# Genetics and Modern Human Origins

JOHN H. RELETHFORD

The past decade has brought considerable debate on the subject of modern human origins. The nature of the transition from *Homo erectus* to archaic *Homo sapiens* to modern *H. sapiens* has been examined primarily in terms of the relative contribution of archaic populations to later moderns, both within and among geographic regions. The recent African origin model proposes that modern humans appeared first in Africa between 100,000 and 200,000 years ago, and then spread through the rest of the Old World, replacing preexisting populations.[1–6] This model has been referred to by a variety of names, including "replacement," "Garden of Eden," "Noah's Ark," and "out of Africa." The recent African origin model contrasts with the multiregional model, which proposes a species-wide transition to modern humans throughout the Old World during the past million years or more.[7–10] Indeed, some proponents of the multiregional model advocate placing *Homo erectus* and all subsequent species of *Homo* in the evolutionary species *Homo sapiens.*[11] This contrasts with the view that there were multiple hominid species during the Middle Pleistocene. The debate continues.[12,13] Although the multiregional model is often portrayed as proposing a simultaneous transition to anatomically modern humans in different geographic regions, it explicitly allows for varying degrees of continuity across time and space.[10] This model, in the broad sense, does not rule out the possibility that modern human morphology appeared first in Africa and then spread through the rest of the Old World through gene flow. However, not all advocates of the multiregional model adhere to this specific subset of the general model.[9]

Comparison of the African and multiregional models is complicated by considering other, less extreme, hypotheses. Some versions of the recent African origin model imply a speciation event associated with the initial origin of modern humans. Another version, which suggests the possibility of some admixture between "moderns" leaving Africa and preexisting "archaics" elsewhere in the Old World,[14,15] is similar to some variants of the multiregional model, which also suggest that modern morphology appeared first in Africa, but involved admixture with other Old World populations.[16] The major difference between these views appears to be the extent of admixture, although the exact level is never specified. A further complication is the possibility that multiple dispersals from Africa produced a more complicated pattern of worldwide variation.[17]

John Relethford is Professor and Chair of the Department of Anthropology at the State University of New York College at Oneonta. He completed his Ph.D. in anthropology (1980) at SUNY-Albany. His primary research interests are anthropological genetics and human variation, particularly the use of quantitative traits for assessing population history. Much of his research has focused on the population history of Ireland. His recent research has been on the application of population genetics to questions of modern human origins.

relethjh@oneonta.edu.

Key words: modern human origins, population genetics, mitochondrial DNA, demographic history

## THE GENETIC EVIDENCE

In addition to differing interpretations of the fossil and archeological records,[10,14–16,18] considerable attention has been given to evidence from the study of contemporary human genetic variation. Patterns of within-group and among-group genetic variation have been examined in an attempt to reconstruct past evolutionary events or, more specifically, to determine the predicted patterns of contemporary human variation under different models of modern human origins and how they fit observed patterns of variation.

Can we trace backward from contemporary populations in order to reconstruct the past? In principle, it has been claimed that both the recent African origin and multiregional models make such distinctively different predictions that resolution of the debate could proceed directly from an examination of the patterns of within-group and among-group genetic variation in modern human populations. In reality, it is not that simple.

To date, four lines of genetic evidence have been used to support a recent African origin model. First, the relative degree of among-group variation in the human species today is relatively low. Second, many traits show higher levels of within-group variation in sub-Saharan African populations. Third, genetic distances show that sub-Saharan African populations are the most genetically divergent. Fourth, the estimated average effective population size of the human species over the past 100,000 to 200,000 years is very small. Each of these observations is considered here, in terms of its support for the recent African origin model and for alternative interpretations.

## THE RELATIVE HOMOGENEITY OF THE HUMAN SPECIES

A basic and highly important parameter of genetic variation is the proportion of genetic variation that exists among groups relative to total variation, also known as Wright's $F_{ST}$. Lewontin's pioneering work[19] showed that $F_{ST}$ is relatively low among the major geographic regions of modern humanity. His estimate of $F_{ST} = 0.06$ indicates that only six percent of the total variation in the human species is the result of among-group variation where the groups were defined in terms of traditional races.

Since his initial study, other researchers, using larger data sets and improved methods, have come to essentially the same general conclusion. Although $F_{ST}$ varies somewhat according to the number and composition of regions and the number of loci, estimates from classic genetic polymorphisms[5,20–23] have a limited range: $F_{ST} = 0.09–0.11$. Other estimates from genetic polymorphisms include an $F_{ST}$ 0.14 for 100 DNA polymorphisms[24] and 0.11 for dinucleotide repeat loci.[25] A somewhat higher estimate has been derived from polymorphic *Alu* insertions;[26] the average $F_{ST}$ over four loci is 0.17, although this level includes one locus (PV 92) with a much higher value (0.28). The average over the other three loci is 0.13. Estimates of $F_{ST}$ from craniometries, using an average heritability of 0.55, range from 0.11 to 0.14, depending on the number of groups in the analysis.[27] Overall, most of these estimates based on a wide range of traits cluster around an $F_{ST}$ of 0.10 to 0.11. Thus, roughly 10% of modern human genetic variation can be accounted for among groups, whereas 90% of the variation occurs within groups. This level of differentiation appears low relative to that of animals[28] such as kangaroo rats (0.67) and brown trout (0.29), but is similar to that of the house mouse (0.12). Mitochondrial DNA (mtDNA) shows even less among-group variation; one study[29] found an $F_{ST}$ of 0.063, which is roughly equivalent to an $F_{ST}$ of 0.017 for nuclear genes.[30] The lower $F_{ST}$ for mitochondrial DNA may reflect its higher rate of mutation. Earlier studies that found a higher $F_{ST}$ from mtDNA are now known to be flawed.[30] Comparisons of $F_{ST}$ values from mtDNA and other genetic traits can also be affected by sex differences in migration patterns.[31]

The estimates for $F_{ST}$ for modern humans are often interpreted as supporting a relatively recent common origin for human groups, which is consistent with the recent African origin model. Another interpretation is relatively high rates of migration between groups, which could be incorporated into either the multiregional or recent African origin model. However, in the latter case, the among-group migration takes place after the initial African origin. The key questions here relate to what we mean by "relatively recent common origin" or "relatively high migration rates." A new method[32] allows the estimation of migration from observed patterns of genetic similarity. Analyses of genetic variation using craniometric[33] and genetic marker[33] data show that observed $F_{ST}$ values among four major geographic regions (Africa, Europe, the Far East, and Australasia) require an average of 0.3 to 0.5 migrants per generation between each pair of regions over the past 80,000 years to reach 95% of equilibrium under this model.[32]

Is this a reasonable amount of average migration? If we consider long-distance movement per individual, this value might seem unrealistic. For example, we would need one migrant every 2.5 generations to obtain an average of 0.4 migrants per generation. If we take, for example, a distance of 10,000 km between Europe and East Asia, we would need to postulate that this hypothetical migrant traveled 4,000 km in a generation or, assuming 25 year generations, 160 km per year. Although this value is not inconceivable, it seems unreasonable as an average rate. However, if we consider the likelihood of migrants traveling in groups, the numbers become probable. For example, a group of 20 migrants every 50 generations would produce the same average rate of migration and would require movement over 200 km per generation, or 8 km per year. If we imagine a group of 40 migrants, the corresponding values are 100 km per generation and 4 km per year. These estimates, although crude, fall comfortably within the range of migration of contemporary foragers.[34] Of course, we most keep in mind the usual caveats regarding the use of contemporary data to infer ancient demographics, particularly when using data on short-range dispersal to infer long-range movements.

