PCR-based DNA profiling of human Y chromosomes

Fabrício R. Santos[1], Denise R. Carvalho-Silva[2], Sérgio D.J. Pena[2]

Departamento de Biologia Geral[1] and Departamento de Bioquímica[2],

Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais,

Caixa Postal 486, 30161-970 Belo Horizonte, Brazil

All correspondence and proofs should be addressed to SDJP

Fax : (5531) 227-3792

e-mail: spena@dcc.ufmg.br

Summary

Until recently, the use of human Y chromosome DNA polymorphisms for evolutionary studies and in human identification had been cramped by a scarcity of suitable markers. Fortunately this picture has changed after the discovery of an extensive range of new polymorphisms in the past five years. These are convenient markers, easily amplified by PCR, that fall into two large groups: slow evolving biallelic markers (single nucleotide polymorphisms and retroelement insertion polymorphisms) and fast evolving multiallelic tandem repeat polymorphisms (microsatellites and minisatellites). In this review we provide simplified detection protocols for the typing of the most important PCR-based polymorphisms and discuss their use in population genetic studies and forensic casework, emphasizing the need to adequate the rate of evolution and information content of the marker to the specific application intended.

1 Introduction

Differences between polymorphisms in the Y chromosome and autosomes

DNA polymorphisms in the human genome are excellent markers of human individuality because of the diploidy and their recombination rate. In evolutionary time spans new mutations are quickly shuffled by recombination, generating genetic uniqueness. The situation of autosomal and X-chromosomal markers contrasts markedly with that of the Y chromosome, which exhibits in most of its length (Y-specific part) a haploid state and lack of recombination. Thus, Y chromosomes are transmitted by males to their male offspring and remain unaltered from generation to generation, establishing genetic lineages (patrilineages) that remain stable until a mutation supervenes. Thus, by the study of DNA polymorphisms of Y chromosomes of a given biological sample one can determine with great precision that it belongs to a certain lineage, but not to a specific individual. Human Y chromosomal DNA polymorphisms are consequently lineage markers and hence have special utility in evolutionary studies [1, 2] as well as forensic application in criminal and paternity casework [3-6].

The unusual properties of the Y chromosome derive from its function in driving the male sex development in mammals. In consequence of this specialized function there was loss of genes that were present in the proto-Y chromosome, which was probably almost entirely homologous to the proto-X chromosome [7]. During this evolutionary process, the Y chromosome has decreased in size and accumulated mutations and repeats of several classes including tandem repeats, transposons and endogenous retrovirus [8]. At the same time, the Y acquired by translocation from autosomes some special genes involved in spermatogenesis [9]. In the modern human Y chromosome several loci with X homology remain [8],

particularly in two regions on both Y chromosome extremities completely homologous to the X chromosome ends, named *Yp* and *Yq* pseudo-autosomal regions (PARs). The *Yp* (2.6 Mbp) and *Yq* (0.32 Mbp) PARs allow X-Y pairing and correct segregation during meiosis, when recombination with the X occurs. Between the two PARs, most of the chromosome (~50 Mbp) correspond to the haploid and non-recombining Y-specific region.

DNA variation in the Y chromosome is a product of mutations that accumulate along successive father-to-son generations allowing patrilineages to be traced back to common ancestors using the knowledge of the ancestral and derived allele states for each Y locus [1,2]. The common ancestor for all living Y chromosomes has been dated with large uncertainties between 37,000 to 411,000 years ago [10,11], but still supporting a recent origin for modern humans.

DNA variability of the human Y chromosome specific region

Although the human Y chromosome exhibits considerable amounts of cytogenetic heteromorphisms [12,13], searches for DNA polymorphisms were initially not successful [14-16] giving origin to the idea that this chromosome was virtually molecularly monomorphic. The sequence hypovariability of the Y was confirmed by sequence studies of large regions and explained as either the consequence of selective sweeps [10] or simply as the expected outcome of its reduced effective population size, which is ¼ of autosomes [11].

The first polymorphic markers described on the Y chromosome were relatively complex and depended on Southern blotting for detection [14,17-22]. More recently many convenient PCR-based markers became available [10,11,23-28]. These include several types of polymorphic variation, ranging from polymorphisms created by very rare events such as point mutations and retroelement insertions to the fast evolving microsatellites and minisatellites. These several types of polymorphisms are studied for different purposes. Tandem repeat

polymorphisms are multiallelic and hypervariable, thus being highly informative. Because of this, they are the markers of choice for paternity casework and criminal investigations. However, because of this fast evolution and the fact that microsatellites evolve mostly by forward-backward stepwise mutations, there is frequent homoplasy in human populations, complicating evolutionary studies. In contrast, the single nucleotide polymorphisms (SNPs) and especially insertion polymorphisms are slow evolving markers and each mutation can be considered a unique event, what has great usefulness in evolutionary studies. On the other hand, they are biallelic polymorphisms with low information content. The availability of polymorphisms with a broad range of rates of evolution and informativeness, thus establishes a menu from which one can choose the most suitable. The objective of this review is to survey PCR-based polymorphisms in the Y-specific region of the human Y chromosome and to present the methodology for their use.

Biallelic (binary) markers

DNA sequence polymorphisms (SNPs)

SNPs are a common type of human genetic variation representing a position at which two (or more) alternative bases occur at frequencies > 1 % in human haploid genomes. They are generated by single base substitutions or small insertion/deletion events occurring at a very low rate (μ) of about 10^{-7} or less per site per generation [11]. Thus SNPs are slowly evolving markers usually occurring just once. They are also called binary markers because usually only two alleles are observed, the ancestral (0) and derived (1) states. To determine which allele is ancestral and which is derived we need outgroup species, such as other primates, to ascertain the most likely ancestral state of that locus, which in many cases proves to be the most common allele [10,11,27-29]. The first SNP in the human Y chromosome was A or G at

position 168 of the locus *DYS271* [24] and the ancestral state, as indicated by the analysis of chimpanzees and gorillas, was A. The derived allele G has shown a very restricted geographic dispersion in Africa. In general, the range of geographical distribution of a binary marker is proportional to the time when the mutation has occurred and the migration rate of males from that area after its origin time. A similarly restricted geographic distribution was found with the transition T **à** C in the *Tat* marker that is present in populations of Northern Eurasia [29], and the transition C **à** T in the *DYS199* locus found only in Amerindians [27]. On the other hand, geographically widespread mutations are expected to be older, some of them having appeared before the major out-of-Africa migrations, over 100,000 years ago [30]. Therefore, mutations such as *DYS271, DYS199* and *Tat* are expected to be younger, and indeed the former two were estimated to have originated around 30,000 [27,31] while the latter is more recent, 4000 years ago [29]. Several other SNPs have been described recently [10,11,28,31] with ancestral and derived states determined for all of them. An example of an ancient mutation with widespread geographical distribution is a transversion C **à** G called *M9* [28] that is found in many individuals in all continents except Africa. It probably occurred quite soon after the first emigrants came out-of-Africa, because most of the Y chromosomes from Europe, Asia, Americas and Oceania show the *M9* mutation. Subsequent mutations, such as in the loci *DYS199*, *DYS234* and *DYS273* [28] have a significantly more restricted geographic distribution that is actually a subset of the *M9* distribution. The association of many Y markers arisen in different evolutionary times allows the derivation of binary marker haplotypes that can be linked by sequential mutation steps in a single human phylogenetic tree [1,2,28]. If, however, reverse mutations or recurrent mutation events happen, trees can become complex because of the presence of loops. In a set of 29 binary markers [10,11,28,29] only a transition A **à** G at position -1532 of the locus *SRY* seems to have mutated back to the ancestral state [2].

