

GENETIC PERSPECTIVES ON HUMAN ORIGINS AND DIFFERENTIATION

Henry Harpending and Alan Rogers

Department of Anthropology, University of Utah, Salt Lake City, Utah 84112

Key Words human origins, population expansion, race, Garden of Eden hypothesis, multiregional hypothesis

■ **Abstract** This is a review of genetic evidence about the ancient demography of the ancestors of our species and about the genesis of worldwide human diversity. The issue of whether or not a population size bottleneck occurred among our ancestors is under debate among geneticists as well as among anthropologists. The bottleneck, if it occurred, would confirm the *Garden of Eden* (GOE) model of the origin of modern humans. The competing model, *multiregional evolution* (MRE), posits that the number of human ancestors has been large, occupying much of the temperate Old World for the last two million years. While several classes of genetic marker seem to contain a strong signal of demographic recovery from a small number of ancestors, other nuclear loci show no such signal. The pattern at these loci is compatible with the existence of widespread balancing selection in humans. The study of human diversity at (putatively) neutral genetic marker loci has been hampered since the beginning by ascertainment bias since they were discovered in Europeans. The high levels of polymorphism at microsatellite loci means that they are free of this bias. Microsatellites exhibit a clear almost linear diversity gradient away from Africa, so that New World populations are approximately 15% less diverse than African populations. This pattern is not compatible with a model of a single large population expansion and colonization of most of the Earth by our ancestors but suggests, instead, gradual loss of diversity in successive colonization bottlenecks as our species grew and spread.

INTRODUCTION

This is a review of several important themes and directions in the study of genetic evidence about the history of our species and the origin of broad scale human differences. We have not written an exhaustive review of the relevant literature: instead we focus on evidence about ancient human demographic history and about the origins of human diversity on continental scales.

Although important insights about human population history have appeared in the literature over the last few decades, the idea that gene differences among humans could inform about the dynamics of human origins only became highly visible with the publication of Cann, Stoneking, and Wilson's famous paper (6)

showing that the most recent common ancestor of human mitochondrial DNA existed within the last few hundred thousand years. This paper fanned a debate that was going on among anthropologists about whether the origin of our species was a diffuse worldwide transformation of archaic humans into modern humans (multiregional evolution model, MRE) or whether our species had a focal origin from some unknown small population of archaics, after which it grew and spread over the temperate Old World (Garden of Eden model, GOE).

The mechanics of the origin of our species has important implications for understanding world patterns of human genetic diversity. If something like MRE occurred, then it is easy to understand the origin of human racial differences: they are ancient. But under GOE it is difficult to account for human differences over the Earth since they are too great.

BRIEF HISTORY OF THE FIELD

Genetic Inference About Human Origins: A New Field

In 1972 Haigh and Maynard Smith (20) published an analysis of protein sequence variants found in a sample of hemoglobins from 10,971 individuals. Their conclusion, discussed in more detail below, was that either the neutral theory of molecular evolution was wrong, or humanity had suffered a severe reduction in numbers to several tens of thousands or less at some time during the Pleistocene. This pioneering paper received little notice even from anthropologists.

There were indications by the 1980s that DNA diversity in humans was less than that of other species (40). Under the neutral mutation hypothesis, and under the hypothesis that mutations are rare, diversity should be proportional to population size over a long time ($4N$ generations). The low human diversity was another indication that we did not have a long history as a large successful species, in seeming contradiction to the human fossil record that showed archaic humans occupied much of the temperate Old World for 1 to 2 million years.

Detailed analysis of diversity in human mitochondrial DNA (mtDNA) showing that the top of the mitochondrial gene tree was only several hundred thousand years old brought the old Haigh and Maynard Smith finding out of obscurity. Since mtDNA is haploid and inherited through one sex, the effective number of genes is only one quarter the number of genes at a diploid nuclear locus. Hence the expected time back to the top of the tree under a model of constant population size is N generations, while the expected time for a nuclear locus is $4N$ generations. The mtDNA coalescent at, say, 200,000 years corresponds, with a generation time of 20 years, to an effective number of humans of 10,000. This is in good agreement with inferences from nuclear DNA diversity, which also suggest a human effective size of around 10,000.

As discussed below, the recent mtDNA coalescent, the relatively low nuclear diversity in humans, and the pattern of variation in hemoglobin protein sequences

all suggest that our species' effective size is 10,000. This might mean that the census size of human ancestors was about twice this, 20,000 or so, until very recently in evolutionary time: the end of the Pleistocene and the spread of agriculture 10,000 years or so ago is often imagined. On the other hand, there is not so much archaeological evidence for any major expansion in human numbers at this time. The competing model is that there was a large expansion in human numbers much earlier than this and that the overall low effective size of humans reflects a severe reduction in the number of human ancestors prior to an expansion sometime within the last few hundred thousand years.

Human Genetic Differentiation: An Old Field

While the use of genetic data to study human demographic history is a new endeavor, the use of genetic diversity to examine recent human history and human mating patterns over large areas is older and better established. In the first two thirds of the 20th century the number of genetic markers known in human populations slowly accumulated. Initially these were blood groups, then electrophoretic variants of proteins. The first systematic approaches to the use of these markers for studying human population processes occurred in the 1960s. Two kinds of models were applied to the accumulating world data, one that presumed that human populations have a tree-like history analogous to that of species, and the other that presumed that human differences reflected in situ differentiation in a geographically structured population. These two opposite views of the genesis of human diversity persist, and a major challenge today remains how to untangle historical processes from the effects of current structure.

The view that human populations represent tips of a tree of descent of populations was implicit in the work of Cavalli-Sforza and Edwards (7, 8). They developed methods to reconstruct the tree of descent under special assumptions about population history. Their method assumed that population fissions occurred, then that all daughter populations were equal in size after the fissions. Gene flow subsequent to fission and unequal sizes of daughter populations all would distort the results. Nevertheless, this was the beginning of an important tradition of reconstructing population history under the assumption that differences among human groups have been generated by cascading fissions. An important issue was whether the root of the tree, thought to be the fundamental and deepest division of humanity, was between East (Asia) and West (Europe and Africa) or North (Eurasia) and South (Africa).

The other important view of human differences was implicit in the work of Newton Morton and his colleagues on human population structure (43). In this tradition the key assumption is that genetic differences are the result of geographically restricted gene flow—*isolation by distance*—and the goal is to relate the relationship between genetic and geographical distance to rates of local gene exchange. Equilibrium between local drift, which leads to the accumulation of differences, and gene flow, which erodes differences, is assumed to have been reached.

Two models were prominent in data analysis in this tradition. In the first, gene frequencies were assumed to be samples from a homogeneous random field and the goal was to estimate the spatial autocorrelation function and the local migration rate that generated it. In the second, migration among a discrete set of populations was described by a matrix of migration rates, and a fit was sought between genetic distances observed among groups and those predicted from migration.

These approaches to understanding human genetic diversity were based on diametrically opposed assumptions, and neither could tell us very much about human history when the assumptions were violated. The tree approach took no account of migration, and the migration-drift equilibrium approach took no account of the history of population fissions. Neither approach was likely to give accurate answers when these assumptions were violated.

Given the large-scale movement and admixture of human DNA that we know about from history, any reconstruction of a tree of populations must be only an exercise in method and in computing. Much of the Spanish-speaking population of the New World, for example, carries both DNA that was in the New World and DNA that was in Europe a few centuries ago. Similarly, DNA in sub-Saharan Africa was borne both by very dark West Africans and very light Khoisan-speaking southern Africans several millennia ago. Any population tree generated from genetic differences among such populations cannot possibly have any meaning at all. Episodic range and size expansions also mean that gene frequencies will be far from spatial equilibrium between local gene flow and drift over large areas, so spatial autocorrelations are also difficult or impossible to interpret in general.

WAS THERE A BOTTLENECK?

The Issue

The word “origin” appears often in discussions of human evolutionary history and in review papers such as this one. Currently there is essentially no evidence from genetics about functional changes that occurred to change any archaic population(s) into anatomically modern humans. The best that current knowledge provides is evidence about the demographic history of our species. Debates about human origins are today only debates about that history, not about functional biological change.

