Our study shows the rapid expansion of the parasite population concurrent with the emergence of agricultural societies in humans, and speciation in the African mosquito vectors. However, the TMRCA estimates (9) for South American and Asian populations, which showed very little growth, were similar to if slightly more recent than the TMRCA for the worldwide population (Table 1). All three populations appear to be older than the African expansion event. This suggests that the parasite migrated from Africa before the recent expansion, perhaps during the Pleistocene expansion in humans, which was followed by migration out of Africa 40,000 to 130,000 years ago (24–27).

Our data provide a detailed picture of mtDNA diversity and genealogical relationships for a worldwide sample of *P. falciparum*, made possible by the lack of recombination in the parasite mtDNA. We show that the parasite mtDNA is more diverse than previously believed (5), and we provide strong evidence for a recent and rapid population expansion in Africa followed by migration to other regions in agreement with recent predictions (1, 2). However, our data reject the claim that the parasite originated 6000 years ago, based on evidence that the world and some of the regional populations appear to be much older. In contrast, because exponential growth predicts much faster growth in the recent than the distant past, it is possible that the worldwide population remained relatively small for a considerable amount of time, even as it spread to other regions. The genetic consequences of exponential growth can be seen clearly in the position of the vast majority of mutations near the tips of the gene tree (Fig. 3A). Interestingly, the 10 most recent mutations are from Africa, suggesting further that this population is growing faster than the others. Figure 3B summarizes our model of the evolutionary history of the *P. falciparum* parasite with the caveat that the mtDNA is a single locus and, therefore, represents only one observation of the parasite evolutionary history. Finally, our data show that historical changes in the hosts—both migration and changes in population size—have had a major impact on parasite demography.

References and Notes

1. M. Coluzzi, *Parassitologia* **41**, 277 (1999).
2. ———, A. Sabatini, A. della Torre, M. A. Di Deco, V. Petrarca, *Science* **298**, 1415 (2002).
3. S. M. Rich, M. C. Licht, R. R. Hudson, F. J. Ayala, *Proc. Natl. Acad. Sci. U.S.A.* **95**, 4425 (1998).
4. S. K. Volkman et al., *Science* **293**, 482 (2001).
5. D. J. Conway et al., *Mol. Biochem. Parasitol.* **111**, 163 (2000).
6. A. L. Hughes, F. Verra, *Proc. R. Soc. Lond. B Biol. Sci.* **268**, 1855 (2001).
7. J. Mu et al., *Nature* **418**, 323 (2002).
8. J. E. Feagin, *Int. J. Parasitol.* **30**, 371 (2000).
9. Methods and additional supporting results are available as supplementary material on *Science* Online.
10. M. Nei, T. Gojobori, *Mol. Biol. Evol.* **3**, 418 (1986).
11. J. H. McDonald, M. E. Kreitman, *Nature* **351**, 652 (1991).
12. A. A. Escalante, D. E. Freeland, W. E. Collins, A. A. Lal, *Proc. Natl. Acad. Sci. U.S.A.* **95**, 8124 (1998).
13. M. Brunet et al., *Nature* **418**, 145 (2002).
14. S. Schneider, D. Roessli, L. Excoffier, Arlequin software (2000). Available at: http://lgb.unige.ch/arlequin/.
15. T. J. Anderson et al., *Mol. Biol. Evol.* **17**, 1467 (2000).
16. J. Castelloe, A. R. Templeton, *Mol. Phyl. Evol.* **3**, 102 (1994).
17. J. C. Wootton et al., *Nature* **418**, 320 (2002).
18. Y.-x. Fu, *Genetics* **147**, 915 (1997).
19. F. Tajima, *Genetics* **123**, 585 (1989).
20. Y.-x. Fu, W.-H. Li, *Genetics* **133**, 693 (1993).
21. A. R. Rogers, H. Harpending, *Mol. Biol. Evol.* **9**, 552 (1992).
22. M. K. Kuhner, J. Yamato, J. Felsenstein, *Genetics* **149**, 429 (1998).
23. O. G. Pybus, A. Rambaut, P. H. Harvey, *Genetics* **155**, 1429 (2000).
24. R. L. Cann, M. Stoneking, A. C. Wilson, *Nature* **325**, 31 (1987).
25. M. Ingman, H. Kaessmann, S. Paabo, U. Gyllensten, *Nature* **408**, 708 (2000).
26. R. Thomson, J. K. Pritchard, P. Shen, P. J. Oefner, M. W. Feldman, *Proc. Natl. Acad. Sci. U.S.A.* **97**, 7360 (2000).
27. P. A. Underhill et al., *Nature Genet.* **26**, 358 (2000).
28. M. Bahlo, R. C. Griffiths, *Theoret. Pop. Biol.* **57**, 79 (2000).
29. We thank A. Clark, T. Anderson, and B. Griffiths for advice and discussion, and B. R. Marshall for editorial assistance.