We must keep in mind that other patterns of migration and dispersal could give rise to the same general pattern. One possibility, raised by Lahr and Foley,[17] is a series of dispersals from Africa. It is also possible that the same net amount of migration could have been produced by a demic diffusion process[35] in which a network of small populations were connected through short-range gene flow. This possibility remains to be investigated. However, this type of scenario must factor in evidence for a small species population size and its likely effects on population distribution and migration.

In any case, the actual pattern of migration most likely varied across both time and space. I am not claiming that a gradualist model is useful as an exact representation of reality: what is mathematically convenient does not necessarily ensure the truth. The point, however, is that the levels of migration needed to produce our observed $F_{ST}$ values are not unrealistic. Our species' relatively low level of $F_{ST}$ can be replicated by a simple migration-drift model. The recent African origin model cannot be ruled out, but it is not necessary to explain this pattern of genetic variation. To complicate matters further, we must also consider the possibility that genetic data reflect both a recent African origin and migration-drift equilibrium.

## HIGHER GENETIC VARIATION IN SUB-SAHARAN AFRICA

A potentially stronger genetic claim in support of the recent African origin model comes from the observation that the highest levels of within-group variation typically are found in sub-Saharan African populations. This pattern has been observed in mitochondrial DNA,[1,36–38] polymorphic microsatellite data,[38] dinucleotide repeat loci,[25] short tandem repeat polymorphisms,[39] and craniometrics.[32] Higher diversity of classical genetic markers or nuclear DNA restriction fragment length polymorphisms (RFLPs)[39] has not been observed. One possible reason for these exceptions might be that many polymorphisms were originally selected because of their higher diversity in European groups, and that this biased the results.[5,39,40] Recent investigation suggests that such bias may be one factor, but not the only one.[39] Another possibility arises from differences in mutation rates under given patterns of demographic history. If our species has recently undergone rapid growth, then differences in within-group diversity across different traits might reflect differences between high-mutation systems (such as mtDNA) and low-mutation systems (such as classic genetic markers). This difference might also explain the low $F_{ST}$ obtained from mitochondrial DNA.[30]

The higher variation of many traits in African populations has frequently been interpreted as supporting the recent African origin model. The logic used here is that because many of these traits are considered to be neutral, the higher variation reflects the greater accumulation of mutations and hence, greater time. Stoneking and Cann[36] succinctly stated this view: "If one accepts that mtDNA mutations are largely neutral, then their occurrence and accumulation are mostly a function of time: the more variability a population possesses, the older it is" (p 22). Because genetic evidence supports higher variation in sub-Saharan African populations, that region of the world is the "oldest." Thus, the higher genetic diversity within sub-Saharan African populations is taken to support an African origin for modern humans.

The main problem with using levels of within-group variation to assess a population's relative age is that the logic used requires a critical assumption that usually is not stated. For this method to work, a newly formed daughter population (such as non-Africans) must experience a bottleneck to reduce its level of within-group variation. Without this "reset," a new daughter population would have the same within-group variation as the parent population. Hence, differences in the magnitude of variation would tell us nothing about the timing of population origins. At first glance, this assumption does not seem unreasonable—we often expect a daughter population to be smaller than its parent population. The problem here has to do with the fact that the level of within-group variation is not reset to zero, but is only slightly reduced from the original level. As a daughter population grows, within-group variation increases, thus erasing the initial loss of diversity caused by the bottleneck. The expected effect of size reduction due to founding is a function of the parental and daughter population sizes, the duration of the bottleneck, the population size after the bottleneck, and the length of the recovery period. Rogers and Jorde[6] have shown that such bottlenecks associated with the formation of a daughter population would have to have been longer and more severe than is realistic in order to produce the expected relationship between within-group variation and the "age" of a population. Although bottlenecks are likely to have been an important part of our species' evolutionary history, bottlenecks associated with initial founding are not likely to have produced a simple relationship whereby we can estimate a population's age from its level of within-group variation. On the other hand, differences in the timing of population growth, rather than population origins, could be reflected in levels of within-group variation.

**Although bottlenecks are likely to have been an important part of our species' evolutionary history, bottlenecks associated with initial founding are not likely to have produced a simple relationship whereby we can estimate a population's age from its level of within-group variation. On the other hand, differences in the timing of population growth, rather than population origins, could be reflected in levels of within-group variation.**

Because most of the literature focuses on the assumed relationship between diversity and age, alternative explanations for the distinctiveness of sub-Saharan Africa have been ignored. Higher diversity in Africa could also have resulted from higher levels of gene flow into Africa, a larger African population, or both. However, higher rates of gene flow are not likely because they would also be expected to reduce the genetic distances between Africa and other regions, which is the precise opposite of what has been observed (see below). The other possibility, that the long-term average population size of sub-Saharan Africa was larger than that in any other region, has gained recent support. This is an attractive alternative because within-group variation is proportional to effective population size. If the African population were larger, it would be expected to include greater genetic diversity, even in the case of equal population ages.

Support for this hypothesis has come from recent analysis of worldwide craniometric data. Several years ago, my colleague John Blangero and I developed a method[41] based on earlier work of Harpending and Ward[42] that allows detection of the effects of differential long-range gene flow on quantitative traits. This method essentially uses the genetic distances of each population from the centroid (the average of all groups) and the within-group variance to predict its expected average within-group variance. Henry Harpending and I then applied this method to a global analysis[32] in which there should be a perfect fit between the observed and expected within-group variances. Figure 1a illustrates the fit of this model based on 57 craniometric traits from the four major Old World regions of interest. The theoretical regression line represents the expected relationship between distance to the centroid and within-group variation. (This line is based on the theoretical model; it is not derived from linear regression of the data points.) We expected that all observed points should fall on the expected line; in reality, Africa shows significantly greater variation than expected and Europe shows significantly less variation than expected.

