J Mol Evol (1992) 34:336–344

Journal of Molecular Evolution

The Evolution of Two West African Populations

O. Colin Stine, G.J. Dover, D. Zhu, and Kirby D. Smith

Department of Medicine, The Johns Hopkins University, 933 Traylor Building, 720 Rutland Avenue, Baltimore, MD 21205, USA

Summary. The identification of genetically coherent populations is essential for understanding human evolution. Among the culturally uniform ethnic groups of west Africa, there are two geographically distinct populations with high frequencies of sickle-cell hemoglobin (HbS). Although the HbS mutation in each group is found on distinguishable chromosomes 11, these populations have been assumed to be parts of a single population. Analysis of mitochondrial DNA (mtDNA) in these populations demonstrated that the two populations identified by alternative chromosomes 11 bearing HbS have distinct distributions of mitochondrial genotypes, i.e., they are maternally separate. These studies also showed that, contrary to expectation, the mtDNA of some individuals is heteroplasmic. For nuclear loci, a comparison of the frequency of alternative alleles established that these populations are genetically distinct. Both the mitochondrial and nuclear data indicate that these populations have been separate for approximately 50,000 years. Although HbS in the two populations is usually attributed to recent, independent mutations, the duration of the separation and the observed geographic distribution of the population allow for the possibility of an ancient origin of HbS. Assuming an ancient mutation and considering the known biogeography, we suggest that HbS protected selected populations from malaria in rain forest refuges during the most recent ice age.

Key words: Humans — Mitochondrial DNA — Nuclear polymorphisms — Heteroplasmy — Genetic differentiation — Sickle cell — Rain forest refuges

Offprint requests to: O.C. Stine

Introduction

Understanding the processes that effect the evolution of populations requires a careful analysis of their genetic structure. This is particularly true in humans where populations have expanded, subdivided, and been subjected to genetic changes caused by migration, natural selection, mutation, and drift. Traditionally, language groups have been used as markers for population subdivisions, and it is generally assumed that genetic differentiation parallels the clustering of the major linguistic families (Excoffier et al. 1987). However, language groupings do not necessarily reflect ancient genetic subdivisions. This study examines the relationship between the genetic and cultural distinctness of ethnic groups with high frequencies of sickle-cell hemoglobin (HbS) in western Africa.

Among four geographically distinct populations (Senegal, Benin, Bantu, and Arab–Indian), which carry the HbS mutation on different chromosomes 11 (Mears et al. 1981; Pagnier et al. 1984), there are two, the Senegal and Benin, that are culturally homogenous with similar agricultural practices and languages from the same linguistic subfamily (Excoffier et al. 1987). The alternative chromosomes 11 are distinguished by the presence and absence of a specific set of restriction sites in a 60-kb region containing the β-globin locus—referred to as the haplotype (Antonarakis et al. 1984). One of the chromosome 11 haplotypes bearing HbS occurs at a frequency of approximately 86% among chromosomes bearing HbS in a large region that includes Senegal, whereas an alternative haplotype associated with HbS occurs at a frequency of 100% in a geographically separate region that includes Benin (Schroder et al. 1990). Despite this clear genetic differentiation at the β-globin locus, these populations

are generally assumed to be a single genetic group. The genetic differentiation at the β-globin locus is explained by assuming that two recent, independent sickle-cell mutations occurred on distinct chromosomes 11, with insufficient time (ca. 3000 years) for gene flow to have homogenized the HbS haplotype frequencies within the population (Kurnit 1979; Mears et al. 1981; Pagnier et al. 1984). Based on this reasoning, the anthropological similarity approximates the evolutionary history of this population.

However, it remains possible that these are two genetically and evolutionarily distinct populations that diverged long ago and only recently learned a common culture (Livingstone 1958; Kurnit 1979; Solomon and Bodmer 1979). In this scenario, the sickle mutation(s) could be very old (Kurnit 1979; Solomon and Bodmer 1979). Solomon and Bodmer (1979), using the same haplotype data as above and assuming the existence of a particular restriction site (Hpa II) polymorphism in the founding population, calculated that HbS may be 70,000–150,000 years old. Unfortunately, there is little information about whether HbS is old or new and whether the populations are distinct. The only clues are (1) a population (n = 46) in the Senegal region where a restriction site polymorphism (Hpa II) and HbS were in Hardy–Weinberg equilibrium, consistent with an ancient origin and (2) a population from Benin (n = 26) where they were not in equilibrium, suggesting a more recent origin of either HbS or one of the restriction site morphs (Mears et al. 1981). Clearly, these observed distributions do not exclude either the hypothesis that there is a single large population with a common evolutionary history, or that there are two distinct and anciently separated populations.

These two hypotheses predict different distributions of mitochondrial DNA (mtDNA) and polymorphic nuclear genes in these two west African populations characterized by alternative chromosomal haplotypes associated with HbS. Either the two populations will appear to be replicate samples from the same population or they will be distinct. The analysis of mtDNA is theoretically the most sensitive measure of the genetic subdivision of populations. The sensitivity derives from the empirical observations that mtDNA is variable among humans, has a high rate of mutation, and is maternally inherited (Moritz et al. 1987). It should be noted, however, that because of maternal inheritance and the possibility of differential migration patterns between sexes, there may be cases where a population is subdivided for its mtDNA but not its nuclear genes (Birky et al. 1989). By analyzing both mtDNA and polymorphic nuclear genes, we established that the two west African populations associated with alternative haplotypes of HbS have remained genetically distinct since the most recent ice age.

