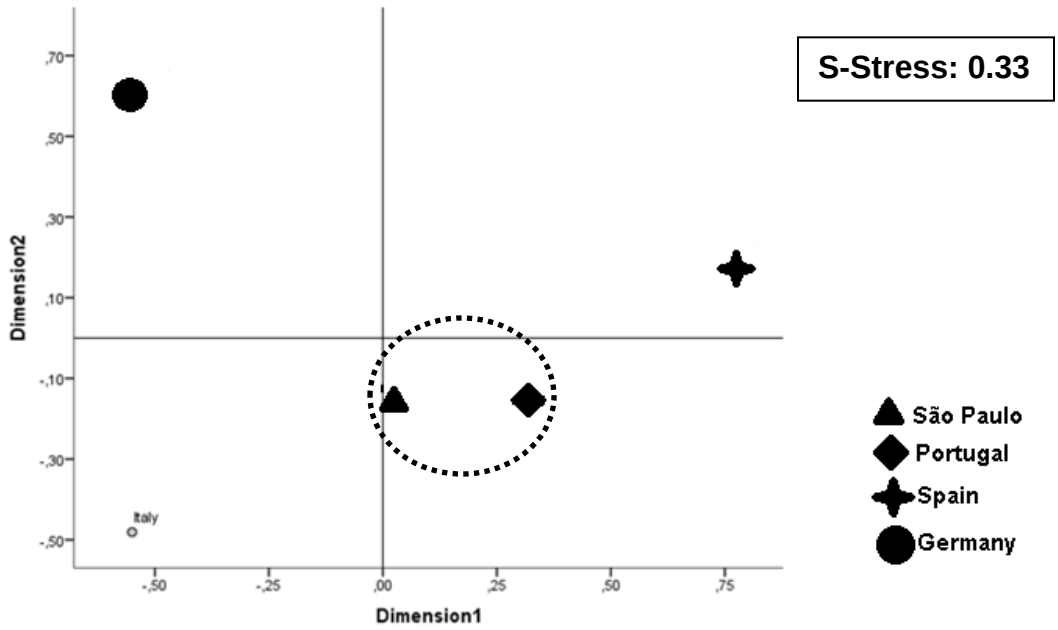

Figure 19 – MDS plot based on R_{ST} values for 17 Y-STR based haplotypes showing relationships among our samples and European populations. The dashed circle highlights a possible main cluster formed. This plot was based on the matrix data of Table S2

The MDS performed for the African populations shows a high S-Stress due to the small number of samples (**Figure 20**) but it also shows clearly the proximity of our sample and the samples from Angola.

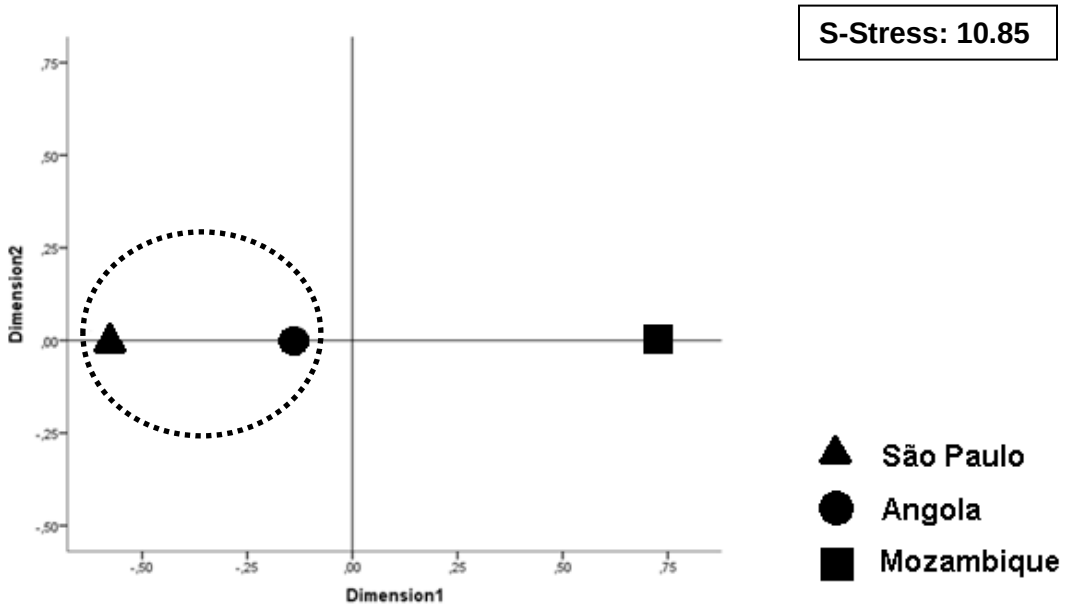

Figure 20 – MDS plot based on R_{ST} values for 17 Y-STR based haplotypes showing relationships among our samples and African populations. The dashed circle highlights a possible main cluster formed. This plot was based on the matrix data of Table S3

5.CONCLUSION

The present work was an effort to evaluate the potential contribution of the use of Y-STRs to Population Genetics, as well as to other applications in Forensic Genetics. The overall analysis performed with the Y-STR data, allows the conclusion of the following aspects:

- The typing of 417 Brazilian male samples, residing in São Paulo state, with Yfiler™ kit corroborated published data and showed a high level of haplotype diversity, indicating applicability of these markers in individual identification and forensic studies,

- Significant genetic distances were found between our samples and Amazonas, Amapá, Pará, Roraima, Distrito Federal, Goiás, Minas Gerais, Santa Catarina e Rio Grande do Sul. By these analyses, special attention on the field of forensic genetics, such as the construction of a specific database, needs to be taken in the future;

- The results obtained with Haplogroup Predictor showed that the major genetic contribution for São Paulo population was European, followed by the African.

- The European Genetic contribution may have come mainly from Portugal and the African genetic contribution, mainly from Angola as no significant genetic distances were found between São Paulo population and Portugal nor Angola.

6.REFERENCES

1. Fisher, R.A., *The genetical theory of natural selection*. 1930, Oxford: Clarendon Press.
2. Haldane, J.B.S., *The causes of evolution*. 1932, London: Longmans Green & Co.
3. Wright, S., *Evolution in Mendelian populations.* Genetics, 1931. **16**(2): p. 97-159.
4. Hedrick, P., *Population Genetics and Ecology. The Princeton guide to ecology*. 2009, Princeton: Princeton University Press.
5. Relethford, J.H.H., R. M, *Population Genetics of Modern Human Evolution. Encyclopedia of life sciences*. 2001, London: Macmillan Publishers Ltd.
6. Chakravarti, A., *Population genetics - making sense out of sequence.* Nature genetics, 1999. **21**: p. 56-60.
7. Cavalli-Sforza, L.L.F., M. W., *The application of molecular genetic approaches to the study of human evolution.* Nature genetics, 2003. **33**: p. 266-275.
8. Avise, J.C., Arnold, J., Ball, R. M., Bermingham, E., Lamb, T., Neigel, J. E., Reeb, C. A. & Saunders, N. C, *Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics.* Annual review of ecology and systematics, 1987. **18**: p. 489-522.
9. Hickerson, M., Carstens, B., Cavender-Bares, J., Crandall, K., Graham, C., Johnson, J., Rissler, L., Victoriano, P. & Yoder, A, *Phylogeography's past, present, and future: 10 years after.* Molecular Phylogenetics and Evolution, 2010. **54**: p. 291-301.
10. Carracedo, A., *Forensic Genetics: History. Forensic Biology*. Advanced forensic sciences series, ed. M.M. Houck. 2016, Waltham: Elsevier Inc.
11. Jeffreys, A.J., Brookfield, J. F., Semeonoff, R, *Positive identification of an immigration test-case using human DNA fingerprints.* Nature, 1985. **317**(6040): p. 818-819.
12. Gill, P., Werrett, D. J, *Exclusion of a man charged with murder by DNA fingerprinting.* Forensic Science International, 1987. **35**(2-3): p. 145-148.
13. Jeffreys, A., J., AllenM., J., Hagelberg E., Sonnberg A, *Identification of the skeletal remains of josef mengele by DNA analysis.* Forensic Science International, 1992. **56**(1): p. 65-76.
14. http://www.innocenceproject.org/.
15. Primorac, D., Schanfield, M. (Eds.), *Forensic DNA Applications: An Interdisciplinary Perspective*. 2014: CRC Press.
16. Amorim, A., *Basic Principles. Forensic Biology*. Advanced Forensic Science Series, ed. M.M. Houck. 2016, Waltham: Elsevier Inc.
17. Landsteiner, K., *Zur Kenntnis der antifermentativen, lytischen und agglutinierenden Wirkungen des Blutserums und der Lymphe.* Zbl bakt, 1900. **27**(10): p. 357-362.
18. Bernstein, F., *Ergebnisse einer biostatistischen zusam-menfassenden betrachtung über die erblichen blutstrukturen des menschen.* Journal of molecular medicine, 1924. **3**(33): p. 1495-1497.
19. Sanger, F.e.a., *DNA sequencing with chain-terminating inhibitors.* Proceedings of the National Academy of Sciences of the United States of America, 1977. **74**: p. 5463–5467.
20. Southern, E.M., *Detection of specific sequences among DNA fragments separated by gel electrophoresis.* Journal of Molecular Biology, 1975. **98**: p. 503–517.
21. Goodwin, W., Linacre, A., & Hadi, S, *An introduction to forensic genetics*. Vol. 2. 2011, Chichester: John Wiley & Sons.
22. Jeffreys, A.J., Wilson, V., & Thein, S. L, *Hypervariable 'minisatellite' regions in human DNA.* Nature, 1985. **314**(6006): p. 67-73.

