Early Origin and Recent Expansion of Plasmodium falciparum

Deirdre A. Joy, *et al.*

Science **300**, 318 (2003);
DOI: 10.1126/science.1081449

Science (print ISSN 0036-8075; online ISSN 1095-9203) is published weekly, except the last week in December, by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005.

sequence on the calibration curve favors the period 1050 to 1010 B.C.E. The ceramic assemblage of phase D-4a is typical Iron Age I-B pottery. Three AMS dates of seeds from locus 1845 of Phase D-4b are consistent and give a weighted average of 2924 ± 22 yr B.P. Though there are four to five possible calibrated range options in the 2σ period 1254 to 1020 calendar years B.C.E, the period 1090 to 1050 B.C.E. seems to fit best in stratigraphic sequence on the calibration curve (Fig. 3).

Phase D-6 (upper) consists of occupation debris above a floor (locus 2836) in which charred olive pits were found. Two consistent dates, both measured by conventional gas counter (GrN-26118, 2920 ± 30 yr B.P.) and AMS (GrA-18826, 2950 ± 50 yr B.P.), give a weighted average of 2928 ± 26 yr B.P. The 1σ calibrated range lists five possible periods in between 1209 to 1050 B.C.E. Considering the stratigraphic sequence and the ceramic assemblage, the period 1130 to 1090 B.C.E. on the calibration curve seems most likely (Fig. 3). The pottery is typical for the second half of the 12th century, probably slightly after the end of New Kingdom Egyptian presence in parts of Canaan, which occurred at some time in the period 1150 to 1135 B.C.E., during the reigns of Pharaohs Ramesses IV to VI (*23*, *25*).

In conclusion, the radiocarbon results, in relation to archaeological, historical, and biblical data, lead us to propose a modified traditional chronology for the Iron Age in the Levant (table S2). The modification is that the Iron Age IIA cultural period includes both the 10th and much of the 9th century B.C.E. (~980 to 835 B.C.E). There is only one known historical candidate that fits the destruction date of Tel Rehov Stratum V, 940 to 900 B.C.E., based on 12 high-quality ^{14}C dates: the invasion of Pharaoh Shoshenq I.

Our research negates an important argument of the low chronology theory, namely, that Iron Age IIA ceramic assemblages should be confined exclusively to the 9th century B.C.E. The ^{14}C dating results imply that it is difficult to distinguish between "Solomonic" and "Omride" pottery. The site of Ta'anach (*27*), about 8 km southeast of Megiddo (Fig. 1), is also mentioned on the Karnak list of places destroyed by Shoshenq. Period II-B pottery at Ta'anach, assigned to 960 to 918 B.C.E. (*27*) and to the 9th century in the low chronology (*28*), is identical to that found in Tel Rehov Stratum V. Period II-B ended in a fierce destruction, which can be related to Shoshenq's campaign in view of our results.

Because Shishak (Shoshenq I) is mentioned as a contemporary of Solomon in biblical texts, we find it plausible to retain the linkage of specified archaeological assemblages (Rehov Stratum V, Ta'anach II-B, Hazor X, Megiddo VB, and perhaps also VA-IVB, etc.) to the United Hebrew Monarchy. Our results also have implications for the chronology of Cyprus and Greece because imported pottery from both countries was found in Tel Rehov Strata V and IV. It appears that the traditional chronology of Greece can be maintained, but for Cyprus, older dates seem appropriate for some pottery groups (*29*, *30*).