There are only three possible reasons for the observed deviations from the expected model. First, mutation rates were not equal across the Old

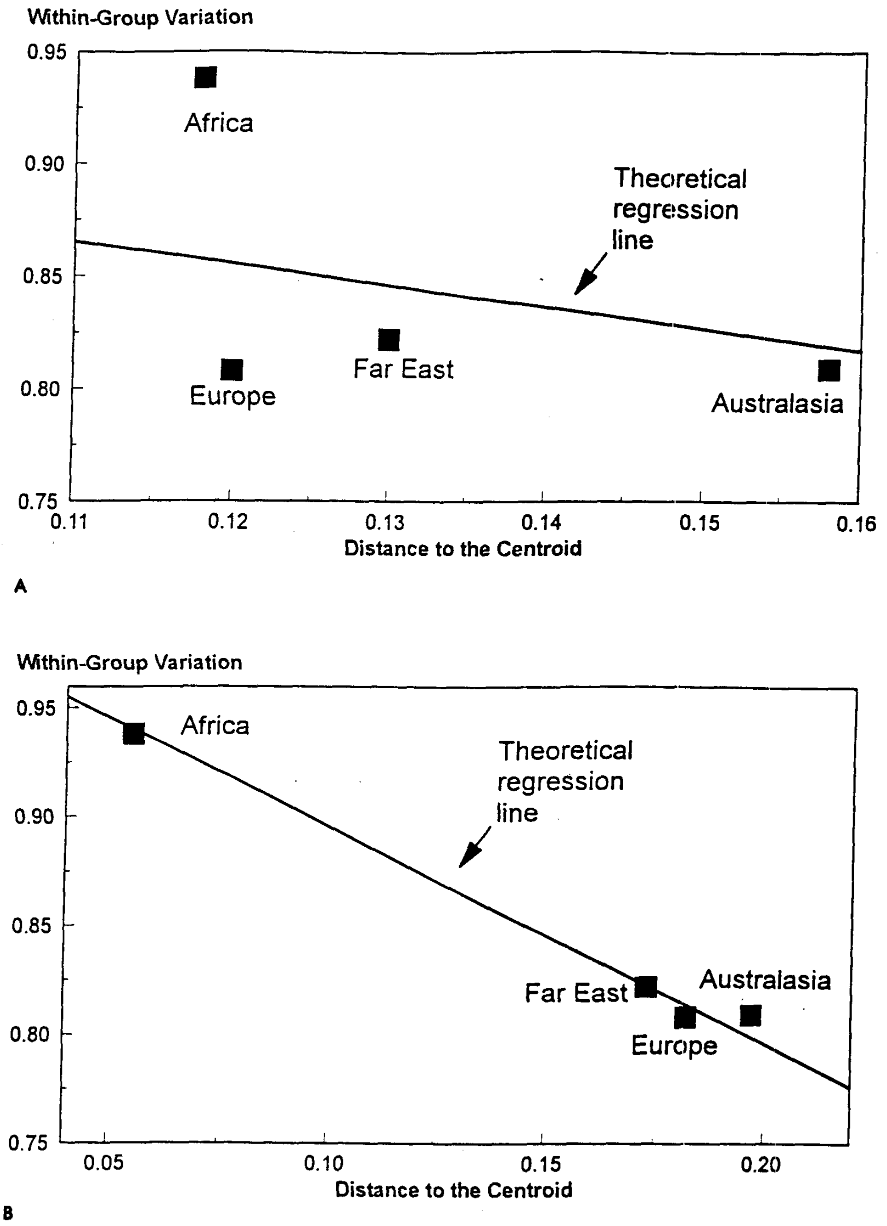

Figure 1. Relationship between average within-group phenotypic variation and genetic distance to the centroid of all groups based on 57 craniometric traits in four major geographic regions and assuming equal effective population sizes.[32] The distance to the centroid is the genetic distance of a population to the mean allele frequencies over all populations (the genetic "center" of a set of populations). The theoretical regression line, derived from the underlying model, is not a linear regression of the data points. (A) Under the assumption of equal effective population sizes, this model shows that Africa is more variable than expected. (B) Setting the effective size of Africa at three times the size of any other region.[32] This model, based on a larger long-term effective population size in Africa, provides an almost perfect fit.

World. This seems highly unlikely and, in any event, such differences would have to have been many orders of magnitude to generate the deviations. Second, different regions would have to have experienced greater gene flow from outside the areas of study. However, because this analysis focused on the global variation of modern humans, the only possibility would have been gene flow from beyond this planet—shades of von Däniken! This possibility can obviously be eliminated. The third possibility is that the assumption of equal effective population sizes used to generate Figure 1A is invalid. As such, the pattern shown in Figure 1A suggests that the African population was larger, Europe smaller, or both. We then looked for the combination of relative population sizes that would minimize the deviations from the expected line. The best fit was found when the population of sub-Saharan Africa was set to three times the size of that in any other region.[32] As shown in Figure 1B, these values produce an almost perfect fit to the expected model. There is no significant difference between observed and expected variances.

Similar results can be obtained using a different approach for mtDNA within-group sequence divergence data. At equilibrium, sequence divergence is directly proportional to female effective population size. The ratio of African to non-African sequence divergence allows an estimate of the relative size of the population in Africa to that in any non-African region. The heterozygosity at the nucleotide level for African mtDNA[38] is 0.0232; the average for non-African populations is 0.0130. The ratio (0.0232:0.0130 = 1.8) provides further support for a larger effective population in sub-Saharan Africa in the past. If the populations were not in equilibrium, the ratio would be lower. This method is not as precise as the Relethford-Harpending method because it does not use a model that incorporates migration among groups, but it is consistent with the finding of a larger average African population size.

The finding of a larger effective population size for Africa does not mean that the African population was always or even initially larger, although these are possibilities. A larger average size could also be caused by differences in the timing of population growth: if Africa grew more quickly than other regions, it would still have a larger average effective size even if the initial and ending population sizes were the same in all regions.

## GREATER GENETIC DIVERGENCE OF SUB-SAHARAN AFRICA

The analysis of among-group variation is another area of study in the application of genetics to the debate about modern human origins. The genetic divergence of different regions or individuals typically has focused on genetic distances or related measures such as among-group mtDNA sequence divergence. Studies using analysis of classic genetic markers,[5,33] mitochondrial DNA,[1,37] polymorphic microsatellites,[38] dinucleotide repeat loci,[25] and craniometrics[32,33] have found sub-Saharan Africa to be the most genetically distinct geographic region.

Studies of genetic distance must consider several sources of among-group differences, including differences in genetic drift, gene flow, and common ancestry. Differences in population size and genetic drift are not typically considered. Recent work shows how this can obscure the underlying pattern of gene flow and common ancestry.[32,33,43] Because of new work suggesting that the average size of the African population over time was larger than that in other regions, it is critical to take differential drift into account in studies of global human variation. A new method for doing so has been developed and applied to allele frequency and craniometric data from different Old World regions. This method causes distances based on genetic markers and craniometrics to become more similar and the distinctiveness of African populations to become more marked.[33] One example, based on genetic marker data, is shown in the dendrogram in Figure 2.

At first glance, Figure 2 seems to support the hypothesis of a recent African origin. A common interpretation of this figure would be an initial split between African and non-African lines, a subsequent split of the European and non-European lines, and finally a split between the Far East and Australasia. Although different studies show variation in the branching order of the non-African populations (thus making premature any attempts to date such splits), the basic division between African and non-African populations is consistent across different data and methods.

The main problem with such an interpretation is that the genetic distances are assumed to reflect primarily a branching process. In other words, the vertical dimension of Figure 2 is taken to reflect time since separation, when in reality genetic distances reflect overall dissimilarity resulting from a variety of causes. A history of bifurcational splits is one possible explanation for Figure 2, but not the only one. The fact that dendrograms look like trees does not mean that they should necessarily be taken as trees. The same patterns could be produced by an appropriate matrix of migration rates among populations. Felsenstein[44] noted a basic problem in interpreting genetic distances—the same distances can be produced by a branching model or by a migration matrix model. A recent simulation study further supports this conclusion.[45] Templeton[46] reached a similar conclusion based on cladistic analysis of mtDNA types. Figure 3 provides an example of this problem. Here, two completely different models produce exactly the same genetic distances.

