Department of Biochemistry, University of California, Berkeley, California 94720, USA

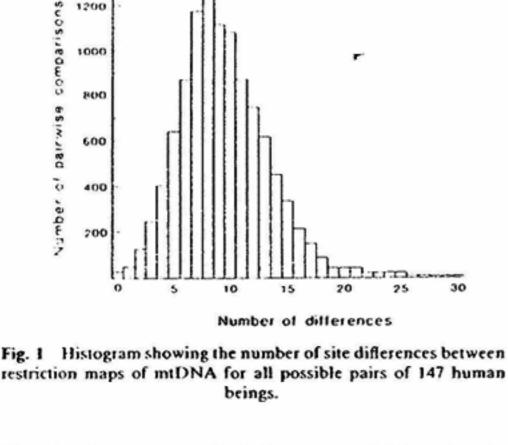
Mitochondrial DNAs from 147 people, drawn from five geographic populations have been analysed by restriction mapping. All these mitochondrial DNAs stem from one woman who is postulated to have lived about 200,000 years ago, probably in Africa. All the populations examined except the African population have multiple origins, implying that each area was colonised repeatedly.

MOLECULAR biology is now a major source of quantitative and objective information about the evolutionary history of the human species. It has provided new insights into our genetic divergence from apes1.8 and into the way in which humans are related to one another genetically 9.14. Our picture of genetic evolution within the human species is clouded, however, because it is based mainly on comparisons of genes in the nucleus. Mutations accumulate slowly in nuclear genes. In addition, nuclear genes are inherited from both parents and mix in every generation. This mixing obscures the history of individuals and Present address: Department of Genetics, University of Hawaii, Honolulu, Hawaii 96822.

allows recombination to occur. Recombination makes it hard to trace the history of particular segments of DNA unless tightly linked sites within them are considered. Our world-wide survey of mitochondrial DNA (mtDNA) adds to knowledge of the history of the human gene pool in three

ways. First, mtDNA gives a magnified view of the diversity present in the human gene pool, because mutations accumulate in this DNA several times faster than in the nucleus15. Second, because mtDNA is inherited maternally and does not recombine16, it is a tool for relating individuals to one another. Third, there are about 1016 mtDNA molecules within a typical human and they are usually identical to one another 17-19. Typical mam--articles -

1.5



mtDNA molecules within each oocyte20. This maternal and haploid inheritance means that mtDNA is more sensitive than nuclear DNA to severe reductions in the number of individuals in a population of organisms 15. A pair of breeding individuals

malian females consequently behave as haploids, owing to a

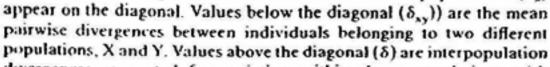
bottleneck in the genetically effective size of the population of

can transmit only one type of mtDNA but carry four haploid sets of nuclear genes, all of which are transmissible to offspring. The fast evolution and peculiar mode of inheritance of mt DNA provide new perspectives on how, where and when the human gene pool arose and grew. Restriction maps MtDNA was highly purified from 145 placentas and two cell lines, HeLa and GM 3043, derived from a Black American and an aboriginal South African (!Kung), respectively. Most placentas (98) were obtained from US hospitals, the remainder coming from Australia and New Guinea. In the sample, there were representatives of 5 geographic regions: 20 Africans (represent-

## ing the sub-Saharan region), 34 Asians (originating from China, Vietnam, Laos, the Philippines, Indonesia and Tonga), 46 Caucasians (originating from Europe, North Africa, and the

Guineans. Only two of the 20 Africans in our sample, those bearing mtDNA types 1 and 81 (see below) were born in sub-Saharan Africa. The other 18 people in this sample are Black Americans, who bear many non-African nuclear genes probably contributed mainly by Caucasian males. Those males would not be expected to have introduced any mtDNA to the Black American population. Consistent with our view that most of Table 1 MtDNA divergence within and between 5 human populations % sequence divergence Population 2 3 4 1. African 0.47 0.04 0.04 0.05 0.06 2. Asian 0.45 0.35 0.01 0.020.04 3. Australian 0.400.31 0.250.030.04 4. Caucasian 0.400.31 0.27 0.05 0.235. New Guinean 0.42 0.34 0.29 0.290.25 The divergence is calculated by a published method?6. Values of the

Middle East), 21 aboriginal Australians, and 26 aboriginal New



mean pairwise divergence between individuals within populations ( $\delta_s$ )

populations, X and Y. Values above the diagonal ( $\delta$ ) are interpopulation divergences, corrected for variation within those populations with equation (1).	m
NATURE VOL. 325 1 JANUARY 1987	CLE
no individual contained more than one type. None of the seven	f
shared types occurred in more than one of the five geographic	C
regions. One type, for example, was found in two Australians.	p

A histogram showing the number of restriction site differences between pairs of individuals is given in Fig. 1; the average number of differences observed between any two humans is 9.5. The distribution is approximately normal, with an excess of pairwise comparisons involving large numbers of differences. From the number of restriction site differences, we estimated the extent of nucleotide sequence divergence26 for each pair of individuals. These estimates ranged from zero to 1.3 substitutions per 100 base pairs, with an average sequence divergence

of 0.32%, which agrees with that of Brown , who examined

Table 1 gives three measures of sequence divergence within

and between each of the five populations examined. These

measures are related to one another by equation (1):

Among Caucasians, another type occurred three times and two

more types occured twice. In New Guinea, two additional types

were found three times and the seventh case involved a type

found in six individuals.

only 21 humans.