There is a long-standing debate among anthropologists about the relationship between archaic human fossils and our species today. Fossils about 2 million years old of tool-using large-brained animals are found in and outside of Africa. The early versions of these have traditionally been lumped into the species *Homo erectus*, but recently it has become fashionable to divide these into a larger number of species. By several hundred thousand years ago some of these populations had larger brains and used a new technology for preparing stone tools [“mode 3,” (12)]

that involved the use of a prepared core from which flakes were detached. These later populations are usually called “archaic *Homo sapiens*,” and they include the Neandertals of Europe, which were morphologically distinct from contemporaries elsewhere in the world.

About 40,000 years ago a radical new technology, the upper Paleolithic, invaded Eurasia, reaching the Atlantic in the West and Siberia in the East within several thousand or tens of thousand years. The new tradition is marked by higher-quality stone tools and extensive use of worked bone. It is usually assumed that this technology is associated with anatomically modern humans since unequivocally modern remains are found after but not before this invasion. This relatively clear archaeological track in Eurasia is not matched anywhere south of the Himalayas. Humans appear in Australia at about this time, so it appears that there was a southern branch of the expansion of modern humans around the Indian Ocean. The technology of this southern branch, however, has no recognizable similarity to the distinctive tradition of the population that invaded Eurasia.

The debate has been about archaic humans and even earlier *Homo erectus* populations and their relationship to later populations of modern humans in each region. The MRE model suggests that there was genetic continuity from the earlier to the later inhabitants, so, for example, DNA from the famous *Homo erectus* population of Peking might still be in our species, especially in East Asia and in the New World. The GOE model, on the other hand, posits that modern humans developed in some restricted and specific region in an isolated subpopulation of archaics, then spread and occupied the Earth. Under this model, Peking *Homo erectus* can have no descendants alive today. The debate is old (62, 30) but it just simmered until the 1980s, when new interpretations of some African fossils heated it up; then the potential of genetics finally caught the attention of anthropologists with the publication of Cann, Stoneking, and Wilson’s paper (6) and the fanfare that accompanied it.

The potential of genetic data to resolve this controversy seems clear since patterns of genetic diversity ought to preserve demographic history. Under MRE, the number of ancestors of our species was large for the last 2 million years or so since they occupied much of the temperate Old World. An estimate (25) is that archaics occupied 25 million square kilometers of the Old World. If densities were like those of contemporary foragers, then there were between 125,000 and 1,000,000 of them on earth. Over this spatial scale there must have been extensive geographic subdivision and isolation, so their effective size ought to have been even greater than that implied by the rule of thumb that the effective size of humans is half census size.

Under GOE, on the other hand, our ancestry was focused in a small population of archaics that became ecologically successful and expanded over the earth. There may have been only a few tens of thousands of them, and their effective size would not have been inflated by subdivision. Under this model we should see in the genetic data a pinch or bottleneck as we look backward in time.

Gene Trees, Mismatch Distributions, and Site Frequency Spectra

What can genetic evidence tell us about the possibility that during the late Pleistocene, the human population passed through a “bottleneck”—a period of reduced population size? Five years ago, our review of this issue concentrated on mitochondrial evidence (56). Today we must integrate evidence from a variety of genetic systems. Five years ago, our discussion focused on the “mismatch distribution,” a histogram of genetic differences between pairs of individuals within a sample. We discuss the mismatch distribution again here, but our emphasis is elsewhere. The mismatch distribution is hard to interpret where recombination is important.

We focus here on what is called the “site frequency spectrum.” A polymorphic nucleotide site is ordinarily present in only two states within a sample, one of which is ancestral and the other “mutant.” The spectrum shows the fraction of the polymorphic sites at which the mutant form is present in one copy, in two copies, and so on. It is often impossible to determine the ancestral state, and in such cases we count the rarest form—referred to as the minor allele—rather than the mutant. In the first case, the spectrum is said to be “folded” since high-frequency and low-frequency sites are folded into a single category. When mutants can be recognized, the spectrum is said to be “unfolded.”

Population expansions generate a distinctive signature both in the mismatch distribution and in the spectrum. The mismatch distribution exhibits a unimodal wave, and the spectrum exhibits an excess of sites at which the minor allele is very rare. These principles form the basis for our discussion of the molecular data that have accumulated during the past 5 years. The latter principle was also at the heart of the argument of Haigh and Maynard Smith (20) almost 30 years ago. They studied variant protein sequences in approximately 22,000 human hemoglobins. There were 10 variants, i.e. sequences that were different from the common type. They ignored 3 of the 10 since they were known to be common and thought to be under strong selection. Of the remaining 7, 4 were found in a single copy in the sample and the other 3 in two copies. From this pattern of polymorphism they deduced that either the neutral theory of molecular evolution was wrong or that our species had undergone a severe bottleneck sometime before the end of the Pleistocene. They suggested that the effective number of ancestors before the expansion was about 10,000, corresponding to a census size of 40,000 or so under the extremely high mortality that they imagined to be characteristic of preagricultural humans.

In essence their argument was that all seven of the mutant forms of hemoglobin occurred at very low frequencies so they must be recent. There were no old mutations. Therefore our ancestral population must have passed through a bottleneck in which any polymorphisms at the loci responsible for hemoglobin were lost.

Today their argument would be phrased differently. Whereas they considered gene frequencies in a population and how they would have been affected by demography, the current style is to reason about the sample of genes and the morphology

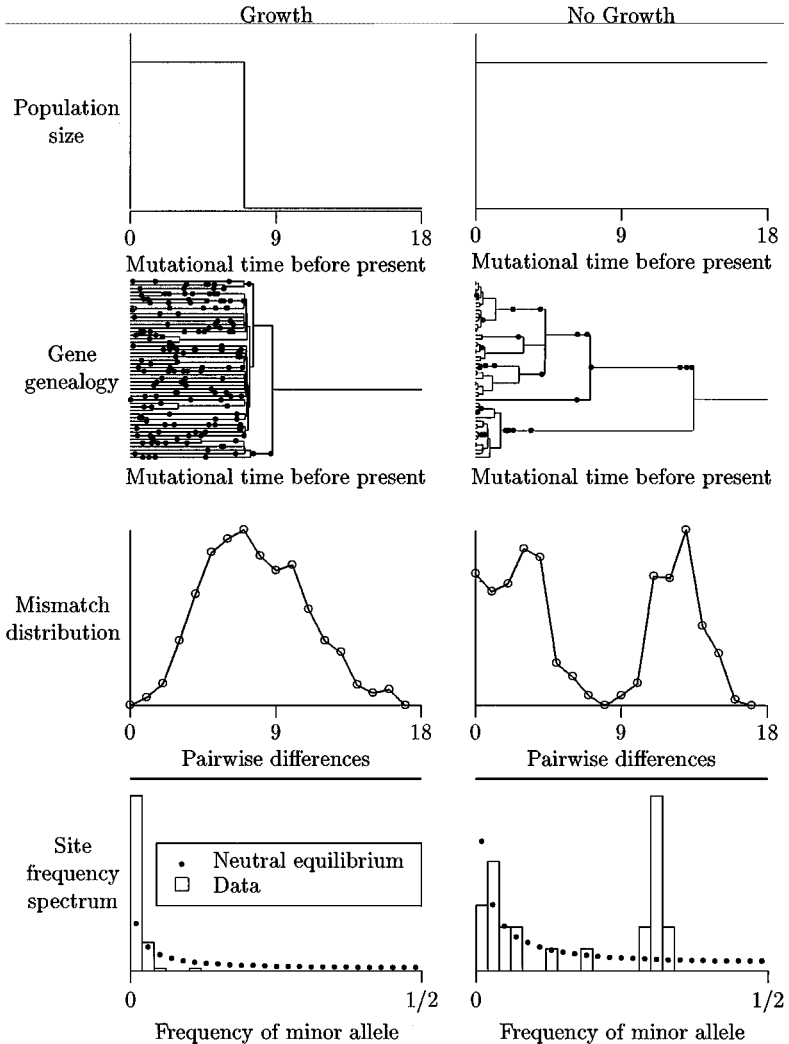


Figure 1 How demographic history influences the gene genealogy, the mismatch distribution, and the site frequency spectrum. Note that the vertical scales are different in the left and right bottom panels: the neutral equilibrium distributions are the same in each panel.

of the tree of descent joining them. The new style of reasoning is illustrated in Figure 1. The two columns in the figure refer to two hypothetical populations. The top panel in each column shows the history of population size. The horizontal axis measures time backward from the present in units of $1/(2u)$ generations, where u is the mutation rate per DNA sequence per generation. The vertical axis measures population size. The graph shows that the left-hand population was small prior

to 7 units of time ago, then grew suddenly, and has been large ever since. The population on the right has never changed in size.