Chromosomal Speciation and Molecular Divergence—Accelerated Evolution in Rearranged Chromosomes

Arcadi Navarro[1]* and Nick H. Barton[2]

Humans and their closest evolutionary relatives, the chimpanzees, differ in ~1.24% of their genomic DNA sequences. The fraction of these changes accumulated during the speciation processes that have separated the two lineages may be of special relevance in understanding the basis of their differences. We analyzed human and chimpanzee sequence data to search for the patterns of divergence and polymorphism predicted by a theoretical model of speciation. According to the model, positively selected changes should accumulate in chromosomes that present fixed structural differences, such as inversions, between the two species. Protein evolution was more than 2.2 times faster in chromosomes that had undergone structural rearrangements compared with colinear chromosomes. Also, nucleotide variability is slightly lower in rearranged chromosomes. These patterns of divergence and polymorphism may be, at least in part, the molecular footprint of speciation events in the human and chimpanzee lineages.

If speciation processes have left any molecular footprints, detecting them could not only shed light on speciation processes along the human lineage, but also would help to identify the specific genomic regions responsible for the separation of humans and other primates and bring us closer to identifying the genetic differences that may underlie the morphological, behavioral, and cognitive differences between us. The role of chromosomal rearrangements in speciation is particularly well supported by several lines of evidence (1–6). Classical models of chromosomal speciation state that, because heterozygous individuals are partly sterile (i.e., underdomi-nant), chromosomal changes act as genetic barriers to gene flow between populations fixed for different arrangements (1, 3) and, thus, facilitate reproductive isolation. However, these models are burdened by a paradox that renders them weak and unconvincing (2): If underdominance were strong, it would be very unlikely that new rearrangements could get established. On the other hand, if underdominance were weak enough for fixation to be likely, chromosomal rearrangements would be very poor barriers to gene flow and, thus, unlikely to contribute to speciation. Recently, a new class of models has been proposed that suggest that chromosomal changes are strong genetic barriers because they reduce recombination in heterokaryotypes and not because of underdominance (3, 4, 6, 7). Such strong barriers would facilitate divergence in the rearranged region during the time when the diverging populations are in parapatry, i.e., have limited gene flow. Their effects would be especially pronounced if

[1]Departament de Ciències Experimentals i de la Salut, Universitat Pompeu Fabra, Doctor Aiguader 80, 08003 Barcelona, Catalonia, Spain. [2]Institute of Cell, Animal and Population Biology, University of Edinburgh, West Mains Road, Edinburgh EH9 3JT, Scotland, UK.