References and Notes

1. M. Balter, *Science* **287**, 31 (2000).
2. I. Finkelstein, *Levant* **28**, 177 (1996).
3. A. Mazar, *Levant* **29**, 157 (1997).
4. S. Gitin, A. Mazar, E. Stern, Eds. *Mediterranean Peoples in Transition—Thirteenth to Early Tenth Centuries BCE* (Israel Exploration Society, Jerusalem, 1998).
5. I. Finkelstein, *Near East. Archaeol.* **62**, 35 (1999).
6. A. Ben-Tor, *Bull. Am. Schools Orient. Res.* **317**, 9 (2000).
7. M. Miller, J. Hayes, *A History of Ancient Israel and Judah* (Westminster, Philadelphia, 1986).
8. J. M. Weinstein, *Radiocarbon* **26**, 297 (1984).
9. G. W. Pearson, M. Stuiver, *Radiocarbon* **28**, 839 (1986).
10. M. Stuiver, B. Becker, *Radiocarbon* **28**, 863 (1986).
11. 1993 Calibration issue, M. Stuiver, A. Long, R. S. Kra, J. M. Devine, Eds., *Radiocarbon* **35**, 35–65 (1993).
12. INTCAL 98: Calibration Issue, M. Stuiver, J. van der Plicht, Eds., *Radiocarbon* **40** (1998).
13. S. W. Manning, B. Kromer, P. I. Kuniholm, M. W. Newton, *Science* **294**, 2532 (2001).
14. H. J. Bruins, W. G. Mook, *Radiocarbon* **31**, 1019 (1989).
15. Near East Chronology: Archaeology and Environment, H. J. Bruins, I. Carmi, E. Boaretto, Eds., *Radiocarbon* **43** (2001).
16. J. van der Plicht, H.J. Bruins, *Radiocarbon* **43**, 1155 (2001).
17. A. Mazar, I. Carmi, *Radiocarbon* **43**, 1333 (2001).
18. A. Gilboa, I. Sharon, *Radiocarbon* **43**, 1343 (2001).
19. Materials and methods are available as supporting material on *Science* Online.
20. S. W. Manning, B. Weninger, *Antiquity* **66**, 636 (1992).
21. B. Weninger, *Radiocarbon* **37**, 443 (1995).
22. A. Mazar, *Isr. Explor. J.* **49**, 1 (1999); www.rehov.org.
23. K. A. Kitchen, *The Third Intermediate Period in Egypt (1100 BC)* (Aris & Phillips, Warminster, ed. 2, 1986).
24. Owing to length restrictions, we are unable to present a more detailed description of each sample in its stratigraphic context, ceramic assemblages, or a discussion of published radiocarbon dates.
25. K. A. Kitchen, in *The Synchronisation of Civilisations in the Eastern Mediterranean in the Second Millennium B.C.* (Österreichischen Akademie der Wissenschaften, Vienna, 2000), pp. 39–52.
26. C. Bronk Ramsey, *Radiocarbon* **37**, 425 (1995). We used OxCal version 3.5 (2000) with a resolution of 4 and without rounded-off ranges.
27. W. E. Rast, *Ta'anach I, Studies in the Iron Age Pottery* (ASOR, Cambridge, 1978).
28. I. Finkelstein, *Tel Aviv* **25**, 208 (1998).
29. J. Smith, in preparation.
30. N. Coldstream, A. Mazar, *Isr. Explor. J.*, in press.

31. We are grateful to J. Camp for supporting the excavations at Tel Rehov. Our thanks to H.-J. Streurman, A. T. Aerts-Bijma, and S. Wijma for carefully preparing and measuring the radiocarbon samples.

Supporting Online Material

www.sciencemag.org/cgi/content/full/300/5617/315/DC1
Materials and Methods
References
Tables S1 and S2

27 January 2003; accepted 11 March 2003

Early Origin and Recent Expansion of *Plasmodium falciparum*

Deirdre A. Joy,[1]* Xiaorong Feng,[1] Jianbing Mu,[1] Tetsuya Furuya,[1] Kesinee Chotivanich,[2] Antoniana U. Krettli,[3] May Ho,[4] Alex Wang,[5] Nicholas J. White,[2] Edward Suh,[5] Peter Beerli,[6] Xin-zhuan Su[1]

The emergence of virulent *Plasmodium falciparum* in Africa within the past 6000 years as a result of a cascade of changes in human behavior and mosquito transmission has recently been hypothesized. Here, we provide genetic evidence for a sudden increase in the African malaria parasite population about 10,000 years ago, followed by migration to other regions on the basis of variation in 100 worldwide mitochondrial DNA sequences. However, both the world and some regional populations appear to be older (50,000 to 100,000 years old), suggesting an earlier wave of migration out of Africa, perhaps during the Pleistocene migration of human beings.