 $\delta = \delta_{xy} - 0.5 \left( \delta_x + \delta_y \right)$ (1) where  $\delta_x$  is the mean pairwise divergence (in percent) between individuals within a single population (X),  $\delta_y$  is the corresponding value for another population (Y),  $\delta_{xy}$  is the mean pairwise divergence between individuals belonging to two different populations (X and Y), and  $\delta$  is a measure of the interpopulation divergence corrected for intrapopulation divergence. Africans

Indeed, the variation within the African population is as great

as that between Africans and any other group ( $\delta_{xy} = 0.40-0.45$ ).

The within-group variation of Asians ( $\delta_x = 0.35$ ) is also compar-

between groups slightly exceeds that within groups.

able to that which exists between groups. For Australians, Caucasians, and New Guineans, who show nearly identical amounts of within-group variation (
$$\delta_x = 0.23-0.25$$
), the variation

When the interpopulational distances ( $\delta_{xy}$ ) are corrected for intrapopulation variation (Table 1), they become very small  $(\delta = 0.01 - 0.06)$ . The mean value of the corrected distance among populations ( $\delta = 0.04$ ) is less than one-seventh of the mean distance between individuals within a population (0.30). Most of the mtDNA variation in the human species is therefore shared between populations. A more detailed analysis supports this runctional constraints Figure 2 shows the sequence divergence ( $\delta_x$ ) calculated for each population across seven functionally distinct regions of the mtDNA genome. As has been found before 24.27,28, the most variable region is the displacement loop ( $\delta_x = 1.3$ ), the major noncoding portion of the mtDNA molecule, and the least variable region is the 16S ribosomal RNA gene ( $\delta_x = 0.2$ ). In general, Africans are the most diverse and Asians the next most, across

A tree relating the 133 types of human mtDNA and the reference

sequence (Fig. 3) was built by the parsimony method. To interpret this tree, we make two assumptions, both of which have

extensive empirical support: (1) a strictly maternal mode of

ple mtDNA genomes. We can therefore view the tree as a

genealogy linking maternal lineages in modern human popula-

tions to a common ancestral female (bearing mtDNA type a).

miDNA transmission (so that any variant appearing in a group of lineages must be due to a mutation occurring in the ancestral lineage and not recombination between maternal and paternal genomes) and (2) each individual is homogeneous for its multi-

100

110

120

all functional regions.

Evolutionary tree

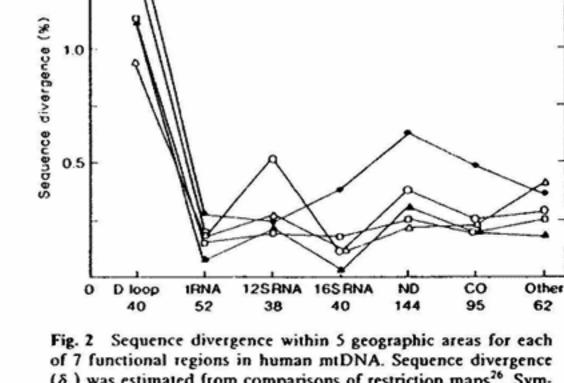
Many trees of minimal or near-minimal length can be made from the data; all trees that we have have examined share the British or opposition and the

· Africa

□ Europe

Australia A New Guinea

O Asia



 $(\delta_s)$  was estimated from comparisons of restriction maps<sup>26</sup>. Symbols for the 5 races are: ●, Africa; O, Asia; △, Australia; □, Caucasian; A, New Guinea. Along the horizontal axis are the numbers of restriction sites in each functional region (D loop, transfer RNA genes, 12S and 16S ribosomal RNA genes, NADH dehydrogenase subunits 1-5, cytochrome oxidase subunits and other protein-coding regions). these 18 people are a reliable source of African mtDNA, we

are mainly from the Eastern Highlands of Papua New Guinea23. Each purified mtDNA was subjected to high resolution mapping22-24 with 12 restriction enzymes (Hpal, Avall, FnuDll, Hhal, Hpall, Mbol, Taql, Rsal, Hinfl, Haelll, Alul and Ddel). Restriction sites were mapped by comparing observed fragment patterns to those expected from the known human mtDNA sequence25. In this way, we identified 467 independent sites, of which 195 were polymorphic (that is, absent in at least one individual). An average of 370 restriction sites per individual. were surveyed, representing about 9% of the 16,569 base-pair human mtDNA genome.