*To whom all correspondence should be addressed. E-mail: arcadi.navarro@cexs.upf.es

divergence is through the accumulation of incompatible alleles, as proposed by Bateson, Dobzhansky, and Muller (8). The spread of new favorable alleles through a species will be delayed if they are linked to chromosomal differences that are already established within that species. This gives different alleles time to spread through the rest of the species range; if these different alleles are incompatible with each other, then they will be trapped at the chromosomal barrier and will strengthen reproductive isolation. In contrast, favorable alleles that are not impeded by the barrier will spread quickly through the whole species, making parapatric divergence less likely. The accumulation of incompatibilities facilitated by chromosomal differences generates genetic barriers of growing strength that, eventually, produce complete reproductive isolation and, therefore, speciation (fig. S1). This accumulation of positively selected alleles in chromosomes presenting rearrangements underlies the key prediction of this and similar models of chromosomal speciation (3): that the molecular signature of positive selection should be stronger across chromosomes carrying rearrangements than in colinear chromosomes (7).

Within-population polymorphism for rearrangements has rarely been described in mammals. However, parapatric coexistence of different arrangements is common (1) and, in particular, has been described in *Pongo* (9) and other primates (10). Human and chimpanzee karyotypes are highly homologous, but some major chromosomal changes differentiating the two species can be detected in banded metaphase chromosomes. Human chromosomes 1, 4, 5, 9, 12, 15, 16, 17, and 18 are separated from their chimpanzee's homologs by pericentric inversions, and human chromosome 2 is the result of the tandem fusion of two acrocentric chromosomes common to the rest of great apes (11, 12). If one or more of these rearrangements were polymorphic while human and chimpanzee lineages were diverging, or during the successive speciation events that led to modern humans and chimpanzees, and if speciation took place along the lines proposed in the

model, we might be able to detect the molecular signature of positive selection on rearranged chromosomes. Previous work on the X-4 fusion in *Drosophila americana* provides evidence that selectively favorable variants can accumulate on rearranged chromosomal segments in a natural population (13).

The rate of protein evolution of a gene can be used to uncover the footprints of past positive selection (14). It is measured as the rate of nonsynonymous nucleotide substitution per nonsynonymous site (K_A) relative to the underlying neutral mutation rate, which is given by the rate of synonymous substitution per synonymous site (K_S). Amino acid changes are usually deleterious because of the action of purifying selection (15) and, thus, $K_A \ll K_S$. However, if positive selection is strong enough, K_A may increase and even exceed K_S. Therefore, K_A/K_S ratios larger than 1 suggest positive selection.

We gathered human-chimpanzee divergence measures for the coding regions of 115 annotated, autosomal genes (16). We calculated K_S and K_A values for these genes by the method of Li (17) implemented in the program DAMBE (18). The cytological position in the human genome of every gene was ascertained by means of OMIM and MapViewer. Chromosomes were classified as rearranged or colinear, in correspondence to the presence or absence of chromosomal rearrangements (11, 12).
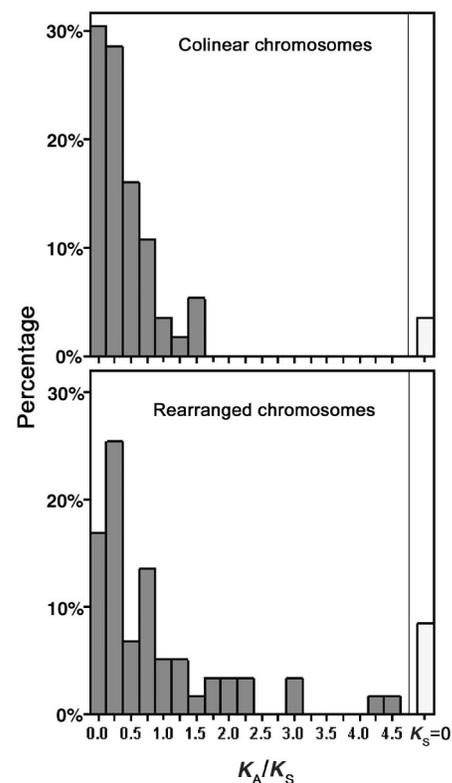
The K_S and K_A values of the 115 genes are listed in table S1. Both the average K_S (1.53%) and K_A (0.76%) were in the range of previous measures of coding sequence divergence (19). They were somewhat high because this data set was compiled to examine K_A/K_S ratios and, thus, genes with no divergence had to be excluded. The average K_A/K_S ratio was 0.61. This figure is based on 108 genes only, because seven genes with a K_S value of 0% were excluded because they had a K_A/K_S ratio of infinity, which also precluded their use in most of our analysis. The full data set, however, was used in a simple test. We built a contingency table by classifying genes according to two categories: their cytological po-

