CHAPTER 30

The human Y chromosome male-specific polymorphisms and forensic genetics

Leonor Gusmão[1], María Brión[2] and Iva Gomes[1,2]

[1] *IPATIMUP, Institute of Molecular Pathology and Immunology of the University of Porto, 4200-465 Porto,Portugal*
[2] *Grupo de Medicina Xenómica, CeGen-Institute of Legal Medicine, University of Santiago de Compostela, 15782 Santiago de Compostela, Spain*

30.1 INTRODUCTION

30.1.1 Y-chromosome structure

The Y chromosome is one of the smallest human chromosomes, with an estimated average size of 60 million base pairs (Mb) (Fig. 30.1). During male meiosis recombination only takes place in the pseudoautosomal regions at the tips of both arms of Y and X chromosomes (PAR1, with 2.6 Mb, and PAR 2, with 0.32 Mb). Along ~95% of its length the Y chromosome is male-specific and effectively haploid, since it is exempt from meiotic recombination. Therefore, this Y-chromosome segment where X-Y crossing over is absent has been designated as the non-recombining region of the Y chromosome or NRY. Because of the high non-homologous recombination occurring within this Y chromosome specific region, a more appropriately name of male-specific region or MSY is nowadays used to designate it [1].

The MSY is a mosaic of heterochromatic and euchromatic regions. Besides the centromeric heterocromatin, a large heterochromatic region is located on the distal long arm of the Y chromosome (Yq) and constitutes more than half of the chromosome in some normal males, but is virtually undetectable in others [2]. A third heterochromatic region was recently discovered by Skaletsky *et al.* [1], interrupting the euchromatic sequences of proximal Yq (see Fig. 30.1). These regions are composed of highly repeated sequences of non-functional DNA: DYZ1, DYZ2, DYZ3, DYZ17, DYZ18, and DYZ19.

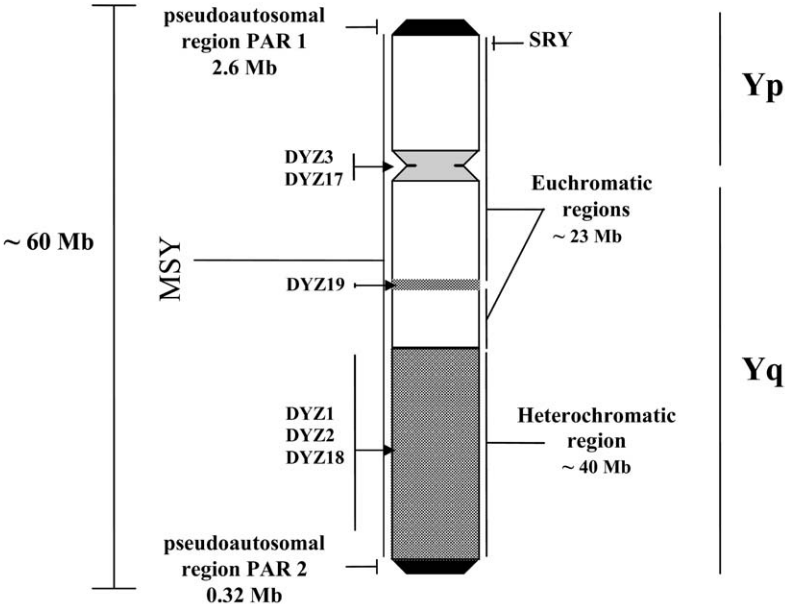

Fig. 30.1. Y-Chromosome structure.

The euchromatin is a constant size region and includes sequences homologous to the X chromosome, Y-specific repetitive sequences, and all the genes identified in the Y chromosome, which include the now identified 27 distinct protein-coding genes or gene families. Near-complete sequence of the Y-chromosome euchromatin has been recently revealed by Skaletsky *et al.* [1] that classifies the euchromatic sequences into three categories. First, the X-transposed, consisting in a stretch recently transposed from the X chromosome – ~3–4 million years ago, that still presents 99% homology to their X-chromosome counterparts. Second, the X-degenerate, consisting of a class of sequences more distantly related to the X chromosome – remnants of ancient autosomal sequences from which the modern X and Y derive. And at last, the ampliconic class composed largely of sequences that exhibit as much as 99.9% identity to other sequences in the MSY, maintained by frequent Y–Y gene conversion events. These sequences are located in seven segments scattered across the long and proximal short arms, and the most striking structural feature are eight massive palindromes located in the ampliconic regions of Yq, six of which carry testis genes.

30.1.2 The evolution of sex chromosomes

The similarities between the X and Y chromosome sequences are consistent with the hypothesis of a common origin. The mammalian advanced sex chromosome systems originated 300 million years ago from systems in which the X and Y were initially largely genetically homologous [3,4]. The evolution of sex chromosomes involved

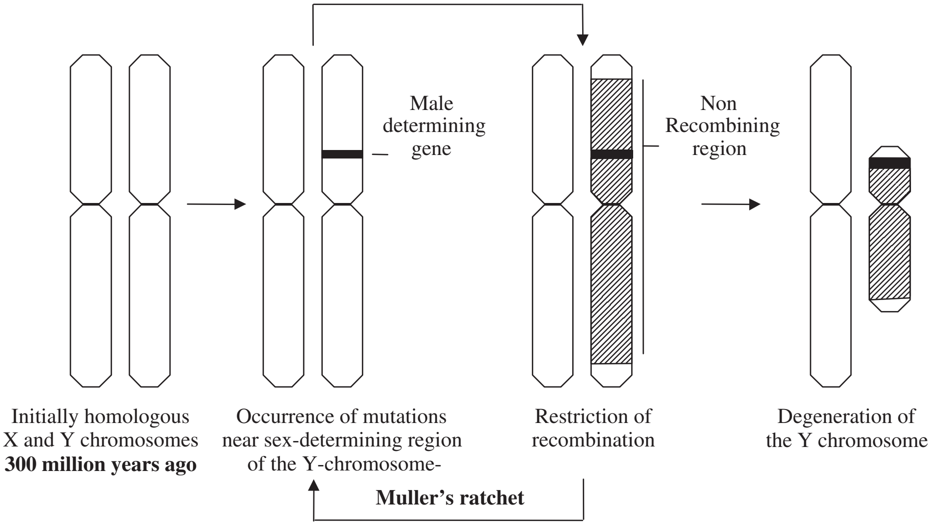

Fig. 30.2. Differentiation of the initially morphologically homogeneous X and Y chromosomes.

mechanisms of restriction of gene recombination, transposition, and translocation. The sequence of events that induced the morphological and genetic differentiation of the X and Y chromosomes and the genetic inactivation of the Y-chromosome genes is still not completely understood. The presently accepted explanation of the differentiation of the initially morphologically homogeneous X and Y chromosomes invokes successive processes where alternated steps of mutation and restriction of recombination were involved (Fig. 30.2).

In time, the Y chromosome comes to carry genes that are beneficial to the male but not to the female sex. If linked to the sex-determining region of the Y chromosome, those genes, favored in males and selected negatively in females, will tend to spread through the population. In order to keep this genetic heterogeneity between X and Y chromosomes, restriction of recombination involves sex determination genes and loci controlling secondary sexual characteristics being promoted by selection mechanisms. In a process referred to as "Muller ratchet," the lack of exchange through all or part of the originally homologous X and Y chromosomes will accumulate deleterious recessive mutations, since they are not restricted by selection. If there is no recombination, some mutations are more liable to be lost from the population and spread of a favorable Y-linked mutant allele through a population that would allow for the fixation of deleterious alleles at other loci. The accumulation of recessive deleterious alleles on the Y chromosome favors a selection for increased activity of the homologous loci on the X chromosome. On the other hand, with the reduction of Y chromosome genetic activity, there will be weak selection against insertions into the Y chromosome. In the absence of gene exchange and selective pressures, transposable elements and tandem repeat sequences are expected to accumulate, leading to a step-by-step reduction of the Y activity.

Sequencing of the MSY provided the opportunity to reexamine the model of evolution of the human sex chromosomes, showing that it is a consequence of two opposed evolutionary dynamics acting on the Y chromosome: gene decay versus gene acquisition and conservation [1,5].

Because of the presence of MSY gene pairs in the ampliconic sequences of the euchromatin which are subject to frequent gene conversion [1,5], and of the little or no X-degenerate gene loss or decay observed during the last 6 million years of human evolution [6], the predictions that the Y chromosome would be vanished completely in the next 10 million years seem to no longer have support.

30.1.3 The Y-chromosome inheritance

As a result of the 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, the MSY is transmitted from father to son unchanged unless a mutational event takes place. For this reason, the Y chromosome contains a record of all the mutational events that occurred among his ancestors, reflecting the history of paternal lineage. Therefore, all modern Y chromosomes have a single paternal ancestor, on their male-specific region.

30.1.4 Y-chromosome-specific polymorphisms

In 1985, Casanova *et al.* [7] undertook the first search for Y-linked restriction fragment length polymorphisms (RFLPs) in humans, with the report of two Y-specific polymorphisms. This and latter surveys on Y-specific markers by RFLP studies [8–10] and sequence analysis [11,12] emphasized the low level of polymorphism of this chromosome, compared with other chromosomes [13]. The attempt to identify new Y-specific polymorphisms in different population samples, mainly in Caucasians [8,9] and Africans [10], showed that the Y chromosome is apparently devoid of polymorphic genetic markers. Jakubiczka *et al.* [8] estimated a frequency of less than 1 point mutation in 18,000 nucleotides, and Malaspina *et al.* [9] calculated less than 1 per 46,515 nucleotides. Spurdle and Jenkins [10] screened a 20,808 bp segment for Y-specific RFLPs and did not detect any new polymorphism.

The low variation found in the Y chromosome was unexpected in view of its origin and is best explained, simply, by its presence at one quarter of the frequency of the autosomes, in diploid populations. Therefore, it is especially subject to drift that will be reflected in a corresponding reduced diversity [14]. The effective population size of the Y chromosome can also be reduced by a particular pattern of mating behavior found in specific populations. The lack of recombination also explains the low Y chromosome variation found, due to the effect of selective pressure in which a whole haplotype is involved instead of a specific allele [15].

Despite the scarcity of polymorphisms, with the availability of the complete sequence of the Y chromosome and with the improvement and applications of new techniques, a high number of Y-specific polymorphisms are now well characterized and available to population, evolutionary, and forensic genetics e.g. [16,17].

30.2 Y CHROMOSOME IN FORENSIC APPLICATIONS

Y-specific polymorphisms have been proven to be especially useful in routine forensic casework. The male specificity and especial inheritance features, distinct from autosomal, X-chromosomal, and mitochondrial markers transmission, determined the relevance of the Y-specific polymorphic markers, first in evolutionary and population genetics and then, based on the population studies, in forensic field applications e.g. [18–22].

30.2.1 Paternity testing

The pattern of inheritance along the male lineage makes the Y-STR polymorphisms suitable in paternity testing. However, the applicability of these markers is limited to approximately one half of paternity cases (those in which a male descendency is in question). It must be considered that a result based exclusively on Y-chromosome STRs does not exclude as father any male relative in the same patrilineage. Whenever possible, autosomal (AS) markers should be also used in order to avoid or reduce this possibility.

The probability of excluding in paternity for Y-chromosome markers is equal to the probability of having two different alleles in the population, which means that it is identical to the probability of discrimination in criminal identification. Thus, Y-linked polymorphisms have a much higher exclusion power than equally variable autosomal loci [19].

The possibility of using these markers is especially important in deficiency cases. In these situations, namely when the alleged father is deceased, it is possible to access his complete Y-chromosome information using male relatives in patrilineage.

30.2.2 Stain analysis

Y-chromosome STRs are useful in the discrimination of stains in forensic investigation when a male suspect is involved. This is the case of most violent crimes including sexual offences [18].

Mixtures of body fluids from different individuals are frequent in forensic casework. Y-chromosome analysis can be particularly helpful to detect male DNA fraction in stains involving male/female mixtures [23,24], the most common biological material available in sexual assaults.