Map comparisons The 147 mtDNAs mapped were divisible into 133 distinct types. Seven of these types were found in more than one individual; Table 2 Clusters of mtDNA types that are specific to one geographic Number of Mean pairwise Average Geographic region-specific distance within age of region clusters clusters\* clusters !

90-180

53-105

43-85

23-45

28-55

## 15 Australia 0.17 Europe 0.09 36 0.11

	ers represented by two or more individuals (and calculated ils, not for mtDNA types) in Fig. 3.
	age in thousands of years based on the assumption that
	rgence occurs at the rate of 2-4% per million years 15,30.
	PINAL AIMPAIC THE COURCE There is only one Atmosp clusion.
+ 1452011111	g that Africa is the source, there is only one African cluster.
4 7/350mm	g that Africa is the source, there is only one African cluster.
FS	g that Africa is the source, there is only one African cluster.

Among the trees investigated was one consisting of five

primary branches with each branch leading exclusively to one

of the five populations. This tree, which we call the population-

specific tree, requires 51 more point mutations than does the tree of minimum length in Fig. 3. The minimum-length tree requires fewer changes at 22 of the 93 phylogenetically-informative restriction sites than does the population-specific tree, while the latter tree required fewer changes at four sites; both trees require the same number of changes at the remaining 67 sites. The minimum-length tree is thus favoured by a score of 22 to

4. The hypothesis that the two trees are equally compatible with

the data is statistically rejected, since 22:4 is significantly

different from the expected 13:13. The minimum-length tree is thus significantly more parsimonious than the populationspecific tree. We infer from the tree of minimum length (Fig. 3) that Africa is a likely source of the human mitochondrial gene pool. This inference comes from the observation that one of the two primary branches leads exclusively to African mtDNAs (types 1-7, Fig. 3) while the second primary branch also leads to African mtDNAs (types 37-41, 45, 46, 70, 72, 81, 82, 111 and 113). By postulating that the common ancestral mtDNA (type a in Fig. 3) was African, we minimize the number of intercontinental migrations needed to account for the geographic distribution of mtDNA types. It follows that b is a likely common ancestor of

all non-African and many African mtDNAs (types 8-134 in

The second implication of the tree (Fig. 3)—that each non-

In the same way, we calculate the minimum numbers of female ranging from 15 to 36 (Tables 2 and 3), will probably rise as more types of human mtDNA are discovered. Tentative time scale A time scale can be affixed to the tree in Fig. 3 by assuming that

mtDNA sequence divergence accumulates at a constant rate in

humans. One way of estimating this rate is to consider the extent

of differentiation within clusters specific to New Guinea (Table

2; see also refs 23 and 30), Australia30 and the New World31.

People colonised these regions relatively recently: a minimum

of 30,000 years ago for New Guinea<sup>32</sup>, 40,000 years ago for

Australia<sup>33</sup>, and 12,000 years ago for the New World<sup>34</sup>. These

times enable us to calculate that the mean rate of mtDNA

divergence within humans lies between two and four percent

NATURE VOL. 325 1 JANUARY 1987 ----- ARTICLES -----D LOOP: 8 2 5 8 9 73 81-83 85 86 118, 64 (16494) 26; 1341 3; 207h 128-134; 2551 39: 259a 106; 340 112 12S IRNA: 663e 43 48, 53; 712 90; 740 16-19; 748b 49; 1240a 7-9 37 110; 1403a (1448) 11-18 20-29 34 49 50 55-57; 1463e 14; 1484e 2 4 5 8 11 14 17 19 23 25 30 32 35 41 43 45 52 58 61 107 110 120; 15361 31 VAL IRNA: 1610a 2 3 44 45 104, 1637c 47 102 108, 1667c 47 16S rBNA: 1715c 27 23 27 123; 1917a 70, 2208a 87; 2223a (2635) 37; 2384a (2472) 106; 2390; 2 3 8 9 49 79 93; 2734a 18 22 54 73 101, 2758k 1-8, 2849k 112;

3123k 119 ND 1: 3315e 11; 3337k 94 118 133; 3391e 41 105 122; 3537a 44; 3592h 2-7 9 37-41; 3698t 130-132; 3842e 44, 3849e 5 46; 3899t 14 18 22 23 32 71 92 102 110 120, 3930c 110, 39441 12; 4092g 112 MET IRNA: 4411a 23 35 36 57; 4464k 129 ND 2: 4481b (10933) 111 112, 4631a 2; 4643k 87 100; 4732k

ANCESTOR-0.4 0.4 0.2 0.2 Sequence divergence % Sequence divergence % elsewhere30. This rate is similar to previous estimates from animals as disparate as apes, monkeys, horses, rhinoceroses, mice, rats, birds and fishes15. We therefore consider the above estimate of 2%-4% to be reasonable for humans, although additional comparative work is needed to obtain a more exact calibration. As Fig. 3 shows, the common ancestral mtDNA (type a) links mtDNA types that have diverged by an average of nearly 0.57%. Assuming a rate of 2%-4% per million years, this implies that the common ancestor of all surviving mtDNA types existed

140,000-290,000 years ago. Similarly, ancestral types b-j may

% Sequence divergence Fig. 4 Genealogical tree relating the nucleotide sequences of D loops from seven human mtDNAs. This tree, which requires fewer mutations than any other branching order, was constructed by the PAUP computer program from the 900-bp sequences determined by Greenberg et al.28 and was rooted by the midpoint method.

