

An Extensive Analysis of Y-Chromosomal Microsatellite Haplotypes in Globally Dispersed Human Populations

Manfred Kayser,¹ Michael Krawczak,² Laurent Excoffier,³ Patrick Dieltjes,⁵ Daniel Corach,⁶ Vincente Pascali,⁷ Christian Gehrig,⁴ Luigi F. Bernini,⁵ Jørgen Jespersen,⁸ Egbert Bakker,⁵ Lutz Roewer,⁹ and Peter de Knijff⁵

¹Department of Evolutionary Genetics, Max Planck Institute for Evolutionary Anthropology, Leipzig; ²Institute of Medical Genetics, University of Wales College of Medicine, Cardiff; ³Genetics and Biometry Lab, Department of Anthropology and Ecology, University of Geneva, and ⁴Institute of Legal Medicine, Geneva; ⁵MGC-Department of Human and Clinical Genetics, Leiden University Medical Center, Leiden; ⁶Servicio de Huellas Digitales Genéticas, Universidad Buenos Aires, Buenos Aires; ⁷Istituto di Medicina Legale, Università Cattolica del S. Cuore, Rome; ⁸Department of Thrombosis Research, University of South Denmark, Esbjerg; and ⁹Institut für Rechtsmedizin, Charité, Humboldt-Universität Berlin, Berlin

The genetic variance at seven Y-chromosomal microsatellite loci (or short tandem repeats [STRs]) was studied among 986 male individuals from 20 globally dispersed human populations. A total of 598 different haplotypes were observed, of which 437 (73.1%) were each found in a single male only. Population-specific haplotype-diversity values were .86–.99. Analyses of haplotype diversity and population-specific haplotypes revealed marked population-structure differences between more-isolated indigenous populations (e.g., Central African Pygmies or Greenland Inuit) and more-admixed populations (e.g., Europeans or Surinamese). Furthermore, male individuals from isolated indigenous populations shared haplotypes mainly with male individuals from their own population. By analysis of molecular variance, we found that 76.8% of the total genetic variance present among these male individuals could be attributed to genetic differences between male individuals who were members of the same population. Haplotype sharing between populations, Φ_{ST} statistics, and phylogenetic analysis identified close genetic affinities among European populations and among New Guinean populations. Our data illustrate that Y-chromosomal STR haplotypes are an ideal tool for the study of the genetic affinities between groups of male subjects and for detection of population structure.

Introduction

Variability at microsatellite or short tandem repeat (STR) loci is being used, in various species, for linkage analysis (Weissenbach et al. 1992), individual identification (Hammond et al. 1994), and population-genetic analyses (for an overview, see Bruford and Wayne 1993). Autosomal STR loci have also been successfully applied to reconstruct human evolutionary history (Bowcock et al. 1994; Goldstein et al. 1995a, 1995b; Mountain and Cavalli-Sforza 1997). The resulting phylogenetic trees reveal evolutionary relationships similar to those based on mtDNA sequence variation, and it has been a long-lasting wish to add to these trees the one strictly based on Y chromosome-specific markers. Approximately 5 years ago, the number of published STR loci on the human Y chromosome was <15 (Roewer et al. 1992; Chen et al. 1994;

Mathias et al. 1994; Jobling and Tyler-Smith 1995), and only one Y chromosome-specific minisatellite was known (Jobling 1994). The number of known Y-chromosomal single-nucleotide polymorphisms (SNPs) was <10 (Nakahori et al. 1989; Jobling 1994; Mathias et al. 1994; Seielstad et al. 1994; Hammer 1995; Whitfield et al. 1995), and only one *Alu*-insertion polymorphisms had been discovered (Hammer 1994). This picture, however, has totally changed, owing to the recent introduction of many Y SNPs (Underhill et al. 1997, 2000) and STRs (White et al. 1999; Ayub et al. 2000), and additional new markers are to be expected.

Recently, Y-STR variability has been used both for the dating of SNP mutations, in order to draw conclusions about the origins and history of human populations (Underhill et al. 1996; Zerjal et al. 1997; Bianchi et al. 1998; Lahermo et al. 1999; Kayser et al. 2000a, 2001), and for human identification in forensic casework (Kayser et al. 1997; Prinz et al. 1997; Honda et al. 1999). Nevertheless, global studies of Y-STR variability are still rare (Deka et al. 1996; Seielstad et al. 1999; Jorde et al. 2000), and most of these analyses have been based on combined single-locus information. The major advantage of analyzing the nonrecombining

Received December 7, 2000; accepted for publication January 29, 2001; electronically published March 16, 2001.

Address for correspondence and reprints: Dr. Peter de Knijff, MGC-Department of Human and Clinical Genetics, Leiden University Medical Center, P.O. Box 9503, 2300 RA Leiden, The Netherlands. E-mail: P.de_knijff@lumc.nl

© 2001 by The American Society of Human Genetics. All rights reserved. 0002-9297/2001/6804-0021\$02.00

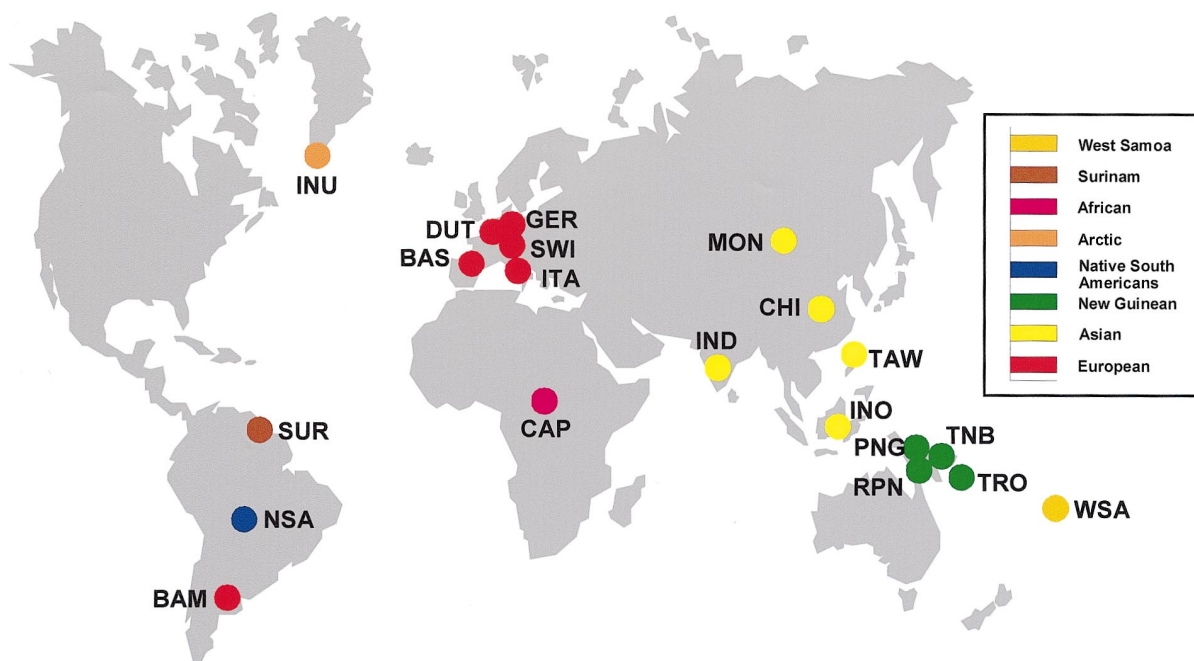


Figure 1 Approximate geographic locations of 20 study populations. The color codings differentiate the eight groups of populations used for AMOVA analyses.

part of the Y chromosome is that single-locus information can be used to construct compound haplotypes that allow male lineages to be characterized in a much more detailed fashion. It has been questioned whether the tracing of human migration history can be achieved solely on the basis of Y STRs (de Knijff et al. 1997). Intuitively, since STRs harbor a much higher degree of genetic variation than do SNPs, this appears to be possible. However, preliminary studies have revealed that, most likely because of the relatively high mutation frequency of Y STRs (Heyer et al. 1997; Kayser et al. 2000b), Y haplotypes can be shared identical by state (IBS) that are not identical by descent (IBD) (de Knijff 2000b). To explore the use of Y-STR haplotypes in more detail, we analyzed the genetic variance, at seven loci, among 986 male individuals from 20 globally dispersed human populations, some of which are closely related (e.g., Dutch, Germans, Swiss, and Italians) and others of which are distantly related (e.g., native South Americans, central African Pygmies, and Papua New Guineans). This enabled us to investigate the utility that these Y STRs have for different evolutionary time scales.

Subjects, Material, and Methods

DNA Samples

DNA samples were obtained from 986 male individuals from 20 different populations: 470 either from Eu-

rope or of European ancestry (88 Germans [GER], 88 Dutch [DUT], 64 Swiss [SWI], 100 Italians [ITA], 100 Buenos Aires Metropolitan [BAM], and 30 Basques [BAS]); 200 from Asia (36 Han-Chinese [CHI], 25 pooled samples from Indians [IND] including 15 Dora and 10 Reddi, 69 pooled samples from Indonesians [INO] including 56 samples from a rural area near Jakarta and 13 from Banjarmasin on South Borneo, 40 Khalkh-Mongolians [MON], and 30 pooled samples from Taiwanese Aborigines [TAW] including 13 Ami, 6 Atayal, 5 Bunun, and 6 Paiwan); 113 from New Guinea (26 Papua New Guineans [PNG], 12 Roro [RPN] from the southern coast of Papua New Guinea, 59 Trobriand Islanders [TRO], and 16 Tolai from New Britain [TNB]); 46 from America (pooled samples from native South Americans [NSA] including 6 Wichi, 12 Tehuelche and 16 Mapuche from Argentina, and 12 Yanomami from Brazil); 62 from the Arctic (Inuit [INU] from Nanortalik in southwestern Greenland; and 31 from Africa (central African Pygmies [CAP])). In addition, 54 Surinamese (SUR) and 10 male individuals from West Samoan (WSA) were included. Figure 1 illustrates the geographical location of all 20 populations. Paternal family history was studied, and care was taken to type unrelated male individuals only.

Genetic Screening

All male individuals were genotyped for six tetranucleotide Y STRs—DYS19, *DYS389I*, *DYS389II*,

DYS390, DYS391, and DYS393 and for one trinucleotide Y STR—DYS392—as described in detail elsewhere (Kayser et al. 1997). It may be noteworthy here that DYS394 is a synonym for DYS19 and refers to the same Y-STR locus amplified with a different primer set. DNA sequencing of a number of alleles amplified with the DYS394, DYS19, and alternative primers revealed that all nucleotide differences between the DNA sequences of DYS394 and DYS19 reported in GENBANK are due to sequencing errors (M.K., unpublished data). According to the recommendations of the International Society of Forensic Genetics, the alleles are designated in terms of the number of variable repeats that they contain (see “DNA Recommendations—1994 Report Concerning Further Recommendations of the DNA Commission of the ISFH Regarding PCR-Based Polymorphisms in STR (Short Tandem Repeats) Systems” 1995). Simultaneous running of allelic ladders for all loci ensured consistent allele designation in the different laboratories. Further information on the markers is available, on request, from the authors and can be found at the Forensic Laboratory for DNA Research and the Y-STR Haplotype Reference Database. Depending on the type of analysis, haplotypes were defined by either combining the number of variable repeats (i.e., the allele name) or using a coded version (i.e., calling the shortest observed allele of each locus “1” and numbering upward for each additional repeat unit), for each locus.

Statistical and Phylogenetic Analysis

For each of the 20 populations, locus-specific allele frequencies were estimated by simple gene counting. The standard error (SE) of allele frequencies was calculated as $SE(p_i) = \sqrt{[(1-p_i)p_i]/N}$, where p_i denotes the frequency of the i th allele at any given locus and N equals the total number of individuals screened at this locus.

The intrapopulation locus-specific variance, V_L , was estimated by $V_L = [1/(n-1)]\sum_{i=1}^n (X_i - \bar{X})^2$, where X_i is the size of the allele on the i th chromosome, \bar{X} is estimated as $\bar{X} = (1/n)\sum_{i=1}^n X_i$, and n is the number of chromosomes sampled for the population. The intrapopulation genetic variance, V_p , was subsequently estimated by averaging across m loci, as $V_p = (1/m)\sum_{j=1}^m V_{L_j}$. Subsequently, a Y-STR haplotype comprising seven loci was constructed for each male, for all analyses, with the loci in the order DYS19-DYS389I-DYS389II-DYS390-DYS391-DYS392-DYS393. An unbiased estimate of haplotype diversity, h , and its variance, $V(h)$, were calculated according to the method of Nei (1987, formulas 8.5 and 8.13 therein). The SE of h , $SE(h)$, was calculated by taking the square root of $V(h)$. Single-locus gene-diversity values were calculated in the same way. Numbers of shared haplotypes were determined for each of the 190 possible pairs of populations, by a simple count-

ing scheme. The probability of identity, p , between these 190 population pairs (which reflects the haplotype-sharing index) was estimated, according to the method of Melton et al. (1995), as $p = \sum_{i,j} x_i x_j$, where x_i and x_j are, respectively, the frequencies of a haplotype in populations i and j , summed over the n haplotypes in the two populations.

In an initial pairwise analysis, allele frequencies at the seven loci were compared between all pairs of populations, by the Fisher exact test-based, genic-comparison option included in GENEPOP (Raymond and Rousset 1995).

Genetic relationships between the different populations, based on the seven-locus Y-STR haplotypes, were further explored by analysis of molecular variance (AMOVA), as implemented in Arlequin (see the ARLEQUIN: A Software For Population Genetic Data Analysis website). AMOVA allows a hierarchic analysis of three genetic-variance components—those due to genetic differences (i) between individuals within populations, (ii) between populations within groups, and (iii) between groups (Excoffier et al. 1992; Excoffier and Smouse 1994). For AMOVA, the following eight groups, containing all 20 populations were defined: (1) all European populations (the GER, the DUT, the SWI, the ITA, the BAM, and the BAS); (2) all Asian populations (the CHI, the IND, the INO, the MON, and the TAW), excluding the New Guinean samples; (3) all mainland and island New Guinean samples (the PNG, the RPN, the TRO, and the TNB); (4) the NSA; (5) the INU; (6) the CAP; (7) the SUR; and (8) the WSA. The genetic structure among our population samples was analyzed with consideration for the molecular differences between individual haplotypes, in addition to differences in haplotype frequencies, resulting in estimates of Φ_{ST} (or R_{ST}), an F_{ST} analogue. Significance levels of the genetic-variance components as well as Φ_{ST} values were estimated by use of 10,000 permutations.

Pairwise genetic distances between populations were computed as a linearization of Φ_{ST} values; that is, $D = \Phi_{ST}/(1 - \Phi_{ST})$ (Slatkin 1995). On the basis of these adjusted Φ_{ST} values, a neighbor-joining (NJ) tree was constructed by PHYLIP, version 3.57c. The resulting tree was visualized by TREEVIEW 1.6.1 (see Rod Page's Home Page). The same adjusted Φ_{ST} values were used for a multidimensional-scaling analysis (Kruskal 1964).

An NJ tree connecting all 598 different Y-STR haplotypes was constructed by means of PHYLIP. The resulting complex NJ tree was reduced to 17 clusters of related haplotypes, on the basis of their positions in this tree. For all 17 major clusters, the relative contribution of haplotypes from each of the eight population groups (as defined above for AMOVA analyses) was estimated. This allows a comparison of the distribution of region-specific haplotypes between the 17 clusters and, at

the same time, shows the relative contribution of each of the 17 clusters to the total number of haplotypes. On the basis of all seven-locus haplotypes with a total frequency of 0.5% (i.e., observed in at least five male individuals) or higher ($n = 24$), a modified reduced median network (Bandelt et al. 1995) was constructed. Comparisons of allele-frequency distributions between regions were performed by nonparametric exact-test procedures embedded in the program StatXact (Cytel-Software). Significance levels were estimated by the Monte-Carlo simulation mode with 10,000 randomizations.

