PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Evolution of Modern Humans: Evidence from Nuclear DNA Polymorphisms

Joanna L. Mountain, Alice A. Lin, Anne M. Bowcock and L. Luca Cavalli-Sforza

Phil. Trans. R. Soc. Lond. B 1992 **337**, 159-165
doi: 10.1098/rstb.1992.0093

Evolution of modern humans: evidence from nuclear DNA polymorphisms

JOANNA L. MOUNTAIN,[1] ALICE A. LIN,[1] ANNE M. BOWCOCK[2] AND L. LUCA CAVALLI-SFORZA[1]

[1] *Department of Genetics, Stanford University, Stanford, California 94305, U.S.A.*
[2] *Department of Pediatrics, University of Texas Southwestern Medical Center, Dallas, Texas 75235, U.S.A.*

SUMMARY

Previously we have described studies of the evolution of modern humans based upon data for classical genetic markers and for nuclear DNA polymorphisms. Such polymorphisms provide a different point of view regarding human evolution than do mitochondrial DNA sequences. Here we compare revised dates for major migrations of anatomically modern humans, estimated from archaeological data, with separations suggested by a genetic tree constructed from classical marker allele frequencies. Analyses of DNA polymorphisms have now been extended and compared with those of classical markers; genetic trees continue to support the hypothesis of an initial African and non-African divergence for modern humans. We have also begun testing non-human primates for a set of human DNA polymorphisms. For most polymorphisms tested so far, humans share a single allele with other primates; such shared alleles are likely to be ancestral. Populations living in humid tropical environments have significantly higher frequencies of ancestral alleles than do other populations, supporting the hypothesis that natural selection acts to maintain high frequencies of particular alleles in some environments.

1. INTRODUCTION

Although genetic data have recently begun to play an important role in clarifying our understanding of the evolution of modern humans, results derived from currently available data are far from conclusive. While mitochondrial DNA (mtDNA) data are being generated at a rapid rate, the mitochrondrial genome represents only a small fraction of an individual's genetic material and may not be representative of the whole. Allele frequencies for 'classical' nuclear genetic markers (blood group, protein, and HLA polymorphisms) are available for thousands of human populations, but these markers are very few. Although nuclear DNA polymorphisms are much more numerous, they have been tested on a limited scale, for only a few populations. For these reasons, among others, current genetic data sets often fail to provide answers to any but the most simple of hypotheses. Additional data, along with novel approaches to data analysis, are beginning to provide more robust conclusions, as well as insight into more complex issues.

One means of evaluating the robustness of results is to examine consistency among various data sets, both genetic and non-genetic (i.e. archaeological, paleoanthropological, linguistic, etc.). Consistency between data sets is examined here for two cases. First, separation points of a genetic tree are compared with dates of human migrations estimated from archaeological data. Second, genetic distances and tree topologies estimated from two independent genetic data sets are examined.

A new perspective on human DNA polymorphisms is provided by data on non-human primates. Testing of human DNA polymorphisms in primates reveals a high degree of sharing among the species and indicates that humans are more closely related to chimpanzees than they are to any other species. Frequencies of the alleles shared by humans and chimpanzees, assumed to be ancestral, provide insight into evolutionary rates within human populations. In addition, distributions of ancestral allele frequencies in different populations indicate that natural selection may have acted on some loci.

2. CLASSICAL MARKERS

(*a*) *Tree of 42 populations*

Previously we published a tree relating forty-two aboriginal populations according to genetic differences (Cavalli-Sforza *et al.* 1988). These differences, called genetic distances, were estimated using population frequencies for 120 alleles associated with 44 polymorphisms. The data represent an abstraction of a large collection of classical marker allele frequencies (Cavalli–Sforza *et al.* 1993). A condensed version of the genetic tree is shown in figure 1. The first split in this tree separates Africans from non-Africans, lending support to the hypothesis that humans originated in

Phil. Trans. R. Soc. Lond. B (1992) **337,** 159–165
Printed in Great Britain

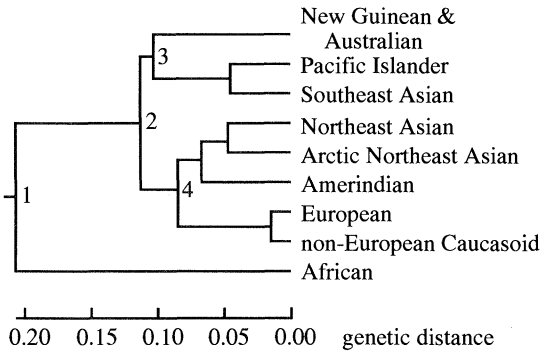


Figure 1. Condensed version of a genetic tree for 42 aboriginal populations (see Cavalli-Sforza *et al.* 1988). Each of the nine populations represents many samples. The tree was constructed according to the average linkage algorithm (Sokal *et al.* 1958) from F_{ST} genetic distances (Reynolds *et al.* 1983) based on frequencies of 120 alleles associated with 44 classical markers (blood group, HLA, and protein polymorphisms). Numbers indicate nodes comparable with those of figure 4*a*.

Africa, and spread from there throughout the rest of the world. The second split separates southeast Asians and Pacific Islanders from northeast Asians, Caucasoids, and Amerindians, suggesting two main migrations into Asia. This tree was compared to a classification of languages; a test of significance demonstrated that there is a highly significant correlation between genetic and linguistic evolution (Cavalli-Sforza *et al.* 1992).

(*b*) *Comparison with archaeological data*

Major separations in this tree can be compared to dates of human migrations, estimated from archaeological data (see Cavilli-Sforza *et al.* (1988) and figure 2). Anatomically modern humans are assumed to have migrated from Africa roughly 100 ka ago (Braüer 1989; Clark 1989; Stringer 1990). Expansion from southeast Asia to Australia and New Guinea is assumed to have taken place approximately 55 ka ago (Roberts *et al.* 1990), and migration of modern humans into Europe is assumed to have taken place about 43 ka ago (Straus 1989). The ratio of genetic distance (G) to time (T, in years) for these three events is roughly constant ($G/T = 2.07 \pm 0.15$). While no time scale is provided by the genetic tree, consistency between the archaeological and genetic data lends support to the hypothesis that the genetic tree corresponds roughly to the evolution of human populations over the last 100 ka.