23. Jeffreys, A.J.e.a., *Individual-specific fingerprints of human DNA.* Nature, 1985. **316**: p. 76–79.
24. Mullis, K., Faloona, F., Scharf, S., Saiki, R., Horn, G. & Erlich, H, *Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction.* Cold Spring Harbor Symposia on Quantitative Biology, 1986. **51**: p. 263-273.
25. Butler, J.M., *Advanced Topics in Forensic DNA Typing: Methodology*. 2011, San Diego: Academic Press.
26. Butler, J.M., *Advanced topics in forensic DNA typing: interpretation*. 2014: Academic Press.
27. Bisso-Machado, R., Bortolini, M. C., & Salzano, F. M, *Uniparental genetic markers in South Amerindians.* Genetics and molecular biology, 2012. **35**(2): p. 365-387.
28. Underhill, P.A., Passarino, G., Lin, A. A., Shen, P., Mirazon Lahr, M., Foley, R. A., Oefner, P. J. & Cavalli-Sforza, L. L, *The phylogeography of Y chromosome binary haplotypes and the origins of modern human populations.* Annals of human genetics, 2001. **65**: p. 43-62.
29. Jobling, M.A.T.-S.C., *Fathers and sons: the Y chromosome and human evolution.* Trends in Genetics, 1995. **11**: p. 449-456.
30. Noordam, M.J., & Repping, S, *The human Y chromosome: a masculine chromosome.* Current opinion in genetics & development, 2006. **16**(3): p. 225-232.
31. Bogusz, M.J., *Handbook of Analytical Separations. Forensic Science*. Vol. 6. 2008, Amsterdam: Elsevier.
32. Jobling, M.A., Tyler-Smith, C, *New uses for new haplotypes: the human Y chromosome, disease and selection.* Trends in Genetics, 2000. **16**(8): p. 356-362.
33. Jobling, M.A., Pandya, A. R., Tyler-Smith, C. H, *The Y chromosome in forensic analysis and paternity testing.* International journal of legal medicine, 1997. **110**(3): p. 118-124.
34. Skaletsky H., e.a., *The Male-Specific Region of the Human Y Chromosome is a Mosaic of Discrete Sequence Classes.* Nature, 2003. **423**: p. 825-837.
35. Gusmão, L., Brion, M., González-Neira A., Lareu, M., Carracedo, A, *Y Chromosome Specific Polymorphisms in Forensic Analysis.* Legal Medicine, 1999. **1**: p. 55-60.
36. Graves, J.A., et al, *The origin and evolution of the pseudoautosomal regions of human sex chromosomes.* Human Molecular Genetics, 1998. **7**: p. 1991–1996.
37. Iida, R., Kishi K, *Identification, Characterization and Forensic Application of Novel Y-STRs.* Legal Medicine, 2005. **7**: p. 255-258.
38. Navarro-Costa, P., *Sex, rebellion and decadence: The scandalous evolutionary history of the human Y chromosome.* Biochimica et Biophysica Acta (BBA)-Molecular Basis of Disease, 2012. **1822**(12): p. 1851-1863.
39. Waters, P.D., Wallis, M.C., Marshall Graves, J.A, *Mammalian Sex - Origin and Evolution of the Y Chromosome and SRY.* 2007. **18**: p. 389- 400.
40. Ohno, S., *Sex Chromosomes and Sex-linked Genes*. Annals of Internal Medicine. Vol. 68. 1967: Springer-Verlag.
41. Charlesworth, B., *The evolution of sex chromosomes.* Science, 1991. **251**: p. 1030–1033.
42. Muller, H.J., *Genetic variability, twin hybrids and constant hybrids, in a case of balanced lethal factors.* Genetics, 1918. **3**(5): p. 422-499.
43. Charlesworth, B., Charlesworth, D, *The degeneration of Y chromosomes.* Philosophical transactions of the Royal Society of London Series B Biological Sciences, 2000. **355**: p. 1563–1572.
44. Graves, J.A., *Sex Chromosome Specialization and Degeneration in Mammals.* Cell, 2006. **124**: p. 901-914.

45. Schlegel, P.N., *The Y Chromosome.* Reproductive BioMedicine Online, 2002. **5**: p. 22-25.
46. Aitken, R.J.G., J. A. M, *Human spermatozoa: the future of sex.* Nature, 2002. **415**: p. 963.
47. de Knijff, P., *Messages through bottlenecks: on the combined use of slow and fast evolving polymorphic markers on the human Y chromosome.* The American Journal of Human Genetics, 2000. **67**(5): p. 1055-1061.
48. Butler, J.M., *Forensic DNA typing: biology, technology, and genetics of STR markers*. 2005, USA: Elsevier Academic Press.
49. Sachidanandam, R., et al, *A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms.* Nature, 2001. **409**: p. 928-933.
50. Venter, J.C., Adams, M.D., Myers, E.W., Li, P.W., Mural, R.J., Sutton, G.G., Smith, H.O., Yandell, M., Evans, C.A., Holt, R.A. and Gocayne, J.D, *The sequence of the human genome.* Science, 2001. **291**: p. 1304-1351.
51. Phillips, C., *Applications of autosomal SNPs and indels in forensic analysis. Forensic DNA Analysis: Current Practices and Emerging Technologies*. 2013: CRC Press.
52. Nachman, M.W.a.C., S.L, *Estimate of the mutation rate per nucleotide in humans.* Estimate of the mutation rate per nucleotide in humans, 2000. **Genetics**(156): p. 1.
53. Kondrashov, A.S., *Direct estimates of human per nucleotide mutation rates at 20 loci causing Mendelian diseases.* Human Mutation, 2003. **21**: p. 12-27.
54. Hammer, M.F., *A recent insertion of an alu element on the Y chromosome is a useful marker for human population studies.* Molecular Biology and Evolution, 1994. **11**: p. 749-761.
55. *http://www.isogg.org/tree/index.html*.
56. Underhill, P.A., Shen, P., Lin, A. A., Jin, L., Passarino, G., Yang, W. H.,Kauffman, E., Bonné-Tamir, B. Bertranpetit, J., Francalacci, P., Jenkins, T., Kidd, J., Mehdi, S.Q., Seielstad, M. T., Weels, R.S., Piazza, A., Davis, R. W., Feldman, M.W., Cavalli-Sforza, L.L., Oefner, P. & Ibrahim, M, *Y chromosome sequence variation and the history of human populations.* Nature genetics, 2000. **26**(3): p. 358-361.
57. Hammer, M.F., & Zegura, S. L, *The role of the Y chromosome in human evolutionary studies.* Evolutionary Anthropology: Issues, News, and Reviews, 1996. **5**(4): p. 116-134.
58. Wang, Y., Liu, C., Zhang, C. C., Li, R., Li, Y., Ou, X. L., & Sun, H. Y, *Analysis of 17 Y-STR loci haplotype and Y-chromosome haplogroup distribution in five Chinese ethnic groups.* Electrophoresis, 2015. **36**(20): p. 2546-2552.
59. Athey, T.W., *Haplogroup prediction from Y-STR values using an allele frequency approach.* Journal of Genetic Genealogy, 2005. **1**: p. 34–39.
60. Jobling, M.A., & Tyler-Smith, C, *The human Y chromosome: an evolutionary marker comes of age.* Nature Reviews Genetics, 2003. **4**(8): p. 598-612.
61. Jobling, M.A., *Y-chromosomal SNP haplotype diversity in forensic analysis.* Forensic science international, 2001. **118**: p. 158-162.
62. *www.familytreedna.com*. (04-13-2016).
63. Park, M.J., Lee, H. Y., Yang, W. I., Shin, K. J, *Understanding the Y chromosome variation in Korea—relevance of combined haplogroup and haplotype analyses.* International Journal of Legal Medicine, 2012. **126**: p. 589–599.
64. Schlecht, J., Kaplan, M. E., Barnard, K., Karafet, T., Hammer, M. F., Merchant, N. C, *Machine-learning approaches for classifying haplogroup from Y chromosome STR data.* PLOS Computational Biology, 2008. **4**(6).
65. Athey, T.W. *http://www.hprg.com*. 2004 (04-14-2016).