Retroelement insertion polymorphisms

Another source of human genome variability are insertion and deletion polymorphisms of retroelements such as SINEs, LINEs and endogenous retroviruses (or fragments thereof, such as LTRs). Among these, the most common are Alu insertion polymorphisms [32]. Alu sequences represent the largest family of short interspersed nucleotide elements (SINEs) in primates with an excess of 500,000 copies per haploid genome. Among all Alu sequences, a very small proportion, numbering 500-2000, is relatively human specific and a subset of these, only recently inserted into the human genome, is polymorphic. Their rate of insertion in a specific site of the human genome is expected to be much smaller than mutation rates for nucleotide substitutions [11]. Alu insertions are similar to SNPs in being biallelic and thus having small information content. However, the sequencing of an outgroup species is not necessary because the molecular nature of the retrotransposition event already points to chromosomes without the insertion as the ancestral. Another important characteristic is that reverse mutations are virtually impossible because there is no known mechanism capable of excising precisely the retroelement after the insertion [32]. The only published polymorphic retroposon in the human Y chromosome is the YAP marker, an Alu element of the active Sb2 subfamily with a likely origin around 135,000 years ago, probably in Africa [31], although an Asian origin has been recently suggested [33]. A new polymorphic L1 insertion in the human Y centromere has been recently isolated (F.R. Santos and C. Tyler-Smith, unpublished data) and there may be some few other polymorphic retroelements in the human Y chromosome still to be discovered.

Applications of binary Y chromosomal markers

Analyses of Y chromosome haplotypes using binary markers are very useful for human evolutionary studies as they reflect the ancient male migration paths that gave rise to modern

geographic groups [1,2]. The recent discovery of several new binary markers [28] hopefully will allow resolution of the patrilineages colonizing all continents. In addition, they can be used to estimate population admixture levels [34] and distinct geographic contributions to the present population of countries such as Brazil ([35], D.R. Carvalho-Silva and S.D.J. Pena, unpublished data). Y chromosome studies can be readily compared to mtDNA analysis, thus contrasting paternal and maternal contributions [36].

Their low informativeness and geographic correlation limit the application of binary markers in human identification for forensic and paternity cases. Thus these markers are more useful for identification of human groups than individuals. Nevertheless, in populations with a well known haplotype distribution, there may be cases where the likely geographic origin can be used in investigations [6].

Microsatellites

The first microsatellite described in the human Y chromosome was a GATA tetranucleotide repeat in the locus *DYS19* [23] that displayed a considerable variation in extensive worldwide surveys [37]. Several new Y microsatellites were described recently [1,5,22,38]. Some of these, such as *DYS389* [39], have a complex tandem repeat structure that involves distinct repeat blocks displaying apparently independent evolution.

The average mutation rate (μ) for Y microsatellites was estimated around 2×10^{-3} per locus per generation [40]. Such high levels inevitably will lead to the occurrence of several recurrent microsatellite haplotypes (homoplasies). These properties indicate that microsatellites will be most informative for recent evolutionary events, before a mutation-drift equilibrium has been reached and while their variability is still proportional to the evolution time [41]. Considering $\mu = 2 \times 10^{-3}$ [40] and an effective population size (N_e) for human Y chromosomes around

4500 individuals [29], the divergence of seven Y microsatellite loci measured as the allele size variance [41] would be proportional to time until 2,128 generations (95 % confidence level) or 42,560 years considering a generation time of 20 years. Previous phylogenetic analysis using microsatellite haplotypes [42-44] has revealed complex evolutionary networks without strong correlation with geography or population groups. A more successful approach was obtained associating microsatellite haplotypes to the *Tat* mutation [29]. In that case, Y chromosomes sharing the same derived allele C in the *Tat* marker were typed for several Y microsatellites and a network of these related chromosomes was used to investigate the migration involving males carrying the *Tat* mutation. A date of origin for *Tat* was calculated around 4000 years ago, a short time to allow divergence in the microsatellites without generating much recurrence as displayed in the compact evolutionary network obtained [29]. The most striking results were the clustering of microsatellite haplotypes of the same populations and the suggestion of a migration route of chromosomes bearing the *Tat* C mutation. However, the identification of the ancestral Y microsatellite haplotype seems to be still the major difficulty. Thus the association of microsatellites to slower evolving markers, especially those arising in a recent evolutionary time such as *Tat* [29], can be very informative for studying human patrilineages.

Recently, we used the tetranucleotide microsatellite *DYS19* in combination with the αh alphoid DNA polymorphism [26] to establish the existence of a major founder Y chromosome haplotype (allele A – 186 bp – at *DYS19* and type II at αh) in South and North Amerindians [45-47]. Our results were confirmed by a subsequent study [27], which has shown that besides the allele A at *DYS19*, the founder haplotype was also characterized by a specific single base change (C → T) at the *DYS199* locus. This discovery was only possible because after more than 10,000 years of the Asian migration into the Americas, *DYS19* remained virtually monomorphic in Amerindians, an observation that runs contrary to the expectations of fast

evolution for *DYS19*, since studies in other populations had revealed high levels of variability [37]. The reason for this slow evolution of *DYS19* in Amerindian patrilineages has not been understood in spite of intense studies (D. R. Carvalho-Silva, F.R. Santos and S.D.J. Pena, unpublished data).

Utilization of Y-linked microsatellites for human identification

In paternity testing Y-linked microsatellites are especially useful in cases where the possible father is already deceased and in which the proband is male [4,48]. However, the use of Y chromosome microsatellites in these applications is subject to the *caveat* that they are lineage markers and do not provide individualization. Thus, although absolutely reliable for exclusions, in case of a match they only indicate that somebody in the patrilineage of the tested suspect is the possible perpetrator. Moreover, even the combined use of several Y-linked microsatellites does not allow absolute discrimination (for Y-linked markers the discrimination probability, i.e. the probability that two randomly chosen individuals differ in typing, the exclusion power, i.e. the probability of excluding the paternity of a falsely accused individual, and the gene diversity are numerically identical). For example, with the use of *DYS19*, *DYS390*, *DYS391* and *DYS393*, which can be easily co-amplified and run as a tetraplex system (Fig.1), we, reach a discrimination probability of 0.95 in the Brazilian population. If we add *DYS389A* and *DYS389B* we increase the discrimination probability to 0.98 and if we add further the alphoid system (see below) we basically maintain the same discrimination probability of 0.98 (D.R. Carvalho-Silva, F.R. Santos and S.D.J. Pena, unpublished results). Thus, even if we add several other markers it is unlikely that we will approach certainty much closer, for two main reasons. First, patrilineages are far reaching and the frequency of unsuspected non-paternity in the population is considerable. Thus, two individuals who are apparently quite unrelated may actually belong to the same patrilineage.

Second, two males presenting the same microsatellite haplotype may represent homoplasy of two distinct patrilineages. In the latter case the use of slowly evolving markers can define the real haplogroups that they belong to [6]. A recent multicenter study [5] evaluated the use of a set of seven Y-linked microsatellites *DYS19*, *DYS389*A, *DYS389*B, *DYS390*, *DYS391*, *DYS392* and *DYS393* in several European populations and found discrimination probabilities ranging from 0.74 to 0.90.

Fig.1

Finally we should remember that obviously Y-linked microsatellites could only be used in paternity testing when the child is male, i.e. in 50% of the cases. There is basically nothing to be gained by typing Y chromosome polymorphisms in routine paternity trios. In contrast, in deficiency cases in which the alleged father is deceased, typing of Y-linked microsatellites in his brothers or other male relatives belonging to the same patrilineage can be instrumental in avoiding the need for exhumations [4,5,48].

The MSY1 minisatellite

Minisatellites are tandem arrays with repeat unit size larger than 6 bp [49]. The first human haploid minisatellite MSY1, described recently in the Y-linked locus *DYF155S1*, is composed of 25 bp tandem repeat units that are variable in sequence and block size [50]. It is the most variable locus in the human Y chromosome, reaching a gene diversity of about 99.9 % due to its high mutation rate ranging from 0.02 to 0.11 per generation [50]. The mutation events generating MSY1 diversity seem to be very complex and include a biased repair mechanism [51] reflecting the different modular structures and repeat block size. The applications of MSY1 in paternity and forensic casework as well as human evolution are similar to other tandem repeats such as microsatellites (see above). However the inherent complexities of the MVR-PCR technique [50] will certainly limit the utilization of this marker into routine laboratory work.