Results

Allele Frequencies and Haplotype Distributions

A global genetic survey was performed with respect to seven Y-STR loci. Human populations were chosen so as to encompass two subsets of closely related groups of male individuals and a number of genetically more distinct groups of male individuals. In total, we screened 986 male individuals from 20 different populations.

Across regions, the allele-frequency distribution differs for most loci (see Appendix A), although the results are not consistent between loci. Figure 2 illustrates the combined allele-frequency distribution for all seven loci. With the exception of DYS392, all loci have a unimodal distribution with one frequent allele and with the less-frequent alleles differing from the most-frequent allele

Table 1

Analysis of Differentiation, by Fisher's Exact Test, between 20 Populations, Based on Single-Locus Allele Frequencies of Seven Y-STR Loci

Locus	NO. (%) OF POPULATION PAIRS			
	Significant			Not Significant: $P > .05$
	$P < .01$	$.01 < P < .05$	$P < .05$	
DYS390	173 (91.1)	10 (5.3)	183 (96.3)	7 (3.7)
DYS392	170 (89.5)	9 (4.7)	179 (94.2)	11 (5.8)
DYS19	170 (89.5)	5 (2.6)	175 (92.1)	15 (7.9)
DYS389II	152 (80.0)	13 (6.8)	165 (86.8)	25 (13.2)
DYS393	137 (72.1)	14 (7.4)	151 (79.5)	39 (20.5)
DYS389I	114 (60.0)	14 (7.4)	128 (67.4)	62 (32.6)
DYS391	83 (43.7)	32 (16.8)	115 (60.5)	75 (39.5)

by a single repeat unit. DYS392 is the only locus that clearly has a bimodal allele-frequency distribution.

With regard to the possible 190 pairwise allele-frequency comparisons between populations, DYS19, DYS392, and DYS390 have highly significant differences ($P < .01$) in ~90% of all comparisons, whereas for DYS391 ~40% of all comparisons ($n = 75$) were not significant (table 1).

Locus-specific genetic differences between populations are also reflected by the intrapopulation (locus-specific) genetic variances (fig. 3). Among the CAP, for example, only two of eight DYS392 alleles were observed, resulting in a markedly reduced genetic variance for this locus. In contrast, the same population harbors seven of

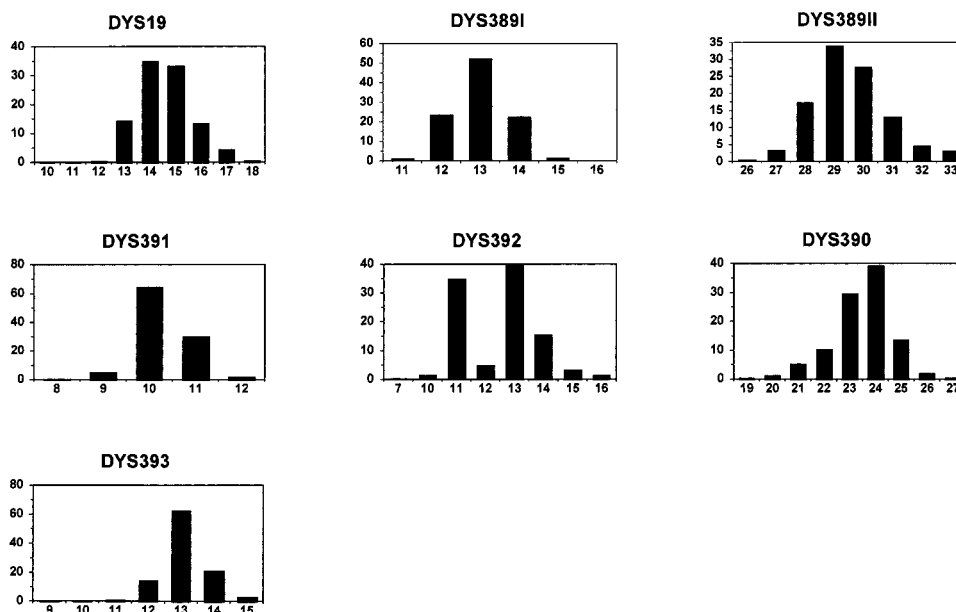


Figure 2 Allele-frequency distributions of the seven Y-STR loci, among all 986 male individuals combined. For each locus, the allelic designation (in number of repeats) is indicated on the X-axis, and the observed frequency (in %) is shown on the Y-axis.

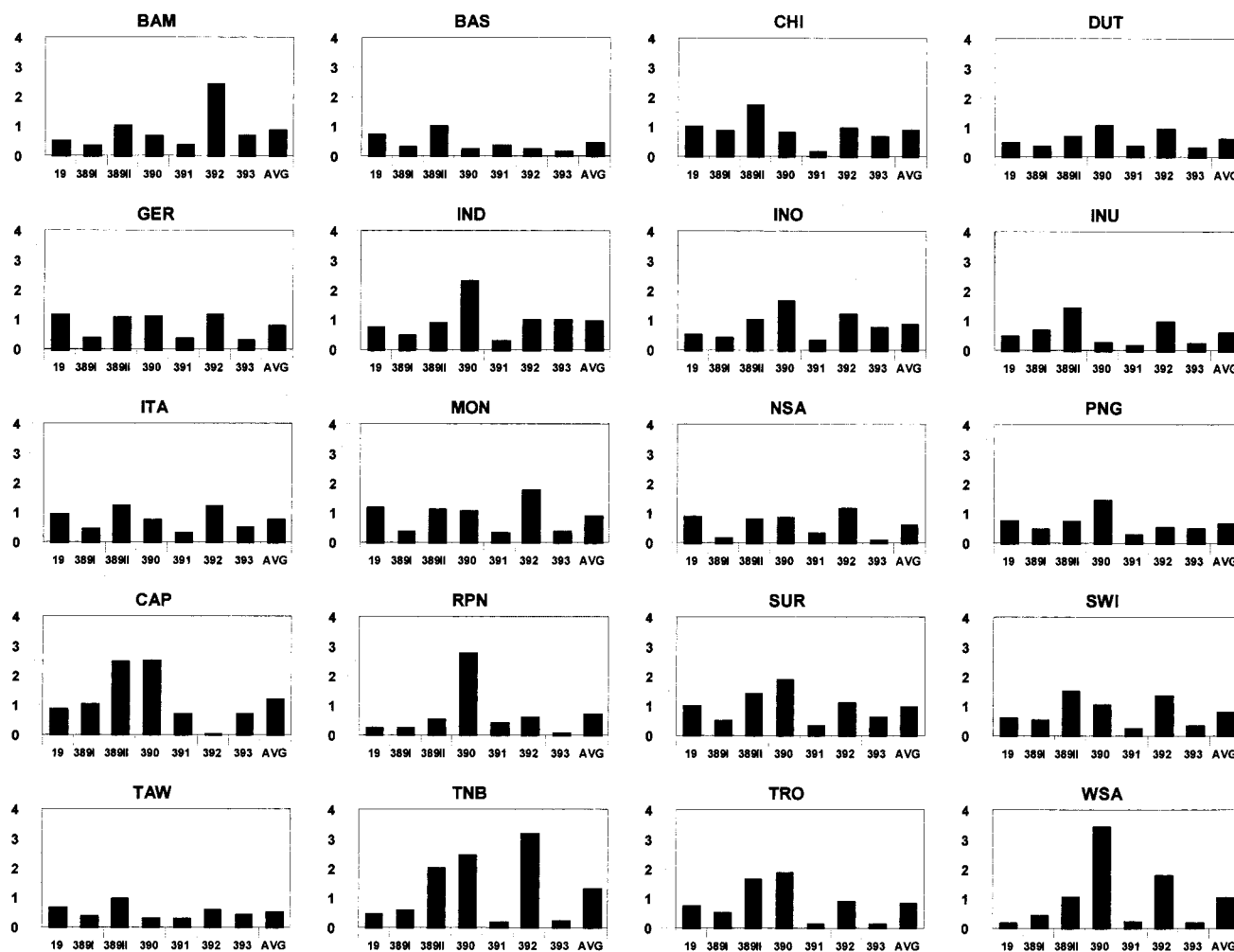


Figure 3 Estimated genetic variances for each the seven Y-STR loci, and the average across the seven loci, within each of the 20 different population samples.

the eight alleles at *DYS389II*—hence the high genetic variance for this locus. Similar intrapopulation locus-specific variance differences can be observed among the IND, the RPN, and the TNB. Among the European populations (except the BAM), the genetic variance is rather evenly distributed across all loci.

A total of 598 different compound Y-STR haplotypes were observed among 986 individuals (see Appendix B). Haplotype-diversity values varied from .99, in the GER, the SWI, the ITA, the CHI, and the SUR, to .86, in the RPN (table 2). A total of 437 haplotypes (73.1%) were observed in just a single male in a single population (these haplotypes are designated “single unique”). A further 68 (11.4%) haplotypes were shared only by male individuals within a single population (these haplotypes are designated “multiple unique”). Single- and multiple-unique haplotypes combined (i.e., the “total unique”) are the number of haplotypes that are specific to a single population (table 2). The remaining haplotypes ($n = 93$

[15.5%]) were observed in multiple male individuals in multiple populations (i.e., are nonunique) and thus are shared by populations. Not surprisingly, the number of single-unique haplotypes was not distributed evenly across the populations. Among the CHI and the TNB, a large number (79%) of male individuals displayed a single unique Y haplotype, whereas the INU (23%) and the BAS (29%) displayed the lowest level of single-unique haplotypes. The INU and the CAP displayed the highest rates of multiple-unique haplotypes (32% and 33%, respectively). High ($\geq 90\%$) frequencies of population-specific haplotypes were observed among the MON, the NSA, and the TNB. Low ($\leq 50\%$) frequencies were noted only in a number of European groups (the GER, the DUT, and the SWI). The BAS, with only 35% population-specific haplotypes, displayed the lowest number of unique haplotypes. For all population samples except the GER, the DUT, the SWI, and the BAS, the proportion of nonunique haplotypes was either higher than or nearly equal to that

Table 2

Y-STR Haplotype-Sharing Statistics

	GER	DUT	SWI	ITA	BAM	BAS	CHI	IND	INO	MON
No. of individuals	88	88	64	100	100	30	36	25	69	40
No. of haplotypes	77	65	51	82	76	17	34	16	53	29
Discrimination (%)	87.5	73.9	79.7	82.0	76.0	56.7	94.4	64.0	76.8	72.5
Haplotype Class:	ff									
Unique:										
Single:										
No.	37	32	20	49	49	5	27	9	38	21
Proportion	.48	.49	.39	.60	.64	.29	.79	.56	.72	.72
Multiple unique:										
No.	2	0	2	4	5	1	1	4	5	5
Proportion	.03	.00	.04	.05	.07	.06	.03	.25	.09	.17
Total unique:										
No.	39	32	22	53	54	6	28	13	43	26
Proportion	.51	.49	.43	.65	.71	.35	.82	.81	.81	.90
Nonunique:										
No.	38	33	29	29	22	11	6	3	10	3
Proportion	.49	.51	.57	.35	.29	.65	.18	.19	.19	.10
Ratio (unique:nonunique)	1.03	.97	.76	1.83	2.45	.54	4.67	4.33	4.30	8.67
Haplotype diversity	.9963	.983	.9921	.9941	.9877	.9379	.9968	.95	.9868	.9705
Haplotype-diversity SE	.0012	.004	.002	.0016	.0033	.0162	.0021	.0147	.0036	.0099
	TAW	PNG	RPN	TRO	TNB	NSA	INU	CAP	SUR	WSA
No. of individuals	30	26	12	59	16	46	62	31	54	10
No. of haplotypes	25	22	7	40	14	34	22	18	47	8
Discrimination (%)	83.3	84.6	58.3	67.8	87.5	73.9	35.5	58.1	87.0	80.0
Haplotype Class:										
Unique:										
Single:										
No.	18	15	4	22	11	25	5	9	35	6
Proportion	.72	.68	.57	.55	.79	.74	.23	.50	.74	.75
Multiple:										
No.	1	0	0	10	2	6	7	6	6	1
Proportion	.04	.00	.00	.25	.14	.18	.32	.33	.13	.13
Total unique:										
No.	19	15	4	32	13	31	12	15	41	7
Proportion	.76	.68	.57	.80	.93	.91	.55	.83	.87	.88
Nonunique:										
No.	6	7	3	8	1	3	10	3	6	1
Proportion	.24	.32	.43	.20	.07	.09	.45	.17	.13	.13
Ratio (unique/nonunique)	3.17	2.14	1.33	4.00	13.00	10.33	1.20	5.00	6.83	7.00
Haplotype diversity	.9839	.9846	.8636	.9825	.9833	.9826	.9053	.9462	.9951	.9333
Haplotype-diversity SE	.0069	.0073	.045	.0037	.0096	.0047	.014	.0148	.0016	.041

of unique haplotypes. Of the 190 pairwise population comparisons, 121 (63.7%) had no shared haplotype, 35 (18.4%) had one, 10 (5.3%) had two, 4 (2.1%) had three, 4 (2.1%) had four, and 16 (8.4%) had five or more (table 3).

AMOVA and Genetic Distances

For AMOVA, populations from the same geographic region were pooled together to form eight groups: (1) all European populations (the GER, the DUT, the SWI, the ITA, the BAM, and the BAS); (2) all Asian populations (the CHI, the IND, the INO, the MON, and the TAW) except the New Guinean samples; (3) all mainland and island New Guinean samples (the PNG, the RPN, the TRO, and the TNB); (4) the NSA; (5) the INU; (6)

the CAP; (7) the SUR; and (8) the WSA. Using AMOVA, we estimated the relative contribution to the total observed genetic variance of (i) the genetic variance between individuals within populations, (ii) the genetic variance between populations within groups, and (iii) the genetic variance between groups (table 4). The contribution of genetic variance between populations, in light of molecular differences between haplotypes and differences in haplotype frequencies combined (i.e., Φ_{ST}), was 23.2% ($P < .00001$). Thus, by far, most (76.8%) of the genetic variance present among our male individuals could be explained by intrapopulation differences, whereas only 16.8% of the genetic variance was due to genetic differences between groups.

The matrix of pairwise Φ_{ST} values is shown in table

Table 3

Number of Shared Haplotypes (below the Diagonal) and Probability of Identity (above the Diagonal), for All 190 Possible Population Pairings

	GER	DUT	SWI	ITA	BAM	BAS	CHI	IND	INO	MON	TAW	PNG	RPN	TRO	TNB	NSA	INU	CAP	SUR	WSA
GER		.0089	.0051	.0038	.0027	.0072	.0003	.0000	.0000	.0003	.0000	.0004	.0047	.0000	.0000	.0015	.0031	.0011	.0008	.0000
DUT	20		.0096	.0048	.0083	.0182	.0003	.0005	.0000	.0000	.0000	.0000	.0028	.0002	.0000	.0044	.0073	.0000	.0029	.0000
SWI	17	15		.0056	.0081	.0120	.0004	.0000	.0000	.0000	.0000	.0000	.0000	.0003	.0000	.0041	.0033	.0000	.0009	.0000
ITA	13	7	8		.0053	.0117	.0000	.0000	.0000	.0003	.0000	.0004	.0000	.0002	.0000	.0033	.0027	.0000	.0007	.0000
BAM	5	8	7	7		.0167	.0003	.0000	.0009	.0000	.0003	.0023	.0000	.0008	.0000	.0065	.0044	.0000	.0000	.0000
BAS	5	8	5	4	6		.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0087	.0065	.0000	.0000	.0000
CHI	1	1	1	1	0	0		.0000	.0004	.0014	.0019	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0028
IND	0	1	0	0	0	0	0		.0035	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
INO	0	0	0	0	2	0	1	2		.0000	.0034	.0067	.0000	.0052	.0000	.0000	.0000	.0005	.0000	.0000
MON	1	0	0	1	0	0	1	0	0		.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0024	.0000	.0000
TAW	0	0	0	0	1	0	2	0	3	0		.0000	.0000	.0023	.0000	.0000	.0005	.0000	.0000	.0000
PNG	1	0	0	1	3	0	0	0	4	0	0		.0000	.0026	.0000	.0000	.0000	.0000	.0000	.0000
RPN	2	1	0	0	0	0	0	0	0	0	0	0		.0000	.0000	.0000	.0000	.0000	.0000	.0000
TRO	0	1	1	1	2	0	0	0	2	0	2	1	0		.0011	.0000	.0003	.0000	.0003	.0000
TNB	0	0	0	0	0	0	0	0	0	0	0	0	0	1		.0000	.0000	.0000	.0000	.0000
NSA	1	1	2	2	3	1	0	0	0	0	0	0	0	0	0		.0032	.0000	.0000	.0000
INU	5	4	4	5	1	1	0	0	0	0	1	0	0	1	0	1		.0000	.0006	.0000
CAP	1	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0		.0006	.0000
SUR	1	1	2	2	0	0	0	0	0	0	0	0	0	1	0	0	1	1		.0000
WSA	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	

5. The lowest interpopulation variances ($\Phi_{ST} < .05$) were observed between some European populations (i.e., between the SWI and the DUT, between the SWI and the ITA, and between the BAM and the BAS) and between a number of Asian and Papua New Guinean populations—that is, between pairs of closely related populations; in contrast, high values ($\Phi_{ST} > .5$) were observed between pairs of distinct populations, such as INU/WSA, INU/CAP, or IND/INU. For most of the pairwise population comparisons, the interpopulation differences were significant. Only nine population pairs have non-significant ($P > .05$) Φ_{ST} values; three of those pairs include only European populations, and the remaining six include only populations from Southeast Asia, mainly New Guinea.