The second panel in each column shows a hypothetical genealogy of 50 genes. Each was produced by a computer simulation that assumed the history shown in the upper panel. The horizontal axes measure time in the same mutational time units. Although humans have two parents, each gene is derived either from the mother or the father and consequently has only a single parent. As we trace the lineages backward into the past, pairs of lineages occasionally coalesce to form a single lineage. These events—called coalescent events—occur when two genes that are ancestral to genes in our sample happen to be copies of a single gene in the previous generation. This happens rarely in large populations, since two random individuals in such a population are unlikely to be siblings. Coalescent events are therefore rare during the period of large population size. Conversely, coalescent events are common during the period of small population size. The result is that in the expanded population coalescent events pile up just to the right of the population expansion. The genealogy of the constant-sized population looks very different. If we could observe gene genealogies directly, it would be easy to recognize expanded populations.

But we cannot observe the gene genealogy of any real population. We can, however, use genetic differences to draw inferences about its structure. The dots in the genealogies in Figure 1 were scattered at random along the branches, and their numbers are comparable with the number of mutational differences that we see in human mitochondrial D-loop sequence data. With real data, some mutations will obscure mutations that came before. But this happens rarely over the time scale that matters here, so we do not lose much by assuming that each mutation strikes a different site [the so-called model of infinite sites (35)]. The number of copies of each mutation is then equal to the number of leaves to its left along the genealogy. The third panel in each column of Figure 1 shows the mismatch distribution. The left-hand distribution shows a unimodal peak at 7, indicating that many of the pairs of genes in the sample differed by 7 mutations. This reflects the fact that, in the genealogy, many pairs differ by 7 units of mutational time. The right-hand mismatch distribution is much more ragged and has several peaks. This pattern reflects the pattern in the gene genealogy, which has coalescent events distributed much more widely in time. A smooth, unimodal peak in the mismatch distribution is the signature of a population expansion.

Population expansions also produce distinctive signatures in the site frequency spectra, which are shown in the bottom panel of each column of Figure 1. The spectra there are folded and show the observed spectrum as a series of open rectangles. The bold dots show the values that are expected under neutrality with constant population size. Notice the excess of singletons in the expanded population—the left-most rectangle rises far above the bold dot, indicating an excess of sites at which the minor allele occurs only once in the sample. This excess reflects the lengths of the terminal (tip-most) branches in the genealogy. The terminal branches are stretched out and comprise a large fraction of the genealogy. This

reflects the low rate of coalescent events in the expanded population. Because these terminal branches are so long, they capture a disproportionate fraction of the mutations. And the mutations that they capture are all singletons. The situation is reversed in the right-hand genealogy. There, the terminal branches comprise a small fraction of the genealogy and capture a correspondingly small fraction of the mutations. Consequently, singletons are much rarer. In summary, a population expansion produces a unimodal mismatch distribution and exaggerates the frequency of singletons within the set of polymorphic sites.

These principles make it possible to rephrase the argument of Haigh and Maynard Smith. Were they writing today, they might argue that the genealogy of their sample of 22,000 hemoglobins must look like the one on the left side of Figure 1 rather than the one on the right, since all 7 variants were extremely rare. This genealogy suggests a bottleneck in the history of human population size. Similar inferences have been drawn from the pattern seen in human mitochondrial DNA (4, 13, 25, 31, 41, 55, 60, 63). In recent years, the mitochondrial pattern has been used to infer not only that an expansion occurred but also that it occurred between 30,000 and 130,000 years ago, beginning from a population containing only a few thousand females and resulting in a population hundreds or thousands times larger (51–53).

The Selection Hypothesis

There is a problem, however, with the mitochondrial evidence: It could have been produced as easily by natural selection as by population growth. Suppose that a favorable mutation had occurred in, say, the mitochondrial genome of some human who lived 100,000 years ago. This mutant genome could have increased in frequency by natural selection until it was carried by all humans. If this had happened, all modern mitochondria would be descendants of this single mutant, and the gene genealogy would look a great deal like the one in the left-hand column of Figure 1. Because the gene tree implied by this hypothesis is so similar to the one implied by the expansion hypothesis, we cannot distinguish the two by looking at mitochondrial data alone.

There are approaches, however, that may allow us to distinguish these hypotheses. The first approach involves comparing the human mitochondrial data with mitochondrial data from other species. The second approach involves comparing different parts of the human genome.

Human and Chimp Mismatch Distributions

Figure 2 compares the mismatch distributions of homologous mitochondrial D-loop sequence in chimpanzees and African humans. The two distributions are remarkably similar. The human sequence is narrower and smoother, but the difference between them looks small compared with differences one observes in random samples simulated from the same history. The two distributions suggest that the mitochondria of humans and chimpanzees are each descended from

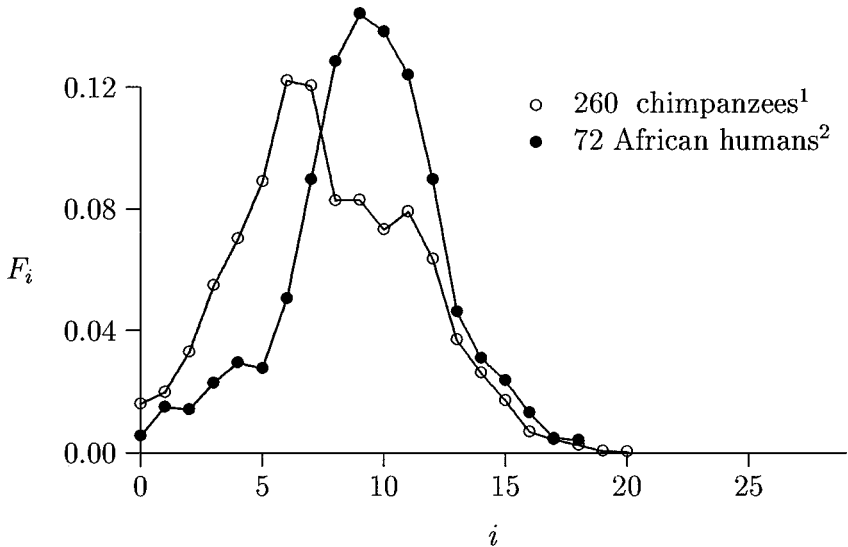


Figure 2 Chimpanzee and African human mitochondrial mismatch distributions. The chimpanzee and human distributions are based on 371 and 370 sites, respectively, of HVS1. (1) From Reference 19a; (2) From Reference 32.

a small population of mitochondria that increased in number during the late Pleistocene.

These patterns could have been caused by increases in the sizes of the chimpanzee and human populations or by selective sweeps of favorable mutations occurring in each population. But whatever the cause of these increases, they must have been roughly simultaneous. It seems unlikely that favorable mutations would occur at so nearly the same time in both species, but it is not hard to imagine conditions that would lead to simultaneous changes in population size. An environmental catastrophe such as an ice age or the Toba supervolcano (47, 48) could have reduced the population of each species. When conditions improved, the two populations could have recovered together, producing the record of roughly simultaneous growth that we see in Figure 2.

This evidence is suggestive, but it is not altogether convincing. In the first place, we have only two species. An environmental catastrophe should have produced the same pattern in several species. We should not expect to see it in every species, for a species that survives the catastrophe in two or more refugia will retain a far larger fraction of its genetic variance than one that survives in only one. The chimpanzee-human pattern should characterize only those species that survived in a single refugium (1). Thus, a few species that showed a different pattern would not refute the hypothesis of a Pleistocene bottleneck, but a few showing the same pattern would greatly strengthen it.

There are also other weaknesses in the chimpanzee-human argument. Suppose that slightly deleterious alleles are held by selection at low frequencies in both populations. An environmental change could cause a previously deleterious allele to become advantageous in both species, leading to simultaneous selective sweeps. Thus, the selection argument does not need to posit simultaneous favorable mutations. To us, this version of the selection hypothesis seems contrived and implausible, but pending some rigorous test, it cannot be dismissed.