Symbols: ●, African origin (Black American); □, Caucasian.

types by geographic origin that they observed. (By contrast, our

methods distinguish between mtDNAs that differ by 0.03%.)

Although Johnson et al. favoured an Asian origin, they too

found that Africans possess the greatest amount of mtDNA

sequences, and (3) that the common ancestor was African. Nuclear DNA studies Estimates of genetic distance based on comparative studies of nuclear genes and their products differ in kind from mtDNA estimates. The latter are based on the actual number of mutational differences between mtDNA genomes, while the former rely on differences in the frequencies of molecular variants peasured between and within populations. Gene frequencies

can be influenced by recombination, genetic drift, selection, and

No. of descendant lineages or clusters specific to a region Ancestor Total Africa Asia Australia 7 2 20 14 19 10 30

10

erectus in Asia was replaced without much mixing with the

Studies of mtDNA suggest a view of how, where and when modern humans arose that fits with one interpretation of

evidence from ancient human bones and tools. More extensive

molecular comparisons are needed to improve our rooting of

the mtDNA tree and the calibration of the rate of mtDNA divergence within the human species. This may provide a more

reliable time scale for the spread of human populations and

better estmates of the number of maternal lineages involved in

4. Ferris, S. D., Wilson, A. C. & Brown, W. M. Proc. nam. Acad. Sci. U.S.A. 78, 2432-2436

It is also important to obtain more quantitative estimates of

27

22

134

invading Homo sapiens from Africa.

founding the non-African populations.

2. Sarich, V. M. & Wilson, A. C. Science 158, 1200-1203 (1967).

3. King, M. C & Wilson, A. C. Science 188, 107-116 (1975).

Received 17 March; accepted 7 November 1986.

Goodman, M. Hum. Biol. 35, 377-424 (1963).

Sci U.S.A. (in the press).

(1981).

Conclusions and prospects

All

14608c 5 CYT b: 14749e 87; 14869j 86 126; 14956l 115; 15005g 56 57; 15172e2 10; 15195; (15221) 44; 15234g 113; 15238c 84 100; 15250c 14-18 20 21 24 25 54 81 90 111; 15606a 65 127-134; 15723g 43; 15790j (16373) 82; 15883e 7 45 58 59 112 THR IRNA: 15897k 97 129; 15907k 46; 15912 81; 15925 15 81 90 PRO IRNA: 15996c/16000g 48 D LOOP: 16049k 15 22 46 74 81; 16065g 1213 122; 16089k 21; 16096k 119; 16125k 3 9 74 78 88-91 93 105 119 120 122 16178 26-29; 16208k 11 15 51 52 55; 16217 95; 16246g 46; 16254a 127 1325 16303k 36 58 59 82 89-93 111 112 131 132; 16310k 1-3 5 6 8-14 20 26-31 58 73 78 82 94-97 99-102 120 125 126 133 134; 16389g/16390b 11 35-41 82 89g 16398e 5 11 34 37 66 118 119; 16490g 10; 16517e 2 11-13 15-24 26-29 51-553 Fig. 3 a, Genealogical tree for 134 types of human mtDNA (133 restriction maps plus reference sequence); b, comprehensive list of polymorphic restriction sites used. The tree accounts for the site differences observed between restriction maps of these mtDNAs with 398 mutations. No other order of branching tested is more parsimonious than this one. This order of branching was obtained (using the computer program PAUP, designed by Dr David Swofford) by ignoring every site present in only one type of mtDNA or absent in only one type and confining attention to the remaining 93 polymorphic sites. The computer program produces an unrooted network, which we converted to a tree by placing the root (arrow) at the midpoint of the longest path connecting the two lineages (see ref. 58). The numbers refer to mtDNA types, no. 1 being from the aboriginal South African (!Kung) cell line (GM 3043), no. 45 being from the HeLa cell line and no. 110 being the published human sequence25. Black bars, clusters of mtDNA types specific to a given geographic region; asterisks, mtDNA types found in more than 1 individual: type 134 was in six individuals, types 29, 65 and 80 each occurred thrice, and other types flagged with asterisks occurred twice. To place the nodes in the tree relative to the percent divergence scale, we took account of the differences observed at all 195 polymorphic sites. b, The numbering of sites is according to the published human sequence25, with 12 restriction enzymes indicated by the following single letter code: a, Alul; b, Avall; c, Ddel; d, FnuDll; e, Haelll; f, Hhal; g, Hinfl; h, Hpal; i, Hpall; j, Mbol; k, Rsal; l, Tagl. Italicized sites are present in the indicated mtDNA types and nonitalicized sites are absent in the indicated types; parentheses refer to alternative placements of inferred sites; sites separated by a slash are polymorphic for two different restriction enzymes caused by a single inferred nucleotide substitution; letters in bold face refer to noncoding regions and genes for transfer RNA, ribosomal RNA and proteins (ND, NADH dehydrogenase; CO, cytochrome oxidase; ATP, adenosine triphosphatase; CYT b, cytochrome b). For example, 8j indicates a Mbol site beginning at nucleotide position 8 in the D loop that was not found in mtDNA type 1 but was present in type 2, etc. Note that since this site is not present in the reference sequence (type 110), the sequence beginning at position 8 is actually a semisite, differing from the Mbol recognition sequence at one position (see ref. 23 for a detailed description of the method of mapping such inferred sites). Not all sites were scored in all individuals: 8j, 1484e and 7750c were not determined for types 1, 31, 59, 63, 68, and all of the New Guinea antDNAs; mtDNAs 114 and 121 could not be typed with Rsal. The locations of some sites differ from those reported before 23.74, as do the individuals in which some sites occur; these revisions are based on re-examination of previously-studied mtDNAs.