Phylogenetic Analysis

On the basis of the linearized Φ_{ST} distances, an NJ tree was drawn (fig. 4). When the tree topology and the Φ_{ST} significances (table 5) are compared, the NJ tree provides a reasonable “fit,” with most of the nonsignificant population differences corresponding to tight clusters, one containing the European population samples and the other containing the Asian and Papua New Guinean populations. Only the close clustering of the BAS with both the INU and the NSA and the close clustering of the MON with the GER seem counterintuitive. An almost identical picture was obtained by multidimensional-scaling analysis (fig. 5).

An alternative way of presenting our results, instead of focusing on differences between populations, is to

concentrate on the genetic relationships between haplotypes. An NJ tree (fig. 6) was constructed including all 598 different seven-locus Y-STR haplotypes observed among the 986 male individuals from all 20 populations. We defined 17 major clusters of related haplotypes, on the basis of their positions within the tree. Every major cluster was analyzed with respect to the geographic origin of the haplotypes that it contains. As was done for the AMOVA analyses, all 20 populations were grouped into eight classes. We find that the haplotypes observed among New Guineans, for example, are primarily restricted to a number of closely related haplotype clusters (1, 11, and 13–15). Also, the African haplotypes are preferentially present in just three clusters (4, 6, and 7), and cluster 4 also contains a large proportion of SUR haplotypes. Cluster 12 contains a very large proportion of Arctic haplotypes, together with a considerable number of Amerindian haplotypes. On the other hand, haplotypes observed among Europeans are present, in all 17 clusters, at low frequencies. A modified reduced me-

Table 4

AMOVA in 20 Populations, Based on Seven Y-STR Loci

SOURCE OF MOLECULAR VARIATION	Φ_{ST} STATISTIC	
	Variance (%)	P
Between groups	16.8	<.00001
Between populations within groups	6.4	<.00001
Between populations	23.2	<.0001
Within populations	76.8	<.00001

dian network connecting only those haplotypes that were observed at a frequency $\geq 0.5\%$ is shown in figure 7. On the basis of this network, it appears that the majority of the INU-specific haplotypes and European haplotypes form two distinct clusters. It is also noteworthy that these haplotypes are connected by single mutation steps. In contrast, Asian and New Guinean haplotypes are separated by multiple mutation steps and are located in different parts of the network, possibly reflecting a more ancient origin of these groups.

Discussion

The major advantage of analysis of polymorphic loci from the nonrecombining part of the Y chromosome is that they facilitate a simple construction of highly informative compound haplotypes that will characterize each distinct male lineage in detail. In addition, because of its (almost) strictly paternal mode of inheritance and lack of recombination, the Y chromosome is extremely sensitive to genetic drift. These two characteristics render the Y chromosome potentially very informative for the study of human evolution. Nevertheless, the tracing of human migration events solely on the basis of Y STRs has been questioned (de Knijff et al. 1997); in particular, because of the relatively high mutation frequency of Y STRs (Heyer et al. 1997; Kayser et al. 2000b), Y haplotypes can be shared IBS without being IBD (de Knijff 2000b). To explore the utility of Y-STR haplotypes in more detail, we analyzed the genetic variance of seven (six tetrameric and one trimeric) Y-STR loci

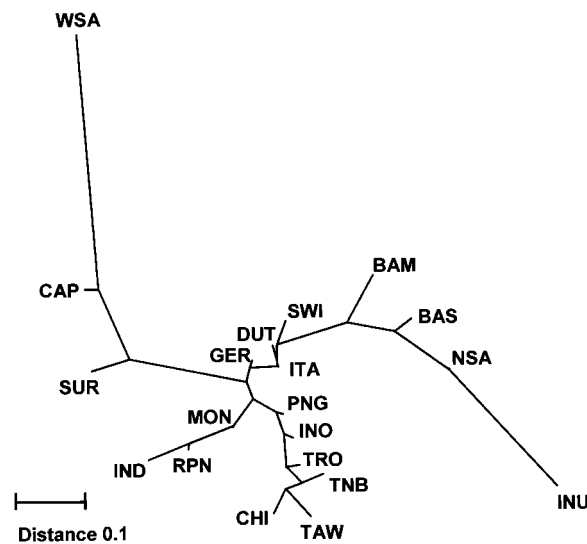


Figure 4 Unrooted NJ tree, based on linearized Φ_{ST} distances derived from AMOVA connecting all 20 population samples.

among 986 male individuals from 20 globally dispersed human populations.

All loci used in this study are located in the nonrecombining part of the human Y chromosome and thus are completely linked, lacking any recombination. Nevertheless, we observed locus-specific differences in both the intra- and interpopulation genetic variance (fig. 1 and Appendix A). These observations can easily be explained by mutation-rate differences between the loci

Table 5

Φ_{ST} Values (below the Diagonal) and Their Significance Levels (above the Diagonal)

	GER	DUT	SWI	ITA	BAM	BAS	CHI	IND	INO	MON	TAW	TNG	RPN	TRO	TNB	NSA	INU	CAP	SUR	WSA
GER		.000	.002	.000	.000	.000	.000	.000	.000	.022	.000	.000	.003	.000	.000	.000	.000	.000	.000	.000
DUT	.079		<u>.174</u>	.006	.000	.000	.000	.000	.000	.000	.000	.000	.001	.000	.000	.000	.000	.000	.000	.000
SWI	.038	.007		.152	.027	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
ITA	.049	.023	.008		.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
BAM	.112	.057	.022	.061		<u>.079</u>	.000	.000	.000	.000	.000	.000	.002	.000	.000	.000	.000	.000	.000	.000
BAS	.150	.169	.090	.152	.022		.000	.000	.000	.000	.000	.000	.001	.000	.000	.000	.000	.000	.000	.000
CHI	.225	.208	.189	.231	.167	.226		.000	.021	.000	<u>.191</u>	.005	.002	.022	<u>.668</u>	.000	.000	.000	.000	.000
IND	.188	.217	.221	.260	.274	.343	.177		.006	.000	.000	.000	.006	.000	.013	.000	.000	.000	.000	.000
INO	.149	.119	.126	.172	.152	.216	.033	.079		.000	.032	.022	.008	.000	<u>.130</u>	.000	.000	.000	.000	.000
MON	.030	.134	.107	.103	.183	.246	.184	.103	.112		.000	.000	.022	.000	<u>.001</u>	.000	.000	.000	.000	.000
TAW	.296	.261	.248	.305	.221	.359	.016	.201	.030	.275		.000	.000	.004	<u>.413</u>	.000	.000	.000	.000	.000
PNG	.106	.145	.090	.159	.095	.159	.085	.173	.044	.124	.128		.018	<u>.061</u>	.031	.000	.000	.000	.000	.002
RPN	.105	.206	.155	.214	.170	.195	.143	.069	.096	.091	.252	.124		.000	.023	.000	.000	.000	.000	.001
TRO	.220	.229	.178	.237	.142	.181	.036	.263	.085	.218	.071	.031	.189		<u>.050</u>	.000	.000	.000	.000	.000
TNB	.242	.190	.177	.229	.142	.204	.015	.131	.029	.189	.000	.075	.128	.049		.000	.000	.000	.000	.005
NSA	.327	.263	.202	.268	.085	.146	.229	.441	.264	.398	.289	.220	.370	.162	.179		.000	.000	.000	.000
INU	.342	.366	.264	.325	.165	.166	.435	.524	.420	.440	.506	.324	.423	.335	.397	.191		.000	.000	.000
CAP	.172	.265	.227	.220	.324	.411	.308	.227	.243	.132	.366	.231	.294	.325	.293	.486	.538		.002	.001
SUR	.102	.135	.087	.070	.146	.243	.269	.304	.229	.145	.342	.165	.292	.244	.258	.316	.356	.100		.001
WSA	.468	.503	.433	.452	.427	.588	.371	.481	.400	.446	.469	.380	.547	.336	.301	.498	.606	.264	.263	

NOTE.—Nonsignificant P values ($P > .05$) are underlined; P values between .01 and .05 are in boldface italic.

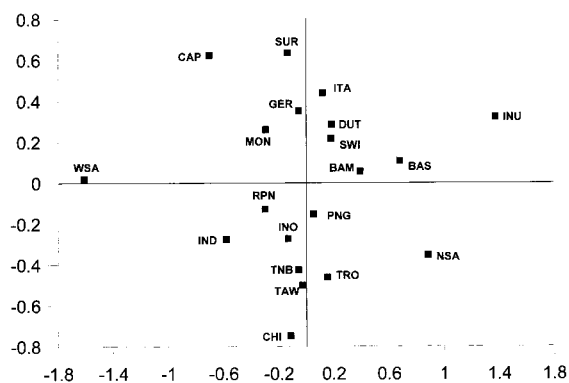


Figure 5 Genetic map based on multidimensional-scaling analysis and a matrix of the pairwise linearized Φ_{ST} distances derived from AMOVA.

and by the different population histories and genetic drifts; for example, DYS390 has the highest variance in 10 of 20 populations (present study) and was found to have the highest mutation rate among 15 Y STRs, including the loci studied here (Kayser et al. 2000b). Locus-specific differences were also demonstrated by a population-differentiation test, in which all but 7 of the 190 population pairs could be significantly distinguished on the basis of DYS390 alone. In contrast, 75 of 190 pairwise comparisons were not significant for DYS391 (table 1). This supports the observations, by others (Jorde et al. 2000), that, especially among European populations, Y STRs are very powerful in the detection of genetic differences between populations, compared with autosomal STRs. This can be attributed to the greater sensitivity of nonrecombining Y-chromosomal markers to founder effects and genetic drift.

Haplotype Analyses

The marked genetic variation of Y-STR haplotypes across global populations is reflected in the haplotype-diversity values, which are $>.98$ in 13 of 20 populations (table 2); values $>.99$ in 3 of the 5 European populations indicate a heterogeneous European gene pool. Haplotype-diversity values were also $>.99$ in the CHI and in the SUR; in the CHI this might be explained by the large population size, in the SUR by the well-known admixture of four distinct populations (i.e., Dutch, African, Asian, and native South American) within the gene pool. Lower haplotype-diversity values were observed in samples from small and isolated populations such as the INU, the BAS, and the CAP. Elsewhere (Perez-Lezaun et al. 1997), a reduced haplotype diversity for the BAS (compared with that in Catalans) has also been described, on the basis of a smaller number of Y microsatellites. The

lowest haplotype diversity, .86 in the RPN, might be due to small sample size and inbreeding. The relatively high (.93) haplotype diversity in the very small ($n = 10$) sample of the WSA seems puzzling, but it is not without precedent (Flint et al. 1999); however, one could also argue that it is simply due to small sample size ($n = 10$). Recently, for the same set of Y STRs that have been used in the present study, it has been shown that, compared with that in 17 eastern Asian/Southeast Asian, Melanesian, and Australian populations, the haplotype diversity in a sample ($n = 28$) from the Cook Islands of Polynesia was low (.89) (Kayser et al. 2001). This was explained by a means of a scenario postulating a recent bottleneck during the colonization of the Pacific. Our results on haplotype diversity are in contrast with those of studies, using different types of markers, that reported the highest genetic diversity in Africa. Recently, this has also been reported for Y STRs, but such differences in non-Africans are not significant (Seielstad et al. 1999; Jorde et al. 2000). In both data sets, the number of African populations pooled to obtain the high diversity values was greater than that for other continents. Especially in the study by Seielstad et al. (1999), in which 25 African populations with small sample sizes were pooled and compared with three to seven populations from other continents, this might have strongly influenced the diversity values obtained. In both studies, diversity was calculated on the basis of combined single-locus data rather than on the basis of haplotype data. If we calculate single-locus gene-diversity values on the basis of our data, we obtain the highest average diversity, .63, for Africa (one population), compared with .60 for Asia (five populations pooled), .56 for Europe (five populations pooled), and .54 for New Guinea (four populations pooled).

Investigation of haplotype sharing (or identity) within populations (multiple-unique haplotypes) and of population-specific haplotypes (single- and multiple-unique haplotypes) allows some insight into population structure (tables 2 and 3). High amounts of haplotype sharing within populations and/or of population-specific haplotypes, such as those observed among the CAP, the INU, the IND, the TRO, and the NSA, indicate small and/or isolated population; on the other hand, high numbers of nonunique haplotypes and consequent haplotype sharing between populations indicate a close relationship between populations. This was found to be true for all European populations studied here. With our current knowledge, it is difficult to say what proportion of STR-haplotype sharing is due to recurrent mutation (i.e., IBS sharing) and what proportion is due to a genuine shared (recent) common ancestry (i.e., IBD sharing); however, it seems plausible to assume that the sharing of one or two haplotypes—as observed here, for example, between the CAP and the GER or between the TRO and the

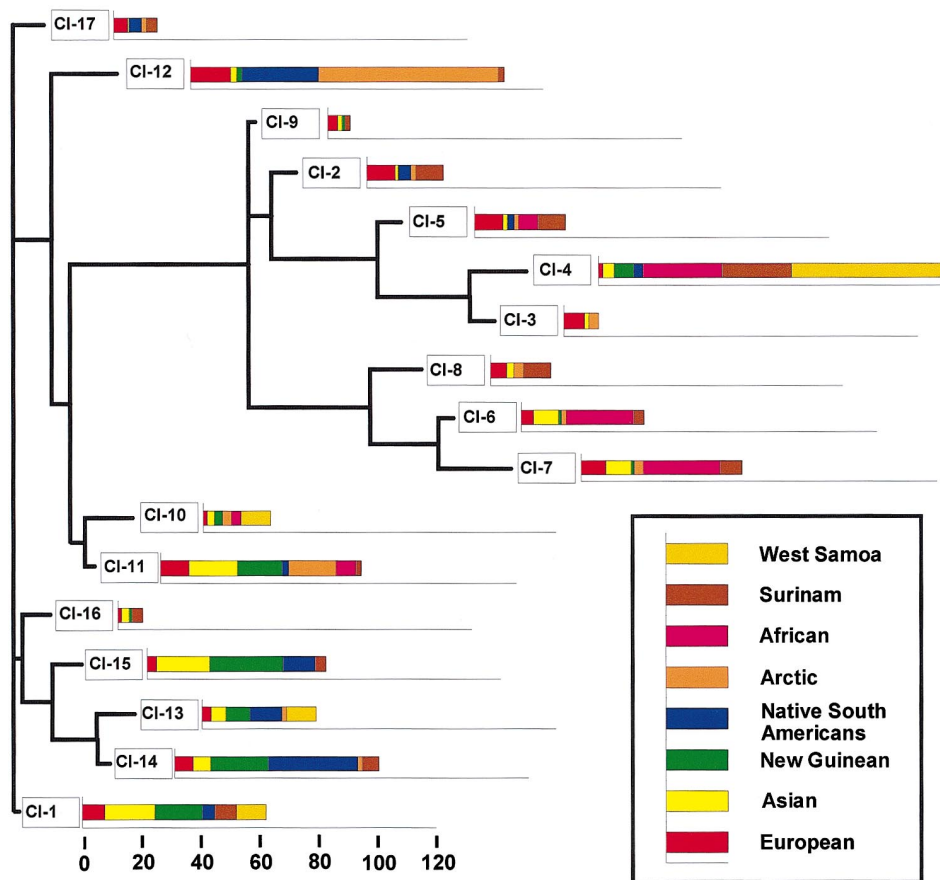


Figure 6 Unrooted NJ tree connecting all 598 distinct seven-locus Y-STR haplotypes. The complex topology of the tree was reduced by delineation of 17 major clusters of related haplotypes. The relative contribution of haplotypes of each of the eight population groups is indicated for each cluster; that is, cluster 1 contains 7.7% of the haplotypes observed among Europeans, 17% of all Asian haplotypes, 15.9% of all New Guinean haplotypes, 4.3% of all the NSA haplotypes, no haplotypes from the Arctic (i.e., the INU) and African (i.e., the CAP) populations, 7.4% of all the SUR haplotypes, and 10% of all the WSA haplotypes. For each of the eight groups of populations, the total sum of haplotypes is 100%; thus, cluster 1 contains 62.3% of 800% (i.e., 7.8%) of all haplotypes. The color codes correspond to the eight groups of populations used for AMOVA and are the same as those used in figures 1 and 7.