Methods

Mitochondrial DNA Genotypes. The genotype of mtDNA in each individual was identified by restriction analysis. Whole cell DNA was prepared from blood samples (Maniatis et al. 1982), and aliquots were digested with each of six restriction enzymes (Ava II, Hpa I, Hae II, Hae III, Hinf I, Msp I) used individually. Restriction fragments were separated by electrophoresis through 1% or 2% agarose gels (Maniatis et al. 1982), transferred to Gene Screen Plus in alkali (Reed and Mann 1985), and mtDNA fragments were detected by Southern blot analysis (Maniatis et al. 1982). Prehybridization and hybridization were carried out in 1 M NaCl, 50 mM Tris-HCl, pH 8.0, 10% dextran sulfate, and 1% SDS. Radioactive probes were prepared by primer extension (Feinberg and Vogelstein 1983) of seven cloned mtDNA fragments that cover 82% of the mitochondrial genome, from nucleotide positions 16,345 to 13,364 (Anderson et al. 1981). Southern blots were washed with 2× Standard saline citrate (SSC: 0.15 M NaCl, 0.015 M sodium citrate), 1% SDS at 45°C for 5 min and then with 0.1× SSC at 45°C for 15 min and exposed to x-ray film for 1–5 days. Most of the observed restriction fragments were those expected from the published sequence of human mtDNA (Anderson et al. 1981). The unexpected mtDNA fragments could be explained by gains or losses of single restriction sites, and in most cases the specific sites could be identified by comparison with previous restriction analyses (Johnson et al. 1983; Cann et al. 1984). For each individual, the presence and absence of the restriction sites were compiled and referred to as the mitochondrial genotype.

Evolutionary Comparisons. The relationships among mitochondrial genotypes are represented by networks (Hudson and Kaplan 1985), as networks minimize the required number of changes and include all the observed single site changes between mitochondrial genotypes. Networks, unlike the standard evolutionary tree, do not require knowledge of an ancestral genotype or make any assumption about the timing or nature of mutational events. If recombination occurs (a possibility that has never been ruled out for mtDNA, see below), then a network must be used to analyze the data (Hudson and Kaplan 1985). Even if recombination does not occur, a network is the most conservative (unlikely to find differences) method of measuring the distance between two populations because it minimizes the number of steps between all possible pairs of genotypes. Thus, if differences between populations are found, they are likely to be legitimate.

Polymerase Chain Reaction (PCR). The oligonucleotide primers used to initiate PCR amplification were chosen because they flank specific restriction sites. The PCR cycle, consisting of 1 min each at 92°C, 50°C, and 72°C, was repeated 30 times. Half the amplified sample was digested with an appropriate restriction enzyme, and each half of the sample was analyzed on 2% agarose gels (Maniatis et al. 1982). Two experimental protocols were routinely used as controls for the PCR. The first, reactions without template DNA, were performed to rule out possible contamination of the primers. The second was that gels were routinely Southern blotted and hybridized with an appropriate radioactive probe to check for quantities of DNA that were not detectable by ethidium staining.

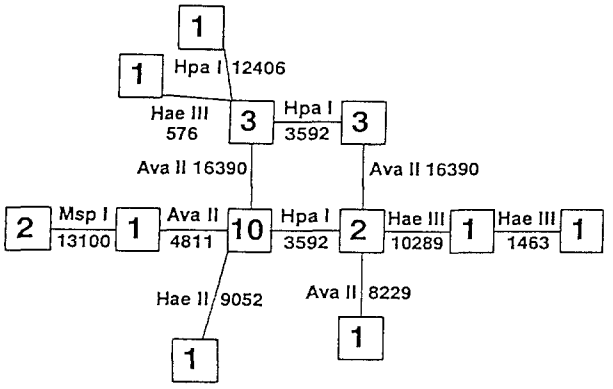

Fig. 1. The network of relationships between the observed mtDNA genotypes among Jamaicans. Each square represents a distinct mitochondrial genotype as defined by the presence or absence of 75 observed restriction sites. The number inside the square is the number of individuals in the sample with that genotype. Each line connecting two squares represents a change in one restriction site. Each line is labeled with the restriction enzyme and the number of the first nucleotide of the restriction site in the published mtDNA sequence (Anderson et al. 1981).

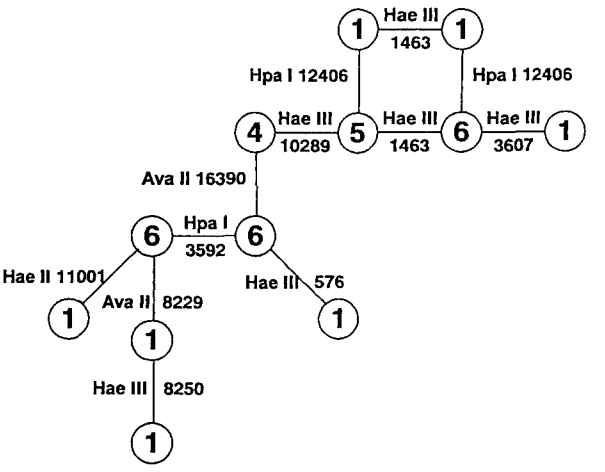

Fig. 2. The network of relationships between the observed mtDNA genotypes among Senegalese. The conventions are as described for Fig. 1.

Results and Discussion

Mitochondrial DNA

We compared a population characterized by the Benin haplotype associated with HbS to a population characterized by the Senegal haplotype. The population with the Benin haplotype came from Jamaica, whereas the one with the Senegal haplotype came from Senegal. The Jamaicans are considered to be representative of the Benin region for three reasons. First, 78% of their HbS-bearing chromosomes have the Benin haplotype. Second, historical records of the slave trade indicate that most of the estimated 750,000 individuals imported from Africa to Jamaica came from the region near the Bight of Benin (Curtin 1969). Third, although this population has unquestionably been subjected to admixture, the African women have contributed their genes including mtDNA disproportionately to the current inhabitants of Jamaica because some of them were afforded preferential treatment (Curtin 1955). In addition, the admixture in the Jamaican population is likely to obscure differences between the Senegal and Benin population because, based on records from the slave trade, half of the Africans imported to Jamaica who were not from the Benin region came from the Senegal region (Curtin 1969). Hence, if these two populations are statistically different the observation is likely to be real.