Finally, there is evidence that the genetic variation in human mitochondria is not selectively neutral (17, 42, 64). The first two of these studies, however, reject a hypothesis that assumes not only that mitochondrial variation is neutral but also that the population's size has remained constant. Thus, the results may reflect population growth rather than selection. The study of Wise et al (64) is more troubling. They reject the neutral hypothesis unambiguously by showing that within a coding region of mtDNA, intrahuman variation includes a disproportionate fraction of the sites that cause amino acid replacements. This argues that the human mitochondrial gene pool includes some fraction of deleterious variants. But it is hard to see how selection against such variants could produce the pronounced waves that consistently appear in human mitochondrial mismatch distributions (see Figure 2). It seems likely that the wave is telling us about something else. Nonetheless, the evidence for selection reduces the confidence that we can have in historical inferences drawn from the mitochondrion. Other data are needed.

Nuclear DNA

Figure 3 shows all of the site frequency spectra that we were able to cull from the literature. In each panel, the open rectangles show data and the dots show the expected values under a model of neutral evolution and stationary population size. The top row contains spectra that, like the spectrum on the left side of Figure 1, show an excess of singletons relative to the stationary neutral model. The middle row contains spectra that fit the stationary model rather well, and the bottom row contains spectra with a deficit of singletons.

It seems unlikely that the variation among these systems is all sampling noise. Authors who have described the data sets in the first row have generally found significant departures from the stationary neutral model (23; S Wooding, AR Rogers, submitted for publication). There is also plenty of evidence for selection in the data sets shown in rows two and three, but this evidence does not in general lie in the site frequency spectrum.

Had all these spectra showed a pattern like that in the mitochondrion, the case for a population expansion would be clear. But these loci show a great variety of patterns, suggesting that natural selection has affected the loci in different ways (27, 29). To understand these differences, we need to consider how the spectrum responds to natural selection.

There are three forms of natural selection to consider: selection against deleterious alleles, selection for advantageous alleles, and balancing selection.

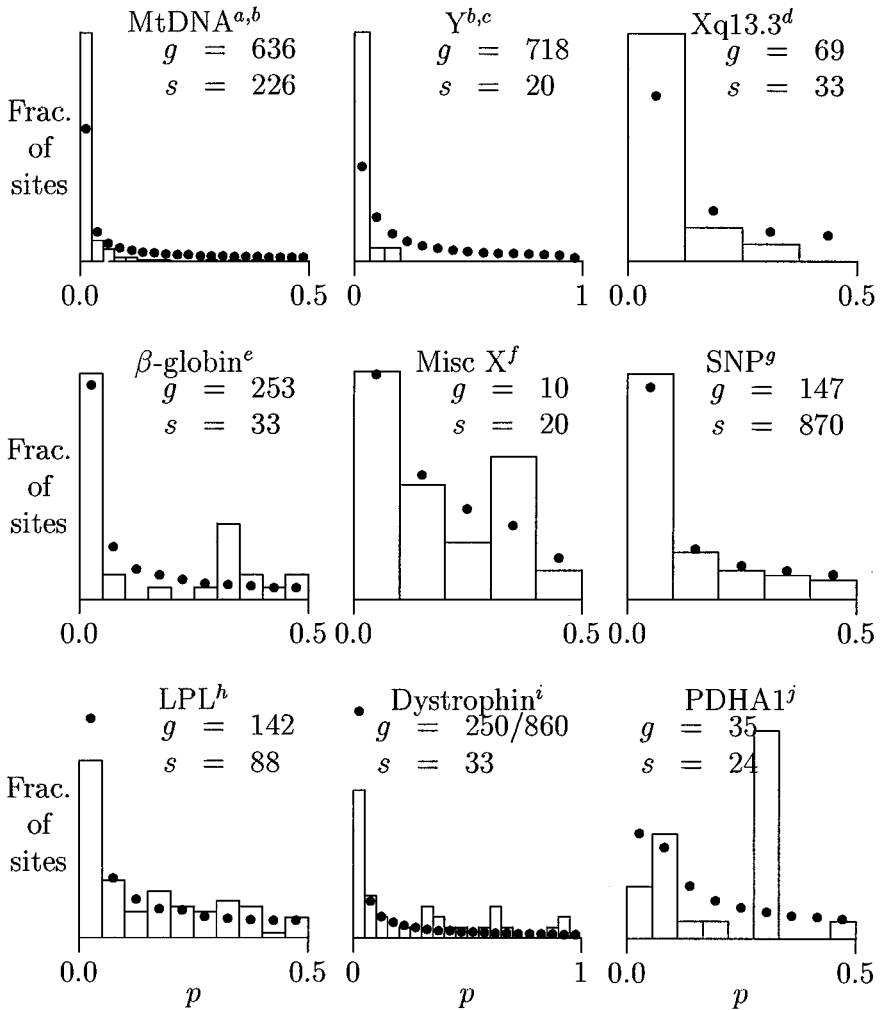


Figure 3 Site frequency spectra. The open rectangles in each panel show observed spectra; the bold dots show the spectra expected under the infinite sites model with no selection and constant population size. g is the number of chromosomes sampled, s is the number of segregating sites, and p is the frequency of the mutant allele (where ancestral state could be determined) or the rarest allele (where ancestral state is unknown). Sources: ^aJorde et al (32), ^bHarpending et al (23), ^cUnderhill et al (61), ^dKaessmann et al (34), ^eHarding et al (22), ^fNachman et al (44), ^gHalushka et al (21), ^hClark et al (11), ⁱZietkiewicz et al (69), ^jHarris & Hey (in preparation, 28).

Deleterious alleles arise from time to time by mutation and may survive within the population for some time before their eventual removal. But selection prevents them from becoming common, so they usually appear as singletons within a sample. Thus, deleterious mutations will inflate the frequency of singletons within the sample of sites. This occurs not only if the deleterious mutations appear within our sample of sites, but also if they are merely linked to sites within our sample (10). However, the effect of background selection on the site frequency spectrum is relatively weak. It seems unlikely that this process accounts for the pattern seen in the top row of Figure 3.

Advantageous alleles may also arise from time to time and sweep through the population under the influence of natural selection. To understand how this affects the spectrum, consider the rate at which an advantageous allele increases in frequency. This rate is high for alleles of intermediate frequency but is much lower for alleles with frequencies near 0 or 1. An advantageous allele spends a long time reaching a frequency of, say, 0.1, then increases rapidly to 0.9, then spends a long time increasing from 0.9 to 1.0. If we observe this process at a random time, we are disproportionately likely to find one allele very rare and the other very common. Consequently, this process also inflates the fraction of singletons in our sample of sites (3). This process could also produce the pattern seen in the top row of Figure 3.

Balancing selection refers to any of several processes that tend to maintain two or more alleles at polymorphic frequencies. For example, selection may favor the heterozygote over either homozygote or it may favor whichever allele is rare. Either way, we end up with an excess of alleles with frequencies near 1/2 and a deficit of singletons. This process might account for the pattern seen in the bottom row of Figure 3.

Thus, we can account for all of the patterns seen in Figure 3 by invoking various forms of selection to explain deviations from the stationary neutral expectation. But there are problems. First, it is suspicious that the loci that show the strongest evidence of selection (those in the top row of Figure 3) are exactly the loci that, on a priori grounds, we would have expected to be neutral. None of these loci are known to code for anything, and the sites assayed in the Y and Xq13.3 data are far from any known coding region. Paradoxically, it is the sequences that code for protein that look most nearly neutral.

And there is more. Not only do the data sets in the top row indicate that some expansion or selective event occurred, they can also be used to estimate the time of this event. The mitochondrial data provide a date of 30,000–130,000 years BP (51), and the Xq13.3 data provide a date of 130,000 BP (S Wooding, AR Rogers, submitted for publication). In addition, the 60 STR loci (32, 33; data not shown) imply a date between 50,000 and 100,000 BP (AR Rogers, LB Jorde, unpublished data). These dates are remarkably similar, especially in view of the notoriously uncertain mutation rate estimates that underlie them.¹

¹No data are yet available from the Y-chromosome since we do not yet know how to estimate a mutation rate.

