



Genes and languages in Europe: an analysis of mitochondrial lineages.

A Sajantila, P Lahermo, T Anttinen, et al.

Genome Res. 1995 5: 42-52

Access the most recent version at doi:[10.1101/gr.5.1.42](https://doi.org/10.1101/gr.5.1.42)

References

This article cites 23 articles, 15 of which can be accessed free at:
<http://genome.cshlp.org/content/5/1/42.refs.html>

Article cited in:

<http://genome.cshlp.org/content/5/1/42.abstract.html#related-urls>

Email alerting service

Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or [click here](#)

To subscribe to *Genome Research* go to:
<http://genome.cshlp.org/subscriptions/>

RESEARCH

Genes and Languages in Europe: An Analysis of Mitochondrial Lineages

Antti Sajantila,^{1,2,11} Päivi Lahermo,³ Tiiu Anttinen,¹ Matti Lukka,⁴
Pertti Sistonen,⁵ Marja-Liisa Savontaus,³ Pertti Aula,³ Lars Beckman,⁶
Lisbeth Tranebjaerg,⁷ Tobias Gedde-Dahl,⁸ Laurie Issel-Tarver,⁹
Anna DiRienzo,¹⁰ and Svante Pääbo¹

¹Zoological Institute, Ludwig Maximilians University, Munich, Germany; ²Department of Human Molecular Genetics and ⁴Department of Immunology, National Public Health Institute, Helsinki, Finland; ³Department of Medical Genetics University of Turku, Turku, Finland; ⁵Finnish Red Cross Blood Transfusion Service, Helsinki, Finland; ⁶Department of Medical Genetics, University of Umeå, Umeå, Sweden; ⁷Department of Medical Genetics, University of Tromsø, Tromsø, Norway; ⁸Institute of Forensic Medicine, University of Oslo, Oslo, Norway; ⁹Department of Molecular and Cell Biology, University of California, Berkeley, California 94720; ¹⁰Northwestern University, Evanston, Illinois 60208-1310

When mitochondrial DNA sequence variation is analyzed from a sample of 637 individuals in 14 European populations, most populations show little differentiation with respect to each other. However, the Saami distinguish themselves by a comparatively large amount of sequence difference when compared with the other populations, by a different distribution of sequence diversity within the population, and by the occurrence of particular sequence motifs. Thus, the Saami seem to have a long history distinct from other European populations. Linguistic affiliations are not reflected in the patterns of relationships of mitochondrial lineages in European populations, whereas prior studies of nuclear gene frequencies have shown a correlation between genetic and linguistic evolution. It is argued that this apparent contradiction is attributable to the fact that genetic lineages and gene frequencies reflect different time perspectives on population history, the latter being more in concordance with linguistic evolution.

European populations show a relatively uniform pattern of nuclear gene frequencies when blood groups and enzyme polymorphisms are analyzed. However, for example, the Basques in the northern part of the Iberian Peninsula, the Sardinians in the Mediterranean, the Icelanders in the Atlantic Ocean, the Finns in eastern Fenno-Scandinavia, and the Saamis in northern Fenno-Scandinavia have been shown to be outliers (Guglielmino et al. 1990; Cavalli-Sforza et al. 1994). The fact that the Basques, the Finns, and the Saami speak non-Indo-European languages (Ruhlen 1987) supports the notion that genes and languages coevolve, which has been shown to be the case when human history is reconstructed on a worldwide scale from nuclear gene frequencies (Cavalli-Sforza et al. 1988, 1992).

The mitochondrial DNA (mtDNA) has several properties useful for the reconstruction of human population history. In particular, because it is almost exclusively maternally inherited (Giles et al. 1980), each individual carries only one type of mtDNA. Consequently, the analysis of molecular lineages is not perturbed by recombination. Furthermore, the high rate of evolution of the mitochondrial DNA (Brown et al. 1979; Ward et al. 1991) permits the accumulation of substantial numbers of substitutions over time periods relevant for human population history. To date, studies concerning mitochondrial diversity in European populations have generally been limited to Indo-European-speaking populations. We have determined a 360-bp-long, hypervariable part of the mitochondrial control region from five Finno-Ugric speaking groups: the Saami, the Finns, the Karelians, the Estonians, and the Volga-Finnic and compared those with four Indo-European speaking groups: the Icelanders, the Sardinians (DiRienzo and Wilson

¹¹Corresponding author. Present address: Zoological Institute, Ludwig Maximilians University, D-80021 Munich, Germany.
E-MAIL sajanti@zi.biologie.uni-muenchen.de; FAX 49-89-5902474.

1991), the British (Piercy et al. 1993), and the Swiss (Pult et al. 1994), as well as the Basques (Bertranpetit et al. 1995), who form a linguistic isolate (Table 1).

RESULTS

Pairwise Comparisons of mtDNA Sequences in Europe

Table 2 shows the mean pairwise sequence differ-

ences within as well as among the populations studied. The pairwise differences within populations vary between 3.24 and 5.03; those among pairs of populations vary between 3.49 and 5.29. Thus, the overall amount of sequence difference within populations is not different from that between populations. This is in contrast to what is observed, for example, in African populations, where differences within three populations vary between 3.68 and 8.74 and those between populations range from 9.07 to 9.95 (Vigilant et al.

