Ann. Hum. Genet. (1996), **60**, 401–408
Printed in Great Britain

Geographic clustering of human Y-chromosome haplotypes

A. RUIZ LINARES[1,2], K. NAYAR[1], D. B. GOLDSTEIN[3,4,5], J. M. HEBERT[1], M. T. SEIELSTAD[1], P. A. UNDERHILL[1], A. A. LIN[1], M. W. FELDMAN[3] AND L. L. CAVALLI SFORZA[1].

[1] *Department of Genetics and* [3] *Department of Biology, Stanford University, Stanford, CA 94305, USA*
[2] *Departamento de Bioquimica, Facultad de Medicina, Universidad de Antioquia, A. A. 1226, Medellin, Colombia*
[4] *Department of Biology, The Pennsylvania State University, University Park, PA 16802, USA*
[5] *Present address: Department of Zoology, University of Oxford, Oxford OX1 3PS, UK*

(Received 10.5.96. Accepted 21.6.96)

SUMMARY

Five polymorphic markers on the Y-chromosome (mostly microsatellites) were typed in 121 individuals from 13 populations around the world. With these markers 78 different haplotypes were detected. Haplotypes present more than once tend to be shared by individuals from the same population or continent. A reconstruction of haplotype phylogeny also indicates significant geographic structure in the data. Based on the similarity of the haplotypes, population relationships were examined and found to be largely concordant with those obtained with other markers. Even though the sample size and the number of markers are small, there is very signficant clustering of the haplotypes by continent of origin.

INTRODUCTION

A considerable amount of genetic data tends to support an expansion out of Africa of modern humans probably on the order of 100,000 years ago (1–8). In addition to the extensive literature on gene frequency analysis of autosomal genes (1–4), a great deal of work has been devoted to mitochondrial DNA (5–8). This molecule has been extensively studied due to the ease with which it can be examined, the clearcut nature of most of its genetic variation and its high substitution rate relative to autosomes. Furthermore, it is inherited only through females and undergoes no recombination. Therefore mtDNA behaves effectively as a single locus, allowing the analysis of female lineages.

The Y-chromosome is seen as the male counterpart to mtDNA as it is inherited exclusively through the male lineage and most of it does not recombine (9,10). However, two problems have delayed progress in using the Y-chromosome in evolutionary studies. First, polymorphic markers have been difficult to find on this chromosome (11–13). Second, until recently, most available Y-specific polymorphisms have been laborious to type and of a complex molecular nature, so that collection of data and the design of models with which to interpret them have been difficult (10, 14–16).

Thus, Y-chromosome markers that are easy to type and of a clearcut molecular nature would be desirable for population studies. Fortunately, such markers are now becoming available. Here we show that five well-defined Y-specific polymorphisms, mostly microsatellites, identify numerous Y-chromosome haplotypes in a sample of world populations. The geographic distribution of these haplotypes shows a significant level of population structure. When populations are compared, relationships are obtained that are largely concordant with those seen with larger data sets of autosomal markers. These results justify optimism for future studies based on

Table 1. *Number of chromosomes examined and of Y-chromosome haplotypes detected in 13 aboriginal populations*

Population	Number of chromosomes	Number of haplotypes
(1) Zaire Pygmy	10	10
(2) C.A.R. Pygmy	12	8
(3) Lisongo	4	4
(4) European	16	13
(5) Chinese	13	11
(6) Japanese	11	10
(7) Cambodian	15	12
(8) Melanesian	4	3
(9) New Guinean	6	5
(10) Australian	2	2
(11) Karitiana	11	4
(12) Surui	8	3
(13) Mayan	9	9
Total	121	78

microsatellite and other simple Y-chromosome polymorphisms.

MATERIALS AND METHODS

Populations and loci examined.

