another way to view basins such as, for example, the yellow basin of the pendulum example is as if it were a lake (the basin cell) with three rivers draining into the lake.

This intriguing phenomenon of Wada basins can be found in many applications. Many physical, biological, or economic systems can be described by differential equations analogous to those of the pendulum. For various choices of parameter values for the system, there are likely to be several coexisting basins. If the boundaries are fractal, it is likely that basin cells can be found. Here we showed three coexisting basins, but there is no limit to the number

of coexisting basins that are Wada basins, all with the same boundary. Such boundaries are uncertain; every boundary point is arbitrarily near every basin. Every boundary point can be perturbed arbitrarily slightly into any of the basins.

## REFERENCES AND NOTES

1. C. Grebogi, E. Ott, J. A. Yorke, *Science* **238**, 632 (1987).
2. E. G. Gwinn and R. M. Westervelt, *Phys. Rev. A* **33**, 4143 (1986); S. W. McDonald, C. Grebogi, E. Ott, J. A. Yorke, *Physica D* **17**, 125 (1985); C. Mira, *C. R. Acad. Sci. Paris A* **288**, 591 (1979); F. C. Moon and G.-X. Li, *Phys. Rev. Lett.* **55**, 1439 (1985); S. Takesue and K. Kaneko, *Prog. Theor. Phys.* **71**, 35 (1984).
3. E. Kasner and J. R. Newman, *Mathematics and the Imagination* (Microsoft Press, Redmond, WA, 1989).
4. J. G. Hocking and G. S. Young, *Topology* (Addison-Wesley, Reading, MA, 1961).
5. J. Kennedy and J. A. Yorke, *Physica D* **51**, 213 (1991).
6. H. E. Nusse and J. A. Yorke, *ibid.*, in press.
7. _____, *Ergodic Theory Dyn. Syst.*, in press.
8. _____, *Dynamics: Numerical Explorations* (Applied Mathematical Sciences 101, Springer-Verlag, New York, 1994).
9. The computer-assisted pictures were made with the program Dynamics (*8*). We thank M. Roberts and J. Tempkin for their comments. This work was supported in part by the Department of Energy (Scientific Computing Staff Office of Energy Research) and the National Science Foundation, and the W. M. Keck Foundation supported our Chaos Visualization Laboratory.

# Global Patterns of Linkage Disequilibrium at the CD4 Locus and Modern Human Origins

S. A. Tishkoff, E. Dietzsch, W. Speed, A. J. Pakstis, J. R. Kidd, K. Cheung, B. Bonné-Tamir, A. S. Santachiara-Benerecetti, P. Moral, M. Krings, S. Pääbo, E. Watson, N. Risch, T. Jenkins, K. K. Kidd

Haplotypes consisting of alleles at a short tandem repeat polymorphism (STRP) and an Alu deletion polymorphism at the CD4 locus on chromosome 12 were analyzed in more than 1600 individuals sampled from 42 geographically dispersed populations (13 African, 2 Middle Eastern, 7 European, 9 Asian, 3 Pacific, and 8 Amerindian). Sub-Saharan African populations had more haplotypes and exhibited more variability in frequencies of haplotypes than the Northeast African or non-African populations. The Alu deletion was nearly always associated with a single STRP allele in non-African and Northeast African populations but was associated with a wide range of STRP alleles in the sub-Saharan African populations. This global pattern of haplotype variation and linkage disequilibrium suggests a common and recent African origin for all non-African human populations.

Several models for the origin of *Homo sapiens sapiens* have been proposed. The "multiregional origin" model suggests that there was no single origin for all modern humans (*1, 2*). After the radiation of *Homo erectus* from Africa into Europe and Asia 800,000 to 1.8 million years ago (*3*), there was a continuous transition among regional populations from *H. erectus* to *H. sapiens*. Such "parallel evolution" among geographically dispersed populations could have been achieved by considerable amounts of gene flow between populations (*1, 2*). By contrast, the "out of Africa" model suggests that all non-African human populations descend from an anatomically modern *H. sapiens* ancestor that evolved in Africa approximately 100,000 to 200,000 years ago and then spread and diversified throughout the rest of the Earth, supplanting any *Homo* populations still present outside of Africa (*1,

4*). Migration out of Africa may have occurred in a single or in multiple waves (*5*).

The best-known genetic evidence used to support the out of Africa hypothesis has come from studies of mitochondrial DNA (mtDNA) in which it was proposed that all modern mtDNA can be traced back through the maternal lineage to a single ancestor that existed in Africa between 100,000 and 300,000 years ago (*6, 7*). The analysis and interpretation of these data have continued to be debated (*8*). Recent mtDNA (*9*) and Y chromosome (*10*) studies support the original findings of a recent origin of all modern humans. We present data from the nuclear autosomal genome that strongly support the out of Africa model of human origins and provide a different and independent estimate, based on linkage disequilibrium, of the recency of the emigration from Africa.

## Genetic Systems Studied

We studied alleles from two tightly linked markers, located ~9.8 kb apart, within non-coding regions of the CD4 gene on the short arm of chromosome 12 (*11–13*) (Fig. 1). These polymorphic markers are of two types that evolve with differing rates. The first is a short tandem repeat polymorphism (STRP). This class of markers consists of tandemly repeated blocks of two to five nucleotides; STRPs often have multiple alleles (defined by the number of repeats) and moderate to high mutation rates (*14*). Many researchers consider them particularly useful as markers for reconstructing recent evolutionary history (*15*). The STRP at the CD4 locus consists of the pentanucleotide sequence TTTTC repeated between 4 and 15 times (*12, 13*); the products (including flanking sequence) of the polymerase chain reaction (PCR) range in size from 80 base pairs (bp) for a 4-repeat allele to 135 bp for a 15-repeat allele (*16*). Most of the 12 alleles seen in humans are found primarily in Africa. Outside of Africa only three alleles (the 85-, 90-, and 110-bp alleles) ever occur at a frequency greater than 10%. Genotype frequencies for all populations are close to predicted Hardy-Weinberg expectations. We have also amplified the CD4 STRP in common chimpanzee ($n = 22$), pygmy chimpanzee ($n = 5$), gorilla ($n = 5$), orangutan ($n = 3$), and gibbon ($n = 4$). Most hominoid species are polymorphic, but alleles range only from three to six repeats (75 to 90 bp) (*17*).

The second polymorphism results from the deletion of 256 bp of a 285-bp Alu element (Fig. 1) (*13*). This type of mutation is unlikely to have occurred more than once; DNA sequence analysis of several Alu deletion chromosomes from African and non-African individuals (*18*) revealed that all chromosomes contain the identical deletion, so that common ancestry can be assumed. The Alu deletion allele [Alu(−)] was typed through use of published primers and protocols (*13*); it was found to be rare or

absent in nine Asian, eight New World, and three Pacific island populations examined and occurs at approximately a 25 to 30% frequency in the seven European and two Middle Eastern populations examined. In 13 African populations the frequency of the Alu(−) allele ranges from 7 to 28%. Genotype frequencies for all populations are close to predicted Hardy-Weinberg expectations. The Alu(−) allele has not been detected in chimpanzees (n = 23), gorillas (n = 2), orangutans (n = 3), or gibbons (n = 4), indicating that the full-length Alu[Alu(+)] is the ancestral state of this polymorphism and that the deletion event likely occurred after the divergence of humans from the great apes 4 to 6 million years ago (19).

## Haplotype Analysis

Both markers were typed in more than 1600 unrelated individuals from 42 geographically dispersed populations (Table 1) (20). The particular combination of alleles found together on a single chromosome is called a haplotype. Haplotype frequencies were estimated either with the HAPLO program (21) or by simple gene counting when ambiguous samples were haplotyped directly with an allele-specific primer and long-range PCR (22) (Fig. 1). Mean haplotype frequencies and haplotype heterozygosity levels of populations grouped by geographic region are shown in Table 1. Haplotype frequencies for several individual populations are represented graphically in Fig. 2. Only the 85-bp Alu(+) and the 110-bp Alu(+) haplotypes occur at a frequency greater than 10% in all non-African populations. The 90-bp Alu(−) haplotype occurs at a high frequency (23 to 34%) in the Middle Eastern and European populations and is rare or absent in all Asian, Pacific island, and New World populations (0 to 9% frequency) (Fig. 2, A to D). Among non-African populations heterozygosity is

highest in the Middle East (0.75 in the Israeli Druze), decreases from west to east across Eurasia, and is lowest in the New World (0.37 in the Brazilian Ticuna) and Australo-Melanesia (0.28 in the Nasioi Melanesians) (Table 1).

