Worldwide diversity of the Y-chromosome tetra-local microsatellite *DYS464*

F.S.G. Kehdy and S.D.J. Pena

Departamento de Bioquímica e Imunologia, Universidade Federal de Minas Gerais, Belo Horizonte, MG, Brasil

Corresponding author: S.D.J. Pena
E-mail: spena@dcc.ufmg.br

ABSTRACT. Of all DNA markers on the human Y-chromosome, the tetra-local Y-linked microsatellite *DYS464* is the most polymorphic. We genotyped *DYS464* in 677 male samples collected worldwide, maintained in the HGDP-CEPH Human Genome Diversity Cell Line Panel. Fourteen different alleles were found, with allele lengths varying from 9 to 23 repeats. One hundred and seventy-five different genotypes were detected, of which 90 appeared to be continent-specific. The region with the highest percentage of unique genotypes was Africa. Genotype diversity was 0.98 for Europe, 0.97 for Central and East Asia, 0.95 for Africa, 0.94 for Oceania, 0.92 for the Middle East, and 0.90 for the Americas. A hierarchical analysis of molecular variance showed low levels of worldwide genetic structure; 88.42% of the genetic variance was found within populations, 9.62% between populations within regions and 1.96% between regions. Since the four *DYS464* repeats are identical, one cannot assign each peak in the electropherogram to a specific locus. Thus, the same genotype may correspond to several haplotypes, with different permutations of alleles. Consequently, genotypes are degenerate, which limits phylogeographical analyses. Yet, because of its high variability, *DYS464* still constitutes an informative tool for population and evolutionary studies.

Key words: Y-chromosome; Worldwide populations; Microsatellites; DNA; Population genetics; *DYS464*

INTRODUCTION

Y-linked loci are haploid and paternally inherited and with the exception of genes in the pseudo-autosomal regions, there is no recombination (reviewed by Jobling and Tyler-Smith, 2003). Thus, Y-chromosomal markers are transmitted together as haplotypes. Therefore, each male individual has the same Y-chromosome haplotype as his father, brothers, paternal grandfather, paternal uncles, etc., thus establishing a patrilineage.

These characteristics render Y-linked polymorphisms extremely useful as genetic tools for paternity testing (Santos et al., 2003), forensic medicine (Gusmão et al., 2006), ancestry studies (Carvalho-Silva et al., 2001, 2006), and human evolutionary genetics (Chiaroni et al., 2009).

More than 300 microsatellite loci have been described on the Y-chromosome (Hanson and Ballantyne, 2006). Among them, the most polymorphic is the tetralocal microsatellite *DYS464* (Redd et al., 2002; Berger et al., 2003).

DYS464 consists of four identical copies (a, b, c, and d) of the tetranucleotide repeat $(CCTT)_n$ on the palindromic AZF region (Kuroda-Kawaguchi et al., 2001) in band Yq11.223 of the Y-chromosome (Redd et al., 2002) (Figure 1). It is believed that the tetra-local structure evolved from two sequential duplications of an original monomer, promoted by mitotic recombination events (Kuroda-Kawaguchi et al., 2001; Skaletsky et al., 2003; Rozen et al., 2003).
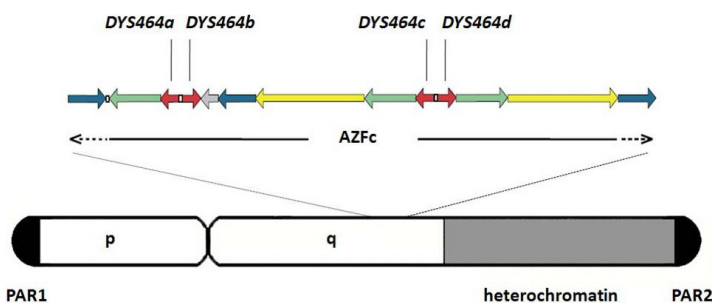

Figure 1. Diagram of the human Y-chromosome, indicating the AZFc (azoospermia factor c) region with its highly palindromic structure, which harbors the four loci of the tetra-local microsatellite *DYS464*. On the Y-chromosome are shown the two pseudo-autosomal regions (PAR1 and PAR2), the short (p) and long (q) arms and the large heterochromatic region. The figure was drawn based on data from Kuroda-Kawaguchi et al. (2001).