INU—reflects IBS sharing rather than IBD sharing. The analysis of additional Y STRs might reveal nonidentity; in contrast, the sharing of ≤ 20 haplotypes between European groups (including the BAS) is more likely to reflect IBD sharing. Interestingly, the BAM share five to eight haplotypes with European groups but share only three with the NSA. This probably reflects their recent European ancestry, combined with little or no admixture with the NSA, which has also been revealed by autosomal markers (Sala et al. 1997). Also, a more recent study of Brazilian populations reports no evidence for male Amerindian admixture, on the basis of Y-SNP analysis (Carvalho-Silva et al. 2001). Interestingly, the INU share either four or five haplotypes with each of the ITA, the SWI, the DUT, and the GER, which might indicate recent European gene flow into Greenland. Within our second group of closely related populations (those from

mainland and island Papua New Guinea), we only sporadically observed haplotype sharing. This can be explained by the much lower sample sizes for the New Guinean groups compared with the European groups.

To study the interpopulation genetic affinities in more detail, we performed an AMOVA, based on Y STR-defined haplotypes (table 4). All but 9 of the 190 population pairs can be differentiated significantly on the basis of the seven-locus Y-STR haplotypes employed (table 5); most of these 9 population pairs for which a nonsignificant P value for Φ_{ST} was obtained include populations in which close relationship can be assumed—that is, among European populations or among New Guinean populations. The results of the haplotype sharing and Φ_{ST} analyses were not always correlated (tables 3 and 5). Of the 15 strictly European population pairs with high numbers of shared haplotypes, 12 have sig-

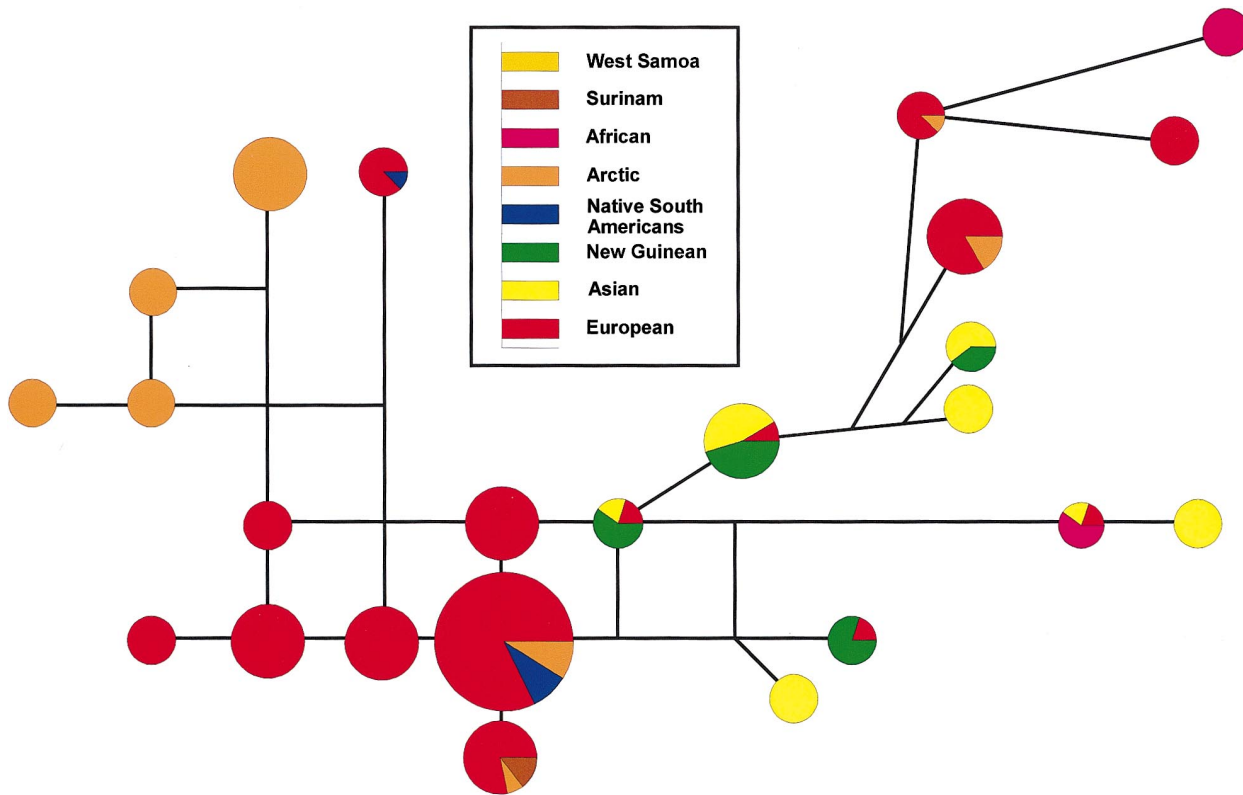


Figure 7 Modified reduced median network based only on haplotypes observed in at least five male individuals (0.5%). The diameter of each circle corresponds to a categorical absolute frequency ($n = 5-9$, $n = 10-14$, and $n > 14$). Multicolored pie charts denote haplotypes found in different population groups.

nificant Φ_{ST} values, whereas population pairs with either no or one haplotype shared—for example, the New Guinean populations—have nonsignificant Φ_{ST} values. This is explainable on the basis of the number of mutational differences between nonidentical haplotypes. Interestingly, two of the three European pairs with nonsignificant Φ_{ST} values include the SWI, suggesting a strong Y-chromosome affinity between the SWI and other Europeans (probably because of their central geographic position and their multicultural composition), and the same observation was made for mtDNA (Pult et al. 1994); on the other hand, despite 20 haplotypes shared by the GER and the DUT, their Φ_{ST} value is statistically significant, which confirms earlier findings, which were based on four-locus Y-STR haplotypes (Roewer et al. 1996). Overall, we find that $\sim 23\%$ of the total genetic variance observed among the male individuals of this study was due to interpopulation differences. This is exactly the same value as was found by Poloni et al. (1997) in 58 globally distributed populations that were analyzed by use of the Y chromosome-based p49a,f/*TaqI* polymorphism.

Phylogenetic Analyses

The unrooted NJ tree, based on linearized pairwise Φ_{ST} values, shows a topology with most populations belonging to either of two groups of related populations—Europeans and New Guineans—forming distinct tight clusters (fig. 4). The New Guineans are located close to the CHI and the TAW. The grouping of the RPN outside the New Guinea cluster may be due to its small sample size. The BAS and the BAM are distant from the European cluster but group together, which may reflect their Iberian origin. Furthermore, the clustering of the BAM on the same branch with the NSA and the INU may indicate some but little Amerindian admixture within the Argentinean sample. The clustering of the BAS next to the NSA is hard to evaluate and needs further investigation; however, it has been suggested elsewhere (Ruhlen 1994) that the BAS and the Na-Dene are remnant members of the Dene-European language group, but this view is not shared by many linguists. The grouping of the CAP together with the SUR most likely reflects the relatively high African admixture in

the Surinamese gene pool. It is also noteworthy that the first branch of the obtained trees does not separate Africans from non-Africans, as has been clearly observed in many phylogenetic trees using different genetic systems, including mtDNA (Cann et al. 1987; Vigilant et al. 1991), autosomal STRs (Bowcock et al. 1994), autosomal minisatellites (Armour et al. 1996), autosomal *Alu*-insertion polymorphisms (Stoneking et al. 1997), and, recently, Y STRs (Seielstad et al. 1999). This has often been interpreted as the indication for a recent common African origin of anatomically modern humans. The difference between the results of our study and those of all others could be the result of the our study's inclusion of only a single African population, whereas 25 pooled African populations were used by Seielstad et al. (1999). A multidimensional-scaling analysis clearly corroborates these conclusions, in that most population samples from the two groups of closely related populations are clustered (fig. 5).

Phylogenetic analysis of the 598 distinct haplotypes revealed a complex picture reflecting the high amount of diversity (fig. 6). However, we were able to identify a number of major clusters that preferentially contain certain populations. Cluster 12 contained most of the INU. Also, most of the CAP haplotypes were found in only 3 of 17 major clusters—that is, in clusters 4, 6, and 7, one of which (i.e., cluster 4) also contains the majority of SUR haplotypes. The INU and the CAP contain a large number of population-specific haplotypes, most probably because they are small and isolated populations. This is also reflected in the Y-STR haplotype tree. On the other hand, Europeans with small numbers of specific haplotypes appear in all of the 17 major clusters, which could be due to the overrepresentation (48% [470/986]) of European-origin male individuals in our study.

It could be argued that all haplotypes that are observed only once are not phylogenetically informative, since they represent recent migration or recurrent mutation events. Therefore, we performed a restricted analysis using only those haplotypes that were observed at a frequency $\geq 0.5\%$ (fig. 7). The common clustering of most of the European haplotypes, as well as their connection by single mutation steps, reflects their close relationship. A similarly tight clustering was observed for common haplotypes observed among the INU, among Asians, and among New Guineans.

The present study demonstrates that, on their own, Y

STRs are a powerful tool for the study of human evolutionary processes. A similar conclusion was reached recently by Forster et al. (2000), on the basis of a phylogenetic approach only. The use of Y STRs allows the simple construction of highly variable haplotypes. With these haplotypes, it is possible to analyze differences in population structure by a comparison of haplotype diversity and of the number of population-specific haplotypes. The use of Y STRs also allows the simultaneous analysis of closely related and distantly related populations, by haplotype-sharing analyses, AMOVA, and phylogenetic analysis. Because of their relatively high mutation rate, Y STRs are polymorphic in potentially all human populations and allow human migration processes to be traced on a historical timescale. Consequently, the number of Y-STR loci needed to obtain a reasonable amount of information can be substantially smaller than what is required when Y SNPs are used. On the other hand, because Y SNPs have an $\sim 100,000 \times$ -lower mutation rate, they are ideal for the study of human migration at an evolutionary—rather than a historical—timescale (Rosser et al. 2000; Semino et al. 2000; Underhill et al. 2000). As has been suggested elsewhere (de Knijff et al. 1997, 2000a), it will be the dual approach—that using Y STRs as well as Y SNPs—that renders the maximum amount of information. This was recently demonstrated in reports on the colonization of the Pacific (Hurles et al. 1998; Kayser et al. 2000a, 2001) and of Iceland (Helgason et al. 2000). In the end, which type of Y-chromosomal polymorphism one chooses will depend on the timescale that one wishes to cover, a luxury that we only dreamed about only 5 years ago.

Acknowledgments

We would like to thank the original donors of the DNA samples, as well as Chris Tyler-Smith, Wulf Schiefenhövel, Erika Hagelberg, Gabriele Herzog-Schröder, and the late P. Mheera Khan for providing samples. L. L. Cavalli-Sforza kindly allowed one of us (L.F.B.) to use the CAP samples. Heike Zimdahl and Arpita Pandya are acknowledged for the DNA extractions, and Peter Henneman and Rene Mieremet are acknowledged for their expert technical assistance. L.E. was supported by Swiss National Science Foundation grant 31-56755.99. Finally, we want to acknowledge the constructive remarks of two anonymous reviewers of a previous version of the manuscript.

Appendix A

Table A1

Allele Frequencies of Seven Y-STR Loci, among Male Individuals from 20 Globally Dispersed Populations

ALLELE	MEAN (SE) FREQUENCY IN				
	BAM (<i>n</i> = 100)	BAS (<i>n</i> = 30)	CHI (<i>n</i> = 36)	DUT (<i>n</i> = 88)	GER (<i>n</i> = 88)
DYS19:					
10
11
12033 (.033)
13	.220 (.041)045 (.022)	.091 (.031)
14	.560 (.050)	.900 (.055)	.250 (.072)	.705 (.049)	.386 (.052)
15	.190 (.039)278 (.075)	.193 (.042)	.239 (.045)
16	.030 (.017)389 (.081)	.034 (.019)	.227 (.045)
17067 (.046)	.056 (.038)	.023 (.016)	.057 (.025)
18028 (.027)
DYS389I:					
11028 (.027)
12	.110 (.031)	.033 (.033)	.528 (.083)	.273 (.047)	.193 (.042)
13	.610 (.049)	.433 (.090)	.139 (.058)	.602 (.052)	.602 (.052)
14	.280 (.045)	.533 (.091)	.306 (.077)	.125 (.035)	.205 (.043)
15
16
DYS389II:					
26
27033 (.033)	.083 (.046)
28	.130 (.034)306 (.077)	.307 (.049)	.159 (.039)
29	.370 (.048)	.400 (.089)	.194 (.066)	.409 (.052)	.386 (.052)
30	.320 (.047)	.400 (.089)	.194 (.066)	.250 (.046)	.295 (.049)
31	.140 (.035)	.100 (.055)	.222 (.069)	.034 (.019)	.114 (.034)
32	.040 (.020)	.067 (.046)045 (.022)
33
DYS390:					
19
20
21	.020 (.014)028 (.027)	.011 (.011)	...
22	.120 (.032)056 (.038)	.182 (.041)	.080 (.029)
23	.280 (.045)	.167 (.068)	.528 (.083)	.386 (.052)	.318 (.050)
24	.540 (.050)	.733 (.081)	.250 (.072)	.318 (.050)	.284 (.048)
25	.040 (.020)	.100 (.055)	.139 (.058)	.068 (.027)	.261 (.047)
26023 (.016)	.057 (.025)
27011 (.011)	...
DYS391:					
8
9	.050 (.022)	.033 (.033)034 (.019)	.034 (.019)
10	.550 (.050)	.333 (.086)	.778 (.069)	.580 (.053)	.568 (.053)
11	.380 (.049)	.600 (.089)	.222 (.069)	.352 (.051)	.364 (.051)
12	.020 (.014)	.033 (.033)034 (.019)	.034 (.019)
DYS392:					
7011 (.011)
10	.050 (.022)
11	.280 (.045)	.067 (.046)	.083 (.046)	.386 (.052)	.591 (.052)
12	.030 (.017)056 (.038)	.057 (.025)	.068 (.027)
13	.450 (.050)	.933 (.046)	.306 (.077)	.545 (.053)	.284 (.048)
14	.030 (.017)500 (.083)	.011 (.011)	.034 (.019)
15	.110 (.031)056 (.038)011 (.011)
16	.050 (.022)
DYS393:					
9	.020 (.014)

Table A1 (continued)