Table 1. Presence of "Saami-specific motif" and substitution at position 1298 in various populations

Language family	Ethnic group	No. of individuals	Motif		16298								
			(no.)	(%)	(no.)	(%)							
Finno-Ugric	Saamic	115	43	37.4	53	46.1							
							Finland						
							Inari Lake	22	7	31.8	2	9.1	
							Skolt	47	17	36.2	28	59.6	
							Sweden						
							Norrbotnen	25	8	32.0	15	60.0	
							Norway						
							Karasjok	21	11	52.0	8	38.1	
							Finnic	161	6	3.7	11	6.8	
							Finnish	50	1	2.0	6	12.0	
Karelian	83	5	6.0	5	6.0								
Estonian	28	0	0.0	0	0.0								
	Maris, Moksha												
Volga-Finnic		34	0	0.0	0	0.0							
Indo-European	Germanic	113	0	0.0	4	1.5							
							German	Swiss ^a	74	0	0.0	4	5.4
							Icelandic	Icelanders	39	0	0.0	0	0.0
							Italic	169	0	0.0	4	2.6	
							English	British ^b	100	0	0.0	3	3.0
							Sardinian	Sardinians ^c	69	0	0.0	1	1.4
							Basque	Basque^d	45	0	0.0	5	11.1
							Afroasiatic, Nilo-Saharan, Niger-Kordofanian, and Khoisan	12 populations ^e	362	0	0.0	1	<0.1
							Na-Dene	2 populations ^f	62	0	0.0	4	6.4
							Eskimo-Aleut	4 populations ^f	53	0	0.0	0	0.0
Amerind	2 populations ^f	103	0	0.0	15	14.6							
Chukchi-Kamchatkan	Chukchi ^f	48	0	0.0	5	10.4							
Asian	Japanese ^g	47	0	0.0	6	12.8							

^aPult et al. (1994); ^bPiercy et al. (1993); ^cDi Rienzo and Wilson (1991); ^dBertranpetit et al. (1995); ^eVigilant et al. (1991) and E. Watson et al. (unpubl.); ^fWard et al. (1991, 1993; and unpubl.); ^gHorai et al. (1990).

Table 2. Mean pairwise differences within (diagonal) and between (above diagonal) populations, and genetic distances $\times 100$ (below diagonal) between populations

	saa	fin	kar	est	vol	swi	bri	sar	ice	bas
saa	3.99	4.48	4.42	4.72	4.89	4.44	4.89	4.76	5.29	4.20
fin	54	3.90	3.96	4.14	4.21	3.80	4.17	4.10	4.56	3.62
kar	44	3	3.96	4.13	4.19	3.84	4.22	4.14	4.59	3.65
est	54	1	3	4.36	4.33	4.02	4.39	4.28	4.72	3.85
vol	72	8	3	3	4.36	4.03	4.38	4.33	4.71	3.88
swi	63	3.5	4.5	2.5	3.5	3.63	4.02	3.96	4.42	3.49
bri	71	3.5	5.5	2.5	1.5	2	4.37	4.31	4.70	3.85
sar	66	4	5	1	4	3.5	1.5	4.22	4.68	3.78
ice	78	9.5	9.5	2.5	1.5	9	0	5.5	5.03	4.23
bas	58	5	5	5	8	5.5	4.5	5	9.5	3.24

Population abbreviations are as follows: (saa) Saami; (fin) Finns; (kar) Karelians; (est) Estonians; (vol) Volga-Finnic; (swi) Swiss; (bri) British; (sar) Sardinians; (ice) Icelanders; (bas) Basques.

1991; Pult et al. 1994). However, the Saami tend to stand out in the amount of sequence variation that they display when compared with other European populations. When the distances between populations are corrected for the within-population differences (Table 2), this becomes particularly apparent: The distances between the Saami and other groups vary between 44 and 78, whereas distances between the others range from 0 to 9.5. When a permutation-based test for detecting genetic differentiation between populations (Hudson et al. 1992) is performed, the Saami are found to be different from the other European populations in eight of the nine comparisons (P of $K_{st} \leq 0.01$), the only exception being the Basques (P of $K_{st} = 0.059$). The mitochondrial gene pool of the Saami is therefore distinct from that of other European populations.

Distributions of Pairwise Sequence Differences Within Populations

The distributions of pairwise sequence differences within populations are shown in Figure 1. With the exception of the Saami, all populations have distributions that tend to have a single mode at three to five substitutions. In contrast, in spite of the relatively large number of sequences determined, the Saami exhibit no defined mode.

Distributions of pairwise differences reflect the demographic history of populations in ways that are still not fully understood. However, one scenario that can give rise to a unimodal distribution of pairwise sequence differences is an ex-

pansion of population size, whereas populations that have been of constant size over a long period of time can exhibit multimodal distributions (Slatkin and Hudson 1991; Marjoram and Donnelly 1994). Although other demographic scenarios might also give rise to unimodal and multimodal distributions (Marjoram and Donnelly 1994), the distributions of pairwise sequence differences indicate that the demographic history of the Saami may be different from that of the other European populations.

A further difference between the Saami and other European populations is that 18% of the pairwise comparisons among Saami are identical, whereas the highest fraction of identical comparisons among the other European populations is 5.6% (Fig. 1). The observation that the relative amount of identical lineages is greater in Saami supports the idea that the Saami may have not experienced an expansion in size, whereas other European populations have done so.