time. However, it is equally possible that the exodus occurred as recently as 23-105 thousand years ago (Table 2). The mtDNA results cannot tell us exactly when these migrations took places

Two previous studies of human mtDNA have included African

individuals21,28; both support an African origin for the humas

mtDNA gene pool. Johnson et al.21 surveyed ~40 restriction

sites in each of 200 mtDNAs from Africa, Asia, Europe and the

New World, and found 35 mtDNA types. This much smalles

number of mtDNA types probably reflects the inability of their

methods to distinguish between mtDNAs that differ by less that

0.3% and may account for the greater clustering of mtDN

differences are between Africans and other populations, suggesting an African origin for the human nuclear gene pool11,12,35 More recent studies of restriction site polymorphisms in nuclear

Our tentative interpretation of the tree (Fig. 3) and the associated time scale (Table 3) fits with one view of the fossil record: that the transformation of archaic to anatomically modern forms of Homo sapiens occurred first in Africa43-45, about 100,000-140,000 years ago, and that all present-day humans are descendants of that African population. Archaeologists have observed that blades were in common use in Africa 80-90 thousand years ago, long before they replaced flake tools in Asia or Europe.46,47 But the agreement between our molecular view and the evidence from palaeoanthropology and archaeology should be treated

DNA14,36-42 support these conclusions.

Relation to fossil record

cautiously for two reasons. First, there is much uncertainty about the ages of these remains. Second, our placement of the common ancestor of all human mtDNA diversity in Africa 140,000-

Other mtDNA studies

280,000 years ago need not imply that the transformation to anatomically modern Homo sapiens occurred in Africa at this time. The mtDNA data tell us nothing of the contributions to this transformation by the genetic and cultural traits of males and females whose mtDNA became extinct. An alternative view of human evolution rests on evidence that Homo has been present in Asia as well as in Africa for at least one million years 48 and holds that the transformation of archaic to anatomically modern humans occurred in parallel in different parts of the Old World 33,49. This hypothesis leads us to expect genetic differences of great antiquity within widely separated parts of the modern pool of mtDNAs. It is hard to reconcile the mtDNA results with this hypothesis. The greatest divergences within clusters specific to non-African parts of the World correspond to times of only 90,000-180,000 years. This might imply

Perhaps the non-African erectus population was replaced by sapiens migrants from Africa; incomplete fossils indicating the possible presence of early modern humans in western Asia at Zuttiyeh (75,000-150,000 years ago) and Qafzeh (50,000-70,000 years ago) might reflect these first migrations 45.53 If there was hybridization between the resident archaic forms in Asia and anatomically modern forms emerging from Africa, we should expect to find extremely divergent types of mtDNA in present-day Asians, more divergent than any mtDNA found in Africa. There is no evidence for these types of mtDNA among the Asians studied21,54-56. Although such archaic types of mtDNA could have been lost from the hybridizing population, the probability of mtDNA lineages becoming extinct in an expanding population is low57. Thus we propose that Homo Table 3 Ancestors, lineages and extents of divergence in the geneaological tree for 134 types of human mtDNA Europe N. Guinea % divergence Age\* 0.57 143-285 0.45 112-225 0.43108-215

0.27

0.26

0.25

68-135

65-130

62-125

NATURE VOL. 325 1 JANUARY 1987.

that the early Asian Homo (such as Java man and Peking man) contributed no surviving mtDNA lineages to the gene pool of our species. Consistent with this implication are features, found recently in the skeletons of the ancient Asian forms, that make it unlikely that Asian erectus was ancestral to Homo sapiens 50-52.