sition (with genes mapping in rearranged and in colinear chromosomes falling in different classes) and the value of their K_A/K_S ratio (with different classes for genes with $K_A/K_S > 1$ and genes with $K_A/K_S \leq 1$). Genes with $K_A/K_S > 1$ tended to cluster in rearranged chromosomes [Table 1, all genes (top) G test, P = 0.0024]. This pattern was also visible when only the 108 genes with $K_S > 0\%$ were considered [Table 1 (bottom), G test, P = 0.0043].

The distribution of K_A/K_S ratios in rearranged and colinear chromosomes is shown in Fig. 1. The average ratio for rearranged chromosomes was 0.84, more than twice the average ratio of colinear chromosomes ($K_A/K_S = 0.37$). These two ratios are significantly different, as detected by both Mann-Whitney's U test (one-tailed P = 0.008) and a permutation test (P = 0.0005, for which the difference in average K_A/K_S between rearranged and colinear chromosomes was used as the measured statistic). Genes in rearranged chromosomes can be further classified according to their position. Genes mapping within a chromosomal rearrangement or in the same chromosomal region (20) were classified as "close" to chromosomal changes. Genes mapping in different chromosomal regions were classified as "far" from chromosomal changes. The average K_A/K_S ratio was 0.68 for genes far from chromosomal

Fig. 1. Distribution of K_A/K_S ratios in colinear and rearranged chromosomes.

Table 1. Contingency tests. For all genes, G = 11.223, P = 0.0024. (Genes with $K_S = 0\%$ are considered to have $K_A/K_S > 1$.) For genes with $K_S > 0$, G = 8.415, P = 0.0043.

K_A/K_S	Chromosomes		
	Rearranged	Colinear	Total
All genes			
>1	20	6	26
≤1	39	50	89
Total	59	56	115
Genes with $K_S > 0$			
>1	15	4	19
≤1	39	50	89
Total	54	54	108

changes and 0.91 for genes close to them. This trend was not significant ($P = 0.23$ in a permutation test), but it certainly is in the direction expected under the model of chromosomal speciation discussed here.

If lower gene flow in rearranged chromosomes triggered the process that produced higher K_A/K_S ratios in rearranged chromosomes, then this lower gene flow might also have resulted in higher neutral divergence in these chromosomes (21). In our sample, the average K_S value of genes in rearranged chromosomes was 1.60% (1.67% in colinear) if only the genes for which K_A/K_S ratios were available were included in the analysis. When all genes were included, average K_S became 1.46% in rearranged and 1.61% in colinear chromosomes. Neither of these differences is significant, and their standard errors are large. Moreover, as explained above, we gathered this data set to measure K_A/K_S ratios and not the absolute level of divergence. To be able to perform tests on the level of neutral divergence, we used two other divergence data sets (22–24). In both data sets, the values of K, the number of nucleotide substitutions per 100 sites, are higher for rearranged than for colinear chromosomes (1.40% versus 1.15% in the data from (22) and 1.25% versus 1.23% in the data from (23). Although differences are only significant for the first data set ($P < 0.001$ in a permutation test), the measures are indeed in the expected direction if gene flow had been lower in rearranged chromosomes.