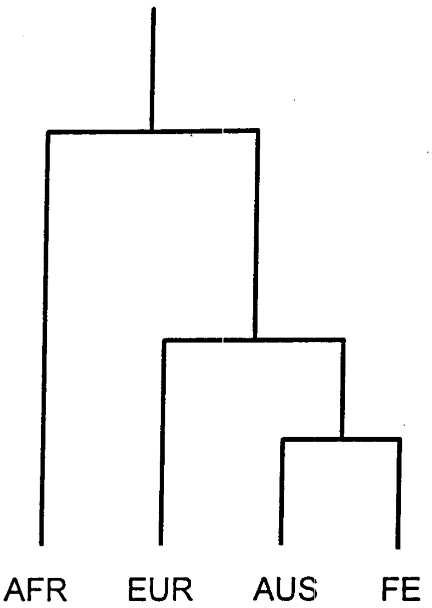

Figure 2. Dendrogram of the genetic distances between four major geographic regions based on 93 allele frequencies for 37 genetic marker loci.[33] The distances are derived from genetic relationship matrices that have been scaled by effective population size in order to remove problems of interpretation due to differences in genetic drift.[32,33,43] Effective size for Africa has been set to three times that of any other region. The difference between African and non-African populations is even more apparent after application of this scaling method. Additional dendrograms and description of the analysis are presented elsewhere.[33] AFR = Africa, EUR = Europe, AUS = Australasia, FE = Far East.

How, then, can we distinguish between these models when looking at the genetic distances among modern human populations? More specifically, does the genetic distinctiveness of Africa reflect an early split of African and non-African populations, smaller rates of gene flow with Africa, or both? Henry Harpending and I[32] have developed a new method that allows estimates of the number of migrants necessary to reproduce a given pattern of genetic relationships among populations. Our method does not prove that a migration model is the best fit, but instead provides an idea of the degree of migration needed to reconstruct a set of genetic distances.

Figure 4 shows estimates of the number of migrants per generation between four major geographic regions; these estimates are derived from craniometrics, genetic markers, and the two data sets pooled.[32,33] The 95% confidence intervals are also shown. In spite of some variation across data sets and among populations, most of the estimates have confidence intervals that fall in the range of 0.2 to 0.6 migrants per generation, clustering around 0.3 to 0.4 migrants per generation. Apart from the differences between the craniometric and genetic marker estimates of the number of migrants between Australasia and the Far East, all of the remaining estimates have confidence intervals that include an average value of 0.35 migrants per generation. It must be kept in mind that these are estimates of the per-generation average number of migrants, and as such represent "effective migrant numbers." An actual pattern of migration could be much more complex, fluctuating over time and space. Nonetheless, these estimates show that a migration-drift model can explain observed genetic distances among human populations on different continents.

There does not seem to be any major regional variation in the number of migrants per generation. If the number of migrants into or out of Africa was about the same as for other regions, how can a migration model explain the genetic distinctiveness of Africa? The important point here is that Figure 4 shows the number of migrants. Because the rate of migration is proportional to this number divided by effective population size, migration rates for larger populations will necessarily be smaller when the migrant numbers are symmetric. Our evidence for a larger African population fits in divergence

## An Example of How A Branching Model and a Migration Model Can Produce the Same Genetic Distances

Model 1: Completely separate populations
(A, B, C, D) of equal size diverging at
different times

Dates of divergence (generations):

| | A | B | C |
|---|---|---|---|
| B | 1000 | | |
| C | 1000 | 500 | |
| D | 1000 | 500 | 200 |

**Model 2: All populations are of equal size**
**begin at the same point in time, and**
**are connected by gene flow**

**Migration matrix (entries indicate the probability of being in a column and originating in a row):**

| | A | B | C | D |
|---|---|---|---|---|
| A | 0.99954 | 0.00008 | 0.00019 | 0.00019 |
| B | 0.00008 | 0.99924 | 0.00034 | 0.00034 |
| C | 0.00019 | 0.00034 | 0.99869 | 0.00079 |
| D | 0.00019 | 0.00034 | 0.00079 | 0.99869 |

Dendrogram of genetic distances:

Dendrogram of genetic distances:

Figure 3. An example of how two different processes, a branching model and a migration model, can produce the same genetic distances. Model 1 is based on a highly simplistic model of drift in populations of equal size that have diverged at different times.[56] Genetic distance is obtained from the expected kinship after *t* generations of drift. Although overly simplistic, this model provides the type of distances expected given differential times of origin. Model 2 is based on a migration model where all populations are of equal size and have exchanged migrants according to the stated migration matrix. The genetic distances were derived from the genetic relationship matrix expected at equilibrium using the Rogers-Harpending migration matrix method.[57] Note that the two models produce the same dendrograms. Therefore, if we have only a dendrogram, we have no way of distinguishing which type of underlying model produced the distances. This type of problem is discussed in further detail by Felsenstein.[44]

with this observation. If the long-term average population of Africa was three times the size of that in any other region, then the migration rate into Africa would have been one-third the rate into any other region, given equal numbers of migrants. Note that this model requires symmetry in migrant numbers, which does not seem unreasonable for this level of migration when comparing continental regions.

In sum, the genetic distinctiveness of Africa could have resulted from small numbers of migrants between each region combined with a larger

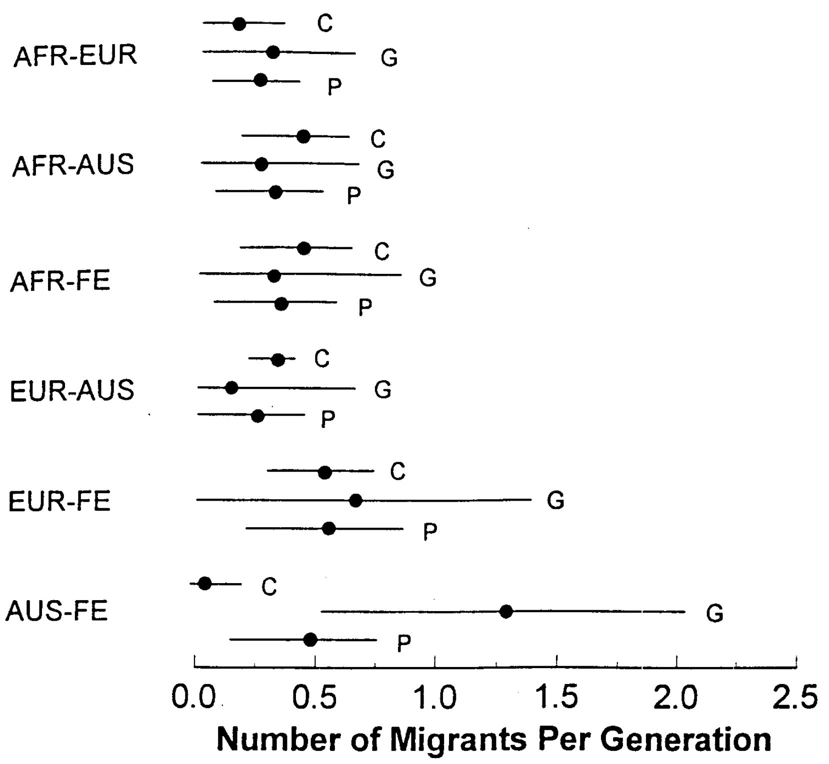

Figure 4. Estimates of the number of migrants per generation needed to replicate the observed genetic distances between populations based on craniometric measures (C), genetics markers (G), and pooled traits (P). The dots indicate the estimates; the horizontal lines indicate the 95% confidence intervals obtained by jackknifing. Each distance is indicated by the symbols for both members of the pair; for example, AFR-EUR indicates the distance between Africa and Europe. AFR = Africa, EUR = Europe, AUS = Australasia, FE = Far East. Details of these analyses are presented elsewhere.[32,33]