**Fig. 1.** Location and distance of polymorphic markers at the CD4 locus on 110-bp Alu(+) and 90-bp Alu(−) chromosomes. Solid boxes represent exons 1 to 3 of the CD4 gene. The Alu element is represented by a hatched box, and the open box shows the location of the deletion that encompasses 239 bp of the Alu element and 17 bp of flanking sequence (13). Arrows show the location and direction of primers used for amplification by PCR of the STRP (A and B) (12) and Alu polymorphism (C and D) (13). Primer E (5′-GCGCAAGATGCACAGCCTG-3′) is designed to be specific for the Alu(−) allele such that long-range PCR amplification with primer A occurs only on chromosomes that contain the Alu deletion (22). A second round of amplification with primers A and B with 1 µl of the 9.9-kb PCR product allows determination of the size of the STRP allele on the Alu(−) chromosome.
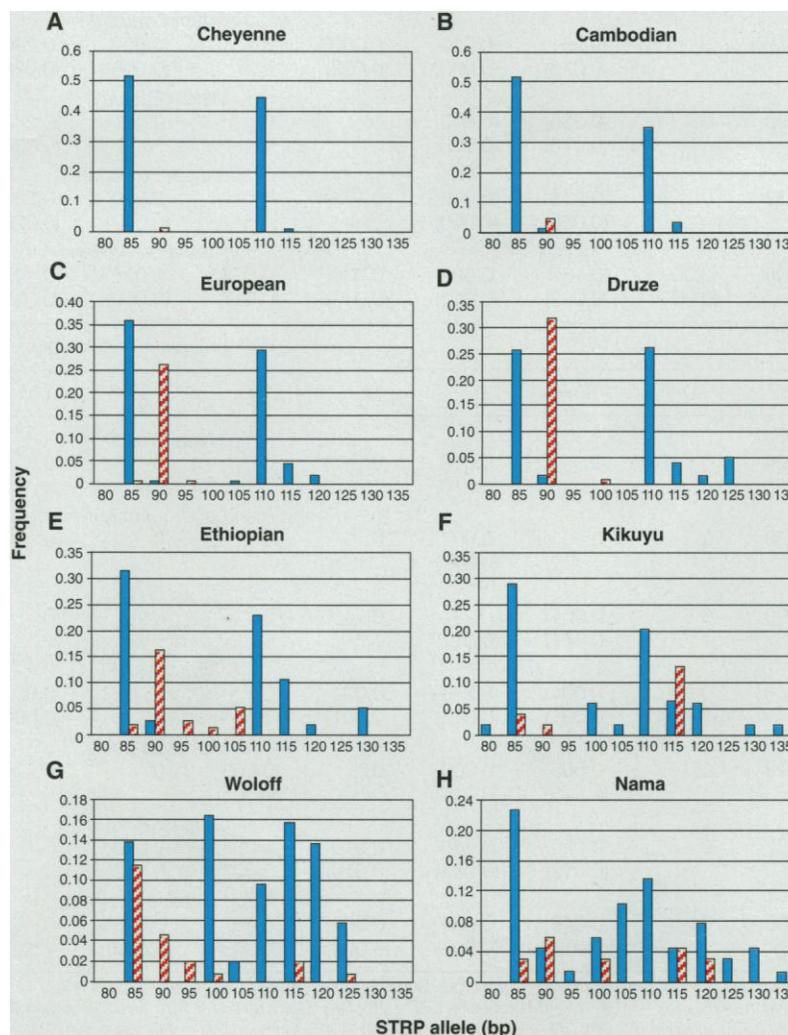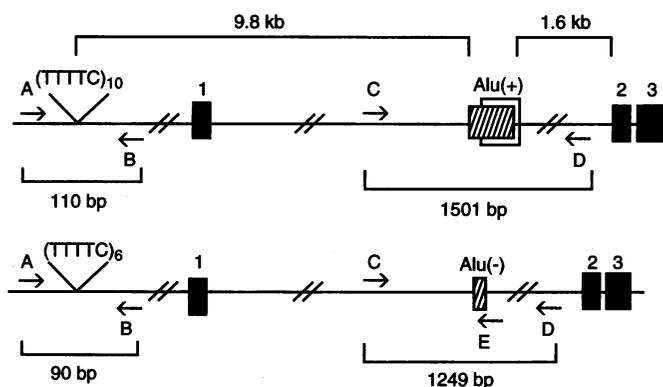
S. A. Tishkoff, W. Speed, A. J. Pakstis, J. R. Kidd, K. Cheung, and K. K. Kidd are in the Department of Genetics, Yale University School of Medicine, 333 Cedar Street, New Haven, CT 06520–8005, USA. E. Dietzsch and T. Jenkins are in the Department of Human Genetics, School of Pathology, South African Institute for Medical Research, and University of the Witwatersrand, Post Office Box 1038, Johannesburg 2000, South Africa. B. Bonné-Tamir is in the Department of Human Genetics, Sackler Faculty of Medicine, Tel Aviv University, 69978 Ramat Aviv, Israel. A. S. Santachiara-Benerecetti is in the Buzzati Traverso Department of Genetics and Microbiology, University of Pavia, Via Abbiategrasso 207, 27100 Pavia, Italy. P. Moral is in the Department of Animal Biology, University of Barcelona, Avinguda Diagonal 645, 08028 Barcelona, Spain. M. Krings and S. Pääbo are at the Institute of Zoology, University of Munich, P.O. Box 202136, D-80021 Munich, Germany. E. Watson is at the Institute of Zoology, University of Munich, Luisenstrasse 14, 80333 Munich 2, Germany, and School of Biological Sciences, Massey University, Palmerston North, New Zealand. N. Risch is in the Department of Genetics, Stanford University Medical School, Stanford, CA 94305, USA.

**Fig. 2.** CD4 STRP-Alu(+/−) haplotype frequencies. Solid blue bars, Alu(+) chromosomes; striped red bars, Alu(−) chromosomes. Haplotype frequencies are shown for four diverse non-African (**A** to **D**) and four diverse African (**E** to **H**) populations selected from the geographic regions shown in Table 1.

Haplotype frequencies in Egyptians do not differ significantly from those in Middle Eastern populations (23), consistent with linguistic evidence suggesting a recent common ancestry. The Ethiopian Jewish (Fig. 2E) and Somali populations show a pattern of haplotype frequencies similar to, but significantly different from, the non-African populations, with slightly more haplotypes (mean heterozygosity of 0.80) (24).

The pattern of haplotype frequencies in sub-Saharan African populations differs from that in the Northeast African and non-African populations (Fig. 2, F to H). Compared to non-African populations, the sub-Saharan populations have more haplotypes that exist at a higher frequency; nine haplotypes occur at a frequency greater than 10% in at least one population, and heterozygosity ranges from 0.82 to 0.90. Frequencies of haplotypes vary considerably among populations as illustrated by the Kikuyu, Woloff, and Nama populations (Fig. 2, F to H).