ALLELE	MEAN (SE) FREQUENCY IN				
	IND (<i>n</i> = 25)	INO (<i>n</i> = 69)	INU (<i>n</i> = 62)	ITA (<i>n</i> = 100)	MON (<i>n</i> = 40)
10
11
12	.160 (.037)	.100 (.055)	.528 (.083)	.114 (.034)	.068 (.027)
13	.660 (.047)	.833 (.068)	.306 (.077)	.750 (.046)	.773 (.045)
14	.140 (.035)	.067 (.046)	.139 (.058)	.114 (.034)	.125 (.035)
15	.020 (.014)028 (.027)	.023 (.016)	.034 (.019)
DYS19:					
10010 (.010)	...
11
12
13029 (.020)	.774 (.053)	.130 (.034)	...
14	.040 (.039)	.058 (.028)	.161 (.047)	.520 (.050)	.275 (.071)
15	.680 (.093)	.667 (.057)	.032 (.022)	.260 (.044)	.325 (.074)
16	.080 (.054)	.203 (.048)	.032 (.022)	.050 (.022)	.300 (.072)
17	.200 (.080)	.043 (.025)030 (.017)	.050 (.034)
18050 (.034)
DYS389I:					
11025 (.025)
12	.280 (.090)	.493 (.060)	.081 (.035)	.240 (.043)	.100 (.047)
13	.520 (.100)	.420 (.059)	.161 (.047)	.570 (.050)	.700 (.072)
14	.200 (.080)	.087 (.034)	.597 (.062)	.180 (.038)	.175 (.060)
15145 (.045)	.010 (.010)	...
16016 (.016)
DYS389II:					
26	.040 (.039)010 (.010)	...
27	.240 (.085)	.087 (.034)030 (.017)	.050 (.034)
28	.360 (.096)	.275 (.054)	.065 (.031)	.140 (.035)	.225 (.066)
29	.320 (.093)	.362 (.058)	.145 (.045)	.380 (.049)	.400 (.077)
30	.040 (.039)	.232 (.051)	.339 (.060)	.310 (.046)	.200 (.063)
31043 (.025)	.242 (.054)	.090 (.029)	.125 (.052)
32194 (.050)	.040 (.020)	...
33016 (.016)
DYS390:					
19
20
21	.080 (.054)	.116 (.039)010 (.010)	...
22	.080 (.054)	.058 (.028)	.048 (.027)	.150 (.036)	.075 (.042)
23	.200 (.080)	.246 (.052)	.065 (.031)	.390 (.049)	.250 (.068)
24	.120 (.065)	.319 (.056)	.855 (.045)	.390 (.049)	.375 (.077)
25	.360 (.096)	.246 (.052)	.032 (.022)	.050 (.022)	.275 (.071)
26	.160 (.073)	.014 (.014)010 (.010)	...
27025 (.025)
DYS391:					
8
9	.040 (.039)	.101 (.036)040 (.020)	.150 (.056)
10	.640 (.096)	.652 (.057)	.806 (.050)	.650 (.048)	.675 (.074)
11	.320 (.093)	.246 (.052)	.194 (.050)	.300 (.046)	.175 (.060)
12010 (.010)	...
DYS392:					
7
10	.040 (.039)	.014 (.014)	.016 (.016)
11	.240 (.085)	.145 (.042)	.113 (.040)	.580 (.049)	.725 (.071)
12072 (.031)	.016 (.016)	.090 (.029)	.025 (.025)
13	.720 (.090)	.536 (.060)	.565 (.063)	.260 (.044)	.075 (.042)
14188 (.047)	.274 (.057)	.060 (.024)	.150 (.056)
15029 (.020)	.016 (.016)
16014 (.014)010 (.010)	.025 (.025)
DYS393:					
9

Table A1 (continued)

ALLELE	MEAN (SE) FREQUENCY IN				
10
11014 (.014)030 (.017)	...
12	.280 (.090)	.217 (.050)280 (.045)	.225 (.066)
13	.080 (.054)	.391 (.059)	.484 (.063)	.570 (.050)	.625 (.077)
14	.560 (.099)	.333 (.057)	.516 (.063)	.110 (.031)	.150 (.056)
15	.080 (.054)	.043 (.025)010 (.010)	...
	NSA (n = 46)	PNG (n = 26)	CAP (n = 31)	RPN (n = 12)	SUR (n = 54)
DYS19:					
10
11
12019 (.018)
13	.500 (.074)111 (.043)
14	.261 (.065)	.192 (.077)	.065 (.044)	.083 (.080)	.222 (.057)
15	.217 (.061)	.692 (.091)	.419 (.089)	.750 (.125)	.500 (.068)
16038 (.038)	.258 (.079)	.167 (.108)	.111 (.043)
17	.022 (.022)	.038 (.038)	.258 (.079)037 (.026)
18038 (.038)
DYS389I:					
11129 (.060)
12	.022 (.022)	.231 (.083)	.323 (.084)	.083 (.080)	.185 (.053)
13	.826 (.056)	.538 (.098)	.419 (.089)	.750 (.125)	.537 (.068)
14	.152 (.053)	.231 (.083)	.065 (.044)	.167 (.108)	.259 (.060)
15065 (.044)019 (.018)
16
DYS389II:					
26019 (.018)
27032 (.032)	.083 (.080)	.019 (.018)
28	.043 (.030)	.033 (.033)	.355 (.086)056 (.031)
29	.326 (.069)	.433 (.090)	.258 (.079)	.750 (.125)	.241 (.058)
30	.391 (.072)	.300 (.084)	.161 (.066)	.167 (.108)	.389 (.066)
31	.217 (.061)	.200 (.073)	.065 (.044)204 (.055)
32	.022 (.022)	.033 (.033)	.065 (.044)074 (.036)
33065 (.044)
DYS390:					
19
20038 (.038)
21	.043 (.030)	.077 (.052)	.323 (.084)	.083 (.080)	.259 (.060)
22	.217 (.061)065 (.044)	.083 (.080)	.167 (.051)
23	.261 (.065)	.115 (.063)	.290 (.082)296 (.062)
24	.478 (.074)	.654 (.093)	.065 (.044)	.167 (.108)	.130 (.046)
25115 (.063)	.258 (.079)	.250 (.125)	.148 (.048)
26417 (.142)	...
27
DYS391:					
8032 (.032)
9	.022 (.022)	.077 (.052)	.129 (.060)	.083 (.080)	.056 (.031)
10	.587 (.073)	.692 (.091)	.452 (.089)	.167 (.108)	.685 (.063)
11	.370 (.071)	.231 (.083)	.355 (.086)	.750 (.125)	.241 (.058)
12	.022 (.022)032 (.032)019 (.018)
DYS392:					
7
10074 (.036)
11	.109 (.046)	.115 (.063)	.968 (.032)	.167 (.108)	.630 (.066)
12032 (.032)019 (.018)
13	.152 (.053)	.808 (.077)833 (.108)	.241 (.058)
14	.609 (.072)	.077 (.052)037 (.026)
15	.130 (.050)
16
DYS393:					
9

Table A1 (continued)

ALLELE	MEAN (SE) FREQUENCY IN				
10
11
12	.022 (.022)065 (.044)185 (.053)
13	.891 (.046)	.577 (.097)	.419 (.089)	.917 (.080)	.463 (.068)
14	.087 (.042)	.308 (.091)	.355 (.086)	.083 (.080)	.315 (.063)
15115 (.063)	.161 (.066)037 (.026)
	SWI (<i>n</i> = 64)	TAW (<i>n</i> = 30)	TNB (<i>n</i> = 16)	TRO (<i>n</i> = 59)	WSA (<i>n</i> = 10)
DYS19:					
10
11
12
13	.125 (.041)	.033 (.033)068 (.033)	...
14	.578 (.062)	.033 (.033)	.125 (.083)	.085 (.036)	...
15	.219 (.052)	.700 (.084)	.750 (.108)	.492 (.065)	.200 (.126)
16	.078 (.034)	.133 (.062)	.063 (.061)	.339 (.062)	.800 (.126)
17100 (.055)	.063 (.061)	.017 (.017)	...
18
DYS389I:					
11	.031 (.022)017 (.017)	...
12	.203 (.050)	.600 (.089)	.250 (.108)	.322 (.061)	.200 (.126)
13	.563 (.062)	.333 (.086)	.438 (.124)	.492 (.065)	.600 (.155)
14	.203 (.050)	.067 (.046)	.313 (.116)	.169 (.049)	.200 (.126)
15
16
DYS389II:					
26	.016 (.016)
27	.047 (.026)	.033 (.033)	.188 (.098)	.017 (.017)	...
28	.109 (.039)	.367 (.088)	.313 (.116)	.153 (.047)	.100 (.095)
29	.391 (.061)	.333 (.086)	.125 (.083)	.288 (.059)	...
30	.250 (.054)	.200 (.073)	.313 (.116)	.136 (.045)	.600 (.155)
31	.141 (.043)	.067 (.046)339 (.062)	.200 (.126)
32	.047 (.026)063 (.061)	.068 (.033)	.100 (.095)
33
DYS390:					
19063 (.061)100 (.095)
20085 (.036)	.500 (.158)
21	.031 (.022)063 (.061)	.017 (.017)	.100 (.095)
22	.141 (.043)125 (.083)	.034 (.024)	...
23	.281 (.056)	.667 (.086)	.125 (.083)	.322 (.061)	.100 (.095)
24	.391 (.061)	.300 (.084)	.500 (.125)	.339 (.062)	.200 (.126)
25	.156 (.045)	.033 (.033)	.125 (.083)	.203 (.052)	...
26
27
DYS391:					
8
9033 (.033)	.063 (.061)	.068 (.033)	.100 (.095)
10	.656 (.059)	.767 (.077)	.813 (.098)	.847 (.047)	.800 (.126)
11	.328 (.059)	.167 (.068)	.125 (.083)	.085 (.036)	.100 (.095)
12	.016 (.016)	.033 (.033)
DYS392:					
7
10
11	.453 (.062)	.033 (.033)	.188 (.098)	.085 (.036)	...
12	.063 (.030)	.033 (.033)700 (.145)
13	.391 (.061)	.267 (.081)	.438 (.124)	.390 (.063)	.100 (.095)
14	.063 (.030)	.633 (.088)	.063 (.061)	.458 (.065)	.100 (.095)
15	.031 (.022)	.033 (.033)	.063 (.061)	.068 (.033)	...
16250 (.108)100 (.095)
DYS393:					
9

Table A1 (continued)

ALLELE	MEAN (SE) FREQUENCY IN				
10
11	.016 (.016)
12	.063 (.030)	.100 (.055)	.188 (.098)	.017 (.017)	...
13	.734 (.055)	.633 (.088)	.750 (.108)	.847 (.047)	.200 (.126)
14	.172 (.047)	.233 (.077)	.063 (.061)	.136 (.045)	.800 (.126)
15	.016 (.016)	.033 (.033)

Appendix B

Table B1

Absolute Frequencies of All Seven-Locus Y-STR Haplotypes, among Male Individuals from 20 Globally Dispersed Populations

HAPLOTYPE (CODE)	NO. OF INSTANCES IN																			TOTAL	COUNT		
	BAM	BAS	CHI	DUT	GER	IND	INO	INU	ITA	MON	NSA	PNG	CAP	RPN	SUR	SWI	TAW	TNB	TRO			WSA	
1 (10-12-30-24-10-11-13)	-	1	1	1	
2 (12-13-30-24-10-11-13)	1	1	1
3 (12-13-30-24-11-13-13)	...	1	1	1
4 (13-12-28-22-10-16-11)	1	1	1
5 (13-12-29-21-09-14-13)	1	1	1
6 (13-12-29-22-10-15-11)	1	1	1
7 (13-12-29-23-10-11-13)	1	1	2	2
8 (13-12-29-24-10-11-12)	1	1	1
9 (13-12-29-24-10-15-13)	1	1	1
10 (13-12-30-23-10-13-13)	1	1	1
11 (13-12-30-24-10-11-14)	1	1	1
12 (13-12-30-24-11-11-13)	1	1	1
13 (13-13-26-24-11-13-12)	1	1	1
14 (13-13-28-23-10-13-13)	1	1	1
15 (13-13-28-23-11-14-13)	1	1	1
16 (13-13-28-23-12-13-13)	1	1	1
17 (13-13-29-22-10-15-13)	3	3	1
18 (13-13-29-23-10-13-13)	2	2	1
19 (13-13-29-23-10-15-13)	1	1	1
20 (13-13-29-24-10-11-13)	1	1	1
21 (13-13-29-24-10-13-14)	2	2	1
22 (13-13-29-24-10-15-13)	1	1	1
23 (13-13-29-24-11-13-13)	1	1	1
24 (13-13-29-27-11-13-13)	1	1	1
25 (13-13-30-22-10-14-13)	4	4	1
26 (13-13-30-22-11-13-14)	2	2	1
27 (13-13-30-23-10-11-13)	1	1	1
28 (13-13-30-23-10-12-13)	1	1	1
29 (13-13-30-23-10-14-13)	1	1	1
30 (13-13-30-23-10-14-14)	1	1	1
31 (13-13-30-24-09-12-13)	1	1	1
32 (13-13-30-24-10-11-13)	1	3	...	1	3	8	4
33 (13-13-30-24-10-11-14)	2	1	3	2
34 (13-13-30-24-10-14-13)	3	3	1
35 (13-13-30-24-11-11-13)	2	2	1
36 (13-13-30-24-11-13-12)	1	1	1
37 (13-13-30-24-11-13-13)	1	1	1
38 (13-13-30-25-09-11-14)	1	1	1
39 (13-13-30-25-10-11-13)	1	1	2	2
40 (13-13-30-25-11-11-13)	1	1	1
41 (13-13-31-22-10-14-13)	2	2	1
42 (13-13-31-23-09-14-13)	1	1	1

Table B1 (continued)

HAPLOTYPE (CODE)	NO. OF INSTANCES IN																			TOTAL	COUNT		
	BAM	BAS	CHI	DUT	GER	IND	INO	INU	ITA	MON	NSA	PNG	CAP	RPN	SUR	SWI	TAW	TNB	TRO			WSA	
43 (13-13-31-23-10-14-13)	1	1	1	
44 (13-13-31-23-10-15-13)	1	1	1
45 (13-13-31-24-11-11-13)	1	1	1
46 (13-13-31-25-10-11-13)	1	1	1
47 (13-14-28-22-10-13-13)	1	1	1
48 (13-14-29-23-10-14-14)	1	1	1
49 (13-14-29-23-12-14-13)	1	1	1
50 (13-14-29-25-11-10-14)	1	1	1
51 (13-14-30-22-10-15-13)	1	1	1
52 (13-14-30-23-10-13-13)	1	1	1
53 (13-14-30-23-10-16-14)	1	1	1
54 (13-14-30-24-09-11-13)	1	1	1
55 (13-14-30-24-10-13-14)	1	1	2	2
56 (13-14-30-24-10-14-14)	14	14	1
57 (13-14-30-24-10-15-13)	1	1	1
58 (13-14-30-24-10-15-14)	1	1	2	2
59 (13-14-30-24-11-13-12)	1	1	1
60 (13-14-30-24-11-14-14)	3	3	1
61 (13-14-30-25-09-13-14)	1	1	1
62 (13-14-30-25-10-14-13)	1	1	1
63 (13-14-30-25-10-15-13)	1	1	1
64 (13-14-31-22-10-13-13)	1	1	1
65 (13-14-31-22-10-14-13)	1	1	1
66 (13-14-31-23-10-16-14)	3	3	1
67 (13-14-31-24-10-11-13)	1	1	2	2
68 (13-14-31-24-10-13-13)	6	6	1
69 (13-14-31-24-10-13-14)	8	8	1
70 (13-14-31-24-10-14-13)	1	1	1
71 (13-14-31-24-11-14-14)	2	2	1
72 (13-14-31-25-11-11-13)	1	1	1
73 (13-14-32-24-10-11-13)	1	1	1
74 (13-14-32-24-10-13-14)	2	2	1
75 (13-14-32-25-10-11-14)	1	1	1
76 (13-15-32-24-10-13-13)	9	9	1
77 (13-16-33-24-10-13-13)	1	1	1
78 (14-11-27-22-10-11-13)	2	2	1
79 (14-12-27-23-10-14-12)	1	1	1
80 (14-12-27-24-11-13-13)	1	1	1
81 (14-12-28-21-10-11-13)	1	1	2	2
82 (14-12-28-21-11-11-12)	1	1	1
83 (14-12-28-22-10-11-09)	1	1	1
84 (14-12-28-22-10-11-12)	1	1	1
85 (14-12-28-22-10-11-13)	7	2	2	1	12	4
86 (14-12-28-22-10-11-14)	1	1	1
87 (14-12-28-22-11-11-13)	1	1	1
88 (14-12-28-23-09-11-12)	1	1	1
89 (14-12-28-23-10-11-13)	1	4	1	2	8	4
90 (14-12-28-23-10-11-14)	1	1	1
91 (14-12-28-23-10-13-12)	1	1	1
92 (14-12-28-23-10-14-12)	1	1	1
93 (14-12-28-23-11-13-13)	1	1	1
94 (14-12-28-24-10-11-12)	1	1	1
95 (14-12-28-24-10-11-13)	1	1	2	2
96 (14-12-28-24-10-13-13)	1	1	1
97 (14-12-28-24-10-14-12)	1	1	1
98 (14-12-28-24-10-15-12)	1	1	1
99 (14-12-28-24-11-13-13)	1	1	1
100 (14-12-28-24-11-15-13)	1	1	1
101 (14-12-28-25-10-11-13)	1	1	1

