J Mol Evol (1999) 49:204–214

JOURNAL OF MOLECULAR EVOLUTION

Divergent Human Y-Chromosome Microsatellite Evolution Rates

Denise R. Carvalho-Silva,[1,2] Fabrício R. Santos,[3] Mara H. Hutz,[4] Francisco M. Salzano,[4] Sérgio D.J. Pena[1,2]

[1] Departamento de Bioquímica e Imunologia, Universidade Federal de Minas Gerais, 31270-910 Belo Horizonte, Brazil
[2] Núcleo de Genética Médica de Minas Gerais (GENE/MG), 30130-909 Belo Horizonte, Brazil
[3] Departamento de Biologia Geral, Universidade Federal de Minas Gerais, 31270-910 Belo Horizonte, Brazil
[4] Departamento de Génetica, Instituto de Biociências, Universidade Federal do Rio Grande do Sul, 91501-970 Porto Alegre, Brazil

Received: 7 November 1998 / Accepted: 9 April 1999

Abstract: In this work, we analyze several characteristics influencing the low variability of the microsatellite *DYS19* in the major founder Amerindian Y chromosome lineage containing the point mutation *DYS199*-T. Variation of *DYS19* was compared with that of five other Y-linked tetranucleotide repeat loci (*DYS389*A, *DYS389*B, *DYS390*, *DYS391*, and *DYS393*) in the *DYS199*-T lineage. All the other microsatellites showed significantly higher levels of variability than *DYS19* as measured by gene diversity and repeat number variance. Moreover, we had previously shown that *DYS19* had high diversity in Brazilians and in several other populations worldwide. Thus, the slow *DYS19* evolution in the *DYS199*-T lineage seems to be both locus and allele specific. To understand the slow *DYS19* evolutionary rate, the microsatellite loci were compared according to their mapping on the Y chromosome and also on the basis of structural aspects such as the base composition of the repeat motif and flanking regions and the degree of perfection and size (repeat number) of the variable blocks. The only observed difference that might be related to the low *DYS19* variability is its small average number of repeats, a value expected to be closer to the founder *DYS19* allele in the *DYS199*-T lineage. These data were also compared to other derived Y lineages. The Tat-C lineage displayed a lower *DYS19* variability correlated to a small average repeat number, while in the *DYS234*-G lineage, a high *DYS19* variability was found associated to a larger average repeat number. This approach reveals that evolution of Y microsatellites in lineages defined by slowly evolving markers, such as point mutations, can be greatly influenced by the size (number of repeats of the variable block) of the founder allele in each microsatellite locus. Thus lineage-dating methods using microsatellite variation should be practiced with great care.

Key words: Human Y chromosome — Microsatellites — Mutation rate — Lineages — Human evolution

Correspondence to: S.D.J. Pena; *e-mail:* spena@dcc.ufmg.br

Introduction

Microsatellites are blocks of tandem repeat units of 1 to 6 base pairs (bp) that are ubiquitous, abundant, and highly polymorphic in eukaryotic genomes due to variation in the number of repeating units between alleles (Litt and Luty 1989; Tautz 1989; Weber and May 1989). Their polymorphism is believed to be primarily the result of strand slippage during replication (Levinson and Gutman 1987; Schlötterer and Tautz 1992), although other DNA turnover mechanisms may also play a minor part in generating microsatellite variability (Pena et al. 1994; Dover 1995; Rubinsztein et al. 1995). In general, the population dynamics of microsatellites fits well with the stepwise-mutation model (Valdes et al. 1993; Di Rienzo et al. 1994; Kimmel and Chakraborty 1996).

Besides constituting valuable tools in the characterization of human genetic individuality for forensic and clinical purposes (Pena and Chakraborty 1994; Pena et al. 1995a), microsatellites have been used increasingly to

study the genetic structure of human populations and human evolution (Bowcock et al. 1994; Jorde et al. 1997). If we are to use microsatellites to establish the dates of human evolutionary events (Goldstein et al. 1996; Zerjal et al. 1997), it becomes critical to try to understand the exact mechanisms that cause mutations in these loci and the rate at which these events occur. Despite intense efforts, little knowledge has been gathered so far in both these interrelated areas, primarily because mutations are rare events whose rates are not large enough to be reliably measured by direct means. The mutation rate of microsatellites in the germline has been estimated to be of the order of 10^{-3}–10^{-4} per generation but its range is probably much wider than that (Weber and Wong 1993; Mahtani and Willard 1993; Talbot et al. 1995; Heyer et al. 1997; Brinkmann et al. 1998). Factors that have been implicated in influencing this rate are the length and base composition of the repeat motif (Schlötterer and Tautz 1992; Weber and Wong 1993; Chakraborty et al. 1997), number of repeats (Weber 1990; Gastier et al. 1995; Jin et al. 1996, Brinkmann et al. 1998), base composition of flanking sequences (Bichara et al. 1995; Andreassen et al. 1996; Macaubas et al. 1997), and degree of perfection of the repeats (Weber 1990; Heale and Petes 1995; Bichara et al. 1995; Goldstein and Clark 1995; Sturzeneker et al. 1998).

Recently, we used the tetranucleotide microsatellite *DYS19* in combination with the αh alphoid DNA polymorphism (Santos et al. 1995a) to establish the existence of a major founder Y chromosome haplotype (the type II at αh linked to allele A of 186 bp at *DYS19*) in South and North Amerindians (Pena et al. 1995b; Santos et al. 1995b; Santos et al. 1996a). Our results were confirmed by Underhill et al. (1996), who showed that the founder haplotype was also characterized by a C→T transition at the *DYS199* locus (*DYS199*-T lineage) occurring specifically in Americas that could be as old as 30,000 years before present. Again, the *DYS19* A allele appeared in most of the *DYS199*-T chromosomes. These data raised a molecular conundrum. Why, after more than 10,000 years of the Asian migration into the Americas, did *DYS19* remain almost-invariable in the Amerindian *DYS199*-T lineage, while in other populations (Santos et al. 1996b) it revealed high levels of variability? Do other microsatellites appear at higher diversities than *DYS19* in the *DYS199*-T lineage? If so, is there something special about the *DYS19* mutation rate that makes it different from other Y-linked microsatellites? Could this be related to the location of *DYS19* in the Y chromosome? Or could there be something special about the sequence structure or allele size (repeat number) of the *DYS19* locus in the *DYS199*-T lineage? To try to answer these questions we decided to compare the rate of evolution of *DYS19* with that of five other Y-linked tetranucleotide microsatellites (reviewed by Jobling and Tyler-Smith 1995) in the Amerindian *DYS199*-T lineage. For that purpose, our experimental strategy was the following: by limiting our observations to males displaying the T allele at *DYS199* (*DYS199*-T lineage), we ensured that all Y chromosomes included in the study were descendants of a single Amerindian ancestor that existed probably about 30,000 years ago (Underhill et al. 1996). Since there is no recombination in the major part of the Y chromosome, *DYS19*, *DYS389*A, *DYS389*B, *DYS390*, *DYS391*, and *DYS393* are linked, i.e., all microsatellite loci in each haplotype bearing the T allele at *DYS199* (the *DYS199*-T lineage) have the same effective population size and have been passed through the same number of generations and the same evolutionary constraints such as drift, migration, expansion, and so on. Thus, we expect that independent forward–backward mutations at these different microsatellite loci should accumulate and be the main factor responsible for generating distinct variability. In other words, the relative values of the gene diversity and the repeat number variance at different microsatellite loci should reflect mainly their relative mutation rates. Thus our null hypothesis is that all microsatellite loci in the *DYS199*-T lineage are evolving at the same mutation rate (Heyer et al. 1997; Caglia et al. 1997), which implies that they should produce similar variance values (Zhivotovsky and Feldmann 1995; Goldstein et al. 1996; Kimmel et al. 1998). By making comparisons of gene diversity and/or repeat number variance across Y-linked loci within Y lineages and across *DYS19* in three distinct Y lineages, we should be able to test our null hypothesis and to ascertain, in relative terms, the tempo and mode of evolution of different Y-linked microsatellite loci in the *DYS199*-T lineage.