66. Petrejčíková, E., Čarnogurská, J., Hronská, D., Bernasovská, J., Boroňová, I., Gabriková, D., Bôžiková, A., Mačeková, S, *Y-SNP analysis versus Y-haplogroup predictor in the Slovak population.* Anthropologischer Anzeiger, 2014. **71**(3): p. 275-285.
67. Kayser, M., Roewer, L., Hedman, M., Henke, L., Henke, J., Brauer, S., Krüger, C., Krawczak, M., Nagy, M., Dobosz, T., Szibor, R, *Characteristics and frequency of germline mutations at microsatellite loci from the human Y chromosome, as revealed by direct observation in father/son pairs.* The American Journal of Human Genetics, 2000. **66**(5): p. 1580-1588.
68. Kayser, M., Sajantila, A, *Mutations at Y-STR loci: implications for paternity testing and forensic analysis.* Forensic Science International, 2000. **118**(2): p. 116-121.
69. Roewer, L., Amemann, J., Spurr, N. K., Grzeschik, K. H., Epplen, J. T, *Simple repeat sequences on the human Y chromosome are equally polymorphic as their autosomal counterparts.* Human Genetics, 1992. **89**(4): p. 389-394.
70. Kayser, M., et al, *A Comprehensive Survey of Human Y-Chromosomal Microsatellites.* American journal of human genetics, 2004. **74**: p. 1183-1197.
71. Lim, S.K., et al, *Variation of 52 new Y-STR loci in the Y Chromosome Consortium worldwide panel of 76 diverse individuals.* International Journal of Legal Medicine, 2007. **121**: p. 124–127.
72. Roewer, L., Krawczak, M., Willuweit, S., Nagy, M., Alves, C., Amorim, A., Anslinger, K., Augustin, C., Betz, A., Bosch, E. and Caglia, A, *Online reference database of European Y-chromosomal short tandem repeat (STR) haplotypes.* Forensic Science International, 2001. **118**(2): p. 106-113.
73. Kayser, M., et al, *Evaluation of Y-chromosomal STRs: a Multicenter Study.* International journal of legal medicine, 1997. **110**: p. 125-133.
74. Pascali, V.L., Dobosz. M., Brinkmann. B, *Coordinating Y-Chromosomal STR Research for the Courts.* International journal of legal medicine, 1998. **112**(1): p. 1.
75. Roewer, L., Kayser, M., de Knijff, P., Anslinger, K., Betz, A., Caglia, A., Corach, D., Füredi, S., Henke, L., Hidding, M. and Kärgel, H. J, *A new method for the evaluation of matches in non-recombining genomes: application to Y-chromosomal short tandem repeat (STR) haplotypes in European males.* Forensic Science International, 2000. **114**(1): p. 31-43.
76. *yhrd.org*. (04-16-2016).
77. Salazar, R.A.P., *Extending Y-STR Loci in Portugal for Forensic and Population*, in *Faculty of Sciences*. 2013, University of Porto: Porto.
78. Dupuy, B.M., Stenersen, M., Egeland, T., Olaisen, B, *Y-Chromosomal Microsatellite Mutation Rates: Differences in Mutation Rate Between and Within loci.* Human mutation, 2004. **23**: p. 117-124.
79. Bao, W., et al, *MSY2: a Slowly Evolving Minisatellite on the Human Y Chromosome Which Provides a Useful Polymorphic Marker in Chinese Populations.* Gene, 2000. **244**: p. 29-33.
80. Jobling, M.A., Heyer. E., Dieltjes, P., de Knijff, P, *Y-Chromosome-Specific Microsatellite Mutation Rates Re-examined Using a Minisatellite, MSY1.* Human molecular genetics, 1999. **8**: p. 2117-2120.
81. *http://www.ncbi.nlm.nih.gov/*. (04-16-2016).
82. Gusmao, L., Gonzalez-Neira, A., Pestoni, C., Brion, M., Lareu, M. V., Carracedo, A, *Robustness of the Y STRs DYS19, DYS389 I and II, DYS390 and DYS393: optimization of a PCR pentaplex.* Forensic Science International, 1999. **116**(3): p. 163-172.

83. Prinz, M., Ishii, A., Sansone, M., Baum, H., Shaler, R, *Y chromosome specific STR testing and the US legal system.* Excerpta Medica International Congress Series, 2000. **1193**: p. 591-594.
84. Buckleton, J.S., Bright,J., Taylor, D, *Forensic DNA evidence interpretation*. 2nd ed. 2016: CRC Press.
85. Thompson, J.M., et al, *Developmental Validation of the PowerPlex® Y23 System: A Single Multiplex Y-STR Analysis System for Casework and Database Samples.* Forensic Science International: Genetics, 2013. **7**: p. 240-250.
86. Olofsson, J.J., et al, *Forensic and population genetic analyses of Danes, Greenlanders and Somalis typed with the Yfiler® Plus PCR amplification kit.* Forensic Science International: Genetics, 2015. **16**: p. 232-236.
87. 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.* Journal of forensic sciences, 2006. **51**(1): p. 64-75.
88. Crocitti, J.J., Santa Barbara, M. M. V, *Brazil Today: An Encyclopedia of Life in the Republic*. 2010, Santa Barbara: Abc-clio.
89. Neves, W.A., Pucciarelli, H. M, *Morphological affinities of the first Americans an exploratory analysis based on early South American human remains.* Journal of Human Evolution, 1991. **21**: p. 261–273.
90. Oliveira, A.M., Domingues, P. M., Gomes, V., Amorim, A., Jannuzzi, J., de Carvalho, E. F., Gusmão, L78, *Male lineage strata of Brazilian population disclosed by the simultaneous analysis of STRs and SNPs.* Forensic Science International: Genetics, 2014. **13**: p. 264-268.
91. 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**(7).
92. Carvalho-Silva, D.R., Santos, F.R., Rocha, J., Pena, S. D. J, *The phylogeography of Brazilian Y-Chromosome lineages.* The American Journal of Human Genetics, 2001. **68**: p. 281-286.
93. Curtin, P.D., *The Atlantic slave trace: a census*. 1969, Milwaukee: The University of Wisconsen Press.
94. Reis, J.J., Gomes, F, *Liberdade por um fio. História dos quilombos no Brasil*. 1996, São Paulo: Companhia das Letras.
95. IBGE Instituto Brasileiro de Geografia Estatística, *Brasil: 500 anos de povoamento*. 2000, Rio de Janeiro: IBGE.
96. Fausto, B., *História Concisa do Brasil*. 2006, São Paulo: EDUSP.
97. *www.d-maps.com*.
98. Palacin, L., Moraes, M. A. S, *História de Goiás*. 1994, Goiânia: UCG Goiânia.
99. IBGE. *http://www.ibge.gov.br/apps/populacao/projecao/*. (04-17-2016).
100. *http://saopaulobrasil.net/PT/Silveira/Sao-Paulo/Sao-Paulo-Historia-Geografia-11*.
101. Ribeiro, D., *O povo brasileiro*. 1996, São Paulo: Companhia das Letras.
102. Sao-Bento, M., Carvalho, M., Bento, A. M., Andrade, L., Lopes, V., Serra, A.,Balsa, F., Anjos, M.J., Corte-Real, F, *Y-chromosome SNP analysis in the Brazilian population of Sao Paulo state (Ribeirao Preto).* Forensic Science International: Genetics Supplement Series, 2009. **2**(1): p. 427-428.
103. Tack, L.C., Thomas, M., Reich, K, *Automated forensic DNA purification optimized for FTA card punches and identifiler STR-based PCR analysis.* Clinics in laboratory medicine, 2007. **27**(1): p. 183-191.

104. Purps, J., et al, *A global analysis of Y-chromosomal haplotype diversity for 23 STR loci.* Forensic Science International: Genetics, 2014. **12**: p. 12-23.
105. Xu, H., et al, *Inferring population structure and demographic history using Y-STR data from worldwide populations.* Molecular Genetics and Genomics, 2015. **290**(1): p. 141-150.
106. Francez, P.A.D.C., Ramos, L. P. V., Palha, T. D. J. B. F., Santos, S. E. B. D, *Haplotype diversity of 17 Y-str loci in an admixed population from the Brazilian Amazon.* Genetics and molecular biology, 2012. **35**(1): p. 45-52.
107. Chemale, G., et al, *Y-chromosomal STR haplotypes in a sample from Brasília, Federal District, Brazil.* forensic Science International: Genetics, 2014. **9**: p. e3-e4.
108. Figueiredo, R.D.F., Ambrosio, I. B., Braganholi, D. F., Chemale, G., Martins, J. A., Gomes, V., Gusmão, L., Cicarelli, R. M, *Male-specific contributions to the Brazilian population of Espirito Santo.* International journal of legal medicine, 2016. **130**(3): p. 679-681.
109. Schwengber, S.P., Kommers, T., Matte, C. H., Raimann, P. E., Carvalho, B. A., Leite, F. P., Medeiros, M.A., Souza, L.F., Castro, C.S., Chassot, F.G., Bonatto, S. L, *Population data of 17 Y-STR loci from Rio Grande do Sul state (South Brazil).* Forensic Science International: Genetics, 2009. **4**(1): p. e31-e33.
110. Cainé, L.M., de Pancorbo PhD, M. M, *Y-chromosomal STR haplotype diversity in males from Santa Catarina, Brazil.* Journal of forensic and legal medicine, 2010. **17**(2): p. 92-95.
111. Excoffier, L., Lischer, H. E, *Arlequin Suite ver 3.5: a New Series of Programs to Perform Population Genetics Analyses Under Linux and Windows.* Molecular Ecology Resources, 2010. **10**: p. 564-567.
112. StatSoft, I.N.C., *STATISTICA (data analysis software system), version 6.* Tulsa, USA, 2001.
113. Qamar, R., Ayub, Q., Mohyuddin, A., Helgason, A., Mazhar, K., Mansoor, A., Zerjal, T., Tyler-Smith, C., Mehdi, S.Q, *Y-chromosomal DNA variation in Pakistan.* The American Journal of Human Genetics, 2002. **70**(5): p. 1107-1124.
114. Cardena, M.M., Mansur, A.J., Pereira, A.C., Fridman, C, *Y haplotype variability in a population of São Paulo state, Brazil.* Forensic Science International: Genetics, 2016. **23**: p. 22-23.
115. Chiaroni, J., Underhill, P. A., Cavalli-Sforza, L. L, *Y chromosome diversity, human expansion, drift, and cultural evolution.* Proceedings of the National Academy of Sciences of the United States of America, 2009. **106**: p. 20174–20179.
116. Cinnioglu, C., et al, *Excavating Y-chromosome haplotype strata in Anatolia.* Human Genetics, 2004. **114**: p. 127-148.
117. Wells, R.S., et al, *The Eurasian heartland: a continental perspective on Y-chromosome diversity.* Proceedings of the National Academy of Sciences, 2001. **98**: p. 10244-10249.
118. Gusmão, L., Brión, M., Gomes, I, *The human Y chromosome male-specific polymorphisms and forensic genetics.* Handbook of Analytical Separations, 2008. **6**: p. 969-1000.
119. Semino, O., Magri, C., Benuzzi, G., Lin, A.A., Al-Zahery, N., Battaglia, V., Maccioni, L., Triantaphyllidis, C., Shen, P., Oefner, P.J. and Zhivotovsky, L.A, *Origin, diffusion, and differentiation of Y-chromosome haplogroups E and J: inferences on the neolithization of Europe and later migratory events in the Mediterranean area.* The American Journal of Human Genetics, 2004. **74**(5): p. 1023-1034.
120. Poznik, G.D., Xue, Y., Mendez, F. L., Willems, T. F., Massaia, A., Sayres, Sayres, M.A.W., Ayub, Q., McCarthy, S.A., Narechania, A., Kashin, S. and Chen, Y,