Neutral divergence is not randomly distributed in the genome. It has been shown that genomic regions with more CpG sites experienced higher rates of divergence between humans and chimpanzees (23); that linked genes have similar K_A and K_S values in human-rodent comparisons (25); that genes are frequently associated with CpG-rich islands (26); that proteins of linked genes have similar K_A/K_S ratios (27, 28); and that different chromosomes have different protein evolution rates (25). Evidence for this clustering of the rates of protein evolution is also found in our data. We built a list of all possible pairs of genes within a chromosome. The cytological (in number of subbands) and physical (in Mb) distances between each pair of genes were measured and, in turn, correlated with their K_A/K_S ratios. The ratios tended to be more similar for closer pairs of genes (Spearman Rank-Correlation: 0.252, $P < 0.01$ for cytological distances and 0.204, $P < 0.01$ for physical distances). Because comparing all possible pairs of genes artificially inflates sample size (25), a permutation test was performed on the value of Spearman Rank-correlation, randomly shuffling the positions of genes. The test is significant ($P < 0.001$) for both cytological and physical distances. The same tests were performed separately for rearranged and colinear chromosomes. For both physical and cytological distance, clustering turns out to be significant for colinear chromosomes ($P = 0.002$ and $P = 0.005$), but it is only marginally significant for rearranged chromosomes ($P = 0.096$ and $P = 0.074$). To assess the effect of clustering in our primary result of higher K_A/K_S ratios in rearranged chromosomes, we performed two further permutation tests. The tests were as described above, except that instead of randomly shuffling individual genes between rearranged and colinear chromosomes, we shuffled blocks of genes. In the first test, we used blocks defined by genomic regions and by chromosomal bands in the second. In both cases, the probability of obtaining a difference in K_A/K_S ratios between rearranged and colinear chromosomes as big as in our data is $P < 0.001$. Therefore, we can conclude that differences between rearranged and colinear chromosomes are not explained by clustering of K_A/K_S values.

Even though differential K_A/K_S values are not explained by clustering of evolutionary rates of proteins, they may be affected by genomic variables such as levels or mutation or recombination. GC content has been shown to correlate positively with the levels of synonymous substitution in the mammalian genome (22, 28, 29), and the amount of GC at fourfold redundant sites (GC4) correlates negatively with K_A/K_S values in mouse-rat comparisons (27). Thus, the higher K values found in the two divergence data sets analyzed here (22, 23) could be due to more GC. Analysis of these data sets shows that, far from being higher, the amount of GC is actually lower in rearranged than in colinear chromosomes [47.80% versus 47.96% for (22) and 36.50% versus 40.17% for (23)]. Therefore, the tendency of rearranged chromosomes to have higher neutral divergence values cannot be explained by their having more GC. Furthermore, the analysis of the amount of GC4 in the set of 115 genes that we used to ascertain K_A/K_S ratios (30) shows that there are no significant differences between the two chromosomal classes. However, the amount of GC4 is slightly lower in genes mapping in rearranged chromosomes (56% versus 62%). Even though these differences are not significant, it is plausible that the K_A/K_S ratio in rearranged chromosomes may be inflated by the lower amount of GC4. To make sure that the significance of our primary result is not affected by a covariation in GC4 content, we performed a permutation test in which the genomic distribution of GC4 was kept constant between permutations. Following the method in (25), we divided our data set in classes of similar GC4 content, with each class containing 10% of the full data set. Then, we swapped only genes within classes. The test shows that, even when we controlled for GC4 content, genes in rearranged chromosomes have significantly higher K_A/K_S ratios ($P = 0.0076$). Finally, the amount of GC4 is thought to correlate with recombination rates in mice (31), and rates of protein evolution might reflect, in part, the local strength of purifying selection, which depends on recombination rates (27). To check whether rearranged and colinear chromosomes have different recombination rates, we used the high-resolution recombination map of the human genome in (32). The average value of centimorgans per megabase is 1.10 in rearranged and 1.17 in colinear chromosomes. Differences are not significant. Therefore, we can conclude that differential levels of mutation and recombination do not explain the differential K_A/K_S ratios between rearranged and colinear chromosomes.