Table B1 (continued)

HAPLOTYPE (CODE)	NO. OF INSTANCES IN																			TOTAL	COUNT
	BAM	BAS	CHI	DUT	GER	IND	INO	INU	ITA	MON	NSA	PNG	CAP	RPN	SUR	SWI	TAW	TNB	TRO		
102 (14-12-28-25-11-11-13)	2	2	1
103 (14-12-28-25-11-14-12)	1	1	1
104 (14-12-29-22-10-11-13)	1	1	1
105 (14-12-29-23-09-11-13)	1	1	1
106 (14-12-29-23-10-11-13)	1	1	2	2
107 (14-12-29-23-10-12-12)	1	1	1
108 (14-12-29-23-11-13-13)	1	1	2	2
109 (14-12-29-24-11-13-13)	1	1	1
110 (14-12-29-25-10-11-13)	1	1	1
111 (14-12-30-23-11-11-13)	1	1	1
112 (14-12-31-24-10-13-14)	1	1	1
113 (14-13-26-24-10-13-13)	1	1	1
114 (14-13-27-23-10-14-12)	1	1	1
115 (14-13-28-22-10-13-12)	2	2	1
116 (14-13-28-22-11-13-13)	1	1	1
117 (14-13-28-23-10-10-13)	1	1	1
118 (14-13-28-23-10-10-14)	3	3	1
119 (14-13-28-23-10-11-13)	1	1	1
120 (14-13-28-23-10-13-13)	1	1	2	2
121 (14-13-28-23-10-14-13)	1	1	1
122 (14-13-28-23-11-13-13)	1	1	1
123 (14-13-28-23-12-13-13)	1	1	1
124 (14-13-28-24-09-14-14)	1	1	1
125 (14-13-28-24-10-10-13)	1	1	1
126 (14-13-28-24-10-13-13)	1	2	3	2
127 (14-13-28-24-10-15-12)	1	1	1
128 (14-13-28-24-11-14-12)	1	1	2	2
129 (14-13-28-24-11-15-13)	2	1	3	2
130 (14-13-29-21-10-11-14)	1	1	1
131 (14-13-29-22-11-13-13)	1	1	1
132 (14-13-29-23-09-13-13)	1	1	1
133 (14-13-29-23-10-07-13)	1	1	1
134 (14-13-29-23-10-10-14)	1	1	1
135 (14-13-29-23-10-11-12)	1	5	6	2
136 (14-13-29-23-10-11-13)	2	2	1
137 (14-13-29-23-10-12-13)	1	1	1
138 (14-13-29-23-10-13-13)	1	1	1	1	4	4
139 (14-13-29-23-10-13-14)	1	1	1
140 (14-13-29-23-10-14-12)	1	1	1
141 (14-13-29-23-10-15-13)	1	...	1	1
142 (14-13-29-23-11-11-12)	1	1	2	2
143 (14-13-29-23-11-13-12)	1	1	1
144 (14-13-29-23-11-13-13)	7	2	1	1	2	1	14	6
145 (14-13-29-23-11-13-15)	1	1	1
146 (14-13-29-23-11-14-13)	1	1	2	2
147 (14-13-29-23-12-13-13)	1	1	1	3	3
148 (14-13-29-24-09-11-13)	1	1	1
149 (14-13-29-24-10-11-09)	1	1	1
150 (14-13-29-24-10-11-13)	1	1	1
151 (14-13-29-24-10-13-12)	1	1	2	2
152 (14-13-29-24-10-13-13)	2	2	...	1	1	3	3	12	6
153 (14-13-29-24-10-13-14)	1	1	2	2
154 (14-13-29-24-10-14-13)	1	2	3	2
155 (14-13-29-24-10-15-13)	1	1	1
156 (14-13-29-24-10-16-14)	1	1	1
157 (14-13-29-24-11-11-14)	1	1	1
158 (14-13-29-24-11-12-13)	1	1	1
159 (14-13-29-24-11-13-13)	9	4	...	6	2	3	4	...	3	3	34	8
160 (14-13-29-24-11-13-14)	...	1	...	1	2	2

Table B1 (continued)

HAPLOTYPE (CODE)	NO. OF INSTANCES IN																			TOTAL	COUNT	
	BAM	BAS	CHI	DUT	GER	IND	INO	INU	ITA	MON	NSA	PNG	CAP	RPN	SUR	SWI	TAW	TNB	TRO			WSA
161 (14-13-29-24-11-14-13)	1	1	2	2
162 (14-13-29-24-11-15-12)	1	1	1
163 (14-13-29-24-12-13-12)	1	1	1
164 (14-13-29-24-12-13-13)	1	1	2	2
165 (14-13-29-25-10-13-12)	...	3	...	1	4	2
166 (14-13-29-25-10-13-13)	1	1	2	2
167 (14-13-29-25-11-13-12)	1	1	2	2
168 (14-13-29-25-11-13-13)	1	2	3	2
169 (14-13-29-25-11-15-13)	1	...	1	1
170 (14-13-30-22-10-11-12)	2	2	1
171 (14-13-30-22-10-11-14)	1	1	1
172 (14-13-30-22-10-13-13)	1	1	1
173 (14-13-30-23-09-11-13)	1	1	1
174 (14-13-30-23-10-11-12)	1	1	2	2
175 (14-13-30-23-10-11-13)	1	1	1
176 (14-13-30-23-10-12-14)	2	2	1
177 (14-13-30-23-10-13-13)	1	1	1
178 (14-13-30-23-10-14-13)	1	1	1
179 (14-13-30-23-10-15-13)	1	...	1	1
180 (14-13-30-23-11-11-12)	2	2	1
181 (14-13-30-23-11-11-13)	1	1	1
182 (14-13-30-23-11-13-12)	1	1	1
183 (14-13-30-23-11-13-13)	1	1	1
184 (14-13-30-23-11-14-14)	1	1	1
185 (14-13-30-23-11-16-13)	1	1	1
186 (14-13-30-24-09-14-13)	1	1	1
187 (14-13-30-24-10-11-11)	1	1	1
188 (14-13-30-24-10-11-13)	1	1	1
189 (14-13-30-24-10-12-13)	1	1	1
190 (14-13-30-24-10-13-13)	2	1	3	2
191 (14-13-30-24-10-14-13)	1	1	1
192 (14-13-30-24-11-13-13)	5	1	...	2	2	10	4
193 (14-13-30-24-11-13-14)	1	...	1	2	2
194 (14-13-30-25-10-13-12)	1	1	1
195 (14-13-30-25-11-11-13)	1	1	1
196 (14-13-30-25-11-13-13)	1	1	1
197 (14-13-30-25-11-14-13)	1	1	1
198 (14-13-30-26-10-13-13)	1	1	1
199 (14-13-30-27-11-11-12)	1	1	1
200 (14-13-31-22-10-11-14)	2	2	1
201 (14-13-31-22-10-14-13)	3	3	1
202 (14-13-31-23-10-11-12)	1	1	2	2
203 (14-13-31-24-10-11-13)	1	1	1
204 (14-13-31-24-11-11-13)	1	1	1
205 (14-13-31-25-11-13-12)	1	1	1
206 (14-13-32-24-10-13-12)	1	1	1
207 (14-13-32-24-10-13-13)	1	1	1
208 (14-14-29-23-10-10-14)	1	1	1
209 (14-14-29-23-10-11-12)	1	1	1
210 (14-14-29-23-10-13-13)	1	1	1
211 (14-14-29-23-10-14-13)	1	1	1
212 (14-14-29-23-11-12-14)	1	1	1
213 (14-14-29-23-11-13-13)	1	1	1
214 (14-14-29-24-11-13-13)	1	1	2	2
215 (14-14-30-21-10-13-13)	1	1	1
216 (14-14-30-23-10-10-14)	1	1	1
217 (14-14-30-23-10-11-12)	1	1	1
218 (14-14-30-23-10-11-13)	1	1	1
219 (14-14-30-23-10-13-13)	1	1	1	1	4	4

Table B1 (continued)

HAPLOTYPE (CODE)	NO. OF INSTANCES IN																			TOTAL	COUNT	
	BAM	BAS	CHI	DUT	GER	IND	INO	INU	ITA	MON	NSA	PNG	CAP	RPN	SUR	SWI	TAW	TNB	TRO			WSA
220 (14-14-30-23-10-14-14)	1	1	1
221 (14-14-30-23-10-16-12)	1	1	1
222 (14-14-30-23-11-13-13)	...	1	...	3	2	1	7	4
223 (14-14-30-23-11-13-14)	...	1	1	1
224 (14-14-30-23-11-14-14)	1	1	1
225 (14-14-30-24-09-11-13)	1	1	1
226 (14-14-30-24-09-13-13)	...	1	1	1
227 (14-14-30-24-10-11-12)	1	1	1
228 (14-14-30-24-10-13-13)	2	1	...	1	1	2	7	5
229 (14-14-30-24-10-13-14)	1	1	1
230 (14-14-30-24-11-11-12)	1	1	1
231 (14-14-30-24-11-13-12)	1	...	1	2	2
232 (14-14-30-24-11-13-13)	...	6	...	2	1	2	11	4
233 (14-14-30-25-10-13-13)	1	1	1
234 (14-14-30-25-11-13-12)	1	1	1
235 (14-14-31-20-10-11-13)	2	...	2	1
236 (14-14-31-23-10-11-12)	1	1	1
237 (14-14-31-23-10-13-13)	1	1	1
238 (14-14-31-23-10-14-13)	1	1	1
239 (14-14-31-23-10-15-13)	1	1	1
240 (14-14-31-23-11-13-13)	...	1	1	1
241 (14-14-31-23-11-14-14)	1	1	1
242 (14-14-31-24-10-11-12)	1	1	1
243 (14-14-31-24-10-11-13)	1	1	1	3	3
244 (14-14-31-24-10-13-13)	1	1	1
245 (14-14-31-24-11-13-13)	1	2	3	2
246 (14-14-31-25-11-11-13)	1	1	1
247 (14-14-32-22-10-11-14)	1	1	1
248 (14-14-32-24-10-11-12)	1	1	1
249 (14-14-32-24-10-11-13)	1	1	1
250 (14-14-32-24-10-13-13)	...	2	2	1
251 (15-11-27-24-10-13-13)	1	...	1	1
252 (15-11-28-22-10-11-14)	1	1	1
253 (15-11-28-22-10-14-12)	1	1	1
254 (15-11-28-25-11-11-14)	1	1	1
255 (15-11-28-25-12-11-14)	1	1	1
256 (15-12-26-24-10-13-13)	1	1	1
257 (15-12-26-25-10-13-14)	1	1	1
258 (15-12-27-22-10-11-12)	1	1	1
259 (15-12-27-23-10-11-13)	1	1	1
260 (15-12-27-23-10-14-13)	5	5	1
261 (15-12-27-23-10-14-14)	1	1	1
262 (15-12-27-24-09-11-12)	1	1	1
263 (15-12-27-24-10-11-12)	1	1	2	2
264 (15-12-27-24-10-11-14)	1	1	1
265 (15-12-27-24-10-13-13)	1	1	1
266 (15-12-27-25-10-13-13)	1	1	1
267 (15-12-27-25-11-13-13)	1	1	1
268 (15-12-28-21-11-11-12)	1	1	1
269 (15-12-28-22-10-11-12)	1	1	1
270 (15-12-28-22-10-11-14)	3	1	4	2
271 (15-12-28-22-10-13-14)	2	...	2	1
272 (15-12-28-22-10-14-13)	1	1	1
273 (15-12-28-23-09-11-12)	1	1	1
274 (15-12-28-23-09-14-13)	1	1	1
275 (15-12-28-23-10-11-11)	1	1	1
276 (15-12-28-23-10-11-13)	1	2	3	2
277 (15-12-28-23-10-11-15)	1	1	1
278 (15-12-28-23-10-12-14)	1	1	1

Table B1 (continued)

HAPLOTYPE (CODE)	NO. OF INSTANCES IN																			TOTAL	COUNT	
	BAM	BAS	CHI	DUT	GER	IND	INO	INU	ITA	MON	NSA	PNG	CAP	RPN	SUR	SWI	TAW	TNB	TRO			WSA
279 (15-12-28-23-10-13-13)	1	1	1
280 (15-12-28-23-10-14-13)	1	1	1	3	3
281 (15-12-28-23-11-14-13)	1	1	2	2
282 (15-12-28-24-09-11-13)	1	1	1
283 (15-12-28-24-10-11-12)	1	1	2	2
284 (15-12-28-24-10-13-12)	1	1	1
285 (15-12-28-24-10-13-13)	1	5	1	4	...	11	4	
286 (15-12-28-24-10-13-14)	1	1	1
287 (15-12-28-24-10-14-13)	1	1	1
288 (15-12-28-24-11-13-12)	1	1	1
289 (15-12-28-24-11-13-14)	1	1	1
290 (15-12-29-21-09-12-13)	1	1	1
291 (15-12-29-21-10-11-13)	1	1	1
292 (15-12-29-21-10-11-15)	1	1	1
293 (15-12-29-21-11-11-12)	1	1	1
294 (15-12-29-21-11-11-13)	1	1	1
295 (15-12-29-21-11-11-15)	1	1	1
296 (15-12-29-22-10-11-13)	1	2	1	4	3
297 (15-12-29-22-10-11-14)	1	1	2	2
298 (15-12-29-22-10-14-12)	1	1	1
299 (15-12-29-22-11-11-14)	1	1	1
300 (15-12-29-23-10-12-12)	1	1	1
301 (15-12-29-23-10-13-14)	1	1	1
302 (15-12-29-23-10-14-13)	1	3	...	1	5	3
303 (15-12-29-23-11-13-13)	1	1	2	2
304 (15-12-29-23-11-14-13)	1	1	1
305 (15-12-29-24-10-11-12)	1	1	1
306 (15-12-29-24-10-11-13)	1	1	1
307 (15-12-29-24-10-13-13)	1	2	3	2
308 (15-12-29-24-10-13-14)	1	1	1
309 (15-12-29-24-11-11-14)	1	1	1
310 (15-12-29-24-11-13-13)	1	1	1
311 (15-12-29-24-12-13-12)	1	1	1
312 (15-12-29-25-11-11-13)	1	1	1
313 (15-12-29-25-11-11-14)	1	1	1
314 (15-12-30-21-11-11-13)	1	1	1
315 (15-12-30-22-10-11-13)	1	1	1
316 (15-12-30-23-10-12-12)	1	1	1
317 (15-12-30-23-10-12-13)	1	1	1
318 (15-12-30-23-10-12-14)	1	1	1
319 (15-12-30-23-10-13-13)	1	1	...	2	2
320 (15-12-30-23-10-14-13)	1	1	1
321 (15-12-30-23-11-11-13)	1	1	1
322 (15-12-30-23-11-11-14)	1	1	1
323 (15-12-30-24-10-13-13)	1	1	1
324 (15-12-30-25-10-12-14)	1	1	1
325 (15-12-31-23-10-14-12)	1	1	1
326 (15-12-31-23-10-14-13)	1	1	1
327 (15-12-31-24-10-13-13)	1	...	1	1
328 (15-13-28-22-10-11-13)	1	1	1
329 (15-13-28-23-10-11-13)	1	1	2	2
330 (15-13-28-23-10-11-14)	2	2	1
331 (15-13-28-23-10-14-13)	1	1	1
332 (15-13-28-24-10-11-13)	1	1	1
333 (15-13-28-24-10-13-13)	1	1	2	2
334 (15-13-28-24-10-13-14)	1	...	1	2	2
335 (15-13-28-24-11-11-13)	1	1	1
336 (15-13-28-24-11-13-14)	1	1	1
337 (15-13-28-24-11-14-13)	2	2	1