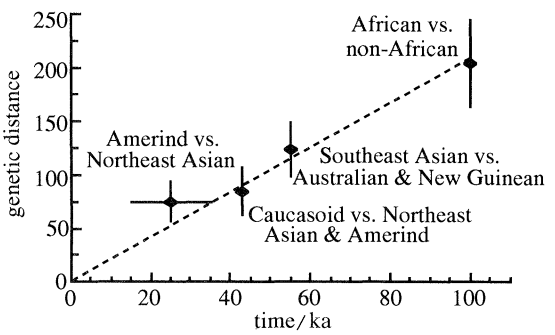


Figure 2. Graphical comparison of genetic distances and dates of early intercontinental migrations of anatomically modern humans, estimated from archaeological data. Genetic distances are those suggested by the tree of figure 1, $\times 10^3$. Standard errors for these distances were obtained from 100 bootstraps of the genetic distances (Efron 1982). Dates are given in ka. The date for migration to the Americas is given as a range (15–35 ka BP). The dashed line passes through both the origin and the weighted average of the three points other than that representing migration into America.

The time of the first migration of modern humans to the Americas remains controversial; although estimates range from 15 ka to 35 ka before present (BP). (Fagan 1987), these genetic data support a date at the upper end of that range (figure 2). The more detailed set of nuclear genetic data agree with the three major migrations suggested by Greenberg (Cavalli-Sforza *et al.* 1993). MtDNA restriction site data have recently confirmed two of the major migrations leading to the peopling of the Americas (Torroni *et al.* 1992).

3. DNA MARKERS: COMPARISONS OF HUMAN POPULATIONS

Over the past ten years new methods for detecting genetic differences have been developed. These detect variation at the DNA sequence level, whether or not this variation is expressed at the peptide or protein level. One such method involves the comparison of lengths of DNA fragments; individuals may differ from one to another in the lengths of specific fragments. Thousands of these restriction fragment length polymorphisms (RFLPS) are currently available for testing. A second method involves a technique called the polymerase chain reaction (PCR). While sequence data such as those analysed for mitochondria can be obtained for nuclear genes, nuclear sequence data are more difficult to interpret at the intraspecific level because, unlike mtDNA, nuclear DNA undergoes frequent recombination. The resulting gene tree is therefore more complex, being reticulated (Hudson 1991). Instead, allele frequencies for DNA polymorphisms are usually employed to measure genetic variation across human populations.

(*a*) *Data for eight populations*

Previously we described a study of 100 DNA polymorphisms in five aboriginal populations (Bowcock *et al.* 1991*a*). Although based on a small number of population samples, the analysis improved on earlier studies of classical markers because a larger number of genetic markers were analysed: it has been shown that the number of independent markers is extremely important in obtaining accurate estimates of genetic distances and in tree reconstruction (Astolfi & Cavalli-Sforza 1981; Nei 1987). The populations considered were Pygmies of the Central African Republic (C.A.R.) and of Zaire, Melanesians of the Bougainville Islands, Chinese, and northern Europeans. This study is currently being extended to

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

include additional populations. To date 80 of the 100 DNA polymorphisms mentioned above have been tested in the five populations, plus Japanese, Australians, and New Guineans. Allele frequency data for these three populations are to be published elsewhere.

(*b*) *Comparison of genetic distances*

Genetic distances, calculated using a standardised measure of the variance of gene frequencies, F_{ST} (Reynolds *et al.* 1983), were compared with distances between similar samples calculated from classical marker data (figure 3). While the points are expected to fall along a straight line, they deviate substantially: the Melanesians, Australians, and New Guineans are relatively more distant from Japanese and Chinese populations according to the DNA data than they are according to the classical marker data. This may be due in part to differences in population samples. The Chinese population sample for the DNA study, for instance, includes both northern and southern Chinese, while the classical marker sample includes only southern Chinese.

The deviations may also be an indication that more markers are necessary to obtain valid estimates of genetic distance; 44 polymorphisms were considered for the classical marker study, while 80 were considered for the DNA marker study. Furthermore, sample sizes for classical markers are much larger than those for DNA polymorphisms. A larger number of markers plus increased sample sizes should lead to smaller standard errors for estimates of genetic distances, and consequently to greater consistency between independent studies.

(*c*) *Tree of eight populations*

An average linkage tree (Sokal & Michener 1958) relating the eight populations was constructed based on the genetic distances estimated from allele frequencies for 80 DNA polymorphisms. The tree (figure 4*a*) indicates an initial separation of Africans and non-Africans, followed by a separation of Melanesians, Australians and New Guineans from Japanese, Chinese, and Europeans. While the population samples tested for these 80 markers differ somewhat from those tested for classical markers, common components of the two trees can be compared (figures 1, 4*a*). Although branch lengths differ, all four possible divergences which can be compared occur in an identical order, indicating a high degree of consistency for the two data sets at this level of resolution. In particular, the greatest degree of divergence is between Africans and non-Africans. Thus, differences between genetic distance estimates are too small to translate into different tree topologies.

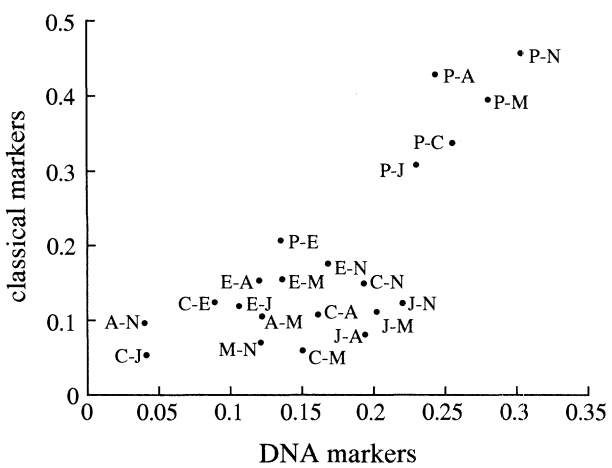


Figure 3. Comparison of two sets of genetic distances among seven populations. Classical marker distances were estimated from frequencies of 120 alleles associated with 44 blood group, protein, and HLA polymorphisms (Cavalli-Sforza *et al.* 1988). DNA distances were estimated from allele frequencies for 80 of the 100 nuclear DNA polymorphisms described by Bowcock *et al.* (1991*a*). A, Australians; C, Chinese; E, Europeans; J, Japanese; M, Melanesians; N, New Guineans; P, Zairean Pygmies.