> 0.39 98-195 0.3485-170 0.30 75-150 0.2870-140

> > 7

BODE STOP DESCRIPTION OF THE STORY OF THE ST ----- LETTERSTONATURE

We thank the Foundation for Research into the Origin of Man, the National Science Foundation and the NIH for support We also thank P. Andrews, K. Bhatia, F. C. Howell, W. W. Howells, R. L. Kirk, E. Mayr, E. M. Prager, V. M. Sarich, G. Stringer and T. White for discussion and help in obtaining placentas. 29. Horai, S., Gojobori, T. & Matsunaga, E. Jap. J. Genet. 61, 271-275 (1986).

the overall extent of nuclear DNA diversity in both human and

African ape populations. By comparing the nuclear and mitochondrial DNA diversities, it may be possible to find out

whether a transient or prolonged bottleneck in population size accompanied the origin of our species15. Then a fuller interes

action between palaeoanthropology, archaeology and molecul

lar biology will allow a deeper analysis of how our species

30. Stoneking, M., Bhatia, K. & Wilson, A. C. Cold Spring Harb. Symp. quant. Biol. 51, 433-4 31. Wallace, D. C., Garrison, K. & Knowler, W. C. Am. J. phys. Anthrop 68, 149-155 (1985) 32. Jones, R. A. Rev. Anthrop. 8, 445-466 (1979). 33. Wolpoff, M. H., Wu, X. Z. & Thorne, A. G. in Origins of Modern Humans: A World Sur of the Fossil Evidence (eds Smith, F. H. & Spencer, F.) 411-484 (Liss, New York, 198 34. Owen, R. C. in Origins of Modern Humans: A World Survey of the Fossil Evidence (e Smith, F. H. & Spencer, F.) 517-564 (Liss, New York, 1984). 35. Mourant, A. E. et al. The Distribution of the Human Blood Groups and Other Polymorp (Oxford University Press, Oxford, 1978).

36. Murray, J. C. et al. Proc. natn. Acad. Sci. U.S.A. 81, 3486-3490 (1984).

37. Cooper, D. N. & Schmidtke, J. Hum. Genet. 66, 1-16 (1984). 38. Cooper, D. N. et al. Hum. Genet. 69, 201-205 (1985). Hill, A. V. S., Nicholls, R. D., Thein, S. L. & Higgs, D. R. Cell 42, 809-819 (1985). 40. Chapman, B. S., Vincent, K. A. & Wilson, A. C. Genetics 112, 79-92 (1986). 41. Chakravarti, A. et al Proc. notn. Acad Sci. U.S.A. 81, 6085-6089 (1984). 42. Chakravarti, A. et al. Am. J. hum. Genet. 36, 1239-1258 (1984). 43. Rightmire, G. P. in Origins of Modern Humans: A World Survey of the Fossil Evidence ( Smith, F. H. & Spencer, F.) 295-326 (Liss, New York, 1984).

44. Brauer, G. Courier Forsch. Int. Senckenberg 69, 145-165 (1984). F. H. & Spencer, F.) 327-410 (Liss, New York, 1984).

45. Braver, G. in Origins of Modern Humans: A World Survey of the Fossil Evidence (eds Se 46. Isaac, G. Phil Trans. R. Soc. Lond. B292, 177-188 (1981). 47. Clark, J. D. Proc. Br. Acad. Lond. 67, 163-192 (1981). 48. Pope, G. G. Proc. nam. Acad. Sci. U.S.A. 80, 4988-4992 (1983). 49. Coon, C. S. The Origin of Races (Knopf, New York, 1962).

 Brown, W. M., Prager, E. M., Wang, A. & Wilson, A. C. J. molec. Evol 18, 225-239 (1982). 6. Sibley, C. G. & Ahlquist, J. E. J. molec. Evol. 20, 2-15 (1984). 7. Bianchi, N. O., Bianchi, M. S., Cleaver, J. E. & Wolff, S. J. molec. Evol. 22, 323-333 (1985). 8. O'Brien, S. J. et al. Nature 317, 140-144 (1985).

(1983).

Vol. 2 Biogenesis (eds E. Quagliariello, E. C. Slater, F. Palmieri, C. Saccone & A. M. Kroon) 49-60 (Elsevier, New York, 1986). 21. Johnson, M. J. et al. J. molec. Evol. 19, 255-271 (1983). 22. Cann, R. L. Ph. D. Thesis, Univ. California, Berkeley, California (1982). 23. Stoneking, M., Bhatia, K. & Wilson, A. C. in Genetic Variation and Its Maintenance in

27. Whittam, T. S., Clark, A. G., Stoneking, M., Cann, R. L. & Wilson, A. C. Proc. natn. Acad.

28. Greenberg, B. D., Newbold, J. E. & Sugino, A. Gene 21, 33-49 (1983).

Tropical Populations (eds Roberts, C. F. & Destefano, G.) 87-100 (Cambridge University Press, Cambridge, 1986). 24. Cann, R. L., Brown, W. M. & Wilson, A. C. Genetics 106, 479-499 (1984). 25. Anderson, S. et al. Nature 290, 457-465 (1981). 26. Nei, M. & Tajima, F. Genetics 105, 207-217 (1983).