The first population consisted of 27 unrelated Jamaicans (Boyer et al. 1984). Based on the presence or absence of 75 restriction sites, 12 distinct mitochondrial genotypes were identified that differed from each other by a series of changes in single restriction sites. Combining these 12 mitochondrial genotypes in a network revealed that 4 of the 5 genotypes, found more than once, are clustered together in the center with the uniquely found genotypes surrounding them (Fig. 1).

The second population consisted of 34 unrelated Senegalese; 78% of their HbS chromosomes had the Senegal haplotype. Restriction analysis of the same 75 sites used to define the Benin mitochondrial genotypes, revealed 12 mitochondrial genotypes differing from each other by a series of changes in single restriction sites. When these data were organized into a network, the five most common mitochondrial genotypes were adjacent to other commonly found genotypes and centrally located (Fig. 2). Many of the variable sites are the same as those seen in another sample of Senegalis (Scozzari et al. 1988).

As shown in Fig. 2, this network relationship indicates that the Hpa I site at 12,406 [the nucleotide number in the published sequence (Anderson et al. 1981)] and the Hae III site at 1463 are found in all possible combinations of presence and absence. Similarly, in the Benin population the Ava II site at 16,390 and the Hpa I site at 3592 are found in all combinations. [Others (Johnson et al. 1983; Brega et al. 1986; Santachiara-Benerecetti et al. 1988) have reported this phenomenon among Orientals and Caucasians.] The standard interpretation suggests that two independent mutations affected the same site in two mtDNA lineages. However, if the mutation rate is 10^{-5}–10^{-7} for mtDNA (Moritz et al. 1987), then two such independent mutations in the same site are very unlikely in a small population. Alternatively, this combination of restriction sites may be the result of recombination between two distinguishable mtDNAs in a heteroplasmic individual.

It is generally accepted that humans are not het-

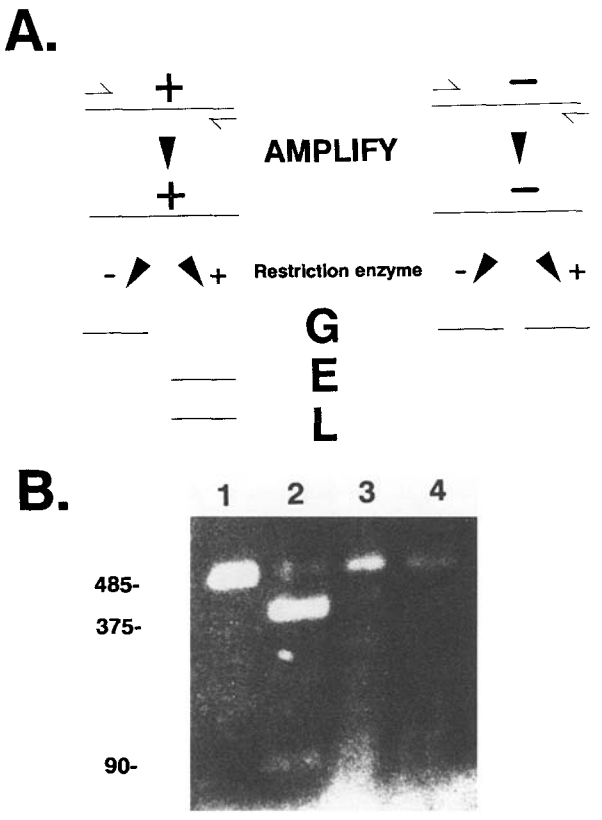

Fig. 3. A test for heteroplasmy. After digesting the DNA with the appropriate restriction enzyme, PCR amplification was performed between primer pairs that flank the Hae III site at 1463 and the Hpa I site at 12,406. Half the amplified DNA was digested with the appropriate restriction enzyme, and the divided samples were run in separate lanes. **A** A diagram of the expected result on the left side if the amplified fragment results from mtDNA that has the site but was not cleaved during the first digestion: the second digestion will cleave the amplified fragment, and two fragments will appear on the gel. The right side of the diagram shows what will happen if the amplified fragment results from mtDNA that does not have the restriction site: the second digestion will not cleave the fragment and both lanes on the gel will have the same size fragment. **B** In lanes 1 and 2, the DNA was not digested before amplification. In lane 2, the amplified fragment is cleaved by subsequent digestion with Hae III. In lanes 3 and 4, the DNA of the same individual was digested to completion with Hae III and then amplified. In lane 4, the amplified fragment was not cleaved by subsequent digestion, thus demonstrating the presence of heteroplasmy.

eroplasmic (Moritz et al. 1987). The exceptions are associated with diseases: Kearns–Sayre and Pearsons syndromes (Holt et al. 1988; Rotig et al. 1989), Leber optic atrophy (Holt et al. 1989), and an unusual mitochondrial myopathy (Holt et al. 1990). However, heteroplasmy has been observed with length variation in many species (Bermingham et al. 1986; Rand and Harrison 1986; Stine 1989) and with restriction site differences in cattle where such variation is postulated to be common (Olivo et al. 1983).

Because Southern blot analysis of the samples examined in this study indicates the presence of a single mitochondrial genotype in each individual, any alternative mitochondrial genotype must be rare. Thus, the demonstration of heteroplasmy necessary to associate the genotypic variation in Fig. 2 with a recombination event rather than independent mutations at the same site requires a sensitive technique, like the PCR, for detecting the rare form.