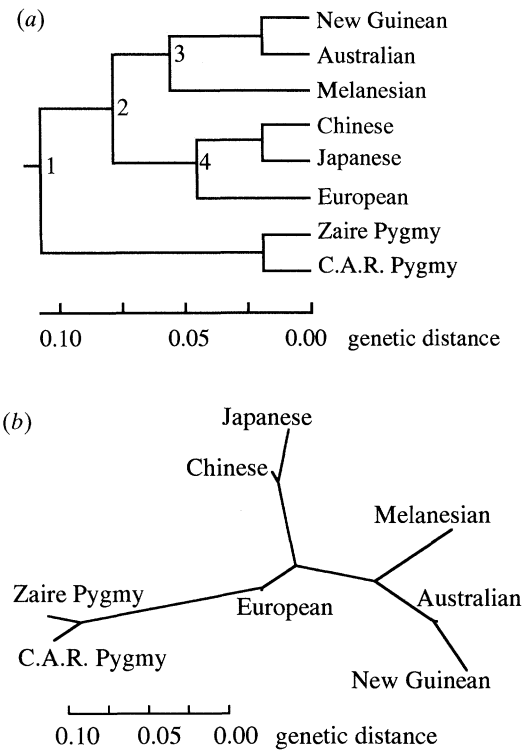


Figure 4. Genetic trees relating eight aboriginal populations. Each tree was constructed based on genetic distances estimated from allele frequencies for 80 nuclear DNA markers. Two methods of reconstruction were employed. (*a*) Average linkage, or UPGMA (Sokal *et al.* 1958); and (*b*) neighbour-joining (Saitou & Nei 1987). The former assumes a constant evolutionary rate. Numbers in (*a*) indicate nodes comparable with those in figure 1.

(*d*) *European admixture*

These nuclear DNA data are inconsistent with the bifurcating, constant rate model of the evolution of human populations; Europeans are much closer genetically to the Africans than are any of the other non-African populations, whereas all non-African populations are expected to be equidistant from the Africans if the evolutionary history suggested by the tree of figure 4*a* is reasonably accurate. The discrep-

ancy becomes apparent through construction of a tree using a method which does not assume a constant evolutionary rate (figure 4*b*). The method, called 'neighbour-joining', finds the shortest tree representing the genetic distances among the populations. The position of the European population in this tree differs from that in the average linkage tree (figure 4*a*) and the branch length leading to Europeans is extremely short. Such a short branch can arise when one population has been formed as an admixture of two or more others. The previous study of five populations indicated that Europeans arose as an admixture of ancestral Africans and ancestral Asians (Bowcock *et al.* 1991*b*). Other hypotheses may also be consistent with this tree structure; evidence against the hypothesis that the short branch length is due to a lower evolutionary rate among Europeans is presented below.

4. COMPARISON WITH OTHER PRIMATES

In order to place the human gene frequencies into a broader context, we have begun testing other primate species (chimpanzee, gorilla, and orangutan) for the same set of DNA polymorphisms as has been tested in the human populations. Humans have been found to share a polymorphism (at least two shared alleles at a genetic locus) with one of these closely related species in only two cases; they share one polymorphism with chimpanzees and another with orangutans. For many other polymorphisms, however, humans share exactly one allele with other primates. These shared alleles indicate sequence similarity for a small stretch of DNA. Considering only those 71 polymorphisms for which all three of the non-human primate species have been tested, in 57 out of 71 cases chimpanzees and humans share a single allele, in 39 cases gorillas and humans share one allele, and in 41 out of 71 cases orangutans and humans share one allele.

(*a*) *Tree of primate species*

Although these polymorphisms were initially detected in humans, and are therefore a biased sample, the number of cases of shared alleles provides some insight into primate evolution. Given that chimpanzees and humans share more alleles, humans and chimpanzees are significantly more closely related to each other than humans and gorillas ($\chi^2 = 10.4$, 1 d.f., $p < 0.005$) or humans and orangutans ($\chi^2 = 8.4$, 1 d.f., $p < 0.005$). These data therefore represent some of the strongest evidence available in support of the hypothesis that humans and chimpanzees are more closely related to each other than either is to any other species (Goodman *et al.* 1989; Tajima 1992). Furthermore, given that gorillas and orangutans share similar numbers of alleles (although often different alleles) with humans, a tree placing these two species in separate clusters is most consistent with the data.

(*b*) *Ancestral alleles*

We have reexamined the frequencies within the human populations for those alleles shared with chimpanzees. For 60 out of 80 polymorphisms studied for all eight human populations, chimpanzees share a single allele with humans. These 60 alleles are assumed to have existed from the time of the separation of ancestral humans and ancestral chimpanzees, and to represent the ancestral alleles at their respective genetic loci. Mean frequencies of these 60 ancestral alleles for each population were calculated (table 1). The standard error estimates are inappropriate for assessing the significance of differences between the populations: they encompass not only variation between populations but also variation between markers. A more appropriate approach to the question of whether mean frequencies differ significantly among populations is an analysis of variance. The results of this analysis (table 2) indicate that the frequencies of ancestral alleles differ very significantly from population to population ($p < 0.005$). This implies, at the very least, that the two Pygmy populations have significantly higher frequencies of ancestral alleles than do the two Asian populations.