It is hard to reconcile these dates with the selection hypothesis. One could propose, as we did above for mitochondria, that environmental change caused previously deleterious alleles to become advantageous simultaneously at several loci. This might affect even neutral sites, provided that they were linked to sites under selection. Yet the Xq13.3 sequence is far from the nearest known coding region. It is hard, moreover, to imagine that the STR data reflect a selective sweep, since none of these loci are within coding regions and they are scattered widely throughout the genome.

Let us consider, therefore, an alternative hypothesis. Suppose that the loci in the top row of Figure 3 are telling the truth—that in the late Pleistocene the human population did pass through a relatively narrow bottleneck and then expand. This hypothesis implies that an excess of singletons should exist at each of the other loci, yet this we do not see. If the hypothesis is correct, then some other process must have acted to reduce the singletons at each protein-coding locus. The only selective process capable of producing such a reduction is balancing selection. Thus, the bottleneck hypothesis implies that balancing selection has been important at all six of data sets in our sample that include protein-coding regions. Extrapolating to the whole genome, we are led to propose that balancing selection may be pervasive throughout the human genome.

This point of view was once common. It was advocated, for example, by Ford (18). But it has been out of fashion for 30 years. The first question that is raised by this proposal is that of genetic load. Since selection can only operate if some individuals are less able to survive and reproduce than others, the mean fitness of a population is reduced by each locus that is maintained at an overdominant equilibrium. To compensate for this reduction, a certain amount of excess fertility is required. If a large number of loci were to be maintained in this fashion, it appears that this excess fertility would have to be implausibly large. For this reason, many authors have concluded that balancing selection can account for only a small fraction of the polymorphic loci that we see in nature. This argument has been enormously influential but has also attracted its share of critics (for contrasting reviews of this subject, see 16, 36, 45). We think that the argument has misled us all by focusing attention on the fitness of a nonexistent optimal genotype. The issue is more usefully framed by asking how many overdominant loci can be maintained by a given amount of genetic variance in fitness. Gillespie (19, pp. 270–72) has developed a model that does this. He shows that if genotypic fitness is formed multiplicatively from the fitnesses of individual loci, then the coefficient of variation in fitness is

$$cv = \sqrt{ns},$$

where n is the number of loci and s is the coefficient of selection against homozygotes at each locus. If selection is to be effective at each locus, then it must also be true that $s > 1/(2N)$, where N is the effective population size. This gives

$$n < (2N \cdot cv)^2.$$

For any given population size and coefficient of variation in fitness, this provides an upper bound on the number of overdominant polymorphisms that can be maintained. To get an idea of the magnitudes involved, suppose that each woman has 10 children of which 2 survive and that this mortality is the only contribution to variance fitness. Then

$$cv = \frac{\sqrt{\text{variance}}}{\text{mean}} = \frac{\sqrt{0.2 \times 0.8}}{0.2} = 2.$$

If $N = 100,000$, then inequality 1 becomes $n < 1.6 \times 10^{11}$. Even if the cv were only 0.01, the model would still allow for 4 million polymorphic loci. If polymorphisms are maintained by frequency-dependent rather than overdominant selection, the number of loci that can be maintained is even larger since such loci contribute to genetic variance in fitness only when they are away from equilibrium (16). In summary, it appears that genetic load imposes no obstacle to the view that a large fraction of the loci in the human genome are influenced by some form of balancing selection.

There are, of course, other objections. To begin with, comparisons between species reveal that substitutions occur more rapidly at sites that have no effect on protein than at sites that change proteins (39). This is as one would expect if most substitutions are neutral, but is not predicted by a model in which most substitutions are selectively advantageous alleles. Yet this observation is not germane here, since substitutions occur in spite of balancing selection, not because of it.

There are other observations that are more germane. To choose just one example, Yamazaki and Maruyama (67) did an analysis with classical polymorphisms that is analogous to contemporary analyses of site frequency spectra. Their procedure differed only in that they counted alleles rather than sites, and in that they weighted the count of alleles with frequency p by the heterozygosity, $2p(1-p)$. Deep down, however, the principles at work were the same as those influencing the site frequency spectra in our Figure 3.² Their results were consistent with a neutral model. But their neutral model assumed, in addition to selective neutrality, that the population was constant in size. Thus, their results may reflect the antagonistic effects of a bottleneck (which increases the frequency of rare alleles) and balancing selection (which decreases that frequency). It is possible that much of the apparent evidence in favor of the neutral theory of molecular evolution is artifactual, produced by the confounded effects of these two forces.

Gillespie (19, p. 63) observes that

If the equilibrium models that we apply to our data are valid, we have to assume that populations have been stable for periods of time that greatly

²Under the stationary neutral model, a locus that is polymorphic within the sample has minor allele frequency p with probability proportional to $1/p + 1/(1-p) = 1/(p(1-p))$. This formula was used to produce the bold dots in Figures 1 and 3. Yamazaki and Maruyama weighted the count of alleles with frequency p by $p(1-p)$ in order to obtain a statistic that did not vary with p .

exceed the time scale of frequency change. Species must maintain their population sizes, migration rates, mutation rates, and, most distressingly, not experience any strong selection events at loci linked to those used in our enzyme studies. Otherwise, the populations will not be in equilibrium. This is precisely what I feel is going on.

So do we.

DIFFERENTIATION

Humans look different in different parts of the world. Anyone could look around and immediately know where he was if given the choices of, say, Lusaka, Tokyo, and Paris. A lot of effort by anthropologists early in the century to emulate taxonomists and to untangle human differences by inferring racial histories never amounted to anything substantial. With the availability of large numbers of genetic markers in the last few decades, efforts to reconstruct biological history and to understand the genesis of human differences have become prominent in the literature again. Human neutral genes tell an interesting story that we review here, but their story does not correspond very closely to the stories told by physical appearance and language. Below we discuss reasons why these different indicators might not agree with each other.

World F_{st}

A standard statistic that is a measure of genetic differentiation among populations is called F_{st} , although there are many variants that in effect measure the same thing. If H_o is a measure of average genetic diversity (heterozygosity or some equivalent) within populations and H_e is genetic diversity in the same populations if they were to mate at random with each other, F_{st} is just $(H_e - H_o)/H_e$. At equilibrium between gene flow among populations and local genetic drift, the value of F_{st} should depend just on the number of migrants in each generation that move among populations. If migration among a set of populations is described by a matrix of migration rates, and if any population is reachable ultimately from any other, there is still a single number that describes the amount of migration and hence the equilibrium value of F_{st} (54).

Many studies have shown that the value of this statistic among samples of world populations is 10% to 15%. Two early presentations of this value reached opposite conclusions. Wright (66) commented that if racial differences this large were seen in another species, they would be called subspecies. Lewontin (38), on the other hand, argued that 10% was a small number, that humans are therefore remarkably homogeneous, and that human differences are trivial. Most later authors have adopted Lewontin's point of view.

The figure of 10% to 15% has been confirmed many times in the subsequent decades, but the conclusion that 10% is a small number has been misleading for

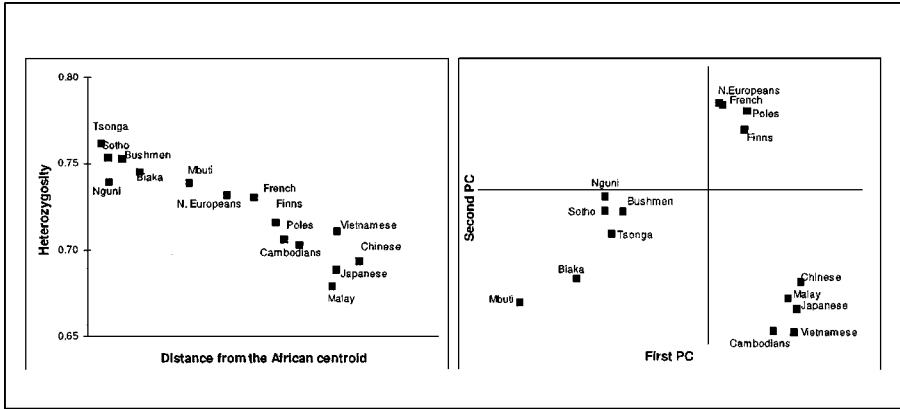


Figure 4 Patterns in microsatellite markers in 15 populations analyzed in Lynn Jorde's laboratory.

two reasons. First, this is a measure of human diversity in neutral genes. There are reasons detailed below to suspect that human differences that are visible, for example in Paris, Lusaka, and Peking, do not follow the same dynamics. Second, the figure of 10% is difficult to reconcile with the GOE model in which our ancestors underwent expansion in numbers as they populated the Earth. It is too large to accommodate such a simple model (56).