121 individuals from 13 populations were assayed (Table 1): 3 from sub Saharan Africa (Pygmies from Zaire and from the Central African Republic (C.A.R.). and the Lisongo population from the C.A.R.): a sample from Europe (mostly Northern Europeans): 3 from East Asia (Chinese, Japanese and Cambodian): 3 from Oceania (Melanesians, Australians and New Guinea Highlanders) and 3 from America (Mayans, and Surui and Karitiana from the Amazon basin in Brazil). In samples from these populations we typed 4 microsatellites, an Alu insertion-deletion (DYS287) and a nucleotide polymorphism (DYS271). Three of the microsatellites typed (YCAI, YCAII and YCAIII) were (CA)n repeats (16) and the third was a tetranucleotide repeat (DYS19) (17). In the populations examined here, YCAI was seen to be monomorphic. YCAII was typed by standard PCR (30 amplification cycles of 94 °C for 30 s, 52 °C for 1 min and 72 °C for 1 min were performed) while YCAIII and the tetranucleotide were typed using 'touch down' PCR: annealing temperatures were decreased from 62 °C by 0·5 °C successively for each of the first 14 cycles. The remaining 20 cycles were: 94 °C for 30 s, 55 °C for 45 s and 72 °C for 1 min. Amplification of both YCAII and YCAIII usually produced two Y specific bands each (with 7 and 9 different sizes, respectively), therefore they were assumed to be duplicated loci with inseparable alleles [K. Nayar *et al.*, in preparation]. Microsatellite allele sizes were determined using an ABI 373 automatic sequencer with dye-labeled primers. The Alu indel and the nucleotide polymorphism were typed as reported in (18,19).

Phylogenetic analyses

Genetic distances between haplotypes were estimated as one minus the proportion of shared alleles (or bands) between them (20). YCAII and YCAIII were given double weight in the calculations. Distances between population pairs were calculated as the average distance between haplotypes of the population pair considered, minus the average distance between haplotypes within each of the populations being compared. It is known that this distance has a lower variance than other genetic distances, when applied to microsatellite data (21). Trees were constructed from distance matrices using neighbor-joining NJ (22) or UPGMA (23). Tree reconstruction from distance matrices and parsimony analyses were performed using programs of the PHYLIP package (25). Other calculations were made using programs specifically written for those calculations.

RESULTS

Using the 5 markers, a total of 78 haplotypes were distinguished (Table 2). With the exception of the two Brazilian Amerind populations, all other populations have a number of haplotypes close to the number of chromosomes sampled, indicating a considerable level of within-population genetic variation. 61 haplotypes were seen only once, while of the 17 haplotypes present more than once, eight represented chromosomes

from the same population, seven were seen in chromosomes from two populations and two were in chromosomes from three populations. Only three haplotypes were found in populations from different continents: two were in populations from Asia and one was shared by a Mayan and a European. Thus, haplotype sharing is structured according to the geographic proximity of the sampled populations, with nearly complete structuring at the continental level.

HAPLOTYPE PHYLOGENY

Trees relating distinct haplotypes were constructed based upon the proportion of alleles shared between them (20). Figure 1 shows the Neighbour Joining tree relating the 78 haplotypes observed in our sample. The mid-point root of this tree separates a cluster [1 in Figure 1], containing 12 out of the 18 African haplotypes observed, from most other haplotypes. The segments that separate the clusters in Figure 1 are very short and the separation of the clusters is not very sharp. However, the composition of the clusters strengthens their validity considerably because it represents a clear separation by continents as shown by the contingency table in Table 3. In fact, it can clearly be seen that cluster 1 contains 12 of the 18 (67%) African haplotypes in the sample. The second cluster has a slight dominance of European haplotypes, although Europeans are also found in the third and fifth clusters. Almost half of the second cluster comprises Africans and Asians in equal numbers. The third cluster has 12 of the 15 American natives in the sample, while the fourth is almost entirely Asian. The fifth is mostly Asians and Oceanians.

Table 3 produces a highly significant chi-square [$p < 10^{-8}$, calculated as G with Williams correction (Sokal & Rohlf 1995, p. 738)]. Parsimony analysis using geographic state as the character also produces significant clustering of haplotypes by geographic origin (data not shown). It is likely that with more markers the separation into continents would be very much sharper, as seen for instance using 30 autosomal microsatellite markers from a sample of 148 individuals of both sexes from which the males greatly overlapped those in the present sample (20). It is possible, however, that the difference in structure results at least in part from linkage. As noted by Goldstein *et al.* (1995*b*), with unlinked markers, averaging within individuals will tend to place members of a population in the same cluster even with a substantial degree of admixture.