Punctuated bursts in human male demography inferred from 1,244 worldwide Y-chromosome sequences. Nature Genetics, 2016. **48**(6): p. 593-599.

121. Rootsi S., M.C., Kivisild T., Benuzzi G., Help H., Bermisheva M., Kutuev I., Barac L., Pericic M., Balanovsky O., Pshenichnov A., Dion D., Grobei M., Zhivotovsky L.A., Battaglia V., Achilli A., Al-Zahery N., Parik J., King R., Cinnioglu C., Khusnutdinova E., Rudan P., E.Balanovska, Scheffrahn W., Simonescu M., Brehm A., Goncalves R., Rosa A., Moisan J.P., Chaventre A. , Ferak V., Furedi S., Oefner P.J., Shen P., Beckman L., Mikerezi I., Terzic R., Primorac D., Cambon-Thomsen A., Krumina A., Torroni A., Underhill P.A., Santachiara-Benerecetti A.S., Villems R., Semino O, *Phylogeography of Y-chromosome haplogroup I reveals distinct domains of prehistoric gene flow in Europe.* American journal of human genetics, 2004. **75**: p. 128-137.
122. Hammer, M.F., Karafet, T. M., Park, H., Omoto, K., Harihara, S., Stoneking, M., & Horai, S, *Dual origins of the Japanese: common ground for hunter-gatherer and farmer Y chromosomes.* Journal of human genetics, 2006. **51**(1): p. 47-58.
123. Semino, O., et al, *The genetic legacy of paleolithic Homo sapiens sapiens in extant Europeans: a Y chromosome perspective.* Science, 2000. **290**: p. 1155–1159.
124. Rootsi, S., Myres, N.M., Lin, A.A., Järve, M., King, R.J., Kutuev, I., Cabrera, V.M., Khusnutdinova, E.K., Varendi, K., Sahakyan, H. and Behar, D.M, *Distinguishing the co-ancestries of haplogroup G Y-chromosomes in the populations of Europe and the Caucasus.* European journal of human genetics, 2012. **20**(12): p. 1275-1282.
125. *http://www.eupedia.com/*.

SUPPLEMENTARY DATA

Sample code	DYS392	DYS391	DYS390	DYS393	DYS389I	DYS389II	DYS19	DYS385a	DYS385b	DYS458	DYS439	GATA_H4	DYS456	DYS437	DYS635	DYS448	DYS438
H001	13	10	23	13	13	16	14	11	15	16	12	12	16	15	23	19	12
H002	11	10	25	14	12	17	14	16	17	16	12	11	15	14	24	20	10
H003	13	11	24	13	14	16	14	11	14	17	12	12	16	15	23	19	12
H004	12	11	23	13	13	16	14	11	13	16	13	12	17	15	23	19	12
H005	13	11	23	13	13	16	14	11	14	17	11	11	16	15	23	19	12
H006	13	11	24	13	12	16	14	11	14	16	12	11	16	15	23	19	12
H007	13	11	25	12	13	16	15	13	14	18	11	11	16	15	23	19	10
H008	11	11	23	12	13	15	18	12	12	19	11	12	14	14	22	21	10
H009	12	11	22	13	14	17	15	15	16	17	11	11	13	15	21	20	10
H010	14	11	25	13	13	17	14	12	14	18	12	12	15	15	23	18	12
H011	13	11	24	13	13	17	15	11	14	16	13	12	17	15	23	19	12
H012	11	11	22	14	14	17	15	14	14	17	13	14	17	16	21	21	10
H013	13	11	24	13	13	16	14	11	14	17	11	13	16	15	23	19	12
H014	13	11	23	13	13	16	15	11	14	17	12	12	16	14	23	18	12
H015	11	10	23	13	13	15	17	12	12	17	11	12	14	15	21	21	10
H016	13	10	24	13	13	16	14	11	14	17	12	11	16	14	23	18	12
H017	11	10	22	13	11	16	16	13	14	15	11	11	14	16	21	20	10
H018	13	10	25	12	14	16	14	12	15	16	12	12	15	14	23	19	12
H019	11	10	22	12	13	16	14	13	16	14	11	12	15	15	21	20	9
H020	13	11	24	13	13	16	14	11	14	20	11	10	17	15	23	19	12
H021	11	10	22	14	13	17	15	17	19	16	12	11	17	14	21	20	11
H022	11	9	24	13	14	16	13	14	14	17	10	12	16	14	21	20	10
H023	13	11	24	12	13	17	14	11	13	16	11	12	15	15	23	18	12
H024	13	11	24	13	13	16	14	11	14	18	13	12	17	15	23	19	12
H025	13	11	24	13	14	17	14	11	14	17	12	12	17	15	23	19	13
H026	11	10	24	12	13	18	13	16	17	18	12	12	15	14	23	20	10
H027	11	10	22	12	12	16	14	14	14	15	11	11	14	16	24	20	10
H028	12	10	24	15	13	17	15	14	15	16	11	11	14	15	19	20	10
H029	13	10	24	13	13	15	14	11	13	16	12	11	15	14	23	19	12
H030	11	9	23	12	14	17	14	13	14	19	11	12	17	14	21	20	10
H031	11	10	25	13	13	16	13	15	17	16	12	11	17	14	23	21	10
H032	11	10	21	15	13	16	17	17	18	15	11	11	16	14	21	21	11
H033	13	11	24	13	13	16	15	11	14	18	12	13	15	15	23	19	12
H034	13	11	24	13	13	16	15	11	14	16	12	12	15	15	23	19	12
H035	16	10	22	13	13	17	13	15	17	16	12	10	15	14	22	20	11
H036	11	10	23	12	12	16	18	14	18	16	12	11	15	16	21	19	91
H037	11	10	23	13	12	16	14	13	15	15	11	11	14	16	21	20	10
H038	13	11	25	13	14	16	14	11	14	17	12	11	16	15	23	19	12
H039	11	10	22	13	12	16	14	13	14	15	11	11	15	16	21	20	10
H040	11	10	24	13	13	17	13	16	17	16	11	13	16	14	22	20	10