Estimating the timing of historical demographic events is central to resolving the question of *P. falciparum* age and genetic diversity. Recently it has been hypothesized, on the basis of an analysis of polytene chromosomes in the mosquito vector, that the African parasite population expanded dramatically ~6000 years ago due to a series of changes involving the emergence of agricultural societies and increased mosquito transmission to humans (*1*, *2*). A related hypothesis (Malaria's Eve) posits that the worldwide parasite population is only ~6000 years old, either as a result of a severe bottleneck or because the population was chronically small until that time (*3*, *4*). Furthermore, a previous study of parasite mitochondrial (mt) DNA supports a recent origin (*5*). However, this view has been challenged by recent findings that suggest the current population is much older (100,000 to 400,000 years) (*6*, *7*). To determine whether the size of the African parasite population increased dramatically ~6000 years ago and whether this event marked the beginning of the current worldwide popu-

lation, we examined mtDNA sequence variation for 100 worldwide parasite isolates.

We amplified and sequenced the 6-kilobase mt genome—consisting of three protein-coding genes (cox III, cox I, cyt b); 20 small, fragmented ribosomal RNA (rRNA) sequences (*8*); and 493 base pairs (bp) of intron sequence—from 96 parasite isolates (fig. S1) (*9*). From the aligned sequences of 100 independent isolates [including four reported previously (*5*)], we identified one insertion and 30 single nucleotide polymorphisms (SNP) (fig. S2). By sequencing a large number of isolates, we found many more substitutions than previously reported for the malaria parasite mtDNA (*5*). Among the SNPs, 17 singletons were verified by two additional amplification and sequencing rounds. Three nonsynonymous and five synonymous substitutions are from the cox III gene, two nonsynonymous and four synonymous substitutions are from the cox I gene, and six synonymous substitutions are from the cyt b gene.

The parasite mtDNA shows no signs of recombination or strong selection and has been evolving at a relatively constant rate, making it ideal for studying the evolutionary history of the parasite. We found no correlation between the linkage disequilibrium measure D' and the distance between sites ($R^2 = 0.0038$; $P = 0.112$), indicating a lack of recombination. For the three protein-coding genes, the number of synonymous substitutions per synonymous site (K_S) compared with the number of nonsynonymous substitutions per nonsynonymous site (K_N) (*10*) did not deviate from neutral expectations ($K_S > K_N$; $P < 0.0001$). Additionally, the McDonald-Kreitman test (*11*) revealed no differences in the ratios of nonsynonymous to synonymous changes within (R_P) and between (R_F) species, in this case *P. falciparum* and *Plasmodium reichenowi* (Fisher's exact test: cox III, $P = 0.83$; cox I, $P = 0.085$; cyt b, $P = 1.000$), suggesting that these genes are not under strong selection. The mtDNA has been evolving at a relatively constant rate across *P. falciparum* and *P. reichenowi* lineages (likelihood ratio test: $\chi^2 = 55.236$; $P > 0.05$), allowing us to estimate the neutral mutation rate (μ) by comparing the number of silent substitutions in the protein-coding genes and noncoding regions for the two species using the methods of (*3*). *P. falciparum* and *P. reichenowi* are thought to have diverged in Africa along with their respective human and chimpanzee hosts (*12*) as early as 7 million years ago (Ma) (*13*). However, because we cannot rule out the possibility that the parasites diverged after the human-chimpanzee split, we used both 7 Ma and 5 Ma, giving neutral mutation rates of 4.91×10^{-9} and 6.88×10^{-9} substitutions per site per year, respectively.

The distribution of South American haplotypes in the worldwide minimum-spanning network (9, *14*) suggests multiple independent colonization events (Fig. 1A). Indeed, analysis of molecular variance (AMOVA) indicated significant population subdivision within South America ($\Phi_{st} = 0.741$; $P < 0.0001$) but not the other regions, in agreement with microsatellites (*15*). Multiple unrelated founding events is one possible explanation for this pattern. In addition, a single dominant haplotype is present in both Asia and South America, although the identity of this haplotype differs between the two regions (Fig. 1, B and C). The vast majority of haplotypes are restricted to a single region, with the notable exception of haplotype 1, which is shared among all four regions (Fig. 1, B to E). Haplotype 1 is most common in Africa, where it is positioned at the center of the network (Fig. 1D). Taken together, these data suggest that haplotype 1 spread throughout the world from Africa. A separate migration from Africa is associated with haplotype 10, which has single-step connections to both South