Materials and Methods

DNA Samples

DNA samples from 57 Amerindian males from South, Central, and North Americas [Surui ($n = 5$), Karitiana ($n = 6$), Muskoke ($n = 6$), Maya ($n = 8$), Ticuna ($n = 32$)] were a kind gift from Dr. Judith Kidd (Department of Genetics of Yale University) and were typed using all markers described below. Moreover, DNA samples previously typed for *DYS19* were studied at all other loci: 37 Amerindian males from the Amazon Region (Santos et al. 1995b), 8 Amerindian males from the Coriell Institute for Medical Research (Camden, NJ, USA; Santos et al. 1996b), and 100 white males from the Southeast region of Brazil (Santos et al. 1993).

PCR Amplification

All samples (102 Amerindians and 100 white Brazilians) were scored for *DYS199* using a new method based on PCR amplification followed by digestion with *Mfe*I according to Santos et al. (1999a). The uncut T alleles (202 bp) were resolved from the cut C alleles (181 + 21 bp) on 6% polyacrylamide gels and silver-stained according to Santos et al. (1993). Ninety-three of the 102 Amerindians and 2 of the 100 Brazilian males carried the T allele at *DYS199* (from now on, all these 95 indi-

Table 1. Details of the six Y microsatellites analyzed in this work

Locus	Repeat sequence[a]	GenBank accession No.	Size range (bp)[b]	Allele nomenclature[c]
DYS19	$(GATA)_3GGTA(GATA)_{12}$	X77751	186–194	194 bp: 12 repeats
*DYS389*A	$(GATA)_{12}(GACA)_6$[d]	G09600	108–124	120 bp: 12 repeats
*DYS389*B	$(GATA)_9(GACA)_3GATA$	—	247–255	247 bp: 9 repeats
DYS390	$(GATA)_4(GACA)(GATA)_{10}(GACA)_8(GATA)_2$	G09611	212–224	212 bp: 10 repeats
DYS391	$(GATA)_{10}(GACA)_3(GATA)$	AF055718	279–287	283 bp: 10 repeats
DYS393	$(GATA)_{13}$	G09601	119–131	123 bp: 13 repeats

[a] Sequence of the allele shown in the last column. These data were obtained from either GenBank or our own sequencing data.
[b] Allele size range observed in this survey.
[c] Alleles are labeled by the repeat number of the largest block.
[d] Allele sequence according to Cooper et al. (1996).

viduals are mentioned as belonging to the *DYS199*-T lineage) and were further typed with six Y-linked tetranucleotide microsatellite loci, namely, *DYS19*, *DYS390*, *DYS391*, *DYS393*, *DYS389*A, and *DYS389*B (Table 1). The first four loci were PCR-amplified together as a tetraplex and *DYS389* was typed separately (Santos et al. 1999b). The primer sequences were obtained from the Genome Database (http://gdbwww.gdb.org/). *DYS389* is actually composed of two microsatellites (*DYS389*A and *DYS389*B, the former apparently having arisen from a partial duplication of the latter) and both are amplified by the same set of primers (Cooper et al. 1996). The allele size of *DYS389*A was calculated by subtracting the size of the small amplicon from that of the larger one. All PCR amplifications were performed according to Santos et al. (1999b).

Alleles were resolved in an automatic fluorescent DNA sequencer (ALF*Express;* Pharmacia, Uppsala, Sweden) and their size established using the Fragment Manager software, version 1.2 (Pharmacia) (Santos et al. 1999b). Alleles were named according to the repeat number of their largest GATA block, which were determined from the microsatellite sequence in GenBank (Table 1). The sequence for *DYS391* in GenBank contained too many ambiguities and was redone (see below). Allele ladders for each microsatellite were constructed by PCR amplification of a pool of DNA from 100 Brazilian males.

Chromosomal Mapping of the Six Y Microsatellite Loci

In order to map physically *DYS19*, *DYS389*, *DYS390*, *DYS391*, and *DYS393* on the human Y chromosome, we employed two approaches. The first, originally described by Tyler-Smith et al. (1993), made use of a deletion panel composed of 33 cell lines carrying varied deletions of the Y chromosome. Each cell line was analyzed by PCR amplification and the products were resolved on 6% polyacrylamide gels and silver-stained. The second approach made use of the Stanford G3 human–hamster radiation hybrid somatic cell panel (Cox et al. 1990; Stewart et al. 1997), which was purchased from Research Genetics (Huntsville, AL, USA). The amplification products were resolved using a denaturing 6% polyacrylamide gel on the ALF*Express* sequencer (Pharmacia). Results of the radiation hybrid panel were analyzed through the World Wide Web at http://www-shgc.stanford.edu/.

Cloning and Sequencing

As an attempt to access the molecular and structural basis of the microsatellite loci studied here, we cloned and sequenced alleles at the *DYS19* and *DYS391* loci (the other sequences were obtained from GenBank). PCR products were cloned either into pCRII vector using the T-A Cloning Kit-R (Invitrogen, San Diego, USA) or into pUC18 vector using the Sureclone kit (Pharmacia). The products of ligation were used to transform competent DH5α *Escherichia coli* cells. Sequencing was accomplished by the dideoxy-chain termination method with the Cycle Sequencing kit (Amersham Life Science, Buckinghamshire, England) by employing M13 primers labeled at their 5′ ends with fluorescein. The sequencing products were run on an ALF automatic fluorescent laser DNA sequencer (Pharmacia).

Statistical Analysis

Allele frequencies for each of the six Y microsatellites were scored by single-gene counting procedures. Microsatellite diversity was measured as the simple Nei's (1987) gene diversity using $D = 1 - \sum i^2$. This index is not sensitive to artificial population agglomeration (Chakraborty et al. 1988) being quite suitable in our case, where individuals from distinct populations were studied. In order to access the significance of the distinct diversity values observed at all loci that could be due to stochastic sampling errors, we established 99% confidence limits by doing 10,000 bootstrap simulations using the GENETIX software (Belkhir et al. 1998). The repeat number variance was also calculated (Goldstein et al. 1996) and the significant differences between microsatellite loci were evaluated by performing a nonparametric Friedman ANOVA test.