DYS464 genotypes are established in a straightforward fashion when four peaks of identical area are present in electropherograms. However, since one cannot assign each peak to a specific locus, the same genotype (combination of alleles - order is ignored) may correspond to different haplotypes (permutations of alleles - order is significant). In this fashion, genotypes are degenerate. Let us take, for instance, the situation in which we identify four peaks, corresponding to four different alleles, say, 13, 14, 15, and 16. This is scored as genotype 13-14-15-16, which could in fact be any of 12 different haplotypes that contain these four alleles in different permutations, i.e., 13-14-16-15, 13-16-15-14, etc. Moreover, when less than four peaks are seen, the relative peak areas have to be used to estimate the number of copies of a particular allele (Butler and Schoske, 2005).

Although *DYS464* has been studied in some human populations (Redd et al., 2002; Berger et al., 2003; Butler and Schoske, 2005), a more complete study of this marker in worldwide human population has not yet been reported. In this article, we describe our results on the typing of *DYS464* in all male samples from HGDP-CEPH Human Genome

Diversity Cell Line Panel (Cann et al., 2002) and discuss its diversity and phylogeography in different human populations.

MATERIAL AND METHODS

Samples

DNA samples from all 677 unrelated males from HGDP-CEPH Human Genome Diversity Cell Line Panel (Cann et al., 2002; http://www.cephb.fr/fr/hgdp/diversity.php) were analyzed in this study. The individuals were sampled across all five continents and assigned to 52 different populations from seven regional groups (Africa, Europe, Middle East, Central/South Asia, East Asia, Oceania, and Americas).

DNA typing

For polymerase chain reaction (PCR) amplification of *DYS464*, the following primers were used: *DYS464*-F: 5'-TTACGAGCTTTGGGCTATG-3' with a tail of the M13-40 17-oligonucleotide GTTTTCCCAGTCACGAC and *DYS464*-R: 5'-CCTGGGTAACAGAGAGACTCTT-3'.

PCR was performed using 2 U *Taq* DNA polymerase (Phoneutria, Belo Horizonte, Brazil) and 200 μM dNTPs in 10 mM Tris-HCl buffer, pH 8.4, with 50 mM KCl, 1.5 mM MgCl$_2$, 0.1% Triton. For fragment separation, we used a MegaBACE 1000 DNA sequencer (GE Healthcare) followed by analyses using Genetic Profiler (version 2.2) and Fragment Profiler (version 1.2) programs (GE Healthcare).

Electropherogram profiles for *DYS464* consisted of one, two, three, or four peaks, which were converted to genotypes based on the estimated repeat number and relative areas of the four alleles present in each sample (Butler and Schoske, 2005). We always assumed that four loci were present, unless there was a compelling reason to believe differently.

Data analysis

The genetic structure of populations and basic population parameters including analyses of molecular variance (AMOVA), genotype diversity, genotype frequency, and genotype sharing were calculated using the Arlequin 2.0 software (Schneider et al., 2000) with 10,000 steps in the Markov chain.

RESULTS

We genotyped *DYS464* in all 677 male samples of the HGDP-CEPH panel, representing 52 different populations from seven regional groups worldwide (Sub-Saharan Africa, Europe, Middle East, Central Asia, East Asia, Oceania, and Americas). Fourteen different alleles were found, having lengths varying from 9 to 23 repeats. However, allele 21 was not found in any sample.

A total of 175 genotypes were identified: 49 in Africa, 53 in Europe, 29 in the Middle East, 72 in Central Asia, 83 in East Asia, 13 in Oceania, and 16 in Americas (Tables 1 and 2). It is important to remember that each genotype could correspond to many different haplotypes containing the same alleles in different orders.

Table 1. Frequency of 175 different *DYS464* genotypes in 677 male samples from the HGDP-CEPH Human Genome Diversity Cell Line Panel belonging to seven geographical groups (Africa, Middle East, Central/South Asia, East Asia, Oceania, Europe, and Americas).