Although a differential lysis can be applied in stains where semen from the offender is mixed with cells from the victim, this is laborious and is not always successful in achieving complete separation of the two cellular fractions. Differential extraction can fail in very small and degraded samples or when the fraction of semen is much reduced, leading to loss of sperm DNA [25,26]. Differential lysis is supposed to fail in 10–20% of forensic cases although Y-specific amelogenin can be detected in most of them. In other mixtures of body fluids from different individuals, such as blood–blood or blood–saliva, differential extraction cannot be applied in order to separate the DNA fraction from male and female sources. In such cases, when AS markers are used, preferential amplification of the major component of the mixture (usually female DNA from the victim) can mask the genetic profile of the assailant. By typing Y markers, even a minor male DNA component in a mixed male/female stain of a sexual assault yields a male-specific profile that can be compared with the DNA of suspects [27]. Prinz *et al.* [23] identified Y-specific STR alleles in male/female DNA mixture in a ratio of 1:2,000, compared with a 1:50 maximum ratio to detect a minor component for an autosomal locus. Another application is in rape cases committed by azoospermic individuals. They represent around 1–2% of all rapes but this percentage is increasing due to the fact that vasectomy is now more frequent. With Y-specific analysis only the male component is detected, and this allows a direct determination of the Y haplotype.

Y-specific STRs can also be useful in the determination of the number of semen contributors in multiple rape cases and as a screening method for linking rape series or for excluding suspects.

30.2.3 Y's counterpart: the X chromosome in forensic genetics

The X chromosome has a particular place in forensic genetics, regardless of its limitations in identity testing, when compared with the Y chromosome and the autosomes. The unique properties of genetic inheritance of this chromosome, consequence of the sex chromosomes evolution, are responsible for its role in forensics. In males, one single copy and the absence of recombination allows direct haplotyping; therefore the single male copy is entirely transmitted to female descendants. This characteristic itself can lead to paternity exclusion if the alleged father is unavailable and two sisters or half sisters are under investigation, since they will share the same paternal alleles [28]. This is also true in kinship analysis when accessing an alleged paternal grandmother/granddaughter relationship. X-linked genetic markers singly handed or supplementing the analysis of Y chromosomal or AS markers, are mainly useful in complex cases of kinship analysis, with female offspring, representing one of the major advantages of the use of X-chromosome markers [28–34]. In fact, in paternity deficiency cases where paternal relatives must be studied, X-STRs can occasionally resolve these cases more efficiently than common AS loci. When evaluating the forensic efficiency of X and AS loci in trios involving a daughter, X-chromosome markers are more efficient than AS markers, since the mean exclusion chance (MEC) is higher for these specific markers [28]. On the other hand, the X chromosome recombines in females providing an interesting system of genetic variation in each generation. The combination of these features, recombination in females and direct haplotyping in males, also makes the X chromosome an interesting system for historical studies [35].

At the present time, studies using X-STRs are of great interest in the fields of population genetics, as well as in forensics and kinship testing and, several X-STRs have been recently validated. Large PCR multiplexes for X-linked genetic markers make population studies and databasing more efficient and need to be designed and optimized. Several X-STR multiplexes have been described in the literature e.g. [36–38]; however, reference to amplifications in one single PCR reaction containing a high number of STR loci has not so far been common. Nevertheless, further studies are still needed on allele frequency distributions in different populations, mutation rates, and linkage disequilibrium, in order to establish reference population databases [36,38,39]. A forensic chromosomal X-STR database compiling population data previously published on allele frequencies and forensic parameters is also available [40]. In the case of X-STRs it is particularly important to study simultaneous transmission of different markers in order to screen possible associations between loci (this can be especially important in admixed and possibly substructured populations). It is expected that linkage disequilibrium intervals are longer on the X chromosome because recombination occurs only in females; therefore only two-thirds of the X chromosome recombines in each generation [35]. Although the lack of association between X-STRs contributes to increase the power of discrimination of a set of markers, haplotype analysis has been demonstrated to be a valuable tool in pedigree-based-kinship testing [29–31,37]. Reliable estimates of mutation rates are also important in kinship analysis; but few X-STRs segregation studies have been reported for these markers. Even so, after combining data previously reported [28,41–44], a total of 18 mutations were compiled in 8,698 allele transmissions, with an average mutation rate of 2.07×10^{-3} (95% CI $1.23–3.27 \times 10^{-3}$) per locus/meiosis. This value is inside the 95% CI of other mutation rates estimates published for Y-chromosomal and autosomal STRs e.g. [45–48].

The common application of AS, Y-chromosomal markers, and the mtDNA genome in forensics exceed the use of X chromosome genetic markers in this field. However, the X chromosome can be an important additional tool in human identification, unravelling some of the challenges presented in forensic genetics. A promising future for its application in this field can only be expected.

30.3 Y-CHROMOSOME MARKERS IN FORENSICS

The non-pseudoautosomal region of a Y chromosome contains different kinds of polymorphisms. These different loci have different mutation rates and consequently it is possible to select appropriate Y polymorphisms to study evolution events over different time scales [15]. STRs seem to be the most suitable Y-chromosome markers in forensic genetics due to their levels of diversity when compared to the biallelic polymorphisms. Taking into account the polymorphic levels, the minisatellite MSY1 is potentially the most informative Y-specific marker in forensic casework. However, in degraded DNA samples only a few repeats from each end of the allele can be typed [18]. Thus, its applicability in criminal casework becomes much reduced, since in most cases it is necessary to deal with poor-quality DNA. Moreover, there are some limitations in the use of this system in paternity investigations due to its high mutation rate.

Although the Y-STRs show moderate levels of polymorphisms when compared with the autosomal STRs, when studied together they can be used to construct highly discriminative haplotypes providing a useful tool in forensic investigation [49]. Therefore, population surveys based on haplotypes including as much Y polymorphisms as possible seem to be the best strategy [19,50] in population databasing for forensic identification or paternity testing purposes. The largest forensic worldwide population database available on Y-chromosome haplotypes is the YHRD (Y chromosome Haplotype Reference Database; www.yhrd.org).

30.3.1 Binary polymorphisms

During the last few years, sequencing efforts and the development of the new genotyping technologies have allowed the identification of a vast number of single nucleotide polymorphisms. However, the special features of the Y chromosome require to be studied in different ways, considering the loss of recombination and the uniparental inheritance.

In 2002, the Y-chromosome consortium published the single most parsimony phylogeny, considering the 153 Y chromosome binary haplogroups, observed with 245 binary markers. In order to solve problems comparing published data with the different nomenclatures available at that time, a simple set of rules were developed to unambiguously label the different clades described in the tree.

In 2003, an update of the haplogroup tree was published considering all the new binary markers described and located in the phylogeny [51].

The global architecture of the Y-chromosome phylogeny includes 18 major haplogroups or clades, named with capital letters from A to R (Table 30.1). In addition, letter Y was assigned to the most inclusive haplogroup comprising haplogroups A–R. Subclades nested within each major haplogroup are named using an alternating alphanumeric system, with lower-case letters (i.e. O3e1). Alternatively, a second nomenclature was proposed that retains the major haplogroup information followed by the name of the terminal mutation that defines a given haplogroup (i.e. O-M117).

Looking at the many published studies, considering the described markers [52–55] and newly discovered haplogroups that refine the tree [56,57], we can confirm that the Y-chromosome phylogeny is well established. It is possible that some important lineages remain undiscovered because many areas of the world have to be more deeply sampled; however, it seems we have available a good representation of the Y-chromosome lineages present all around the world.

TABLE 30.1
Y-CHROMOSOME HAPLOGROUPS AND TERMINAL SNPS

Haplogroup	SNP[a]
A	M91
A1	M31
A2	M6, M14, M23, M49, M71, M135, M141, M196, M206, M212, MEH1, P3, P4, P5, PK1
A2a	M14
A2b	P28
A3	M32
A3a	M28, M59
A3b	M144, M190, M220
A3b1	M151
A3b2	M13, M63, M127, M202, M219, M305
A3b2a	M171
A3b2b	M118
BR	$SRY_{10831.1}$ (SRY1532), M42, M94, M139, M299
B	M60, M181
B1	M236, M288
B1a	M146
B2	M182
B2a	M150
B2a1	M218
B2a1a	M109, M152, P32
B2a2	M108.1
B2a2a	M43
B2b	M112, M192, 50f2
B2b1	P6
B2b2	M115, M169
B2b3	M30, M129
B2b3a	M108.2
B2b4	P7
B2b4a	P8
B2b4b	MSY2.1, M211
C	RPS4Y (M130), M216
C1	M8, M105, M131
C2	M38
C2a	P33
C2b	M208
C3	M217, P44
C3a	M93
C3b	P39
C3c	M48, M77, M86
C3d	M407
C4	M210
C5	M356
DE	YAP (M1), M145, M203
D	M174
D1	M15
D2	M55, M57, M64.1, M179, P12, P37.1, P41.1 (M359.1), 12f2.2
D2a	M116.1
D2a1	M125

TABLE 30.1
CONTINUED

Haplogroup	SNP[a]
D2a1a	P42
D2a2	M151
D3	P47
E	SRY_{4064} (SRY8299 or M40), M96, P29
E1	M33, M132
E1a	M44
E2	M75
E2a	M41
E2b	M54, M90, M98, M85
E2b1	M85
E2b1a	M200
E3	P2, DYS391p
E3a	M2 (SY81 orDYS271), M180, P1,
E3a1	M58
E3a2	M116.2
E3a3	M149
E3a4	M154
E3a5	M155
E3a6	M10, M66, M156, M195
E3a7	M191, U186, U247
E3a7a	U174
E3a8	U175, U209
E3a8a	U290
E3a8a1	U181
E3b	M215
E3b1	M35
E3b1a	M78
E3b1a1	V12
E3b1a1a	M224
E3b1a1b	V32
E3b1a2	V13
E3b1a2a	V27
E3b1a3	V22
E3b1a3a	M148
E3b1a3b	V19
E3b1b	M81
E3b1b1	M107
E3b1b2	M183, M165
E3b1c	M123
E3b1c1	M34
E3b1c1a	M136, M84
E3b1c1b	M290
E3b1d	M281
E3b1e	V6
E3c	M329
F	P14, M89, M213, M235
F1	P91, P104
F2	M427, M428
G	M201

TABLE 30.1
CONTINUED

Haplogroup	SNP[a]
G1	M285, M342
G1a	P20
G2	P15
G2a	P16
G2a1	P17, P18
G2b	M286
G3	M287
G4	M377
H	M69
H1	M52
H1a	M82
H1a1	M36, M197
H1a2	M97
H1a3	M39, M138
H1b	M370
H2	Apt
I	M170, M258, P19, U179
I1	P38
I1a	M253, M307, P30, P40
I1a1	M227
I1a2	M21
I1a3	M72
I1b	S31
I1b1	P37.2
I1b1a	P41.2 (M359.2)
I1b1b	M26
I1b1b1	M161
I1b2	S23, S30, S32, S33
I1b2a	M223, S24
I1b2a1	M284
I1b2a2	M379
I1b2a3	P78
I1b2a4	P95
J	12f2.1, M304
J1	M267
J1a	M62
J1b	M365
J1c	M367, M368
J1d	M369
J1e	M390
J2	M172
J2a	M410
J2a1	DYS413$\leqslant$18
J2a1a	M47, M322
J2a1b	M67 (S51)
J2a1b1	M92, M260
J2a1b1a	M327
J2a1b2	M163, M166
J2a1c	M68

TABLE 30.1
CONTINUED

Haplogroup	SNP[a]
J2a1d	M137
J2a1e	M158
J2a1f	M289
J2alg	M318
J2alh	M319
J2a1i	M339
J2a1j	M419
J2a2	M340
J2b	M12, M314, M221
J2b1	M102
J2b1a	M241
J2b1a1	M99
J2b1a2	M280
J2b1a3	M321
J2b1b	M205
K	M9
K1	M353, M387
K1a	SRY9138 (M177)
K2	M70, M184, M193, M272
K2a	M320
K3	M147
K4	M230
K4a	M254
K4a1	M226
K5	P117
L	M11, M20, M22, M61, M185, M295
L1	M27, M76
L2	M317
L2a	M274
L2b	M349
L3	M357
L3a	PK3
M	M4 (M296), M5, M106, M186, M189, P35
M1	P34
M1a	P51
M2	P87
M2a	M104 (P22)
M2a1	M16
M2a2	M83
NO	M214
N	LLY22g, M231
N1	M128
N2	P43
N3	Tat (M46)
N3a	M178
N3a1	P21
O	M175
O1	MSY2.2
O1a	M119