Sequence Motifs Common in Saami

Of 115 Saami mitochondrial sequences, 43 (37%) share a combination of three substitutions at positions 16144, 16189, and 16270 (cf. Andersson et al. 1981). This sequence motif was found in Saami living in Finland and Sweden as well as in Norway at frequencies ranging from 32% to 52%. When other populations from around the World (Table 1) were investigated for the presence of this lineage, it was found in 1 of 50 Finns and 5 of 83 Karelians but in 0 of 62 other Finno-Ugric speakers. It is also absent among 282 Indo-

EUROPEAN mtDNA DIVERSITY

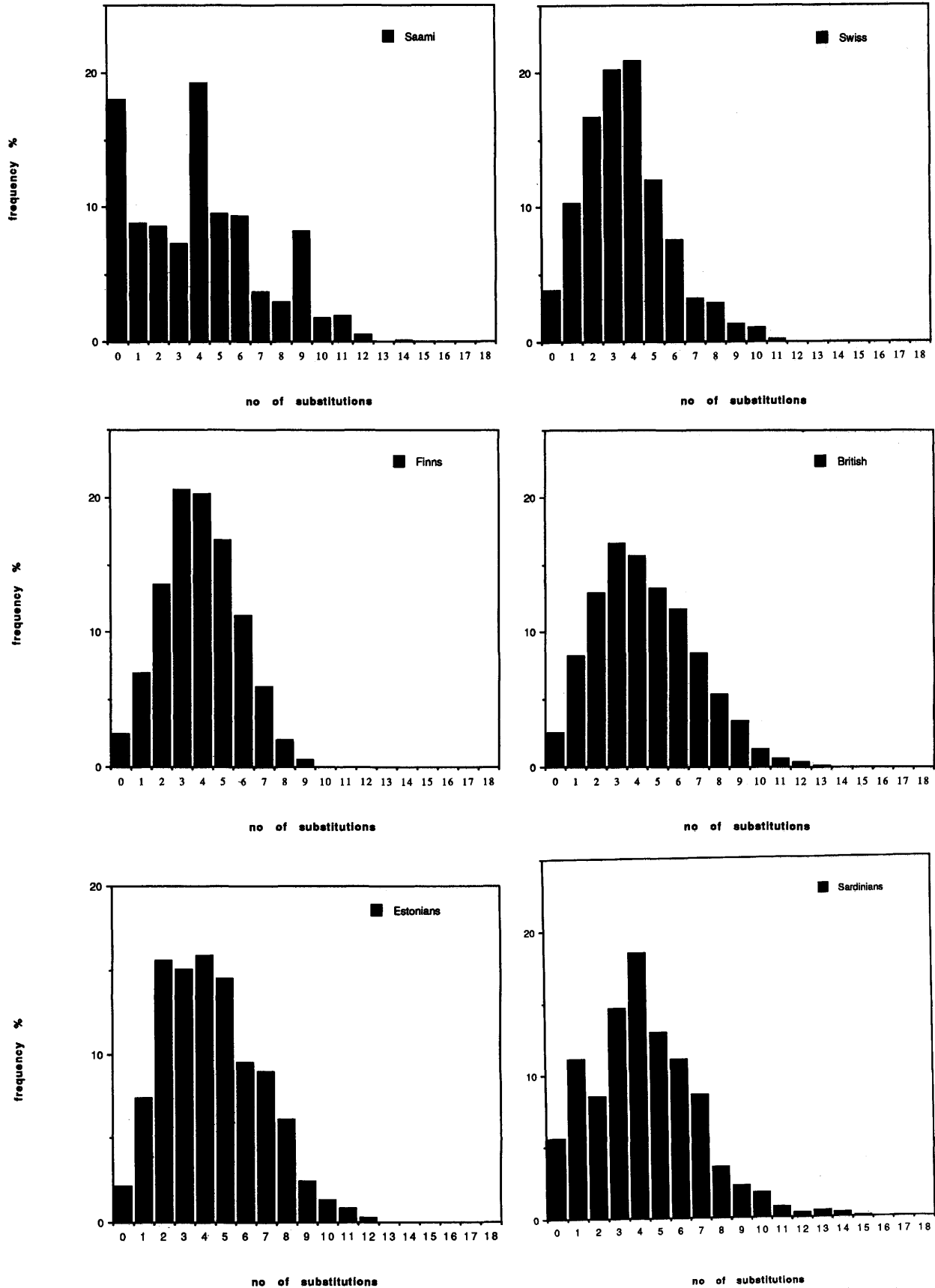


Figure 1 (Fig. 1 and legend continued on following page.)

SAJANTILA ET AL.

European speakers, 45 Basques, 101 indigenous Siberians, 47 Japanese, 165 Native Americans, and 362 Africans. Thus, the presence of this sequence motif at high frequencies seems to be restricted to Saami.

Of those Saami who do not carry sequences characterized by the three substitutions, 74% share a single substitution at position 16298 (C → T transition). This substitution is found in Saami populations with frequencies varying from

9% to 60% (Table 1). However, it exists in other populations as well, although generally at a lower frequency (0%–12.8%), a fact that can be attributable to parallel substitutions. It is noteworthy, though, that it exists also in Finns and Karelians at 12% and 6%, respectively.

In total, 83.5% of the Saami carry either the sequence motif or the substitution at position 16298, an indication that the Saami share a common history distinct from that of other groups