Using PCR analysis, heteroplasmy for a restriction site can be demonstrated most easily if the mtDNA that does not have the site is the rare form. As diagrammed in Fig. 3A, the DNA sample is first digested to completion with the appropriate restriction enzyme. A PCR initiated with primers flanking the restriction site will only amplify DNA that is not cleaved by the enzyme. Partial digestion can be excluded by digesting the amplified DNA. If the amplified DNA results from DNA with an uncleaved site, the amplified product will be cleaved. However, if the amplified DNA is the result of DNA without the site, it will not be cleaved.

Even if heteroplasmy is ubiquitous, heteroplasmy for any given restriction site is unexpected. Thus, candidate sites must be identified from the population analysis. Sites that appeared in all combinations of presence and absence, the Hpa I site at 12,406 and the Hae III site at 1463, were assayed. For the Hpa I site, two of eight individuals were heteroplasmic (Fig. 3B). Heteroplasmy at the Hae III site was not found in the four individuals tested. The presence of heteroplasmy in two individuals in this small sample makes it unlikely that the observed heteroplasmy is the result of new mutations. However, it is possible that this is the result of paternal transmission of mtDNA, a recently demonstrated possibility (Gyllensten et al. 1991). Regardless of the mechanisms by which heteroplasmy occurs and although these samples are too small to estimate how many rare types of mtDNA are present in a single individual or how often heteroplasmy occurs, the results are consistent with heteroplasmy being common and widespread.

Although the demonstration of heteroplasmy does not prove the occurrence of recombination in mtDNA, its presence makes recombination a potential source for some of the observed variation in the population networks. Thus, recombination may account for the various combinations of the Hpa I site at 12,406 and the Hae III site at 1463 seen in Fig. 3. Hence, phylogenetic analysis of the mtDNA should be based on networks (Hudson and Kaplan 1985).

The mitochondrial genotypes of the Senegal and Benin HbS-bearing populations were compared (Fig. 4). Three mitochondrial genotypes, found frequently in both populations, are shared; each of the other mitochondrial genotypes is unique to one or the

other of the populations. Two statistical methods were used to test whether the populations represent two samples drawn from a single homogenous population or not. First, a chi-square test was performed. The unique mitochondrial genotypes were clustered with the nearest frequently found genotype in order to avoid small cell sizes. When a unique genotype could be paired with two frequently found genotypes, the unique genotype was clustered in such a way as to minimize the χ^2 value, in an attempt to make this as conservative (unlikely to find differences) an approach as possible. The χ^2 for this comparison is 16.03 (df = 5, $P < 0.01$), indicating that the populations are significantly different, in agreement with the HbS haplotype data on these same samples. This result is clearly statistically significant even though the χ^2 calculation used is the least sensitive for distinguishing two populations (Roff and Bentzen 1989). However, because several mitochondrial genotypes are shared between the two populations, and after clustering there are no mtDNA genotypes unique to the Senegal population, we subjected the data to an additional statistical test.

The second statistic test is a general method for partitioning variance associated with mtDNA. It assumes that the mtDNA diversity is neutral and occurs primarily between individuals, as is the case in our study. The method relates the variation within and between populations to a parameter, *Nm,* defined as the product of the effective population size (*N*) and the migration rate (*m*). This procedure has the advantage that an *Nm* calculated from mitochondrial data, in appropriate cases, can be used to make predictions about the distribution of nuclear genes within and between populations (Birky et al. 1983, 1989).

The interpretation of an estimate of *Nm* has been established both theoretically and empirically (Birky et al. 1983; Rand and Harrison 1989). An *Nm* of less than 1 indicates that the populations are subdivided for both the mitochondrial and nuclear genomes. If *Nm* is greater than 1 but less than 4 [because of the maternal inheritance of mtDNA, the upper limit may be somewhat higher if males migrate more often than females (Birky et al. 1989)], the populations must be divided for their mitochondrial genomes but no conclusion regarding population subdivision can be drawn about their nuclear genes. And if *Nm* is greater than 4, there is a single population without subdivision of either mitochondrial or nuclear genes (Birky et al. 1983). These ranges are dependent on uniparental inheritance, a lack of heteroplasmy, and uniform male and female migration rates. Thus, the numerical limits (1 and 4) of genetic distinctness must be interpreted cautiously.

Because of the uncertainty of the assumptions,

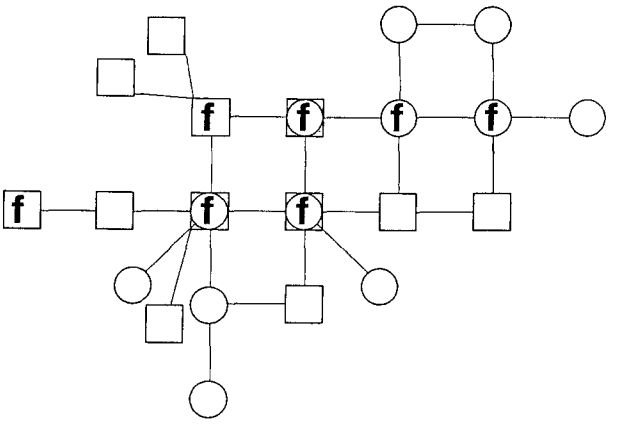

Fig. 4. The network of the relationships between the observed mtDNA genotypes in the two populations characterized by alternative chromosome 11 haplotypes associated with HbS. The conventions are as in Figs. 2 and 3 with the squares representing those mitochondrial genotypes found in the population characterized by the Benin chromosome 11 haplotype and the circles representing the mitochondrial genotypes in the population with the Senegal chromosome 11 haplotype. Mitochondrial genotypes found frequently in one or both populations are marked with an F. Shared mitochondrial genotypes were frequent in both populations.