The nonrandom association between the CD4 STRP alleles and the Alu(−) allele is referred to as linkage disequilibrium

**Table 1.** Unweighted mean CD4 STRP-Alu frequency estimates and standard errors (in parentheses) for seven regions of the world. Haplotypes are represented by the STRP allele and the Alu allele. n, number of populations; 2N, number of chromosomes sampled. No 135-bp Alu(−) haplotypes were observed. Estimated heterozygosities (Het.) for the regions were calculated for each population as 1 - $\Sigma$[(frequency of each haplotype)$^2$] and then averaged across populations.

| | | STRP (bp)–Alu(+) haplotypes | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| n | 2N | 80+ | 85+ | 90+ | 95+ | 100+ | 105+ | 110+ | 115+ | 120+ | 125+ | 130+ | 135+ |
| *New World** | | | | | | | | | | | | | |
| 8 | 704 | 0 | 0.553 (0.047) | 0.001 (0.001) | 0 | 0.001 (0.001) | 0.001 (0.001) | 0.421 (0.051) | 0.004 (0.004) | 0.001 (0.001) | 0 | 0 | 0 |
| *Pacific island and Australo-Melanesian†* | | | | | | | | | | | | | |
| 3 | 130 | 0 | 0.727 (0.094) | 0.013 (0.007) | 0 | 0 | 0.007 (0.007) | 0.227 (0.088) | 0.013 (0.007) | 0 | 0 | 0 | 0 |
| *Asian‡* | | | | | | | | | | | | | |
| 9 | 600 | 0 | 0.603 (0.033) | 0.011 (0.005) | 0.003 (0.002) | 0.003 (0.003) | 0.004 (0.003) | 0.310 (0.033) | 0.023 (0.011) | 0.002 (0.001) | 0 | 0 | 0 |
| *European§* | | | | | | | | | | | | | |
| 7 | 658 | 0 | 0.361 (0.038) | 0.010 (0.003) | 0.003 (0.002) | 0 | 0.007 (0.006) | 0.289 (0.030) | 0.037 (0.009) | 0.017 (0.005) | 0 | 0 | 0 |
| *Middle Eastern‖* | | | | | | | | | | | | | |
| 2 | 198 | 0 | 0.335 (0.075) | 0.015 (0.005) | 0 | 0 | 0 | 0.240 (0.030) | 0.045 (0.005) | 0.020 (0.000) | 0.025 (0.025) | 0 | 0 |
| *Northeast African¶* | | | | | | | | | | | | | |
| 3 | 232 | 0 | 0.273 (0.058) | 0.010 (0.010) | 0 | 0 | 0.020 (0.012) | 0.280 (0.026) | 0.113 (0.047) | 0.063 (0.038) | 0.006 (0.006) | 0.027 (0.007) | 0 |
| *Sub-Saharan African#* | | | | | | | | | | | | | |
| 10 | 806 | 0.002 (0.002) | 0.187 (0.022) | 0.023 (0.012) | 0.030 (0.013) | 0.074 (0.017) | 0.026 (0.009) | 0.172 (0.026) | 0.156 (0.026) | 0.107 (0.019) | 0.024 (0.007) | 0.021 (0.007) | 0.005 (0.002) |

| | | STRP (bp)–Alu(−) haplotypes | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| n | 2N | 80− | 85− | 90− | 95− | 100− | 105− | 110− | 115− | 120− | 125− | 130− | Het. |
| *New World** | | | | | | | | | | | | | |
| 8 | 704 | 0 | 0 | 0.018 (0.010) | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.48 (0.018) |
| *Pacific island and Australo-Melanesian†* | | | | | | | | | | | | | |
| 3 | 130 | 0 | 0 | 0.007 (0.007) | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.40 (0.080) |
| *Asian‡* | | | | | | | | | | | | | |
| 9 | 600 | 0 | 0.001 (0.001) | 0.037 (0.012) | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.52 (0.034) |
| *European§* | | | | | | | | | | | | | |
| 7 | 658 | 0 | 0.001 (0.001) | 0.276 (0.017) | 0.001 (0.001) | 0 | 0 | 0.001 (0.001) | 0 | 0 | 0 | 0 | 0.69 (0.016) |
| *Middle Eastern‖* | | | | | | | | | | | | | |
| 2 | 198 | 0 | 0.005 (0.005) | 0.305 (0.015) | 0 | 0.005 (0.005) | 0 | 0 | 0 | 0 | 0 | 0 | 0.73 (0.025) |
| *Northeast African¶* | | | | | | | | | | | | | |
| 3 | 232 | 0 | 0.020 (0.012) | 0.160 (0.044) | 0.010 (0.010) | 0.004 (0.003) | 0.013 (0.013 | 0 | 0 | 0 | 0 | 0 | 0.78 (0.025) |
| *Sub-Saharan African#* | | | | | | | | | | | | | |
| 10 | 806 | 0 | 0.046 (0.011) | 0.042 (0.008) | 0.009 (0.003) | 0.007 (0.003) | 0.003 (0.003) | 0 | 0.054 (0.014) | 0.007 (0.004) | 0.005 (0.002) | 0.001 (0.001) | 0.85 (0.010) |

*Ticuna (Brazil; 2N = 124), Karitiana (Brazil; 2N = 100), Surui (Brazil; 2N = 78), Quechua (Peru; 2N = 44), Maya (Yucatan; 2N = 98), Jemez Pueblo (USA; 2N = 86), Pima (USA; 2N = 70), Cheyenne (USA; 2N = 104).   †Micronesian (2N = 46), New Guinean (2N = 40), Nasioi (Bougainville; 2N = 44).   ‡Malaysian (2N = 24), Atayal (Taiwan; 2N = 74), Ami (Taiwan; 2N = 68), Kochari (Assam; 2N = 32), Cambodian (2N = 48), Chinese (2N = 88), Japanese (2N = 86), Yakut (Siberia; 2N = 72), Asiatic Indian (2N = 108).   §Adygei (Caucasus; 2N = 98), Europeans (mixed; 2N = 178), Finns (2N = 68), Danes (2N = 38), Basque (Spain; 2N = 124), Ashkenazi Jews (2N = 102), Roman Jews (2N = 50).   ‖Druze (Israel; 2N = 96), Yemenite Jews (2N = 102).   ¶Egyptian (2N = 68), Ethiopian Jews (2N = 114), Somali (2N = 50).   #Kikuyu (Kenya; 2N = 44), Yoruba (Nigeria; 2N = 36), Wolof (Senegal; 2N = 102), Mbuti (Zaire; 2N = 74), Biaka (Central African Republic; 2N = 106), Herero Bantu-speakers (Namibia; 2N = 98), Zu/Wasi !Kung San (Namibia; 2N = 88), Sekele !Kung San (Namibia; 2N = 104) Bantu-speakers (mixed South African; 2N = 88), Nama (Namibia; 2N = 66).

and can be estimated from the parameter $\delta$ (25), where $\delta_{(-,x)} = [p_{(-)}(x) - p_{(+)}(x)]/[1 - p_{(+)}(x)]$, $p_{(-)}(x)$ is the proportion of STRP allele $x$ on Alu($-$) chromosomes, and $p_{(+)}(x)$ is the proportion of STRP allele $x$ on Alu($+$) chromosomes (Table 2). Outside of Africa only one value is large: $\delta_{(-,90)}$ = 0.98 for combined non-African populations, a value close to the maximum of 1.0, indicating numerically the obvious finding (Table 1) that the Alu deletion allele almost always occurs on a chromosome with the 90-bp STRP allele. The largest $\delta_{(-,x)}$ values for the Egyptian, Ethiopian, and Somali populations are also with the 90-bp STRP allele [$\delta_{(-,90)}$ values of 1.0, 0.61, and 0.67, respectively). Among the sub-Saharan African populations, examples can be found of the Alu($-$) allele positively associated with several STRP alleles. However, only the 90-bp STRP allele consistently (in 9

outof 10 populations) exhibits a positive association with the Alu($-$) allele; it also has the highest median $\delta_{(-,x)}$ value (0.24), suggesting that the Alu deletion originally occurred on a chromosome containing a 90-bp STRP allele and there has not been sufficient time for complete equilibrium to be attained. Figure 3 summarizes the three categories of $\delta_{(-,90)}$ values: strong in non-Africans (and Egyptians), intermediate in Ethiopian and Somali populations, and still present but weak in sub-Saharan African populations.