Table 1. *Mean frequencies, for each of eight populations, of 60 alleles present among both humans and chimpanzees*

(Shared alleles are assumed to be the ancestral at their respective loci. Populations are listed in decreasing order of mean frequency. See text for discussion of standard errors (s.e.).)

population	mean frequency ± s.e.
Zaire Pygmies	0.648 ± 0.038
C.A.R. Pygmies	0.618 ± 0.039
Melanesians	0.578 ± 0.046
New Guineans	0.564 ± 0.046
Europeans	0.563 ± 0.035
Australians	0.560 ± 0.042
Japanese	0.545 ± 0.042
Chinese	0.534 ± 0.039

(*c*) *Distributions of ancestral allele frequencies*

To further explore the differences among human populations, histograms of the ancestral allele frequencies were plotted for each (figure 5). The variation among mean frequencies of ancestral alleles, as shown in table 1 and evaluated in table 2, is reflected in these

Table 2. *Comparison of means of 60 ancestral allele frequencies across eight populations.*

(Analysis of variance was carried out after angular transformation of the allele frequencies (Sokal & Rohlf 1969). Mean frequencies for each population are given in table 1)

source	d.f.	mean square	*F*-test	*P* value
Among alleles	59	0.9688	17.653	< 0.0001
Among populations	7	0.1633	3.080	0.0036
residual	413	0.0530		
total	479			

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

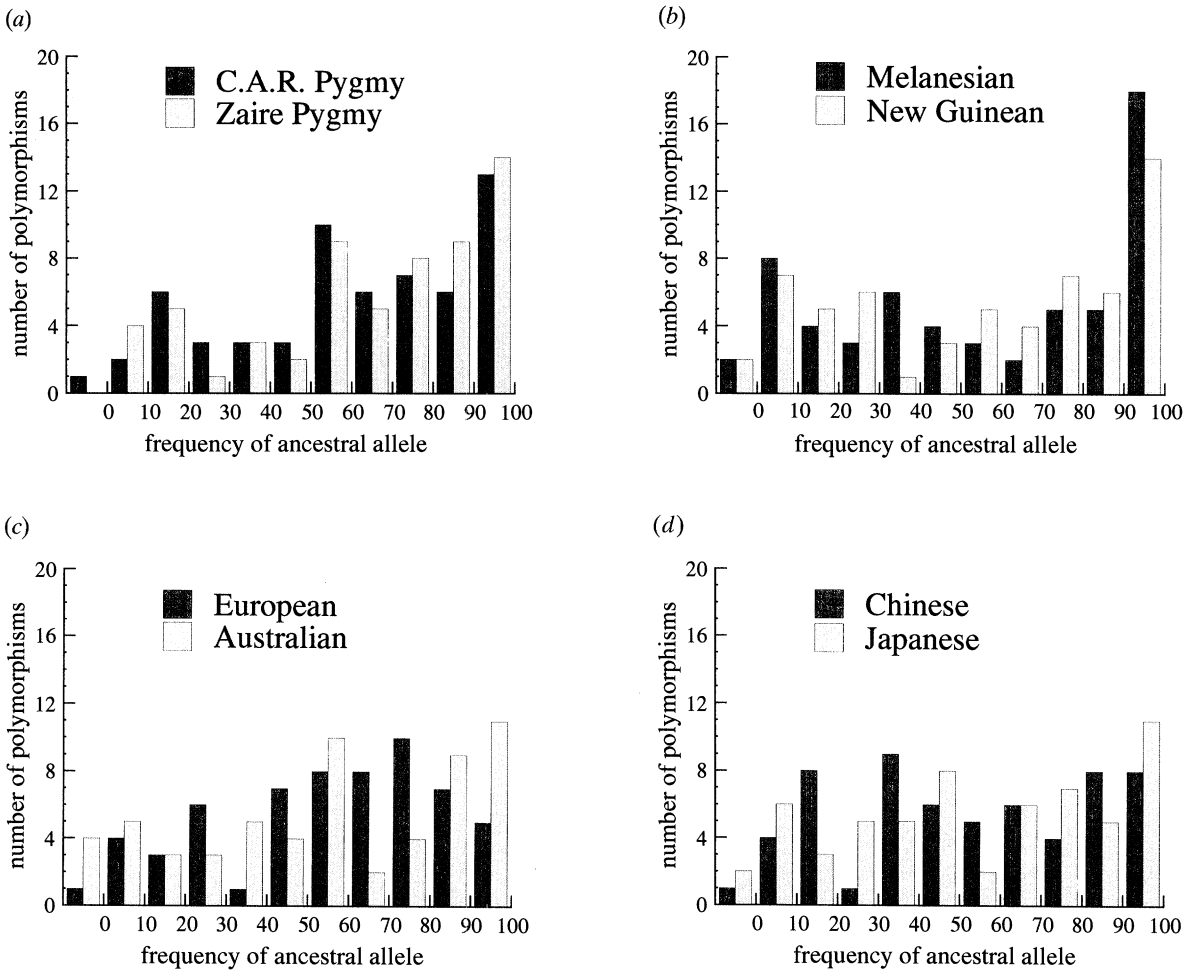


Figure 5. Histograms recording, for each of eight populations, the number of ancestral alleles with frequencies in each of 11 classes. Alleles shared with chimpanzees are considered ancestral. The first class includes only alleles with frequencies of 0%. All other classes represent alleles with frequencies in a 10% range (0–10%, 10–20%, etc.), excluding the lower value. Two populations are given per panel. (*a*, *b*) Distributions for these four populations (C.A.R. Pygmies, Zaire Pygmies, Melanesians, and New Guineans) differ significantly from equality; the populations tend to have high frequencies for more ancestral alleles than do the other four populations. (*c*, *d*) Distributions for these four populations (Australians, Caucasians, Japanese, and Chinese) do not differ significantly from equality.

histograms. Some populations have a very uneven distribution, with many ancestral alleles in the highest class (frequencies of 90–100%). Figure 5*a*,*b* represents histograms which deviate significantly from equality; a larger proportion of ancestral alleles have high frequencies in these populations (C.A.R. Pygmies, Zaire Pygmies, Melanesians, and New Guineans) than in the remaining four populations (Australians, Europeans, Japanese, and Chinese, figure 5*c*,*d*).

5. DISCUSSION

(*a*) *Dates for origin of modern humans*

DNA polymorphisms are currently being examined in a number of aboriginal human populations. Although genetic distances estimated from data for the eight populations discussed here differ from those estimated from classical marker data, the resulting genetic trees are similar. Both indicate an initial separation between Africans and non-Africans. The suggestion of an African origin for humans could reflect a migration out of Africa either 1–1.5 Ma ago, 100 ka ago, or both. Nuclear genetic data alone provide no timescale. Internal calibration can be carried out, however, through comparison with archaeological data. Classical polymorphisms indicate a correlation between genetic and archaeological data sets, lending support to the hypothesis that the migration out of Africa corresponds to the more recent date.