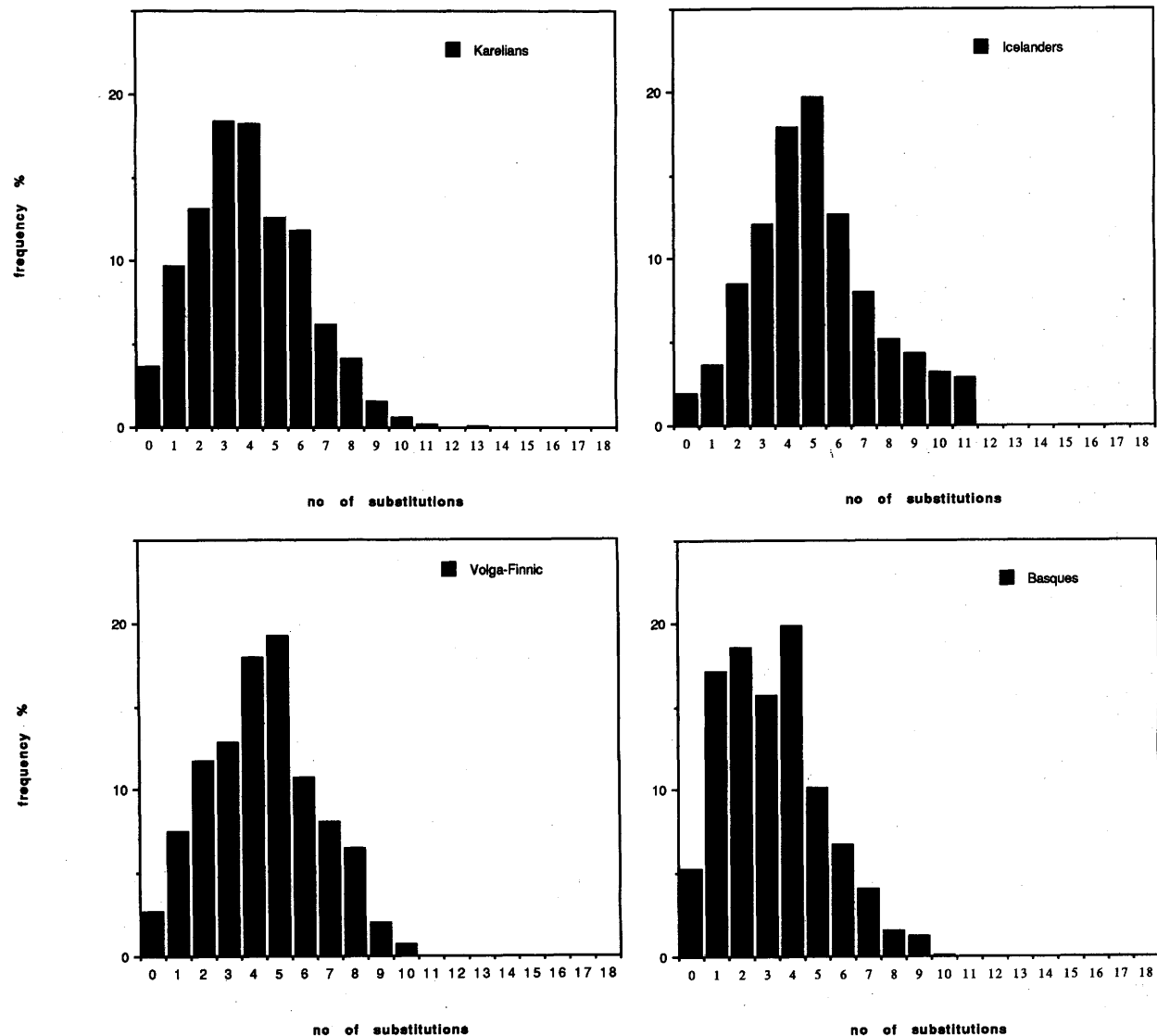


Figure 1 Distribution of pairwise sequence differences within 10 European populations. The comparisons were performed for 115 Saami, 50 Finns, 83 Karelians, 28 Estonians, 34 Volga-Finnic speakers, 100 British, 74 Swiss, 39 Icelanders, 69 Sardinians, and 45 Basques. The abscissa gives the number of substitutions in a pair, and the ordinate gives the relative frequency (%) of pairs. Sequences are from Di Rienzo and Wilson (1991; Sardinians), Piercy et al. (1993; British), Pult et al. (1994; Swiss), and Bertranpetit et al. (1995; Basques). Previously unpublished Saami, Finnish, Karelian, Estonian, Volga-Finnic, and Icelandic sequences have been submitted to the EMBL data library under accession numbers X90001–X90115, X90116–X90165, X90166–X90248, X90249–X90276, X90277–X90310, and X90311–X90349, respectively.

included in this study. The presence of the motif as well as the substitution at position 16298 at low frequency in the Finns and the Karelians are most likely the result of admixture between these populations and the Saami.

mtDNA Sequence Comparisons Among Saami Groups

When mitochondrial DNA sequences were compared within the Saami groups, the distributions of pairwise differences in all groups were similar in that they did not show a distinctive mode and a had high proportion of identical sequences (data not shown). However, Saami from Inari Lake (Finland) showed a mean of 6.52 substitutions within the population, whereas the mean numbers of substitutions within Skolt Saami (Finland), and Saami from Norrbotten (Sweden) and from Karasjok (Norway) were 3.12, 3.25, and 2.76, respectively (Table 3). In comparisons between the groups, Inari Lake Saami stand out with mean distances of 5.13–5.71, whereas comparisons between other Saami groups vary between 3.08 and 3.19. Thus, individual Saami groups can differ substantially from other groups with respect to mitochondrial sequence diversity. When tested for genetic differentiation (Hudson et al. 1992), the Inari Lake Saami were found to differ from the Skolt Saami and the Saami from Norrbotten (P of $K_{st} \leq 0.01$), whereas the other Saami groups did not reveal any significant differences. The differentiation that exists among Saami groups represents a further distinction between those and other European populations, which show little if any differentiation among them and very similar amounts of intrapopulation differences.

Table 3. Mean pairwise differences within (diagonal) and between (above diagonal), plus genetic distances $\times 100$ between different Saami populations

	Inari	Skolt	Norrbotten	Karasjok
Inari	6.52	5.63	5.71	5.13
Skolt	81	3.12	3.13	3.08
Norrbotten	82	5.5	3.25	3.19
Karasjok	49	14	18	2.76

Population abbreviations are as follows: (Inari) Saami from Inari Lake region, Finland; (Skolt) Skolt Saami from eastern Finland; (Norrbotten) Saami from Norrbotten, Sweden; (Karasjok) Saami from Karasjok, Norway.