Genotype	Africa (N = 98)	Middle East (N = 79)	Central Asia (N = 175)	East Asia (N = 173)	Oceania (N = 21)	Europe (N = 88)	Americas (N = 43)
9-12-15-16				0.005			
10-11-14-16	0.010						
11-11-11-15			0.005				
11-11-11-16				0.017			
11-11-12-16			0.040	0.005			
11-11-14-15			0.022				
11-12-12-12				0.005			
11-12-12-15				0.005			
11-12-12-16				0.005			
11-12-12-17				0.005			
11-12-12-18				0.005			
11-12-13-15			0.011				
11-12-13-16				0.005			
11-12-14-15				0.005			
11-12-16-18			0.005				
11-13-13-14			0.022				
11-13-15-16				0.005			
11-14-14-14						0.011	
11-14-14-15			0.005			0.045	
11-14-14-16						0.011	
11-14-15-15			0.005	0.005		0.011	
11-14-15-16			0.028			0.011	
11-14-15-18						0.011	
11-14-16-18		0.012					
11-14-17-17	0.010						
11-15-15-16			0.005	0.005		0.011	
11-15-16-17						0.011	
11-16-16-16			0.005				
11-16-17-17	0.010						
12-12-12-12				0.005	0.047		
12-12-12-13				0.011			
12-12-12-15			0.005			0.011	
12-12-13-13		0.025	0.005	0.011			
12-12-13-14				0.005			
12-12-14-14		0.012					
12-12-14-15			0.022	0.011			
12-12-14-16			0.028				
12-12-15-15				0.005			
12-12-15-18	0.010						
12-12-16-16				0.017			
12-12-16-17				0.005			
12-12-17-17					0.142		
12-13-13-13				0.017			
12-13-13-14		0.050				0.011	0.046
12-13-13-15		0.012	0.011	0.005			
12-13-13-16			0.005	0.005			
12-13-14-14		0.113		0.011		0.011	
12-13-14-15			0.017	0.040			
12-13-14-16			0.017	0.005		0.011	
12-13-14-17		0.012		0.005			
12-13-15-16	0.020	0.012	0.011	0.023		0.011	
12-13-15-17				0.005			
12-13-15-18							0.069
12-13-16-16			0.005	0.011		0.011	
12-14-14-14			0.005				

Table 1. Continued.

Genotype	Africa (N = 98)	Middle East (N = 79)	Central Asia (N = 175)	East Asia (N = 173)	Oceania (N = 21)	Europe (N = 88)	Americas (N = 43)
12-14-14-15			0.011			0.011	
12-14-14-16			0.011	0.005		0.011	
12-14-15-15		0.025	0.022	0.005		0.011	
12-14-15-16	0.010		0.034	0.017		0.034	
12-14-15-17	0.010	0.037		0.005			
12-14-16-16				0.005		0.011	
12-14-16-17			0.011			0.022	
12-14-16-18		0.164					
12-14-17-17			0.011				
12-15-15-15			0.011				
12-15-15-16		0.025	0.137	0.005		0.034	
12-15-15-17			0.022	0.011			
12-15-15-18	0.010		0.034				
12-15-16-16	0.010		0.017			0.034	
12-15-16-17	0.010		0.005				
12-15-16-18	0.040				0.047		
12-16-16-16			0.011			0.011	
12-16-16-17					0.142		
12-16-17-17		0.012					
12-17-17-17					0.142		
13-13-13-13				0.011			
13-13-13-14				0.034		0.034	
13-13-13-15				0.017			
13-13-13-16				0.005			
13-13-13-17				0.005			
13-13-14-14				0.005		0.022	
13-13-14-15			0.017	0.005			
13-13-14-16				0.011			
13-13-14-17			0.005	0.005	0.047		
13-13-15-15	0.010		0.005	0.028			
13-13-15-16			0.011				
13-13-15-17			0.011	0.023	0.095		
13-13-16-16				0.005			
13-13-16-17	0.010						
13-13-16-18	0.010						
13-13-17-17			0.005				
13-13-17-18				0.011			
13-13-18-18	0.010						
13-14-14-14	0.010	0.012	0.005	0.011			
13-14-14-15				0.005			0.046
13-14-14-16	0.030			0.017			
13-14-14-17		0.012		0.005			
13-14-14-20	0.010						
13-14-15-15			0.005	0.017			
13-14-15-16			0.022	0.028	0.047		
13-14-15-17				0.005			
13-14-15-18	0.010						0.023
13-14-15-19	0.020						
13-14-16-16			0.005	0.005		0.011	
13-14-16-17		0.063		0.017	0.047		0.023
13-14-16-18			0.011				
13-14-17-17							0.023
13-14-17-18			0.005				
13-15-15-15			0.005	0.005			
13-15-15-16			0.011			0.011	
13-15-15-17							0.139
13-15-15-18	0.010						0.139
13-15-16-16	0.020		0.011				
13-15-16-17	0.040	0.012	0.005	0.011			0.046