9. Cavalli-Sforza, L. L. Proc. R. Soc. Lond. B164, 362-379 (1966). Cavalli-Sforza, L. L. Sci. Am. 231(3), 81-89 (1974). 11. Nei, M. & Roychoudhury, A. K. Fool Biol 14, 1-59 (1982). 12. Nei, M. in Population Genetics and Molecular Evolution (eds Ohta, T. & Aoki, K.) 41-64 (Japan Sci. Soc. Press, Tokyo, 1985). 13. Constans, J. et al. Am. J. phys. Anthrop. 68, 107-122 (1985). 14. Wainscoat, J. S. et al. Nature 319, 491-493 (1986). 15. Wilson, A. C. et al Biol. J. Linn. Soc. 26, 375-400 (1985). 16. Olivo, P. D., Van de Walle, M. J., Laipis, P. J. & Hauswirth, W. W. Nature 306, 400-402 17. Brown, W. M. Proc. natn. Acad. Sci. U.S.A. 77, 3605-3609 (1980). 18. Monnat, R. J. & Loeb, L. A. Proc. natn. Acad. Sci. U.S.A. 82, 2895-2899 (1985). 19. Monnat, R. J., Maxwell, C. L. & Loeb, L. A. Cancer Res 45, 1809-1814 (1985). 20. Hauswirth, W. W. & Laipis, P. J. in Achievements and Perspectives in Mitochondrial Research,

50. Stringer, C. B. Courier Forsch. Inst. Senckenberg 69, 131-143 (1984). 51. Andrews, P. Courier Forsch. Inst. Senckenberg 69, 167-175 (1984). 52. Andrew, P. New Scient. 102, 24-26 (1984). 53. Stringer, C. B., Hublin, J. J. & Vandermeersch, B. in Origins of Modern Humans: A Survey of the Fossil Evidence (eds Smith, F. H. & Spencer, F.) 51-136 (Liss, New

found that 12 of them bear restriction site markers known21 to occur exclusively or predominantly in native sub-Saharan Africans (but not in Europeans, Asians or American Indians) nor, indeed, in all such Africans). The mtDNA types in these

## Africa 14 0.36 Asia 27 0.21 New Guinea 7

‡ Assuming that Africa is the source, there is only	one African cluster.
ES	33
following features with Fig. 3. (1) two primal composed entirely of Africans, the other included populations studied; and (2) each population ple lineages connected to the tree at widely distince submission of this manuscript, Horai of for our samples of African and Caucasian population of a Japanese population by another in	stems from multi- spersed positions. et al. <sup>29</sup> built a tree pulations and their

## African origin as a group are more variable ( $\delta_x = 0.47$ ) than other groups.

Multiple lineages per race

Fig. 3).

shares these two features.

African population has multiple origins—can be illustrated most simply with the New Guineans. Take, as an example, mtDNA type 49, a lineage whose nearest relative is not in New Guinea, but in Asia (type 50). Asian lineage 50 is closer genealogically to this New Guinea lineage than to other Asian mtDNA lineages. Six other lineages lead exclusively to New Guinean mtDNAs, each originating at a different place in the tree (types 12, 13, 26-29, 65, 95 and 127-134 in Fig. 3). This small region of New Guinea (mainly the Eastern Highlands Province) thus seems to have been colonised by at least seven maternal lineages (Tables 2 and 3). lineages that colonised Australia, Asia and Europe (Tables 2 and 3). Each estimate is based on the number of region-specific clusters in the tree (Fig. 3, Tables 2 and 3). These numbers,

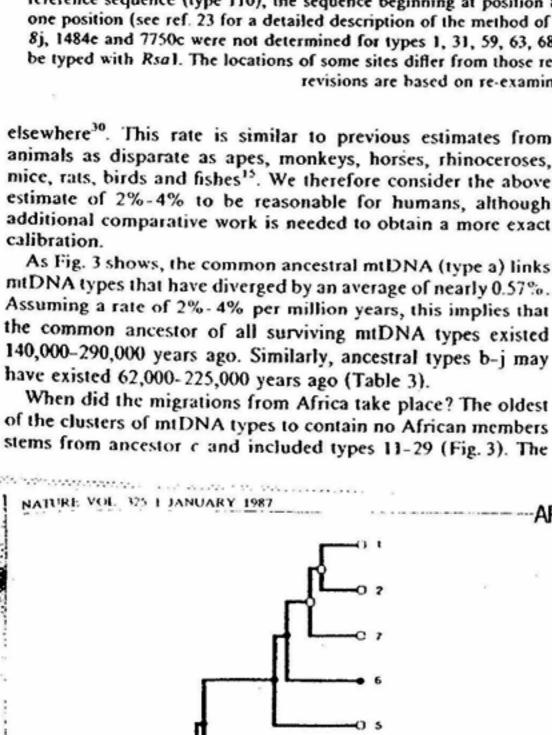
# per million years; a detailed account of this calculation appears