The alphoid heteroduplex (αh) polymorphic system

The αh polymorphism [26] is an unique system that allows the detection of variants in the Y chromosome alphoid DNA. It consists basically of a multilocal PCR amplification of specific divergent alphoid units present in the Y centromere. Two types of loci are co-amplified, one left side locus (αhL locus) and several loci in the right side (αh# loci) of the alphoid array. The locus αhL in the left side of the alphoid block is usually invariable among individuals [26,52] and is 4 bp smaller than the right side αh# loci (285 bp). Consequently, when the two types of loci are co-amplified there is presence of a short loop in heteroduplexes formed between them (Fig. 2). This loop is responsible for a large shift in mobility of the heteroduplexes relative to the homoduplexes, best seen in non-denaturing polyacrylamide gels (Fig. 2). Distinct point mutations are present in each αh# locus and cause the formation of further bubbles (mismatches) in heteroduplex, further affecting the heteroduplex mobility already retarded by the loop. Thus, each combination αhL + αh# (# = 1 to 39) generates a pair of heteroduplexes, named h1 to h39, that have different mobilities (Fig. 2). The mobility shift of each heteroduplex pair reflects the sequence of each corresponding αh# locus. In any single male, there can be one to six heteroduplex pairs corresponding to the same number of αh# loci variants, giving a heteroduplex pattern called αh type.

Fig.2 Fifty three αh types (αhI to αhLIII) have been found [26,52,53] which are resulted from distinct combinations of 39 different αh# (αh1 to αh39). Another source of variation is the occurrence of deletions of αh loci that can generate some of the αh types. Point mutations usually are unique events, but deletions can produce some recurrent αh types. The simplest type is αhI, a pattern with no heteroduplex formation that can be subtyped by artificial mixing experiments and is more often generated by deletion events involving the αhL locus [26,46,52] regardless of the αh# constitution. The combination of the αh system with binary

markers has revealed that simpler types such as αhII and αhIII, with few heteroduplex pairs, can be recurrently derived by deletion events from more complex types such as αhV and αhIX [52]. The applications of the αh polymorphic system are basically the same of the binary markers, especially for human evolutionary studies, as demonstrated previously [26, 45-47,52].

2 Technical procedures

Protocols for analysis of binary markers

An updated collection of binary markers amplified by PCR is listed below (Table 1). We are going to describe simple analysis methods for many of the markers, mostly RFLPs that can be readily incorporated to laboratory screening routine.

Table 1 Some other markers will not be discussed in detail because they were recently published and the sequence information for designing alternative protocols is not available. They include polymorphisms analyzed by more complex methods such as DHPLC [28] and SSO typing [31].

Protocol 1

General protocols – unless differently stated in the marker-specific sections (Table 2)

Protocol 1.1 PCR (mix)

Using 200 µl PCR tubes it is prepared 12.5 µl reactions in standard buffers 1 x provided by manufacturers of the enzyme *Taq* DNA polymerase (Perkin-Elmer, Promega), 1.5 mM $MgCl_2$, 200 µM of each dNTP, 1 µM of each primer (Table 1), 0.5 unit of enzyme and 10 to 50 ng of genomic DNA. No oil or paraffin is added to the tube.

Protocol 1.2 PCR (cycling)

The best results were obtained in PCR machines with hot lid to allow oil free amplification

such as PTC100 or PTC200 (MJR Research), GeneAmp 9600 (Perkin Elmer) and Genius (Techne). In addition to the cycling program there is an initial denaturation step at 94º C for 2 minutes and final extension at 72º C for 5 minutes with subsequent cooling to 4º C. Usually 30 cycles yielded sufficient amplified products however an increase to 35 cycles is recommended for bad quality DNAs.

Protocol 1.3 Digestion

Performed in a 10 µl volume containing 5 µl of PCR products and final 1 x digestion buffer provided by the restriction enzyme manufacturer. Boehringer Manhein, NEB and Promega provided restriction enzymes.

Protocol 1.4 Electrophoresis and visualization of bands

Table 2

Use a horizontal apparatus with 10 cm length of agarose gel in standard TAE or TBE running buffers [56]. Usually 5 µl of PCR products or 10 µl digested PCR products are loaded onto the gel mixed with standard loading buffer [56]. Some gels use also NuSieve (FMC) agarose to increase resolution. Alternatively, all products can be separated in non denaturing PAGE in TBE 1 x followed by silver staining [35], particularly when short PCR products have to be visualized. Ethidium bromide is usually incorporated in the agarose gel and running buffer [56]. After running, the gel is photographed (Polaroid) under UV irradiation.

Protocols for analysis of microsatellites

The first analyses of Y microsatellites were done in standard denaturing sequencing gels of fragments radioactively labeled and visualized in a photographic film [23] or in non denaturing PAGE followed by silver staining [35]. These methods were time consuming procedures that have been replaced by less laborious automated protocols using laser DNA sequencers [5,43]. Thus, recent improvements in Y microsatellite typing have gradually enabled the multiplexing of as many loci as possible during PCR and/or electrophoresis, in

order to reduce their handling time. In this review we are going to describe some improved protocols for Y microsatellite typing for some loci listed bellow (Table 3) that can be applied to any other microsatellite. Additional protocols and special cares with Y microsatellite typing using many different technologies were published recently in a large multicenter study [5].

Table 3 **Protocol 2 PCR amplification of Y linked tetranucleotide repeat loci**

Protocol 2.1 PCR tetraplex *DYS19* + *DYS390* + *DYS391* + *DYS393* (mix and cycling)

Using 500 µl tubes, in a 12.5 µl reaction volume add 1 x buffer (provided by the *Taq* manufacturer), 200 µM dNTPs, 1.5 mM $MgCl_2$, 1.6 µM of *DYS19* and *DYS391* primers, 1 µM of *DYS390* primers, 0.5 µM of *DYS393* primers, 0.625 unit of *Taq* polymerase and 20 ng DNA. This mixture was submitted to 30 cycles at 51º C 30 s, 72º C 90 s, 94º C 30 s.

Protocol 2.2 PCR *DYS389* (mix and cycling)

Use the same conditions as for the tetraplex with 1 µM of *DYS389* primers. *DYS389* is typed as a duplex because it is actually composed of at least two independent repeat blocks, *DYS389*A and *DYS389*B; the former apparently having arisen from a partial duplication of the latter [39] and both are amplified by the same set of primers (Table 3). PCR amplification then yields two products of different size: the larger amplicon (~350 bp) encompasses both microsatellites (*DYS389*A+B) while the smaller one (~250 bp) corresponds to *DYS389*B. The allele size of *DYS389*A can be calculated by subtracting the size of the small amplicon from that of the larger one.

Protocol 2.3 Electrophoresis (gel and running conditions)

When the primer is labeled with a fluorophor, the PCR products may be run in an automated electrophoresis system with laser detection using either fluorescein (ALF® - Pharmacia, Uppsala, Sweden) or a fluorescent carbocyanine dye Cy5* (ALF Express®, Pharmacia, Uppsala, Sweden) at the 5'end of either forward or reverse primer. Both reactions (tetraplex

and duplex) can be carried out as described in protocols 2.1 and 2.2 and resolved by running at 1500 V, 45 mA, 34 W, 45º C, for at least 2 hours (ALF express) or 6 hours (ALF). In the absence of an automated system, it is feasible to employ 8% PAGE (in TBE 1 x) silver stained [35]. In such cases, a non-labeled primer is employed in single locus PCR amplification and ran at constant current of 30 mA for at least 4 hours.

Protocol 2.4 Typing of alleles

15-repeat allele = 190 bp (*DYS19*), 108 bp (*DYS389A*), 255 bp (*DYS389B*), 287 bp (*DYS391*); 25-repeat allele at *DYS390* = 212 bp. We confirmed the sequence at DYS19 and DYS391 alleles. The sequence of the other loci was obtained in GenBank. To ensure a correct allele designation at all loci, an allelic ladder was constructed by amplifying a pool of 200 individuals (Fig. 1). In addition, an internal lane standard may be run so that each sample may be compared on the same scale and an external one allows us to calculate the size, in base pairs, of the peak.

Protocols for analysis of complex markers

The systems listed in Table 4 display complex polymorphisms amplified by PCR in the human Y chromosome. The recently described minisatellite MSY1 is detected by a MVR-PCR approach [50] that detects both size and unit sequence differences. Due to its complexity, MVR-PCR is not a protocol executable in most laboratories. A further complex polymorphism, the alphoid heteroduplex system, can be readily set up in a laboratory interested especially in human evolutionary studies and will be discussed below.