If speciation involved the differential accumulation of incompatible favorable alleles in chromosomes with rearrangements, then the selective sweeps produced by the fixation of such alleles should have decreased neutral polymorphism in these chromosomes (14). This effect should be detectable if the fixation of favorable alleles were recent. In the human-chimpanzee case, we expect the effect to be very weak, if present at all, because the original separation of the two lineages occurred between 6 or 7 million years ago and because, although the dates of successive speciation events along the two lineages are unclear, they are certainly not recent (33, 34), and variability is likely to have recovered. Also, recent demographic events might have erased the signature of any ancient selective sweeps (35). Nevertheless, an analysis of the data set in (36) shows that variability is lower in rearranged chromosomes (37). As can be seen in Table 2, haplotype heterozygosity is weakly, but significantly, lower in rearranged chromosomes ($P = 0.014$, Mann-Whitney's U test) and all the other measures, including Tajima's D, are in the expected direction if old sweeps did determine higher K_A/K_S ratios in rearranged chromosomes. These trends are compelling but not conclusive, as evidence

Table 2. Variability measures from (36).

	Chromosomes	
	Rearranged	Colinear
Heterozygosity*	0.52	0.57
π	$5.5 \cdot 10^{-4}$	$6.1 \cdot 10^{-4}$
θ	$9.4 \cdot 10^{-4}$	$9.7 \cdot 10^{-4}$
Tajima's D	−0.99	−0.96

*$P = 0.014$, Mann-Whitney's U test.

from *Drosophila* (*38*) shows that the fixation of chromosomal rearrangements could produce a selective sweep that could have similar effects on linked variability.

Our observation that the relative rate of amino acid substitution is substantially higher in chromosomes that have undergone rearrangements during the divergence between chimpanzees and humans may be explained, at least in part, as a consequence of the barrier to gene flow generated by those rearrangements. Such an evolutionary explanation implies, first, that chromosomal changes are a major factor determining local rates of protein evolution; second, that the biological processes that led to the separation of our two lineages and/or to successive speciations, were facilitated, or even triggered, by chromosomal changes; third, that these processes involved periods of hybridization before speciation was gradually completed; and, finally, that the key genomic regions involved in these crucial changes are the ones spanned by rearrangements. Nevertheless, a twofold difference in K_A/K_S ratio between rearranged and colinear chromosomes is so substantial that it raises the question of whether it can be exclusively due to the speciation mechanism proposed here (*7*). Even taking into account that in our particular sample the lower amount of GC4 in genes mapping in rearranged chromosomes is slightly inflating their K_A/K_S values, a twofold difference seems unlikely under the hypothesis of a time-constant rate of adaptive substitutions, because it could only be explained if rearrangements had been barriers in parapatry for at least half the time of divergence of the two lineages. Crucially, the evolutionary rate of a protein is not always constant through time (*39*, *40*) and episodes of rapid change have been shown to be associated with speciation and the dispersal of new species (*41*, *42*). As mentioned above, other speciation-related processes, which have not yet been explored, may have contributed to the difference. For example, by preventing gene flow, rearrangements might facilitate the accumulation of alleles that are under weak divergent selection. Moreover, favorable alleles that do not take place in incompatibilities might accumulate in the low gene flow regions defined by rearrangements because they spread through the population much faster than neutral variants. Even so, explanations unrelated to speciation that would inflate the difference in protein evolution rates cannot be completely ruled out. It is possible, for example, that adaptive changes in certain genes or regions facilitate the accumulation of further adaptive changes in the same genes or regions (*43*); that the accumulation of positively selected alleles in a new arrangement facilitates its fixation; that the reduction in population size associated with the fixation of a new rearrangement allows for a relaxation of purifying selection; or that changes in recombinational context associated with rearrangements might move linked regions to regions with a different equilibrium base composition and thus lead to changes in mutation rates. Anyhow, it is clear that several fascinating questions are raised by our results. Did all the chromosomal changes currently separating the two species became fixed because of some speciation event? In which lineage and time did a given speciation and the associated adaptation took place? What is the functional content of the genomic areas involved in that adaptation? When more data become available, it will be possible to address them and to take a further step into linking the structure and function of the genome.