Based on the mean number of alleles shared between haplotypes, genetic distances among populations were calculated (21) and the possible phylogenetic relationships among populations inferred (Figure 2). In the tree constructed using UPGMA (Figure 2A), the first split separates African from non-African populations, consistent with the observation that the greatest genetic distance is seen between these groups (see also the Neighbour-Joining [NJ] tree in Figure 2B). Among non-African populations, the two Brazilian Amerind groups split first, followed by Europeans and Mayans, and finally Asian and Oceanians. It is likely that the apparent early split of the Karitiana and Surui might relate to a high level of drift in these populations which have been through recent population constrictions (this can also be seen by the long branches they show in the NJ tree). In addition, some level of Caucasoid admixture (particularly for the Mayans) is likely. Such demographic factors are known to affect the topologies of trees obtained by different methods (26). A major difference between the NJ and UPGMA trees is that the position of the Japanese is closer to Africans in the NJ tree. Furthermore, in the tree of haplotypes, several Japanese haplotypes are seen to cluster with African haplotypes (cluster 1 of Figure 1). These observations are difficult to explain but are probably related to the high frequency of the Alu insertion in Y-chromosomes from Japan and Africa (18,19,27).

DISCUSSION

It is encouraging that with a few well defined Y-specific polymorphisms, mainly microsatel-

Table 2. *Composite haplotypes of 121 individuals from 13 populations.*

(1–10: Zaire Pygmy. 11–22: C.A.R. Pygmy. 23–26: Lisongo. 27–42: European. 43–55: Chinese. 56–66: Japanese. 67–81: Cambodian. 82–85: Melanesian. 86–91:New Guinean. 92–93: Australian. 94–104: Karitiana. 105–112: Surui. 113–121: Mayan.Numbers for microsatellite loci (YCAI to DYS19) indicate allele sizes. For YCAII and YCAIII (a) and (b) respectively indicate the smallest and the largest band detected. For DYS287 (+) and (−) indicates presence or absence of the Alu insertion, respectively. For DYS271 A or G indicates the nucleotide present at the polymorphic site. N.D. = no data.)

Individual	YCAI	YCAIIa	YCAIIb	YCAIIIa	YCAIIIb	DYS19	DYS287	DYS271
1 P1G	128	151	153	193	195	192	+	A
2 P10G	128	151	151	193	197	196	+	G
3 P20G	128	151	155	193	197	188	+	G
4 P37G	128	151	155	197	201	188	+	G
5 P44G	128	151	155	193	197	200	+	G
6 P45G	128	151	153	191	191	192	−	N.D.
7 P94G	128	155	155	193	193	204	−	A
8 P102G	128	151	155	191	191	200	+	G
9 P103G	128	151	153	191	193	192	+	A
10 P106G	128	155	155	189	197	N.D.	+	A
11 P26	128	151	155	N.D.	N.D.	204	+	G
12 P73	128	151	155	199	203	200	+	G
13 P33	128	151	155	193	197	200	+	G
14 P115	128	151	155	193	197	204	+	G
15 P201	128	151	151	191	201	200	+	G
16 P202	128	151	151	199	205	200	+	G
17 P205	128	151	155	193	199	196	+	A
18 P206	128	151	155	193	197	204	+	G
19 P233	128	151	151	193	197	200	+	G
20 P235	128	151	155	193	197	200	+	G
21 P262	128	151	155	193	197	204	+	G
22 P272	128	151	155	193	199	200	−	A
23 F18	128	151	151	193	197	196	+	G
24 F19	128	151	151	193	197	200	+	G
25 F21	128	155	155	197	197	196	−	A
26 F30	128	151	155	193	197	200	+	G
27 5–158	128	151	159	199	199	196	−	A
28 7–195	128	N.D.	N.D.	197	197	200	−	A
29 8–52	128	151	157	197	201	188	+	N.D.
30 10–50	128	151	155	199	201	192	−	A
31 13–134	128	151	155	197	199	192	−	A
32 16–434	128	151	159	199	199	192	−	A
33 16–462	128	151	155	199	201	196	−	A
34 22–246	128	151	157	199	201	192	−	A
35 24–281	128	151	155	195	199	192	−	A
36 31–439	128	151	159	199	199	192	−	A
37 38–444	128	151	155	191	197	196	−	A
38 38–447	128	151	155	197	203	196	−	A
39 50–106	128	151	155	191	197	192	−	A
40 50–470	128	151	155	197	203	192	−	A
41 51–474	128	151	159	199	199	196	−	A
42 ITA8	128	151	159	197	199	192	−	A
43 CH4	128	151	155	189	189	192	−	A
44 CH5	128	151	159	191	191	196	−	A
45 CH10	128	151	157	197	197	196	−	A
46 CH12	128	151	157	191	191	196	−	A
47 CH17	128	151	157	N.D.	N.D.	200	−	A
48 CH31	128	151	155	N.D.	N.D.	192		
49 CH37	128	149	157	N.D.	N.D.	196	N.D.	A
50 CH38	128	159	159	191	193	196	N.D.	A
51 CH40	128	157	157	N.D.	N.D.	196	−	A
52 CH43	128	153	155	189	189	196	N.D.	A
53 CH44	128	151	153	N.D.	N.D.	192	−	A
54 CH47	128	157	157	197	199	196	−	A
55 CH49	128	151	155	N.D.	N.D.	196	−	A
56 JA3	128	151	155	195	199	196	+	A
57 JA4	128	151	159	195	195	196	−	A
58 JA11	128	151	157	N.D.	N.D.	200	−	A