**Table 2.** $\delta_{(-,90)}$ values for the three CD4 STRP alleles most frequently associated with the Alu($-$) allele. African populations are listed individually, and all non-African populations containing the Alu($-$) allele are combined for a single estimate. A value of $+1$ indicates complete association with the Alu($-$) allele, 0 indicates no association, and a dash indicates that the STRP allele is more common on Alu($+$) than on Alu($-$) chromosomes. No other STRP allele had a $\delta_{(-,x)}$ value greater than 0.20 in any population.

| Population | STRP allele (bp) | | |
|---|---|---|---|
| | 85 | 90 | 115 |
| *Non-African* | | | |
| Combined | – | 0.98 | – |
| *Northeast African* | | | |
| Egyptian | – | 1 | – |
| Ethiopian Jewish | – | 0.61 | – |
| Somalian | 0.18 | 0.67 | – |
| *Sub-Saharan African* | | | |
| Woloff | 0.41 | 0.22 | – |
| Yoruba | – | 0.20 | 0.18 |
| Mbuti | 0.46 | 0.19 | 0.00 |
| Biaka. | 0.38 | – | – |
| Kikuyu | – | 0.11 | 0.64 |
| Bantu-speakers | – | 0.26 | 0.29 |
| Herero | 0.01 | 0.38 | 0.13 |
| Nama | – | 0.27 | 0.18 |
| Sekele !Kung San | – | 0.26 | 0.13 |
| Zu/wasi !Kung San | – | 0.19 | 0.03 |

## Implications of Linkage Disequilibrium

Distributions of haplotype frequencies and linkage disequilibrium in extant populations are the result of several processes: mutation at the STRP, recombination between the two sites, random genetic drift, and gene flow among populations. We begin our analysis by reconstructing a plausible scenario to explain the contrasting patterns of linkage disequilibrium in Africans and non-Africans.

The CD4 STRP apparently started as a small-sized repeat, still seen in nonhuman primates, and has accumulated variation in the human lineage during the 4 to 6 million years since humans diverged from the great apes (19). At some time during that period the Alu was partially deleted from a single chromosome which, through genetic drift, eventually became common. Initially, there must have been complete linkage disequilibrium between the Alu($-$) allele and the "ancestral" nearby STRP allele (most likely a 90-bp allele, as noted above). Over time, because of the effects of mutation and recombination, that linkage disequilibrium was largely lost in the ancestral African population.

Then, through random genetic drift, the Alu($-$) chromosome containing the 90-bp STRP allele became the predominant Alu($-$) chromosome in one or more populations in the Northeastern periphery of Africa. Further reduction in genetic variation most likely occurred when members of this ancestral population left Africa and populated the rest of the world, bringing with them the chromosome containing the 90-bp STRP allele and the Alu($-$) allele. The presence of almost complete linkage disequilibrium between these two alleles in

all non-African populations surveyed suggests that the Alu($-$) chromosome containing the 90-bp STRP allele was the only Alu($-$) chromosome to remain in the populations that left Africa. The migration from Africa was also accompanied by a further reduction in the number of Alu($+$) haplotypes; only those Alu($+$) chromosomes containing an 85- or 110-bp STRP allele remain at a significant frequency outside of Africa. This finding supports studies of both mtDNA and nuclear DNA, suggesting that the ancestral population that left Africa was small (26, 27), but contrasts with results from the major histocompatibility complex (MHC) locus, for which there remains considerable variation outside of Africa (28). These contrasting patterns may be due to selection at the MHC locus that favors maintenance of allelic lineages (29).

The intermediate pattern of genetic variation seen in the Ethiopian Jewish and Somali populations could be explained either by recent admixture with nearby Middle Eastern populations or a recent common ancestry with non-African populations (30). Because these two Northeastern African populations have only a subset of the haplotype diversity present in sub-Saharan Africa, admixture between sub-Saharan and Middle Eastern populations does not seem a good explanation of their haplotype frequencies; rather, these populations may represent the modern descendants of the ancestral population that spawned the migration from Africa.

The high level of haplotype diversity observed within all 10 sub-Saharan African populations of diverse geographic origin suggests that much of the variation observed today was present in the ancestral African population from which they descended, with differences among these pop-
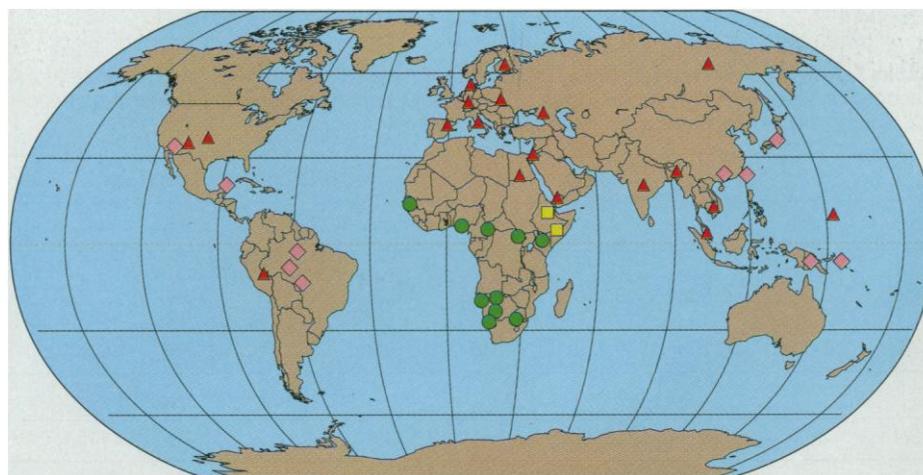


**Fig. 3.** Global distribution of populations used in the CD4 STRP-Alu haplotype study and the associated $\delta_{(-,90)}$ values for the 90-bp STRP and Alu($-$) alleles. (◆) Populations with no Alu($-$) allele detected; (▲) populations with $\delta_{(-,90)}$ values ranging from 0.90 to 1.0; (■) populations with $\delta_{(-,90)}$ values ranging from 0.63 to 0.80; (●) populations with $\delta_{(-,90)}$ values less than 0.38.

ulations due to recent effects of random genetic drift in each region. Thus, these data would suggest that sub-Saharan Africans have maintained a large effective population size and that no recent bottleneck events have occurred during the evolution of *H. sapiens* in Africa.

The much greater variation in both number and frequency of CD4 STRP-Alu haplotypes in sub-Saharan Africans is consistent with the out of Africa hypothesis (4). The clinal pattern of decreasing heterozygosities west to east outside of Africa is consistent with a migration event of modern humans out of Northeast Africa into the Middle East and Europe, and east into Asia, the Pacific islands, and the New World. In addition, all non-African populations share a similar pattern of haplotype variation and linkage disequilibrium, which implies that they all derive from a single ancestral gene pool. This pattern is similar to that seen in Northeast African populations but is distinct from that in all sub-Saharan African populations. Patterns of genetic variation distinguishing between African and non-African populations have been documented in data of both mtDNA (6, 31) and nuclear DNA (27, 32, 33). In addition, however, the retention of essentially complete linkage disequilibrium outside of Africa implies that the common origin of the non-African populations is recent because there has not been sufficient time for linkage disequilibrium to be diminished by the shuffling effects of mutation and recombination.

## Dating the Emergence from Africa

Assuming the above scenario, we can attempt to estimate how long ago the ancestors of modern non-Africans existed as a single distinct population with complete linkage disequilibrium of the 90-bp and Alu(−) alleles. All Alu(−) alleles appear to be descended from a single original mutation event that occurred in the human lineage a maximum of 4 to 6 million years ago (19). There has been enough time in sub-Saharan Africa since the origin of the Alu deletion for mutation and recombination to increase the allelic variation at the STRP on Alu(−) chromosomes, but not as much time since the introduction of the Alu(−) chromosome outside of Africa for variation to be regenerated.

To assess the relative effect of recombination and mutation on reducing levels of linkage disequilibrium between the Alu deletion and the STRP, we contrast the distributions of STRP alleles on Alu(+) and Alu(−) chromosomes in sub-Saharan African populations. High levels of recombination will make those distributions similar within populations. For comparison, we contrast the distributions of STRP alleles on chromosomes with the same Alu allele from different sub-Saharan populations. Genetic drift will tend to make those distributions different, whereas gene flow will tend to make them similar. By any of several statistical methods (34) we see high levels of similarity among Alu(+) chromosomes and among Alu(−) chromosomes from different populations and much less similarity between Alu(+) and Alu(−) chromosomes taken from the same or from different populations. These results imply a recent divergence of the populations or considerable gene flow among them (or both) and suggest that recombination is playing only a negligible role in equalizing the allele frequency distributions on Alu(+) versus Alu(−) chromosomes within populations. This conclusion is supported by the observation that no Alu(−) chromosomes (out of 132) in Sub-Saharan African population and only 1 chromosome (out of 286) outside of Africa carries a 110-bp allele, despite this allele having a frequency of 21% on Alu(+) chromosomes across Sub-Saharan Africa and 39% outside of Africa (Table 3). In addition, most of the variance of STRP alleles on Alu(−) chromosomes in the Northeast African and non-African populations consists of alleles smaller than 110 bp. This pattern of variation is consistent with a model of mutation involving primarily small changes in repeat size away from the ancestral 90-bp allele (35). Thus, both inside and outside Africa, the level of linkage disequilibrium now observed appears to result primarily from mutation occurring over time, with recombination having a smaller effect.