Results

Microsatellite Variability Within the DYS199-*T Lineage*

The allele distributions for all microsatellite loci in the *DYS199*-T lineage ($n = 95$) are shown in Table 2. *DYS19* was quite invariable in this lineage, with only two alleles found. The only individual who did not show the 10-repeat allele (A, or 186 bp) at *DYS19* was a Mayan who had the 12-repeat allele (a 2-repeat difference in the major GATA block). High values of *DYS19* diversity were observed in the sample of white Brazilians (Santos et al. 1993).

As measures of diversity we used two parameters: repeat number variance and gene diversity indexes. The latter does not take into account the stepwise mutation model. Our results for the Amerindian *DYS199*-T lineage are shown in Table 3. The smaller value for the variance

Table 2. Allele frequency distribution of six Y-tetranucleotide loci in the *DYS199*-T lineage

Locus	Allele[a]	Frequency in Amerindians ($n = 95$)
DYS19	10	0.99
	12	0.01
*DYS389*A	09	0.01
	10	0.42
	11	0.40
	12	0.11
	13	0.06
*DYS389*B	09	0.06
	10	0.73
	11	0.21
DYS390	10	0.15
	11	0.55
	12	0.29
	13	0.01
DYS391	09	0.01
	10	0.87
	11	0.12
DYS393	12	0.04
	13	0.70
	14	0.06
	15	0.20

[a] Alleles are named by the repeat number of the largest $(GATA)_n$ block.

was detected at *DYS19* (0.04), while larger values were seen for all other loci, with the highest at *DYS389*A (0.78). The variance differences observed suggest that in Y chromosomes bearing allele T at *DYS199*, microsatellite loci are evolving at very distinct rates. When we performed a nonparametric ANOVA test, we obtained significant differences ($p < 0.05$) for each pair of microsatellite loci studied.

We also used 10,000 bootstrap simulations (Fig. 1) to establish 99% confidence limits (Table 3) for the gene diversity in all six Y microsatellite loci. The simulated *DYS19* diversity distribution was significantly distinct ($p < 0.01$) from all other microsatellite distributions (Table 3). Thus, in the *DYS199*-T lineage, locus *DYS19* seems to be evolving at a much reduced mutation rate compared with the other Y-linked microsatellite loci. It is important to make clear that if we perform the same simulations with only 48 of 57 *DYS199*-T Amerindians not previously typed with *DYS19* (see Materials and Methods), we obtain find the same results, indicating that we have no significant ascertainment bias.

The previous analyses allow us to reject the null hypothesis of no differences in the mutation rates at six Y-linked microsatellite loci in the Amerindian *DYS199*-T lineage. Data published elsewhere for different Amerindian samples (Underhill et al. 1996, Deka et al. 1996; Rodriguez-Delfin et al. 1997) and simulations with a bootstrap resampling method (Belkhir et al. 1998) indicate that our results cannot be explained by sampling or stochastic events. Nor can the low mutation rate for *DYS19* (that we observed by measuring the repeat number variance and gene diversity) in the *DYS199*-T lineage be explained by its presence on the nonrecombining part of the Y chromosome, since the other microsatellite loci located in the same region are much more variable. The high variability observed in other microsatellites also excludes the founder effect in the Amerindian population as a single evolutionary fact explaining the low *DYS19* variability. In view of these facts, we decided to examine the microsatellites under the light of several distinct features as follow.

Mapping of the Microsatellites on the Y Chromosome

Since no published mapping data were available for these six Y microsatellite loci, we examined if there was a correlation of the relative mutation rates of these microsatellites with their location on the Y chromosome.

We used two mapping methods to locate the microsatellite loci on the human Y chromosome: a deletion panel and a radiation hybrid panel. The deletion panel (Tyler-Smith et al. 1993) divides the human Y chromosome into 24 intervals, with the centromere located in interval 8. This means that from interval 1 to interval 7, this panel represents the short arm of the Y, while the long arm lies between intervals 9 and 24. Our results for

Table 3. Gene diversity (D) and allele size variance (s^2) values for six Y microsatellite loci in the *DYS199*-T ($n = 95$), Tat-C[a] ($n = 60$), and *DYS234*-G[b] ($n = 30$) lineages

	DYS199-T lineage		Tat-C lineage		*DYS234*-G lineage	
Locus	D (CI)[c]	s^2	D (CI)	s^2	D (CI)	s^2
DYS19	0.02 (0–0.09)	0.04	0.03 (0–0.13)	0.02	0.68 (0.48–0.79)	0.88
*DYS389*A	0.65 (0.58–0.71)	0.78	0.53 (0.35–0.66)	0.70	0.35 (0.07–0.58)	0.28
*DYS389*B	0.43 (0.29–0.54)	0.25	0.15 (0–0.31)	0.08	0.37 (0.13–0.59)	0.23
DYS390	0.59 (0.49–0.66)	0.46	0.35 (0.16–0.51)	0.23	0.67 (0.44–0.78)	1.17
DYS391	0.22 (0.11–0.36)	0.12	0.51 (0.41–0.57)	0.28	0.53 (0.33–0.62)	0.31
DYS393	0.47 (0.33–0.60)	0.73	0.21 (0.04–0.4)	0.17	0.38 (0.13–0.57)	0.27

[a] From Zerjal et al. (1997).
[b] From Hurles et al. (1998).
[c] Simulated 99% confidence intervals (Belkhir et al. 1998) for gene diversity based on the resampling of observed allele frequencies.

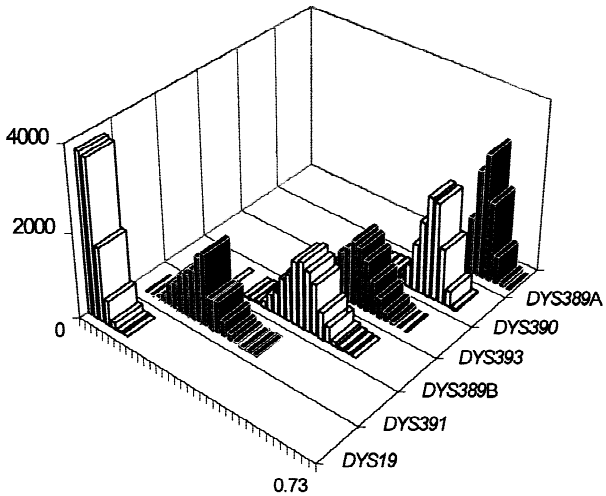

Fig. 1. Class distribution (0.02 interval on the *X* axis) of 10,000 bootstrap simulations (Belkhir et al. 1998) of gene diversity values for each of the six Y-linked microsatellite loci in the *DYS199*-T lineage.

the five microsatellites are shown in Table 4. The presence of the locus is indicated by a plus sign. These results were compared to the mapping interval determined by 56 loci according to the original article (Tyler-Smith et al. 1993). These results indicate that *DYS19* and *DYS393* both mapped to the short arm, while *DYS389, DYS390,* and *DYS391* were located on the long arm. Moreover, *DYS19* and *DYS390* were unambiguously assigned to intervals 5 (where the *50f2/B* locus is situated) and 14 (together with the *GMG XY3* and *25H4* loci), respectively (Tyler-Smith et al. 1993). On the other hand, we could locate *DYS389* between the *14A3C* (interval 11) and the *RBF5* (interval 12) loci, thus creating another interval not determined before (Tyler-Smith et al. 1993). *DYS393* was found to be associated with the *RBF6* and *LLY25h* loci (interval 3). *DYS391* was assigned to interval 11 (together with the *14A3C* locus). For the last three microsatellites, some discrepancies in the amplification results were found that could be due to chromosome rearrangements frequently observed in these cells (Tyler-Smith et al. 1993).