Patterns of Sharing of Mitochondrial Lineages

To investigate the pattern of sharing of mitochondrial sequences between European populations, the frequency of sequences shared between groups was plotted for the groups in which the sequences occur. Some representative examples are shown in Figure 2. It can be seen that the sequences shared between the Saami and other populations tend to be substantially more frequent in one of the groups than in the other. This is the case in comparisons both with groups that are similar to the Saami in that they speak Finno-Ugric languages and in comparison with other groups. In contrast, in comparisons among European populations that do not involve Saami, the number of shared sequences is greater and the relative proportion of each lineage in the respective population is more equal than for lineages shared with the Saami. Whereas the latter pattern is consistent with the fact that the populations are derived from a common gene pool, the pattern observed in comparison between Saami and the other European populations may reflect sharing of lineages resulting from a limited amount of gene flow between populations and/or convergent evolution of the DNA sequences. Thus, the pattern of lineage sharing between Saami and other European populations is more consistent with gene flow, whereas that among other populations is more consistent with shared ancestry or extensive gene flow.

DISCUSSION

The Genetic History of Saami

The Saami mitochondrial gene pool differs substantially from that of other populations studied.

In contrast, other European populations, including Finno-Ugric speaking groups other than the Saami, Indo-European speaking groups, and the Basques, show no major characteristics that are not also found in the other populations. The data thus argue that the Saami have a history distinct from that of other groups in Europe.

The Saami speak several mutually unintelligible languages and have different current subsistence patterns (Sinor 1988). They are thus linguistically and culturally diverse,

SAJANTILA ET AL.

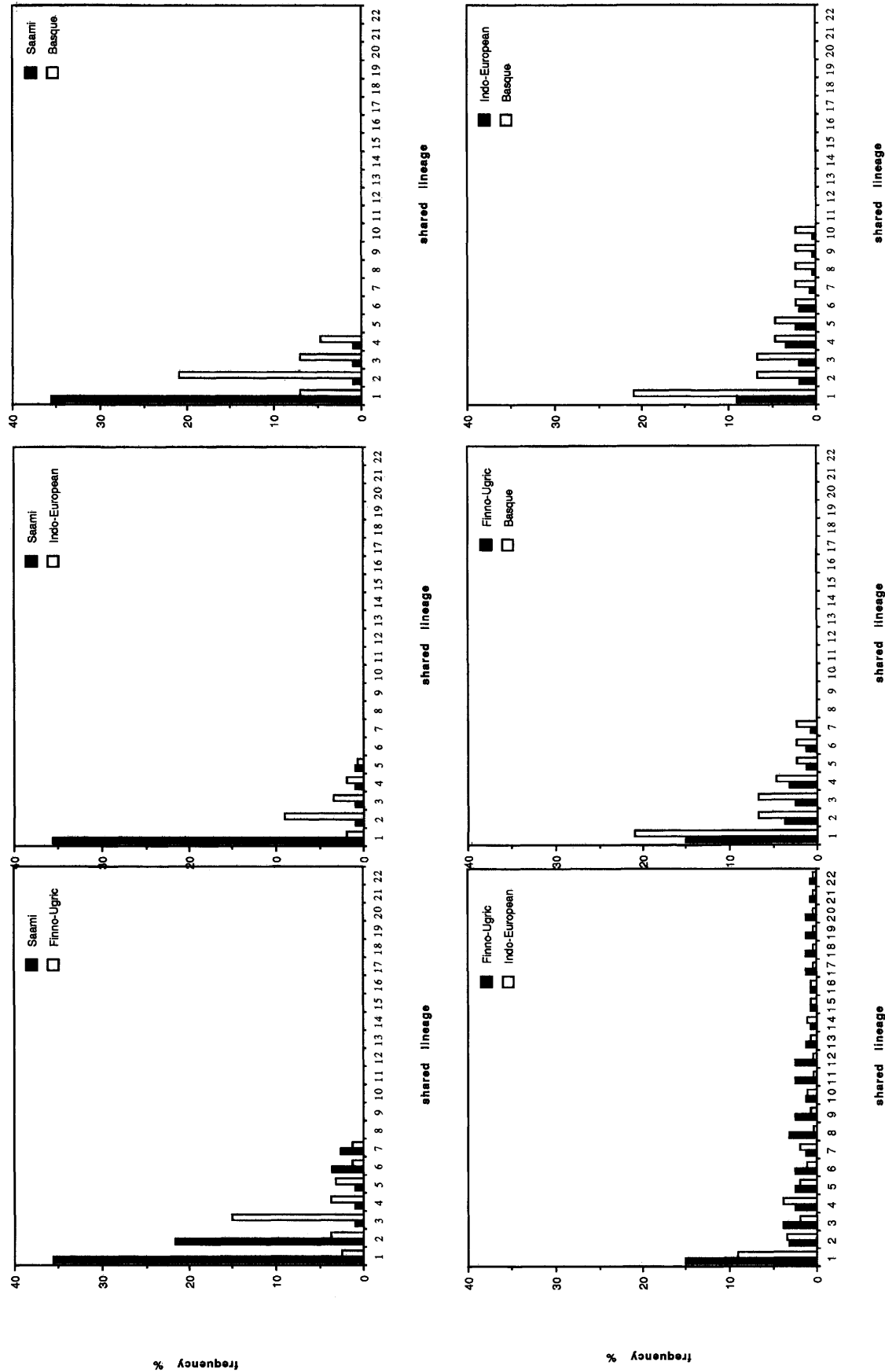


Figure 2 Pattern of mitochondrial lineage sharing among 115 Saamis, 183 Finno-Ugric speakers, excluding the Saamis (50 Finns, 83 Karelians, 28 Estonians, and 34 Volga-Finnic speakers), 282 Indo-European speakers, and 45 Basques. The lineages shared between the groups analyzed are given on the abscissa starting with the most common lineage. The relative proportion of each of the lineages in the respective populations is shown on the ordinate.