References and Notes

1. M. King, *Species Evolution* (Cambridge Univ. Press., Cambridge, 1993).
2. F. Spirito, in *Endless Forms* D. J. Howard, S. H. Berlocher, Eds. (Oxford Univ. Press, New York, 2000), pp. 320–329.
3. L. H. Rieseberg, *Trends Ecol. Evol.* **16**, 351 (2001).
4. M. A. F. Noor, K. L. Grams, L. A. Bertucci, J. Reiland, *Proc. Natl. Acad. Sci. U.S.A.* **98**, 12084 (2001).
5. C. A. Machado, R. M. Kliman, J. A. Markert, J. Hey, *Mol. Biol. Evol.* **19**, 472 (2002).
6. D. Ortiz-Barrientos, J. Reiland, J. Hey, M. A. Noor, *Genetica* **116**, 167 (2002).
7. A. Navarro, N. H. Barton, *Evolution* **57**, 447 (2003).
8. H. A. Orr, *Genetics* **144**, 1331 (1997).
9. D. N. Janczewski, D. Goldman, S. J. O'Brien, *J. Hered.* **81**, 375 (1990).
10. S. Consigliere, R. Stanyon, U. Koehler, G. Agoramoorthy, J. Wienberg, *Chromosome Res.* **4**, 264 (1996).
11. J. J. Yunis, O. Prakash, *Science* **215**, 1525 (1982).
12. D. Haig, *Philos. Trans. R. Soc. London Ser. B Biol. Sci.* **354**, 1447 (1999).
13. J. Vieira, B. F. McAllister, B. Charlesworth, *Genetics* **158**, 279 (2001).
14. R. Hudson, *Oxford Surv. Evol. Biol.* **7**, 1 (1990).
15. M. Kimura, *The Neutral Theory of Molecular Evolution* (Cambridge Univ. Press. Cambridge, 1983).
16. As shown in supporting online material, the measures had different sources. Most of them (86) were obtained by analyzing sequences retrieved from Gen-Bank; 17 by reanalyzing sequences provided by authors, and 12 directly from the literature. Whenever possible, literature data or sequences provided by authors were checked against GenBank sequences. When different sources provided different divergence estimates for the same gene or when more than one allele was available, the pairs of sequences with the smallest distances were selected for analysis, so the most ancestral variants were used. Genes with no divergence or for which ancestral polymorphism has been reported were excluded from this analysis. Also, to avoid redundant information, duplicated genes in tandem for which gene conversion has been reported were also excluded.
17. W. H. Li, *J. Mol. Evol.* **36**, 96 (1993). The maximum likelihood method implemented in the package PAML (*44*) performs better than Li's method, which we use here. Yang's method could not be used with our full data set, because the divergence measures obtained directly from the literature used Li's method. The methods, however, should not produce different results for such a short distance as the time separating humans and chimpanzees. This was confirmed by repeating our primary calculations using PAML on the reduced data set. The results obtained in this way did not differ from the ones presented here.
18. X. Xia, Z. Xie, *J. Hered.* **92**, 371 (2001).
19. F. C. Chen, W. H. Li, *Am. J. Hum. Genet.* **68**, 444 (2001).
20. The cytological positions of the breakpoints of the chromosomal changes separating the genomes of humans and chimpanzees (see table S2) and a classification of chromosomal regions were determined as previously described (*11*, *12*).
21. Note, however, that K_S patterns should be different, because the value of K_S depends on the proportion in which gene flow is reduced, whereas incompatibilities leading to higher K_A can only be established below a given threshold of gene flow.