Table 4 **Protocol 3 Heteroduplex analysis of the** a**h system**

Protocol 3.1 PCR (mix)

In a 200 µl tube, 12.5 µl volume reactions are prepared with 1 x buffer (provided by the *Taq*

manufacturer), 200 µM dNTPs, 1.5 mM $MgCl_2$, 1 µM of primers (Table 4), 1 unit of *Taq* per tube and 50 ng of genomic DNA;

Protocol 3.2 PCR (cycling)

An initial denaturation at 94º C for 3 min, followed by 40 cycles at 65º C 30 s; 72º C 60 s, 94º C 30 s. After cycling it is used a 10 min extension at 72º C followed by a slow cooling of 1º C per 10 s until 4º C, to allow a better heteroduplex formation. At the end add 2.5 µl of MDE loading buffer (FMC) to the 12.5 µl volume reaction and load 6 µl onto the gel.

Protocol 3.3 Electrophoresis

Use a 20 x 20 cm glass plate vertical electrophoresis system, with 1 mm width spacers and combs (24 or 30 wells). Prepare 1 x MDE polyacrylamide gel (FMC) according to the manufacturer. Use 1 x TBE buffer and run at 160 V for 15 hours. After running visualize the bands by silver staining [52].

Protocol 3.4 Typing of alleles

For heteroduplex typing the individual patterns should be compared to a reference marker such as an individual with a pattern already characterized or a ladder of heteroduplexes that can be amplified from clones ([26,53], F.R. Santos upon request). The scoring can be also done comparing with the schemes (Fig. 2) showing the individualized heteroduplex pairs and αh types previously found [26,52,53]. Some new heteroduplex pairs that can appear in further analysis may need cloning and artificial mixing experiments [26,46,52] to be ascertained correctly. Some problems for scoring the heteroduplexes can appear due to the co-migration of bands, either the superior or inferior, belonging to different heteroduplex pairs. As each pair of heteroduplexes result from a unique combination of two loci (αhL + αh#), for the stoichiometry of heteroduplex formation they are expected to appear in the same intensity, and the heteroduplex pairs do vary in intensities as their source loci can be in single copy or at least duplicated [26]. If bands are co-migrating they should also appear thicker on the gel as

the sum of the intensities of each band (Fig. 2). However individuals with the same αh type and some difference in intensity of bands have only a quantitative difference that is not used for discrimination of αh types.

Acknowledgments

We thank Chris Tyler-Smith and Arpita Pandya for access to unpublished data. Our research was supported by CNPq (Conselho Nacional de Pesquisas) and FAPEMIG from Brazil and Leverhulme Trust from UK.

References

1 Jobling MA, Tyler-Smith C (1995) Fathers and sons - the Y chromosome and human evolution. *Trends in Genetics* 11: 449-456

2 Santos FR, Tyler-Smith C (1996) Reading the human Y chromosome: the emerging DNA markers and human genetic history. *Braz J Genet* 19: 665-670

3 Roewer L, Epplen JT (1992) Rapid and sensitive typing of forensic stains by PCR amplification of polymorphic simple repeat sequences in case work. *Forensic Sci Int* 53: 163-171

4 Santos FR, Epplen JT, Pena SDJ (1993) Testing deficiency paternity cases with a Y-linked tetranucleotide repeat polymorphism. in *DNA fingerprinting: State of the Science*. ed. by Pena SDJ, Chakraborty R, Epplen JT and Jeffreys AJ. pages 261-265

5 Kayser M, Caglia A, Corach D, Fretwell N, Gehrig C, Graziosi G, Heidorn F, Herrmann S, Herzog B, Hidding M, Honda K, Jobling M, Krawczak M, Leim K, Meuser S, Meyer E, Oesterreich W, Pandya A, Parson W, Penacino G, Perez-Lezaun A, Piccinini A, Prinz M, Schmitt C, Roewer L (1997) Evaluation of Y-chromosomal STRs: a multicenter study. *Int J Legal Med* 110: 125-133

6 Jobling MA, Pandya A, Tyler-Smith C (1997) The Y chromosome in forensic analysis and paternity testing. *Int J Legal Med* 110: 118-124

7 Marshall-Graves JA (1995) The origin and function of the mammalian Y chromosome and Y-borne genes - an evolving understanding. *BioEssays* 17: 311-321

8 Vogt PH, Affara N, Davey P, Hammer M, Jobling MA, Lay YFC, Mitchell M, Schempp W, Tyler-Smith C, Williams G, Yen P, Rappold GA (1997) Report of the third international workshop on Y chromosome mapping 1997. *Cytogenet Cell Genet* 79: 1-20

9 Lahn BT, Page DC (1997) Functional coherence of the human Y chromosome. *Science* 278: 675-680

10 Whitfield LS, Sulston JE, Goodfellow PN (1995) Sequence variation of the human Y chromosome *Nature* 378: 379-380

11 Hammer MF (1995) A recent common ancestry for human Y chromosomes. *Nature* 378: 376-378

12 Denver Conference (1965) A proposed standard system of nomenclature of the human mitotic chromosomes. *Lancet* 1: 1063-10065

13 Erdtmann B, Salzano FM, Mattevi M (1981) Size variability of the Y chromosome distal C-band in brazilian indians and caucasoids. *Ann Hum Biol* 8: 415-424

14 Jakubiczka S, Arnemann J, Cooke HJ, Krawczak M, Schmidtke J (1989) A search for restriction fragment length polymorphism on the human Y chromosome. *Hum Genet* 84: 86-88

15 Malaspina P, Persichetti F, Novelletto A, Iodice C, Terrenato L, Wolfe J, Ferraro M, Prantera G (1989) The human Y chromosome shows a low level of DNA polymorphism. *Ann Hum Genet* 54: 297-305

16 Dorit RL, Akashi H, Gilbert W (1995) Absence of polymorphism at the ZFY locus on the human Y chromosome. *Science* 268: 1183-1185

17 Casanova M, Leroy P, Boucekkine C, Weissenbach J, Bishop C, Fellous M, Purrello M, Fiori G, Siniscalco M (1985) A human Y-linked DNA polymorphism and its potential for estimating genetic and evolutionary distance. *Science* 230: 1403-1406

18 Lucotte G, Ngo NY (1985) p49f, A highly polymorphic probe, that detects Taq1 RFLPs on the human Y chromosome. *Nucleic Acids Res* 13: 8285

19 Oakey R, Tyler-Smith C (1990) Y chromosome DNA haplotyping suggests that most European and Asian men are descended from one of two males. *Genomics* 7: 325-330

20 Spurdle AB, Jenkins T (1992) The Y chromosome as a tool for studying human evolution. *Curr Opin Genet Dev* 2: 487-491

21 Jobling MA (1994) A survey of long-range DNA polymorphisms on the human Y chromosome. *Hum Mol Genet* 3: 107-114

22 Mathias N, Bayés M, Tyler-Smith C (1994) Highly informative compound haplotypes for the human Y chromosome. *Hum Mol Genet* 3: 115-124

23 Roewer L, Arnemann J, Spurr NK, Grzeschik K-H, Epplen JT (1992) Simple repeat sequences on the human Y chromosome are equally polymorphic as their autosomal counterparts. *Hum Genet* 89: 389-394

24 Seielstad MT, Herbert JM, Lin AA, Underhill PA, Ibrahim M, Vollrath D, Cavalli-Sforza LL (1994) Construction of human Y-chromosomal haplotypes using a new polymorphic A to G transition. *Hum Mol Genet* 3: 2159-2161

25 Hammer MF, Horai S (1995) Y chromosomal DNA variation and the peopling of Japan. *Am J Hum Genet* 56: 951-962

26 Santos FR, Pena SDJ, Tyler-Smith C (1995) PCR haplotypes for the human Y chromosome based on alphoid satellite variants and heteroduplex analysis. *Gene* 165: 191-198

27 Underhill PA, Jin L, Zemans R, Oefner PJ, Cavalli-Sforza LL (1996) A pre-Colombian Y chromosome-specific transition and its implications for human evolutionary history. *Proc Natl Acad Sci USA* 93: 196-200