African population. The fact that this model fits the observed data does not, of course, make it true. However, the fact that a reasonable pattern of migration and population sizes could just as easily explain the larger genetic distances to Africa shows that these distances do not automatically support the recent African origin model. The common finding of African uniqueness might reflect demographic patterns, not a branching history. The reverse, however, is also true. Our results do not prove a migration-drift model, but show only that it is possible. Traditional genetic distance analysis may continue to be plagued by this indeterminacy. Other approaches are needed.

In addition to traditional genetic distance analyses, some recent studies have employed a different approach to the problem of modern human origins. Comparison of human RFLPs[40] and polymorphic *Alu* insertion data[26] with corresponding data for nonhuman primates has allowed estimation of ancestral allele frequencies. Comparison of allele frequencies from contemporary human populations with these estimates supports a recent African origin because sub-Saharan African samples are genetically closest to the reconstructed ancestral gene pool. These results are suggestive but not conclusive: it is not clear whether such patterns can be replicated by a migration-drift model. For example, if non-African populations were smaller and drifted more, they might be expected to lie further from the reconstructed ancestral gene pool. This is an area needing further investigation, which should not be difficult because most of the original theory regarding migration-drift balance was initially worked out for comparison of contemporary and ancestral allele frequencies.

The analysis of mtDNA variation among human groups has also looked at the distribution of various mtDNA types. In addition to providing evidence of a basic African versus non-African difference, such studies also show that mtDNA variation outside of Africa is a subset of the variation existing within Africa. This finding also supports a recent African origin model, but it is not clear whether the same pattern could be replicated by a suitable migration-drift model. This question needs to be addressed in future work.

## SMALL ESTIMATED SPECIES POPULATION SIZE

Few studies have aroused so much interest and debate as Cann, Stoneking, and Wilson's analysis[1] of human mitochondrial DNA and their conclusion that all living humans can trace their mitochondrial ancestry back to a single female who lived in Africa roughly 200,000 years ago. Their position was strengthened with further analysis, particularly with larger samples that also included living Africans[37] rather than African-Americans, where European admixture could potentially confound results. The later finding[47,48] of flaws in the computer methodology used to generate trees of descent caused many to question the age and significance of "mitochondrial Eve."

Other methods, however, have produced similar dates. Stoneking and colleagues[49] used archeological data on the colonization of Papua, New Guinea to calibrate mtDNA trees. This study led them to estimate an age of coalescence at 135,000 years ago, with a range of 60,000 to 400,000 years. More recently, a comparison of human mtDNA sequences with those from the great apes has produced a similar date.[50] Here, the tree was calibrated by taking 13 million years as the divergence of African hominoid and orangutan lineages, yielding an estimate of 143,000 years for the common human mtDNA ancestor.

What is the significance of these dates? It is now recognized that the age of "Eve" tells us nothing about population origins, but rather that humans were few in number at that time. It is estimated that the effective total size of the human population at this time was roughly 1,000 to 10,000 females, or a total effective size of 2,000 to 20,000 people.[3] These estimates are based on the mathematical relation-

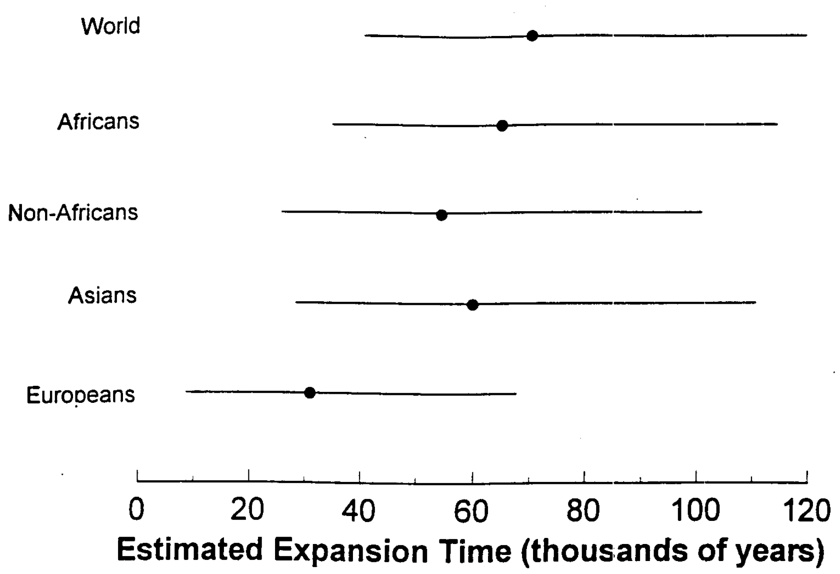

Figure 5. Estimates of expansion times of world populations derived from mtDNA mismatch analysis. The dots indicate the estimates and the horizontal lines indicate approximate 95% confidence intervals. Further details are presented elsewhere.[54]

ship between genetic diversity, age of coalescence, and population size. They are considerably less than estimates based on ecological and demographic models,[3,51] which suggest a total Old World population of *Homo erectus* of between 125,000 and 500,000 years. Thus, the genetic estimates provide indirect support for the recent African origin model. It seems unlikely that so small a population could have been spread over the Old World for more than a million years and still be connected by the gene flow required under a multiregional model. Such a small population implies that the entire species was in a limited geographic area, such as Africa.

It is important to keep in mind that such estimates refer to an average long-term species population size, which in turn could result from numerous demographic scenarios such as constant size, early small size and later growth, or a bottleneck. It is also important to note that for population genetic purposes (e.g., the expectation of genetic drift), the average population size over time will be much closer to the minimum population size than to the arithmetic average.[52]

## A PLEISTOCENE POPULATION EXPLOSION

The issue of population size is further complicated by genetic evidence of a major demographic shift during the past 100,000 years. This evidence is derived from a new method[3,53] of mtDNA analysis based on mismatch distributions, which are simply histograms of the number of differences between pairs in a sample of mtDNA sequences drawn from the same population. For example, if the mtDNA of a pair of individuals in a population is compared and found to differ by two sequences, then this pair would be plotted as 2 on a histogram that ranged from 0 to the maximum number of mtDNA sequence differences.