Table B1 (continued)

HAPLOTYPE (CODE)	NO. OF INSTANCES IN																			TOTAL	COUNT	
	BAM	BAS	CHI	DUT	GER	IND	INO	INU	ITA	MON	NSA	PNG	CAP	RPN	SUR	SWI	TAW	TNB	TRO			WSA
338 (15-13-28-25-10-13-14)	2	1	3	2
339 (15-13-28-25-11-13-14)	4	4	1
340 (15-13-29-21-10-11-13)	1	1	1
341 (15-13-29-21-11-11-14)	1	1	1
342 (15-13-29-22-10-11-12)	1	1	1
343 (15-13-29-22-10-11-13)	1	1	1
344 (15-13-29-22-10-13-12)	2	2	1
345 (15-13-29-22-11-11-13)	1	1	1
346 (15-13-29-22-11-11-14)	1	1	1
347 (15-13-29-23-09-11-12)	1	1	2	2
348 (15-13-29-23-09-13-13)	1	1	1
349 (15-13-29-23-10-11-12)	1	1	1
350 (15-13-29-23-10-11-13)	1	1	1	3	3
351 (15-13-29-23-10-12-15)	1	1	1
352 (15-13-29-23-10-14-13)	1	1	...	1	3	3
353 (15-13-29-23-11-11-12)	1	1	1
354 (15-13-29-23-11-11-15)	1	1	1
355 (15-13-29-23-11-13-13)	1	1	1	3	3
356 (15-13-29-24-09-11-13)	1	1	1
357 (15-13-29-24-09-13-12)	1	1	1
358 (15-13-29-24-10-11-12)	1	1	1
359 (15-13-29-24-10-11-13)	3	3	1
360 (15-13-29-24-10-12-15)	1	1	1
361 (15-13-29-24-10-13-13)	1	1	3	5	3
362 (15-13-29-24-10-13-14)	2	1	3	2
363 (15-13-29-24-10-15-14)	1	1	1
364 (15-13-29-24-10-16-13)	1	1	1
365 (15-13-29-24-11-13-12)	1	1	1
366 (15-13-29-24-11-13-13)	1	1	1	3	3
367 (15-13-29-24-11-13-14)	1	1	1
368 (15-13-29-24-11-14-13)	1	1	1
369 (15-13-29-25-09-13-14)	1	1	1
370 (15-13-29-25-10-11-13)	1	1	3	5	3
371 (15-13-29-25-10-13-13)	1	...	1	1
372 (15-13-29-25-10-13-14)	1	1	1
373 (15-13-29-25-10-13-15)	1	1	1
374 (15-13-29-25-11-13-13)	1	3	4	2
375 (15-13-29-25-11-13-14)	1	4	5	2
376 (15-13-29-25-11-13-15)	1	1	1
377 (15-13-29-26-10-11-13)	1	1	1
378 (15-13-29-26-11-11-13)	1	1	1
379 (15-13-29-26-11-13-13)	1	4	5	2
380 (15-13-29-26-11-13-14)	1	1	1
381 (15-13-30-20-09-11-14)	3	...	3	1
382 (15-13-30-21-10-11-13)	1	...	1	2	2
383 (15-13-30-21-10-11-15)	1	1	1
384 (15-13-30-21-11-11-12)	1	1	1
385 (15-13-30-22-10-11-12)	1	1	1
386 (15-13-30-22-10-11-13)	1	1	1
387 (15-13-30-22-10-11-14)	1	1	1
388 (15-13-30-23-09-13-12)	1	1	1
389 (15-13-30-23-10-11-12)	1	1	1
390 (15-13-30-23-10-12-14)	1	1	1
391 (15-13-30-23-10-12-15)	1	1	1
392 (15-13-30-23-10-13-13)	1	1	1
393 (15-13-30-23-10-14-13)	2	2	1
394 (15-13-30-23-11-13-13)	1	1	1
395 (15-13-30-23-11-14-13)	1	1	1
396 (15-13-30-23-11-14-14)	1	1	1

Table B1 (continued)

HAPLOTYPE (CODE)	NO. OF INSTANCES IN																			TOTAL	COUNT	
	BAM	BAS	CHI	DUT	GER	IND	INO	INU	ITA	MON	NSA	PNG	CAP	RPN	SUR	SWI	TAW	TNB	TRO			WSA
397 (15-13-30-24-09-16-13)	1	1	1
398 (15-13-30-24-10-11-13)	1	1	1
399 (15-13-30-24-10-13-13)	1	2	3	2
400 (15-13-30-24-10-13-14)	1	1	1
401 (15-13-30-24-10-13-15)	1	1	1
402 (15-13-30-24-10-14-13)	1	1	1
403 (15-13-30-24-10-15-13)	1	1	1
404 (15-13-30-24-11-11-13)	1	1	1
405 (15-13-30-24-11-13-13)	1	1	1
406 (15-13-30-24-11-13-14)	2	2	1
407 (15-13-30-24-11-14-13)	1	1	1
408 (15-13-30-24-11-15-13)	1	1	1
409 (15-13-30-24-12-11-13)	1	1	1
410 (15-13-30-25-11-11-13)	1	1	2	2
411 (15-13-30-25-11-13-13)	1	1	1	1
412 (15-13-30-25-11-13-14)	2	2	1
413 (15-13-30-26-11-11-13)	1	1	1
414 (15-13-30-26-11-13-13)	1	1	1	1
415 (15-13-31-21-10-11-13)	2	2	1
416 (15-13-31-21-11-11-13)	2	2	1
417 (15-13-31-23-10-11-13)	1	1	1
418 (15-13-31-23-10-13-14)	1	1	1
419 (15-13-31-23-10-14-13)	3	3	1
420 (15-13-31-23-11-13-13)	1	1	1
421 (15-13-31-24-10-14-13)	1	...	1	1
422 (15-13-31-24-11-11-13)	1	1	1
423 (15-13-31-25-10-11-13)	1	1	1
424 (15-13-31-25-10-14-13)	2	2	1
425 (15-13-31-25-11-11-13)	1	1	1
426 (15-13-32-22-10-11-13)	1	1	1
427 (15-13-32-24-12-13-13)	1	1	1
428 (15-13-32-25-10-11-14)	1	1	1
429 (15-13-32-25-10-14-13)	1	1	1
430 (15-14-29-22-08-11-13)	1	1	1
431 (15-14-29-23-10-11-14)	1	1	1
432 (15-14-29-24-10-11-14)	1	1	1
433 (15-14-29-24-10-13-14)	1	1	1
434 (15-14-29-24-11-14-13)	1	1	1
435 (15-14-29-26-10-13-14)	2	2	1
436 (15-14-30-19-10-11-13)	1	1	1
437 (15-14-30-20-10-11-14)	1	1	1
438 (15-14-30-21-10-11-12)	1	1	1
439 (15-14-30-22-10-10-15)	1	1	1
440 (15-14-30-22-10-11-12)	1	1	1
441 (15-14-30-22-10-11-14)	1	1	1
442 (15-14-30-23-09-11-12)	1	1	1
443 (15-14-30-23-10-12-14)	1	1	1
444 (15-14-30-24-10-13-13)	1	1	1
445 (15-14-30-24-10-13-14)	1	1	1
446 (15-14-30-24-10-16-13)	2	2	1
447 (15-14-30-24-11-14-13)	2	2	1
448 (15-14-30-25-10-11-13)	1	1	1
449 (15-14-30-25-10-13-14)	1	1	1
450 (15-14-30-25-11-11-13)	1	1	1
451 (15-14-30-25-11-13-12)	1	1	1
452 (15-14-31-21-10-11-14)	1	1	1
453 (15-14-31-22-10-11-12)	1	1	1
454 (15-14-31-22-10-11-13)	1	1	1
455 (15-14-31-23-10-10-14)	1	1	1

Table B1 (continued)

HAPLOTYPE (CODE)	NO. OF INSTANCES IN																			TOTAL	COUNT	
	BAM	BAS	CHI	DUT	GER	IND	INO	INU	ITA	MON	NSA	PNG	CAP	RPN	SUR	SWI	TAW	TNB	TRO			WSA
456 (15-14-31-23-10-12-14)	1	1	2	2
457 (15-14-31-23-11-14-14)	1	1	1
458 (15-14-31-24-10-13-13)	1	1	1
459 (15-14-31-24-10-14-13)	1	1	1
460 (15-14-31-25-10-11-13)	2	2	1
461 (15-14-31-25-11-11-13)	1	1	1
462 (15-14-31-25-11-13-13)	1	1	1
463 (15-14-32-21-10-11-13)	1	1	1
464 (15-14-32-21-10-11-14)	1	1	1
465 (15-14-32-23-10-11-14)	1	1	1
466 (15-14-32-23-10-12-13)	1	1	1
467 (15-14-32-23-10-12-14)	1	1	2	2
468 (15-14-32-23-10-12-15)	1	1	2	2
469 (15-14-32-23-10-14-13)	1	1	1
470 (15-14-32-25-10-14-13)	1	2	...	2	1
471 (15-15-31-23-09-11-12)	1	1	1
472 (15-15-32-23-10-13-13)	1	1	1
473 (15-15-33-23-09-11-13)	2	2	1
474 (16-11-27-24-09-11-12)	1	1	1
475 (16-11-28-24-09-11-12)	1	1	1
476 (16-12-27-23-10-14-13)	1	1	2	2
477 (16-12-27-24-09-13-13)	1	1	1
478 (16-12-27-25-10-15-12)	1	1	1
479 (16-12-28-21-10-11-12)	1	1	1
480 (16-12-28-22-10-14-13)	1	1	1
481 (16-12-28-23-10-11-12)	1	1	1
482 (16-12-28-23-10-11-13)	1	1	1
483 (16-12-28-23-10-13-12)	1	1	1
484 (16-12-28-23-10-14-12)	1	1	1
485 (16-12-28-23-10-14-13)	1	1	1
486 (16-12-28-23-10-14-14)	1	1	1
487 (16-12-28-24-10-13-12)	1	1	1
488 (16-12-28-24-10-13-13)	1	1	1
489 (16-12-28-24-10-14-13)	1	...	1	1
490 (16-12-28-25-09-13-12)	1	1	1
491 (16-12-28-26-10-11-13)	1	1	1
492 (16-12-29-21-10-11-15)	1	1	1
493 (16-12-29-23-10-13-13)	2	...	2	1
494 (16-12-29-23-10-14-12)	1	1	1
495 (16-12-29-23-10-14-13)	1	...	1	1
496 (16-12-29-23-11-13-13)	1	1	1
497 (16-12-29-23-11-14-13)	2	2	1
498 (16-12-29-24-10-11-12)	1	1	1
499 (16-12-29-24-10-14-13)	1	1	1
500 (16-12-29-25-10-12-12)	1	1	1
501 (16-12-29-25-10-14-13)	1	1	1
502 (16-12-30-21-09-12-13)	1	1	1
503 (16-12-30-22-09-12-13)	1	1	1
504 (16-12-30-23-10-12-14)	1	1	1
505 (16-12-30-23-10-14-13)	1	1	1
506 (16-12-31-25-11-12-14)	1	1	1
507 (16-13-28-21-10-14-12)	1	1	1
508 (16-13-28-24-09-11-12)	1	1	1
509 (16-13-28-24-10-13-12)	1	1	1
510 (16-13-29-21-10-11-15)	3	3	1
511 (16-13-29-22-10-11-13)	1	1	1
512 (16-13-29-22-11-11-13)	1	1	1
513 (16-13-29-23-10-12-14)	1	1	1
514 (16-13-29-23-10-13-14)	1	1	1

Table B1 (continued)

HAPLOTYPE (CODE)	NO. OF INSTANCES IN																			TOTAL	COUNT	
	BAM	BAS	CHI	DUT	GER	IND	INO	INU	ITA	MON	NSA	PNG	CAP	RPN	SUR	SWI	TAW	TNB	TRO			WSA
515 (16-13-29-23-11-13-13)	1	1	1
516 (16-13-29-24-09-11-13)	3	3	1
517 (16-13-29-24-10-11-13)	1	1	1
518 (16-13-29-24-10-11-15)	1	1	1
519 (16-13-29-24-10-14-12)	1	1	1
520 (16-13-29-24-11-13-13)	1	1	2	2
521 (16-13-29-25-10-11-13)	6	6	1
522 (16-13-29-25-10-13-13)	4	4	1
523 (16-13-29-25-10-13-14)	2	2	1
524 (16-13-29-25-11-10-13)	1	1	1
525 (16-13-29-25-11-11-13)	1	1	2	2
526 (16-13-29-25-11-13-14)	1	1	1
527 (16-13-29-26-10-11-13)	1	1	1
528 (16-13-30-19-10-12-14)	1	1	1
529 (16-13-30-20-10-12-14)	3	3	1
530 (16-13-30-21-10-11-14)	2	2	1
531 (16-13-30-21-10-11-15)	1	1	2	2
532 (16-13-30-22-09-11-14)	1	1	1
533 (16-13-30-22-10-12-15)	1	1	1
534 (16-13-30-23-10-11-13)	1	1	1
535 (16-13-30-24-09-16-14)	1	1	1
536 (16-13-30-24-10-14-13)	1	1	1
537 (16-13-30-25-10-11-13)	1	1	2	2
538 (16-13-30-25-11-11-13)	1	1	2	2
539 (16-13-30-26-10-12-11)	1	1	1
540 (16-13-31-23-10-13-13)	1	...	1	1
541 (16-13-31-23-10-14-13)	1	...	1	1
542 (16-13-31-24-10-14-13)	4	4	1
543 (16-13-31-24-11-11-13)	1	1	2	2
544 (16-13-31-25-10-11-13)	1	1	1
545 (16-13-31-25-10-13-12)	1	1	1
546 (16-13-31-25-10-13-13)	1	1	1
547 (16-13-31-25-11-11-13)	2	2	1
548 (16-13-31-25-11-11-14)	1	1	1
549 (16-13-32-20-10-12-14)	1	...	1	1
550 (16-14-29-21-10-11-15)	2	2	1
551 (16-14-29-23-10-13-13)	1	1	1
552 (16-14-30-20-10-12-14)	1	...	1	1
553 (16-14-30-22-10-11-12)	1	1	1
554 (16-14-30-24-12-13-13)	1	1	1
555 (16-14-30-25-10-11-13)	1	1	1
556 (16-14-31-21-10-11-15)	1	1	1
557 (16-14-31-21-10-12-14)	1	...	1	1
558 (16-14-31-22-10-13-12)	1	1	1
559 (16-14-31-22-11-11-12)	1	1	1
560 (16-14-31-23-10-13-12)	1	1	1
561 (16-14-31-23-10-13-14)	1	1	1
562 (16-14-31-24-09-11-13)	1	1	1
563 (16-14-31-24-10-13-12)	2	1	3	2
564 (16-14-31-24-10-13-13)	1	1	1
565 (16-14-31-24-12-11-13)	1	1	1
566 (16-14-31-25-10-11-13)	1	1	2	2
567 (16-14-32-21-10-14-12)	1	1	1
568 (16-14-32-23-10-11-13)	1	1	1
569 (16-14-32-24-10-12-13)	1	1	1
570 (16-14-32-25-10-11-13)	1	1	1
571 (17-12-27-22-10-12-12)	1	1	1
572 (17-12-27-23-10-11-13)	...	1	1	1
573 (17-12-27-23-10-13-12)	4	4	1