TABLE 30.1
CONTINUED

Haplogroup	SNP[a]
O1a1	M101
O1a2	M50, M103, M110
O2	P31, M268
O2a	M95
O2a1	M88, M111
O2a1a	PK4
O2a2	M297
O2b	SRY465 (M176)
O2b1	P49
O2b1a	47z
O3	M122
O3a	M324
O3a1	M121, DYS257
O3a2	M164
O3a3	LINE1, M159
O3a4	M7
O3a4a	M113, M188, M209
O3a4a1	N4
O3a4a2	N5
O3a5	M134
O3a5a	M117, M133
O3a5a1	M162
O3a6	M300
O3a7	M333
P	92R7, M45, M74 (N12), P27
Q	MEH2, M242, P36
Q1	M120, N14 (M265)
Q1a	M378
Q2	M25, M143
Q3	M3
Q3a	M19
Q3b	M194
Q3c	M199
Q4	M323
Q5	M346
R	M207 (UTY2)
R1	M173
R1a	$\mathrm{SRY}_{10831.2}$ (SRY1532)
R1a1	M17, M198
R1a1a	M56
R1a1b	M157
R1a1c	M64.2, M87, M204
R1b	M343
R1b1	P25
R1b1a	M18
R1b1b	M73
R1b1c	M269, S3, S10, S13, S17
R1b1c1	M37
R1b1c2	M65

TABLE 30.1
CONTINUED

Haplogroup	SNP[a]
R1b1c3	M126
R1b1c4	M153
R1b1c5	M160
R1b1c6	SRY2627 (M167)
R1b1c7	M222
R1b1c8	U152
R1b1c9	U106
R1b1c9a	U198
R1b1d	M335
R2	M124

[a] Includes data from Jobling and Tyler-Smith [51]; Mohyuddin *et al.* [138]; Shen *et al.* [139]; Cruciani *et al.* [57,58,67]; Underhill *et al.* [53]; Sengupta *et al.* [71]; Hammer *et al.* [66]; Sims *et al.* [56]; Semino *et al.* [140]; Cinnioglu *et al.* [69]; Rootsi *et al.* [72]; Karlsson *et al.* [141]; Scheinfeldt *et al.* [142]; Di Giacomo *et al.* [73]; HYPERLINK "http://isogg.org/tree/ISOGG_HapgrpK07.html" Kayser *et al.* [63,143]; Regueiro *et al.* [144]; Shi *et al.* [77]; Deng *et al.* [145]; and the Ethnoancestry (Genealogy-DNA Rootsweb List, 11 January 2006).

Usually, there is a strong correlation between the Y-chromosome variation and the geography [51]. Some of the haplogroups are confined to specific areas or continents, like haplogroups A and B, confined to Africa. While other clades show an extensive geographic representation. This condition has been used to explore patterns of migrations, population substructure, and mixture among diverse populations [58–60]. In addition, this geographical distribution can be used to make inferences on the possible geographic origin of any sample of interest. In a forensic scenario this may serve as a tool to predict the origin of the paternal lineage of any stain contributor.

30.3.1.1 Global distribution of Y haplogroups

Y DNA haplogroup A represents the oldest branch of the Y-chromosome phylogeny. Like haplogroup B, it only appears in Africa, with the highest frequency among the hunter-gatherer groups in Ethiopia and Sudan [58,61]. A3b1 is a Khoisan exclusive haplogroup.

Clade C was found in Central Asia, South Asia, and East Asia [53,62–65]. C1 lineage is found exclusively in Japan. C2 is found in New Guinea, Melanesia, and Polynesia. C3 lineage is believed to have originated in Southeast or Central Asia, spreading from there into northern Asia, the Americas and Central Europe. C4 appears to be restricted among aboriginal Australians and is dominant in that population. C5 has a significant presence in India.

Haplogroup D appears in Central Asia, Southeast Asia, and in Japan [62,76] showing the highest frequencies in Tibet and Japan (50% and 35%, respectively).

Haplogroup E is one of the most branched, with many subhaplogroups described [57,58,67]. E1 and E2 were described in Northeast Africa, and E3 shows a wide geographic distribution, with two main clades: E3a, present all around Africa and among African-Americans; and E3b, present in Western Europe, North Africa, and the Near East.

Haplogroup F is the parent of haplogroups from G to R; however excluding these common haplogroups, the minor clades F*, F1, and F2, seem to appear in the Indian continent [68].

The highest frequencies of haplogroup G appear in the Caucasus region; however it also shows significant frequencies in the Mediterranean areas and the Middle East [69,70].

Until now, haplogroup H has not been well studied, members of this haplogroup were mainly found in the Indian continent [68,71].

Haplogroup I is a clear European haplogroup; it is one of the most frequent haplogroups among northwestern European populations [72].

It is generally agreed that haplogroup J was dispersed by the westward movement of people from the Middle East to North Africa, Europe, Central Asia, Pakistan, and India [54,71]. However, Di Giacomo *et al.* [73] also consider it as a signature of the expansion of the Greek world, with an accompanying novel quota of genetic variation produced during its demographic growth.

Haplogroup K is the ancestral haplogroup of major groups L to R, but, in addition, also includes the minor K* and K1 to K5 haplogroups, which are present at low frequencies in dispersed geographic regions all around the world [63].

Haplogroup L is found mainly in India and Pakistan, as well as in the Middle East and, very occasionally, in Europe, particularly in Mediterranean countries [68,71].

The highest frequencies of haplogroup M are shown in Melanesia, being restricted to the geographical distribution of Papuan languages [74–76].

The Y-DNA haplogroup N has a wide distribution, primarily in northern Eurasia, and is often associated (but not necessarily) with current and earlier Uralic speakers [72].

Lineage O represents nearly 60% of chromosomes in East Asia. The O3 haplogroup has the highest frequency, being absent outside East Asia. The O1 and O2 haplogroups appear in Malaysia, Vietnam, Indonesia, South China, Japan, and Korea [66,76,77].

The P clade is the parent of haplogroups Q and R, and is rarely found. It has been detected at low frequencies in the Caucasus and India [52,70].

Haplogroup Q is found in Asia, the Americas, Europe, and the Middle East. One of its sub-clades, group Q3 is almost exclusively associated with the Native Americans [78,79].

Finally, the last clade of the Y-chromosome tree is the extensive haplogroup R, which is mainly represented by two lineages – R1a and R1b [64,69,80,81]. The members of R1b are believed to be the descendants of the first modern humans who entered Europe, and is now the most common Y haplogroup in Europe. More than half of men of European descent belong to R1b. Haplogroup R1a is currently found in central and western Asia, India, and in Slavic populations of Eastern Europe.

30.3.1.2 Y-SNPs typing technologies

A number of different technologies have been described for the analysis of SNP variation; however the decision on the appropriate method depends mainly on the number of SNPs to be analyzed.

Most of the multiplex reactions described during these last years are based on minisequencing or base extension reactions, followed by capillary electrophoresis. Some examples are the 6 PCR multiplex reactions used by Onofri *et al.* [82] for typing 37 Y-chromosome SNPs, or the 29-plex reaction described by the SNP for ID European working group (Fig. 30.3) [81].

An alternative approach is the use of MALDI-TOF MS (matrix assisted laser desorption ionization time-of-light mass spectrometry) for detection of minisequencing products [83]. Other authors have also used microarrays for Y-SNP typing [84,85].

Finally, there are several kits commercially available based on the Luminex platform (The Signet$^{\mathrm{TM}}$ Y-SNP Genotyping Kits, Marligen Biosciences, Inc). These kits enable the detection of up to 97 SNPs, distributed in 12 multiplexes [86].

30.3.2 Minisatellite (MSY1-DYF155S1)

The most variable Y-specific marker is the minisatellite MSY1. An extremely high degree of structural variation is observed in this minisatellite using an MVR-PCR strategy, with a virtual heterozygosity of 99.9% [87–89]. From the observed diversity, a high mutation rate is expected in human minisatellites [90,91]. Based on the high degree of structural variation found in this locus, Jobling *et al.* [88] estimates the mutation rate to be between 2 and 11% per generation.

Despite the high informative potential of this minisatellite, the MVR-PCR method is technically complex and therefore the analysis of a large number of samples in population surveys is difficult [92]. Thus, the use of this minisatellite in population or in forensic genetics was not widespread.

MSY2 (DYS440) was the second minisatellite described on the Y chromosome [93]; however compared with the MSY1, it shows a very low diversity, with only 3 types of long repeat units.

30.3.3 Alphoid satellite DNA

Alphoid satellite DNA sequences are tandemly repeated arrays present in the centromere region. In the Y chromosome the alphoid DNA (Yα1, DYZ3) seems to be a functional part of the centromere [94]. A large number of alphoid patterns can be distinguished by the combination of restriction enzymes and allows the identification of haplogroup types that contribute to the definition of paternal lineages [95–98].

The variation found in alphoid haplotypes is mainly due to point mutations and insertion/deletion events. It can be detected as RFLPs and analyzed using different

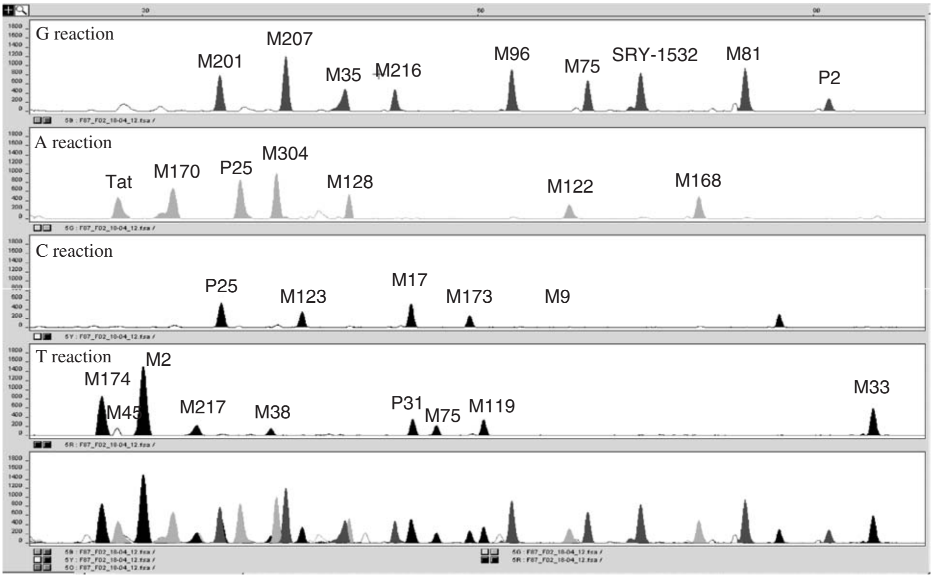

Fig. 30.3. Y-chromosome SNP 29-plex reaction described by the SNP for ID European working group.

techniques, namely conventional gel electrophoresis, pulse-field gel electrophoresis [95], or even by the identification of heteroduplex patterns of PCR-amplified fragments [99].

Although a high variability is present in these sequences, the methods usually used are not suitable for the analysis of forensic DNA samples, except those described by Santos *et al*. [95] using PCR amplification and heteroduplex analysis. Despite the importance of the information disclosed by PCR variant alphoid satellite DNA subunits in the study of human populations, this methodology can be considered quite complex for forensic purposes when compared with the existing ones for the study of other type of markers, namely the microsatellites.

30.3.4 STRs (short tandem repeats) or microsatellites

Although different kinds of polymorphisms were described on the haploid male-specific portion of the human Y chromosome, due to their levels of diversity and typing simplicity, the STRs are the most used markers in the forensic field. All the advantages already emphasized in the forensic application of autosomal STRs when compared with the study of other kind of markers are also applied to the Y-chromosome-specific STRs. In particular, they can be studied using very simple and reliable PCR techniques and, moreover, most PCR-STRs produce small amplicons with the advantage that they can be studied in degraded samples.

Since the first report of a Y-specific tetranucleotide repeat polymorphism, DYS19 (Y27H39), by Roewer and Epplen [100], over 200 Y-STR polymorphisms have been described [19,101–106], and Y-chromosome-specific STR analysis has been extensively applied to human migrations and evolution as well as to forensics e.g. [18,107].

Regarding the Y chromosome polymorphic STRs described until now, DYS19, DYS385, DYS389I, DYS389II, DYS390, DYS391, DYS392, DYS393, DYS438, and DYS439 have been selected by The Scientific Working Group on DNA Analysis Methods (SWGDAM) as the core set for forensic DNA analysis in the U.S. This same set of Y-STRs is included in the Y-STR Haplotype Reference Database (YHRD; http://www.yhrd.org), known as extended haplotype (ExtHt). With the help of collaborative efforts made in the construction of this database, these markers are the best studied, concerning amplification performance and specificity, multiplex amplification strategies, sequence structure and nomenclature, mutation rates, as well as worldwide allele frequencies distribution. Therefore, this group of markers has been, until now, the most used in population and forensic genetics.