Mismatch distributions from human populations typically show a single mode greater than zero. This type of distribution contrasts with the theoretical distribution of an equilibrium population with constant population size, which has a mode at 0 sequence differences.[53] However, the observed mismatch distributions do fit a model of population expansion in which the peak of the observed distribution reflects the time of the expansion. As a consequence, it is possible to estimate the date of the expansion and its overall magnitude, thus providing insight into ancient population dynamics. To date, 23 of 25 samples from human populations fit the expansion model.[54]

These analyses have shown that the human population underwent a dramatic expansion from an initial size of between 1,000 and 10,000. Based on worldwide data, the estimated date of this expansion is roughly 60,000 years ago, with a likely range between 33,000 and 150,000 years ago.[6] This range incorporates a standard error on the estimates, as well as a range of estimates for the nucleotide divergence rate. This method has also been applied to samples formed by combining individuals in major geographic regions, resulting in estimates averaging about 60,000 years ago.[54] The expansion times for Africans is slightly greater than for non-Africans,[3,54] although the confidence intervals around these estimates overlap (see Fig. 5). Expansion times have also been estimated using samples from local populations—most of these estimates range from 30,000 to 65,000 years ago, with an average of roughly 40,000 years.[54] Although estimates differ with sample composition and the types of mtDNA data used, an average expansion of 50,000 years is consistent with all of these analyses. This average date is interesting because it corresponds in timing with the "creative explosion" proposed by some archeologists,[18] thus suggesting that cultural and demographic events were linked.

In order to relate these findings to models of modern human origins, it is necessary to consider the likely population dynamics underlying this evidence of dramatic population growth. The mismatch distributions could result from two somewhat different demographic hypotheses. First, the preexpansion species population size was always small, perhaps reflecting an initial origin at some point not much earlier than the expansion. Second, the human population might have been much larger, but experienced a bottleneck at some point prior to the expansion. The mitochondrial DNA evidence can take us back only as far as a small initial population; it does not allow us to reconstruct the demographic events leading to that small size. Any demographic hypothesis must also deal with the observation of a larger average effective population size in Africa than elsewhere. A larger African population implies that the African population was initially larger, that Africa expanded first, and/or that Africa suffered less of a bottleneck. There obviously are many different models that could accommodate the small species size, the population expansion, and a larger African population. Future studies on the genetics of modern human origins should focus on further definition and testing of these hypotheses, rather than emphasizing phylogenetic debate.

## IMPLICATIONS FOR MODELS OF MODERN HUMAN ORIGINS

The ability of data on modern human genetic variation to resolve competing models of modern human origins is not as clear-cut as it was thought to be just a few years ago. It is also apparent that extreme forms of the multiregional and recent African origin models are not as useful as once thought, except as starting points for the further development and testing of hypotheses. The oft-cited statements that Africa's high within-group variation and greater genetic distances from other regions support the recent African origin model are premature. Each of these findings can be explained in terms of population dynamics that are consistent with both the multiregional and recent African origin models.

It is clear that the genetic evidence tells us much less directly about phylogenetic relationships than was once thought. Instead, the genetic evidence tells us things about ancient population dynamics. The major findings to date are a major expansion in population size during the late Pleistocene, a very small preexpansion size, and a larger African population. These observations allow us to examine origin models indirectly. In other words, these findings, rather than providing a direct test of phylogenetic relationships, allow us to estimate certain parameters of population dynamics and use them to assess the relative likelihood of the origin models.

What are the implications of these findings? There are two possibilities for the small estimated species population: first, that the human species originated in a small local region, as postulated by the recent African origin model; second, that the species underwent a bottleneck at some point in the last 100,000 years or so. These two hypotheses are not mutually exclusive—a restricted origin could have been followed by a species-wide bottleneck.

The small estimated preexpansion species size argues against the multiregional model. Based on estimates to date, it seems unlikely that such a small population could be spread over three continents and remain connected by gene flow for a long time. The evidence of our species' dramatic expansion, combined with other information, does allow us to rule out the classic replacement model, which suggests that modern humans arose in Africa and replaced other human populations while expanding. Mismatch distributions based on comparisons among populations (intermatch distributions) suggest that population divergence occurred before population expansion. Preliminary estimates further suggest that regional divergence started roughly 100,000 years ago. Based on these results, Harpending and colleagues[3] distinguished between a "strong Garden of Eden" model involving simultaneous divergence and expansion and a "weak Garden of Eden" model in which these two processes are uncoupled. Under the weak Garden of Eden model, modern humans arose in Africa and, starting about 100,000 years ago, separated into small regional populations weakly connected by gene flow. Several tens of thousands of years later, these groups experienced population growth or recovery from a bottleneck. Thus, population divergence and expansion occurred at different times.

**Contrary to popular belief, the genetic data give us little direct information about phylogeny.**

The other possible explanation for a low estimated species size is that there was a bottleneck that reduced the size of the human population at some time prior to the population expansion. In other words, the prebottleneck population could have been large enough to accommodate the multiregional model. Then a species-wide bottleneck could have occurred, followed by the population explosion of the Late Pleistocene. In that case, the mtDNA mismatch distributions would show a small population after the bottleneck and the population expansion, but could not tell us anything about the size of the prebottleneck population. None of this means that this scenario did occur, but rather that it could have occurred.

To consider further possible implications, we must examine the likelihood of a species-wide bottleneck. One possibility is that there was a worldwide shift in climate. If so, then the mtDNA of nonhuman species should show a similar effect. There is some supporting evidence of this from chimpanzee mtDNA, which shows a similar pattern of population expansion from a small initial size at roughly the same time as that indicated for humans.[6] Similar analysis of many other nonhuman species is an obvious next step. Further analysis of the timing of a possible species-wide bottleneck is another possible avenue of investigation. One possibility that has been suggested[6,33] is the eruption of the super-volcano Toba in Sumatra, which took place 73,500 years ago. Estimates of its likely impact on climate suggest that nontropical climates would have been affected more severely than tropical climates.[55]

Any of these demographic models must also deal with the finding of a larger African average effective population size. If the species population was initially small because of a geographically limited origin in Africa, then a larger African population could have resulted from initially smaller non-African daughter populations or later population expansion outside of Africa or, perhaps, both. These possibilities are compatible with the weak Garden of Eden replacement model. The mtDNA mismatch analyses suggest an earlier expansion in Africa, but they are not conclusive due to the large confidence intervals surrounding these estimates (Fig. 5). Recent work has suggested that comparison of high-mutation and low-mutation genetic systems supports an earlier expansion in Africa.[30]

If the species population was initially larger and the mtDNA analyses are picking up recovery from a bottleneck, then a larger African population could have resulted if the bottleneck had a smaller initial impact in Africa or if recovery from it occurred earlier in Africa than elsewhere. If the bottleneck occurred and was produced by a climatic catastrophe such as the Toba eruption, we would expect Africa to be affected less, to recover faster, and hence have a larger population. Although this is consistent with what we know about ancient population dynamics, it by no means proves that things happened this way. Other demographic hypotheses must be developed and tested.