By typing the Stanford G3 radiation hybrid panel (Table 5), we found links for each one of the five microsatellites, with the exception of *DYS19* (that fortunately had been assigned previously by us using deletion mapping). For the first time, a detailed map of the microsatellites on the human Y chromosome is shown (Fig. 2) but no apparent association between chromosomal location and distinct variability for the studied loci could be detected.

Structural Features of the Y Microsatellite Loci

DYS19 *Flanking Regions and Repeat Blocks.* To ascertain whether other sequence variations, outside or inside the repeat region, might be seen in the *DYS19* locus present in the *DYS199*-T lineage, we sequenced three 10-repeat alleles from distinct Amerindian males (Fig. 3). No difference in sequence was observed compared with the Caucasian 10-repeat allele (Roewer et al. 1992). When two Caucasian alleles (11- and 12-repeat) and one Bushman allele (9-repeat) were compared with the Amerindian 10-repeat alleles, the only differences observed were in the number of GATA repeats. These comparisons also showed that the largest GATA block is the region where insertions or deletions of GATA repeat units preferentially occur.

Repeat Block Structure, Size, and Variability. Many factors have been implicated in influencing the rate of autosomal microsatellite mutation but no similar approach was devoted to Y-linked microsatellites. In Table 1, some parameters for each locus are displayed, such as the structure of the repeat blocks with their degree of perfection, the size of the largest repeat blocks, the allele size range, and the correspondence of allele length (bp) and repeat number of the largest repeating block. All the loci studied were tetranucleotide microsatellites displaying mixtures of two repeat motifs: GATA and GACA. *DYS393* was the only perfect tetranucleotide, displaying 13 GATA repeats for the 123-bp allele. *DYS390, DYS391, DYS389*A, and *DYS389*B were the more complex, with a $(GATA)_n(GACA)_n$ arrangement, *DYS19* had a $(GATA)_n$ array interrupted by a GGTA motif. This interruption, probably the fruit of an A→G transition, might influence the mutation rate of this microsatellite (see Macaubas et al. 1997). For instance, it is known that in the expansion-prone trinucleotide microsatellites involved in the etiology of both spinocerebellar ataxia (Chung et al. 1993) and *FRAXA* (Hirst et al. 1994), alleles with motif interruptions appear to be much more stable than perfect ones. For *DYS19* it seems that only the large block (see Fig. 3) is involved in generating the microsatellite variability.

An association of repeat number variance values with the average repeat number of the largest block was observed in the *DYS199*-T lineage, especially for *DYS19* and *DYS393,* the microsatellite loci with the simplest repeat blocks (Table 1, Fig. 4A). The perfect microsatellite *DYS393* showed one of the greatest values for variance and gene diversity (Table 3). *DYS389*A and *DYS390* showed the highest variance and gene diversity values, probably due to their complex array of repeat motifs, where both have another GACA block with at least six repeats. These complex microsatellites can have distinct evolving repeat blocks, which increase their overall mutation rate. This would result in larger variance and gene diversity values. At least for *DYS390,* the high mutation rate due to variation in distinct repeat blocks was already noted (Forster et al. 1998).

Comparison of Microsatellite Variability Between Distinct Y Lineages

In Table 3, the microsatellite variability of the *DYS199*-T lineage is compared with literature data for two other Y

Table 4. Amplification data obtained in 33 deleted Y-chromosome cells for five Y microsatellites

No.	Cell line	Description	*DYS19*	*DYS389*	*DYS390*	*DYS391*	*DYS393*
1	4532	46,XX, man	–[a]	–	–	–	–
2	RFY1	46,XX, man	–	–	–	–	–
3	89030291	46,XX, man	+[a]	–	–	–	+
4	Q998-8 der	Y-hybrid	+	–	–	–	+
5	CHoP	45,Xt(Y;18), man	+	–	–	–	+
6	842968	46,X,I(Yp)	+	–	–	–	+
7	KIB	46,X,del(Y)(q11.2)	+	–	–	–	+
8	1491/76R	45,X,t(Y;18), man	+	–	–	–	+
9	M-6R-1	46,X,del(Yq)	+	–	–	+	+
10	GM2668	45,X/46,X,del(Y)(q12)	+	+	–	+	+
11	MSM	46,X,del(Y)(q11.23)	+	+	–	+	+
12	TAP	46,X,del(Y)(q11.1)	+	+	–	+	+
13	BIT	46,X,del(Yq)	+	+	+	+	+
14	JB	46,X,idic(Y)(q11.22)	+	+	+	+	+
15	DL-2782/80R	46,X,idic(Y)(q11)	+	+	+	+	+
16	WSM184	45,X/46,X,idic(Y)(q11)	+	+	+	+	+
17	863829	45,X/46,X,idic(Yp)(q11.21)	+	+	+	+	+
18	863712	45,X/46,X,idic(Yp)(q11.21)	+	+	+	+	+
19	870740	46,X,idic(Yp)(q11.22)	+	+	+	+	+
20	DAA	46,X,del(Y)(q11.3)	+	+	+	+	+
21	JOW	45,X/46,X,r(Y)	+	+	+	+	+
22	PIP	46,X,idic(Y)(q12)	+	+	+	+	+
23	134/89	45,X/46,X,idic(Y)(q12)	+	+	+	–	+
24	BRG	46,X,del(Yq)	+	+	+	+	+
25	GM6967	45,X/46,X,idic(Y)(p11)/47,X,idic(Y),idic(Y)	+	+	+	+	+
26	861748	48,XX,idic(Y),idic(Y)(p11.2)	+	+	+	+	–
27	870377	45,X/46,X,idic(Y)(p11)	+	+	+	+	+
28	M188044	45,X/46,X,r(Y)	+	+	+	–	–
29	651	46,XY, woman	–	+	+	+	–
30	GM2103	46,X,t(X;Y)(Yq11>Yqter)	–	–	–	–	+
31	GM2469	46,X,t(X;Y)(Yq11>Yqter)	–	–	–	–	–
32	GM0118	46,XX,t(Y;15)(Yq11.Yqter)	–	–	–	–	+
33	GM7970	47,XX,+t(Yq12.Yqter)	–	–	–	–	+

[a] Absence (−) or presence (+) of amplification.