30.3.4.1 Guidelines on the use of Y-STRs in forensic analysis

The use of Y-STRs in the forensic field has been greatly improved by coordinating efforts concerning, simultaneously, typing methodologies, nomenclature, and databasing. The use of a consensus nomenclature is crucial to allow for second opinions, proficiency testing, exchange of data, and databasing. Sequence data on Y-STRs are important in the establishment of allele nomenclature and, although for some STRs, with simple repeat structure, it is easy to find a consensus nomenclature; in others, with a complex structure, it becomes more difficult [108]. For example, for the first described STRs, nomenclature changes were made for DYS19, DYS390, and DYS389 in order to include non-repetitive motives or motives that were found to be variable with the accumulation of new sequence data. In order to clarify some confusion that still exists in the field, mainly as a consequence of the large number of new markers that have been introduced in recent years, the DNA Commission of International Society for Forensic Genetics (ISFG) recently provided recommendations on the nomenclature of Y-STR loci and alleles, as well as on its use in forensic genetics [109]. Some Y-STR markers for which sequence information is available, and a nomenclature based on the recommendations of the DNA Commission of the ISFG has already been published, are listed in Table 30.2.

30.3.4.2 Y-STR typing strategies

Multiplex PCR amplification techniques can be used in order to increase the information content of the Y-STR haplotype typing approach, and also to reduce sample quantity in forensic cases, since working with minimal amounts of material is common in casework. A great effort has been done in order to develop STR multiplex systems including a large number of markers, which will greatly improve the power of discrimination between unrelated individuals, as well as minimize costs and labor.

Nowadays, there are many different PCR multiplexing strategies described for the amplification of a large number of Y-STRs e.g. [110–114]. In addition widely used

TABLE 30.2
Y-STRS REPEAT STRUCTURE AND NOMENCLATURE

GDB Locus Name	STR Reference	Repeat Structure	Nomenclature Reference
YCAII ^{MC}	Mathias *et al.* [96]	$(CA)_n$	Schmidt *et al.* [129]
YCAIII ^{MC}	Mathias *et al.* [96]	$(CA)_n$	Kayser *et al.* [19] De Knijff *et al.* [101]
DYS19/DYS394	Roewer and Epplen [101]	$(TAGA)_3tagg(TAGA)_n$	Kayser *et al.* [19] De Knijff *et al.* [101]
DYS385	Kayser *et al.* [19] De Knijff *et al.* [101]	$(aagg)_{6-7}(GAAA)_n$	Kayser *et al.* [19] De Knijff *et al.* [101]
DYS388	Kayser *et al.* [19] De Knijff *et al.* [101]	$(ATT)_n$	Butler *et al.* [111]
DYS389 I	Kayser *et al.* [19] De Knijff *et al.* [101]	$(TCTG)_3(TCTA)_n$	Kayser *et al.* [19] De Knijff *et al.* [101]
DYS389 II	Kayser *et al.* [19] De Knijff *et al.* [101]	$(TCTG)_n(TCTA)_nN_{28}(TCTG)_3(TCTA)_n$	Kayser *et al.* [19] De Knijff *et al.* [101]
DYS390	Kayser *et al.* [19] De Knijff *et al.* [101]	$(tcta)_2(TCTG)_n(TCTA)_n(TCTG)_n(TCTA)_ntca(tcta)_2$	Kayser *et al.* [19] De Knijff *et al.* [101]
DYS391	Kayser *et al.* [19] De Knijff *et al.* [101]	$(tctg)_3(TCTA)_n$	Kayser *et al.* [19] De Knijff *et al.* [101]
DYS392	Kayser *et al.* [19] De Knijff *et al.* [101]	$(TAT)_n$	Kayser *et al.* [19] De Knijff *et al.* [101]
DYS393/DYS395	Kayser *et al.* [19] De Knijff *et al.* [101]	$(AGAT)_n$	Kayser *et al.* [19] De Knijff *et al.* [101]
DYS426	Jobling *et al.* [106]	$(GTT)_n$	Butler *et al.* [111]
DYS434	Ayub *et al.* [103]	$(TAAT)_{1-2}(CTAT)_n$	Gusmão *et al.* [108]
DYS435	Ayub *et al.* [103]	$(TGGA)_n$	Gusmão *et al.* [108]
DYS436	Ayub *et al.* [103]	$(GTT)_n$	Gusmão *et al.* [108]
DYS437	Ayub *et al.* [103]	$(TCTA)_n(TCTG)_{1-3}(TCTA)_4$	Gusmão *et al.* [108]
DYS438	Ayub *et al.* [103]	$(TTTTC)_1(TTTTA)_{0-1}(TTTTC)_n$	Gusmão *et al.* [108]
DYS439 (GATA A4)	Ayub *et al.* [103]	$(GATA)_n$	Gusmão *et al.* [108]
DYS441	Iida *et al.* [104]	$(TTCC)_n$	[a]
DYS442	Iida *et al.* [104]	$(TATC)_2\ (TGTC)_3\ (TATC)_n$	[a]

TABLE 30.2
CONTINUED

GDB Locus Name	STR Reference	Repeat Structure	Nomenclature Reference
DYS443	Iida *et al.* [104]	$(TTCC)_n$	Iida *et al.* [104]
DYS444	Iida *et al.* [104]	$(ATAG)_n$	[a]
DYS445	Iida *et al.* [104]	$(TTTA)_n$	Iida *et al.* [104]
DYS446	Redd *et al.* [105]	$(TCTCT)_n$	Redd *et al.* [105]
DYS447	Redd *et al.* [105]	$(TAATA)_n(TAAAA)_1(TAATA)_n(TAAAA)_1(TAATA)_n$	Redd *et al.* [105]
DYS448	Redd *et al.* [105]	$(AGAGAT)_n\ N_{42}(AGAGAT)_n$	Redd *et al.* [105]
DYS449	Redd *et al.* [105]	$(TTTC)_n\ N_{50}(TTTC)_n$	Redd *et al.* [105]
DYS450	Redd *et al.* [105]	$(TTTTA)_n$	Redd *et al.* [105]
DYS452	Redd *et al.* [105]	$(TATAC)_2(TGTAC)_2(TATAC)_n(CATAC)_1(TATAC)_1(CATAC)_1$ $(TATAC)_{3-4}(CATAC)_{0-2}(TATAC)_{0-3}(CATAC)_1(TATAC)_3$	Redd *et al.* [105]
DYS453	Redd *et al.* [105]	$(AAAT)_n$	Redd *et al.* [105]
DYS454	Redd *et al.* [105]	$(AAAT)_n$	Redd *et al.* [105]
DYS455	Redd *et al.* [105]	$(AAAT)_n$	Redd *et al.* [105]
DYS456	Redd *et al.* [105]	$(AGAT)_n$	Redd *et al.* [105]
DYS458	Redd *et al.* [105]	$(GAAA)_n$	Redd *et al.* [105]
DYS459 ^{MC}	Redd *et al.* [105]	$(TAAA)_n$	Redd *et al.* [105]
DYS460 (GATA A7.1)	White *et al.* [102]	$(ATAG)_n$	Gusmão *et al.* [108]
DYS461 (GATA A7.2)	White *et al.* [102]	$(TAGA)_n(CAGA)$	Gusmão *et al.* [108]
DYS462 (G09411)	Bosch *et al.* [110]	$(TATG)_n$	Bosch *et al.* [110]
DYS463	Redd *et al.* [105]	$(AAAGG)_n(AAGGG)_n(AAGGA)_2$	Redd *et al.* [105]
DYS464^{MC}	Redd *et al.* [105]	$(CCTT)_n$	Redd *et al.* [105]
DYS485	Kayser *et al.* [17]	$(TTA)_n$	Butler *et al.* [130]
DYS490	Kayser *et al.* [17]	$(TTA)_n$	Butler *et al.* [130]
DYS495	Kayser *et al.* [17]	$(AAT)_n$	Butler *et al.* [130]
DYS504	Kayser *et al.* [17]	$(TCCT)_n$	Butler *et al.* [130]
DYS505	Kayser *et al.* [17]	$(TCCT)_n$	Butler *et al.* [130]
DYS508	Kayser *et al.* [17]	$(TATC)_n$	Butler *et al.* [130]
DYS510	Kayser *et al.* [17]	$(TAGA)_3(TACA)(TAGA)(TACA)(TAGA)_n$	Dai *et al.* [131]
DYS513	Kayser *et al.* [17]	$(TATC)_n$	Dai *et al.* [131]
DYS520	Kayser *et al.* [17]	$(ATAG)_n\ (ATAC)_n$	Butler *et al.* [130]
DYS522	Kayser *et al.* [17]	$(GATA)_n$	Butler *et al.* [130]

DYS525	Kayser *et al.* [17]	(TAGA)$_n$	Butler *et al.* [130]
DYS532	Kayser *et al.* [17]	(CTTT)$_n$	Butler *et al.* [130]
DYS533	Kayser *et al.* [17]	(ATCT)$_n$	Butler *et al.* [130]
DYS534	Kayser *et al.* [17]	(CTTT)$_n$	Butler *et al.* [130]
DYS540	Kayser *et al.* [17]	(TTAT)$_n$	Butler *et al.* [130]
DYS542	Kayser *et al.* [17]	(ATAG)$_2$ ATAA (ATAG)$_n$	Butler *et al.* [130]
DYS544	Kayser *et al.* [17]	(GATA)$_3$ GATG (GATA)$_n$	Dai *et al.* [131]
DYS552	Kayser *et al.* [17]	(TCTA)$_3$ TCTG (TCTA)$_n$ N$_{40}$ (TCTA)$_n$	Dai *et al.* [131]
DYS556	Kayser *et al.* [17]	(AATA)$_n$	Butler *et al.* [130]
DYS557	Kayser *et al.* [17]	(TTTC)$_n$	Butler *et al.* [130]
DYS561	Kayser *et al.* [17]	(GATA)$_n$ (GACA)$_4$	Dai *et al.* [131]
DYS570	Kayser *et al.* [17]	(TTTC)$_n$	Butler *et al.* [130]
DYS575	Kayser *et al.* [17]	(AAAT)$_n$	Butler *et al.* [130]
DYS576	Kayser *et al.* [17]	(AAAG)$_n$	Butler *et al.* [130]
DYS587	Kayser *et al.* [17]	(ATACA)$_n$ [(GTACA)(ATACA)]$_3$	Dai *et al.* [131]
DYS593	Kayser *et al.* [17]	(AAAAC)$_2$ AAAAT (AAAAC)$_4$ (AAAAT)$_n$	Dai *et al.* [131]
DYS594	Kayser *et al.* [17]	(TAAAA)$_n$	Butler *et al.* [130]
DYS632	Kayser *et al.* [17]	(CATT)$_n$	Butler *et al.* [130]
DYS635 (GATA-C4)	White *et al.* [102]	(TCTA)$_4$(TGTA)$_2$(TCTA)$_2$(TGTA)$_2$(TCTA)$_2$(TGTA)$_{0.2}$(TCTA)$_n$	Gusmão *et al.* [108]
DYS641	Kayser *et al.* [17]	(TAAA)$_n$	Butler *et al.* [130]
DYS643	Kayser *et al.* [17]	(CTTTT)$_n$	Butler *et al.* [130]
GATA-A10	White *et al.* [102]	(TCCA)$_2$(TATC)$_n$	Gusmão *et al.* [108]
GATA-H4	White *et al.* [102]	(AGAT)$_4$ CTAT(AGAT)$_2$(AGGT)$_3$(AGAT)$_n$ N$_{24}$(ATAG)$_4$(ATAC)$_1$(ATAG)$_2$	Gusmão *et al.* [108]
GATA-H4.1	White *et al.* [102]	(AGAT)$_4$ CTAT(AGAT)$_2$(AGGT)$_3$(AGAT)$_n$	Gusmão *et al.* [108]

Note: ^{MC} Multi-copy Y-STR.

[a] Modified in order to observe the ISFG recommendations.

commercial kits are also available allowing the simultaneous typing of as much as 12 (PowerPlex® Y System [115]; Promega Corporation) and even 17 markers (AmpFlSTR YFiler Amplification Kit [116]; AB Applied Biosystems).

30.3.4.3 Population genetics and databasing

The determination of Y-STR haplotype frequencies in different populations is a main point for the correct interpretation of the genetic profile matches in kinship analysis and forensic casework.