## CONCLUSIONS

Genetic data have been increasingly used in the continuing debate regarding phylogenetic interpretations of modern human origins—multiregional evolution versus recent African origin. Recent research, reviewed here, shows that many of the findings once thought to support a branching model of human origins can also be explained by a migration-drift model. Neither higher African diversity or greater African divergence necessarily requires a branching tree based on the recent African origin model. These and other findings can instead by explained by a larger average long-term African population.

Contrary to popular belief, the genetic data give us little direct information about phylogeny. The success of genetic data in issues such as the ape-human split does not necessarily translate when we look at variation among groups of modern humans, all of whom belong to the same species. The difficulties in applying species-based phylogenetic models to these data are well known, but are not often heeded. What, then, do the genetic data tell us? Quite simply, they reflect the demographic history of our species.

## ACKNOWLEDGMENTS

Some of the research discussed in this paper was supported in part by a SUNY Graduate Research Initiative Grant. I thank Dr. W. W. Howells for providing generous access to his craniometric data. I thank Henry Harpending, Alan Rogers, and Milford Wolpoff for many stimulating conversations on this topic. Thanks also to Ken Korey for sharing a copy of his AAPA paper with me, and to Milford Wolpoff for sharing preprinted material from his forthcoming book.

## REFERENCES

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

**2** Stringer CB, Andrews P (1988) Genetic and fossil evidence for the origin of modern humans. Science *239*:1263–1268.

**3** Harpending HC, Sherry ST, Rogers AR, Stoneking M (1993) Genetic structure of ancient human populations. Curr Anthropol *34*:483–496.

**4** Stoneking M (1993) DNA and recent human evolution. Evol Anthropol *2*:60–73.

**5** Cavalli-Sforza LL, Piazza A, Menozzi P (1994) *History and Geography of Human Genes*. Princeton: Princeton University Press.

**6** Rogers AR, Jorde LB (1995) Genetic evidence on modern human origins. Hum Biol *67*:1–36.

**7** Wolpoff MH (1989) Multiregional evolution: The fossil alternative to Eden. In Mellars P, Stringer C (eds), *The Human Revolution: Behavioural and Biological Perspectives on the Origins of Modern Humans*. Princeton: Princeton University Press, pp 62–108.

**8** Thorne AG, Wolpoff MH (1992) The multiregional evolution of humans. Sci Am *266*:76–83.

**9** Wolpoff MH, Thorne AG, Smith FH, Frayer DW, Pope GG (1994) Multiregional evolution: A world-wide source for modern human populations. In Nitecki MH, Nitecki DV (eds), *Origins of Anatomically Modern Humans*. New York: Plenum Press, pp 175–199.

**10** Wolpoff MH (1995) *Paleoanthropology*. New York: McGraw Hill, in press.

**11** Wolpoff MH, Thorne AG, Jelinek J, Yinyun Z (1994) The case for sinking *Homo erectus*. 100 years of *Pithecanthropus* is enough! Cour Forsch-Inst Senckenberg *171*:341–361.

**12** Tattersall I (1994) What do we mean by human—and why does it matter? Evol Anthropol *3*:114–116.

**13** Wolpoff M (1994) What do we mean by human—and why does it matter? Evol Anthropol *3*:116–117.

**14** Bräuer G (1992) Africa's place in the evolution of *Homo sapiens*. In Bräuer G, Smith FH (eds), *Continuity or Replacement: Controversies in Homo sapiens Evolution*. Rotterdam: A. A. Balkema, pp 83–98.

**15** Stringer CB (1994) Out of Africa—A personal history. In Nitecki MH, Nitecki DV (eds), *Origins of Anatomically Modern Humans*. New York: Plenum Press, pp 149–172.

**16** Smith FH (1994) Samples, species, and speculations in the study of modern human origins. In Nitecki MH, Nitecki DV (eds), *Origins of Anatomically Modern Humans*. New York: Plenum Press, pp 227–249.

**17** Lahr MM, Foley R (1994) Multiple dispersals and modern human origins. Evol Anthropol *3*:48–60.

**18** Klein RG (1992) The archaeology of modern human origins. Evol Anthropol *1*:5–14.

**19** Lewontin RC (1974) The apportionment of human diversity. Evol Biol *6*:381–398.

**20** Latter BDH (1980) Genetic differences within and between populations of the major human groups. Am Natural *116*:220–237.

**21** Nei M, Roychoudhury AK (1982) Genetic relationship and evolution of human races. Evol Biol *14*:1–59.

**22** Ryman N, Chakraborty R, Nei M (1983) Differences in the relative distribution of human gene diversity between electrophoretic and red and white cell antigen loci. Hum Hered *33*:93–102.

**23** Livshits G, Nei M (1990) Relationships between intrapopulational and interpopulational genetic diversity in man. Ann Hum Biol *6*:501–513.

**24** Bowcock AM, Kidd JR, Mountain JL, Hebert JM, Carotenuto L, Kidd KK, Cavalli-Sforza LL (1991) Drift, admixture, and selection in human evolution: A study with DNA polymorphisms. Proc Natl Acad Sci USA *88*:839–843.

**25** Deka R, Shriver MD, Yu LM, DeCroo S, Hundrieser J, Bunker CH, Ferrell RE, Chakraborty R (1995) Population genetics of dinucleotide (dC-dA)$_n$•(dG-dT)$_n$ polymorphisms in world populations. Am J Hum Genet *56*:461–474.

**26** Batzer MA, Stoneking M, Algeria-Hartman M, Bazan H, Kass DH, Shaikh TH, Novick GE, Ioannou PA, Scheer WD, Herrera RJ, Deininger PL (1994) African origin of human-specific polymorphic *Alu* insertions. Proc Natl Acad Sci USA *91*:12288–12292.

**27** Relethford JH (1994) Craniometric variation in modern human populations. Am J Phys Anthropol 95:53–62.

**28** Nei M (1987) *Molecular Evolutionary Genetics*. New York: Columbia University Press.

**29** Whittam TS, Clark AG, Stoneking M, Cann RL (1986) Allelic variation in human mitochondrial genes based on patterns of restriction site polymorphism. Proc Natl Acad Sci USA *83*:9611–9615.

**30** Harpending HC, Relethford JH, Sherry ST (n.d.) Methods and models for understanding human diversity. In Boyce AJ, Maseic-Taylor CGN (eds), *Molecular Biology and Human Diversity*. Cambridge: Cambridge University Press, in press.

**31** Melnick DJ, Hoelzer GA (1993) What is mtDNA good for in the study of primate evolution? Evol Anthropol *2*:2–10.

**32** Relethford JH, Harpending HC (1994) Craniometric variation, genetic theory, and modern human origins. Am J Phys Anthropol *95*:249–270.

**33** Relethford JH, Harpending HC (1995) Ancient differences in population size can mimic a recent African origin of modern humans. Curr Anthropol *36*:667–674.

**34** Kingdon J (1993) *Self-Made Man: Human Evolution from Eden to Extinction?* New York: John Wiley & Sons.

**35** Weiss KM (1988) In search of times past: Gene flow and invasion in the generation of human diversity. In Mascie-Taylor CGN, Lasker GW (eds) *Biological Aspects of Human Migration*. Cambridge: Cambridge University Press, pp 130–166.