we have estimated *Nm* in three ways. The *Nm* calculated using the mtDNA data from the west African populations is 2.6 (*K* statistics: Birky et al. 1989), 5.9 (*K* and G_{st}: Rand and Harrison 1989), and between 0.3 and 2.6 with 95% confidence (phylogenetic tree analysis: Slatkin and Maddison 1989). Although these estimates vary, *Nm* is always low [*Nm* values as high as 49 have been reported for geographically separate groups of a single population (Brown and Chapman 1991)], indicating that there are two distinct populations. However, all of these estimates are close to the theoretical range in which populations are subdivided for their mitochondrial genomes, but not necessarily for their nuclear genes (Birky et al. 1983), and a firm conclusion about their overall genetic distinctness cannot be drawn.

Both methods (*Nm* and χ^2) indicate that there are two populations in terms of their maternal (mtDNA) lineages, in agreement with the HbS haplotype data on the same samples. However, neither method is conclusive enough to demonstrate an ancient separation between the west African populations associated with alternative haplotypes. Thus, a firm conclusion regarding their genetic distinctness rests on an analysis of nuclear genes.

Nuclear Genes

Recently, an independent method for comparing the nuclear genes between agglomerated populations has been developed. It was shown that even though a

Table 1. Summary of the statistics for nuclear alleles

Locus	Number of alleles	χ^2	P^a
ABO	235,551	47.28	***
A_1A_2BO	7811	37.00	***
MN	13,750	150.0	***
MNSs	2451	73.79	***
Rh	11,883	139.99	***
Kell	5964	18.80	***
P system	2944	28.56	***
GC	1766	8.49	***
Kidd	235	25.15	***
Lutheran	572	0.13	n.s.
Lewis RBC	850	0.29	n.s.
Lewis saliva	850	0.0006	n.s.
6-Phosphogluconate	1926	4.0	*
Ag-β-lipoprotein	1600	5.16	*
PGM-1	2446	5.43	*
Glucose-6-phosphate dehydrogenase	6573	89.42	***
Transferrin	6540	79.46	***
Hp O	4422	103.6	***
Hp	6336	7.9	***

[a] n.s., not significant; * $P < 0.05$; *** $P < 0.005$

broad grouping may combine heterogenous samples, the comparisons between two agglomerated populations are legitimate if they are based on polymorphisms (Chakraborty et al. 1988; Neel et al. 1988). In order to determine if the west African populations are distinct for their nuclear genes as well as their mtDNA, χ^2 was calculated for polymorphic nuclear genes using data taken from the literature (Mourant et al. 1976; Tills et al. 1983). The identification of test populations was based on the geographic distribution of HbS documented by Livingstone (1958). The population from the Senegal region included ethnic groups from Senegal, Gambia, Guinea, and western Liberia, whereas the population from the Benin region included ethnic groups from Nigeria, Benin, and Ghana. These two groups are separated by intervening ethnic groups that do not have an appreciable frequency of the sickle-cell gene (Livingstone 1958). Recently, Schroder et al. (1990) have compiled the existing haplotype data for HbS-bearing chromosomes in Africa and concluded that these two agglomerated populations have the appropriate alternative haplotypes. Thus, our mtDNA samples and these agglomerated populations can be legitimately considered two independent samples of the same populations.

Published tables (Mourant et al. 1976; Tills et al. 1983) list data on blood group and enzyme polymorphisms for a total of 316,236 alleles from 17 loci for the above ethnic groups. As shown in Table 1, the two agglomerated populations were significantly different at 14 of 17 polymorphic loci (χ^2, 3 loci at $P < 0.05$ and 11 loci at $P < 0.005$). Thus, based on their nuclear genes, these two populations are clearly genetically distinct. The genetic separation of these populations associated with alternative HbS haplotypes has now been established by independent statistical methods on two independent data sets (χ^2 and Nm for mtDNA and χ^2 for nuclear genes). The question remains as to when the populations diverged.

Time Scale

A time of divergence can be calculated by several independent methods. If the divergence time estimated from the various methods applied to independent data sets are similar, confidence in the historical separateness of the two west African populations will be strengthened.

Two methods can be used to estimate the genetic distance from mtDNA data. The first measures the genetic distance within and between populations from the average number of restriction site differences (Nei and Tajima 1981). The genetic distances within and between the west African populations were measured using the fewest changes in restriction sites between mitochondrial genotypes in the network. Within the population bearing the Senegal HbS haplotype, the average number of restriction site differences is 1.089; within the population bearing the Benin HbS haplotype, 0.831; and between populations, 1.349. These numbers are similar to those Johnson et al. (1983) observed within and between races that are thought to have diverged 100,000 years ago. However, the present study observed nearly twice the number of nucleotides (400 vs 214). Thus, simplistically, because the same number of changes was observed in twice the number of nucleotides, the time of divergence of the west African populations should be ½ of that for the races, or approximately 50,000 years.

The second method of estimating genetic distance from mtDNA is to use the average number of mutations in population specific clusters to estimate the percent divergence, p (Stoneking et al. 1986). In this study, there are nine clusters of genotypes that are population specific. The average number of mutations necessary to derive each cluster from the nearest shared genotype is 2. Because 400 nucleotides were observed, p equals 0.5%. This value is 1.6 times that seen between Southeast Asians and New Guineans who diverged 30,000–50,000 years ago. Thus by this calculation, the two different HbS-bearing populations separated an estimated 50,000–80,000 years ago.