EUROPEAN mtDNA DIVERSITY

and it may be asked to what extent the mitochondrial gene pool is homogeneous among Saami. As shown above, two mitochondrial lineages, common among Saami but rare or absent elsewhere in Europe, exist in all Saami groups studied. Furthermore, the mean pairwise difference among mitochondrial sequences in the Saami groups (with the exception of the Inari Lake Saami) varies between 3.08 and 3.19, whereas that between other European groups varies between 3.49 and 4.72. Thus, the Saami not only have a history distinct from that of other Europeans but share this history among themselves.

A possible exception is the Inari Lake Saami, for whom the mean pairwise difference both within the population and with other Saami groups is exceptionally high (Table 3). Although the sequences from the Inari Lake Saami have the motif characterized by three substitutions at a frequency comparable with the other Saami, the substitution at position 16298 occurs at a frequency of 9%, whereas it occurs at 38% or more in the other groups. This may be attributable either to genetic drift or to admixture of the Inari Lake Saami with other populations.

The unique position of the Saami in the genetic landscape of Europe could mean that they are an old population in Europe, which diverged from other European populations prior to subsequent linguistic and cultural diversification. Alternatively, they may have come to Europe from another, currently unknown region. The latter possibility does not, however, get support from the survey of human populations that has so far failed to identify the mitochondrial lineages typical of Saami at any appreciable frequencies in other populations (Table 1). However, further sampling, particularly of Asian populations, is necessary to clarify this.

Gene Frequencies, Genetic Lineages, Languages, and Population History

European populations speak languages that belong to three linguistic phyla. However, no distinction according to language can be observed in the mitochondrial gene pool of Europe (Table 2). Rather, although the Saami languages are related most closely to Finnish and Estonian (Ruhlen 1987; Sinor 1988), the mitochondrial lineages in the latter are related most closely to those in populations that speak languages belonging to other linguistic phyla than to those in

Saami. At first glance, this seems to contradict the coevolution of genes and languages demonstrated for nuclear markers (Cavalli-Sforza et al. 1988, 1992). However, findings from gene frequencies and genetic lineages are not necessarily contradictory. Rather, they may reflect two different perspectives on population history.

Gene frequencies can be changed dramatically by reductions in population size. In contrast, genetic lineages that survive through such changes will still reflect the temporally deeper relationships of populations (Fig. 3). An illustrative example of this may be the Finns. Studies of nuclear genes in the Finnish population suggest that Finns have gone through a substantial bottleneck in size (Nevanlinna 1972). This is indicated by >30 recessive autosomal diseases that have high carrier frequencies in Finns but are al-

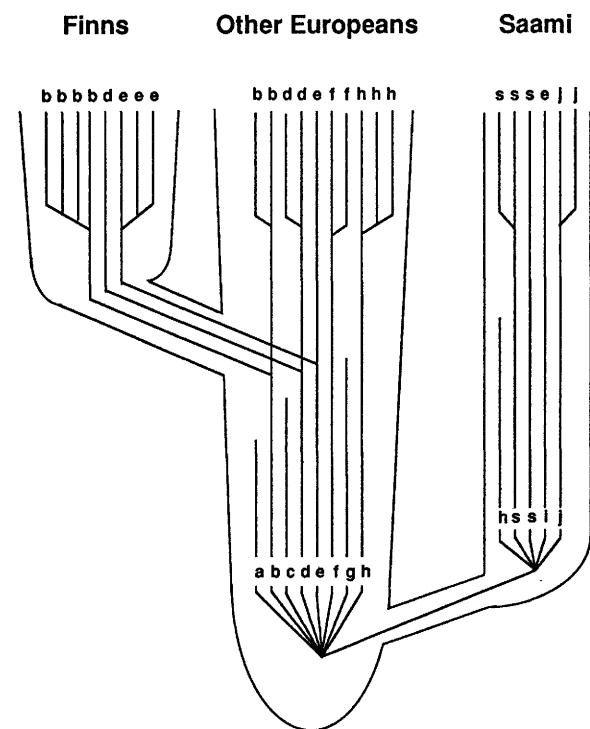


Figure 3 Schematic illustration of one difference between analyses of gene frequencies and molecular lineages. A population (e.g., Finns) that shares its recent ancestry with another one (e.g., most other Europeans) has experienced a bottleneck in population size. This is detectable as differences in gene frequencies when compared with the other populations. However, both populations share many genetic lineages. In contrast, a population that has been separated from the others long ago (e.g., Saami) displays different genetic lineages.

most absent in other populations (de la Chapelle 1993) including, as far as is known, the Saami (Tollersrud et al. 1994). Conversely, some common recessive diseases that have high carrier frequencies in other European populations, such as cystic fibrosis and phenylketonuria, are virtually absent in Finns. Furthermore, in phylogenetic trees based on protein and blood group analyses, Finns appear as outliers (Cavalli-Sforza et al. 1994). In stark contrast, when mitochondrial lineages are analyzed, the number of shared lineages between the Finns and Indo-European speakers is high and the distance between the groups no greater than between other European populations, although the frequencies of particular lineages may vary in the populations. Thus, the analysis of mitochondrial lineages may reflect the affiliation of the Finnish population to other European populations prior to the founder effects that caused gene frequencies to change.