**36** Stoneking M, Cann RL (1989) African origin of human mitochondrial DNA. In Mellars

P, Stringer C (eds), *The Human Revolution: Behavioural and Biological Perspectives on the Origins of Modern Humans*. Princeton: Princeton University Press, pp 17–30.

**37** Vigilant L, Stoneking M, Harpending H, Hawkes K, Wilson AC (1991) African populations and the evolution of human mitochondrial DNA. Science *253*:1503–1507.

**38** Bowcock AM, Ruiz-Linares A, Tomfohrde J, Minch E, Kidd JR, Cavalli-Sforza LL (1994) High resolution of human evolutionary trees with polymorphic microsatellites. Nature *368*:455–457.

**39** Rogers AR, Jorde LB (n.d.) Ascertainment bias in estimates of average heterozygosity. Am J Hum Genet, submitted.

**40** Mountain JL, Cavalli-Sforza LL (1994) Inference of human evolution through cladistic analysis of nuclear DNA restriction polymorphisms. Proc Natl Acad Sci USA *91*:6515–6519.

**41** Relethford JH, Blangero J (1990) Detection of differential gene flow from patterns of quantitative variation. Hum Biol *62*:5–25.

**42** Harpending HC, Ward RH (1982) Chemical systematic and human populations. In Nitecki M (ed), *Biochemical Aspects of Evolutionary Biology*. Chicago: University of Chicago Press, pp 213–256.

**43** Relethford JH (n.d.) Genetic drift can obscure population history: Problem and solution. Hum Biol, in press.

**44** Felsenstein J (1982) How can we infer geography and history from gene frequencies? J Theor Biol *96*:9–20.

**45** Korey K (1995) On pattern and origin of human biological diversity. Presented at the 64th Annual Meeting of the American Association of Physical Anthropologists, Oakland, California.

**46** Templeton AR (1993) The "Eve" hypotheses: A genetic critique and reanalysis. Am Anthropol *95*:51–72.

**47** Hedges SB, Kumar S, Tamura K, Stoneking M (1992) Human origins and analysis of mitochondrial DNA sequences. Science *255*:737–739.

**48** Templeton AR (1992) Human origins and analysis of mitochondrial DNA sequences. Science *255*:737.

**49** Stoneking M, Sherry ST, Redd AJ, Vigilant L (1992) New approaches to dating suggest a recent age for the human mtDNA ancestor. Philos Trans R Soc London B *337*:167–175.

**50** Horai S, Hayasaka K, Kondo R, Tsugane K, Takahata N (1995) Recent African origin of modern humans revealed by complete sequences of hominoid mitochondrial DNAs. Proc Natl Acad Sci USA *92*:532–536.

**51** Weiss KM (1984) On the number of members of the genus *Homo* who have ever lived, and some evolutionary implications. Hum Biol *56*:637–649.

**52** Wright S (1969) *Evolution and the Genetics of Populations, Volume 2: The Theory of Gene Frequencies*. Chicago: University of Chicago Press.

**53** Rogers AR, Harpending HC (1992) Population growth makes waves in the distribution of pairwise genetic differences. Mol Evol Biol 9:552–569.

**54** Sherry ST, Rogers AR, Harpending H, Soodyall H, Jenkins T, Stoneking M (1994) Mismatch distributions of mtDNA reveal recent human population expansions. Hum Biol *66*:761–775

**55** Rampino MR, Self S (1992) Volcanic winter and accelerated glaciation following the Toba super-eruption. Nature *359*:50–52.

**56** Relethford JH (1991) Genetic drift and anthropometric variation in Ireland. Hum Biol *63*:155–165.

**57** Rogers AR, Harpending HC (1986) Migration and genetic drift in human populations. Evolution *40*:1312–1327.

## Books Received

- Morgan E (1994) *The Scars of Evolution. What Our Bodies Tell Us About Human Origins*. New York: Oxford University Press. xii + 196pp. ISBN 0-19-280021-3. $12.95 (paper).
- Nesse RM, Williams GC (1995) *Why We Get Sick. The New Science of Darwinian Medicine*. New York: Times Books. xi + 291pp. ISBN 0-8129-2224-7. $24.00 (cloth).
- Hecht MK, MacIntyre RJ, Clegg MT (eds) (1995) *Evolutionary Biology, Volume 28*. New York: Plenum Press. xiv + 274pp. ISBN 0-306-44927-7. $75.00 (cloth).
- Zimmermann E, Newman JD, Jurgens U (1995) *Current Topics in Primate Vocal Communication*. New York: Plenum Press. x + 286 pp. ISBN 0-306-45064-X (cloth).
- Pahl-Wostl C (1995) *The Dynamic Nature of Ecosystems*. New York: John Wiley & Sons, Inc. xiii + 267 pp. ISBN 0-471-95570-1. $79.95 (cloth).
- Rogers J, Waldron T (1995) *A field Guide to Joint Disease in Archaeology*. New York: John Wiley & Sons, Inc. ISBN 0-471-95506-X $48.00 (paper).
- Dark KR (1995) *Theoretical Archaeology*. Ithaca, NY: Cornell University Press. 256 pp. ISBN 0-8014-8252-6. $19.95 (paper).
- Ospovat D (1995) *The Development of Darwin's Theory: Natural History and Natural Selection, 1838-1859*. Cambridge: Cambridge University Press. xviii + 301 pp. ISBN 0-521-46940-6. (paper).
- Kertzer DI, Laslett P (1995) *Aging in the Past: Demography, Society, and Old Age*. Berkeley: University of California Press. ISBN 0-520-08466-7. (paper).
- Simic A, Omidian PA, Almquist AJ (eds) (1995) *Human Sexuality*. Dubuque, IA: Kendall/Hunt Publishing Company. ix + 422pp. ISBN 0-7872-0524-9. (paper).
- Corbey R, Theunissen B (eds) (1995) *Ape, Man, Apeman: Changing Views since 1600*. Leiden: Department of Prehistory of Leiden University. 411pp. ISBN 90-73368-05-7.
- Stiner MC (1995) *Honor Among Thieves: A Zooarcheological Study of Neandertal Ecology*. Princeton, NJ: Princeton University Press. xxii + 447pp. ISBN 0-691-03456-7. $69.50 (cloth).
- Johnson R, Stam R (eds) (1995) *Brazilian Cinema*. New York: Columbia University Press. 491 pp. 0-231-10267-4. $19.50 (paper).
- Roberts C, Manchester K (1995) *The Archaeology of Disease*. Second Edition. New York: Cornell University Press. x + 243 pp. ISBN 0-8014-3220-0. $39.95 (cloth).
- Kelly RL (1995) *The Foraging Spectrum: Diversity in Hunter-Gatherers Lifeways*. Washington, DC: Smithsonian Institution Press. xvi + 446 pp. ISBN 1-56098-466-X.$24.95 (paper).
- Leonard RJ (1995) *Human Gross Anatomy: An Outline Text*. New York: Oxford University Press. xv + 437 pp. ISBN 0-19-509003-9. $24.95 (paper).
- Kurtèn B (1995) *The Cave Bear Story: Life and Death of a Vanished Animal*. New York: Columbia University Press. xi + 163 pp. ISBN 0-231-10361-1. $11.95 (paper).