The use of Y-STRs as inclusion evidence involves population genetic profile definition, with the elaboration of a large number of databases. Because of the lack of recombination between Y-chromosome-specific markers, they are transmitted as haplotypes in the same way as single locus alleles, and the construction of Y-specific STR databases seems to be more complex than for unlinked AS markers, since the whole haplotype must be typed for each sample. First, the informative content of Y-specific STRs, results from the possibility of constructing highly discriminative haplotypes. The suitability of the Y-STR databases for practical use will be greatly increased with the typing of each individual to as many loci as possible, instead of typing a great number of individuals for a small number of Y-STRs. On the other hand, population substructuring seems to be more severe in the case of Y chromosome than for unlinked AS markers. Due to the lower effective number of Y chromosomes in a population, Y haplotypes/haplogroups tend to present a higher proportion of variation between populations than that observed for other markers located on autosomes or X chromosome. The interpopulational variability on Y-profiles makes the definition of local databases crucial for the application of Y-specific markers in practice. DNA Commission of the International Society of Forensic Genetics (ISFG) recommended the use of regional Y-STR haplotype databases, and pooling data from different regions is only valid after verifying that no population substructure exists [109].

A large amount of Y-haplotype data is, therefore, essential for two main reasons: (a) in match probability calculation, since it is not valid to multiply the allele frequency of each locus, a large number of haplotypes are needed to allow reliable frequency estimates; (b) population substructure analysis is highly dependent on the amount of data available.

Due to their importance in forensics, Y-STR haplotype distributions in populations worldwide have been made available not only through classical publications but also, more recently, through large-scale forensic databases. The YHRD is the most extensive survey available on line and the development of these databases is important not only for haplotype frequency estimation and subsequent application for match probability calculations in forensic studies, but also for performing comparative population analysis.

30.3.4.4 Y-STR mutation rates

According to data on autosomal STRs, the frequency of mutation events in the male germ line is higher than in the female germ line [45,117]. In a survey on Y and

TABLE 30.3
Y-STR MUTATION RATES

	Total[a]			
Locus	Nr. Mutations	Allele Transmissions	Frequency ($\times 10^{-3}$)	95% CI $\times 10^{-3}$
DYS19	13	7314	1.7774	0.947–3.038
DYS389 I	11	5518	1.9935	0.996–3.564
DYS389 II	12	5505	2.1798	1.127–3.805
DYS390	15	6796	2.2072	1.236–3.638
DYS391	23	6744	3.4104	2.163–5.113
DYS392	4	6710	0.5961	0.162–1.526
DYS393	4	5498	0.7275	0.198–1.862
DYS385	24	10207	2.3513	1.507–3.497
DYS437	5	2437	2.0517	0.667–4.781
DYS438	1	2476	0.4039	0.010–2.248
DYS439	12	2451	4.8960	2.532–8.537
GATA A10	4	946	4.228	1.153–10.971
DYS460	5	1109	4.509	1.465–10.490
DYS461	0	873	–	0.000–4.217
DYS635	3	873	3.436	0.709–10.010
GATA H4	3	1036	2.896	0.598–8.439
Total	139	66493	2.090	1.758–2.468

[a] Includes data from Heyer *et al.* [132]; Bianchi *et al.* [133]; Kayser *et al.* [46]; Dupuy *et al.* [47]; Kurihara *et al.* [125]; Góes *et al.* [134]; Budowle *et al.* [135]; Ballard *et al.* [136]; Gusmão *et al.* [48]; Turrina *et al.* [43]; and Domingues *et al.* [137].

X-linked loci, Scozzari *et al.* [118] report an overall higher diversity for the Y-linked loci and suggest a higher rate of accumulation of variants on this chromosome. These results can be explained by the higher number of divisions involved in male gametogenesis when compared with the female one, and should be reflected in a higher mutation rate on Y-STRs.

A large amount of data is necessary to estimate reliable allele specific mutation rates, essential for consistent dating of Y-SNP defined lineages (haplogroups) and data interpretation in kinship analysis. Although studies on Y-STR mutation rates are still scarce and have only considered a restricted number of markers, no significant differences were found between the average Y-STR mutation rates (see Table 30.3) and those found in autosomal STRs [45,119,120]. Data on Y-STR mutations also support that slippage is the mechanism involved, and that there is an agreement with the generally accepted single-step mutation model for microsatellites in which the alleles are known to mutate primarily through the gain and loss of single repeat units [117,121,122].

STR mutation rates present not only inter- but also intra-locus variation, depending on the locus structure and allele length e.g. [45,123]. When compared to the average value, a significantly lower mutation rate was observed at DYS392 and higher at DYS391 and DYS439 loci [48]. For TAGA repeats, a trend for higher mutability was confirmed for longer alleles.

In Y-STRs, repeat gains were found to be twice more frequent than losses [46–48,124,125], as expected for young microsatellites [126]. When comparing repeat gains and losses in different allele classes, Gusmão *et al.* [48] found no evidence for an excess of repeat losses at longer alleles supporting that this mechanism is biased toward microsatellite expansion e.g. [45,127,128] until a certain repeat length at which the rates of expansion and contraction mutations are equal [126].

30.4 ACKNOWLEDGMENTS

IPATIMUP is partially supported by Fundação para a Ciência e a Tecnologia (through POCI, Programa Operacional Ciência e Inovação). Iva Gomes is a recipient of a grant from Fundação para a Ciência e a Tecnologia (SFRH/BD/21647/2005). CeGen (National Genotyping Center of Spain) is funded by "Fundacion Genoma España."

30.5 REFERENCES

1 H. Skaletsky, T. Kuroda-Kawaguchi, P.J. Minx, H.S. Cordum, L. Hillier, L.G. Brown, S. Repping, T. Pyntikova, J. Ali, T. Bieri, A. Chinwalla, A. Delehaunty, K. Delehaunty, H. Du, G. Fewell, L. Fulton, R. Fulton, T. Graves, S.F. Hou, P. Latrielle, S. Leonard, E. Mardis, R. Maupin, J. McPherson, T. Miner, W. Nash, C. Nguyen, P. Ozersky, K. Pepin, S. Rock, T. Rohlfing, K. Scott, B. Schultz, C. Strong, A. Tin-Wollam, S.P. Yang, R.H. Waterston, R.K. Wilson, S. Rozen and D.C. Page, The male-specific region of the human Y chromosome is a mosaic of discrete sequence classes. Nature, 423 (2003) 825–837.

2 S. Foote, D. Vollrath, A. Hilton and D.C. Page, The human Y chromosome: overlapping DNA clones spanning the euchromatic region. Science, 258 (1992) 60–66.

3 B. Charlesworth, The evolution of sex chromosomes. Science, 251 (1991) 1030–1033.

4 B.T. Lahn and D.C. Page, Four evolutionary strata on the human X chromosome. Science, 286 (1999) 964–967.

5 S. Rozen, H. Skaletsky, J.D. Marszalek, P.J. Minx, H.S. Cordum, R.H. Waterston, R.K. Wilson and D.C. Page, Abundant gene conversion between arms of palindromes in human and ape Y chromosomes. Nature, 423 (2003) 873–876.

6 J.F. Hughes, H. Skaletsky, T. Pyntikova, P.J. Minx, T. Graves, S. Rozen, R.K. Wilson and D.C. Page, Conservation of Y-linked genes during human evolution revealed by comparative sequencing in chimpanzee. Nature, 437 (2005) 100–103.

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

8 S. Jakubiczka, J. Arnemann, H.J. Cooke, M. Krawczak and J. Schmidtke, A search for restriction fragment length polymorphism on the human Y chromosome. Hum. Genet., 84 (1989) 86–88.

9 P. Malaspina, F. Persichetti, A. Novelletto, C. Iodice, L. Terrenato, J. Wolfe, M. Ferraro and G. Prantera, The human Y chromosome shows a low level of DNA polymorphism. Ann. Hum. Genet., 54 (1990) 297–305.

10 A. Spurdle and T. Jenkins, The search for Y chromosome polymorphism is extended to negroids. Hum. Mol. Genet., 1 (1992) 169–170.

11 R.L. Dorit, H. Akashi and W. Gilbert, Absence of polymorphism at the ZFY locus on the human Y chromosome. Science, 268 (1995) 1183–1185.

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

13 M.A. Jobling, A survey of long-range DNA polymorphisms on the human Y chromosome. Hum. Mol. Genet., 3 (1994) 107–114.

14 A. Sajantila, A.H. Salem, P. Savolainen, K. Bauer, C. Gierig and S. Pääbo, Paternal and maternal DNA lineages reveal a bottleneck in the founding of the Finnish population. Proc. Natl. Acad. Sci., 93 (1996) 12035–12039.

15 M.A. Jobling and C. Tyler-Smith, Fathers and sons: the Y chromosome and human evolution. Trends Genet., 11 (1995) 449–456.

16 P.A. Underhill, P. Shen, A.A. Lin, L. Jin, G. Passarino, W.H. Yang, E. Kauffman, B. Bonne-Tamir, J. Bertranpetit, P. Francalacci, M. Ibrahim, T. Jenkins, J.R. Kidd, S.Q. Mehdi, M.T. Seielstad, R.S. Wells, A. Piazza, R.W. Davis, M.W. Feldman, L.L. Cavalli-Sforza and P.J. Oefner, Y chromosome sequence variation and the history of human populations. Nat. Genet., 26 (2000) 358–361.

17 M. Kayser, R. Kittler, A. Erler, M. Hedman, A.C. Lee, A. Mohyuddin, S.Q. Mehdi, Z. Rosser, M. Stoneking, M.A. Jobling, A. Sajantila and C. Tyler-Smith, A comprehensive survey of human Y-chromosomal microsatellites. Am. J. Hum. Genet., 74 (2004) 1183–1197.

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

19 M. Kayser, A. Caglià, D. Corach, N. Fretwell, C. Gehrig, G. Graziosi, F. Heidorn, S. Herrmann, B. Herzog, M. Hidding, K. Honda, M. Jobling, M. Krawczak, K. Leim, S. Meuser, E. Meyer, W. Oesterreich, A. Pandya, W. Parson, G. Penacino, A. Perez-Lezaun, A. Piccinini, M. Prinz, C. Schmitt, P.M. Schneider, R. Szibor, J. Teifel-Greding, G. Weichhold, P. de Knijff and L. Roewer, Evaluation of Y-chromosomal STRs: a multicenter study. Int. J. Legal Med., 110 (1997) 125–149.

20 M. Prinz, A. Ishii, A. Coleman, H.J. Baum and R.C. Shaler, Validation and casework application of a Y chromosome specific STR multiplex. Forensic Sci. Int., 120 (2001) 177–188.

21 L. Roewer, M. Krawczak, S. Willuweit, M. Nagy, C. Alves, A. Amorim, K. Anslinger, C. Augustin, A. Betz, E. Bosch, A. Caglia, A. Carracedo, D. Corach, A.F. Dekairelle, T. Dobosz, B.M. Dupuy, S. Furedi, C. Gehrig, L. Gusmão, J. Henke, L. Henke, M. Hidding, C. Hohoff, B. Hoste, M.A. Jobling, H.J. Kargel, P. de Knijff, R. Lessig, E. Liebeherr, M. Lorente, B. Martinez-Jarreta, P. Nievas, M. Nowak, W. Parson, V.L. Pascali, G. Penacino, R. Ploski, B. Rolf, A. Sala, U. Schmidt, C. Schmitt, P.M. Schneider, R. Szibor, J. Teifel-Greding and M. Kayser, Online reference database of European Y-chromosomal short tandem repeat (STR) haplotypes. Forensic Sci. Int., 118 (2001) 106–113.

22 A. Dettlaff-Kakol and R. Pawlowski, First Polish DNA "manhunt": an application of Y-chromosome STRs. Int. J. Legal Med., 116 (2002) 289–291.

23 M. Prinz, K. Boll, H. Baum and B. Shaler, Multiplexing of Y chromosome specific STRs and performance for mixed samples. Forensic Sci. Int., 85 (1997) 209–218.

24 R. Zehner and U. Bohrer, DYS19 and amelogenin in artificial blood stains with defined amounts of male and female cells. Int. J. Legal Med., 111 (1998) 340–342.

25 U.B. Gyllensten, A. Josefsson, K. Schemschat, T. Saldeen and U. Petterson, DNA typing of forensic material with mixed genotypes using allele-specific enzymatic amplification (polymerase chain reaction). Forensic Sci. Int., 52 (1992) 149–160.

26 J. Kreike and A. Lehner, Sex determination and DNA competition in the analysis of forensic mixed stains by PCR. Int. J. Legal Med., 107 (1995) 235–238.

27 A. Betz, G. Babler, G. Dietl, X. Steil, G. Weyermann and W. Pflug, DYS STR analysis with epithelial cells in a rape case. Forensic Sci. Int., 118 (2001) 126–130.

28 R. Szibor, M. Krawczak, S. Hering, J. Edelmann, E. Kuhlisch and D. Krause, Use of X-linked markers for forensic purposes. Int. J. Legal Med., 117 (2003) 67–74.