Table 2. (*cont.*)

Individual	YCAI	YCAIIa	YCAIIb	YCAIIIa	YCAIIIb	DYS19	DYS287	DYS271
59 JA13	128	151	157	189	189	196	–	A
60 JA20	128	151	157	195	199	204	+	A
61 JA29	128	151	159	197	201	196	–	A
62 JA38	128	155	159	195	195	196	–	A
63 JA39	128	151	157	193	197	204	+	A
64 JA40	128	151	151	191	193	188	–	A
65 JA46	128	151	157	193	197	204	+	A
66 JA50	128	151	155	193	197	200	+	A
67 86–157	128	151	157	193	197	196	–	A
68 86–158	128	151	157	197	197	196	–	A
69 86–162	128	151	157	189	189	196	–	A
70 86–174	128	151	155	189	189	192	–	A
71 86–177	128	155	155	N.D.	N.D.	192	–	A
72 86–194	128	151	157	193	197	196	–	A
73 86–200	128	151	157	193	197	196	–	A
74 86–225	128	151	157	193	197	200	–	A
75 86–227	128	151	151	193	195	192	–	A
76 86–229	128	157	159	199	201	196	–	A
77 86–239	128	157	159	193	193	204	–	A
78 86–252	128	151	157	193	199	196	–	A
79 86–272	128	151	151	193	193	192	–	A
80 86–286	128	151	157	199	199	200	–	A
81 86–297	128	151	155	195	195	196	–	A
82 ME9	128	151	157	197	199	196	–	A
83 ME18	128	151	157	199	199	196	–	A
84 ME23	128	151	159	197	197	196	–	A
85 ME27	128	151	157	199	199	196	–	A
86 NG26	128	151	153	195	195	196	–	A
87 NG31	128	151	151	191	197	192	–	A
88 NG32	128	151	157	189	189	196	–	A
89 NG37	128	151	157	195	199	196	–	A
90 NG41	128	151	151	191	197	192	–	A
91 NG52	128	151	157	N.D.	N.D.	200	–	A
92 Aus6	128	151	155	195	195	196	–	A
93 Aus11	128	153	153	197	197	200	–	A
94 BI4	128	151	159	199	201	188	–	A
95 BI8	128	155	155	203	203	192	–	A
96 BI9	128	151	159	199	201	188	–	A
97 BI12	128	151	159	199	203	192	–	A
98 BI17	128	155	155	203	203	192	–	A
99 BI22	128	151	159	199	201	188	–	A
100 BI23	128	151	159	199	201	188	–	A
101 BI37	128	151	159	201	203	188	–	A
102 BI41	128	151	159	199	201	188	–	A
103 BI43	128	151	159	199	201	188	–	A
104 BI48	128	151	159	199	201	188	–	A
105 BI61	128	151	163	199	201	188	–	A
106 BI64	128	151	159	199	201	188	–	A
107 BI73	128	151	159	199	201	188	–	A
108 BI75	128	151	159	199	201	188	–	A
109 BI77	128	151	159	199	201	188	–	A
110 BI79	128	151	159	199	201	188	–	A
111 BI97	128	151	159	199	201	188	–	A
112 BI98	128	151	163	201	201	188	–	A
113 MI4	128	151	159	197	199	192	–	A
114 MI23	128	153	155	195	199	196	+	G
115 MI31	128	151	159	193	199	192	–	A
116 MI32	128	151	159	201	201	192	–	A
117 MI39	128	151	159	199	199	188	–	A
118 MI45	128	151	157	201	201	188	–	A
119 MI46	128	151	157	195	201	188	–	A
120 MI50	128	151	159	189	197	188	–	A
121 MI51	128	151	159	199	203	196	–	A