The estimated time of divergence for nuclear genes can be estimated from Nei's (1972) genetic distance. The genetic distance between the Senegal and Benin

populations was calculated separately for blood group alleles and allozymes (Nei and Roychoudhury 1974). The distance between the Senegal and Benin populations was 0.01513 for blood group alleles and 0.010195 for enzymes. These estimates of genetic distance are approximately one-third of that for blood groups and half of that for proteins previously calculated for the different races (Nei and Roychoudhury 1984). Because the races are thought to have diverged 100,000 years ago, the west African populations are estimated on the basis of their polymorphic nuclear genes to have been separate for 30,000–50,000 years.

Thus, three independent methods of analysis based on either mtDNA or nuclear genes lead to the same conclusion: these populations diverged approximately 50,000 years ago. That is about the time when the most recent ice age began.

Implications

The presence of two genetically distinct, anciently separated populations in western Africa has three implications that are consistent with other independent lines of evidence. The first involves the age and number of HbS mutations. Although the ancient separation between the Benin and Senegal populations does not distinguish between the competing hypotheses of multiple mutations or a single mutation, it does mean that multiple recent HbS mutations are not necessary to explain the observed distribution of alternative haplotypes. An ancient origin of a single or multiple mutations, before the populations separated, would provide sufficient time for the observed differentiation of HbS haplotypes (Solomon and Bodmer 1979). Moreover, an ancient origin is consistent with the observation of numerous distinct chromosomal haplotypes (called atypicals) that are closely related to the common ones (Antonarakis et al. 1984; Srinivas et al. 1988). Chromosomes carrying HbS would be subject to recombination and mutation that would generate such atypicals and allow enough time for them to reach measurable frequencies (see Solomon and Bodmer 1979). These observations are difficult to reconcile with recent mutations.

Second, if HbS is ancient, there must be an environment in which the mutation could have been maintained as a stable polymorphism. One candidate region is the rain forest. *Anopheles funestus,* a mosquito carrier of malaria dwells in the rain forest and malaria would provide the necessary selective pressure to maintain the HbS polymorphism (Wilson 1949). Evidence supporting this notion can be derived from the occurrence of HbS in pygmies. Pygmies are thought to have lived continuously in the rain forests and are subdivided into two genetically distinct groups, the eastern and western pygmies. Both groups have HbS (E. 8%, Cavalli-Sforza et al. 1969; W. 18% Motulsky et al. 1966), and their geographic distribution correlates with rain forest refuges in central Africa. Genetic analysis suggests that the two groups of pygmies diverged an estimated 50,000–115,000 years ago (nuclear genes, Cavalli-Sforza et al. 1969; mtDNA, Vigilant et al. 1989). Significantly, the geographic distribution (Fig. 5) of the west African populations associated with HbS also corresponds to the geographic distribution of particular species and subspecies of birds and butterflies that currently live in the rain forests (Mayr and O'Hara 1986). The latter distribution is thought to reflect the recent expansion of the rain forest from ice age refuges 10,000–50,000 years ago (Fig. 5). The correlation between the human and animal populations of the rain forest includes the presence of the gap between the Senegal and Benin HbS-bearing populations. HbS is virtually absent in ethnic groups from eastern Liberia and Ivory Coast (Livingstone 1958). Additionally, in central Africa, the Bantu who have a distinct HbS haplotype (Pagnier et al. 1984), are geographically correlated with a third rain forest refuge. They are anthropologically (Excoffier et al. 1987) and genetically (mtDNA, Scozzari et al. 1988) distinct from the other populations with HbS. Thus, the presence of HbS in the rain forest environment may have played a significant role in human evolution.

Third, the separation of the west African populations associated with HbS implies a low amount of gene flow between the two populations. This is consistent with the previously described gap in the distribution of HbS. If migration is restricted, the HbS mutation is not expected to have expanded into every population whose current habitat is conducive to malaria, e.g., those of the Ivory Coast and eastern Liberia. Additionally, anthropological data suggest that large-scale migrations have not occurred in west Africa over the past 9000 years (Livingstone 1958). The genetic data presented here suggest that significant migration between the two west African populations associated with HbS may not have occurred for the last 50,000 years. Similarly, eastern and western pygmies, separate for at least 50,000 years, have been inferred to have a low rate of gene flow even though they are linguistically and anthropologically similar (Cavalli-Sforza et al. 1969; Vigilant et al. 1989). Thus, the cultural similarity between the Senegalese and the Beninese appears to be the result of recent learning and not a reflection of their genetic history (Livingstone 1958).

Haldane (1949) has proposed that disease, including parasitic infections, is a major factor in evolution. It is generally accepted that the recent expansion of African populations with HbS is