Random fluctuations in allele frequencies—random genetic drift—confound efforts to make precise estimates of how many generations have elapsed since complete disequilibrium existed. We cannot estimate precisely when the founding event out of Africa occurred without knowing the mutation and recombination rates, which are too small to be estimated by direct observation. However, by comparing the relative breakdown of disequilibrium in sub-Saharan and non-African populations, we can make a conservative estimate of the maximum time since the last common ancestor of the non-African populations (36). For our analysis we make the following assumptions: (1) The Alu deletion originally occurred in Africa as a single event after the separation of humans from the great apes and was originally associated with a single STRP allele (the 90-bp allele). (2) In the founding population for non-Africans, the Alu(−) chromosome was associated with only a single STRP allele (the 90-bp allele). (3) The mutation rate at the STRP is the same inside and outside of Africa. (4) There has been no back mutation to a 90-bp allele after a mutation away from the 90-bp allele.

**Table 3.** Relative weighted mean frequencies of the CD4 STRP alleles in each region standardized separately for Alu(+) and Alu(−) chromosomes. 2N, number of chromosomes sampled. Haplotype diversity $d(x)$ was calculated separately for Alu(+) and Alu(−) chromosomes as $1 - f_i^2$ where $f_i$ is the standardized frequency given in the table for STRP allele $i$.

| | | | | | STRP (bp)–Alu(+) haplotypes | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 2N | 80+ | 85+ | 90+ | 95+ | 100+ | 105+ | 110+ | 115+ | 120+ | 125+ | 130+ | 135+ | $d$[Alu(+)] |
| | | | | | | *Non-African* | | | | | | | |
| 2020 | 0 | 0.553 | 0.010 | 0.002 | 0.002 | 0.005 | 0.390 | 0.027 | 0.009 | 0.002 | 0 | 0 | 0.54 |
| | | | | | | *Sub-Saharan African* | | | | | | | |
| 674 | 0.001 | 0.203 | 0.028 | 0.040 | 0.090 | 0.030 | 0.210 | 0.200 | 0.140 | 0.030 | 0.030 | 0.004 | 0.84 |
| | | | | | STRP (bp)–Alu(−) haplotypes | | | | | | | | |
| 2N | 80− | 85− | 90− | 95− | 100− | 105− | 110− | 115− | 120− | 125− | 130− | 135− | $d$[Alu(−)] |
| | | | | | | *Non-African* | | | | | | | |
| 270 | 0 | 0.011 | 0.978 | 0.004 | 0.004 | 0 | 0.004 | 0 | 0 | 0 | 0 | 0 | 0.04 |
| | | | | | | *Sub-Saharan African* | | | | | | | |
| 132 | 0 | 0.264 | 0.258 | 0.068 | 0.041 | 0.008 | 0 | 0.271 | 0.045 | 0.038 | 0.007 | 0 | 0.78 |

(5) Except for the founding "event," mutation has had a larger effect than genetic drift or recombination on the proportions of most STRP alleles on Alu(+) and Alu(−) chromosomes (37).

Let $P_B$ represent the frequency of the progenitor STRP allele on non-African Alu(−) chromosomes. Assuming a mutation rate of μ per generation, the frequency of chromosomes still bearing the original mutation can be approximated by $e^{-N_B\mu}$, where $N_B$ is the number of generations since the founding. Similarly, within Africa, if $P_A$ is the frequency of chromosomes still bearing the original STRP allele, $P_A$ can be approximated by $e^{-N_A\mu}$. Thus, to obtain the relative age $R$ of the Alu(−) chromosome inside versus outside of Africa, we can calculate

$$R = \frac{\ln(P_A)}{\ln(P_B)} \cong \frac{N_A}{N_B} \qquad (1)$$

The relative frequencies of STRP alleles separately on Alu(+) and on Alu(−) chromosomes are given in Table 3 for the weighted averages of sub-Saharan and of non-African populations (38). Using the pooled frequencies in each region minimizes the effects of recent population-specific genetic drift. Outside Africa it is clear that the 90-bp allele was the progenitor on Alu(−) chromosomes, and its current frequency on such chromosomes is 264/270 = 0.978. Within sub-Saharan Africa, three alleles occur at nearly equal high frequencies on Alu(−) chromosomes—the 85-, 90-, and 115-bp alleles. As previously noted, patterns of linkage disequilibrium suggest that the ancestral Alu(−) chromosome in Africa contained a 90-bp STRP allele (Table 2). The high frequency of the 85-bp allele on Alu(−) chromosomes may be the result of genetic drift or of selective or mechanistic constraint against mutation to alleles smaller than 85 bp. Sequencing analysis of several 115-bp alleles from Alu(−) chromosomes suggests that their presence on Alu(−) chromosomes results from an ancient recombination event with an Alu(+) chromosome and they have reached high frequency through genetic drift (39). Because we restrict our analyses to include only mutation events from the ancestral 90-bp allele and to make our analyses as conservative as possible, we exclude from initial consideration all STRP alleles of 110 bp or greater (1 case outside Africa, 49 cases within Africa). We find $P_A$ = 34/85 = 0.40 and $P_B$ = 264/269 = 0.9814. Therefore, $R$ = $\ln(P_A)/\ln(P_B)$ = 48.8. Assuming a maximum of 5 million years before present (Y.B.P.) for the origin of the Alu(−) chromosome in Africa, this gives a maximum age of 102,000 Y.B.P. [with an upper bound of 217,000 to 313,000 Y.B.P.

(40)] for the introduction of the Alu(−) chromosome outside of Africa. Slightly different approaches to analyzing the data give somewhat different values of $R$ and hence different dates for the founding of the non-African populations (41). However, even with implausibly conservative assumptions, the dates that result from our calculations are much more recent than the dates proposed by the multiregional origin hypothesis.

Our date estimates represent maximum values based on assumption 1 enumerated above, that the Alu(−) allele arose 5 million years ago. It is likely that the Alu deletion event occurred more recently, in which case our estimates of the date of founding of the non-African populations would also be more recent (42). Assumption 2 also gives a conservative estimate because violation of this assumption would result in a larger value of $R$. Regarding assumption 4, back mutation outside of Africa is unlikely to have played a large role because of the very low frequency of alleles other than 90 bp. Back mutation having occurred within Africa is more likely; however, ignoring such back mutation leads to an underestimate of $N_A$, and thus $R$. Hence, assumption 4 also provides a conservative estimate of $R$.

## A Single Recent African Origin

Because this is a single locus subject to the effects of random genetic drift and we are observing only one realization of evolution, it is likely that similar analysis of other loci will give slightly different date estimates. However, the difference in haplotype variation at the CD4 locus between sub-Saharan African and non-African populations is striking and cannot easily be accounted for except by a recent common origin of non-Africans from Africa. We present conservative maximum estimates for the origin of non-African humans that are consistent with the fossil record, which places anatomically modern humans in Africa as early as 120,000 Y.B.P. and dates the earliest modern human fossils in the Middle East at 90,000 to 120,000 Y.B.P. (43). They are also consistent with the estimates for the date for the last common mtDNA ancestor of all modern humans (African and non-African) of 100,000 to 300,000 Y.B.P. (6, 9).