In this analysis Africans are represented by two African Pygmy populations. Two factors, however, suggest that the Pygmy populations are reasonably representative. First, one of these two Pygmy populations, that of the Central African Republic, is known to be highly mixed with non-Pygmy Africans. Despite this level of admixture, the two Pygmy populations cluster fairly closely in the inferred trees. Second,

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

many more African populations have been studied for classical nuclear polymorphisms such as blood groups. These data indicate that while African Pygmies differ somewhat, genetically, from other sub-saharan African populations, they cluster with them in the inferred trees. Bushmen may be slightly more different from other sub-saharan African populations than Pygmies, but they still cluster with them (Cavalli-Sforza *et al.* 1988). In any case, additional African populations are currently under study, and will eventually provide more complete representation.

Mitochondrial DNA data are also consistent with the date of 100 ka, suggesting a date of about 200 ka for the common ancestor of modern mtDNA types. Although estimates of dates based on mtDNA span a broad range, the upper values do not appear to extend far beyond 500 ka (Stoneking & Cann 1989; Stoneking *et al.*, this symposium). These dates, if accurate, provide upper limits to the date of migration from Africa, assuming that the genetic trees herein reflect evolutionary history. Gene trees (such as those relating mtDNA types) and trees representing population histories are unlikely to correspond; the divergences within gene trees necessarily predate those within corresponding population trees. Dates of 200 ka, or even of 500 ka, for the common ancestor are therefore consistent with the hypothesis that modern humans first began expanding and migrating throughout the world around 100 ka ago.

(*b*) *Evolutionary rates*

The testing of non-human primate populations for human nuclear DNA polymorphisms has provided not only reinforcement for the hypothesis that humans and chimpanzees are closely related compared to other species, but also an indication of which alleles are ancestral. The variation in frequencies of ancestral alleles provides evidence against two hypotheses superficially suggested by some evolutionary trees. In both the classical marker and DNA marker trees the African populations are found at the end of a very long branch (figures 1 and 4). The length of this branch may indicate either a higher evolutionary rate within African populations, or an African origin for modern humans. The former would have occurred if African populations remained relatively small, or experienced bottlenecks, after separation from other populations. The African populations, however, have higher frequencies of ancestral alleles, indicating that these populations have experienced a relatively low rate of allele frequency change (tables 1 and 2).

Likewise, the short branch leading to Europeans (figure 4*b*) may indicate either that Europeans have evolved more slowly than other populations, due to a large population size, or that Europeans are an admixture of other populations. Europeans, however, have relatively intermediate frequencies of ancestral alleles, and therefore do not appear to have evolved at a particularly low rate. The hypothesis that Europeans arose as an admixture appears to be more consistent with the data.

(*c*) *Natural selection*

Examination of the distributions of ancestral allele frequencies reveals variation among the eight populations. Four of them tend to have significantly uneven distributions, with high frequencies of many ancestral alleles (figure 5*a*,*b*). This variation among the populations may be due to differences in effective population sizes; if the African populations were larger, then gene frequencies would have changed at a slower rate in these populations than in other populations. This hypothesis is now being examined analytically.

Natural selection may also have acted on some of the polymorphisms, contributing to differences among populations in the frequencies of ancestral alleles. An earlier analysis indicated that up to one-third of the 100 DNA polymorphisms may be under the influence of either stabilizing or disruptive selection (Bowcock *et al.* 1991*b*). Given that the 60 putative ancestral alleles are associated with a subset of these 100 polymorphisms, at least a fraction are likely to be influenced by natural selection. Because the two African, the Melanesian, and the New Guinean populations each appear to have many ancestral alleles with high frequencies, the alleles in the top class (90–100%) were examined in these populations. For significantly many polymorphisms ($\chi^2 = 4.8313$, 1 d.f., $p < 0.05$) Africans (figure 5*a*) have high frquencies of the same alleles as do the Melanesians and New Guineans (figure 5*b*). The number of cases where Africans have high frequencies of the same alleles as do the Chinese and Japanese (figure 5*d*) is not significant ($\chi^2 = 0.891$, 1 d.f.). Thus, the equatorial populations tend to have very high frequencies for the same ancestral alleles. This finding indicates that natural selection may be influencing some allele frequencies, maintaining high frequencies of those alleles shared with chimpanzees. Because Pygmies, Melanesians, New Guineans tend to inhabit a rain forest environment similar to that of chimpanzees, natural selection may be acting on a few polymorphisms within these populations, maintaining high frequencies of ancestral alleles.

Currently there is an effort underway to collect samples for DNA analysis from vanishing peoples around the world (Cavalli-Sforza *et al.* 1991). If successful, this effort will lead to much larger databases for both mitochondrial and nuclear DNA variation. Each of these types of data will continue to make distinct contributions to our understanding of the evolution of modern humans. Any such findings can then be compared to those of other disciplines; additional data should enable us to detect further details and obtain more consistent results at each level of comparison. We may eventually clarify details of major migration events, population size changes, and the magnitude of selection effects.

This investigation has been supported in part by National Institutes of Health grants GM20467 and EY07106. We thank the Royal Society for providing funding (to J.L.M.) for travel to attend this meeting, and the Lucille P. Markey Charitable Trust for providing funding (to L.C.S.) for travel to Africa to collect samples.

REFERENCES

Astolfi, P. & Cavalli-Sforza, L.L. 1981 A comparison of methods for reconstructing evolutionary trees. *Syst. Zool.* **30**, 156–169.

Bowcock, A.M., Hebert, J.M., Mountain, J.L., Kidd, J.R., Rogers, J., Kidd, K.K. & Cavalli-Sforza, L.L. 1991*a* Study of an additional 58 DNA markers in five human populations from four continents *Gene Geog.* **5**, 151–173.

Bowcock, A.M., Kidd, J.R., Mountain, J.L., Hebert, J.M., Carotenuto, L., Kidd, K.K. & Cavalli-Sforza, L.L. 1991*b* Drift, admixture, and selection in human evolution: A study with DNA polymorphisms. *Proc. natn. Acad. Sci. U.S.A.* **88**, 839–843.

Brauer, G. 1989 The evolution of modern humans: a comparison of the African and non-African evidence. In *The human revolution: behavioural and biological perspectives on the origins of modern humans* (ed. P. Mellars & C. Stringer), pp. 123–154. Princeton University Press.