H041	13	11	24	13	13	16	14	11	14	17	11	13	15	15	23	20	12
H042	11	11	24	14	13	15	13	16	18	15	12	12	17	14	22	20	10
H043	13	12	24	13	13	16	14	11	14	15	12	12	15	15	24	19	12
H044	11	10	22	12	13	18	15	13	15	15	11	11	16	14	22	22	9
H045	11	10	26	12	13	16	14	13	13	20	11	11	15	15	21	21	9
H046	11	11	23	12	13	16	14	13	17	18.2	11	12	13	14	22	20	10
H047	13	11	24	13	12	16	14	11	14	18	11	12	16	15	23	19	12
H048	11	11	21	13	12	18	16	15	17	16	11	12	15	14	21	21	11
H049	13	10	24	13	13	15	14	13	13	17	12	12	15	15	23	19	12
H050	11	10	23	12	14	17	14	13	17	18.2	12	11	15	14	20	20	10
H051	13	11	24	13	14	16	14	11	14	16	12	13	15	15	23	19	12
H052	13	10	24	13	13	16	14	10	15	18	11	12	17	15	23	19	12
H053	13	11	25	13	12	16	15	11	14	18	12	12	16	15	23	19	12
H054	11	10	24	12	14	17	14	13	15	19	13	11	13	15	21	19	9
H055	11	11	23	11	13	17	15	14	16	18	12	10	16	16	22	20	10
H056	11	10	22	12	13	17	14	14	15	15	12	11	17	15	20	20	9
H057	11	12	21	14	12	17	16	15	15	16	12	11	16	16	22	22	10
H058	11	11	21	14	13	17	17	17	19	16	13	11	15	14	21	21	11
H059	11	10	22	13	12	17	14	12	15	15	11	11	14	16	24	20	10
H060	13	11	25	13	14	17	14	11	14	16	10	12	17	15	24	19	12
H061	13	11	24	13	13	17	14	11	15	17	12	12	16	15	23	19	12
H062	9	10	24	12	13	16	13	13	16	15	13	12	15	15	22	20	9
H063	11	10	22	13	12	17	14	13	14	17	11	11	14	16	22	20	10
H064	13	11	24	13	13	15	14	11	14	17	12	12	17	15	23	19	11
H065	13	11	24	13	14	16	14	11	14	17	12	11	17	14	23	18	12
H066	13	11	23	13	14	16	14	11	14	17	11	12	16	15	24	19	12
H067	13	11	24	13	14	15	14	11	14	18	12	11	16	15	24	19	11
H068	11	10	24	14	12	16	14	13	14	15	11	11	14	16	23	21	10
H069	11	10	25	12	13	17	13	15	16	17	13	11	15	14	21	20	10
H070	14	11	23	14	14	16	15	11	12	17	10	11	13	14	21	19	10
H071	13	10	24	13	13	16	14	11	14	18	11	12	18	14	23	18	12
H072	13	11	25	13	13	16	13	11	11	17	12	9	15	15	23	19	12
H073	13	11	24	12	13	17	14	11	13	16	11	13	15	15	23	18	11
H074	11	10	23	13	12	16	15	13	17	15	11	11	14	16	21	20	10
H075	13	10	23	13	13	16	14	11	13	18	13	12	17	16	23	20	12
H076	11	10	24	13	13	17	13	16	18	16	12	12	17	14	23	20	10
H077	11	9	24	13	14	16	13	13	13	18	10	12	16	14	21	20	10
H078	11	9	23	12	13	16	16	13	13	14	11	11	16	14	22	21	9
H079	13	10	26	14	13	16	14	12	14	19	12	11	15	15	23	19	12
H080	13	10	23	13	12	16	14	11	14	16	12	12	15	15	23	19	12
H081	13	10	24	13	13	16	14	11	13	17	12	12	17	17	23	19	11
H082	11	10	21	15	13	17	17	17	18	15	13	11	15	14	21	21	11
H083	11	10	21	13	12	17	17	16	17	18	13	12	15	14	22	21	11
H084	14	10	24	12	12	16	14	13	19	18	12	11	15	15	21	20	11

H085	11	11	22	13	13	17	15	14	15	16	13	11	15	17	20	18	8
H086	11	11	23	13	12	16	14	13	14	14	11	11	14	16	21	20	10
H087	13	11	23	13	13	16	14	11	14	16	12	12	15	16	24	19	12
H088	13	11	23	13	13	17	14	11	14	19	12	12	17	15	23	19	12
H089	14	10	23	13	13	18	15	14	18	17	11	12	16	14	21	19	9
H090	13	11	25	13	14	16	14	11	15	16	12	11	16	15	24	19	11
H091	11	10	24	13	13	18	14	17	19	16	12	12	15	14	21	20	10
H092	13	11	24	13	14	17	14	11	14	19	12	13	16	15	23	19	12
H093	13	11	24	13	13	16	14	11	15	18	13	12	15	15	23	19	12
H094	13	11	24	14	13	17	14	11	14	17	12	13	15	15	23	19	12
H095	13	11	23	13	13	16	14	11	14	17	12	11	17	14	23	19	12
H096	16	10	23	12	12	17	13	13	14	18	13	12	16	14	23	20	11
H097	11	9	22	12	13	17	14	13	17	16	11	11	15	15	25	21	9
H098	11	9	24	13	14	16	13	13	14	18	10	11	16	14	22	20	10
H099	13	11	24	13	12	16	14	11	15	16	12	12	17	15	24	19	12
H100	12	11	22	12	13	16	15	11	14	18	11	12	16	15	23	19	12
H101	13	10	24	13	13	17	14	11	14	17	12	12	15	15	24	19	12
H102	11	10	21	13	12	16	14	13	14	15	11	11	14	15	22	20	10
H103	11	10	22	13	12	16	14	13	15	15	11	11	14	16	23	20	10
H104	11	9	23	12	13	16	14	13	17	14	12	11	17	14	22	20	9
H105	12	10	24	13	13	16	15	11	15	18	12	11	15	15	23	19	12
H106	11	10	25	13	13	17	16	11	16	15	10	12	16	15	23	20	11
H107	11	10	22	14	12	17	15	14	14	16	11	12	15	16	20	21	10
H108	13	10	23	13	14	16	17	13	14	17	13	12	16	15	23	19	12
H109	13	11	22	13	13	16	14	11	14	19	12	11	16	14	23	18	12
H110	11	10	24	13	14	18	16	11	11	17	11	12	13	15	23	22	10
H111	12	10	22	14	14	17	15	13	17	15	11	11	13	14	23	20	10
H112	12	10	23	14	14	17	15	16	16	18	12	12	13	15	23	20	10
H113	13	10	24	13	13	16	15	11	15	17	11	11	16	15	23	18	12
H114	11	10	22	14	12	16	15	14	15	15	11	13	15	16	20	21	11
H115	13	11	24	13	14	16	14	12	14	18	12	12	16	15	23	19	12
H116	12	11	23	14	13	16	16	10	15	15	11	11	17	15	20	20	10
H117	11	9	24	13	14	16	13	13	14	18	11	12	16	14	22	20	10
H118	13	10	24	13	13	16	14	11	14	18	12	11	15	15	23	19	12
H119	11	10	21	13	13	18	15	15	15	17	12	12	16	14	22	20	11
H120	13	11	24	13	13	16	14	11	14	16	12	11	15	15	23	19	12
H121	11	10	21	14	14	16	14	13	15	20	10	11	14	17	22	21	10
H122	11	10	23	13	12	16	14	14	14	15	12	12	14	16	23	20	10
H123	13	11	24	13	13	16	15	9	14	18	11	11	16	15	23	19	12
H124	13	11	24	13	14	16	14	11	15	18	12	12	16	15	23	20	12
H125	13	11	24	13	13	16	14	10	15	16	12	12	18	15	23	19	12
H126	11	10	23	12	12	17	14	13	17	17.2	11	12	15	14	20	21	10
H127	11	11	25	14	13	18	15	11	13	15	10	12	17	14	23	20	11
H128	15	11	24	12	13	15	14	11	16	19	12	12	16	15	23	19	12

H129	13	10	25	14	13	16	15	11	14	15	12	12	16	15	23	18	12
H130	13	9	23	13	13	18	15	14	16	18	11	12	15	14	21	19	9
H131	13	11	23	13	13	16	14	11	14	16	13	12	16	15	23	19	12
H132	14	11	25	13	13	16	14	11	14	18	13	12	16	15	24	19	12
H133	13	11	23	14	13	16	14	11	14	17	12	13	15	15	23	19	12
H134	13	11	24	14	13	17	15	11	14	17	12	13	15	15	24	19	12
H135	12	10	23	14	14	18	15	15	15	15	11	11	15	14	21	20	10
H136	13	10	23	13	12	16	14	11	14	18	11	12	15	14	24	19	12
H137	13	11	23	13	13	16	14	11	14	17	13	12	15	15	23	19	12
H138	12	10	24	14	12	17	16	14	14	16	10	12	14	15	19	20	10
H139	13	10	24	13	14	17	14	11	14	17	13	13	16	15	23	19	12
H140	11	10	22	13	12	16	15	13	14	15	11	11	14	15	22	19	10
H141	13	11	24	13	13	16	14	11	14	18	12	12	15	15	23	19	12
H142	13	10	23	14	13	16	16	15	16	17	12	11	14	14	21	19	9
H143	13	10	24	13	13	15	16	12	12	15	13	11	14	14	23	19	11
H144	12	11		13	13	16	14	11	14	18	11	12	16	15	23	19	12
H145	13	11	23	13	14	16	14	11	16	17	11	13	17	14	20	19	13
H146	13	9	23	14	13	16	14	16	17	16	11	12	15	14	21	20	11
H147	11	10	23	13	13	17	13	17	17	15	13	12	16	14	23	20	10
H148	11	10	22	12	13	17	14	14	14	17	10	11	16	15	21	21	9
H149	13	11	24	12	13	17	14	11	15	17	12	11	15	15	23	19	12
H150	13	10	23	13	13	16	14	11	14	17	12	12	16	15	23	19	12
H151	11	10	23	15	13	17	15	14	14	16	13	12	15	15	22	21	10
H152	11	10	23	13	12	17	13	16	17	16	11	11	15	14	21	20	10
H153	11	11	23	13	12	16	14	14	14	16	12	11	14	16	21	20	10
H154	11	10	24	13	14	17	13	16	18	15	13	12	18	14	23	20	10
H155	11	10	21	12	14	18	16	17	17	17	12	11	14	14	20	21	11
H156	13	11	24	13	14	17	14	11	14	16	12	12	17	15	23	19	12
H157	13	11	24	13	14	16	15	12	15	17	13	12	15	14	23	19	12
H158	13	11	24	13	12	14	15	12	14	16	10	12	17	15	24	19	12
H159	13	10	24	13	14	17	14	11	14	21	11	12	16	16	23	19	12
H160	11	10	21	15	13	18	14	17	18	16	12	11	14	14	21	21	11
H161	11	10	25	13	13	16	14	17	18	16	11	12	16	14	22	20	10
H162	13	10	24	13	13	15	16	12	12	15	13	11	14	14	23	19	11
H163	11	10	21	14	13	16	17	18	19	15	13	11	15	14	20	21	11
H164	13	10	24	13	13	16	14	11	13	17	12	11	17	15	23	19	12
H165	11	10	21	15	14	17	16	17	18	15	13	11	16	14	22	18	11
H166	13	11	24	13	13	16	14	11	14	17	12	12	16	15	23	20	12
H167	13	11	23	12	13	16	14	11	14	16	12	13	15	15	23	19	12
H168	11	10	23	12	13	17	14	12	18	18.2	13	11	16	14	20	20	10
H169	11	10	24	14	13	17	16	12	12	18	11	11	15	15	21	20	10
H170	11	10	22	15	13	17	14	12	13	18	14	11	15	15	24	20	10
H171	11	10	24	14	12	17	15	13	18	16	11	11	14	16	21	20	10
H172	12	10	24	13	14	16	14	11	14	16	12	12	15	14	23	19	12