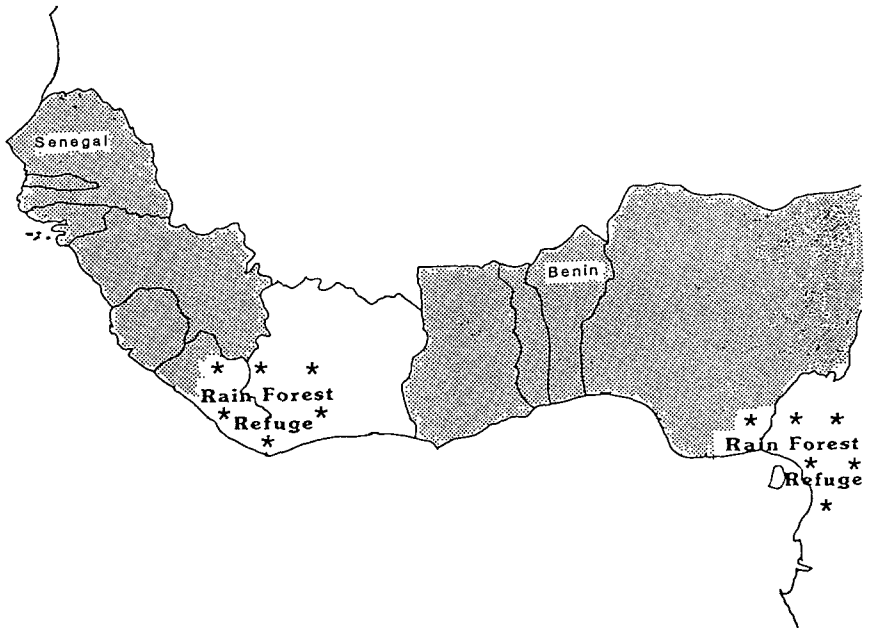

Fig. 5. A map of Africa showing Senegal, Benin, and the surrounding regions. The geographic distribution of tribes with a high frequency of HbS is shaded (after Livingstone 1958). The locations of the rain forest refuges during the ice ages as postulated by Mayr and O'Hara (1986) are marked.

associated with the introduction of the Malaysian agricultural system (Wiesenfeld 1967). Based on our data, HbS may have played a role in human evolution for a much longer time. It is plausible that ancient hemoglobin mutations protected human populations in malaria-infested rain forests during the most recent ice age. The selection of populations isolated in separate rain forest refuges would have led to the formation of genetically distinct groups of west Africans that could have subsequently expanded and recently learned a common culture. Furthermore, it is plausible that this evolutionary scenario may be extended to include rain forest refuges in central Africa, India, and southeast Asia.

Acknowledgments. We acknowledge helpful suggestions and comments from Ronald L. Nagel, Bob Chapman, Carolyn Decker, Terry Bishop, Trefor Jenkins, and Samuel Boyer. We are particularly indebted to Dr. Nagel for some of the samples.

References

Anderson S, Bankier AT, Barrell BG, de Bruijn MHL, Coulson AR, Drouin J, Eperon IC, Nierlich DP, Roe BA, Sanger F, Schreier PH, Smith AJH, Staden R, Young IG (1981) Sequence and organization of the human mitochondrial genome. Nature 290:457–465

Antonarakis SE, Boehm CD, Serjeant GR, Theisen CE, Dover GJ, Kazazian HH (1984) Origin of the β^s-globin gene in blacks: the contribution of recurrent mutation or gene conversion or both. Proc Natl Acad Sci USA 81:853–856

Bermingham E, Lamb T, Avise JC (1986) Size polymorphism and heteroplasmy in the mitochondrial DNA of lower vertebrates. J Hered 77:249–252

Birky CW Jr, Maruyama T, Fuerst P (1983) An approach to population and evolutionary genetic theory for genes in mitochondria and chloroplasts, and some results. Genetics 103: 513–527

Birky CW Jr, Fuerst P, Maruyama T (1989) Organelle gene diversity under migration, mutation, and drift: equilibrium expectations, approach to equilibrium, effects of heteroplasmic cells, and comparison to nuclear genes. Genetics 121: 613–619

Boyer SH, Dover GJ, Sergeant GR, Smith KD, Antonarakis SE, Embury SH, Margolet L, Noyes AN, Boyer ML, Bias WB (1984) Production of F cells in sickle cell anemia: regulation by a genetic locus or loci separate from the β-globin gene cluster. Blood 64:1053–1058

Brega A, Scozzari R, Maccioni L, Iodice C, Wallace DC, Bianco I, Cao A, Santachiara-Benerecetti AS (1986) Mitochondrial DNA polymorphisms in Italy. I. Population data from Sardinia and Rome. Ann Hum Genet 50:327–338

Brown BL, Chapman RW (1991) Gene flow and mitochondrial DNA variation in the killifish, *Fundulus heteroclitus*. Evolution 45:1147–1161

Cann RL, Brown WM, Wilson AC (1984) Polymorphic sites and the mechanism of evolution in human mitochondrial DNA. Genetics 106:479–499

Cavalli-Sforza LL, Zonta LA, Nuzzo F, Bernini L, De Jong WWW, Meera Khan P, Went LN, Siniscalco M, Nijenhuis LE, van Loghem E, Modiano G (1969) Studies on African pygmies. I. A pilot investigation of Babinga pygmies in the Central African Republic (with an analysis of genetic distances). Am J Hum Genet 21:252–274

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

Curtin PD (1955) Two Jamaicas: the role of ideas in a tropical colony 1830–1865. Harvard University Press, Cambridge MA

Curtin PD (1969) The Atlantic slave trade: a census. University of Wisconsin Press, Madison

Excoffier L, Pellegrini B, Sanchez-Mazas A, Simon C, Langaney A (1987) Genetics and history of sub-Saharan Africa. Yearb Phys Anthropol 30:151–194

Feinberg A, Vogelstein B (1983) A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. Anal Biochem 132:6–13

Gyllensten U, Wharton D, Josefsson A, Wilson AC (1991) Paternal inheritance of mitochondrial DNA in mice. Nature 352: 255–257

Haldane JBS (1949) Disease and evolution. Ric Sci [Suppl] 19: 68–75

Holt IJ, Harding AE, Morgan-Hughes JA (1988) Deletion of muscle mitochondrial DNA in patients with mitochondrial myopathies. Nature 331:717–719

Holt IJ, Miller DH, Harding AE (1989) Genetic heterogeneity and mitochondrial DNA heteroplasmy in Leber's hereditary optic neuropathy. J Med Genet 26:739–743

Holt IJ, Harding AE, Petty RKH, Morgan-Hughes JA (1990) A new mitochondrial disease associated with mitochondrial DNA heteroplasmy. Am J Hum Genet 46:428–433