Diversity Gradients

The F_{ST} statistic is a measure of relative diversity independent of absolute diversity. Classical genetic markers, blood groups and protein electrophoretic variants, never revealed strong or interesting world differences in absolute genetic diversity (9). Everyone realized that most of these had been discovered in Europeans and that there was an overall sampling bias in favor of European diversity.

With the widespread typing of microsatellites, however, we now have an excellent multilocus system for assessing neutral diversity differences within our species. Rogers and Jorde (57) showed that ascertainment bias would be negligible with markers with high levels of diversity (i.e. heterozygosity), and microsatellites more than fulfill their criteria (2). There are two large sets of microsatellite typings available that we discuss here, one from Lynn Jorde's laboratory at the University of Utah (32) and the other from Kenneth Kidd's laboratory at Yale University (5).

The left panel of Figure 4 shows average heterozygosity in a sample of 60 polymorphic microsatellite markers from 15 populations in the Old World plotted against genetic distance from the African centroid on the horizontal axis (15, 24). The pattern is unambiguous: there is a linear diversity gradient away from Africa. African populations show the highest heterozygosities, then Europeans, then East Asians, with the lowest heterozygosities in this sample.

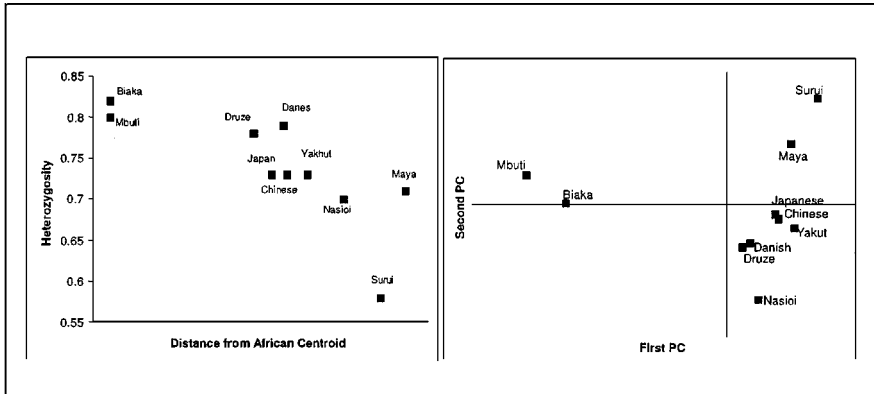


Figure 5 Patterns in microsatellite markers in 10 populations analyzed in Kenneth Kidd's laboratory.

This linear decline in heterozygosity is not quite so apparent in the left panel of Figure 5, which is the same graph but with different populations based on a different set of microsatellite markers. The American Indian populations are genetically the most different from Africans, but one, the Maya, shows relatively high levels of microsatellite diversity, whereas the other, the Surui of Amazonia, shows drastically reduced diversity. There are at least two possible interpretations of this pattern in the New World populations. First, the Maya are part of a large population and so have not lost diversity in the last few millennia, while the Surui are a small isolate that has undergone extensive genetic drift and diversity loss. Second, it is possible that the Maya have incorporated a lot of European DNA and that the Surui represent the precontact neutral diversity in Amerindian populations.

Patterns of Genetic Distance

The right panels of Figures 4 and 5 portray genetic differences among the populations in the form of principal coordinates plots. These are the best (in the sense of least squares) two-dimensional representations of the distances: two populations close to each other on these displays are genetically similar, while distant populations in the displays are genetically different. The data from Jorde's laboratory shown in Figure 4 are mostly from large populations, and there is a clear pattern. There are three clusters that correspond to the traditional three major races. The African populations are relatively dispersed while European and East Asian populations form tight clusters.

We might be led to reify the traditional three races from examining genetic distances in Figure 4, but a glance at the distances in Figure 5 shows that that would be premature. These data, from Kidd's laboratory, include several small populations that have probably been small and isolated for a long time. The effect of drift, greatest in small populations, is to cause the populations to "wander"

on a plot of genetic distances such as this, and the interesting features of this portrayal of distances is the way that small populations are outliers. Taken at face value, this plot suggests that the two Amerindian groups, Maya and Surui, are more different from each other than are Chinese and Danes. This is a correct portrayal of differences in neutral gene frequencies, but it is easy to see that historical processes and recent demography are hopelessly confounded. The Surui are different from the Maya because of a lot of recent drift, but the Japanese are different from Danes because of ancient demographic history.

This exercise suggests that simple studies of genetic distance over large geographic scales may often be misleading. What is needed in this field is a technique for using the information in the left panels of Figures 4 and 5 to adjust the information in the right panels. For example, the low diversity of the Surui suggests that they have undergone extraordinary diversity loss because of drift. Another consequence of this drift would have been to exaggerate their genetic distance from other populations, and there ought to be a way to adjust Surui distances to account for this. A step in this direction has been proposed by Relethford (49), who showed that drift in small Jewish populations in West Asia led to spuriously large measures of genetic distances among them. Since he had information about historical sizes of the isolates, he was able to adjust the distances and show that there had been much less outbreeding with neighboring populations than that suggested by unadjusted distances. This paper of Relethford's should foreshadow a new level of sophistication in human population genetic data analysis.

Genesis of Neutral Differences Among Human Populations

If multiregional evolution is the right portrayal of human history, then our species' ancestry has been dispersed over much of the temperate Old World for a million years or more. Under this model there is no difficulty understanding how genetic differences among populations arose. They are ancient and they reflect isolation by distance in a structured population with, of course, episodic population expansions such as the Bantu expansion into sub-Saharan Africa or the European expansion into the New World overlaid. But if the GOE hypothesis is correct, then it is hard to understand how human differentiation occurred.

To understand the difficulty with human differentiation under GOE, consider a simple model of a homogeneous population that instantly divides into daughter populations, each of effective size N_0 . These daughter populations each exchange M migrants with a mixed pool of emigrants from all the daughter populations. The daughters also grow exponentially at rate r so that after T generations each daughter population is of size $N_0 e^{rT}$. Write $F(T)$ for F_{st} at generation T after the separation event.

First consider the extreme case where the daughter populations were completely isolated precursors of discrete races ($M = 0$) and never grew ($r = 0$). In this case,

$$F(T) = 1 - e^{-T/2N_0}.$$

If the modern human expansion occurred 50,000 years ago, say $T = 2000$ generations, and the effective sizes of the daughter populations were $N_0 = 10,000$, the predicted F today is only 0.096. This extreme and unlikely scenario, with no gene flow among the daughter populations, barely generates the observed diversity of 10% to 15% observed today. Any gene flow among the daughters, of course, would reduce the predicted diversity drastically.

A more likely scenario is that daughter populations were small initially and then grew exponentially in their new environments. If r is positive, then populations continue to grow, but F quickly approaches a limit because it hardly changes in large populations. Suppose that ancient racial precursor populations, each of effective size of $N_0 = 2000$, split and went off to their separate continents. Each grew at the low (for humans) rate of 10% per generation ($r = 0.10$). The new equilibrium

$$F = 1 - e^{1/2N_0r}$$

of $F = 0.0025$ would be reached quickly. This is only one fortieth to one fiftieth of the observed world F . Again, this supposes no gene flow among the daughter populations: F would be slightly smaller with gene flow.

There are at least two other possible scenarios for the origin of continental scale human gene differences. One supposes that the ancestral population of modern humans was subdivided before the expansion so that race differences are older than the ecological success of modern humans. Harpending et al (25) called this the “weak Garden of Eden hypothesis.” This model was independently described by Lahr and Foley (37) on the basis of archaeological and paleontological evidence. There are also hints from reconstruction of phylogenies of genes that race differences are older than the modern human expansion (26).

Another mechanism for generating human neutral diversity may have been cascading bottlenecks as the range of our species increased. If humans in new environments were often descended from small pioneer bands, then each such episode, if these bands were small, would cause a loss of neutral diversity in the colonizer. This mechanism could account for the diversity cline away from Africa that is so clearly shown in the microsatellites.