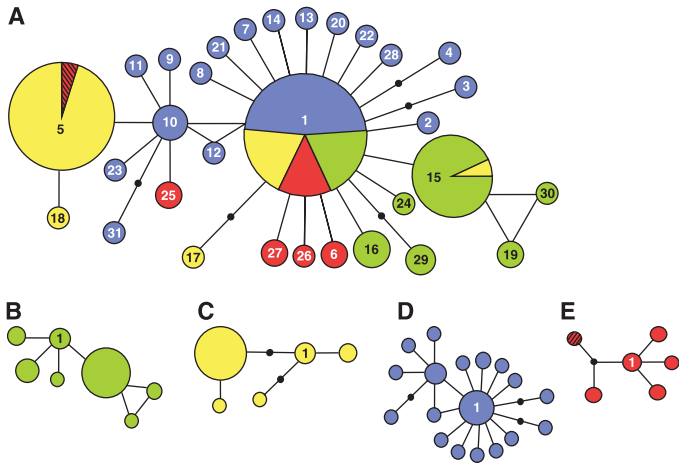

Fig. 1. Minimum spanning networks showing genetic relationships among *P. falciparum* mtDNA haplotypes. (**A**) Worldwide minimum spanning network for 100 *P. falciparum* mtDNA haplotypes. Asian, South American, African, and Papua New Guinean haplotypes are indicated with green, yellow, blue, and red, respectively. Individual networks for haplotypes found in (**B**) Asia, (**C**) South America, (**D**) Africa, and (**E**) Papua New Guinea. The star-like shape of the African network indicates population expansion. Lines represent one mutational step and black dots are hypothetical missing intermediates. Circle size is proportional to haplotype frequency, and hatch marks identify an isolate of questionable origin. Numbers refer to haplotype ID numbers in fig. S2.

Table 1. Demographic model, current effective population size, and generations since the most recent common ancestor (MRCA) or time since expansion event. Numbers in parentheses are 95% confidence intervals; numbers following ± are standard deviations. We present different measures of variance depending on the methods used to calculate the population mutational parameter θ that in turn depended on the demographic model. N_e was calculated from $\theta = N_e\mu$. We used programs to estimate θ under different models: exponential growth and constant model [FLUCTUATE (*22*)], stepwise growth [GENIE (*23*)]. Time to $N_e < 100$ was calculated for the world population using the formula $N_e = N_o e^{-rt}$. TMRCA was calculated for Asia and South America in GENETREE (*28*). Time to expansion for Africa and Papua New Guinea was estimated using τ (*14*).

Population	Model	μ ($\times 10^{-9}$) 5 Ma / 7 Ma	N_e current ($\times 10^5$)	Generations
Global	Exponential	4.91	6.72 ± 0.51	160,000 ± 29,000
		6.88	4.79 ± 0.36	110,000 ± 20,000
South America	Constant	4.91	0.55 ± 0.14	110,000 ± 38,000
		6.88	0.39 ± 0.10	77,000 ± 27,000
Asia	Constant	4.91	1.16 ± 0.36	94,000 ± 35,000
		6.88	0.83 ± 0.28	68,000 ± 25,000
Africa	Stepwise	4.91	1.49 (1.06–2.17)	30,000 (6,000–42,000)
		6.88	1.06 (0.76–1.55)	21,000 (4,000–29,000)
Papua New Guinea	Stepwise	4.91	0.53 (0.30–1.05)	38,000 (0–61,000)
		6.88	0.38 (0.22–0.75)	27,000 (0–43,000)

[1]Laboratory of Malaria and Vector Research, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD 20892–0425, USA. [2]Wellcome Trust-Mahidol University-Oxford Tropical Medicine Research Programme, Mahidol University, 420/6 Rajvithi Road, Bangkok 10400, Thailand. [3]Malaria Laboratory, Centro de Pesquisas Rene Rachou, FIOCRUZAv, Augusto de Lima 1715 Belo Horizonte, MG 30190-002, Brazil. [4]Department of Microbiology and Infectious Disease, University of Calgary, Calgary T2N 1N4, Canada. [5]Center for Information Technology, National Institutes of Health, Bethesda, MD 20892, USA. [6]School of Computational Science and Information Technology (CSIT), Dirac Science Library, Florida State University, Tallahassee, FL 32306–4120, USA.