H173	11	10	22	13	12	16	14	13	14	15	11	11	14	16	22	20	10
H174	11	9	24	13	13	16	13	13	14	17	10	12	15	14	21	20	10
H175	11	10	23	12	12	16	14	13	17	18	11	11	15	14	22	21	9
H176	11	11	22	12	12	17	15	13	16	17	11	11	15	16	23	21	9
H177	13	10	23	13	13	16	14	11	13	16	12	12	18	17	23	19	12
H178	13	10	24	13	13	16	14	11	15	17	12	13	17	15	23	19	12
H179	13	10	24	13	13	16	14	11	14	18	12	12	16	15	23	20	12
H180	13	10	24	12	13	16	14	11	14	17	12	11	16	15	23	19	12
H181	13	11	24	13	13	16	14	11	14	17	14	12	16	15	24	19	12
H182	13	11	24	13	13	16	14	11	15	18	13	12	15	15	24	19	12
H183	13	11	24	14	13	16	14	11	14	17	13	12	15	14	24	19	12
H184	11	10	24	13	13	17	13	16	18	15	13	12	17	14	25	20	10
H185	13	11	24	13	13	16	14	11	14	18	12	12	15	14	23	20	10
H186	11	9	23	13	14	17	13	13	14	17	10	12	15	14	21	20	10
H187	13	11	24	13	13	16	14	11	14	17	12	12	14	15	23	19	12
H188	11	9	23	13	13	17	13	13	14	17	10	12	15	14	21	20	10
H189	13	11	24	15	13	16	14	12	15	18	12	12	16	15	23	19	12
H190	11	10	24	13	13	17	13	16	18	15	12	11	17	14	22	20	10
H191	13	11	24	13	13	16	14	11	14	17	12	12	17	15	21	19	13
H192	11	9	23	12	13	16	15	13	17	14	11	12	15	14	21	21	9
H193	11	9	23	12	13	16	17	13	13	14	11	12	16	14	23	21	9
H194	13	11	24	13	13	16	14	11	14	17	12	12	15	14	23	18	12
H195	11	10	23	14	13	17	13	16	16	16	11	11	16	14	20	21	10
H196	11	10	23	13	13	17	14	13	17	16	13	13	15	15	21	21	9
H197	11	10	23	13	13	17	13	16	17	16	12	11	17	14	21	18	10
H198	13	11	23	13	13	16	13	12	14	16	12	11	16	15	23	19	12
H199	10	10	22	13	13	17	16	15	15	18	12	11	15	16	21	21	10
H200	11	9	24	13	14	16	13	13	13	17	10	12	16	14	21	20	10
H201	11	10	21	15	13	16	15	16	20	16	11	10	17	14	21	21	11
H202	13	11	24	13	13	16	14	11	15	17	12	12	16	15	23	19	12
H203	13	11	24	13	13	16	14	11	13	16	12	12	16	14	23	18	12
H204	11	10	25	12	12	16	16	13	17	15	11	11	13	16	21	19	9
H205	13	11	23	12	13	17	14	11	14	17	12	11	15	15	23	19	12
H206	13	11	25	13	13	17	14	11	14	17	10	12	15	14	23	19	12
H207	13	11	24	13	14	16	14	11	14	18	12	13	16	15	23	19	12
H208	12	10	23	14	12	17	17	14	14	16	10	12	14	15	19	20	10
H209	13	11	24	12	13	16	14	11	14	16	12	13	15	15	23	19	12
H210	13	11	25	13	13	17	14	12	14	18	13	12	17	15	23	19	12
H211	13	10	23	14	13	16	16	15	17	15	11	11	16	15	21	20	8
H212	11	10	24	14	13	17	13	16	18	15	12	11	17	14	23	19	10
H213	11	10	23	12	13	16	13	14	16	16	12	12	15	15	21	21	9
H214	13	9	23	13	12	16	14	13	16	17	12	11	18	14	21	19	9
H215	11	10	23	13	13	15	16	12	13	16	11	12	14	15	22	20	10
H216	13	11	24	13	13	16	14	11	14	16	12	12	16	15	23	19	13

H217	11	10	23	12	13	16	15	13	18	15	12	11	15	15	21	22	9
H218	11	10	23	12	13	16	15	13	16	14	11	12	16	14	22	21	9
H219	13	11	24	13	13	16	14	11	15	18	13	12	15	15	24	19	12
H220	11	9	23	12	13	16	14	13	13	14	11	12	16	14	24	21	9
H221	11	10	23	15	14	16	15	11	20	15	11	11	14	14	20	20	10
H222	14	11	24	13	13	16	13	11	14	17	12	11	16	14	23	19	12
H223	11	10	23	12	13	17	16	13	16	16	10	11	15	15	23	20	9
H224	13	10	24	13	12	15	14	11	14	18	12	12	15	15	24	19	12
H225	13	11	24	13	13	17	14	11	15	17	11	12	15	15	23	19	12
H226	11	10	21	13	12	17	15	13	15	17	11	11	15	16	21	22	10
H227	12	10	24	13	13	17	13	16	17	17	12	12	15	14	20	20	10
H228	13	11	25	13	13	16	14	11	13	17	11	11	15	15	23	19	12
H229	11	10	23	13	12	16	14	14	14	16	12	12	14	16	24	20	10
H230	14	11	24	13	13	17	13	14	16	18	11	11	16	14	22	20	11
H231	11	10	26	12	13	16	15	13	13	20	11	11	15	15	21	21	9
H232	11	10	23	13	12	16	14	13	14	15	12	11	14	16	21	20	9
H233	11	10	21	14	13	17	16	17	17	16	12	11	16	14	21	21	11
H234	13	11	24	13	13	16	14	11	14	19	11	12	15	15	23	19	12
H235	11	10	21	14	13	17	17	17	19	15	11	11	16	14	21	21	11
H236	13	11	22	13	13	16	14	11	14	17	12	12	16	15	23	19	12
H237	13	10	22	13	14	16	16	10	20	18	12	12	15	14	21	18	13
H238	11	10	23	12	13	19	14	13	17	18.2	11	12	15	14	20	20	10
H239	13	12	24	13	14	17	14	11	14	18	12	12	15	15	23	19	12
H240	13	11	23	13	13	16	15	11	14	18	12	12	16	15	23	20	12
H241	14	10	23	13	13	16	14	11	15	17	12	12	15	15	23	19	12
H242	11	10	21	15	13	17	15	16	16	15	12	11	17	13	22	20	11
H243	11	10	23	12	13	16	14	13	17	17	12	12	15	15	22	21	9
H244	11	10	22	13	12	16	14	13	15	16	12	11	14	16	21	20	10
H245	11	10	21	13	12	18	16	17	17	18	11	11	14	14	21	21	11
H246	11	10	26	13	14	17	17	13	17	15	12	11	15	14	21	18	10
H247	11	10	23	13	12	16	14	13	14	15	11	11	14	16	22	20	10
H248	11	10	24	13	13	19	13	17	17	17	12	11	16	14	20	19	10
H249	13	9	24	13	13	16	13	11	14	16	11	12	15	15	23	19	12
H250	11	10	21	14	13	16	17	17	17	17	13	11	17	14	21	21	11
H251	13	10	23	13	13	17	15	14	16	17	11	11	15	14	21	19	9
H252	13	10	23	13	14	16	15	11	14	17	11	11	16	14	23	19	12
H253	13	10	24	12	14	16	14	11	14	17	12	11	15	15	23	19	12
H254	11	11	25	13	12	16	14	14	14	17	11	11	14	14	24	19	11
H255	11	10	23	13	13	18	13	15	16	15	12	11	15	14	21	20	10
H256	13	11	24	12	13	16	14	11	14	18	13	11	15	15	23	19	12
H257	13	11	24	13	12	14	15	12	14	16	10	12	16	15	24	19	12
H258	12	10	23	13	12	16	14	14	14	15	12	12	16	16	20	22	10
H259	13	11	24	15	13	17	15	11	14	17	12	11	15	15	23	20	12
H260	13	11	24	13	12	16	14	11	14	19	13	12	15	15	23	19	12