115; 4769a 110, 4793e 76 85, 5176a 21 25 54-57; 5261e 118; 52691 71; 535# 17-19 TRP IRNA: 55381 130; 5552c 10 L. ORIGIN: 5742i 15 44; 5754i (5755) 81 CO I: 5978a 52: 5984b/5983g 87: 5985W5983g 133 134; 5996a 87: 6022a 74; 61661 (6168) 63, 6211g 37-41 82; 6260e 95; 6356c 5 9; 6377c 68; 64091 (7854) 68; 6501i 78; 6610g 12; 66990 (8719.8723) 78; 6871g 18 19; 6915k 5 49; 6931g 48. 6957e 44 45; 7025a 75 86 93 94 103-112; 7055a 3 6 7; 7241k 72; 7335l 6 7 30 92 120; 7347e 124 SER IRNA: 74611 71; 7474a 121 CO II: 76171 55 73; 7750c 3 9 37 40 41 98, 7859j 120; 7970g 10 24 71-74 77; 8074a 68; 8112i 1; 8150i 1 126 . 8165e 10; 8249 b (8250) /8250e 1 61 132 LYS tRNA: 8299k 126 ATP 8: 8391e 31 120; 8515c 27 ATP 6: 8592j 21 23; 8783g 6 7; 8852l (8854) 8856) 109 110: 8994e60 61; 9009a 15 16; 9053t 52 91-93 95-102 112; 9070 5-7; 9150 7 103 104 CO III: 9266e 28 49 67, 9294e 25; 9342e 27; 9380! 90; 9429k 56 57: 9553e 37 117; 9714e 98; 9746k 62; 97511 48; 9859g 86 GLY IRNA: 10028a 126 ND 3: 10066l 45; 10084l 45; 10352a 110; 10364e 96; 10394c 1-29 39-45 49 50 84-89 94-100 111 120-123 ARG IRNA: 10413a (10536) 16 46 56 7 85 98 ND 4L: 10644k 19 20; 10689e 96; 10694a 104; 10725e 22 46 96 ND 4; 10806g 1-7 50 83 124; 108931 34; 11146c 61 110; 11167i 6 7; 11329e (11690) 2 4; 11350a 15 37 110; 11806a 65; 11922j 7; 12026h 124 ND 5: 12345k (12350,12528) 88 127-134; 12406h 91-93; 12560a 19 52 93; 12795j (12798,12806.13374) 40 90 98 99; 12810k 5-7; 12925g 7; 12990a (13642) 27; 13004j (13018,13182,13194) 20; 13031g 1-6 9-16 18-23 26-35 37 39 40 42 43 45-64 82 83 92 93, 13051e 23 38-41; 13068a 128; 13096k 26; 13100k 107;

13103g 7; 132081 4 5 7 15 75 88; 13268g 65; 13367b 40 88 90 99; 134041 8 9 79 : 13635 (13641) 39 115; 13702e 38-41; 14015a 14; 14050 (14366) 125 ND 6;

14279e 1; 14279j 23; 14322a 92; 14385c 10; 14509a 68 123-126; 14567i 68;

apparent age of this cluster (calculated in Table 3) is 90,000 } 180,000 years. Its founders may have left Africa at about that



variability and that a midpoint rooting of their tree leads to an African origin. Greenberg et al.28 sequenced the large noncoding region, which includes the displacement loop (D loop), from four Caucasians and three Black Americans. A parsimony tree for these seven D loop sequences, rooted by the midpoint method, appears in Fig. 4. This tree indicates (1) a high evolutionary rate for the D loop (at least five times faster than other other mtDNA regions), (2) a greater diversity among Black American D loop

migration, so the direct relationship found between time and mutational distance for mtDNA would not be expected for genetic distances based on nuclear DNA. But studies based on polymorphic blood groups, red cell enzymes, and serum proteins show that (1) differences between racial groups are smaller than those within such groups and (2) the largest gene frequency

15 5 36 Assuming that the mtDNA divergence rate is 2-4% per million years 15,30

arose.

1

54. Horai, S., Goobori, T. & Matsunaga, E. Hum. Genet. 68, 324-332 (1984). 55. Bonne-Tamir, B. et al. Am. J. hum. Genes. 38, 341-351 (1986). 56. Horai, S. & Matsunaga, E. Hum. Genet. 72, 105-117 (1986). 57. Avise, J. C., Neigel, J. E. & Arnold, J. J. molec. Evol. 20, 99-105 (1984). 58. Farris, J. S. Am. Nat. 106, 645-668 (1972).

<sup>12</sup> people are 2-7, 37-41 and 82 (see below). Methods used to purify mtDNA and more detailed ethnographic information on the first four groups are as described17,22; the New Guineans