Cavalli-Sforza, L.L., Minch, E. & Mountain, J.L. 1992 Coevolution of genes and languages revisited. *Proc. natn. Acad. Sci. U.S.A.* **89**, 5620–5624.

Cavalli-Sforza, L.L., Piazza, A. & Menozzi, P. 1993 *History and geography of human genes*. Princeton University Press. (In the press.)

Cavalli-Sforza, L.L., Piazza, A., Menozzi, P. & Mountain, J. 1988 Reconstruction of human evolution: Bringing together genetic, archaeological, and linguistic data. *Proc. natn. Acad. Sci.* **85**, 6002–6006.

Cavalli-Sforza, L.L., Wilson, A.C., Cantor, C.R., Cook-Deegan, R.M. & King, M.-C. 1991 Call for a worldwide survey of human genetic diversity: a vanishing opportunity for the Human Genome Project. *Genomics* **11**, 490–491.

Clark, J.D. 1989 The origin and spread of modern humans: a broad perspective on the African evidence. In *The human revolution: behavioural and biological perspectives on the origins of modern humans* (ed. P. Mellars & C. Stringer) pp. 565–588. Princeton University Press.

Efron, B. 1982 *The jackknife, bootstrap, and other resampling plans*. Philadelphia: Society for Industrial and Applied Mathematics.

Fagan, B.M. 1987 *The great journey: the peopling of ancient America*. London: Thames and Hudson.

Goodman, M., Koop, B.F., Czelusniak, J., Fitch, D.H.A., Tagle, D.A. & Slightom, J.L. 1989 Molecular phylogeny of the family of apes and humans. *Genome* **31**, 316–335.

Hudson, R.R. 1991 Gene genealogies and the coalescent process. *Oxf. Surv. Evol. Biol.* **7**, 1–44.

Nei, M. 1987 *Molecular evolutionary genetics*. New York: Columbia University Press.

Reynolds, J., Weir, B.S. & Cockerham, C.C. 1983 Estimation of the coancestry coefficient: basis for a short-term genetic distance. *Genetics* **105**, 767–779.

Roberts, R.G., Jones, R. & Smith, M.A. 1990 Thermoluminescence dating of a 50,000-year-old human occupation site in northern Australia. *Nature, Lond.* **345**, 153–156.

Saitou, N. & Nei, M. 1987 The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molec. Biol. Evol.* **4**, 406–425.

Sokal, R.R. & Michener, C.D. 1958 A statistical method for evaluating systematic relationships. *Univ. Kansas Sci. Bull.* **38**, 1409–1437.

Sokal, R.R. & Rohlf, F.J. 1969 *Biometry: The principles and practice of statistics in biological research*. San Francisco: W. H. Freeman.

Stoneking, M. & Cann, R.L. 1989 African Origin of Human Mitochondrial DNA. In, *The human revolution: behavioral and biological perspectives on the origins of modern humans* (ed. P. Mellars & C. Stringer), pp. 17–30. Princeton University Press.

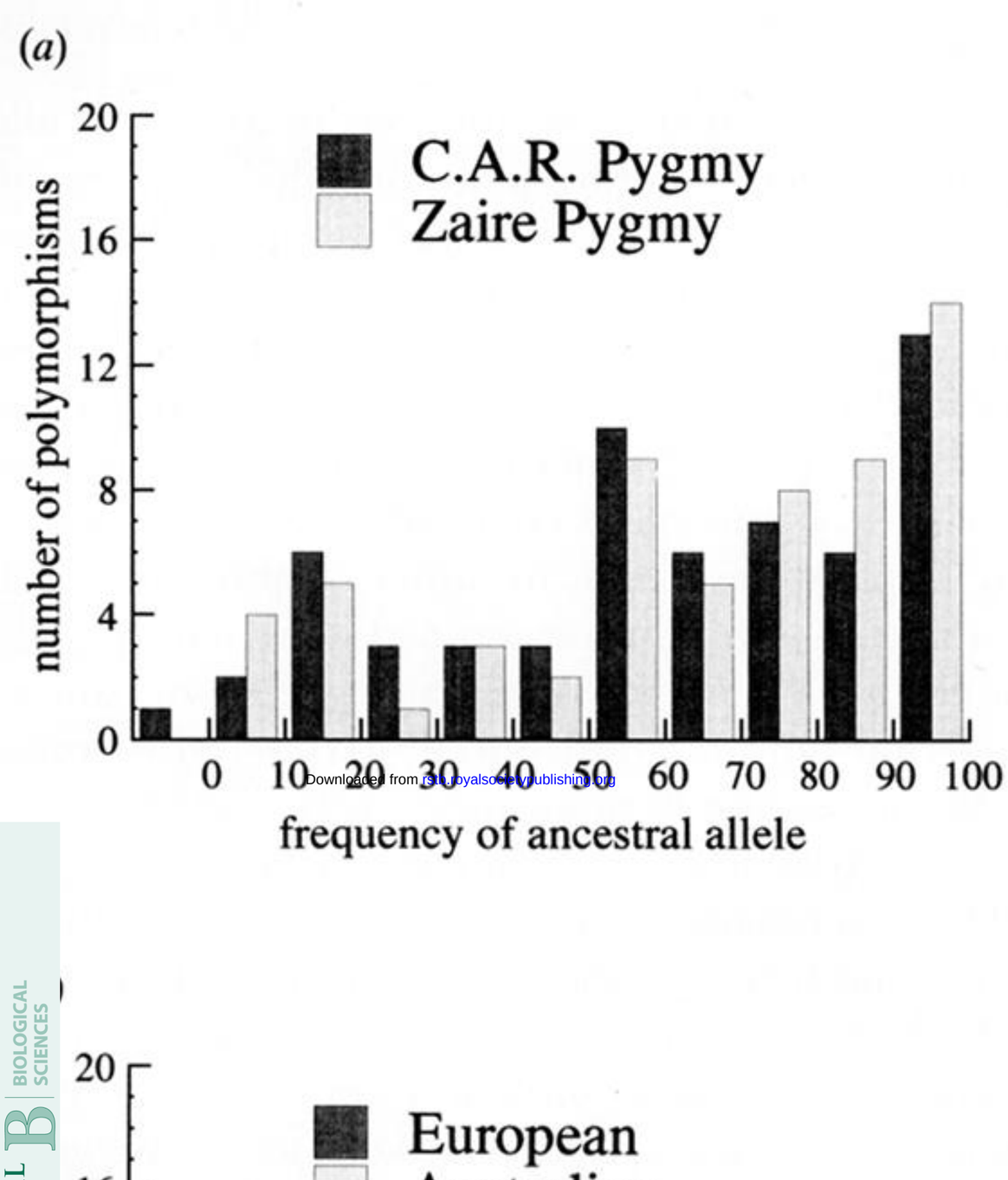
Straus, L.G. 1989 Age of modern Europeans. *Nature, Lond.* **342**, 476–477.