Continued on next page

Table 1. Continued.

Genotype	Africa (N = 98)	Middle East (N = 79)	Central Asia (N = 175)	East Asia (N = 173)	Oceania (N = 21)	Europe (N = 88)	Americas (N = 43)
13-15-16-18	0.061	0.012				0.011	
13-15-17-18			0.005				
13-15-17-20	0.010						
13-16-16-16	0.030			0.034			
13-16-16-17	0.040					0.011	
13-16-16-18	0.204						
13-16-16-19	0.020						
13-16-17-17	0.010						
13-16-17-18	0.010						
14-14-14-14	0.010						
14-14-14-15						0.011	
14-14-14-16			0.022				0.023
14-14-14-17			0.005				
14-14-15-15			0.011	0.115		0.022	
14-14-15-16	0.010		0.011			0.011	
14-14-15-17	0.030						0.023
14-14-15-18							0.023
14-14-16-16	0.010	0.037	0.005	0.011	0.095	0.022	
14-14-16-17		0.012		0.005			
14-14-16-18				0.005			
14-14-16-19				0.005			
14-14-17-17		0.012					
14-15-15-15				0.005	0.047		
14-15-15-16			0.005	0.011			0.232
14-15-15-17	0.020			0.005		0.011	
14-15-15-18				0.005			0.069
14-15-16-16			0.011			0.011	
14-15-16-17			0.011	0.011		0.034	0.046
14-15-16-18	0.020		0.005	0.005			
14-15-17-17	0.020						
14-15-17-18						0.011	
14-16-16-16	0.010	0.037	0.005				
14-16-16-17		0.164				0.011	
14-16-16-18	0.010	0.012					
14-16-17-17	0.010						
14-16-17-18	0.010		0.005			0.011	
14-17-17-18	0.010						
15-15-15-15			0.005	0.046		0.022	
15-15-15-16			0.011	0.005			
15-15-15-17			0.005			0.022	
15-15-15-18			0.005				
15-15-16-16			0.022	0.011		0.022	
15-15-16-17			0.011	0.005		0.045	
15-15-16-18		0.012				0.011	
15-15-16-19			0.005				
15-15-17-17	0.010					0.056	0.023
15-15-17-18						0.079	
15-15-17-19						0.011	
15-16-16-16			0.022		0.047	0.011	
15-16-16-17	0.030	0.012	0.011			0.022	
15-16-16-18		0.025					
15-16-17-17		0.025		0.005		0.022	
15-16-17-18	0.010			0.005			
15-17-17-17	0.010					0.011	
16-16-16-16					0.047	0.011	
16-16-16-17		0.012					
16-16-17-17				0.011			
16-16-18-18				0.017		0.011	
16-17-17-17				0.005			
17-18-18-18	0.020						
20-20-22-23	0.010						

Table 2. Data on the *DYS464* genotypes encountered in this study.

Region	Number of individuals	Number of genotypes	Number of continent-specific genotypes	Genotype diversity (means ± SE)
Africa	98	49	20 (40%)	0.95 ± 0.01
Middle East	79	29	7 (24%)	0.92 ± 0.01
Central Asia	175	72	19 (26.4%)	0.97 ± 0.01
East Asia	173	83	29 (34.9%)	0.97 ± 0.01
Oceania	21	13	3 (23%)	0.94 ± 0.03
Europe	88	53	8 (15%)	0.98 ± 0,01
Americas	43	16	4 (25%)	0.90 ± 0.03
Total	677	175	90 (51.4%)	0.99

Even so, there was considerable geographical differentiation in genotype distribution. Among the 175 genotypes scored, 90 were observed in a single continent. The region with the highest percentage of unique genotypes was Africa (20 in a total of 49 genotypes, i.e., 40.8%).