29 R. Szibor, I. Plate, J. Edelmann, S. Hering, E. Kuhlisch, M. Michael and D. Krause, Chromosome X haplotyping in deficiency paternity testing principles and case report. Int. Congr. Ser., 1239 (2003) 815–820.

30 R. Szibor, S. Hering, E. Kuhlisch, I. Plate, S. Demberger, M. Krawczak and J. Edelmann, Haplotyping of STR cluster DXS6801-DXS6809-DXS6789 on Xq21 provides a powerful tool for kinship testing. Int. J. Legal Med., 119 (2005) 363–369.

31 S. Hering, C. Augustin, J. Edelmann, M. Heidel, J. Dressler, H. Rodig, E. Kuhlisch and R. Szibor, DXS10079, DXS10074 and DXS10075 are STRs located within a 280 kb region of Xq 12 and provide stable haplotypes useful for complex kinship cases. Int. J. Legal Med., 120 (2006) 337–345.

32 C. Toni, S. Presciuttini, I. Spinetti and R. Domenici, Population data of four X-chromosome markers in Tuscany, and their use in a deficiency paternity case. Forensic Sci. Int., 137 (2003) 215–216.

33 S. Turrina, R. Atzei and D. De Leo, Polymorphism of four X-chromosomal STRs: DXS7423, DXS7424, DXS8378 and DXS6809 in a North Italian population sample and their use in kinship testing. Forensic Sci. Int., 168 (2007) 241–243.

34 J. Edelmann, R. Lessig, M. Klintschar and R. Szibor, Advantages of X-chromosomal microsatellites in deficiency paternity testing: presentation of cases. Int. Congr. Ser., 1261 (2004) 257–259.

35 S.F. Schaffner, The X chromosome in population genetics. Nat. Rev. Genet., 5 (2004) 43–51.

36 C. Bini, S. Ceccardi, G. Ferri, S. Pelotti, M. Alu′, E. Rocaglia, G. Beduschi, L. Caenazzo, E. Ponzano, P. Tasinato, C. Turchi, M. Mazzanti, A. Tagliabracci, C. Toni, I. Spinetti, R. Domenico and S. Presciuttini, Development of a heptaplex PCR system to analyse X-chromosome STR loci from five Italian population samples: a collaborative study. Forensic Sci. Int., 153 (2005) 231–236.

37 C. Robino, A. Giolitti, S. Gino and C. Torre, Development of a heptaplex PCR system to analyse X-chromosome STR loci from five Italian population samples: a collaborative study. Int. J. Legal Med., 120 (2006) 315–318.

38 I. Gomes, M. Prinz, P. Pereira, C. Meyers, R.S. Mikulasovich, A. Amorim, A. Carracedo and L. Gusmão, Genetic analysis of three US population groups using an X-chromosomal STR decaplex. Int. J. Legal Med., 121 (2007) 198–203.

39 M.A. Moreno, J.J. Builes, P. Jaramillo, C. Espinal, D. Aguirre, M.M. Pancorbo, L. Gusmão and M.L.J. Bravo, Validation of five X-chromosomal STRs DXS6800, DXS6807, DXS6798, DXS8377 and DXS7423 in an Antioquian population sample. Int. Congr. Ser., 1288 (2006) 295–297.

40 R. Szibor, S. Hering and J. Edelmann, A new Web site compiling forensic chromosome X research is now online. Int. J. Legal Med., 120 (2006) 252–254.

41 M.T. Zarrabeitia, T. Amigo, C. Sanudo, A. Zarrabeitia, D. González-Lamuño and J.A. Riancho, A new pentaplex system to study short tandem repeat markers of forensic interest on X chromosome. Forensic Sci. Int., 129 (2002) 85–89.

42 K.J. Shin, B.K. Kwon, S.S. Lee, J.E. Yoo, M.J. Park, U. Chung, H.Y. Lee, G.R. Han, J.H. Choi and C.Y. Kim, Five highly informative X-chromosomal STRs in Koreans. Int. J. Legal Med., 118 (2004) 37–40.

43 S. Turrina and D. De Leo, Population genetic comparisons of three X-chromosomal STRs (DXS7132, DXS7133 and GATA172D05) in North and South Italy. Int. Congr. Ser., 1261 (2004) 302–304.

44 M. Poetsch, H. Petersmann, A. Repenning and E. Lignitz, Development of two pentaplex systems with X-chromosomal STR loci and their allele frequencies in a northeast German population. Forensic Sci. Int., 155 (2005) 71–76.

45 B. Brinkmann, M. Klintschar, F. Neuhuber, J.B. Hühne and B. Rolf, Mutation rate in human microsatellites: influence of the structure and length of the tandem repeat. Am. J. Hum. Genet., 62 (1998) 1408–1415.

46 M. Kayser, L. Roewer, M. Hedman, L. Henke, J. Henke, S. Brauer, C. Krüger, M. Krawczak, M. Nagy, T. Dobosz, R. Szibor, P. de Knijff, M. Stoneking and A. Sajantila, Characteristics and frequency of germline mutations at microsatellite loci from the human Y chromosome as revealed by direct observation in father/son pairs. Am. J. Hum. Genet., 66 (2000) 1580–1588.

47 B.M. Dupuy, M. Stenersen, T. Egeland and B. Olaisen, Y-chromosomal microsatellite mutation rates: differences in mutation rate between and within loci. Hum. Mutat., 23 (2004) 117–124.

48 L. Gusmão, P. Sánchez-Diz, F. Calafell, P. Martín, C.A. Alonso, F. Álvarez-Fernández, C. Alves, L. Borjas-Fajardo, W.R. Bozzo, M.L. Bravo, J.J. Builes, J. Capilla, M. Carvalho, C. Castillo, C.I. Catanesi, D. Corach, A.M. Di Lonardo, R. Espinheira, E. Fagundes de Carvalho, M.J. Farfán, H.P. Figueiredo, I. Gomes, M.M. Lojo, M. Marino, M.F. Pinheiro, M.L. Pontes, V. Prieto, E. Ramos-Luis, J.A. Riancho, A.C. Souza Go′ es, O.A. Santapa, D.R. Sumita, G. Vallejo, L. Vidal Rioja, M.C. Vide, C.I. Vieira da Silva, M.R. Whittle, W. Zabala, M.T. Zarrabeitia, A. Alonso, A. Carracedo and A. Amorim, Mutation rates at Y chromosome specific microsatellites. Hum. Mutat., 26 (2005) 520–528.

49 V.L. Pascali, M. Dobosz and B. Brinkmann, Coordinating Y-chromosomal STR research for the Courts. Int. J. Legal Med., 112 (1999) 1.

50 M.E. Hurles, C. Irven, J. Nicholson, P.G. Taylor, F.R. Santos, J. Loughlin, M.A. Jobling and B.C. Sykes, European Y-chromosomal lineages in Polynesians: a contrast to the population structure revealed by mtDNA. Am. J. Hum. Genet., 63 (1998) 1793–1806.

51 M.A. Jobling and C. Tyler-Smith, The human Y chromosome: an evolutionary marker comes of age. Nat. Rev. Genet., 4 (2003) 598–612.

52 P.A. Underhill, P. Shen, A.A. Lin, L. Jin, G. Passarino, W.H. Yang, E. Kauffman, B. Bonne-Tamir, J. Bertranpetit, P. Francalacci, M. Ibrahim, T. Jenkins, J.R. Kidd, S.Q. Mehdi, M.T. Seielstad, R.S. Wells, A. Piazza, R.W. Davis, M.W. Feldman, L.L. Cavalli-Sforza and P.J. Oefner, Y chromosome sequence variation and the history of human populations. Nat. Genet., 26 (2000) 358–361.

53 P.A. Underhill, G. Passarino, A.A. Lin, P. Shen, M. Mirazon Lahr, R.A. Foley, P.J. Oefner and L.L. Cavalli-Sforza, The phylogeography of Y chromosome binary haplotypes and the origins of modern human populations. Ann. Hum. Genet., 65 (2001) 43–62.

54 L. Quintana-Murci, C. Krausz, T. Zerjal, S.H. Sayar, M.F. Hammer, S.Q. Mehdi, Q. Ayub, R. Qamar, A. Mohyuddin, U. Radhakrishna, M.A. Jobling, C. Tyler-Smith and K. McElreavey, Y-chromosome lineages trace diffusion of people and languages in southwestern Asia. Am. J. Hum. Genet., 6 (2001) 537–542.

55 O. Semino, G. Passarino, P.J. Oefner, A.A. Lin, S. Arbuzova, L.E. Beckman, G. De Benedictis, P. Francalacci, A. Kouvatsi, S. Limborska, M. Marcikiae, A. Mika, B. Mika, D. Primorac, A.S. Santachiara-Benerecetti, L.L. Cavalli-Sforza and P.A. Underhill, The genetic legacy of paleolithic *Homo sapiens sapiens* in extant Europeans: a Y chromosome perspective. Science, 290 (2000) 1155–1159.

56 L.M. Sims, D. Garvey and J. Ballantyne, Sub-populations within the major European and African derived haplogroups R1b3 and E3a are differentiated by previously phylogenetically undefined Y-SNPs. Hum. Mutat., 28 (2007) 97.

57 F. Cruciani, R. La Fratta, A. Torroni, P.A. Underhill and R. Scozzari, Molecular dissection of the Y chromosome haplogroup E-M78 (E3b1a): a posteriori evaluation of a microsatellite-network-based approach through six new biallelic markers. Hum. Mutat., 27 (2006) 831–832.

58 F. Cruciani, P. Santolamazza, P. Shen, V. Macaulay, P. Moral, A. Olckers, D. Modiano, S. Holmes, G. Destro-Bisol, V. Coia, D.C. Wallace, P.J. Oefner, A. Torroni, L.L. Cavalli-Sforza, R. Scozzari and P.A. Underhill, A back migration from Asia to sub-Saharan Africa is supported by high-resolution analysis of human Y-chromosome haplotypes. Am. J. Hum. Genet., 70 (2002) 1197–1214.

59 M. Brion, B. Quintans, M. Zarrabeitia, A. Gonzalez-Neira, A. Salas, V. Lareu, C. Tyler-Smith and A. Carracedo, Micro-geographical differentiation in Northern Iberia revealed by Y-chromosomal DNA analysis. Gene, 329 (2004) 17–25.

60 M.F. Hammer, V.F. Chamberlain, V.F. Kearney, D. Stover, G. Zhang, T. Karafet, B. Walsh and A.J. Redd, Population structure of Y chromosome SNP haplogroups in the United States and forensic implications for constructing Y chromosome STR databases. Forensic Sci. Int., 164 (2006) 45–55.

61 O. Semino, A.S. Santachiara-Benerecetti, F. Falaschi, L.L. Cavalli-Sforza and P.A. Underhill, Ethiopians and Khoisan share the deepest clades of the human Y-chromosome phylogeny. Am. J. Hum. Genet., 70 (2002) 265–268.

62 T.M. Karafet, L. Xu, R. Du, W. Wang, S. Feng, R.S. Wells, A.J. Redd, S.L. Zegura and M.F. Hammer, Paternal population history of East Asia: sources, patterns, and microevolutionary processes. Am. J. Hum. Genet., 69 (2001) 615–628.

63 M. Kayser, S. Brauer, G. Weiss, W. Schiefenhövel, P. Underhill, P. Shen, P. Oefner, M. Tommaseo-Ponzetta and M. Stoneking, Reduced Y-chromosome, but not mitochondrial DNA, diversity in human populations from West New Guinea. Am. J. Hum. Genet., 72 (2003) 281–302.

64 R.S. Wells, N. Yuldasheva, R. Ruzibakiev, P.A. Underhill, I. Evseeva, J. Blue-Smith, L. Jin, B. Su, R. Pitchappan, S. Shanmugalakshmi, K. Balakrishnan, M. Read, N.M. Pearson, T. Zerjal, M.T. Webster, I. Zholoshvili, E. Jamarjashvili, S. Gambarov, B. Nikbin, A. Dostiev, O. Aknazarov, P. Zalloua, I. Tsoy, M. Kitaev, M. Mirrakhimov, A. Chariev and W.F. Bodmer, The Eurasian heartland: a continental perspective on Y-chromosome diversity. Proc. Natl. Acad. Sci., 98 (2001) 10244–10249.

65 T. Zerjal, R.S. Wells, N. Yuldasheva, R. Ruzibakiev and C. Tyler-Smith, A genetic landscape reshaped by recent events: Y-chromosomal insights into central Asia. Am. J. Hum. Genet., 71 (2002) 466–482.

66 M.F. Hammer, T.M. Karafet, H. Park, K. Omoto, S. Harihara, M. Stoneking and S. Horai, Dual origins of the Japanese: common ground for hunter-gatherer and farmer Y chromosomes. J. Hum. Genet., 51 (2006) 47–58.