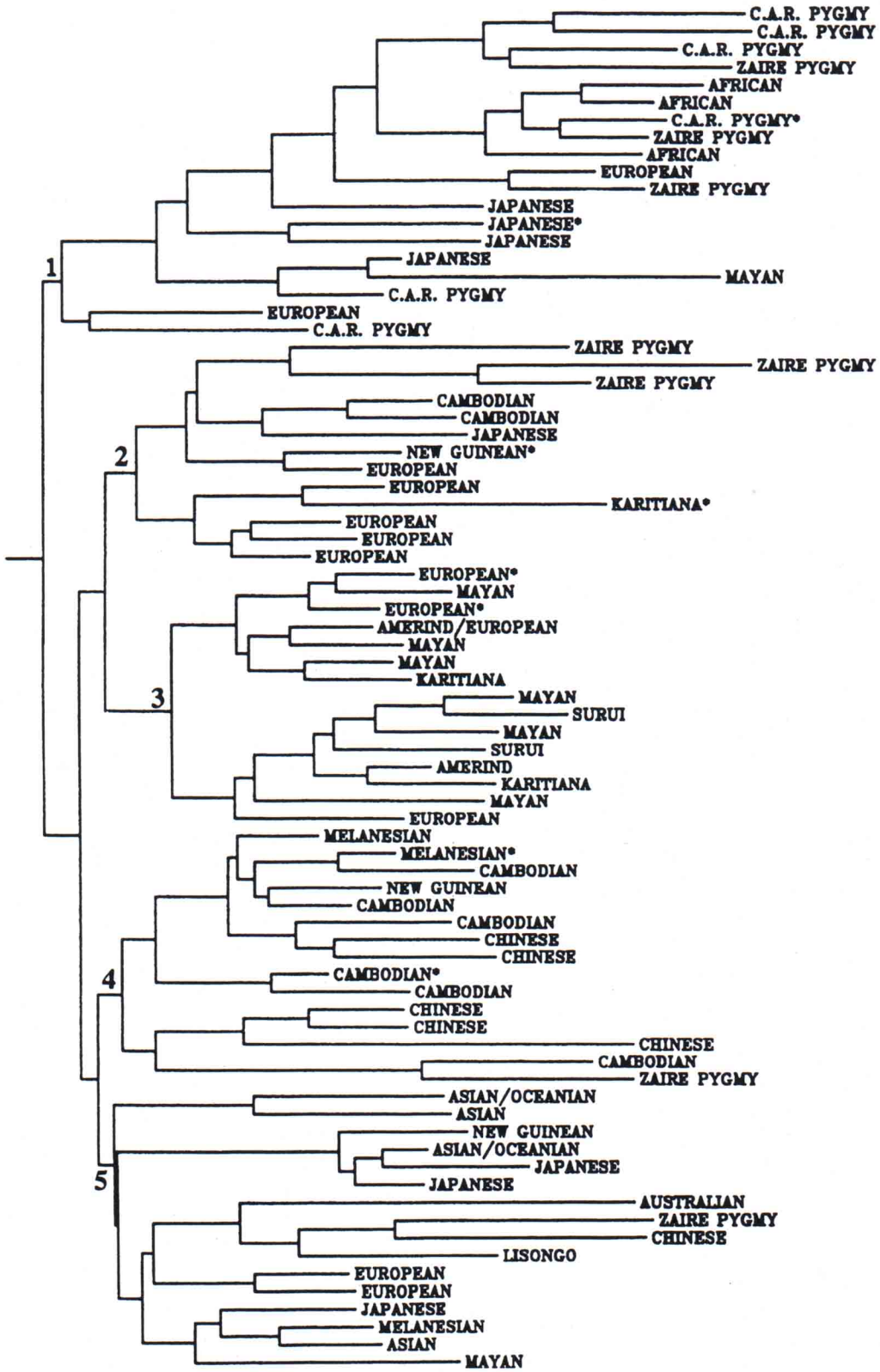

Fig. 1. Neighbor-Joining (22) tree relating 78 Y-chromosome haplotypes. Haplotypes bear the names of the populations where they were detected: with an asterisk if present in more than one chromosome. Haplotypes seen in more than one population within a continent have the continent name, those seen in populations from different continents have both continent names and are assigned to both in Table 3. The tree was rooted by midpoint. UPGMA (23) produced a tree of similar topology.

Table 3. *Continental distribution of the 5 major clusters indicated in Figure 1*

Cluster	AFR	EUR	ASIA	OCE	AM	Total
1	12	2	4	0	1	19
2	3	5	3	1	1	13
3	0	4	0	0	12	16
4	1	0	11	3	0	15
5	2	2	8	5	1	18
Total	18	13	26	9	15	81

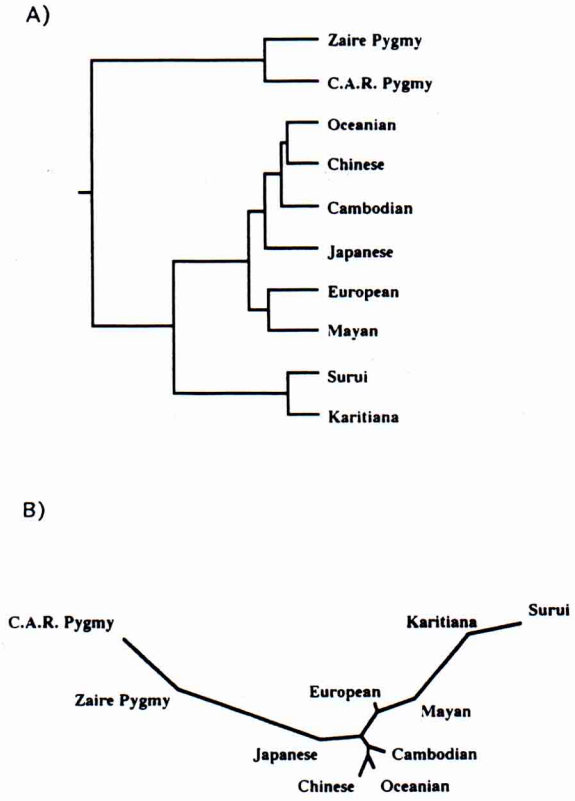

Figure 2. UPGMA (a) and Neighbor-Joining (b) trees relating 10 aboriginal populations based on Y-chromosome haplotypes. In order to increase sample sizes, data from Australia, Melanesia and New Guinea were pooled as Oceania, and the Lisongo data were pooled with those for the C. A. R. Pygmies (this last population is thought to represent an admixture of Pygmy and neighboring populations such as the Lisongo (2)).

lites, we have obtained population relationships largely concordant with those based on much larger data sets of autosomal markers (1, 2, 20, 28). As with other markers, the greatest genetic distance is seen between African and non-African populations; a finding compatible with an African origin for humans. Based on Y-chromosome data, there seems to be a particularly close relationship between Asian (mainly Chinese) and Oceanian populations.