28 Underhill PA, Jin L, Lin AA, Mehdi SQ, Jenkins T, Vollrath D, Davis RW, Cavalli-Sforza LL, Oefner PJ (1997) Detection of numerous Y chromosome biallelic polymorphisms by denaturing high-performance liquid chromatography. *Genome Res* 7: 996-1005

29 Zerjal T, Dashnyam B, Pandya A, Kayser M, Roewer L, Santos FR, Schiefenhovel W, Fretwell N, Jobling MA, Harihara S, Shimizu K, Semjidmaa D, Sajantila A, Salo P, Crawford MH, Ginter EK, Evgrafov OV, Tyler-Smith C (1997) Genetic relationships of Asians and northern Europeans revealed by Y-chromosomal DNA analysis. *Am J Hum Genet* 60: 1174-1183

30 Cavalli-Sforza LL, Menozzi P, Piazza A (1994) *The history and geography of human genes*. Princeton University Press, Princeton, New Jersey, USA

31 Hammer MF, Spurdle AB, Karafet T, Bonner MR, Wood ET, Novelletto A, Malaspina P, Mitchell RJ, Horai S, Jenkins T, Zegura SL (1997) The geographic distribution of human Y chromosome variation. *Genetics* 145: 787-805

32 Batzer MA, Arcot SS, Phinney JW, Alegria-Hartman M, Kass DH, Milligan SM, Kimpton C, Gill P, Hochmeister M, Ioannou PA, Herrera RJ, Boudreau DA, Scheer WD,

Keats BJ, Deininger PL, Stoneking M (1996) Genetic variation of recent Alu insertions in human populations. *J Mol Evol*. 42: 22-29

33 Altheide TK, Hammer MF (1997) Evidence of a possible Asian origin of YAP+ Y chromosomes. *Am J Hum Genet* 61: 462-466

34 Shriver MD, Smith MW, Jin L, Marcini A, Akey JM, Deka R, Ferrell RE (1997) Ethnic-affiliation estimation by use of population-specific DNA markers. *Am J Hum Genet* 60: 957-964

35 Santos FR, Pena SDJ, Epplen JT (1993) Genetic and Population Study of an Y-linked Tetranucleotide Repeat DNA Polymorphism with a Simple Non-Isotopic Technique. *Hum Genet* 90: 655-656

36 Poloni ES, Semino O, Passarino G, Santachiara-Benerecetti AS, Dupanloup I, Langaney A, Excoffier L (1997) Human genetic affinities for Y-chromosome P49a,f/*Taq*I haplotypes show strong correspondence with linguistics. *Am J Hum Genet* 61: 1015-1035

37 Santos FR, Gerelsaikan T, Munkhtuja B, Oyunsuren T, Epplen JT, Pena SDJ (1996) Geographic differences in the allele frequencies of the human Y-linked tetranucleotide polymorphism *DYS19*. *Hum Genet* 97: 309-313

38 Jobling MA, Samara V, Pandya A, Fretwell N, Bernasconi B, Mitchell RJ, Gerelsaikhan T, Dashnyam B, Sajantila A, Salo PJ, Nakahori Y, Disteche CM, Thangaraj K, Singh L, Crawford MH, Tyler-Smith C (1996) Recurrent duplication and deletion polymorphisms on the long arm of the Y chromosome in normal males. *Hum Mol Genet* 5: 1767-1775

39 Cooper G, Amos W, Hoffman D, Rubinsztein DC (1996) Network analysis of human Y microsatellite haplotypes. *Hum Mol Genet* 5: 1759-1766

40 Heyer E, Puymirat J, Dieltjes P, Bakker E, de Knijff P (1997) Estimating Y chromosome specific microsatellite mutation frequencies using deep rooting pedigrees. *Hum Mol Genet* 6: 799-803

41 Goldstein DB, Zhivotovsky LA, Nayar K, Ruiz-Linares A, Cavalli-Sforza LL, Feldman M (1996) Statistical properties of the variation at linked microsatellite loci: implications for the history of human Y chromosomes. *Mol Biol Evol* 13: 1213-1218

42 Ciminelli BM, Pompei F, Malaspina P, Hammer M, Persichetti F, Pignatti PF, Palena A, Anagnou N, Guanti G, Jodice C (1995) Recurrent simple tandem repeat mutations during human Y-chromosome radiation in Caucasian subpopulations. *J Mol Evol* 41: 966-973

43 Roewer L, Kayser M, Dieltjes P, Nagy M, Bakker E, Krawczak M, de Knijff P (1996) Analysis of molecular variance (AMOVA) of Y-chromosome-specific microsatellites in two closely related human populations. *Hum Mol Genet* 5: 1029-1033

44 Deka R, Jin L, Shriver MD, Yu LM, Saha N, Barrantes R, Chakraborty R, Ferrel R (1996) Dispersion of human Y chromosome haplotypes based on five microsatellites in global populations. *Genome Research* 6: 1177-1184

45 Pena SDJ, Santos FR, Bianchi NO, Bravi CM, Carnese FR, Rothhammer F, Gerelsaikhan T, Munkhtuja B, Oyunsuren T (1995) A major founder Y-chromosome haplotype in Amerindians. *Nature Genetics* 11: 15-16

46 Santos FR, Hutz MH, Coimbra CEA, Santos RV, Salzano FM, Pena SDJ (1995) Further evidence for the existence of major founder Y chromosome haplotype in Amerindians. *Braz J Genet* 18: 669-672

47 Santos FR, Rodriguez-Delfin L, Pena SDJ, Moore J, Weiss KM (1996) North and South Amerindians may have the same major founder Y chromosome haplotype. *Am J Hum Genet* 58:1369-1370

48 Pena SDJ, Chakraborty R (1994) Paternity testing in the DNA era. *Trends Genet* 10: 190-194

49 Jeffreys AJ, Wilson V, Thein SL (1985) Hypervariable minisatellite regions on the human Y chromosome. *Nature* 314: 67-73

50 Jobling MA, Bouzekri N, Taylor PG (1998) Hypervariable digital DNA codes for human paternal lineages: MVR-PCR at the Y-specific minisatellite, MSY1 (DYF155S1). *Hum Mol Genet* 7: 643-653

51 Bouzekri N, Taylor PG, Hammer MF, Jobling MA (1998) Novel mutation processes in the evolution of a haploid minisatellite, MSY1: array homogeneization without homogeneization. *Hum Mol Genet* 7: 655-659

52 Santos FR, Bianchi NO, Pena SDJ (1996) Worldwide distribution of human Y chromosome haplotypes. *Genome Research* 6: 601-611

53 Alphoid heteroduplex site - http://www.icb.ufmg.br/~fsantos/ah.html

54 Kwok C, Tyler-Smith C, Mendonca BB, Hughes I, Berkovitz GD, Goodfellow PN, Hawkins JR (1996) Mutation analysis of the 2 kb 5' to *SRY* in XY females and XY intersex subjects. *J Med Genet.* 33: 465-468

55 Bianchi NO, Bailliet G, Bravi CM, Carnese RF, Rothhammer F, Martinez-Marignac VL, Pena SDJ (1997) Origin of Amerindian Y-chromosomes as inferred by the analysis of six polymorphic markers. *Am J Phys Anthropol* 102: 79-89

56 Sambrook J, Fritsch EF, Maniatis T (1989) *Molecular cloning: a laboratory manual*. Cold Spring Harbor Laboratory Press, New York, USA

57 Genome Data Base (GDB) – http://www.gdb.org/

58 Chen H, Lowther W, Avramopoulos D, Antonarakis SE (1994) Homologous loci DXYS156X and DXYS156Y contain a polymorphic pentanucleotide repeat (TAAAA)n and map to human X and Y chromosomes. *Hum Mutat* 4: 208-211

Figure Legends

Figure 1 - Fluorogram with allelic ladders for the Y microsatellite tetraplex (*DYS19, DYS390, DYS391* and *DYS393*). Pooled DNA samples from 200 random Brazilians (males and females) were amplified with protocol 2.1 (see ”Technical procedures”) and run on ALF express® (Pharmacia, Uppsala, Sweden). The alleles at each locus are seen as individual peaks and the areas under each peak are roughly proportional to the frequency of the corresponding allele in that DNA pool. Note that the largest allele of *DYS19* and the smallest one of *DYS390* are very close. In general this does not cause any problems in typing. However, the size of *DYS390* alleles can be increased by 10 base pairs simply by adding a tail of 10 Ts to the lower primer. An additional locus ladder is seen between *DYS390* and *DYS391*. This ladder corresponds to an X-linked tetranucleotide microsatellite that is homologous to *DYS391* (D.R. Carvalho-Silva and S.D.J. Pena, manuscript in preparation). In this run the size of this ladder was exaggerated compared to what we generally see in males because 50% of the pooled DNA amplified originated from females. The fluorogram was visualized with the Fragment Manager software (Pharmacia, Uppsala, Sweden).