67 F. Cruciani, R. La Fratta, P. Santolamazza, D. Sellitto, R. Pascone, P. Moral, E. Watson, V. Guida, E.B. Colomb, B. Zaharova, J. Lavinha, G. Vona, R. Aman, F. Cali, N. Akar, M. Richards, A. Torroni, A. Novelletto and R. Scozzari, Phylogeographic analysis of haplogroup E3b (E-M215) Y chromosomes reveals multiple migratory events within and out of Africa. Am. J. Hum. Genet., 74 (2004) 1014–1022.

68 T. Kivisild, S. Rootsi, M. Metspalu, S. Mastana, K. Kaldma, J. Parik, E. Metspalu, M. Adojaan, H.V. Tolk, V. Stepanov, M. Golge, E. Usanga, S.S. Papiha, C. Cinnioglu, R. King, L. Cavalli-Sforza, P.A. Underhill and R. Villems, The genetic heritage of the earliest settlers persists both in Indian tribal and caste populations. Am. J. Hum. Genet., 72 (2003) 313–332.

69 C. Cinnioglu, R. King, T. Kivisild, E. Kalfoglu, S. Atasoy, G.L. Cavalleri, A.S. Lillie, C.C. Roseman, A.A. Lin, K. Prince, P.J. Oefner, P. Shen, O. Semino, L.L. Cavalli-Sforza and P.A. Underhill, Excavating Y-chromosome haplotype strata in Anatolia. Hum. Genet., 114 (2004) 127–148.

70 I. Nasidze, D. Quinque, I. Dupanloup, S. Rychkov, O. Naumova, O. Zhukova and M. Stoneking, Genetic evidence concerning the origins of South and North Ossetians. Ann. Hum. Genet., 68 (2004) 588–599.

71 S. Sengupta, L.A. Zhivotovsky, R. King, S.Q. Mehdi, C.A. Edmonds, C.E. Chow, A.A. Lin, M. Mitra, S.K. Sil, A. Ramesh, M.V. Usha Rani, C.M. Thakur, L.L. Cavalli-Sforza, P.P. Majumder and P.A. Underhill, Polarity and temporality of high-resolution Y-chromosome distributions in India identify both indigenous and exogenous expansions and reveal minor genetic influence of Central Asian pastoralists. Am. J. Hum. Genet., 78 (2006) 202–221.

72 S. Rootsi, C. Magri, T. Kivisild, G. Benuzzi, H. Help, M. Bermisheva, I. Kutuev, L. Barac, M. Pericic, O. Balanovsky, A. Pshenichnov, D. Dion, M. Grobei, L.A. Zhivotovsky, V. Battaglia, A. Achilli, N. Al-Zahery, J. Parik, R. King, C. Cinnioglu, E. Khusnutdinova, P. Rudan, E. Balanovska, W. Scheffrahn, M. Simonescu, A. Brehm, R. Goncalves, A. Rosa, J.P. Moisan, A. Chaventre, V. Ferak, S. Furedi, P.J. Oefner, P. Shen, L. Beckman, I. Mikerezi, R. Terzic, D. Primorac, A. Cambon-Thomsen, A. Krumina, A. Torroni, P.A. Underhill, A.S. Santachiara-Benerecetti, R. Villems and O. Semino, Phylogeography of Y-chromosome haplogroup I reveals distinct domains of prehistoric gene flow in Europe. Am. J. Hum. Genet., 75 (2004) 128–137.

73 F. Di Giacomo, F. Luca, L.O. Popa, N. Akar, N. Anagnou, J. Banyko, R. Brdicka, G. Barbujani, F. Papola, G. Ciavarella, F. Cucci, L. Di Stasi, L. Gavrila, M.G. Kerimova, D. Kovatchev, A.I. Kozlov, A. Loutradis, V. Mandarino, C. Mammi, E.N. Michalodimitrakis, G. Paoli, K.I. Pappa, G. Pedicini, L. Terrenato, S. Tofanelli, P. Malaspina and A. Novelletto, Y chromosomal haplogroup J as a signature of the post-neolithic colonization of Europe. Hum. Genet., 115 (2004) 357–371.

74 C. Capelli, J.F. Wilson, M. Richards, M.P.H. Stumpf, F. Gratrix, S. Oppenheimer, P. Underhill, V.L. Pascali, T.M. Ko and D.B. Goldstein, A predominantly indigenous paternal heritage for the Austronesian-speaking peoples of insular Southeast Asia and Oceania. Am. J. Hum. Genet., 68 (2001) 432–443.

75 M.E. Hurles, J. Nicholson, E. Bosch, C. Renfrew, B.C. Sykes and M.A. Jobling, Y chromosomal evidence for the origins of Oceanic-speaking peoples. Genetics, 160 (2002) 289–303.

76 M.P. Cox and M.M. Lahr, Y-chromosome diversity is inversely associated with language affiliation in paired Austronesian- and Papuan-speaking communities from Solomon Islands. Am. J. Hum. Biol., 18 (2006) 35–50.

77 H. Shi, Y.L. Dong, B. Wen, C.J. Xiao, P.A. Underhill, P.D. Shen, R. Chakraborty, L. Jin and B. Su, Y-chromosome evidence of southern origin of the East Asian-specific haplogroup O3-M122. Am. J. Hum. Genet., 77 (2005) 408–419.

78 M.C. Bortolini, F.M. Salzano, M.G. Thomas, S. Stuart, S.P. Nasanen, C.H. Bau, M.H. Hutz, Z. Layrisse, M.L. Petzl-Erler, L.T. Tsuneto, K. Hill, A.M. Hurtado, D. Castro-de-Guerra, M.M. Torres, H. Groot, R. Michalski, P. Nymadawa, G. Bedoya, N. Bradman, D. Labuda and A. Ruiz-Linares, Y-chromosome evidence for differing ancient demographic histories in the Americas. Am. J. Hum. Genet., 73 (2003) 524–539.

79 S.L. Zegura, T.M. Karafet, L.A. Zhivotovsky and M.F. Hammer, High-resolution SNPs and microsatellite haplotypes point to a single, recent entry of Native American Y chromosomes into the Americas. Mol. Biol. Evol., 21 (2004) 164–175.

80 C. Capelli, N. Redhead, V. Romano, F. Cali, G. Lefranc, V. Delague, A. Megarbane, A.E. Felice, V.L. Pascali, P.I. Neophytou, Z. Poulli, A. Novelletto, P. Malaspina, L. Terrenato, A. Berebbi, M. Fellous, M.G. Thomas and D.B. Goldstein, Population structure in the Mediterranean basin: a Y chromosome perspective. Ann. Hum. Genet., 70 (2006) 207–225.

81 M. Brion, J.J. Sanchez, K. Balogh, C. Thacker, A. Blanco-Verea, C. Borsting, B. Stradmann-Bellinghausen, M. Bogus, D. Syndercombe-Court, P.M. Schneider, A. Carracedo and N. Morling, Introduction of a single nucleotide polymorphism-based "Major Y-chromosome haplogroup typing kit" suitable for predicting the geographical origin of male lineages. Electrophoresis, 26 (2005) 4411–4420.

82 V. Onofri, F. Alessandrini, C. Turchi, M. Pesaresi, L. Buscemi and A. Tagliabracci, Development of multiplex PCRs for evolutionary and forensic applications of 37 human Y chromosome SNPs. Forensic Sci. Int., 157 (2006) 23–35.

83 S. Paracchini, B. Arredi, R. Chalk and C. Tyler-Smith, Hierarchical high-throughput SNP genotyping of the human Y chromosome using MALDI-TOF mass spectrometry. Nucleic Acids Res., 30 (2002) e27.

84 M. Raitio, K. Lindroos, M. Laukkanen, T. Pastinen, P. Sistonen, A. Sajantila and A.C. Syvanen, Y-chromosomal SNPs in Finno-Ugric-speaking populations analyzed by minisequencing on microarrays. Genome Res., 11 (2001) 471–482.

85 M. Lareu, B. Sobrino, C. Phillips, M. Brión and A. Carracedo, Typing Y-chromosome single nucleotide polymorphisms with DNA microarray technology. Prog. Forensic Genet., 9 (2003) 21–25.

86 J.H. Wetton, K.W. Tsang and H. Khan, Inferring the population of origin of DNA evidence within the UK by allele-specific hybridization of Y-SNPs. Forensic Sci. Int., 152 (2005) 45–53.

87 A.J. Jeffreys, A. MacLeod, K. Tamaki, D.L. Neil and D.G. Monckton, Minisatellite repeat coding as a digital approach to DNA typing. Nature, 354 (1991) 204–209.

88 M.A. Jobling, N. Bouzekri and P.G. Taylor, Hypervariable digital DNA codes for human paternal lineages: MVR-PCR at the Y-specific minisatellite, MSY1 (DYF155S1). Hum. Mol. Genet., 7 (1998) 643–653.

89 N. Bouzekri, P.G. Taylor, M.F. Hammer and M.A. Jobling, Novel mutation processes in the evolution of a haploid minisatellite, MSY1: array homogenization without homogenization. Hum. Mol. Genet., 7 (1998) 655–659.

90 A.J. Jeffreys, K. Tamaki, A. MacLeod, D.G. Monckton, D.L. Neil and J.A. Armour, Complex gene conversion events in germline mutation at human minisatellites. Nat. Genet., 6 (1994) 136–145.

91 A.J. Jeffreys, M.J. Allen, J.A.L. Armour, A. Collick, Y. Dubrova, N. Fretwell, T. Guram, M. Jobling, C.A. May, D.L. Neil and R. Neumann, Mutation processes at human minisatellites. Electrophoresis, 16 (1995) 1577–1585.

92 M. Brion, R. Cao, A. Salas, M.V. Lareu and A. Carracedo, New method to measure minisatellite variant repeat variation in population genetic studies. Am. J. Hum. Biol., 14 (2002) 421–428.

93 W. Bao, S. Zhu, A. Pandya, T. Zerjal, J. Xu, Q. Shu, R. Du, H. Yang and C. Tyler-Smith, MSY2: a slowly evolving minisatellite on the human Y chromosome which provides a useful polymorphic marker in Chinese populations. Gene, 244 (2000) 29–33.

94 C. Tyler-Smith, R.J. Oakey, Z. Larin, R.B. Fisher, M. Crocker, N.A. Affara, M.A. Ferguson-Smith, M. Muenke, O. Zuffardi and M.A. Jobling, Localization of DNA sequences required for human centromere function through an analysis of rearranged Y chromosomes. Nat. Genet., 5 (1993) 368–375.

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

96 N. Mathias, M. Bayes and C. Tyler-Smith, Highly informative compound haplotypes for the human Y chromosome. Hum. Mol. Genet., 3 (1994) 115–123.

97 N.O. Bianchi, G. Bailliet, C.M. Bravi, R.F. Carnese, F. Rothhammer, V.L. Martinez-Marignac and S.D. Pena, Origin of Amerindian Y-chromosomes as inferred by the analysis of six polymorphic markers. Am. J. Phys. Anthropol., 102 (1997) 79–89.

98 F.R. Santos, A. Pandya, C. Tyler-Smith, S.D. Pena, M. Schanfield, W.R. Leonard, L. Osipova, M.H. Crawford and R.J. Mitchell, The central Siberian origin for native American Y chromosomes. Am. J. Hum. Genet., 64 (1999) 619–628.

99 F.R. Santos, S.D. Pena and C. Tyler-Smith, PCR haplotypes for the human Y chromosome based on alphoid satellite DNA variants and heteroduplex analysis. Gene, 165 (1995) 191–198.

100 L. Roewer and J.T. Epplen, Rapid and sensitive typing of forensic stains by PCR amplification of polymorphic simple repeat sequences in case work. Forensic Sci. Int., 53 (1992) 163–171.

101 P. de Knijff, M. Kayser, A. Caglià, D. Corach, N. Fretwell, C. Gehrig, G. Graziosi, F. Heidorn, S. Herrmann, B. Herzog, M. Hidding, K. Honda, M. Jobling, M. Krawczak, K. Leim, S. Meuser, E. Meyer, W. Oesterreich, A. Pandya, W. Parson, G. Penacino, A. Perez-Lezaun, A. Piccinini, M. Prinz, C. Schmitt, P.M. Schneider, R. Szibor, J. Teifel-Greding, G. Weichhold and L. Roewer, Chromosome Y microsatellites: population genetic and evolutionary aspects. Int. J. Legal Med., 110 (1997) 134–149.