Table 5. Physical mapping of Y microsatellites by PCR analysis of the Stanford G3 panel

Microsatellite	Linked marker	LOD score	Distance (cR)[a]	Mapping
DYS19	ND[b]			
*DYS389*A	*DYS271*	13.71	4.34	Yq11.21
*DYS389*B	*DYS148*	11.92	5.05	Yq11.21
DYS390	*DYS280*	5.47	31.70	Yq11.21
DYS391	*DYS271*	9.79	18.17	Yq11.21
DYS393	*DYS252*	11.52	5.43	Yp11.31

[a] One centiray (cR) roughly equals a distance of 30 kb.
[b] Not determined (i.e., no links were found in the radiation hybrid panel).

chromosome lineages: the Tat-C lineage found in North Eurasians (Zerjal et al. 1997) and the *DYS234*-G lineage found in Polynesians (Hurles et al. 1998). The three derived mutations define lineages that originated in slightly distinct times (Underhill et al. 1996; Zerjal et al. 1997; Hurles et al. 1998), and their average microsatellite variance (for the six loci studied here) was 0.44, 0.24, and 0.52 for *DYS199*-T, Tat-C, and *DYS234*-G respectively. Since these values are much lower than the expected value at the mutation drift equilibrium (Goldstein et al. 1996), the microsatellite variance seems to reflect the variability of lineages that originated recently, as displayed by their restricted geographic location (Underhill et al. 1996; Zerjal et al. 1997; Hurles et al. 1998). In Table 3 we can see the magnitude that the variance of the repeat number reaches for each of the six Y-linked loci, within and between Y chromosome lineages. The microsatellite variance values in the *DYS199*-T and Tat-C lineages were compared to their average repeat numbers (Figs. 4A and B). Thus, in the *DYS199*-T and Tat-C lineages, the microsatellite *DYS19* presented small average repeat numbers of 10.02 and 11.02, respectively, and low variance values of 0.04 and 0.02, respectively, while in the *DYS234*-G lineage, *DYS19* displays a larger average repeat number of 12.5 and a high variance of 0.88 (Table 3).

Discussion

The basic question that we tried to answer in this study was why, after more than 10,000 years (i.e. 500 genera-

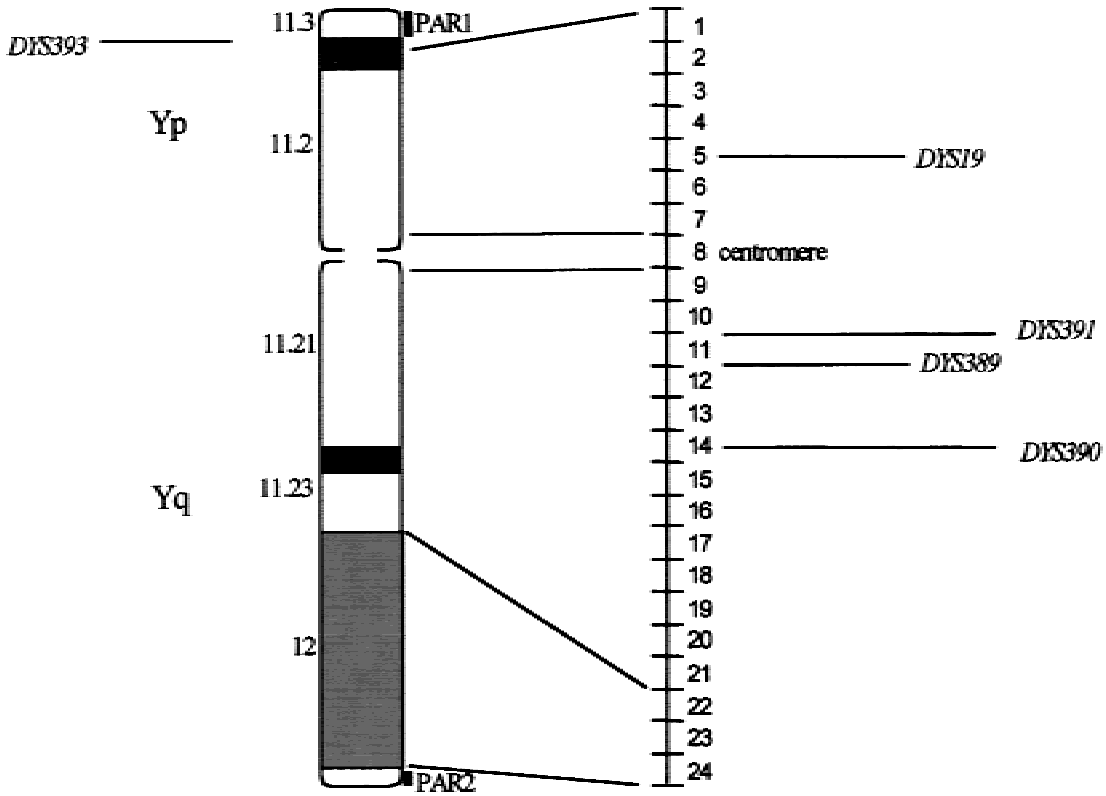

Fig. 2. Regional mapping of the tetranucleotide microsatellite loci on the human Y chromosome.

tions) of the Asian migration into the Americas, did *DYS19* remain almost-invariable, with a high frequency of the 10-repeat allele (186 bp) in the *DYS199*-T lineage? First, we examined the variability of five other tetranucleotide repeat loci (*DYS389*A, *DYS389*B, *DYS390*, *DYS391*, and *DYS393*) in the *DYS199*-T lineage. All microsatellites studied displayed significantly higher levels of variation than *DYS19* (Fig. 1, Tables 2 and 3). Then we rejected the hypothesis of identical mutation rates for the different Y-linked microsatellite loci (indirectly measured by the repeat number variance and gene diversity indexes), a common behavior already noted for human autosomal microsatellites (Jin et al. 1996, Di Rienzo et al. 1998) and *Drosophila* (Schlötterer et al. 1998). Thus we have investigated distinct properties that could be influencing the mutation rates in these Y microsatellites.

Chromosomal Location

We mapped all loci regionally and found that *DYS389*A, *DYS389*B, *DYS390*, and *DYS391* were located on the long arm, while *DYS19* and *DYS393* were on the short arm of the Y chromosome (Fig. 2). Judged by the presence of many regions with different replication timing (Camargo and Cervenka 1982; Schempp et al. 1989), we would expect that chromatin packing of the human Y chromosome should be quite heterogeneous, and this could influence the rate of occurrence of DNA strand slippage or of mismatch repair (or other genome turnover mechanisms) responsible for microsatellite evolution. We could not find any apparent association between distinct mutation rates and localization of the microsatellite loci. However, *DYS19* was found to be located in the large Giemsa-negative Yp11.2 band where the structural organization is poorly understood.