Two alternative hypotheses that might explain the global pattern of haplotype variation and linkage disequilibrium are selection and the multiregional origin hypothesis. Specific selection schemes acting on mutations closely linked to the CD4 STRP can either reduce the variance of alleles at the STRP (a selective "sweep") or result in higher than expected levels of variation at the STRP (overdominant selection) (44). However, no simple selective schema can explain the diversity of haplo-

types in sub-Saharan Africa and the specific pattern of haplotypes seen in all non-African populations (45). Similar patterns of haplotype variation and differential linkage disequilibrium between Africans and non-Africans at several additional unlinked loci (33, 46) also suggest that selection effects could not explain these results.

The second alternative, the multiregional origin hypothesis, seems highly unlikely: It predicts roughly equivalent time depth and genetic diversity in all parts of the world. There has been enough time since the origin of the Alu deletion in Africa for mutation at the STRP, and to a lesser extent recombination between the markers, to generate multiple alleles and nearly randomize the Alu(−) allele with respect to the specific STRP allele on the same chromosome. In contrast, the limited number of STRP alleles and the essentially complete linkage disequilibrium in all non-African populations suggest a recent single population for their origin (47). It is highly unlikely that this pattern could be explained by a recent bottleneck event in non-African populations that have been diverging since H. erectus emerged from Africa 800,000 or more years ago; we would not expect to find such similar patterns of haplotype diversity and linkage disequilibrium. An out of Africa model invoking expansion and spreading of populations from a single common ancestor is most parsimonious.

Because we can trace the inheritance of chromosomes through both males and females, analysis of nuclear DNA markers provide a useful comparison with studies of mitochondrial and Y chromosome DNA variation. Haplotyped loci are especially useful because linkage disequilibrium provides more information beyond that of the single polymorphisms considered as independent markers. By choosing haplotype systems that approach equilibrium with different rates, we can explore different evolutionary time depths. The haplotype system at the CD4 locus has the appropriate low levels of recombination and mutation to be particularly useful for studying the period when modern humans first expanded out of Africa. The decreased variation and strong disequilibrium seen in non-African populations contrast with the quite different pattern in sub-Saharan African populations and together constitute some of the strongest evidence to date from the nuclear genome for a recent common African origin of all modern non-African human populations.

### REFERENCES AND NOTES

1. M. H. Wolpoff, in *The Human Revolution: Behavioral and Biological Perspectives on the Origins of Modern Humans,* P. Mellers and C. B. Stringer, Eds. (Edinburgh Univ. Press, Edinburgh, 1989), vol. 1, pp. 62–108.

2. M. H. Wolpoff, in *Continuity or Replacement: Controversies in Homo sapiens Evolution*, G. Bräuer and F. H. Smith, Eds. (Balkema, Rotterdam, 1992), pp. 25–63.

3. C. B. Stringer, in (2), pp. 9–23.

4. C. C. Swisher III et al., *Science* **263**, 1118 (1994); C. B. Stringer and P. Andrews, *ibid*. **239**, 1263 (1988).

5. M. M. Lahr and R. Foley, *Evol. Anthropol.* **3**, 48 (1995).

6. R. L. Cann, M. Stoneking, A. C. Wilson, *Nature* **325**, 31 (1987); L. Vigilant, M. Stoneking, H. Harpending, K. Hawkes, A. C. Wilson, *Science* **253**, 1503 (1991).

7. Classical analyses of blood group and protein polymorphisms also support the out of Africa hypothesis. See, for example, M. Nei and G Livshits, *Hum Hered.* **39**, 276 (1989).

8. S. B. Hedges, S. Kumar, K. Tamura, M. Stoneking, *Science* **255**, 737 (1992); D. R. Maddison, M. Ruvolo, D. L. Swofford, *Syst. Biol.* **41**, 111 (1992); A. Templeton, *Science* **255**, 737 (1992); *Am. Anthropol.* **95**, 51 (1993).

9. M. Ruvolo, T. R. Disotell, M. W. Allard, W. M. Brown, R. L. Honeycutt, *Proc. Natl. Acad. Sci. U.S.A.* **88**, 1570 (1991); M. Hasegawa, A. Di Rienzo, T. D. Kocher, A. C. Wilson, *J. Mol. Evol.* **37**, 347 (1993); K. Tamura and M. Nei, *Mol. Biol. Evol.* **10**, 512 (1993); S. Horai, K. Hayasaka, R. Kondo, K. Tsugane, N. Takahata, *Proc. Natl. Acad. Sci. U.S.A.* **92**, 532 (1995); D. Penny, M. Steel, P. J. Waddell, M. D. Hendy, *Mol. Biol. Evol.* **12**, 863 (1995).

10. R. L. Dorit, H. Akashi, W. Gilbert, *Science* **268**, 1183 (1995); M. F. Hammer, *Nature* **378**, 376 (1995); P. N. Goodfellow, *ibid.*, p.379.

11. The CD4 gene, which encodes the T4 glycoprotein expressed on the surface of T lymphocytes, is important for proper immune function and has been the site of active investigation; neither of the markers lies within coding or known regulatory elements [M. D. Blum, G. T. Wong, K. M. Higgins, M J. Sunshine, E. Lacy, *J. Exp. Med.* **177**, 1343 (1993); F. P. Gillespie et al., *Mol. Cell. Biol.* **13**, 2952 (1993); D. D. Duncan, A. Stupakoff, S. M. Hedrick, K. B. Marcu, G. Siu, *ibid.* **15**, 3179 (1995)].

12. M. C. Edwards, P. R. Clemens, M. Tristan, A. Pizzulti, R. A. Gibbs, *Nucleic Acids Res.* **19**, 4791 (1991).

13. M. C. Edwards and R. A. Gibbs, *Genomics* **14**, 590 (1992).

14. Mutation rates may vary considerably among loci but usually range between $10^{-5}$ and $10^{-2}$ per generation [J. L. Weber and C. Wong, *Hum. Mol. Genet.* **2**, 1123 (1993); M. M. Mahtani and H. F. Willard, *ibid.*, p. 431; D. Tautz and C. Schlötterer, *Curr. Opin. Genet. Dev.* **4**, 832 (1994)].

15. A. Bowcock et al., *Nature* **368**, 455 (1994); R. Deka et al., *Am. J. Hum. Genet.* **56**, 461 (1995); L.B. Jorde et al., *ibid.* **57**, 523 (1995); D. B. Goldstein et al., *Proc. Natl. Acad. Sci. U.S.A.* **92**, 6723 (1995); N. Risch et al., *Nat. Genet.* **9**, 152 (1995).

16. Others (12, 13) count the 3' flanking sequence CTTCTTTT as two repeats. Sequencing shows that all humans and nonhuman primates sampled share this same flanking sequence; therefore, we have not considered them part of the STRP. STRP alleles were amplified with the primers and protocol described in (12). Our 85-bp allele contains five perfect repeats and corresponds to the 88-bp allele listed in (12).

17. Sequencing analysis shows that chimpanzee and gorilla alleles contain only perfect repeats, whereas some alleles in orangutan and gibbon contain small insertions or deletions within one or two of the repeat units (data not shown).

18. PCR products from nine individuals homozygous for the Alu(−) allele were sequenced with primer R106 (13) and a newly designed primer R107 (5'-TATCT-CAGGCTGTCTCAGTG-3').

19. A. C. Wilson and V. M. Sarich, *Proc. Natl. Acad. Sci. U.S.A.* **63**, 1088 (1969); A. Hill, in *Integrative Paths to the Past: Paleoanthropological Advances in Honor of F. Clark Howell*, R. S. Corruccine and R. L. Ciochon, Eds. (Prentice-Hall, Englewood Cliffs, NJ, 1994), pp. 123–145.

20. Information on origins of population samples, sampling procedures, and preparation of DNA samples can be found by checking from the Kidd Lab Home Page on the World Wide Web at <http://info.med.yale.edu/genetics/kkidd>.

21. M. E. Hawley and K. K. Kidd, *J. Hered.* **86**, 409 (1995). Program and documentation are available via anonymous FTP from paella.med.yale.edu, in the directory pub/haplo.