Hudson RR, Kaplan NL (1985) Statistical properties of the number of recombination events in the history of a sample of DNA sequences. Genetics 111:147–164

Johnson MJ, Wallace DC, Ferris SD, Rattazzi MC, Cavalli-Sforza LL (1983) Radiation of human mitochondrial DNA types analyzed by restriction endonuclease cleavage patterns. J Mol Evol 19:255–271

Kurnit DM (1979) Evolution of sickle variant gene. Lancet i: 104

Livingstone FB (1958) Anthropological implications of sickle cell gene distribution in west Africa. Am Anthropol 60:533–562

Maniatis T, Fritsch EF, Sambrook J (1982) Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor NY

Mayr E, O'Hara RJ (1986) The biogeographic evidence supporting the pleistocene forest refuge hypothesis. Evolution 40: 55–67

Mears JG, Lachman HM, Cabannes R, Amegnizin KPE, Labie D, Nagel RL (1981) Sickle gene: its origin and diffusion from west Africa. J Clin Invest 68:606–610

Moritz C, Dowling TE, Brown WM (1987) Evolution of animal mitochondrial DNA: relevance for population biology and systematics. Annu Rev Ecol Syst 18:269–292

Motulsky AG, Vandepitte J, Fraser GR (1966) Population genetic studies in the Congo. I. G6PD deficiency, hemoglobin S and malaria. Am J Hum Genet 18:514–537

Mourant AE, Kopec AC, Domaniewska-Sobczak K (1976) The distribution of human blood groups and other polymorphisms. Oxford University Press, London

Neel JV, Satoh C, Smouse P, Asakawa J, Takahashi N, Goriki K, Fujita M, Kageoka T, Hazama R (1988) Protein variant in Hiroshima and Nagasaki: tales of two cities. Am J Hum Genet 43:870–893

Nei M (1972) Genetic distance. Am Nat 106:283–292

Nei M, Roychoudhury AK (1974) Genetic variation within and between the three major races of man: Caucasoids, Negroids, and Mongoloids. Am J Hum Genet 26:421–443

Nei M, Tajima F (1981) DNA polymorphisms detectable by restriction endonucleases. Genetics 97:145–163

Olivo PD, Van de Walle MJ, Laipis PJ, Hauswirth WW (1983) Nucleotide sequence evidence for rapid genotypic shifts in the bovine mitochondrial DNA D-loop. Nature 306:400–402

Pagnier J, Mears JG, Dunda-Belkhodja O, Schaefer-Rego KE, Beldjord C, Nagel RL, Labie D (1984) Evidence for the multicentric origin of the sickle cell hemoglobin gene in Africa. Proc Natl Acad Sci USA 81:1771–1773

Rand DM, Harrison RG (1986) Mitochondrial DNA transmission genetics in crickets. Genetics 114:955–970

Rand DM, Harrison RG (1989) Molecular population genetics of mtDNA size variation in crickets. Genetics 121:551–569

Reed KC, Mann DA (1985) Rapid transfer of DNA from agarose gels to nylon membranes. Nucleic Acids Res 13:7207–7221

Roff DA, Bentzen P (1989) The statistical analysis of mitochondrial DNA polymorphisms: χ^2 and the problem of small samples. Mol Biol Evol 6:539–545

Rotig A, Colonna M, Bonnefort JP, Blanche S, Fischer A, Sandubray JM, Munnich A (1989) Mitochondrial DNA deletion in Pearson's marrow/pancreas syndrome. Lancet i:902–903

Santachiara-Benerecetti AS, Scozzari R, Semino O, Torroni A, Brega A, Wallace DC (1988) Mitochondrial polymorphisms in Italy II. Molecular analysis of new and rare morphs from Sardinia and Rome. Ann Hum Genet 52:39–56

Schroder WA, Munger ES, Powers DR (1990) Sickle cell anaemia, genetic variations and the slave trade to the United States. J Afr His 31:163–180

Scozzari R, Torroni A, Semino O, Sirugo G, Brega A, Santachiara-Benerecetti AS (1988) Genetic studies on the Senegal population. I. Mitochondrial DNA polymorphisms. Am J Hum Genet 43:534–544

Slatkin M, Maddison WP (1989) A cladistic measure of gene flow inferred from the phylogenies of alleles. Genetics 123: 603–613

Solomon E, Bodmer WF (1979) Evolution of the sickle variant gene. Lancet i:923

Srinivas R, Dunda O, Krishnamoorthy R, Fabry ME, Georges A, Labie D, Nagel RL (1988) Atypical haplotypes linked to the β^s gene in Africa are likely to be the product of recombination. Am J Hematol 29:60–62

Stine OC (1989) *Cepaea nemoralis* from Lexington, Virginia: the isolation and characterization of their mitochondrial DNA, the implications for their origin and climatic selection. Malacologia 30:305–315

Stoneking M, Bhatia K, Wilson AC (1986) Rate of sequence divergence estimated from restriction maps of mitochondrial DNAs from Papua New Guinea. Cold Spring Harbor Symp Quant Biol 51:433–439

Tills D, Kopec AC, Tills RE, Mourant AE (1983) The distribution of the human blood groups and other polymorphisms. Supplement 1. Oxford University Press, Oxford

Vigilant L, Pennington R, Harpending H, Kocher TD, Wilson AC (1989) Mitochondrial DNA sequence in single hairs from a southern African population. Proc Natl Acad Sci USA 86:9350–9354

Wiesenfeld SL (1967) Sickle-cell trait in human biological and cultural evolution. Science 157:1134–1140

Wilson DB (1949) Malaria incidence in central and south Africa. In: Boyd MF (ed) Malariology. WB Saunders, Philadelphia, pp 800–819

Received June 27, 1991/Revised and accepted November 11, 1991