The most frequent genotypes seen in each continent were the following:

- In Africa, 13-16-16-18 (20.4%) - this genotype seemed widespread in Sub-Saharan Africa, seen in significant frequencies among the Biaka (43.3%), Bantu (18.2%), Yoruba (15.4%), Mandenka (12.5%), and Mbuti (7.7%), but not the San. This genotype was not seen outside Africa.
- In Europe, 15-15-17-18 (8.0%) - this genotype was seen among the Basques of France (31.3%), Orcadian (14.3%) and Bergamo of Italy (12.5%). It was not seen among the French, Sardinian, Tuscan, Adygei of Caucasus, and Russian peoples. This genotype was not seen outside Europe.
- In East Asia, 14-14-15-15 (11.6%) - this genotype was seen among the Yakut of Siberia (88.9%) and the Miaozu (28.6%), Naxi (12.5%) and Han (4.2%) of China. Outside East Asia it was seen among the Russian (12.5%) and the Brahui in Pakistan (8.0%).
- In the Americas, 14-15-15-16 (23.2%) - this genotype was seen among the Pimas (64.3%) and Mayans (33.3%), but not in Colombians (Piapoco and Curripaco), Karitiana and Surui. Outside the Americas, this genotype was seen in the Daur of Mongolia (14.3%), Basques of France (6.3%), Hazara of Pakistan (4.2%), and Han of China (4.2%).
- In Central Asia, 12-15-15-16 (13.7%) - the genotype was seen among the Bahui (16.0%), Baloshi (16.0%), Makrani (15.0%), Pathan (30.0%), and Burusho (23.8%) of Pakistan, but not in the Hazara. Outside Central Asia, it was seen in Bedouin of Israel (7.1%), Uygur of China (12.5%), Sardinians of Italy (6.3%), Adygei of the Caucasus (14.3%), and Russians (6.3%).
- In the Middle East, there was a tie for genotypes 12-14-16-18 and 14-16-16-17 (16.4% each). The former was seen exclusively among the Bedouin of Israel (46.2%), while the latter was seen exclusively among the Mozabite of Algeria (55.0%) and Druze of Israel (14.3%).
- In Oceania, there was a tie for three genotypes: 12-12-17-17, 12-16-16-17 and 12-17-17-17 (14.2% each). The first two were seen in both Papuans (7.7 and 15.4%, respectively) and Melanesians (25.0 and 12.5%, respectively), while the last was observed only among Melanesians (37.5%).

The genotype diversity was 0.98 for Europe, 0.97 for Central and East Asia, 0.95 for Africa, 0.94 for Oceania, 0.92 for the Middle East, and 0.90 for the Americas (Table 2).

The genotypes were submitted to a hierarchical AMOVA using the Arlequin program (Schneider et al., 2000), and the results of this analysis are displayed in Table 3. Our analysis showed low levels of worldwide genetic structure: 88.42% of the genetic variance was found

within populations, 9.62% between populations within regions and 1.96% between regions. Focusing on each region, the within-population component was responsible for more than 87% of the genetic variance, except for the Middle East people and the Americas, which exhibited, lower within-population variance and 21.78 and 33.20% of the variance between populations within continents, respectively.

Table 3. Analysis of molecular variance (AMOVA) for *DYS464*.

Region	Number of regions	Number of populations	Variance components (%)		
			Within populations	Between populations	Between regions
World	1	52	88.71	11.29	-
World	7	52	88.42	9.62	1.96
Africa	1	7	94.55	5.45	-
Middle East	1	4	78.22	21.78	-
Central Asia	1	8	95.91	4.09	-
East Asia	1	18	87.99	12.01	-
Oceania	1	2	94.90	5.10	-
Europe	1	8	97.18	2.82	-
Americas	1	5	66.80	33.20	-

DISCUSSION

DYS464 is indeed highly polymorphic, but it has some characteristics that render the establishment of its evolutionary history difficult. First of all, the four copies are identical and one cannot assign each peak in the electropherogram to a specific locus. Thus, the same genotype may correspond to several haplotypes with different permutation of alleles. In this fashion, genotypes are degenerate. Second, *DYS464* has a high mutation rate, estimated at 2.86×10^{-2}, approximately 10 times higher than the average mutation rate for Y-chromosome microsatellites (Berger et al., 2003). Third, the tetralocal palindromic structure of *DYS464* renders it susceptible to allelic conversion mediated by intrachromosomal mitotic recombination events (Rozen et al., 2003).