22. S. Michalatos-Beloin, S. A. Tishkoff, K. K. Kidd, G. Ruano, in preparation. Among the sub-Saharan African Alu(−) haplotypes, 115 of 132 were unambiguously resolved by allele-specific long-range PCR or homozygosity. The remainder were inferred by HAPLO (21). Haplotypes in all sub-Saharan African populations were >90% resolved except in the Herero (50% resolved), Bantu speakers (79%), and Woloff (88%). Outside of Africa, all six non– 90-bp Alu(−) haplotypes were unambiguously resolved. The genotypes of an additional 30 doubly heterozygous individuals (Alu +/−, 90 bp/non–90 bp) from assorted non-African populations were unambiguously resolved; the 90-bp allele was associated with the Alu(−) allele in all cases.

23. Chi-square heterogeneity test of P. L. Workman and J. D. Niswander, *Am. J. Hum. Genet.* **22**, 24 (1970). $\chi^2 = 7.12$, df = 10, $P > 0.05$.

24. For Ethiopians $\chi^2 = 21.90$, df = 11, $P < 0.05$ and for Somali $\chi^2 = 38.70$, df = 10, $P < 0.001$ with methods described in (23).

25. B. O. Bengtsson and G. Thomson, *Tissue Antigens* **18**, 356 (1981).

26. A. Rogers and L. Jorde, *Hum. Biol.* **67**, 1 (1995).

27. J. S. Wainscoat et al., *Nature* **319**, 491 (1986); J. S. Jones and S. Rouhani, *ibid.*, p. 449.

28. F. J. Ayala, A. Escalante, C. O'hUigin, J. Klein, *Proc. Natl. Acad. Sci. U.S.A.* **91**, 6787 (1994).

29. Y. Satta, C. O'hUigin, N. Takahata, J. Klein, *ibid.*, p. 7184.

30. C. Sheba et al., *Am. J. Phys. Anthropol.* **20**, 167 (1962); I. Hakim, M. Gross, B. Bonné-Tamir, in *Pluridisciplinary Approach of Human Isolates*, A. Chaventre and D. F. Roberts, Eds. (Institut National Etudes Demographiques, Paris, 1990), pp. 43–57; U. Ritte et al., *Hum. Biol.* **65**, 359 (1993); Z. Zoossmann-Diskin et al., *Isr. J. Med. Sci.* **27**, 245 (1991).

31. D. A. Merriwether et al., *J. Mol. Evol.* **33**, 543 (1991); A. Di Rienzo and A. C. Wilson, *Proc. Natl. Acad. Sci. U.S.A.* **88**, 1597 (1991).

32. L. L. Cavalli-Sforza, A. Piazza, P. Menozzi, J. C. Mountain, *Proc. Natl. Acad. Sci. U.S.A.* **85**, 6002 (1988); A. M. Bowcock et al., *ibid.* **88**, 839 (1991); A. M. Bowcock et al., *Gene Geogr.y* **5**, 151 (1991); M. Nei and A. K. Roychoudhury, *Mol. Biol. Evol.* **10**, 927 (1993); M. A. Batzer et al., *Proc. Natl. Acad. Sci. U.S.A.* **91**, 12288 (1994); J. J. Martinson et al., *Am. J. Hum. Genet* **57**, 1186 (1995); H. Zischler, H. Geisert, A. von Haeseler, S. Pääbo, *Nature* **378**, 489 (1995).

33. C. M. Castiglione et al., *Am. J. Hum. Genet.* **57**, 1445 (1995).

34. Analyses with the η coefficient [Q. McNemar, *Psychological Statistics* (Wiley, New York, 1969)] and several genetic distance measures on STRP frequencies normalized on Alu(+) and Alu(−) chromosomes—Fst [J. Reynolds, B. S. Weir, C. C. Cockerham, *Genetics* **105**, 767 (1983)], Rst (48), Nei's index of similarity (I) [M. Nei, *Molecular Evolutionary Genetics* (Columbia Univ. Press, New York, 1987]—all show high levels of similarity in the distributions of STRP alleles on Alu(+) chromosomes in all sub-Saharan populations, an equivalent level of similarity on Alu(−) chromosomes, but much less similarity between Alu(+) and Alu(−) chromosomes whether from the same or different populations. For example, with the η coefficient as a measure of similarity, the means of the pairwise values among the 10 sub-Saharan populations is 0.90 for Alu(+) chromosomes and 0.84 for Alu(−) chromosomes. The lower similarity for Alu(−) chromosomes reflects the smaller number of Alu(−) chromosomes (and hence high variance) in several populations. If we consider only those five populations with more than 10 Alu(−) chromosomes, the mean similarity values are 0.89 for Alu(+) and 0.88 for Alu(−) chromosomes. We contrast these with the means of comparisons of Alu(+) with Alu(−) chromosomes within a population (0.63) and between populations (0.70). Limiting the comparisons to the five populations with more than 10 Alu(−) chromosomes changes these mean values minimally to 0.68 (within populations) and 0.69 (between populations).

35. A. M. Valdes, M. Slatkin, N. B. Freimer, *Genetics* **133**, 737 (1993); M. D. Shriver, L. Jin, R. Chakraborty, E. Boerwinkle, *ibid.* **134**, 983 (1993); A. Di Rienzo et al., *Proc. Natl. Acad. Sci. U.S.A.* **91**, 3166 (1994).

36. Variation at the STRP and linkage disequilibrium between the markers will be affected by the demographic history of the populations [H. C. Harpending, S. T. Sherry, A. R. Rogers, M. Stoneking, *Curr. Anthropol.* **34**, 483 (1993); M. Slatkin, *Genetics* **137**, 331 (1994)]. Because allelic variation may have been lost in a small ancestral population some time after leaving Africa but before rapid population expansion, we could be estimating the time since divergence and expansion of the ancestral non-African population, which may have occurred subsequent to migration out of Africa.

37. Other assumptions may be warranted for other haplotype loci. Specifically, for markers with low mutation rate but proportionally higher recombination rates than evident at CD4, models and analyses based primarily on recombination would be most appropriate.

38. Ethiopian Jews, Somalis, and Egyptians have been excluded from the analysis because of the possibility of admixture with sub-Saharan Africans and non-Africans.

39. Sequencing analysis of more than 100 alleles from chromosomes of humans of diverse geographical origin shows that alleles of 110 bp or larger on Alu(+) chromosomes have an imperfect CTTTC as the fourth repeat from the 5' end. Thus, the ancestral STRP allele was most likely a small-sized perfect repeat of TTTTC (17), and alleles 110 bp or larger are derived from a single mutation that generated a higher repeat-number allele with the imperfect CTTTC. Sequencing analysis of several large-sized alleles from Alu(−) chromosomes demonstrates that they share the point mutation present on Alu(+) chromosomes; thus, these alleles most likely derive from an ancient recombination event transferring a large-sized allele from an Alu(+) chromosome to an Alu(−) chromosome with subsequent mutation by one or few repeats to other large-sized alleles.

40. To estimate a lower bound for R from these data (hence an upper bound maximum age), we note that the variability in the estimate of R depends much more on variability in the denominator than the numerator. According to a Poisson distribution, the value of $P_B$ that corresponds to a probability of 5% for there to appear at most five alleles outside of Africa without the 90 bp STRP on Alu(−) chromosomes is 0.961. This gives a lower 95% confidence limit for R of ln(0.40)/ln(0.961) = 23.0 (or an age maximum of 5 million Y.B.P./23.0 = 217,000 Y.B.P.). The Poisson assumption for non-Africans depends on statistical independence of the chromosomes sampled, a reasonable assumption because at least five of the six non– 90-bp Alu(−) haplotypes derive from distinct populations (one 100-bp in the Druze, one 85-bp in Yemenite Jews, one 85-bp and one 95-bp in mixed European, one 85-bp in Asiatic Indian, and one 110-bp in Adygei). Regarding variability on the value of $P_A$, we examined the P values for the five sub-Saharan populations with more than 10 Alu(−) chromosomes. The mean of these values is 0.40, the same as the combined value we used above. However, even if we use the largest of these values (0.53), along with the lower 95% confidence limit for $P_B$, we obtain a value for R of ln(0.53)/ln(0.961) = 16.0 (or maximum age of 5 million/16.0 = 313,000 Y.B.P.).