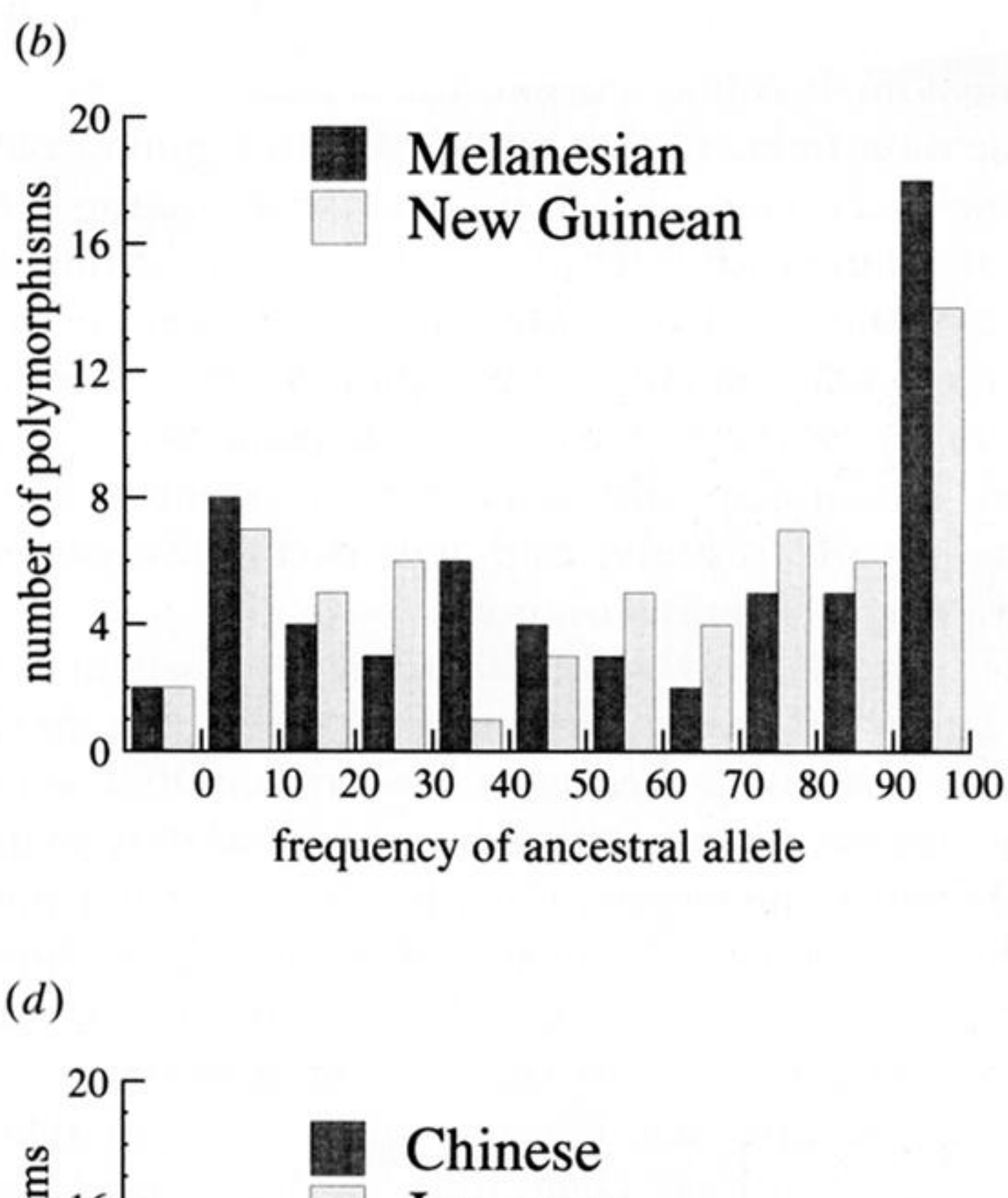
Stringer, C.B. 1990 The emergence of modern humans. *Scient. Am.* **Dec.** 98–104.