Considering these complexities, it is no surprise that the phylogeography of *DYS464* may occasionally appear strange. For instance, the most common genotype in Central Asia, 12-15-15-16, was also seen in the Bedouins of Israel (7.1%), Uygurs of China (12.5%), Sardinians of Italy (6.3%), Adygei of the Caucasus (14.3%), and Russians (6.3%) and in no other populations. This "saltatory" distribution does not make evolutionary sense, unless we consider the possibility that Bedouins, Uygurs and Sardinians could have different haplotypes of the same four loci in different order and/or that one or more conversion events occurred leading to loss of heterozygosity (note the presence of two alleles 15).

On the other hand, many populations exhibit high frequencies of specific genotypes that may prove to be useful in evolutionary studies. For instance, the fact that genotype 14-14-15-15 was seen in 88.9% of the Yakut of Siberia, suggests that the male component of this population could have experienced a population bottleneck or a founder effect. Research done by other groups supports this notion (Pakendorf et al., 2002; Khar'kov et al., 2008).

Africa had the highest percentage of unique genotypes. This is compatible with the view that modern man emerged in Africa and migrated from that continent to populate all other areas of Earth (reviewed in Pena, 2007) and that the movement of modern humanity out-of-Africa was associated with a population size bottleneck and reduction of variability (Yu et al., 2002; Torroni et al., 2006).

However, contrary to most other Y-chromosomal loci studied (Jobling and Tyler-Smith, 2003), *DYS464* did not display its largest genotypic variability in Africa. One probably cannot and should not assign too much significance to this observation, especially if one considers the genotypic degeneracy of *DYS464*, its high mutation rate and also the fact the African samples present in the HGDP-CEPH Human Genome Diversity Cell Line Panel may only be a small portion of the total African variability.

A hierarchical AMOVA was performed, and the results revealed low levels of worldwide genetic structure - the within-population component was responsible for 88.42% of the genetic variability, with 9.62% between populations within regions and 1.96% between regions. These relative levels of within-population variance are higher than those observed in other worldwide surveys with Y-chromosome markers: 76.8% by Kayser et al. (2001) and 66.5% by Wilder et al. (2004). We believe that this can be attributed to the high mutation rate of *DYS464*.

It has been suggested that the differentiation of microsatellite polymorphisms in human populations is driven more by genetic drift than by mutation pressure (Perez-Lezaun et al., 1997). Since the effective population size of the Y-chromosome is around one-quarter of that in autosomes, in theory, genetic drift should have a deep impact on *DYS464* variation. On the other hand, the high mutation rate of *DYS464* will counteract the effects of drift, producing homoplasy and decreasing the between-population variance.

In conclusion, our results demonstrate that the *DYS464* microsatellite on the human Y-chromosome shows indeed very high variability in geographically widespread human populations, with genotype diversities above 0.90 in all continents studied. Even though it only provides the tale of a single "gene" and despite that its tetra-local structure often does not permit detailed phylogeographical inferences, *DYS464* still appears to be an informative tool for population and evolutionary studies.

ACKNOWLEDGMENTS

REFERENCES

Berger B, Niederstatter H, Brandstatter A and Parson W (2003). Molecular characterization and Austrian Caucasian population data of the multi-copy Y-chromosomal STR DYS464. *Forensic Sci. Int.* 137: 221-230.

Butler JM and Schoske R (2005). U.S. population data for the multi-copy Y-STR locus DYS464. *J. Forensic Sci.* 50: 975-977.

Cann HM, de Toma C, Cazes L, Legrand MF, et al. (2002). A human genome diversity cell line panel. *Science* 296: 261-262.

Carvalho-Silva DR, Santos FR, Rocha J and Pena SD (2001). The phylogeography of Brazilian Y-chromosome lineages. *Am. J. Hum. Genet.* 68: 281-286.