Structural Features of Both Repeat and Flanking Regions

We next compared the length, base composition of the repeat motif and flanking regions, number of repeats, and degree of perfection of *DYS19* with those of the other Y microsatellite loci. We could not find any significant structural peculiarities of *DYS19* that could explain its low variability except for a discrete positive correlation between the average repeat size of the large block and the repeat number variance measures (Fig. 4A). Besides, of particular relevance was the observation that the nucleotide sequences of three Amerindian 10-repeat alleles compared to a Bushman 9-repeat allele and to Caucasian 10-, 11-, and 12-repeat alleles have shown differences only in the repeat number of the larger GATA block (Fig. 3). When the average repeat number is smaller (10.02 for *DYS19*, 10.08 for *DYS391*, and 10.14 for *DYS389*B), lower variance values were obtained in the *DYS199*-T lineage (Fig. 4A). However, *DYS389*B and *DYS391*, as

1) 12rp (C): CTACTGAGTTTCTGTTATAGTGTTTTTAATATATATATAGTATTATATATATAGTGTTATATATATATAGTGTTTTA**GATAGATAGATAGGTAGATAGATAGATAGATAGATAGATAGATAGATAGATAGATAGATAGATA**TAGTGACACTCTCCTTAACCCAGATGGACTCCTTGTCCTCACTACATGCCAT

2) 11rp (C): CTACTGAGTTTCTGCAATAGTGTCNTTAAATATATATATAGTATTATATATATAGNGTNANNTATATATAGTNTTTTA**GCTAGATAGATAGGTAGATAGATAGATAGATAGATAGATAGATAGATAGATAGATAGATA**----TAGTGACACTCTCCTNAACCCAGCNGGCCTCCTNNTNCTCACTACATGCCAT

3) 10rp (C): CTACTGAGTTTCTGTTATAGTGTTTTTAATATATATATAGTATTATATATATAGTGTTATATATATATAGTGTTTTA**GATAGATAGATAGGTAGATAGATAGATAGATAGATAGATAGATAGATAGATAGATA**--------TAGTGACACTCTCCTTAACCCAGATGGACTCCTTGTCCTCACTACATGCCAT

4) 10rp (A): CTACTGAGTTTCTGTTATAGTGTTTTTAATATATATATAGTATTATATATATAGTGTTATATATATATAGTGTTTTA**GATAGATAGATAGGTAGATAGATAGATAGATAGATAGATAGATAGATAGATAGATA**--------TAGTGACACTCTCCTTAACCCAGATGGACTCCTTGTCCTCACTACATGCCAT

5) 10rp (A): CTACTGAGTTTCTGTTATAGTGTTTTTAATATATATATAGTATTATATATATAGTGTTATATATATATAGTGTTTTA**GATAGATAGATAGGTAGATAGATAGATAGATAGATAGATAGATAGATAGATAGATA**--------TAGTGACACTCTCCTTAACCCAGATGGACTCCTTGTCCTCACTACATGCCAT

6) 10rp (A): CTACTGAGTTTCTGTTATAGTGTTTTTAATATATATATAGTATTATATATATAGTGTTATATATATATAGTGTTTTA**GATAGATAGATAGGTAGATAGATAGATAGATAGATAGATAGATAGATAGATAGATA**--------TAGTGACACTCTCCTTAACCCAGATGGACTCCTTGTCCTCACTACATGCCAT

7) 9rp (K): CTACTGAGTTTCTGTTATAGTGTTTTTAATATATATATAGTATTATATATATAGTGTTATATATATATAGTGTTTTA**GATAGATAGATAGGTAGATAGATAGATAGATAGATAGATAGATAGATAGATA**------------TAGTGACACTCTCCTTAACCCAGATGGACTCCTTGTCCTCACTACATGCCA

Fig. 3. Comparison of the sequence of different alleles of *DYS19* (classified according to the repeat number of the largest block; rp) in Amerindians (A), Caucasians (C), and a Bushman (K). The repeat array regions are shown in *boldface* type and the interrupting GGTA motif is *underlined*. Sequence 1 is the one found in GenBank under accession number X77751. Sequence 2 was actually described as being an allele from a different locus, *DYS394* (GenBank accession number G10053), which, however, proved to be the same as *DYS19* (Jobling and Tyler-Smith 1995). Sequence 3 is the Caucasian A allele (10 repeats) sequenced in the original description of *DYS19* (Roewer et al. 1992). Sequences 4–6 are the Amerindian A alleles (10 repeats) sequenced in the present study. Sequence 7 is a nine-repeat allele found by us in a Bushman (Santos et al. 1996b). Sequences 6 and 7 were deposited under the respective accession numbers AF057296 (10-repeat allele) and AF057295 (9-repeat allele).

well as *DYS389*A and *DYS390*, show complex structures, with possible variation occurring in other repeat blocks (Table 1). Thus repeat number-dependent variation should be analyzed carefully in complex microsatellite structures. For example, the correlation between average repeat number and variance could not be observed in the *DYS199*-T lineage for *DYS389*A (Fig. 4A) and in the Tat-C lineage for *DYS390*, *DYS391*, and *DYS389*A (Fig. 4B), where other large repeat blocks can be generating further diversity. On the other hand, the perfect microsatellite *DYS393* contains a single repeat block, with an average repeat number of 13.43 (the largest among all microsatellites), as well as high gene diversity and repeat number variance values in the *DYS199*-T lineage (Table 3, Fig. 4A). In the Tat-C lineage (Fig. 4B) *DYS393* is not as variable but considering that this lineage (Zerjal et al. 1997) is more recent than the *DYS199*-T lineage (Underhill et al. 1996), we expect a lower variation for the former.

Our Y microsatellites can be compared with literature data for autosomal microsatellites, where it was observed that a difference of only one repeat unit may be responsible for determining the distinct mutation rate within the same locus (Jin et al. 1996; Brinkmann et al. 1998). Thus, each allele could present an independent evolutionary pace, smaller alleles having low mutation rates and larger ones displaying higher mutation rates (Jin et al. 1996; Brinkmann et al. 1998, Schlötterer et al. 1998). The same might be occurring with the human Y chromosome microsatellites. Then we compared the *DYS19* variance and average repeat number present in each distinct lineage (*DYS199*-T, Tat-C, *DYS234*-G). Presumably, the average repeat number is a closer value to the founder allele, which was present in the Y chromosome where each point mutation defining the lineage has arisen, i.e., the founder microsatellite allele (Santos and Tyler-Smith 1996). In the *DYS199*-T and Tat-C lineages, we found low diversity and variance values at *DYS19* (Table 3, Figs. 4A and B) associated with short alleles (small average repeat number). Conversely, in the *DYS234*-G lineage (Hurles et al. 1998), where the most frequent alleles at *DYS19* are the 12- and 13-repeat ones (average of 12.5 repeats), a higher repeat number variance (0.88) was detected (Table 3). Despite the fact that the Tat-C and *DYS199*-T lineages show similar *DYS19* diversity and variance values, their average repeat numbers (Table 3) differ by about one repeat unit. As the average of the observed variance in all Y microsatellite loci is proportional to the lineage origin time (Goldstein et al. 1996; Zerjal et al. 1997), it seems that *DYS19* in the Tat-C lineage has accumulated a diversity similar to that of the *DYS199*-T lineage but in fewer generations, which could be a reflex of its larger founder allele (Fig. 4B). Therefore, it seems conceivable that the smaller repeat number in the *DYS19* locus of the *DYS199*-T lineage could be responsible for the distinctiveness in gene diversity and variance values demonstrated in this work.