Studies of gene frequencies have shown that in addition to the Saami and the Finns, Icelanders, Sardinians, and Basques are among genetic outliers in Europe (Cavalli-Sforza et al. 1994). However, in contrast to Saami, but analogous to the Finns, mitochondrial lineage analyses show these populations to belong to a gene pool common to most European groups. The discrepancy between these results may be explained if these populations have undergone bottlenecks in population size in prehistoric times. For the Icelanders, such a scenario is easily reconcilable with the known colonization of Iceland from Scandinavia and the British Isles during the Iron Age. Similarly, Sardinia has probably been colonized earlier from mainland European populations. Presumably, in both instances, the colonization and subsequent relative isolation resulted in a reduction in effective population size that is detected in divergent gene frequencies, whereas the original affiliations of these populations are reflected in the close relationship of their mitochondrial lineages to those of most other European populations. In the case of the Basques, the reduction in population size may be the result of their cultural and linguistic isolation. Yet, mitochondrial lineages shared with other populations (Bertranpetit et al. 1995) indicate that they may be of the same origin as other populations in Europe. If this is true, future studies of genetic lineages in the nuclear genome will reveal a pattern similar to that for mitochondrial lineages not only in Finns, Icelanders, and Sardinians but also in Basques.

Genetic and Language Evolution

A tempting hypothesis is that populations that go through bottlenecks in size will tend to accelerated their rate of language evolution as a result of drift-like linguistic process. The result would be a congruence between trees based on languages and gene frequency. In contrast, analyses of genetic lineages would tend to disagree with language trees. Such a disagreement has been observed not only in Europe but also in the New World (Ward et al. 1993). Patterns of relationships of genetic lineages would instead reflect temporarily deeper affinities of populations relatively unperturbed by founder effects and other changes in effective population size. Thus, the analysis of gene frequencies may be the instrument of choice to study population history at a time depth that is in agreement with that of linguistic history, which probably is on the order of a few thousand years. In contrast, analyses of genetic lineages may be used to study the older history of populations. The time depth that a particular genetic lineage reflects will be determined by its evolutionary rate. For the mitochondrial region studied here, estimates of the evolutionary rate vary between approximately one substitution in 13,000 years (Ward et al. 1991; Horai et al. 1995) to one substitution in 5000 years (Lundström et al. 1992). Thus, the temporal perspective in which the distinct history of the Saami, and the shared history of other populations in Europe, has to be seen is on the order of tens of thousands of years.

METHODS

DNA Samples, Amplification, and Sequencing

DNA was extracted according to standard methods from blood samples of 115 Saami, 50 Finns, 83 Karelians, 28 Estonians, 34 Volga-Finnic, and 39 Icelanders. As far as possible, it was assured that individuals donating blood were unrelated. PCR was performed in a programmable heating block (PE 4800, Perkin-Elmer) using 20 nmoles of each dNTP, 50 pmoles of each primer (L15926: 5'-TCA AAC TTA CAC CAG TCT TGT AAA CC-3' and H 16498: 5'-CCT GAA GTA GGA ACC AGA TG-3') and 2 units of *Taq* DNA polymerase (Perkin-Elmer) in 50 μ l of 50 mM Tris-HCl (pH 8.8), 15 mM (NH₄)₂SO₄, 1.5 mM MgCl₂, 0.1% Triton X-100, and 0.01% gelatin. An aliquot from the amplicon was subsequently sequenced according to the Sequenase 2.0 (U.S. Biochemical Corp., Cleveland, OH) protocol using primers L15926 (5'-CAC CAT TAG CAC CCA AAG CT-3') and H16401 (5'-TGA TTT CAC GGA GGA TGG TG-3'). Reaction products were electrophoresed at 75 W for 2.5 hr through a 6% polyacrylamide gel containing 7 M urea.

After electrophoresis, the gels were rinsed with ddH₂O, dried under vacuum, and exposed to Kodak XAR film for 24–72 hr. Sequences were read and aligned using the computer program ESEE (Cabot 1993). Alternatively, an automated DNA sequencer (A.L.F., Pharmacia LKB Biotechnology AB) was used. In this case, two amplifications were performed from each individual, where one of the primers carried a biotin at its 5' end. After separation of the strands using avidin coupled to a solid phase (Hultman et al. 1991), sequences were determined using fluorescein-labeled primers.

Sequence Comparisons

The distributions of pairwise differences within and between populations, lineage sharing between populations, and permutation tests for detecting population subdivision (Hudson et al. 1992) were calculated using an unpublished program by Dr. A. von Haeseler (Ludwig Maximilians University, Munich, Germany). Genetic distances (D) between populations were calculated as follows: $D = |D_{\text{BET}} - (D_{\text{W1}} + D_{\text{W2}})/2|$ where D_{BET} = mean pairwise difference between populations 1 and 2, D_{W1} = mean pairwise difference within population 1, and D_{W2} = mean pairwise difference within population 2.

ACKNOWLEDGMENTS

This study was supported by the Finnish Cultural Foundation (A.S) and the Deutsche Forschungsgemeinschaft (S.P.). L.I.-T. is a Howard Hughes Medical Institute Predoctoral Fellow. We would like to thank Karin Bauer, Arndt von Haeseler, Matthias Krings, Karin Pääbo, and Elisabeth Watson for fruitful discussions.

The publication costs of this article were defrayed in part by payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 USC section 1734 solely to indicate this fact.