22. S. Yi, D. L. Ellsworth, W. H. Li, *Mol. Biol. Evol.* **19**, 2191 (2002).
23. I. Ebersberger, D. Metzler, C. Schwarz, S. Paabo, *Am. J. Hum. Genet.* **70**, 1490 (2002).
24. From table 1 of (*22*), we obtained a set of 45 values of the number of nucleotide substitutions per 100 sites, *K*, between human and chimpanzee sequences, based on 447,330 bp of noncoding, nonrepetitive sequences from different chromosomes. S. Yi kindly provided the amounts of GC for these data. Second, I. Ebersberger kindly provided the *K* values and amounts of GC for the set of 8859 sequence pairs encompassing 1,944,162 bp of sequences from all the genome that were analyzed in (*23*).
25. M. J. Lercher, E. J. B. Williams, L. D. Hurst, *Mol. Biol. Evol.* **18**, 2032 (2001).
26. G. Bernardi, *Gene* **241**, 3 (2000).
27. E. J. B. Williams, L. D. Hurst, *Nature* **407**, 900 (2000).
28. ———, *Mol. Biol. Evol.* **19**, 1395 (2002).
29. J. Castresana, *Nucl. Acids Res.* **30**, 1751 (2002).
30. Their GC4 content was ascertained by means of DNAsp by J. Rozas and R. Rozas [*Bioinformatics* **15**, 174 (1999)]. Because the ancestral sequence is not available, the average GC4 content of the two sequences was used.
31. A. Eyre-Walker, *Proc. R. Soc. London Ser. B Biol. Sci.* **252**, 237 (1993).
32. A. Kong *et al.*, *Nature Genet.* **31**, 241 (2002).
33. M. Brunet *et al.*, *Nature* **418**, 145 (2002).
34. P. Vignaud *et al.*, *Nature* **418**, 152 (2002).
35. M. Przeworski, R. R. Hudson, A. Di Rienzo, *Trends Genet.* **16**, 296 (2000).
36. J. C. Stephens *et al.*, *Science* **293**, 489 (2001).
37. We obtained several conventional measures of diversity (haplotype heterozygosity; the average number of pairwise differences per site, π; and the Watterson's estimate of the population mutation parameter, θ_w) and of the deviation from the frequency spectrum expected under neutrality (Tajima's *D*) for 292 autosomal genes from the data set of variability measures in (*36*).
38. P. Andolfatto, F. Depaulis, A. Navarro, *Genet. Res.* **77**, 1 (2001).
39. M. W. Fitch, *Biochem. Genet.* **5**, 231 (1971).
40. P. Lopez, D. Casane, H. Philippe, *Mol. Biol. Evol.* **19**, 1 (2002).
41. G. S. Van Doorn, P. C. Luttikhuizen, F. J. Weissing, *Proc. R. Soc. London Ser. B* **268**, 2155 (2001).
42. M. Wallis, *J. Mol. Evol.* **53**, 10 (2001).
43. T. Johnson, N. H. Barton, *Genetics* **162**, 395 (2002).
44. Z. Yang, Phylogenetic Analysis by Maximum Likelihood (PAML), version 3.13a (University College London, London, 2002); available at http://abacus.gene.ucl.ac.uk/software/paml.html
45. We thank A. Andrés, J. Bertranpetit, F. Calafell, J. Castresana, B. Charlesworth, D. Charlesworth, F. Depaulis, T. Johnson, P. Keightley, H. Laayouni, K. Lythgoe, G. Marais, X. Maside, T. Vines, and the staff of Biomedical Systems Group for helpful discussion and criticism. We are also grateful to F.-C. Chen, I. Ebersberger, A. Eyre-Walker, J. Fay, P. Keightley, W.-H. Li, S. Paabo and S. Yi for kindly making data available. The detailed comments of J. Bertranpetit, D. Charlesworth, B. Charlesworth and F. Depaulis greatly improved a previous version of the manuscript. This research was funded by the National Environmental Research Council (UK), Ministerio de Ciencia y Tecnologia (Spain, BMC 2001-0772), and the Ramón y Cajal Program (Spain).