Figure 2 – Y chromosome alphoid heteroduplex (αh) polymorphic system. (A) MDE (FMC) polyacrylamide gel stained by silver [52] showing PCR heteroduplex patterns (protocol 3) of artificial clone mixtures and an African individual. In lanes **h1**, **h5**, **h6**, **h7** and **h9** mixtures of PCR products of each of five different αh clones of the right side with a cloned αhL locus are shown [26]. In lane **M** the PCR products mixture of all previous αh# loci (1,5,6,7,9) with the αhL locus is depicted. The pattern in lane **M** reproduces completely the profile observed in lane **I**, the heteroduplex pattern of an African male with the αhX type. A drawing with the αhX profile is represented in column **S** where an arrow points to the heteroduplex band that looks thicker in lanes **M** and **I**, when compared to other bands, because of co-migration of the upper bands of heteroduplex pairs **h7** and **h9**. The homoduplexes are seen at the bottom of the gel. (B) Diagram explaining how the heteroduplexes observed in the lane **I** of Fig. 2A were generated by PCR amplification of the left side locus (αhL) and the five right side loci (αh1, αh5, αh6, αh7 and αh9).

Fig. 1

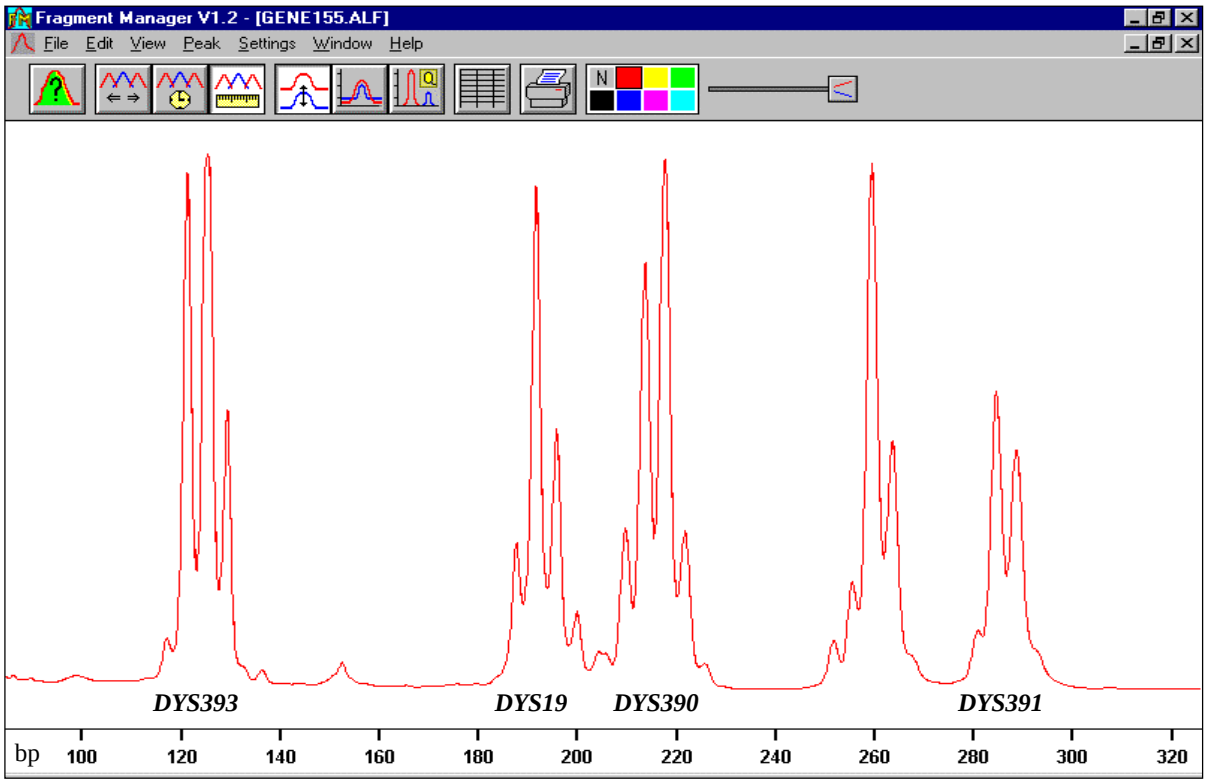
Fragment Manager V1.2 - [GENE155.ALF]
File Edit View Peak Settings Window Help
N
DYS393
DYS19
DYS390
DYS391
bp 100 120 140 160 180 200 220 240 260 280 300 320