REFERENCES

- Andersson, S., A.T. Bankier, B.G. Barrell, M.H.L. de Bruijn, A.R. Coulson, J. Drouin, I.C. Eperon, D.P. Nierlich, B.A. Roe, S. Sanger, P.H. Schreier, A.J.H. Smith, R. Staden, and I.G. Young. 1981. Sequence and organisation of the human mitochondrial genome. *Nature* **290**: 457–465.
- Bertranpetit, J., J. Sala, F. Calafell, P. Underhill, P. Moral, and D. Comas. 1995. Human mitochondrial DNA variation and the origin of the Basques. *Ann. Hum. Genet.* **59**: 63–81.
- Brown, W.M., M. George, and A.C. Wilson. 1979. Rapid evolution of animal mitochondrial DNA. *Proc. Natl. Acad. Sci.* **76**: 1967–1971.
- Cabot, E.L. 1993. ESEE, The Eyeball Sequence Editor, version 3.0. Burnaby, B.C., Canada, V5C 2Y2.
- Cavalli-Sforza, L.L., A. Piazza, P. Menozzi, and J. Mountain. 1988. Reconstruction of human evolution: Bringing together genetic, archeological and linguistic data. *Proc. Natl. Acad. Sci.* **85**: 6002–6006.
- Cavalli-Sforza, L.L., E. Minch, and J.L. Mountain. 1992. Co-evolution of genes and languages revisited. *Proc. Nat. Acad. Sci.* **89**: 5620–5624.
- Cavalli-Sforza, L.L., P. Menozzi, and A. Piazza. 1994. *The history and geography of human genes*. Princeton University Press, Princeton, NJ.
- de la Chapelle, A. 1993. Disease gene mapping in isolated human populations: The example of Finland. *J. Med. Genet.* **30**: 857–865.
- Di Rienzo, A. and A.C. Wilson. 1991. Branching pattern in the evolutionary tree for human mitochondrial DNA. *Proc. Natl. Acad. Sci.* **88**: 1597–1601.
- Giles, R.E., H. Blanc, H.M. Cann, and D.C. Wallace. 1980. Maternal inheritance of human mitochondrial DNA. *Proc. Natl. Acad. Sci.* **77**: 6715–6719.
- Guglielmino, C.R., A. Piazza, P. Menozzi, and L.L. Cavalli-Sforza. 1990. Uralic genes in Europe. *Am. J. Phys. Anthropol.* **83**: 57–68.
- Horai, S. and K. Hayasaka. 1990. Intraspecific nucleotide sequence differences in the major noncoding region of human mitochondrial DNA. *Am. J. Human Genet.* **46**: 828–842.
- Horai S, K. Hayasaka, R. Kondo, K. Tsugane, and N. Takahata. 1995. Recent African origin of modern humans revealed by complete sequences of hominoid mitochondrial DNAs. *Proc. Natl. Acad. Sci.* **92**: 532–536.
- Hudson, R.R., D.D. Boos, and N.L. Kaplan. 1992. A statistical test for detecting geographic subdivision. *Mol. Biol. Evol.* **9**: 138–151.
- Hultman, T., S. Berg, T. Moks, and M. Uhlen. 1991. Bidirectional solid-phase sequencing of *in vitro*-amplified plasmid DNA. *BioTechniques* **10**: 30–34.
- Lundström, R., S. Tavaré, and R.H. Ward. 1992. Estimating substitution rates from molecular data using the coalescent. *Proc. Natl. Acad. Sci.* **89**: 5961–5965.
- Marjoram, P. and P. Donnelly. 1994. Pairwise comparisons of mitochondrial DNA sequences in subdivided populations and implications for early human evolution. *Genetics* **136**: 673–683.
- Nevanlinna, H.R. 1972. The Finnish population structure: A genetic and genealogical study. *Hereditas* **71**: 195–236.
- Piercy, R., K.M. Sullivan, N. Benson, and P. Gill. 1993. The application of mitochondrial DNA typing to the

SAJANTILA ET AL.

study of white Caucasian genetic identification. *Int. J. Leg. Med.* **106**: 85–90.

Pult, I., A. Sajantila, J. Simanainen, O. Georgiev, W. Schaffner, and S. Pääbo. 1994. Mitochondrial DNA sequences from Switzerland reveal striking homogeneity of European populations. *Biol. Chem. Hoppe-Seyler* **375**: 837–840.

Ruhlen, M. 1987. *A guide to the world languages. Vol. 1, classification*. Edward Arnold, London, UK.

Sinor, D. 1988. In *Uralic languages: Description, history and foreign Influences* (ed. Sinor D), xiii–xx. E.J. Brill, Leiden, The Netherlands.

Slatkin, M. and D. Hudson. 1991. Pairwise comparisons of mitochondrial DNA sequences in stable and exponentially growing populations. *Genetics* **129**: 555–562.

Tollersrud, O.K., O. Nilssen, L. Tranebjaerg, and O. Borud. 1994. Aspartylglucosaminuria in northern Norway: A molecular and genealogical study. *J. Med. Genet.* **31**: 360–363.

Vigilant, L., M. Stoneking, D. Harpending, K. Hawkes, and A.C. Wilson. 1991. African populations and the evolution of human mitochondrial DNA. *Science* **253**: 1503–1507.

Ward, R.H., B.S. Frazer, K. Dew, and S. Pääbo. 1991. A single North-American tribal group contains extensive mitochondrial diversity. *Proc. Natl. Acad. Sci.* **88**: 8720–8724.

Ward, R.H., A. Redd, D. Valencia, B. Frazier, and S. Pääbo. 1993. Genetic and linguistic differentiation in the Americas. *Proc. Natl. Acad. Sci.* **90**: 10663–10667.

Received May 12, 1995; accepted in revised form June 19, 1995.