*To whom correspondence should be addressed. E-mail: djoy@niaid.nih.gov

America and Papua New Guinea (PNG).

To determine the geographical origin of the current worldwide population, we used a simple heuristic approximation of exact root probabilities (*16*) to assign outgroup weights to each haplotype (table S1). Three of the four haplotypes with the highest outgroup weights are from Africa (haplotypes 1, 10, 12, and 15) (Fig. 1A). Haplotypes 10 and 12 are found only in Africa and haplotype 1 has a worldwide distribution, indicating that the network has an African root. An African origin also finds support on the basis of mitochondrial (*5*) and microsatellite diversity (*15*, *17*) and of the initial separation of *P. falciparum* from the African chimpanzee parasite *P. reichenowi*.

The starlike shape of the worldwide and African haplotype networks (Fig. 1, A and D), as well as significantly negative Fu's F_s (*18*), Tajima's *D* (*19*), and Fu and Li's *D* and *F* (*20*) values for these two populations (32 out of 36 tests; $0.0001 < P < 0.05$), but not the others ($P > 0.5$, all tests), strongly suggests sudden expansion. A single peak in the worldwide, African, PNG, and Asian mismatch distributions (*21*) (fig. S3) implies that all four populations have experienced rapid growth. However, the maximum likelihood estimates of the growth parameter *g* (*22*) confirmed growth only in the worldwide (11,592), African (75,758), and PNG (30,405) populations, as the approximate 95% confidence interval of *g* for the South American and Asian populations included zero (*9*). This is clearly seen in the generalized skyline plots, which track population size back through time (*9*, *23*) (Fig. 2). Both the Asian and South American plots conform well to the constant size model. The African plot shows a sudden stepwise expansion and contrasts with the exponential growth of the worldwide population. Both the stepwise and logistic growth models provide good fits to the PNG data.

Sudden expansion of the parasite population in Africa corroborates a major prediction of Coluzzi's hypothesis in which elevated rates of malaria transmission accompanied speciation of *Anopheles gambiae* (*1*, *2*). In addition, the position of haplotype 1 at the center of both the African and worldwide expansions (Fig. 1, A and D) suggests that migration to other parts of the world followed expansion in Africa, as was also predicted (*1*). The time since the African expansion began was estimated ($\tau = 1.171$) and compared to the age of the worldwide population (*9*). The time to the most recent common ancestor (TMRCA) of the worldwide population was calculated using synonymous sites from protein-coding genes and noncoding sites. We applied the Jukes-Cantor correction to all distances. The ratio of the average nucleotide difference within *P. falciparum* ($\pi_{syn} = 0.0012$) to the average nucleotide difference between *P. falciparum* and *P. reichenowi* ($K = 0.0856$), multiplied by an estimate of divergence time between the two species, in this case 5 Ma and 7 Ma, yielded a TMRCA of 70,000 years (S.D. ± 37,700) and 98,000 years (S.D. ± 52,750), respectively. These TMRCA estimates do not agree with the assertion that the *P. falciparum* emerged from Africa ~6000 years ago. Using a maximum-likelihood estimate of the current worldwide effective population size (N_e) (*9*, *22*) along with the above assumption of exponential growth, we applied the formula $N_t = N_o e^{-rt}$ to estimate the number of generations since the population was very small ($N_e < 100$) (Table 1). The ratio of TMRCA to time to $N_e < 100$ gave an approximate long-term generation time of about two generations per year.

Given our estimate of the generation time, the timing of the African expansion agrees well with Coluzzi's hypothesis (Table 1).