Fig. 2A

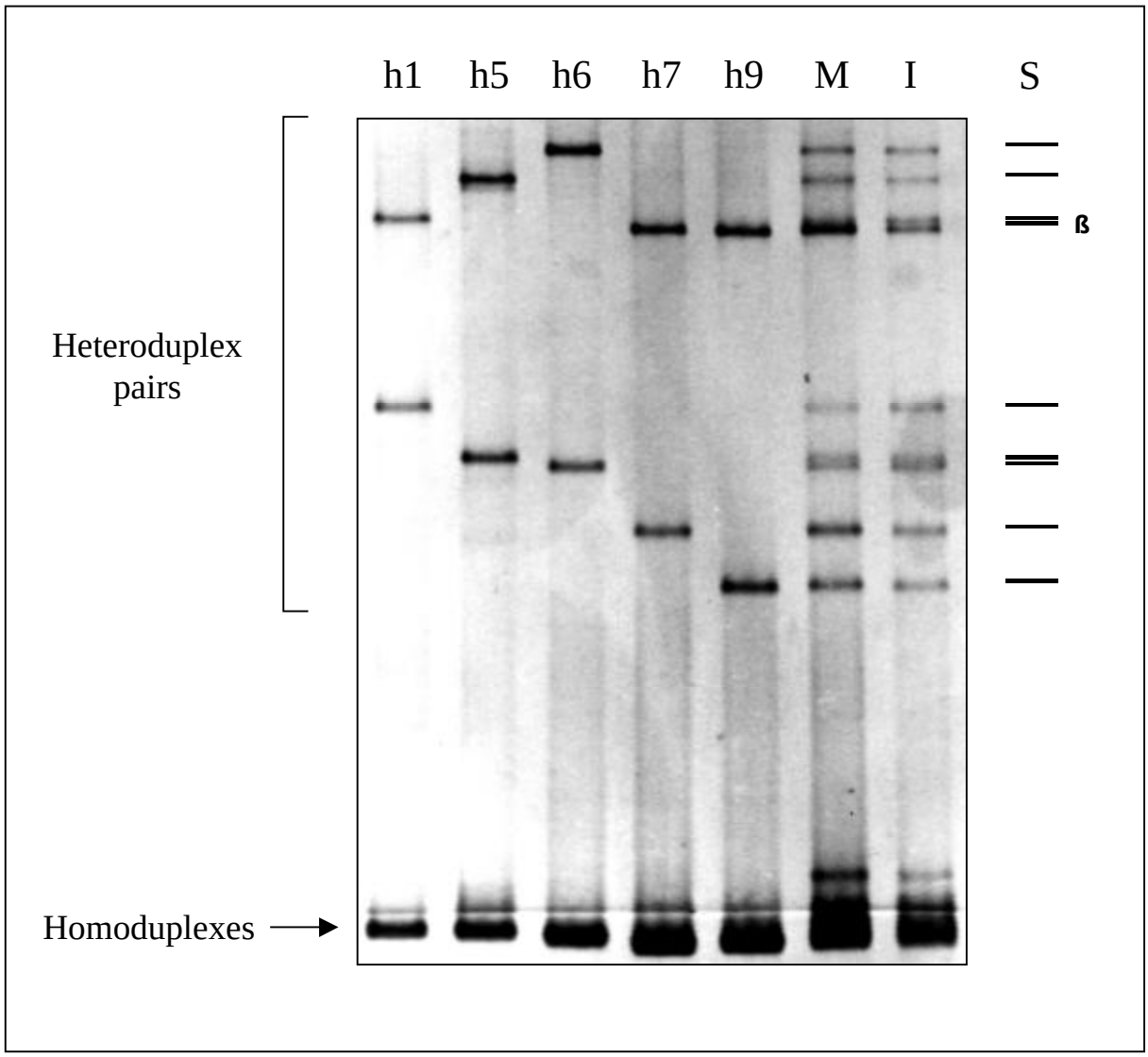

Fig. 2B

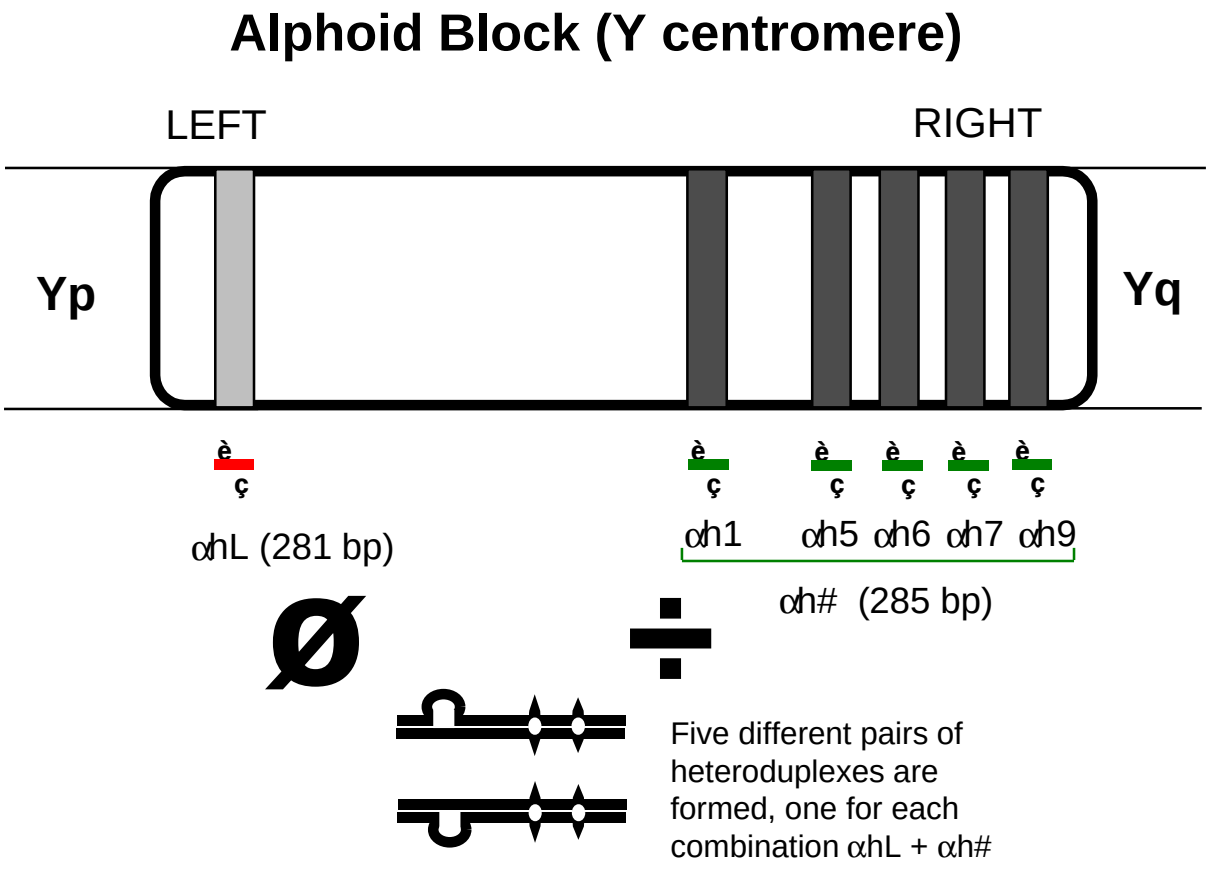
Alphoid Block (Y centromere)
LEFT
RIGHT
Yp
Yq
αhL (281 bp)
αh1
αh5
αh6
αh7
αh9
αh# (285 bp)
Five different pairs of heteroduplexes are formed, one for each combination αhL + αh#

Table 1 – Binary markers amplified by PCR

Locus [marker]	Upper and lower primers (5'→3')	Size (bp)	Variation 0 ® 1	Site* (bp)	Detection methods	Referen
DYS287 [YAP]	actgctaaaaggggatggat caggggaagataaagaaata	155 or 455	→Alu		Gel	[25]
DYS271 [sY81]	aggcactggtcagaatgaag aatggaaaatacagctcccc	209	A→G	168	RFLP DHPLC	[24] [28]
DYZ3 [YαI]	tctgagacacttctttgtggta cgctcaaaatatccactttcac	285	A↔G **	48	RFLP	[26]
SRY [-8299]	acagcacattagctggtatgac tctctttatggcaagacttacg	509	G→A	147	RFLP	[10]
SRY [-1532]	tccttagcaaccattaatctgg aaatagcaaaaaatgacacaaggc	167	A→G ***	115	RFLP	[10,54]
DYS287 [PN1]	tcacataatttcattttccc tagtcctctccttattaacg	?	C→T	?	SSO	[31]
DYS287 [PN2]	gatgcaaatgagaaagaact ctaaaaactggagggagaaa	?	C→T	?	SSO	[31]
DYS287 [PN3]	gatgcaaatgagaaagaact ctaaaaactggagggagaaa	?	G→A	?	SSO	[31]
SRY [-2627]	cgcggctttgaatttcaagctctg ccagggccccgagggactctt	391	C→T	130	RFLP	[55]
[Tat]	gactctgagtgtagacttgtga gaaggtgccgtaaaagtgtgaa	112	T→C	28	RFLP	[29]
DYS199 [M3] ****	taatcagtctcctcccagca aaaattgtgaatctgaaatt**c**aagg	202	C→T	181	RFLP DHPLC	[27] [28]
DYS234 [M4]	tcctaggttatgattacagagcg tgcagaacatttgtactgttcc	273	A→G	88	DHPLC	[28]
DYS214 [M5]	gggtttatactgacctgccaatgtt ttattgggaactttcagggg	325	G→A	73	DHPLC	[28]
DYS198 [M6]	cactaccacatttctggttgg cgctgagtccattctttgag	218	T→C	37	DHPLC	[28]
DYS253 [M7]	actgtgagcgagctgaaaat gcagccttgtgaaccaatta	280	C→G	216	DHPLC	[28]
DYS263 [M8]	cccacccacttcagtatgaa aggctgacagacaagtccac	267	G→T	137	DHPLC	[28]
[M9]	gcagcatataaaactttcagg aaaacctaactttgctcaagc	340	C→G	68	DHPLC	[28]
[M10]	gcattgctataagttacctgc taataaaaattgggtcaccc	343	T→C	156	DHPLC	[28]
[M11]	tctctctgtctgtctctccctcc gagcataaacaagaacttactgagc	222	A→G	44	DHPLC	[28]
DYS260 [M12]	actaaaacaccattagaaacaaagg ctgagcaacatagtgacccc	309	G→T	286	DHPLC	[28]
[M13]	tcctaacctggtggtctttc agccatgattttatccaacc	231	G→C	157	DHPLC	[28]
[M14]	agacggttagatcagttctctg tagataaaagcacattgacacc	287	T→C	180	DHPLC	[28]
[Ml5]	acaaatcctgaacaatcgc aaatgtgttgagtctgggaag	162 or 171	→+9 bp	110	DHPLC	[28]
DYS214 [Ml6]	tgttatgtcatttgaacccag ccgtgtgttgctgggctgtc	266	C→A	38	DHPLC	[28]
[Ml7]	ctggtcataacactggaaatc tgacctacaaatgagaaactc	335 or 334	→- 1 bp	68	DHPLC	[28]
[Ml8]	ctggtcataacactggaaatc tgacctacaaatgagaaactc	335 or 337	→+ 2 bp	63	DHPLC	[28]
[M19]	ctggtcataacactggaaatc tgacctacaaatgagaaactc	335	T→A	131	DHPLC	[28]
[M20]	gattgggtgtcttcagtgct cacacaacaaggcaccatc	413	A→G	118	DHPLC	[28]
[M21]	cttttatttctgactacaggg aacagcagatttgagcagg	415	A→T	357	DHPLC	[28]
DYS273 [M22]	agaagggtctgaaagcaggt gcctactacctggaggcttc	327	A→G	129	DHPLC	[28]