With only five polymorphic markers, very significant geographic clustering of Y-chromosome haplotypes is observed. Using autosomal microsatellites, a very substantial degree of geographic clustering of multilocus genotypes has previously been observed (20). However, in that case, many more markers were used and they were mostly unlinked loci. Linkage of Y-specific markers will make phylogenetic trees very sensitive to migration, apart from the fact that haplotype and population trees need not agree, as has been seen with mtDNA haplotypes (5, 6). The fact that geographic clustering has been detected with only five markers, suggests that microsatellite-based Y-haplotype trees might have more structure than those seen with mitochondrial DNA.

Additional Y-microsatellites should facilitate estimates of the coalescence time for all human Y-chromosomes. Recent studies have attemped to estimate the coalescence time for Y-haplotypes based on sequence data. In one (13), a region of DNA was used in which no polymorphisms were observed, producing a value with a confidence interval ranging from 0 to 800.000 years. Two other studies have given dates ranging from 37·000 to 411·000 (31, 32). With the data reported here, we can already rule out a very recent coalescent event, as there has been enough time for variation to build up at these loci. While the small number of microsatellites included in our study permits only low statistical reliability, the data confirm the more general point that Y-linked microsatellites will allow such calculations, and that the estimated coalescence time will not be very low (Goldstein *et al. Mol. Biol. Evol.*, to appear). Additionally, microsatellites have the advantage that their high level of variation should allow the phylogenetic analysis of closely related populations as well as migration processes and other demographic events that are more difficult to examine with Y-nucleotide polymorphisms, given their rarity and smaller mutation rate.

We thank A. G. Clark for help during the later stages of this work.This work was supported by NIH grants to L. L. Cavalli-Sforza and to M. W. Feldman. A. Ruiz Linares was supported in part by Colciencias and the Programa de Reproduccion. Universidad de Antioquia.