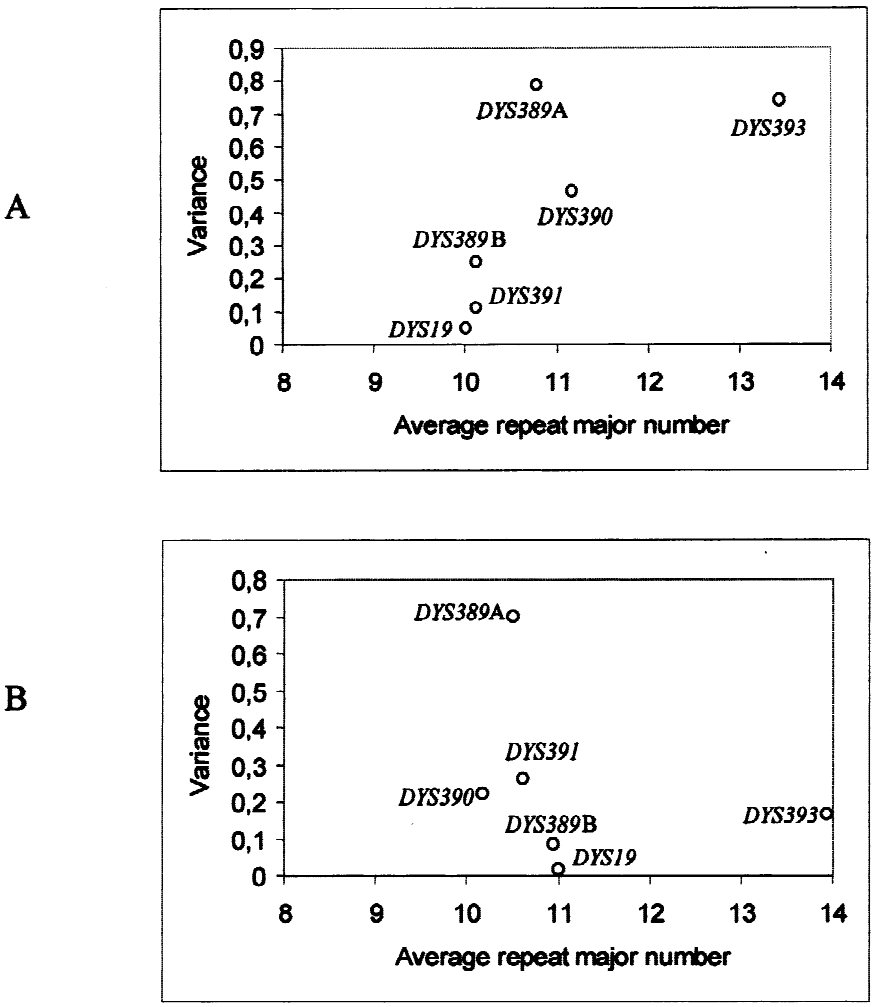

Fig. 4. Association between the repeat number variance values and the average repeat number of the largest blocks of all microsatellites in the *DYS199*-T **(A)** and Tat-C **(B)** lineages.

This study shows that Y-linked microsatellites, inherited as haplotypes, have advantages for the study of characteristics influencing mutation rates, compared to autosomal microsatellites (Brinkmann et al. 1998; Schlötterer et al. 1998). Y microsatellites can be studied within lineages, where distinct loci can be analyzed independently of other evolutionary variables. They can also be studied between lineages where the same microsatellite loci reveal distinct evolutionary pathways depending on their different founder alleles.

In summary, in the present work we observed distinct microsatellite mutation rates generating different diversity and variance values. The difference in the mutation rate per locus could not be explained by distinct Y-chromosome localization or by sequence differences in the repeat and flanking regions. However, an association between the size of the larger repeat block and the locus diversity was observed. When the *DYS19* locus data in different Y-chromosome lineages were compared, smaller founder alleles were associated with lower diversity levels, while larger founder alleles were seen when a high diversity was observed. Thus, the size of the founder allele (repeat number) for each lineage may play a role in the mode and tempo of microsatellite evolution. Besides, it is clear that allele- and locus-specific mutation rates should be specifically determined in each study to allow more precise and reliable datings using microsatellite variation.

Acknowledgments. We are grateful to Dr. Judith Kidd (Department of Genetics of Yale University), Carlos E.A. Coimbra, Jr. (FIOCRUZ), and Ricardo V. Santos (FIOCRUZ and Departamento de Antropologia, Museu Nacional, Rio de Janeiro) for kindly providing us with some Amerindian samples used in this study, Eduardo M. Tarazona-Santos (Departamento de Bioquímica, ICB, UFMG) for helpful statistical advice and François Bonhomme (Genome and Population Laboratory at University of Montpellier) for his availability in solving our doubts about GENETIX. This research was supported by grants from the Centro Brasil–Argentina de Biotecnologia, Fundação de Amparo à Pesquisa do Estado de Minas Gerais, Conselho Nacional de Desenvolvimento Cientifico e Tecnológico of Brazil, Programa de Apoio a Núcleos de Excelência, Financiadora de Estudos e Projetos, Fundação de Amparo à Pesquisa do Estado do Rio Grande do Sul, Wenner–Gren Foundation for Anthropological Research, and McArthur Foundation.

References

Andreassen R, Egeland T, Olaisen B (1996) Mutation rate in the hypervariable VNTR g3 (D7S22) is affected by allele length and a

flanking DNA sequence polymorphism near the repeat array. Am J Hum Genet 59:360–367

Belkhir K, Borsa P, Goudet J, Chikhi L, Bonhomme F (1998) *GENETIX, logiciel sous WindowsTM pour la génétique des populations.* Laboratoire Génome et Populations, CNRS UPR 9060, Université de Montpellier II, Montpellier, France

Bichara M, Schumacher S, Fuchs RPP (1995) Genetic instability within monotonous runs of CpG sequences in *Escherichia coli*. Genetics 140:897–907

Bowcock AM, Ruiz-Linares A, Tomfohrde J, Minch E, Kidd JR, Cavalli-Sforza LL (1994) High resolution of human evolutionary trees with polymorphic microsatellites. Nature 368:455–457

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

Caglia A, Novelletto A, Dobosz M, Malaspina P, Ciminelli BM, Pascali VL (1997) Y-chromosome STR loci in Sardinia and continental Italy reveal islander-specific haplotypes. Eur J Hum Genet 5:288–292

Camargo M, Cervenka J (1982) Patterns of DNA replication of human chromosomes. II. Replication map replication model. Am J Hum Genet 34:757–780

Chakraborty R, Smouse PE, Neel JV (1988) Population amalgamation and genetic variation: Observations on artificially agglomerated tribal populations of Central and South America. Am J Hum Genet 43:709–725

Chakraborty R, Kimmel M, Stivers DN, Davison LJ, Deka R (1997) Relative mutation rates at di-, tri, and tetranucleotide microsatellite loci. Proc Natl Acad Sci USA 94:1041–1046

Chung MY, Ranum LP, Duvick LA, Servadio A, Zoghbi HY, Orr HT (1993) Evidence for a mechanism predisposing to intergenerational CAG repeat instability in spinocerebellar ataxia type I. Nature Genet 5:254–258

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

Cox DR, Burmeister M, Price ER, Kin S, Myers RM (1990) Radiation hybrid mapping: A somatic cell genetic method for constructing high-resolution maps of mammalian chromosome. Science 250: 245–250

Deka R, Jin L, Shriver MD, et al. (1996) Dispersion of human Y chromosome haplotypes based on five microsatellites in global populations. Genome Res 6:1177–1184

Di Rienzo A, Peterson AC, Garza JC, Valdes AM, Slatkin M, Freimer NB (1994) Mutational processes of simple-sequence repeat loci in human populations. Proc Natl Acad Sci USA 91:3166–3170