Carvalho-Silva DR, Tarazona-Santos E, Rocha J, Pena SD, et al. (2006). Y-chromosome diversity in Brazilians: switching perspectives from slow to fast evolving markers. *Genetica* 126: 251-260.

Chiaroni J, Underhill PA and Cavalli-Sforza LL (2009). Y-chromosome diversity, human expansion, drift, and cultural evolution. *Proc. Natl. Acad. Sci. U. S. A.* 106: 20174-20179.

Gusmão L, Butler JM, Carracedo A, Gill P, et al. (2006). 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: 191-200.

Hanson EK and Ballantyne J (2006). Comprehensive annotated STR physical map of the human Y-chromosome: Forensic implications. *Leg. Med.* 8: 110-120.

Jobling MA and Tyler-Smith C (2003). The human Y-chromosome: an evolutionary marker comes of age. *Nat. Rev. Genet.* 4: 598-612.

Kayser M, Krawczak M, Excoffier L, Dieltjes P, et al. (2001). An extensive analysis of Y-chromosomal microsatellite haplotypes in globally dispersed human populations. *Am. J. Hum. Genet.* 68: 990-1018.

Khar'kov VN, Stepanov VA, Medvedev OF, Spiridonova MG, et al. (2008). The origin of Yakuts: analysis of Y-chromosome haplotypes. *Mol. Biol.* 42: 226-237.

Kuroda-Kawaguchi T, Skaletsky H, Brown LG, Minx PJ, et al. (2001). The AZFc region of the Y-chromosome features massive palindromes and uniform recurrent deletions in infertile men. *Nat. Genet.* 29: 279-286.

Pakendorf B, Morar B, Tarskaia LA, Kayser M, et al. (2002). Y-chromosomal evidence for a strong reduction in male population size of Yakuts. *Hum. Genet.* 110: 198-200.

Pena SDJ (2007). The Evolution and Structure of Human Genetic Diversity. In: Pharmacogenomics in Admixed Populations (Suarez-Kurtz G, ed.). Landes Bioscience, Austin, 1-11. Available at [http://www.eurekah.com/chapter/3190]. Accessed…

Perez-Lezaun A, Calafell F, Mateu E, Comas D, et al. (1997). Microsatellite variation and the differentiation of modern humans. *Hum. Genet.* 99: 1-7.

Redd AJ, Agellon AB, Kearney VA, Contreras VA, et al. (2002). Forensic value of 14 novel STRs on the human Y-chromosome. *Forensic Sci. Int.* 130: 97-111.

Rozen S, Skaletsky H, Marszalek JD, Minx PJ, et al. (2003). Abundant gene conversion between arms of palindromes in human and ape Y-chromosomes. *Nature* 423: 873-876.

Santos FR, Epplen JT and Pena SDJ (2003). Testing Deficiency Paternity Cases With an Y-linked Tetranucleotide Repeat Polymorphism. In: DNA Fingerprinting: State of the Science (Pena SDJ, Chakraborty R, Epplen JT and Jeffreys AJ, eds.). Birkhäuser Verlag, Basel, 261-266.

Schneider S, Roessli D and Excoffier L (2000). Arlequin Ver 2.000: A Software for Population Genetics Data Analysis. Genetics and Biometry Laboratory, University of Geneva, Switzerland.

Skaletsky H, Kuroda-Kawaguchi T, Minx PJ, Cordum HS, et al. (2003). The male-specific region of the human Y-chromosome is a mosaic of discrete sequence classes. *Nature* 423: 825-837.

Torroni A, Achilli A, Macaulay V, Richards M, et al. (2006). Harvesting the fruit of the human mtDNA tree. *Trends Genet.* 22: 339-345.

Wilder JA, Kingan SB, Mobasher Z, Pilkington MM, et al. (2004). Global patterns of human mitochondrial DNA and Y-chromosome structure are not influenced by higher migration rates of females versus males. *Nat. Genet.* 36: 1122-1125.

Yu N, Chen FC, Ota S, Jorde LB, et al. (2002). Larger genetic differences within Africans than between Africans and Eurasians. *Genetics* 161: 269-274.