* position of the mutation related to the 5' end of the upper primer

** some individuals display both allele states as the locus is duplicated

Table 2 – Binary marker specific protocols

Marker	PCR		Digestion	Electrophoresis		Typing of alleles	Geographic dispersion
	Mix	Cycling		Gel	Condition		
YAP (*DYS287*)	1 unit of *Taq* per tube; 0.5 µM primers	30 cycles at 51º C 60 s; 72º C 60 s, 94º C 60 s		1.5 % agarose in TAE 0.5 x	100 V, 40 minutes	allele 0 = 155 bp; allele 1 = 455 bp	allele 1 is found in Africa, Japan, Tibet, Middle East and Europe [25,31,33]
***SRY*-8299**		30 cycles at 60º C 30 s; 72º C 60 s, 94º C 30 s	1 unit of *Bsr*BI, 37º C 2 hours	1.5 % agarose in TAE 0.5 x	100 V, 40 minutes	allele 0 = bands at 147 + 362 bp; allele 1 = 509 bp	allele 1 is a subset of YAP allele 1 chromosomes and found in Africa, Middle East and Europe [33]
sY81 (*DYS271*)		30 cycles at 60º C 20 s; 72º C 30 s, 94º C 20 s	1 unit of *Hsp*92II , 37º C 2 hours	3.5 % Nusieve / agarose (2/1.5) in TAE 0.5 x	100 V, 40 minutes	allele 0 = bands at 102 + 65 + 42 bp; allele 1 = bands at 144 + 65 bp	allele 1 is a subset of *SRY*-8299 allele 1 chromosomes found almost exclusively in sub-Saharan Africa [24,31]
***SRY*-2627**		30 cycles at 63º C 30 s; 72º C 60 s, 94º C 30 s	1 unit of *Ban*I, 37º C 2 hours	3.5 % Nusieve / agarose (2/1.5) in TAE 0.5 x	100 V, 60 minutes	allele 0 = bands at 264 + 86+ 41 bp; allele 1 = bands at 350 + 41 bp	allele 1 is found in Iberian populations ([55], Pandya et al. in preparation)
Tat	1 unit of *Taq* per tube	30 cycles at 61º C 20 s; 72º C 30 s, 94º C 20 s	1 unit of *Hsp*92II, 37º C 2 hours	4 % Nusieve / agarose (3/1) in TAE 0.5 x	100 V, 40 minutes	allele 0 = bands at 85 + 27 bp; allele 1 = 112 bp	allele 1 is found in north Europe and Asia [29]
DYS199	1 unit of *Taq* per tube	30 cycles at 61º C 20 s; 72º C 30 s, 94º C 20 s	1 unit of *Mfe*I, 37º C 2 hours	4 % Nusieve / agarose (3/1) in TAE 0.5 x	100 V, 40 minutes	allele 0 = bands at 181 + 21 bp; allele 1 = 202 bp	allele 1 is found in Americas [27]
Y a I (*DYZ3*)	1 unit of *Taq* per tube	30 cycles at 65º C 30 s; 72º C 60 s, 94º C 30 s	1 unit of *Hin*dIII, 37º C 2 hours	1.5 % agarose in TAE 0.5 x	100 V, 40 minutes	* allele A = 285 bp; allele G = bands at 237 + 48 bp	both allele states are found everywhere [26]
***SRY*-1532**		30 cycles at 60º C 20 s; 72º C 30 s, 94º C 20 s	0.5 unit of *Dra*III, 37º C 2 hours	3.5 % Nusieve / agarose (2/1.5) in TAE 0.5 x	100 V 30 minutes	** allele 0 = 167 bp; allele 1 = bands at 112 + 55 bp	allele 1 is found in most individuals everywhere and allele 0 is found in African bushmen and populations of northern Europe and Asia ([1,11], Pandya et al. in preparation)

* in this case the ancestral and derived states are not known as the locus can be duplicated and a single individual may contain both allele states [26]

** the allele 0, corresponding to the nucleotide A, can be a recurrent allele state [2]

Table 3 – Microsatellites in the human Y chromosome

Locus [marker]	Upper and lower primers (5' ® 3')	Allele size variation (bp)	Repeat block **	References
DYS19 [Y-27H39]	ctactgagtttctgttatagt atggcatgtagtgaggaca	174 – 210	$(gata)_3$ (ggta) $(gata)_n$	[23]
DYS288	cattacaaatacctggacactg ttgctttgcttgtcatttcaga	119 – 123	$(ca)_n$	[57]
DYS385	agcatgggtgacagagcta tgggatgctaggtaaagctg	360 – 412	$(gaaa)_n$	[57]
DYS388	gtgagttagccgtttagcga cagatcgcaaccactgcg	126 – 138	$(ata)_n$	[57]
*DYS389A**	ccaactctcatctgtattatctat tcttatctccacccaccaga	105 – 125	$(gata)_n$ $(gaca)_6$	[39,57]
*DYS389B**	see DYS389A	239 – 263	$(gata)_n$ $(gaca)_3$ (gata)	[39,57]
DYS390	tatattttacacatttttgggcc tgacagtaaaatgaacacattgc	191 – 227	$(gata)_2$ $(gaca)_8$ $(gata)_n$	[1,57]
DYS391	ctattcattcaatcatacaccca gattctttgtggtgggtctg	275 – 295	$(gata)_n$ $(gaca)_3$ (gata)	[1,57]
DYS392	tcattaatctagcttttaaaaacaa agacccagttgatgcaatgt	236 – 263	$(att)_n$	[1,57]
DYS393	gtggtcttctacttgtgtcaatac aactcaagtccaaaaaatgagg	108 – 132	$(gata)_n$	[1,57]
[YCAI]	cccatgcctgttctccagatt gagagtgtgacacatcaggta	124 – 132	$(ca)_n$	[22]
[YACII]	tatattaaatagaagtagtga tatcgatgtaatgttatatta	147 – 165	$(ca)_n$	[22]
DYS413 [YACIII]	caacattgtgtgaatgtgtga tcctcagagaaggagaaacta	192 – 204	$(ca)_n$	[22]
DYS425	tggagagaagaagagagaaat agtaattctggaggtaaaatgg	198 – 207	trinucleotide	[38]
DYS426	ggtgacaagacgagactttgtg ctcaaagtatgaaagcatgacc	94 – 97	trinucleotide	[38]
DXYS156Y	gtagtggtctttgcctcc cagataccaaggtgagaatc	160 – 170	$(taaaa)_n$	[58]

* Amplification of locus *DYS389* has two independent microsatellites, A and B [39].

** The most variable repeat block is displayed as $(unit)_n$.

Table 4 – Complex polymorphisms amplified by PCR

System	Locus	Upper and lower primers (5'→3')	Source of variation	Detection method	References
αh system	*DYZ3*	tctgagacacttctttgtggta cgctcaaaatatccactttcac	Sequence polymorphisms and loci deletion	Heteroduplex analysis	[26]
MSY1	*DYF155S1*	ctcaagctaggacaaagggaaagg gaggtagatgctgaagcggtatag	Variable size of repeat block and point mutations in the repeat units	MVR-PCR	[50]