Di Rienzo A, Donnelly P, Toomajian C, et al. (1998) Heterogeneity of microsatellite mutations within and between loci, and implications for human demographic histories. Genetics 148:1269–1284

Dover G (1995) Slippery DNA runs on and on and on . . . . Nature Genet 10: 254–256

Forster P, Kayser M, Meyer E, et al. (1998) Phylogenetic resolution of complex mutational features at Y-STR DYS390 in aboriginal Australians and Papuans. Mol Biol Evol 15:1108–1114

Gastier JM, Pulido JC, Sunden S, et al. (1995) Survey of trinucleotide repeats in the human genome: Assessment of their utility as genetic markers. Hum Mol Genet 4:1829–1836

Goldstein DB, Clark AG (1995) Microsatellite variation in North American populations of *Drosophila melanogaster*. Nucleic Acids Res 23:3882–3886

Goldstein DB, Zhivotovsky LA, Nayar K, et al. (1996) Statistical properties of the variation at linked microsatellite loci: Implications for the history of human Y chromosomes. Mol Biol Evol 13:1213–1218

Heale SM, Petes TD (1995) The stabilization of repetitive tracts of DNA by variant repeats requires a functional DNA mismatch repair system. Cell 83:539–545

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

Hirst MC, Grewal PK, Davies KE (1994) Precursor arrays for triplet repeat expansion at the fragile X locus. Hum Mol Genet 3:1553–1560

Hurles ME, Irven C, Nicholson C, et al. (1998) European Y-chromosomal lineages in Polynesia: A contrast to the population structure revealed by mitochondrial DNA. Am J Hum Genet 63: 1793–1806

Jin L, Macaubas C, Hallmayer J, Kimura A, Mignot E (1996) Mutation rate varies among alleles at a microsatellite locus: Phylogenetic evidence. Proc Natl Acad Sci USA 93:15285–15288

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

Jorde LB, Rogers AR, Bamshad M, et al. (1997) Microsatellite diversity and the demographic history of modern humans. Proc Natl Acad Sci USA 94:3100–3103

Kimmel M, Chakraborty R (1996) Measures of variation at DNA repeat loci under a general stepwise mutation model. Theor Popul Biol 50:345–367

Kimmel M, Chakraborty R, King JP, Bamshad M, Watkins WS, Jorde LB (1998) Signatures of population expansion in microsatellite repeat data. Genetics 148:1921–1930

Levinson G, Gutman GA (1987) Slippage-strand mispairing: A major mechanism for DNA sequence evolution. Mol Biol Evol 4:203–221

Litt M, Luty JA (1989) A hypervariable microsatellite revealed by in vitro amplification of a dinucleotide repeat within the cardiac muscle actin gene. Am J Hum Genet 44:397–401

Macaubas C, Jin L, Hallmayer J, Kimura A, Mignote E (1997) The complex mutation pattern of a microsatellite. Genome Res 7:635–641

Mahtani MM, Willard HF (1993) A polymorphic X-linked tetranucleotide repeat locus displaying a high rate of new mutation: Implications for mechanisms of mutation at short tandem repeat loci. Hum Mol Genet 2:431–437

Nei M (1987) Molecular evolutionary genetics. Columbia University Press, New York, p 177

Pena SDJ, Chakraborty R (1994) Paternity testing in the DNA era. Trends Genet 10:204–209

Pena SDJ, Souza KT, Andrade M, Chakraborty R (1994) Allelic associations of two polymorphic microsatellite in intron 40 of the human von Willebrand factor gene. Proc Natl Acad Sci USA 91:723–727

Pena SDJ, Prado VF, Epplen JT (1995a) DNA diagnosis of human genetic individuality. J Mol Med 73:555–564

Pena SDJ, Santos FR, Bianchi NO, et al. (1995b) A major founder Y-chromosome haplotype in Amerindians. Nature Genet 11:15–16

Rodriguez-Delfin L, Santos SE, Zago MA (1997) Diversity of the human Y chromosome of South American Amerindians: A comparison with blacks, whites, and Japanese from Brazil. Ann Hum Genet 61:439–448

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

Rubinsztein DC, Amos W, Leggo J, et al. (1995) Microsatellite evolution: Evidence for directionality and variation in rate between species. Nature Genet 10:337–343

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

Santos FR, Pena SDJ, Epplen JT (1993) Genetic and population study of a Y-linked tetranucleotide repeat DNA polymorphism with a simple non-isotopic technique. Hum Genet 90:655–656

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

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

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

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

Santos FR, Pandya A, Tyler-Smith C, et al. (1999a) The central siberian origin for native american Y chromosomes. Am J Hum Genet 64:619–628

Santos FR, Carvalho-Silva DR, Pena SDJ (1999b) PCR-based DNA profiling of human Y chromosomes. In: A laboratory guide to DNA fingerprinting/profiling. Birkhäuser (in press)

Schempp W, Weber B, Müller G (1989) Mammalian sex-chromosome evolution: A conserved homoeologous segment on the X and Y chromosomes in primates. Cytogenet Cell Genet 50:201–205

Schlötterer C, Tautz D (1992) Slippage synthesis of simple sequence DNA. Nucleic Acids Res 20:211–215

Schlötterer C, Ritter R, Harr B, Brem G (1998) High mutation rate of a long microsatellite allele in *Drosophila melanogaster* provides evidence for allele-specific mutation rates. Mol Biol Evol 15:1269–1274

Stewart EA, McKusick KB, Aggarwal A, et al. (1997) An STS-based radiation hybrid map of the human genome. Genome Res 7:422–433

Sturzeneker R, Haddad LA, Bevilacqua RA, Simpson AJ, Pena SD (1998) Polarity of mutations in tumor-associated microsatellite instability. Hum Genet 102:231–235

Talbot CC Jr, Avramopoulos D, Gerken S, et al. (1995) The tetranucleotide repeat polymorphism D21S1245 demonstrates hypermutability in germline and somatic cells. Hum Mol Genet 4:1193–1199

Tautz D (1989) Hypervariability of simple sequences as a general source for polymorphic DNA markers. Nucleic Acids Res 17:6463–6471

Tyler-Smith C, Oakey RJ, Larin Z, et al. (1993) Localization of DNA sequences required for human centromere function through an analysis of rearranged Y chromosomes. Nature Genet 5:368–375

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

Valdes AM, Slatkin M, Freimer NB (1993) Allele frequencies at microsatellite loci: The stepwise mutation model revisited. Genetics 133:737–749

Weber JL (1990) Informativeness of human $(dC\text{-}dA)_n$, $(dG\text{-}dT)_n$ polymorphisms. Genomics 7:524–530

Weber JL, May PE (1989) Abundant class of human DNA polymorphism which can be typed using the polymerase chain reaction. Am J Hum Genet 44:388–396

Weber JL, Wong C (1993) Mutation of human short tandem repeats. Hum Mol Genet 2:1123–1128

Zerjal T, Dashnyam B, Pandya A, et al. (1997) Genetic relationships of Asians and Northern Europeans, revealed by Y-chromosomal DNA analysis. Am J Hum Genet 60:1174–1183

Zhivotovsky LA, Feldman MW (1995) Microsatellite variability and genetic distances. Proc Natl Acad Sci USA 92:11549–11552