41. If the 85-bp allele rose to high frequency in sub-Saharan Africa through drift or through mechanistic or selective constraint, we can combine the 85- and 90-bp alleles into a single progenitor group to obtain a highly conservative estimate of R. In this case, using alleles less than 110 bp in size, $P_A = 69/85 = 0.812$; applying Eq. 1 gives R = 11.1 for a maximal age of 450,000 years for the origin of the Alu(−) chromosome outside of Africa. Assuming the 115-bp allele on Alu(+) chromosomes is the result of an early recombination event (39), the 90- and 115-bp alleles can be considered as a single progenitor group, and STRP alleles of all other sizes as the mutants, giving $P_A = 70/132$, $P_B = 264/270$, R = 28.2, for a maximal age of 177,000 years. Various

models that assume stepwise mutation of the STRP imply that the variation of allele sizes increases linearly with time (48) [D. B. Goldstein, A. Ruiz Linares, L. L. Cavalli-Sforza, M. W. Feldman, *Genetics* **139**, 463 (1995)]. Thus, we can estimate *R* as the ratio of variances of allele sizes (in base pairs) for alleles less than 110 bp (22.5/0.75 = 30.0), giving a maximum age of 167,000 Y.B.P. Because we have considered only a single locus, the standard error on this variance-ratio estimate is high. However, the apparent lower boundary of 85 bp to the STRP allele size results in an underestimate of the time of origin of the Alu(−) chromosome in Africa, so again this estimate of *R* is conservative.

42. Despite a smaller effective population size of Alu(−) chromosomes (because of a lower frequency), the mean variance of STRP alleles on Alu(−) chromosomes among the 10 sub-Saharan African populations is still substantial [64% of the mean variance of STRP alleles on Alu(+) chromosomes], suggesting that while the origin of the Alu(−) chromosome is more recent than the origin of the Alu(+) chromosome, it is still quite ancient.

43. C. B. Stringer, in *The Origin and Evolution of Humans and Humanness*, D. T. Rasmussen Ed. (Jones and Bartlett, Boston, 1993), pp. 75–94.

44. M. Slatkin, *Mol. Biol. Evol* **12**, 473 (1995).

45. Both the STRP and the Alu deletion polymorphism are located in noncoding regions of the CD4 gene and are unlikely to have any functional significance. Except in the unlikely case of strong positive cis-acting epistasis of functional variants flanking exons 1 and 2 of the 90-bp Alu(−) chromosomes, selection cannot explain the maintenance of the linkage disequilibrium seen in non-African populations because there would be nothing to prevent formation of "non– 90 bp" Alu(−) chromosomes by recombination between the markers or mutation at the STRP. In addition, the very low frequency of the Alu(−) chromosomes in Asia, the Pacific islands, and the New World, as well as the very high frequency of 85-bp Alu(+) and 110-bp Alu(+) chromosomes in all non-African populations, argues against strong positive selection for the Alu(−) chromosome.

46. S. A. Tishkoff, W. C. Speed, J. R. Kidd, K. K. Kidd, *Am. J. Phys. Anthropol. Suppl.* **20**, 211 (1995); S. A. Tishkoff *et al.*, *Am. J. Hum. Genet.* **57S**, A42 (1995); K. K. Kidd *et al.*, *Alcoholism: Clinical and Experimental Research*, in press.

47. Very large and unlikely amounts of gene flow and drift acting in a similar manner in many geographically dispersed populations would be required for the 90-bp Alu(−) haplotype to have been recently introduced into preexisting *H. erectus* populations and achieve the frequencies seen today. The low frequency of the 90-bp Alu(−) chromosome in all Asian, Pacific island, and New World populations argues against high levels of gene flow from European or Middle Eastern populations into these regions before historical times.

48. M. Slatkin, *Genetics* **139**, 457 (1995).

49. We thank A. Deinard for providing nonhuman primate samples. For assistance with collecting African DNA samples, we thank H. Soodyall and G. Sirugo as well as R. Aman from the National Museums of Kenya. We thank M. Hawley for assistance with the HAPLO program. We also thank our many colleagues for helpful comments and criticisms in the course of this study, including J. Cheng, A. G. Clark, A. Di Rienzo, R. Dorit, R. Harding, L. B. Jorde, M. Slatkin, and an anonymous reviewer. Supported in part by the Medical Research Council of South Africa and by a grant to M.K. from Boehringer Ingelheim Fonds, NIH grant HG00348 to N.R., USPHS grants AA09379 and MH39239 to K.K.K., USNSF grant DBS9208917 to K.K.K. and SBR-9408934 to J.R.K., and a grant from the Alfred P. Sloan Foundation to K.K.K.

---

## RESEARCH ARTICLE

# The Exchange of Impact Ejecta Between Terrestrial Planets

Brett J. Gladman, Joseph A. Burns, Martin Duncan, Pascal Lee, Harold F. Levison

Orbital histories of ejecta from the terrestrial planets were numerically integrated to study their transfer to Earth. The properties of the lunar and martian meteorites are consistent with a recurrent ejection of small meteoroids as a result of impacts on their parent bodies. Long-range gravitational effects, especially secular resonances, strongly influence the orbits of many meteoroids, increasing their collision rates with other planets and the sun. These effects and collisional destruction in the asteroid belt result in shortened time scales and higher fluxes than previously believed, especially for martian meteorites. A small flux of mercurian ejecta appears possible; recovery of meteorites from the Earth and Venus is less likely.

The study of meteorites has illuminated the nature of extraterrestrial environments and astrophysical processes, particularly the conditions at the time of our solar system's formation. Most meteorites come from asteroids, but recently a number of objects from the moon and Mars have been recognized. These latter meteorites help to characterize the surfaces of these bodies, especially the martian meteorites, which are our only samples of that planet. To learn more about the parent bodies and the paths that the meteorites traveled before arriving on Earth, we must understand the orbital dynamics governing their transfer. In particular, what is the delivery efficiency, that is, the fraction of escaping ejecta that reach Earth, from different sources?

The SNC meteorites (1) have features that suggest they are derived from Mars (2) and are somehow delivered to Earth. Even though the petrology and young crystallization ages of the SNCs point to an origin on a large parent body with recent geologic activity, Mars was only recently accepted as their source because it was thought unlikely that rocks could survive being blasted off of a planet (3). Another argument against a martian origin was that if these objects were launched from Mars, surely there should be many more meteorites coming from the moon, which is closer and has a lower escape velocity; yet, as of 1982, there were six SNC falls but no lunar meteorites. Then the lunar meteorite ALHA81005 was recognized among Antarctic meteorites (4), and because of our familiarity with the returned Apollo and Luna samples, the origin of this meteorite was immediately accepted. With more Antarctic meteorites (5) being recovered, we currently have a dozen members of each class, although they comprise only 0.1% of all meteorites. The SNCs are now generally acknowledged to have originated on Mars, after further examination of their composition, especially the virtually perfect isotopic match between gases trapped within one of them and the martian atmosphere, as determined by the Viking landers (6). Thus, we use the term "martian meteorites" for the SNCs and ALH84001, the latter being distinct from the SNC classification (7).

We can learn about the dynamics of the inner solar system by comparing the measured transfer ages of such meteorites against orbital histories. Some of the ideas presented below are not new, but our dynamical simulations improve on previous work that, because of computational limitations of the time, used Monte Carlo calculations rather than full orbital integrations. In principle, some aspects of the cratering process itself can also be constrained (8).

A cosmic ray exposure (CRE) age (9) of a meteorite is the time during which the object was bombarded by energetic cosmic rays in space. During this exposure, measurable radioactive isotopes accumulate, allowing the duration of the exposure to be estimated (Table 1). Most lunar meteorites were delivered to the Earth in a far shorter time than any martian meteorite, and the martian meteorites have an average mass 38 times that of the lunar ones. The issue of pairing (10) is unlikely to affect these trends significantly.

B. J. Gladman and P. Lee are in the Department of Astronomy, Cornell University, Ithaca, NY 14853, USA. J. A. Burns is in the Departments of Astronomy and Theoretical and Applied Mechanics, Cornell University, Ithaca, NY 14853, USA. M. Duncan is in the Department of Physics, Queen's University, Kingston, ON K7L 3N6, Canada. H. F. Levison is with Geophysical, Astrophysical, and Planetary Sciences, Southwest Research Institute, 1050 Walnut Street, Suite 409, Boulder, CO 80302, USA.