Non-Neutral Traits

Ten to fifteen percent of human genetic diversity at neutral loci is between populations and the rest is diversity within populations. This number describes neutral variation among human populations and not necessarily variation that is not neutral. For example Relethford (49) estimate that the equivalent F_{st} for skin color, treated as a quantitative trait, is 0.60. This is five to six times greater than the figure for neutral genes. Certainly other visible traits that most humans notice are more like skin color than they are like neutral traits. This was emphasized by Nei & Roychoudhury (46) in their pioneering work with gene differences among races when they said that their results did “not apply to those genes which control morphological characters such as pigmentation and facial structure.”

Diamond (14) reviews evidence about prominent visible morphological differences among human populations. There are “explanations” for many of these traits involving adaptation to the environment—for example, that skin color modulates vitamin D synthesis or that the Asian epicanthic fold blocks glare off the snow. Unfortunately there is little compelling evidence to support any of the explanations. Diamond suggests that Darwin was correct in attributing most visible differences to sexual selection, the maintenance of originally arbitrary preferences by reproductive partner choice. Sexually selected preferences are maintained not because they are adaptive responses to the environment but because the other sex prefers them. Even though they may be bizarre and likely maladaptive to the environment, like the tail on the male peacock, the females may prefer them and, since other females also prefer them, the sons of the preferring female will also have high fitness. Both the preference and the trait can be maintained. The less the trait detracts from fitness in the environment the easier it is to maintain by a system of sexual selection.

If suites of visible human traits have been generated and maintained by sexual selection, there are interesting implications that should be explored. Evolutionary dynamics of visible traits could bear little or no relationship to the dynamics of neutral traits, and this seems to be what we observe. Neutral traits passively reflect demographic history, population size, and gene flow, whereas sexually selected traits could reflect older population-specific characteristics. Imagine, for example, that the small size of African Pygmies was maintained by reproductive preference. Then gene flow from neighboring larger populations could homogenize regional neutral gene frequencies while even a weak preference for small size within Pygmies could maintain the difference. In this way size could provide an older description of the population differences than that given by neutral genes.

There is a parallel with language. Ruhlen (59) discusses the maintenance of Basque, a unique language spoken in a region of Europe between France and Spain. He points out that low-level gene flow into Basques over thousands of years would result in the virtual eradication of gene differences between Basques and their neighbors. Language does not mix in the same way, and the immigrants would simply have learned to speak Basque. There is a common type advantage in this linguistic model, and similarly sexually selected traits may follow common type advantage. Just as the Basque language may tell us about a history that has been erased in the neutral genome, physical appearance may tell us about a deeper history than do neutral genes.

Harpending and Eller (24) discuss morphological similarities between Khoisan-speaking peoples of southern Africa and East Asian populations, speculating that these unlikely similarities may reflect an old shared ancestry, the traces of which have been eradicated by gene flow in neutral markers. Many Khoisan-speakers have yellowish skin, epicanthic folds, shovel-shaped incisors, and “mongoloid spots,” a bluish mark on the lower back in young infants.³ All these are

³In the !Kung language spoken by Bushmen in the northern Kalahari both black Africans and Europeans are called “!ohm,” a category that includes predators and other inedible animals. Asians, on the other hand, are immediately called “zu,” meaning people.

conventionally regarded as East Asian traits, and it is asking a lot of coincidence that they should be shared with a group in the southern third of Africa. A more likely explanation is that these traits are of a complex maintained by sexual selection and that they reveal some ancient shared ancestry that has been obscured in the neutral genome by gene exchange with neighbors.

From this perspective the old ideas that there were more discrete “races” in the distant past does not seem so absurd. This old idea is consistent with the “weak Garden of Eden hypothesis” (25) of ancient subdivision discussed above, and it is consistent with the model that visible differences and language each preserve ancient signatures of history that have been erased in the neutral genome by high levels of local gene flow since the end of the Pleistocene and the spread of agriculture in the last ten millennia or so.

CONCLUSIONS

Genetics and the GOE Hypothesis

Five years ago, we would have said that genetic evidence provided unambiguous support for the GOE model of human origins. Today, the case is far less clear. Several loci indicate that the human population passed through a bottleneck—a period of small population size. These loci seem to support the GOE hypothesis. Yet other loci indicate just as strongly that no bottleneck has occurred within the past several hundred thousand years. Any attempt to use genetic data in unraveling human history must deal with the discrepancy between these sets of loci.

We opt in favor of a bottleneck for two reasons: First, the loci that support a bottleneck are precisely those that on a priori grounds seem most likely to be neutral. Second, the loci that support a bottleneck all agree about when it occurred. It is hard to reconcile this agreement with the alternative hypothesis, which holds that the apparent evidence for expansion is really evidence for selection. But if we accept the bottleneck hypothesis, we must find some way to account for the loci that fail to support the bottleneck hypothesis. This leads us to propose that there is extensive balancing selection throughout the portions of the nuclear genome that code for protein.

Origin of Large-Scale Human Differences Is Not Understood

Five years ago we would have said that the overall measure of difference among worldwide human populations, F_{ST} , of 10% had been known for decades and that accumulating data were only confirming it. Today this remains as true as ever. What is new is data that are not contaminated by ascertainment bias, like microsatellite markers and DNA sequences (68) from world samples. These can be used to compare within-population diversity, and they reveal a clear pattern. Diversity declines smoothly with genetic distance from Africa, so that heterozygosity at microsatellites in the New World is about 15% less than heterozygosity at the same loci in Africans.

It is difficult to understand how F_{st} of 10% could have come about if the GOE hypothesis is correct since such large differences would take too long to accumulate between large populations. One way out of the problem is to posit that race differences are older than the expansion of our species; another is to posit that drift during successive colonization bottlenecks led to worldwide differences along with the worldwide diversity cline.

Visit the Annual Reviews home page at www.AnnualReviews.org

LITERATURE CITED

- Ambrose SH. 1998. Late Pleistocene human population bottlenecks, volcanic winter, and differentiation of modern humans. *J. Hum. Evol.* 34:623–51
- Bowcock AM, Kidd JR, Mountain JL, Hebert JM, Carotenuto L, et al. 1991. Drift, admixture, and selection in human evolution: a study with DNA polymorphisms. *Proc. Natl. Acad. Sci. USA* 88:839–43
- Bravereman JM, Hudson RR, Kaplan NL, Langley CH, Stephan W. 1995. The hitchhiking effect on the site frequency spectrum of DNA polymorphisms. *Genetics* 140:783–96
- Brown WM. 1980. Polymorphism in mitochondrial DNA of humans as revealed by restriction endonuclease analysis. *Proc. Natl. Acad. Sci. USA* 77:3605–9
- Calafell F, Shuster A, Speed W, Kidd J, Kidd K. 1997. Short tandem repeat polymorphism evolution in humans. *Eur. J. Hum. Genet.* 6:38–49
- Cann RL, Stoneking M, Wilson AC. 1987. Mitochondrial DNA and human evolution. *Nature* 325(1):31–36
- Cavalli-Sforza LL, Edwards AWF. 1964. Analysis of human evolution. In *Proc. 11th Int. Cong. Genet.* pp. 923–33
- Cavalli-Sforza LL, Edwards AWF. 1967. Phylogenetic analysis: models and estimation procedures. *Am. J. Hum. Genet.* 19:233–57
- Cavalli-Sforza LL, Menozzi P, Piazza A. 1994. *The History and Geography of Human Genes*. Princeton, NJ: Princeton Univ. Press
- Charlesworth B, Morgan MT, Charlesworth D. 1993. The effect of deleterious mutations on neutral molecular variation. *Genetics* 134:1289–1303
- Clark AG, Weiss KM, Nickerson DA, Taylor SL, Buchanan A, et al. 1998. Haplotype structure and population genetic inferences from nucleotide-sequence variation in human lipoprotein lipase. *Am. J. Hum. Genet.* 63:595–612
- Clark J. 1977. *World Prehistory: A New Perspective*. Cambridge: Cambridge Univ. Press. 554 pp.
- Di Rienzo A, Wilson AC. 1991. Branching pattern in the evolutionary tree for human mitochondrial DNA. *Proc. Nat. Acad. Sci., USA* 88:1597–1601
- Diamond J. 1992. *The Third Chimpanzee*. New York: HarperPerennial. 407 pp.
- Eller E. 1999. Population substructure and isolation by distance in three continental regions. *Am. J. Phys. Anthropol.* 108: 147–59
- Ewens WJ. 1979. *Mathematical Population Genetics*. New York: Springer-Verlag. 325 pp.
- Excoffier L. 1990. Evolution of human mitochondrial DNA: Evidence for departure from a pure neutral model of populations at equilibrium. *J. Mol. Evol.* 30:125–39
- Ford E. 1964. *Ecological Genetics*. London: Methuen. 442 pp.
- Gillespie JH. 1991. *The Causes of Molecular Evolution*. New York: Oxford Univ. Press