Table B1 (continued)

HAPLOTYPE (CODE)	NO. OF INSTANCES IN																			TOTAL	COUNT	
	BAM	BAS	CHI	DUT	GER	IND	INO	INU	ITA	MON	NSA	PNG	CAP	RPN	SUR	SWI	TAW	TNB	TRO			WSA
574 (17-12-28-23-10-14-13)	1	3	4	2
575 (17-12-28-23-11-11-14)	6	6	1
576 (17-12-28-24-10-13-14)	1	1	1
577 (17-12-29-22-10-11-13)	1	1	1
578 (17-13-28-24-10-13-14)	1	1	1
579 (17-13-29-23-10-14-12)	1	1	1
580 (17-13-29-25-10-11-13)	1	1	1
581 (17-13-29-26-11-11-13)	1	1	1
582 (17-13-30-21-10-11-14)	2	2	1
583 (17-13-30-22-11-15-13)	1	1	1
584 (17-13-30-23-10-15-13)	1	1	1
585 (17-13-30-24-10-11-13)	2	2	1
586 (17-13-30-25-10-11-13)	2	2	1
587 (17-13-30-26-10-11-13)	1	1	1
588 (17-13-31-21-10-11-14)	2	2	1
589 (17-13-31-24-10-11-12)	1	1	1
590 (17-13-32-24-11-11-14)	1	1	1
591 (17-14-29-23-10-11-13)	...	1	1	2	2
592 (17-14-29-23-10-12-14)	1	1	1
593 (17-14-29-24-10-11-13)	1	1	1
594 (17-14-31-23-10-14-13)	1	...	1	1
595 (17-14-32-21-10-11-14)	1	1	1
596 (18-12-30-24-11-13-13)	1	1	1
597 (18-13-31-25-11-11-13)	2	2	1
598 (18-14-30-25-10-13-12)	1	1	1
Totals:																						
Haplotypes	76	17	34	65	77	16	53	22	82	29	34	22	18	7	47	51	25	14	40	8	598	
Individuals	100	30	36	88	88	25	69	62	100	40	46	26	31	12	54	64	30	16	59	10	986	

Electronic-Database Information

The URLs for data in this article are as follows:

ARLEQUIN: A Software For Population Genetic Data Analysis, <http://anthropologie.unige.ch/arlequin>

Rod Page's Home Page, <http://taxonomy.zoology.gla.ac.uk/rod/rod.html> (for TREEVIEW 1.6.1)

Forensic Laboratory for DNA Research, <http://www.medfca.leidenuniv.nl/fldo> (for marker information)

PHYLIP, <http://evolution.genetics.washington.edu/phylip.html>
Y-STR Haplotype Reference Database, http://ystr.charite.de/index_gr.html (for marker information)

References

- Armour JA, Anttinen T, May CA, Vega EE, Sajantila A, Kidd JR, Kidd KK, Bertranpetit J, Pääbo S, Jeffreys AJ (1996) Minisatellite diversity supports a recent African origin for modern humans. *Nat Genet* 13:154–160
- Ayub Q, Mohyuddin A, Qamar R, Mazhar K, Zerjal T, Mehdi SQ, Tyler-Smith C (2000) Identification and characterisation of novel human Y-chromosomal microsatellites from sequence database information. *Nucleic Acids Res* 28:e8
- Bandelt H-J, Forster P, Sykes BC, Richards MB (1995) Mi-

tochondrial portraits of human populations using median networks. *Genetics* 141:743–753

Bianchi NO, Catanesi CI, Bailliet G, Martinez-Marignac VL, Bravi CM, Vidal-Rioja LB, Herrera RJ, López-Camelo JS (1998) Characterization of ancestral and derived Y-chromosome haplotypes of New World native populations. *Am J Hum Genet* 63:1862–1871

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

Bruford MW, Wayne RK (1993) Microsatellites and their application to population genetic studies. *Curr Opin Genet Dev* 3:939–943

Cann RL, Stoneking M, Wilson AC (1987) Mitochondrial DNA and human evolution. *Nature* 325:31–36

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

Chen H, Lowther W, Avramopoulos D, Antonarakis SE (1994) Homologous loci DXYS156 X and DXYS156Y contain a polymorphic pentanucleotide repeat (TAAAA)n and map to human X and Y chromosomes. *Hum Mutat* 4:208–211

Deka R, Jin L, Shriver MD, Yu LM, Saha N, Barrantes R, Chakraborty R, Ferrell RE (1996) Dispersion of human Y

- chromosome haplotypes based on five microsatellites in global populations. *Genome Res* 6:1177–1184
- de Knijff P (2000a) Messages through bottlenecks: on the combined use of slow and fast evolving polymorphic markers on the human Y chromosome. *Am J Hum Genet* 67: 1055–1061
- (2000b) Y chromosomes shared by descent or by state. In: Renfrew C, Boyle K (eds) *Archaeogenetics: DNA and the population prehistory of Europe*. The McDonald Institute, Cambridge, pp 301–304
- de Knijff P, Kayser M, Caglia A, Corach D, Fretwell N, Gehrig C, Graziosi G, et al (1997) Chromosome Y microsatellites: population genetic and evolutionary aspects. *Int J Legal Med* 110:134–140
- DNA Commission of the ISFH (1995) DNA recommendations—1994 report concerning further recommendations of the DNA Commission of the ISFH regarding PCR-based polymorphisms in STR (short tandem repeats) systems. *Vox Sang* 69:70–71
- Excoffier L, Smouse PE (1994) Using allele frequencies and geographic subdivision to reconstruct gene trees within a species: molecular variance parsimony. *Genetics* 136:343–359
- Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among mtDNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics* 131:479–491
- Flint J, Bond J, Rees DC, Boyce AJ, Roberts-Thomson JM, Excoffier L, Clegg JB, Beaumont MA, Nichols RA, Harding RM (1999) Minisatellite mutational processes reduce *F_{st}* estimates. *Hum Genet* 105:567–576
- Forster P, Röhl A, Lünemann P, Brinkmann C, Zerjal T, Tyler-Smith C, Brinkmann B (2000) A short tandem repeat-based phylogeny for the human Y chromosome. *Am J Hum Genet* 67:182–196
- Goldstein DB, Ruiz-Linares AR, Cavalli-Sforza LL and Feldman MW (1995a) An evaluation of genetic distances for use with microsatellite loci. *Genetics* 139:463–471
- (1995b) Genetic absolute dating based on microsatellites and the origin of modern humans. *Proc Natl Acad Sci USA* 92:6723–6727
- Hammer MF (1994) A recent insertion of an ALU element on the Y chromosome is a useful marker for human population studies. *Mol Biol Evol* 11:749–761
- (1995) A recent common ancestry for human Y chromosome. *Nature* 378:376–378
- Hammond HA, Jin L, Zhong Y, Caskey CT, Chakraborty R (1994) Evaluation of 13 short tandem repeat loci for use in personal identification applications. *Am J Hum Genet* 55: 175–189
- Helgason A, Sigurðardóttir S, Nicholson J, Sykes B, Hill EW, Bradley DG, Bosnes V, Gulcher JR, Ward R, Stefánsson K (2000) Estimating Scandinavian and Gaelic ancestry in the male settlers of Iceland. *Am J Hum Genet* 67:697–717
- 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
- Honda K, Roewer L, de Knijff P (1999) Male DNA typing from 25-year-old vaginal swabs using Y chromosomal STR polymorphisms in a retrieval request case. *J Forensic Sci* 44: 868–872
- Hurles ME, Irvén C, Nicholson J, Taylor PG, Santos FR, Loughlin J, Jobling MA, Sykes BC (1998) European Y-chromosomal lineages in Polynesians: a contrast to the population structure revealed by mtDNA. *Am J Hum Genet* 63: 1793–1806
- Jobling MA (1994) A survey of long-range DNA polymorphisms on the human Y chromosome. *Hum Mol Genet* 3: 107–114
- Jobling MA, Tyler-Smith C (1995) Fathers and sons: the Y chromosome and human evolution. *Trends Genet* 11:449–456
- Jorde LB, Watkins WS, Bamshad MJ, Dixon ME, Ricker CE, Seielstad MT, Batzer MA (2000) The distribution of human genetic diversity: a comparison of mitochondrial, autosomal, and Y-chromosome data. *Am J Hum Genet* 66: 979–988
- Kayser M, Brauer S, Weiss G, Schiefenhövel W, Underhill PA, Stoneking M (2001) Independent histories of Y chromosomes from Melanesia and Australia. *Am J Hum Genet* 68: 173–190
- Kayser M, Brauer S, Weiss G, Underhill P, Roewer L, Schiefenhövel W, Stoneking M (2000a) Melanesian origin of Polynesian Y chromosomes. *Curr Biol* 10:1237–1246
- Kayser M, Caglia A, Corach D, Fretwell N, Gehrig C, Graziosi G, Heidorn F, et al (1997) Evaluation of Y chromosomal STRs: a multicenter study. *Int J Legal Med* 110:125–133
- Kayser M, Roewer L, Hedman M, Henke L, Henke J, Brauer S, Krüger C, Krawczak M, Nagy M, Dobosz T, Szibor R, de Knijff P, Stoneking M, Sajantila A (2000b) 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: 1580–1588
- Kruskal JB (1964) Nonmetric multidimensional-scaling—a numerical-method. *Psychometrika* 29:115–129
- Lahermo P, Savontaus M-L, Sistonen P, Béres J, de Knijff P, Aula P, Sajantila A (1999) Y chromosomal polymorphisms reveal founding lineages in the Finns and the Saami. *Eur J Hum Genet* 7:447–458
- Mathias N, Bayès M, Tyler-Smith C (1994) Highly informative compound haplotypes for the human Y chromosome. *Hum Mol Genet* 3:115–123
- Melton T, Peterson R, Redd AJ, Saha N, Sofro ASM, Martinson J, Stoneking M (1995) Polynesian genetic affinities with Southeast Asian populations as identified by mtDNA analysis. *Am J Hum Genet* 57:403–414
- Mountain JL, Cavalli-Sforza LL (1997) Multilocus genotypes, a tree of individuals, and human evolutionary history. *Am J Hum Genet* 61:705–718
- Nakahori Y, Tamura T, Yamada M, Nakagome Y (1989) Two 47z[DXYS5] RFLPs on the X and the Y chromosome. *Nucleic Acids Res* 17:2152
- Nei M (1987) *Molecular evolutionary genetics*. Columbia University Press, New York
- Perez-Lezaun A, Calafell F, Seielstad M, Mateu E, Comas D, Bosch E, Bertranpetit J (1997) Population genetics of Y-chromosome short tandem repeats in humans. *J Mol Evol* 45:265–270

- Poloni ES, Semino O, Passarino G, Santachiara-Benerecetti AS, Dupanloup I, Langaney A, Excoffier L (1997) Human genetic affinities for Y-chromosome p49a,f/*TaqI* haplotypes show strong correspondence with linguistics. *Am J Hum Genet* 61:1015–1035
- Prinz M, Boll K, Baum H, Shaler B (1997) Multiplexing of Y chromosome specific STRs and performance for mixed samples. *Forensic Sci Int* 85:209–218
- Pult I, Sajantila A, Simanainen J, Georgiev O, Schaffner W, Pääbo S (1994) Mitochondrial DNA sequences from Switzerland reveal striking homogeneity of European populations. *Biol Chem Hoppe Seyler* 375:837–840
- Raymond M, Rousset F (1995) GENEPOP (version-1.2)—population-genetics software for exact tests and ecumenicism. *J Hered* 86:248–249
- Roewer L, Arnemann J, Spurr NK, Grzeschik K-H, Epplen JT (1992) Simple repeat sequences on the human Y chromosome are equally polymorphic as their autosomal counterparts. *Hum Genet* 89:389–394
- Roewer L, Kayser M, Dieltjes P, Nagy M, Bakker E, Krawczak M, de Knijff P (1996) Analysis of molecular variance (AMOVA) of Y-chromosome specific microsatellites in two closely related human populations. *Hum Mol Genet* 5:1029–1033
- Rosser ZH, Zerjal T, Hurles ME, Adojaan M, Alavantic D, Amorim A, Amos W, et al (2000) Y-chromosomal diversity in Europe is clinal and influenced primarily by geography, rather than by language. *Am J Hum Genet* 67:1526–1543
- Ruhlen M (1994) *The origin of language: tracing the evolution of the mother tongue*. John Wiley & Sons, New York
- Sala A, Penacino G, Corach D (1997) VNTR polymorphism in the Buenos Aires, Argentina, metropolitan population. *Hum Biol* 69:777–783
- Seielstad MT, Hebert JM, Lin AA, Underhill PA, Ibrahim M, Vollrath D, Cavalli-Sforza LL (1994) Construction of human Y-chromosomal haplotypes using a new polymorphic A to G transition. *Hum Mol Genet* 3:2159–2161
- Seielstad M, Bekele E, Ibrahim M, Toure A, Traore M (1999) A view of modern human origins from Y chromosome microsatellite variation. *Genome Res* 9:558–567
- Semino O, Passarino G, Oefner PJ, Lin AA, Arbuzova S, Beckman LE, De Benedictis G, Francalacci P, Kouvatsi A, Limborska S, Marcikiae M, Mika A, Mika B, Primorac D, Santachiara-Benerecetti AS, Cavalli-Sforza LL, Underhill PA (2000) The genetic legacy of paleolithic homo sapiens sapiens in extant Europeans: a Y chromosome perspective. *Science* 290:1155–1159
- Slatkin M (1995) A measure of population subdivision based on microsatellite allele frequencies. *Genetics* 139:457–462
- Stoneking M, Fontius JJ, Clifford SL, Soodyall H, Arcot SS, Saha N, Jenkins T, Tahir MA, Deininger PL, Batzer MA (1997) Alu insertion polymorphisms and human evolution: evidence for a larger population size in Africa. *Genome Res* 7:1061–1071
- Underhill PA, Jin L, Lin AA, Mehdi SQ, Jenkins T, Vollrath D, Davis RW, Cavalli-Sforza LL, Oefner PJ (1997) Detection of numerous Y chromosome biallelic polymorphisms by denaturing high-performance liquid chromatography. *Genome Res* 7:996–1005
- Underhill PA, Jin L, Zemans R, Oefner PJ, Cavalli-Sforza LL (1996) A pre-Columbian Y chromosome-specific transition and its implication for human evolutionary history. *Proc Natl Acad Sci USA* 93:196–200
- Underhill PA, Shen P, Lin AA, Jin L, Passarino G, Yang WH, Kauffman E, Bonne-Tamir B, Bertranpetit J, Francalacci P, Ibrahim M, Jenkins T, Kidd JR, Mehdi SQ, Seielstad MT, Wells RS, Piazza A, Davis RW, Feldman MW, Cavalli-Sforza LL, Oefner PJ (2000) Y chromosome sequence variation and the history of human populations. *Nat Genet* 26:358–361
- Vigilant L, Stoneking M, Harpending H, Hawkes K, Wilson AC (1991) African populations and the evolution of human mitochondrial DNA. *Science* 253:1503–1507
- Weissenbach J, Gyapay G, Dib C, Vignal A, Morissette J, Millasseau P, Vaysseix G, Lathrop M (1992) A second-generation linkage map of the human genome. *Nature* 359:794–801
- White PS, Tatum OL, Deaven LL, Longmire JL (1999) New, male-specific microsatellite markers from the human Y chromosome. *Genomics* 57:433–437
- Whitfield LS, Sulston JE, Goodfellow PN (1995) Sequence variation of the human Y chromosome. *Nature* 378:379–380
- Zerjal T, Dashnyam B, Pandya A, Kayser M, Roewer L, Santos FR, Schiefenhövel W, Fretwell N, Jobling MA, Harihara S, Shimizu K, Semjidmaa D, Sajantila A, Salo P, Crawford MH, Ginter EK, Evgrafov OV, Tyler-Smith C (1997) Genetic relationships of Asian and northern Europeans, revealed by Y-chromosomal DNA analysis. *Am J Hum Genet* 60:1174–1183