H261	13	11	25	13	13	16	13	11	14	17	12	12	15	15	23	19	12
H262	13	12	24	13	13	16	13	11	15	17	12	13	17	15	23	19	12
H263	11	10	24	12	13	16	14	14	17	20	12	12	15	15	21	18	9
H264	13	11	24	13	13	16	14	11	14	18	13	13	18	15	23	19	12
H265	13	11	24	13	13	16	14	11	14	17	11	11	15	15	23	19	12
H266	13	11	24	13	13	16	14	9	15	18	11	12	16	15	25	19	12
H267	11	9	24	13	14	16	13	13	14	18	10	12	17	14	21	19	10
H268	11	10	21	12	13	18	15	16	17	16	12	12	14	14	21	21	11
H269	13	11	24	13	12	16	15	11	15	19	12	12	16	15	23	19	12
H270	11	10	21	14	13	17	15	15	19	17	12	12	15	14	21	21	11
H271	13	10	23	13	13	18	14	11	14	16	12	11	16	14	23	18	13
H272	11	10	25	13	13	16	16	11	14	17	11	11	17	14	23	20	11
H273	11	9	24	13	14	16	13	13	14	19	10	12	16	14	21	20	10
H274	13	11	23	13	13	16	14	12	14	17	11	12	15	14	23	19	12
H275	11	10	21	14	13	16	15	13	15	16	11	11	15	15	21	22	10
H276	11	10	22	14	12	17	15	13	14	18	11	11	15	16	22	21	10
H277	13	10	25	14	13	16	14	11	14	16	14	12	15	16	23	19	12
H278	13	11	24	12	13	16	15	11	15	20	12	12	15	13	23	19	12
H279	13	10	24	13	13	16	13	11	14	16	12	12	15	15	23	19	12
H280	12	10	23	14	14	17	15	14	15	14	11	11	15	14	21	20	10
H281	11	11	21	14	12	16	15	15	16	15	12	11	17	16	22	22	10
H282	11	10	25	13	12	16	14	14	17	18	12	11	15	14	23	19	11
H283	11	9	25	14	14	17	13	16	17	17	12	12	15	14	23	20	10
H284	13	11	24	13	13	17	14	11	15	17	12	12	15	15	23	19	12
H285	14	10	24	13	13	16	14	11	14	19	13	11	16	14	23	18	12
H286	15	10	24	12	12	17	13	14	14	17	14	12	16	14	23	20	11
H287	13	11	24	13	14	16	14	11	14	17	12	11	16	15	23	19	12
H288	13	11	25	13	14	16	14	11	14	16	12	11	16	15	23	18	12
H289	11	9	24	13	14	16	13	13	14	18	10	12	15	14	21	20	10
H290	11	11	21	12	13	17	15	16	17	18	11	11	15	14	21	21	11
H291	11	10	23	15	14	15	15	10	21	17	12	11	15	14	22	21	10
H292	13	11	24	13	12	18	14	11	15	19	11	11	16	15	24	19	12
H293	13	10	24	13	14	17	15	14	16	16	12	11	15	14	21	19	9
H294	11	10	23	12	13	17	14	13	15	18.2	11	11	15	14	20	20	10
H295	11	11	21	13	13	18	15	16	17	15	11	12	15	14	21	20	11
H296	13	11	24	13	13	16	14	11	15	18	12	12	15	16	23	19	12
H297	11	10	23	12	12	18	14	13	14	16	11	11	15	15	21	20	9
H298	11	11	25	13	13	17	16	11	14	15	11	11	17	14	23	20	11
H299	13	11	24	13	13	16	15	10	15	18	12	12	15	14	23	19	13
H300	11	9	22	13	13	17	13	12	15	15	12	12	17	14	20	22	10
H301	11	11	23	13	12	16	14	14	20	17	11	11	15	14	23	19	11
H302	12	10	23	14	12	18	15	14	15	17	9	12	15	15	21	21	10
H303	11	10	24	12	13	17	13	16	16	16	12	11	17	14	21	20	10
H304	11	10	23	12	12	17	15	13	15	17	10	11	15	15	23	20	9

H305	13	11	24	13	14	15	14	11	14	17	13	12	16	15	24	19	12
H306	13	11	25	13	13	16	15	11	14	16	12	11	17	14	23	18	12
H307	13	11	24	13	13	16	14	11	14	17	13	12	16	15	23	19	12
H308	13	11	24	13	12	14	15	12	14	16	9	12	16	15	24	19	12
H309	11	10	25	12	12	17	13	16	16	18	12	11	15	14	21	21	10
H310	11	10	24	13	13	17	13	16	18	15	11	12	16	14	26	20	10
H311	11	9	24	13	14	16	13	13	14	18	10	12	16	14	21	20	10
H312	13	10	25	14	13	16	14	11	14	17	11	12	15	14	25	19	12
H313	11	10	22	13	12	16	15	13	14	15	11	11	14	15	23	19	10
H314	13	10	23	13	15	17	13	11	11	18	13	12	16	15	23	19	12
H315	13	10	24	13	13	16	14	11	15	20	11	12	16	15	25	19	12
H316	13	11	24	13	14	16	14	10	15	18	11	12	17	15	24	18	13
H317	11	10	23	12	13	16	15	13	13	15	12	12	15	14	21	21	9
H318	11	10	21	14	13	17	17	16	17	16	12	11	16	14	20	21	11
H319	13	10	24	12	13	16	14	11	14	17	12	12	15	15	23	19	12
H320	14	10	23	12	14	16	14	16	18	18	12	11	15	16	23	19	10
H321	11	10	24	13	13	17	13	16	18	16	13	12	15	14	22	19	10
H322	13	11	24	13	13	17	14	11	14	19	12	12	16	15	23	20	12
H323	13	11	24	13	12	16	14	11	14	17	13	12	15	14	24	20	12
H324	13	11	24	13	13	16	14	11	12	18	12	12	15	15	23	19	12
H325	11	11	25	13	13	17	16	11	14	15	11	11	17	14	23	19	11
H326	11	10	21	15	13	17	17	17	18	16	11	11	15	14	21	20	11
H327	13	11	24	13	14	16	14	11	14	16	11	12	16	15	23	19	12
H328	11	10	23	12	13	17	14	13	17	18.2	12	11	15	14	21	21	10
H329	11	10	23	12	12	17	15	13	16	17	11	11	15	15	23	20	9
H330	14	11	23	13	12	17	14	11	14	18	13	11	16	14	23	18	12
H331	13	11	24	13	13	16	14	11	15	16	11	12	16	15	23	19	12
H332	11	10	23	12	13	16	14	13	16	16	12	13	15	14	21	20	9
H333	11	10	21	15	13	17	17	18	19	16	13	12	16	13	21	21	11
H334	11	11	21	14	14	17	17	17	18	15	12	11	17	14	21	21	11
H335	13	10	24	13	13	16	13	12	14	19	12	10	17	15	24	19	12
H336	13	12	24	13	13	16	14	11	13	18	14	12	15	16	23	18	13
H337	11	11	21	14	13	18	15	16	17	16	11	11	15	14	22	21	11
H338	14	10	24	14	13	16	14	11	15	17	12	11	16	15	23	19	13
H339	12	11	23	14	13	15	16	14	15	16	11	11	17	15	20	20	11
H340	11	10	23	13	12	17	15	14	14	19	11	11	15	16	21	20	10
H341	14	11	25	13	13	18	13	14	14	16	12	12	15	15	23	20	11
H342	11	9	24	13	14	16	13	14	14	18	10	12	16	14	21	20	10
H343	11	9	24	14	13	18	13	15	19	16	14	12	15	14	21	20	10
H344	13	11	24	13	13	16	14	11	14	15	12	12	15	14	23	18	13
H345	13	11	26	13	12	15	14	11	14	17	12	12	16	15	24	19	12
H346	13	10	24	13	13	16	14	11	15	16	12	12	16	15	23	19	12
H347	11	10	22	13	12	17	15	14	14	15	11	12	15	16	20	21	10
H348	11	10	24	13	12	19	13	15	17	18	10	12	14	15	23	20	10