- 19a. Goldberg TL. 1996. *Genetics and biogeography of East African Chimpanzees (Pan troglodytes schweinfurthii)*. PhD thesis. Harvard Univ., Cambridge, MA. 269 pp.
20. Haigh J, Maynard Smith J. 1972. Population size and protein variation in man. *Genet. Res. Cambridge* 19:73–89
21. Halushka MK, Fan JB, Bently K, Hsie L, Shen N, et al. 1999. Patterns of single-nucleotide polymorphisms in candidate genes for blood-pressure homeostasis. *Nat. Genet.* 22:239–47
22. Harding RM, Fullerton SM, Griffiths RC, Bond J, Cox MJ, et al. 1997. Archaic African and Asian lineages in the genetic ancestry of modern humans. *Am. J. Hum. Genet.* 60:722–89
23. Harpending HC, Batzer MA, Gurven M, Jorde LB, Rogers AR, Sherry ST. 1998. Genetic traces of ancient demography. *Proc. Nat. Acad. Sci. USA* 95:1961–67
24. Harpending HC, Eller E. 1999. Human diversity and its history. In *The Biology of Biodiversity*, ed. M Kato, Chap. 20, pp. 301–14. Tokyo: Springer-Verlag
25. Harpending HC, Sherry ST, Rogers AR, Stoneking M. 1993. The genetic structure of ancient human populations. *Curr. Anthropol.* 34:483–96
26. Harris E, Hey J. 1999. X chromosome evidence for ancient human histories. *Proc. Nat. Acad. Sci. USA* 96:3320–24
27. Harris EE, Hey J. 1999. Human demography in the Pleistocene: Do mitochondrial and nuclear genes tell the same story? *Evol. Anthropol.* 8:81–86
28. Deleted in proof.
29. Hey J. 1997. Mitochondrial and nuclear genes present conflicting portraits of human origins. *Mol. Biol. Evol.* 14(2): 166–72
30. Howells WW. 1942. Fossil man and the origin of races. *Am. Anthropol.* 44:182–93
31. Jones JS, Rouhani S. 1986. How small was the bottleneck? *Nature* 319:449–50
32. Jorde LB, Bamshad MJ, Watkins WS, Zenger R, Fraley AE, et al. 1995. Origins and affinities of modern humans: a comparison of mitochondrial and nuclear genetic data. *Am. J. Hum. Genet.* 57:523–38
33. Jorde LB, Rogers AR, Bamshad M, Watkins WS, Krakowiak, et al. 1997. Microsatellite diversity and the demographic history of modern humans. *Proc. Natl. Acad. Sci. USA* 94:3100–3
34. Kaessmann H, HeiBig F, von Haeseler A, Pääbo S. 1999. DNA sequence variation in a non-coding region of low recombination on the human X chromosome. *Nat. Genet.* 22:78–81
35. Kimura M. 1971. Theoretical foundation of population genetics at the molecular level. *Theor. Popul. Biol.* 2:174–208
36. Kimura M. 1983. *The Neutral Theory of Molecular Evolution*. New York: Cambridge Univ. Press
37. Lahr M, Foley R. 1994. Multiple dispersals and modern human origins. *Evol. Anthropol.* 3(2):48–60
38. Lewontin RC. 1972. The apportionment of human diversity. In *Evolutionary Biology*, ed. M. Hecht, Vol. 6, pp. 381–98. New York: Appleton-Century-Crofts
39. Li WH. 1997. *Molecular Evolution*. Sunderland, MA: Sinauer
40. Li WH, Sadler LA. 1991. Low nucleotide diversity in man. *Genetics* 129:513–23
41. Maynard Smith J. 1990. The Y of human relationships. *Nature* 344:591–92
42. Merriwether DA, Clark AG, Ballinger SW, Schurr TG, Soodyall H, et al. 1991. The structure of human mitochondrial DNA variation. *J. Mol. Evol.* 33:543–55
43. Morton NE, ed. 1973. *Genetic Structure of Populations*. Honolulu: Univ. Hawaii
44. Nachman MW, Bauer VL, Crowell SL, Aquadro CF. 1998. DNA variability and recombination rates at x-linked loci in humans. *Genetics* 150:1133–41
45. Nei M. 1975. *Molecular Population Genetics and Evolution*. Amsterdam: North-Holland

46. Nei M, Roychoudhury A. 1982. Genetic relationship and evolution of human races. *Evol. Biol.* 14:1–59
47. Rampino MR, Self S. 1992. Volcanic winter and accelerated glaciation following the Toba super-eruption. *Nature* 359:50–52
48. Rampino MR, Self S. 1993. Bottleneck in human-evolution and the Toba eruption. *Science* 262:1955
49. Relethford JH. 1992. Cross-cultural analysis of migration rates: effects of geographic distance and population size. *Am. J. Phys. Anthropol.* 89:459–66
50. Relethford JH, Harpending HC. 1994. Craniometric variation, genetic theory, and modern human origins. *Am. J. Phys. Anthropol.* 95:249–70
51. Rogers AR. 1995. Genetic evidence for a Pleistocene population explosion. *Evolution* 49(4):608–615
52. Rogers AR. 1996. Mitochondrial mismatch analysis is insensitive to the mutational process. *Mol. Biol. Evol.* 13(7): 895–902
53. Rogers AR. 1997. Population structure and modern human origins. In *Progress in Population Genetics and Human Evolution*, eds. P. J. Donnelly, S. Tavaré, pp. 55–79. New York: Springer-Verlag
54. Rogers AR, Harpending HC. 1986. Migration and genetic drift in human populations. *Evolution* 40:1312–27
55. Rogers AR, Harpending HC. 1992. Population growth makes waves in the distribution of pairwise genetic differences. *Mol. Biol. Evol.* 9:552–69
56. Rogers AR, Jorde LB. 1995. Genetic evidence on modern human origins. *Hum. Biol.* 67(1):1–36
57. Rogers AR, Jorde LB. 1996. Ascertainment bias in estimates of average heterozygosity. *Am. J. Hum. Genet.* 58:1033–41
58. Deleted in proof.
59. Ruhlen M. 1994. *The Origin of Language*. New York: Wiley & Sons
60. Sherry S, Rogers AR, Harpending HC, Soodyall H, Jenkins T, Stoneking M. 1994. Mismatch distributions of mtDNA reveal recent human population expansions. *Hum. Biol.* 66(5):761–75
61. Underhill PA, Jin L, Lin AA, Mehdi SQ, Jenkins T, et al. 1997. Detection of numerous Y chromosome biallelic polymorphisms by denaturing high-performance liquid chromatography. *Genome Res.* 7:996–1005
62. Weidenreich F. 1940. Some problems dealing with ancient man. *Am. Anthropol.* 42:375–83
63. Wills C. 1990. Population size bottleneck. *Nature* 348:398
64. Wise C, Sraml M, Eastale S. 1998. Departure from neutrality at the mitochondrial NADH dehydrogenase subunit 2 gene in humans, but not in chimpanzees. *Genetics* 148:409–21
65. Deleted in proof.
66. Wright S. 1978. *Evolution and the Genetics of Populations, Vol. 4, Variability Within and Among Natural Populations*. Chicago, II: Univ. Chicago Press
67. Yamazaki T, Maruyama T. 1974. Evidence that enzyme polymorphisms are selectively neutral, but blood group polymorphisms are not. *Science* 183:1091–92
68. Zietkiewicz E, Votova V, Jarnik M, Koran-Laskowska M, Kidd KK, et al. 1997. Nuclear DNA diversity in worldwide distributed human populations. *Gene* 205:161–71
69. Zietkiewicz E, Votova V, Jarnik M, Koran-Laskowska M, Kidd KK, et al. 1998. Genetic structure of the ancestral population of modern humans. *J. Mol. Evol.* 47:146–55