REFERENCES

1. NEI. M. AND ROYCHOUDHURY. A. R. Genetic relationships and evolution of human races. *Evol. Biol.* **14**, 1–59 (1982).
2. CAVALLI SFORZA L. L., PIAZZA A., AND MENOZZI P. The History and Geography of Human Genes. (Princeton University Press. Princeton NY. 1994).
3. GOLDSTEIN D. B., RUIZ LINARES A., CAVALLI-SFORZA L. L. AND FELDMAN M. W. Genetic absolute dating based on microsatellites and the origin of modern humans *Proc. Natl. Acad. Sci. USA*. **92**, 6723–6727 (1995).
4. NEI. M. Commentary: Genetic support for the Out-of-Africa theory of human evolution. *Proc. Natl. Acad. Sci. USA*. **92**, 6720–6722 (1995).
5. CANN R. L., STONEKING M. AND WILSON A. C. Mitochondrial DNA and human evolution. *Nature* **325**, 31–36 (1987).
6. VIGILANT L., STONEKING M., HARPENDING H., HAWKES K. WILSON A. C. African populations and the evolution of human mitochondrial DNA. *Science* **253**, 1503–1507 (1991).
7. STONEKING M. Mitochondrial DNA and human evolution. *J. Bioenerg. Biomembr.* **26**, 251–259 (1994).
8. HORAI S., HAYASAKA K., KONDO R., TSUGANE K. AND TAKAHATA N. Recent African origin of modern humans revealed by complete sequences of hominoid mitochondrial DNAs. *Proc. Natl. Acad. Sci. USA* **92**, 532–536 (1995).
9. SPURDLE A. B. AND JENKINS T. The Y-chromosome as a tool for studying human evolution. *Current Opinion in Genetics and Development* **2**, 487–491 (1992).
10. JOBLING M. A., TYLER-SMITH C. Fathers and sons – the Y-chromosome and human evolution. *Trends Genet.* **11**, 449–456 (1995).
11. COOPER. D. N. AND SCHMIDTKE. J. DNA retriction fragment length polymorphism and heterozygosity in the human genome. *Hum. Genet.* **66**, 1–16 (1984).
12. MALASPINA P., PERSICHETTI F., NOVELLETTO A., IODICE C., TERRENATO L., WOLFE J., FERRARO M., PRANTERA G. The human Y-chromosome shows a low level of DNA polymorphism. *Ann. Hum. Genet.* **54**, 297–305 (1990).
13. DORIT R. L., AKASHI H., GILBERT W. Absence of polymorphism at the ZFY locus on the human Y-chromosome. *Science* **268**, 1183–1185 (1995).
14. LUCOTTE G., GUERIN P., HALLE L., LOIRAT F., HAZOUT S. Y chromosomal DNA polymorphisms in two african populations. *Am. J. Hum. Genet.* **45**, 16–20 (1989).
15. OAKEY R. AND TYLER-SMITH C. Y-chromosome DNA haplotyping suggests that most European and Asian men are descended from one of two males. *Genomics* **7**, 325–330 (1990).
16. MATHIAS N., BAYES M. AND TYLER-SMITH C. Highly informative compound haplotypes for the human Y-chromosome. *Human Molecular Genetics* **3**, 115 (1994).
17. ROEWER L., ARNEMANN J., SPURR N. K., GRZESCHIK K. H. AND EPPLEN J. T. Simple repeat sequences on the human Y-chromosome are equally polymorphic as their autosomal counterparts. *Human Genetics* **89**, 389 (1992).
18. HAMMER M. F. 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 (1994).
19. SEIELSTAD M. T., HEBERT J. M., LIN A. A., UNDERHILL P. A., IBRAHIM M., VOLLRATH D. AND CAVALLI-SFORZA L. L. Construction of human Y-chromosomal haplotypes using a new polymorphic A to G transition. *Hum. Mol. Gen.* **3**, 2159–61 (1995).
20. BOWCOCK A. M., RUIZ-LINARES A., TOMFOHRDE J., MINCH E., KIDD J. R. AND CAVALLI-SFORZA L. L. High resolution of human evolutionary trees with polymorphic microsatellites. *Nature* **368**, 455–457 (1994).
21. GOLDSTEIN D. B., RUIZ LINARES A., CAVALLI-SFORZA L. L. AND FELDMAN M. W. An evaluation of genetic distances for use with microsatellite loci. *Genetics* **139**, 463–471 (1995).
22. SAITOU N. AND NEI M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **4**, 406–425 (1987).
23. SOKAL R. R. AND MICHENER C. D. *Univ. Kansas Sci. Bull.* **38**, 1409–1438 (1958).
24. CAVALLI-SFORZA. L. L., MINCH E. AND MOUNTAIN J. L. (1992) Coevolution of genes and languages revisited. *Proc. Natl. Acad. Sci. USA* **89**, 5620–5624.
25. PHYLIP PACKAGE. VERSION 3·5c (distributed by J. Felsenstein. Department of Genetics. University of Washington).
26. RUIZ-LINARES A., MINCH. E., MEYER. D. CAVALLI-SFORZA. L. L. Analysis of classical and DNA markers for reconstructing human population history. In *The origins and past of Homo sapiens sapiens as viewed from DNA*. Hanihara. K., Brenner. S., eds. World Scientific Publishers. Singapore (1995).
27. HAMMER. M. F. AND HORAI S. Y chromosomal DNA variation and the peopling of Japan. *Am. J. Hum. Genet.* **56**, 951–962 (1995).
28. NEI M. AND ROYCHOUDHURY A. K. Evolutionary relationships of human populations on a global scale. *Mol. Biol. Evol.* **10**, 927–943 (1993).
29. WEBER J. L. AND WONG C. Mutation of human short tandem repeats. *Hum. Mol. Gen.* **2**, 1123–1128 (1993).
30. HUDSON R. R. Gene genealogies and the coalescent process *Oxford Surv. Evol. Biol.* **7**, 1–44 (1988).
31. HAMMER. M. F. A recent common ancestry for human Y chromosomes. *Nature* **378**, (1995) 376–378.
32. WHITFIELD. L. S., SULSTON. J. E. AND GOODFELLOW. P. N. Sequence variation of the human Y-chromosome. *Nature* **378**, (1995) 379–380.
33. SOKAL. R. R. AND ROHLF. F. J. *Biometry: The Principles and Practice of Statistics in Biological Research*. Third edition. W. H. Freeman and Company, New York (1995).