102 P.S. White, O.L. Tatum, L.L. Deaven and J.L. Longmire, New, male-specific microsatellite markers from the human Y chromosome. Genomics, 57 (1999) 433–437.

103 Q. Ayub, A. Mohyuddin, R. Qamar, K. Mazhar, T. Zerjal and C. Tyler-Smith, Identification and characterisation of novel human Y-chromosomal microsatellites from sequence database information. Nucleic Acids Res., 28 (2000) e8.

104 R. Iida and K. Kishi, Identification, characterization and forensic application of novel Y-STRs. Legal Med., 7 (2005) 255–258 (Tokyo).

105 A.J. Redd, A.B. Agellon, V.A. Kearney, V.A. Contreras, T. Karafet, H. Park, P. de Knijff, J.M. Butler and M.F. Hammer, Forensic value of 14 novel STRs on the human Y chromosome. Forensic Sci. Int., 130 (2002) 97–111.

106 M.A. Jobling, V. Samara, A. Pandya, N. Fretwell, B. Bernasconi, R.J. Mitchell, T. Gerelsaikhan, B. Dashnyam, A. Sajantila, P.J. Salo, Y. Nakahori, C.M. Disteche, K. Thangaraj, L. Singh, M.H. Crawford and C. Tyler-Smith, Recurrent duplication and deletion polymorphisms on the long arm of the Y chromosome in normal males. Hum. Mol. Genet., 5 (1996) 1767–1775.

107 L.A. Zhivotovsky, P.A. Underhill, C. Cinnioglu, M. Kayser, B. Morar, T. Kivisild, R. Scozzari, F. Cruciani, G. Destro-Bisol, G. Spedini, G.K. Chambers, R.J. Herrera, K.K. Yong, D. Gresham, I. Tournev, M.W. Feldman and L. Kalaydjieva, The effective mutation rate at Y chromosome short tandem repeats, with application to human population-divergence time. Am. J. Hum. Genet., 74 (2004) 50–61.

108 L. Gusmão, A. Gonzalez-Neira, C. Alves, M. Lareu, S. Costa, A. Amorim and A. Carracedo, Chimpanzee homologous of human Y specific STRs: a comparative study and a proposal for nomenclature. Forensic Sci. Int., 126 (2002) 129–136.

109 L. Gusmão, J.M. Butler, A. Carracedo, P. Gill, M. Kayser, W.R. Mayr, N. Morling, M. Prinz, L. Roewer, C. Tyler-Smith and P.M. Schneider, DNA Commission of the International Society of Forensic Genetics (ISFG): an update of the recommendations on the use of Y-STRs in forensic analysis. Int. J. Legal Med., 120 (2006) 191–200.

110 E. Bosch, A.C. Lee, F. Calafell, E. Arroyo, P. Henneman, P. de Knijff and M.A. Jobling, High resolution Y chromosome typing: 19 STRs amplified in three multiplex reactions. Forensic Sci. Int., 125 (2002) 42–51.

111 J.M. Butler, R. Schoske, P.M. Vallone, M.C. Kline, A.J. Redd and M.F. Hammer, A novel multiplex for simultaneous amplification of 20 Y chromosome STR markers. Forensic Sci. Int., 129 (2002) 10–24.

112 J.G. Shewale, H. Nasir, E. Schneida, A.M. Gross, B. Budowle and S.K. Sinha, Y-chromosome STR system, Y-PLEX 12, for forensic casework: development and validation. J. Forensic Sci., 49 (2004) 1278–1290.

113 A. Hall and J. Ballantyne, The development of an 18-locus Y-STR system for forensic casework. Anal. Bioanal. Chem., 376 (2003) 1234–1246.

114 S. Beleza, C. Alves, A. González-Neira, M. Lareu, A. Amorim, A. Carracedo and L. Gusmão, Extending STR markers in Y chromosome haplotypes. Int. J. Legal Med., 117 (2003) 27–33.

115 B.E. Krenke, L. Viculis, M.L. Richard, M. Prinz, S.C. Milne, C. Ladd, A.M. Gross, T. Gornall, J.R. Frappier, A.J. Eisenberg, C. Barna, X.G. Aranda, M.S. Adamowicz and B. Budowle, Validation of male-specific, 12-locus fluorescent short tandem repeat (STR) multiplex. Forensic Sci. Int., 151 (2005) 111–124.

116 J.J. Mulero, C.W. Chang, L.M. Calandro, R.L. Green, Y. Li, C.L. Johnson and L.K. Hennessy, Development and validation of the AmpFlSTR Yfiler PCR amplification kit: a male specific, single amplification 17 Y-STR multiplex system. J. Forensic Sci., 51 (2006) 64–75.

117 J.L. Weber and C. Wong, Mutation of human short tandem repeats. Hum. Mol. Genet., 2 (1993) 1123–1128.

118 R. Scozzari, F. Cruciani, P. Malaspina, P. Santolamazza, B.M. Ciminelli, A. Torroni, D. Modiano, D.C. Wallace, K.K. Kidd, A. Olckers, P. Moral, L. Terrenato, N. Akar, R. Qamar, A. Mansoor, S.Q. Mehdi, G. Meloni, G. Vona, D.E. Cole, W. Cai and A. Novelletto, Differential structuring of human populations for homologous X and Y microsatellite loci. Am. J. Hum. Genet., 61 (1997) 719–733.

119 L. Henke and J. Henke, Mutation rate in human microsatellites. Am. J. Hum. Genet., 64 (1999) 1473–1474.

120 A. Sajantila, M. Lukka and A.C. Syvanen, Experimentally observed germline mutations at human micro- and minisatellite loci. Eur. J. Hum. Genet., 7 (1999) 263–266.

121 A. Di Rienzo, A.C. Peterson, J.C. Garza, A.M. Valdes, M. Slatkin and N.B. Freimer, Mutational processes of simple-sequence repeat loci in human populations. Proc. Natl. Acad. Sci., 91 (1994) 3166–3170.

122 L.A. Zhivotovsky and M.W. Feldman, Microsatellite variability and genetic distances. Proc. Natl. Acad. Sci., 92 (1995) 11549–11552.

123 A. Di Rienzo, P. Donnelly, C. Toomajian, B. Sisk, A. Hill, M.L. Petzl-Erler, G.K. Haines and D.H. Barch, Heterogeneity of microsatellite mutations within and between loci, and implications for human demographic histories. Genetics, 148 (1998) 1269–1284.

124 G. Cooper, N.J. Burroughs, D.A. Rand, D.C. Rubinsztein and W. Amos, Markov chain Monte Carlo analysis of human Y-chromosome microsatellites provides evidence of biased mutation. Proc. Natl. Acad. Sci., 96 (1999) 11916–11921.

125 R. Kurihara, T. Yamamoto, R. Uchihi, S.L. Li, T. Yoshimoto, H. Ohtaki, K. Kamiyama and Y. Katsumata, Mutations in 14 Y-STR loci among Japanese father–son haplotypes. Int. J. Legal Med., 118 (2004) 125–131.

126 X. Xu, M. Peng and Z. Fang, The direction of microsatellite mutations is dependent upon allele length. Nat. Genet., 24 (2000) 396–399.

127 C.R. Primmer, N. Saino, A.P. Moller and H. Ellegren, Directional evolution in germline microsatellite mutations. Nat. Genet., 13 (1996) 391–393.

128 W. Amos, S.J. Sawcer, R.W. Feakes and D.C. Rubinsztein, Microsatellites show mutational bias and heterozygote instability. Nat. Genet., 13 (1996) 390–391.

129 U. Schmidt, N. Meier and S. Lutz, Y-chromosomal STR haplotypes in a population sample from southwest Germany (Freiburg area). Int. J. Legal Med., 117 (2003) 211–217.

130 J.M. Butler, A.E. Decker, P.M. Vallone and M.C. Kline, Allele frequencies for 27 Y-STR loci with U.S. Caucasian, African American, and Hispanic samples. Forensic Sci. Int., 156 (2006) 250–260.

131 H.L. Dai, X.D. Wang, Y.B. Li, J. Wu, J. Zhang, H.J. Zhang, J.G. Dong and Y.P. Hou, Characterization and haplotype analysis of 10 novel Y-STR loci in Chinese Han population. Forensic Sci. Int., 145 (2004) 47–55.

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

133 N.O. Bianchi, C.I. Catanesi, G. Bailliet, V.L. Martinez-Marignac, C.M. Bravi, L.B. Vidal-Rioja, R.J. Herrera and J.S. López-Camelo, Characterization of ancestral and derived Y-chromosome haplotypes of New World native populations. Am. J. Hum. Genet., 63 (1998) 1862–1871.

134 A.C.S. Góes, E.F. Carvalho, I. Gomes, D.A. da Silva, E.H.F. Gil, A. Amorim and L. Gusmão, Population and mutation analysis of 17 Y-STR loci from Rio de Janeiro (Brazil). Int. J. Legal Med., 119 (2005) 70–76.

135 B. Budowle, M. Adamowicz, X. Aranda, C. Barna, R. Chakraborty, D. Cheswick, B. Dafoe, A. Eisenberg, R. Frappier, A. Gross, L.C. Ladd, H. Lee, S.C. Milne, C. Meyers, M. Prinz, M.L. Richard, G. Saldanha, A.A. Tierney, L. Viculis and B.E. Krenke, Twelve short tandem repeat loci Y chromosome haplotypes: genetic analysis on populations residing in North America. Forensic Sci. Int., 150 (2005) 1–15.

136 D.J. Ballard, C. Phillips, C.R. Thacker, C. Robson, A.P. Revoir and D. Syndercombe Court, A study of mutation rates and the characterisation of intermediate, null and duplicated alleles for 13 Y chromosome STRs. Forensic Sci. Int., 155 (2005) 65–70.

137 P.M. Domingues, L. Gusmão, D.A. da Silva, A. Amorim, R.W. Pereira and E.F. de Carvalho, Sub-Saharan Africa descendents in Rio de Janeiro (Brazil): population and mutational data for 12 Y-STR loci. Int. J. Legal Med., 121 (2007) 238–241.

138 A. Mohyuddin, Q. Ayub, P.A. Underhill, C. Tyler-Smith and S.Q. Mehdi, Detection of novel Y SNPs provides further insights into Y chromosomal variation in Pakistan. J. Hum. Genet., 51 (2006) 375–378.

139 P. Shen, T. Lavi, T. Kivisild, V. Chou, D. Sengun, D. Gefel, I. Shpirer, E. Woolf, J. Hillel, M.W. Feldman and P.J. Oefner, Reconstruction of patrilineages and matrilineages of Samaritans and other Israeli populations from Y-chromosome and mitochondrial DNA sequence variation. Hum. Mutat., 24 (2004) 248–260.

140 O. Semino, C. Magri, G. Benuzzi, A.A. Lin, N. Al-Zahery, V. Battaglia, L. Maccioni, C. Triantaphyllidis, P. Shen, P.J. Oefner, L.A. Zhivotovsky, R. King, A. Torroni, L.L. Cavalli-Sforza, P.A. Underhill and A.S. Santachiara-Benerecett, 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. Am. J. Hum. Genet., 74 (2004) 1023–1034.

141 A.O. Karlsson, T. Wallerström, A. Götherström and G. Holmlund, Y-chromosome diversity in Sweden: a long-time perspective. Eur. J. Hum. Genet., 14 (2006) 963–970.

142 L. Scheinfeldt, F. Friedlaender, J. Friedlaender, K. Latham, G. Koki, T. Karafet, M. Hammer and J. Lorenz, Unexpected NRY chromosome variation in Northern Island Melanesia. Mol. Biol. Evol., 23 (2006) 1628–1641.

143 M. Kayser, S. Brauer, R. Cordaux, A. Casto, O. Lao, L.A. Zhivotovsky, C. Moyse-Faurie, R.B. Rutledge, W. Schiefenhoevel, D. Gil, A.A. Lin, P.A. Underhill, P.J. Oefner, R.J. Trent and M. Stoneking, Melanesian and Asian origins of Polynesians: mtDNA and Y chromosome gradients across the Pacific. Mol. Biol. Evol., 23 (2006) 2234–2244.

144 M. Regueiro, A.M. Cadenas, T. Gayden, P.A. Underhill and T.J. Herrera, Iran: tricontinental nexus for Y-chromosome driven migration. Hum. Hered., 61 (2006) 132–143.

145 W. Deng, B. Shi, X. He, Z. Zhang, J. Xu, B. Li, J. Yang, L. Ling, C. Dai, B. Qiang, Y. Shen and R.E. Chen, Evolution and migration history of the Chinese population inferred from Chinese Y-chromosome evidence. J. Hum. Genet., 49 (2004) 339–348.