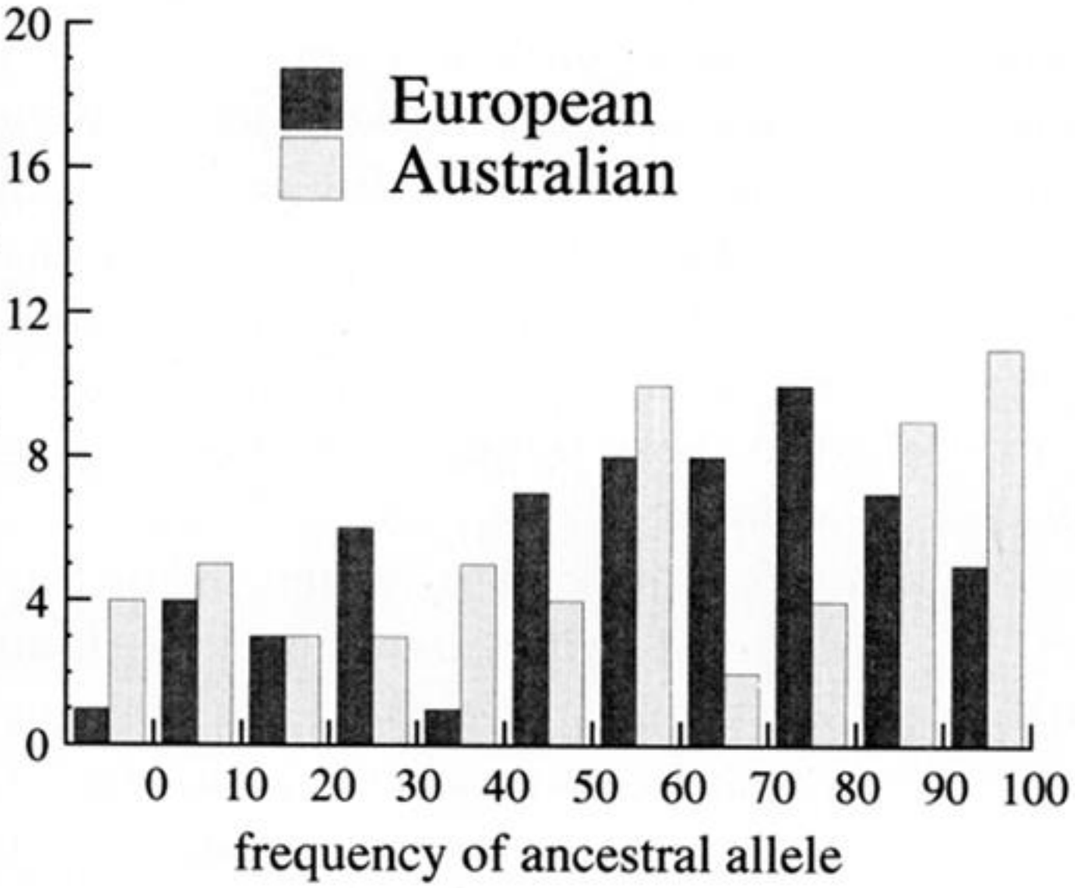
Tajima, F. 1992 Statistical method for estimating the standard errors of branch lengths in a phylogenetic tree reconstructed without assuming equal rates of nucleotide substitution among different lineages. *Molec. Biol. Evol.* **9**, 168–181.

Torroni, A., Schurr, T.G., Yang, C.-C., Szathmary, E.J.E., Williams, R.C., Schanfield, M.S., Troup, G.A., Knowler, W.C., Lawrence, D.N., Weiss, K. & Wallace, D.C. 1992 Native American mitochondrial DNA analysis indicates that the Amerind and Nadene populations were founded by two independent migrations. *Genetics* **130**, 153–162.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES







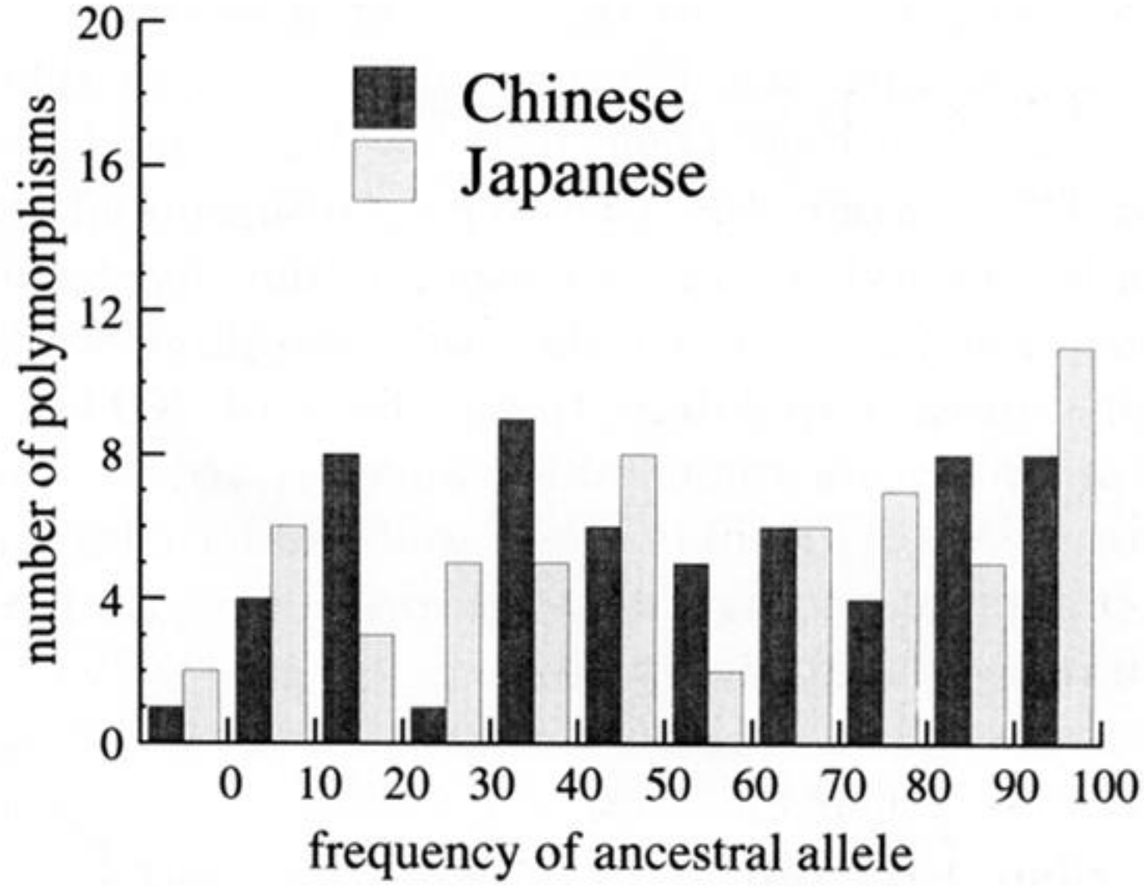


Figure 5. Histograms recording, for each of eight populations, the number of ancestral alleles with frequencies in each of 11 classes. Alleles shared with chimpanzees are considered ancestral. The first class includes only alleles with frequencies of 0%. All other classes represent alleles with frequencies in a 10% range (0–10%, 10–20%, etc.), excluding the lower value. Two populations are given per panel. (*a*, *b*) Distributions for these four populations (C.A.R. Pygmies, Zaire Pygmies, Melanesians, and New Guineans) differ significantly from equality; the populations tend to have high frequencies for more ancestral alleles than do the other four populations. (*c*, *d*) Distributions for these four populations (Australians, Caucasians, Japanese, and Chinese) do not differ significantly from equality.