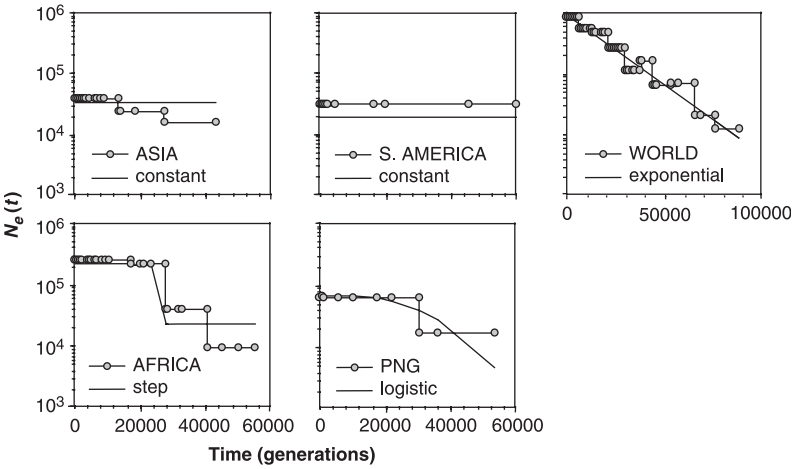

Fig. 2. Generalized skyline plots (*23*) of changes in effective population size backward in time for the regional and worldwide populations. The observed data is plotted along with one parametric model for which there was a good visual fit. Only the world, African, and Papua New Guinean populations produce good fits to expansion models. Maximum-likelihood trees assuming a molecular clock were used as input along with a neutral mutation rate of 4.91×10^{-9}.

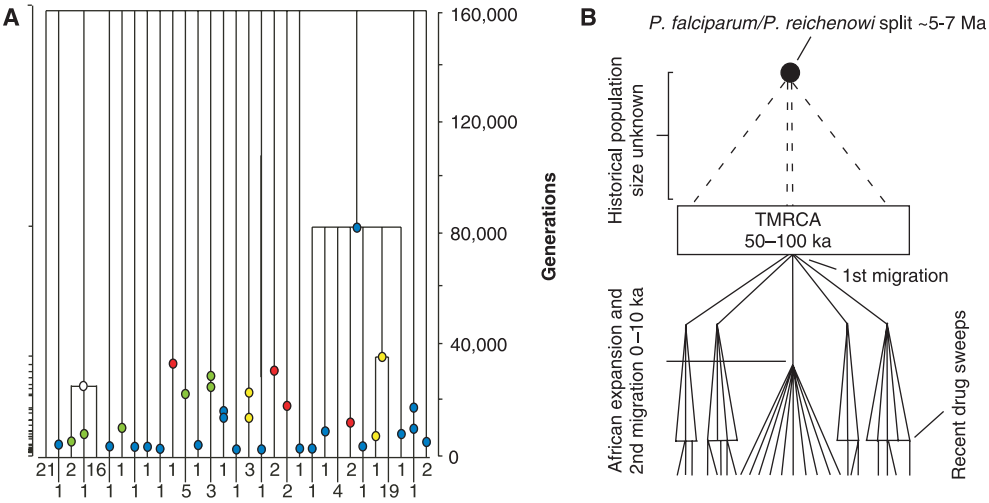

Fig. 3. (**A**) Gene tree of the *P. falciparum* mitochondrial genome. The tree is based on 1,000,000 coalescent simulations (*28*). The time scale on the right shows the TMRCA in generations using a mutation rate of 4.91×10^{-9}. SNPs are designated with a circle and are also shown as tick marks on the left axis. The tree shows the ancestral distribution of mutations and events (TMRCA, recent rapid expansion) in the population history of the parasite. Each haplotype is represented at the tips of the tree by its frequency in the total sample. Mutations that occurred in African, Asian, South American, and Papua New Guinean isolates are represented as blue, green, yellow, and red circles, respectively. One mutation was detected in Asia and South America and is shown as an open circle. (**B**) Proposed recent evolutionary history of *P. falciparum* highlighting fluctuations in population size and migration events.

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.

Supporting Online Material
www.sciencemag.org/cgi/content/full/300/5617/318/DC1
Materials and Methods
Figs. S1 to S3
Table S1

12 December 2002; accepted 11 March 2003

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., underdominant), 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