H349	11	10	24	13	12	17	13	9	17	14	10	11	14	14	21	20	10
H350	13	10	24	13	13	16	14	11	13	17	12	12	17	16	23	19	11
H351	14	10	24	13	15	15	14	11	14	18	12	12	16	15	23	19	12
H352	14	11	24	12	13	15	14	10	15	17	12	12	15	15	23	19	12
H353	13	11	23	13	12	17	15	11	14	18	12	12	16	15	23	19	13
H354	13	10	23	14	13	16	16	15	15	15	11	11	14	15	21	20	8
H355	11	9	24	13	14	16	13	13	14	17	10	12	16	14	21	20	10
H356	11	10	21	14	13	17	15	16	17	16	12	11	16	13	22	20	11
H357	14	10	23	13	13	16	15	11	13	17	11	12	14	14	22	19	10
H358	13	11	24	13	14	16	14	11	14	18	11	12	16	15	23	19	12
H359	13	11	24	13	13	16	14	11	14	17	11	12	16	14	23	19	12
H360	11	11	23	13	13	17	13	15	19	15	12	12	17	14	21	20	10
H361	13	11	24	12	14	16	14	11	15	16	12	12	16	15	23	19	12
H362	13	10	23	13	13	16	14	11	13	17	13	12	17	16	23	20	12
H363	13	11	24	13	14	16	14	11	16	17	12	11	15	15	24	19	12
H364	11	9	23	12	13	17	14	13	17	18.2	11	11	14	14	23	19	10
H365	11	10	23	13	13	17	13	17	18	14	11	11	15	14	21	18	10
H366	13	11	24	14	13	16	14	12	15	17	12	12	15	15	24	19	12
H367	11	9	25	12	14	15	14	13	18	18.2	11	10	15	14	20	20	10
H368	11	10	21	14	12	17	15	13	15	18	13	11	16	16	21	21	10
H369	13	11	23	13	13	16	16	11	14	17	12	12	16	15	23	18	12
H370	11	11	25	13	14	17	16	12	17	14	12	11	15	14	21	20	10
H371	11	9	25	13	13	16	13	14	14	18	10	12	16	14	21	20	10
H372	13	11	24	13	13	16	15	11	14	18	11	12	16	15	23	20	12
H373	13	10	23	13	13	16	14	12	14	17	11	12	16	15	23	19	12
H374	11	10	23	12	13	16	17	14	17	15	10	13	15	16	23	21	9
H375	11	10	22	12	13	18	15	13	15	15	11	11	15	15	22	21	9
H376	11	10	24	13	14	16	13	16	17	16	13	11	15	14	23	20	10
H377	13	11	23	13	13	16	14	11	14	17	12	11	16	15	24	19	12
H378	16	10	22	11	13	16	13	12	16	20	12	11	17	15	21	20	10
H379	13	11	24	13	13	16	14	11	14	18	12	13	16	15	23	19	10
H380	11	10	21	14	13	16	17	17	17	17	13	11	17	14	21	21	11
H381	14	10	23	12	13	17	14	17	18	16	11	11	14	16	22	20	10
H382	11	9	25	13	13	17	13	14	15	18	10	12	16	14	21	20	10
H383	13	11	24	13	13	16	14	12	14	17	11	12	15	15	23	18	12
H384	13	11	23	12	13	16	14	11	14	17	12	13	15	15	23	19	12
H385	12	11	22	13	14	17	15	15	16	17	11	11	13	15	21	20	10
H386	14	10	22	15	14	16	15	15	17	15	11	11	14	14	22	20	10
H387	11	10	25	13	13	16	15	11	15	15	11	12	17	14	23	20	11
H388	11	11	24	12	13	17	14	14	15	18	12	11	15	15	22	19	9
H389	11	11	24	13	14	16	13	13	14	18	11	12	16	14	21	20	10
H390	11	9	24	13	14	16	13	13	14	17	10	12	16	14	21	20	10
H391	11	10	22	14	13	18	15	13	15	18	12	12	15	16	23	21	10
H392	13	10	24	13	13	15	14	11	14	19	12	13	16	15	23	18	12

H393	11	10	23	13	14	17	14	14	15	17	11	11	17	14	23	21	9
H394	11	10	23	12	13	16	14	13	16	16	11	13	15	14	22	20	9
H395	11	10	23	12	14	17	14	12	18	19.2	11	10	16	14	21	21	10
H396	11	9	21	14	13	17	14	17	17	16	12	11	14	14	24	19	11
H397	13	10	24	13	13	16	14	11	14	17	14	12	15	15	23	19	12
H398	11	10	23	14	14	18	13	13	16	18	12	12	15	14	20	18	10
H399	14	11	23	13	13	16	14	11	14	18	13	11	16	15	23	19	12
H400	11	10	21	15	13	16	15	17	19	16	11	11	17	14	21	21	11
H401	13	11	24	13	14	17	14	11	14	17	14	12	17	15	23	19	12
H402	13	11	25	13	13	16	14	11	14	17	12	13	16	15	23	20	12
H403	11	10	21	14	14	18	16	18	19	17	11	11	17	14	21	21	11
H404	11	11	24	12	12	16	16	14	18	18	12	11	13	16	21	19	9
H405	11	10	25	13	13	18	13	16	16	19	12	11	15	14	22	20	10
H406	13	11	25	13	13	16	15	11	14	18	12	12	16	15	23	19	12
H407	13	12	24	13	12	16	15	11	14	16	13	12	16	15	23	19	12
H408	12	11	25	13	13	16	14	11	14	18	11	12	16	15	23	19	12
H409	11	10	22	15	13	17	15	14	15	16	13	12	15	16	21	21	10
H410	12	10	24	12	13	18	14	11	15	16	12	12	16	15	24	20	12
H411	11	10	23	12	13	17	14	13	15	17	11	11	17	15	23	20	9
H412	11	10	21	13	13	17	15	14	18	16	12	12	15	14	21	21	11
H413	13	11	24	13	13	16	14	11	14	18	12	12	15	15	23	19	12
H414	11	10	21	14	13	18	15	15	19	16	12	11	15	13	21	21	11
H415	11	10	24	13	13	19	13	17	17	18	11	11	15	14	21	20	10
H416	13	10	24	13	13	16	14	12	14	18	12	11	15	15	23	18	12
H417	11	10	21	15	14	17	17	17	19	17	11	11	17	14	22	21	11

Table S2 - Matrix of the R_{ST} genetic distances within the samples from present study and other 20 Brazilian population. These calculi were achieved through 17 Y-STRs *loci*.

	Present Study	Amazonas	Acre	Amapa	Distrito Federal	Ceara	Mato Grosso do Sul	Espirito Santo	Rio Grande do Sul	Rio de Janeiro	Pernambuco	Parana	Para	Rondonia	Roraima	Maranhao	Goias	Minas Gerais	Tocantins	São Paulo	Santa Catarina
Present Study	*																				
Amazonas	0.06161	*																			
Acre	0.00712	0.02880	*																		
Amapa	0.00489	0.05419	0.00790	*																	
Distrito Federal	0.01081	0.06610	0.02341	0.00383	*																
Ceara	0.00079	0.06426	-0.00079	0.01047	0.01687	*															
Mato Grosso do Sul	-0.00483	0.06715	0.01282	-0.00038	0.00387	0.00286	*														
EspiritoSanto	0.00003	0.05744	0.00004	0.00363	0.00934	-0.00351	-0.00489	*													
Rio Grande do Sul	0.00492	0.06898	0.01115	0.00145	0.00441	0.00744	-0.00410	0.00273	*												
Rio de Janeiro	-0.00076	0.05354	0.00345	0.00710	0.01234	-0.00212	-0.00365	0.00077	0.00700	*											
Pernambuco	0.01128	0.07355	0.01017	0.01149	0.00384	0.01137	-0.00202	0.00611	0.00125	0.01228	*										
Parana	-0.00583	0.06834	0.00164	0.00445	0.00875	-0.00828	-0.00887	-0.00419	0.00192	-0.00547	0.00566	*									
Para	0.01731	0.07377	0.01898	0.00744	0.01156	0.02474	0.01171	0.02017	0.01052	0.01863	0.01874	0.01343	*								
Rondonia	-0.00121	0.06602	0.01060	0.00166	0.00519	0.00054	-0.00786	-0.00141	0.00209	0.00069	0.01269	-0.00443	0.01192	*							
Roraima	0.02885	0.07662	0.01631	0.01923	0.01856	0.03044	0.01999	0.02652	0.01527	0.02887	0.00526	0.02351	0.01178	0.02695	*						
Maranhao	0.00293	0.07402	0.01155	0.01829	0.02814	0.00159	0.00766	0.00723	0.02064	0.00391	0.02529	-0.00348	0.03126	0.00793	0.04978	*					
Goias	0.01437	0.06743	0.00538	0.00416	0.01162	0.01731	0.00709	0.00995	0.00778	0.01572	0.00213	0.01075	0.01155	0.01165	0.02862	0.02379	*				
Minas Gerais	0.31463	0.34660	0.35853	0.34672	0.32644	0.31119	0.29832	0.31689	0.34890	0.30496	0.36057	0.30722	0.37571	0.31595	0.38033	0.30369	0.37253	*			
Tocantins	0.00411	0.05870	-0.00091	0.02914	0.03436	0.00192	0.01232	0.01189	0.03069	0.00089	0.02753	-0.00193	0.03497	0.01509	0.04354	-0.00076	0.03309	0.30682	*		
São Paulo	0.00060	0.05013	-0.00292	0.00192	0.00982	-0.00058	-0.00332	-0.00069	0.00334	0.00055	0.00468	-0.00296	0.01319	0.00008	0.02045	0.00818	0.00687	0.33137	0.00787	*	
Santa Catarina	0.05632	0.14695	0.10224	0.04849	0.02304	0.06878	0.02573	0.05503	0.03797	0.05818	0.02973	0.05328	0.04835	0.04801	0.06048	0.08083	0.04772	0.33961	0.10850	0.05757	*

Table S3- Matrix of the R_{ST} genetic distances within the samples from present study and European populations. These calculi were achieved through 17 Y-STRs loci.

	Present Study	Portugal	Italy	Spain	Germany
Present Study	*				
Portugal	-0.00010	*			
Italy	0.00754	0.01731	*		
Spain	0.01079	0.00763	0.03358	*	
Germany	0.01854	0.02362	0.02114	0.03002	*

Table S4- Matrix of the RST genetic distances within the samples from present study and African populations. These calculi were achieved through 17 Y-STRs loci.

	Present Study	Angola	Mozambique
Present Study	*		
Angola	0.00276	*	
Mozambique	0.05444	0.01645	*