Meager genetic variability of the human malaria agent *Plasmodium vivax*

M. C. Leclerc*†, P. Durand*†, C. Gauthier*, S. Patot*, N. Billotte‡, M. Menegon§, C. Severini§, F. J. Ayala¶, and F. Renaud*‖

*Unité Mixte de Recherche 2724, Institut de Recherche pour le Développement et Centre National de la Recherche Scientifique, 911 Avenue Agropolis, 34394 Montpellier Cedex 5, France; ‡Centre de Coopération International en Recherche Agronomique pour le Développement, Avenue Agropolis, F-34398 Montpellier Cedex 5, France; §Dipartimento di Malattie Infettive, Parassitarie ed Immunomediate, Istituto Superiore di Sanità, Viale Regina Elena 299, 00161 Rome, Italy; and ¶Department of Ecology and Evolutionary Biology, University of California, Irvine, CA 92697

Contributed by F. J. Ayala, July 22, 2004

Malaria is a major human parasitic disease caused by four species of *Plasmodium* protozoa. *Plasmodium vivax*, the most widespread, affects millions of people across Africa, Asia, the Middle East, and Central and South America. We have studied the genetic variability of 13 microsatellite loci in 108 samples from 8 localities in Asia, Africa, South America, and New Guinea. Only one locus is polymorphic; nine are completely monomorphic, and the remaining three are monomorphic in all but one or two populations, which have a rare second allele. In contrast, *Plasmodium falciparum* displays extensive microsatellite polymorphism within and among populations. We further have analyzed, in 96 samples from the same 8 localities, 8 tandem repeats (TRs) located on a 100-kb contiguous chromosome segment described as highly polymorphic. Each locus exhibits 2–10 alleles in the whole sample but little intrapopulation polymorphism (1–5 alleles with a prevailing allele in most cases). Eight microsatellite loci monomorphic in *P. vivax* are polymorphic in three of five *Plasmodium* species related to *P. vivax* (two to seven individuals sampled). *Plasmodium simium*, a parasite of New World monkeys, is genetically indistinguishable from *P. vivax*. At 13 microsatellite loci and at 7 of the 8 TRs, both species share the same (or most common) allele. Scarce microsatellite polymorphism may reflect selective sweeps or population bottlenecks in recent evolutionary history of *P. vivax*; the differential variability of the TRs may reflect selective processes acting on particular regions of the genome. We infer that the world expansion of *P. vivax* as a human parasite occurred recently, perhaps <10,000 years ago.

Malaria counts among mankind's worst scourges. It is caused by four species of protozoan parasites of the genus *Plasmodium*, affects 300–500 million people, and kills more than one million people every year. Most mortality and morbidity occur in subSaharan Africa, caused by *Plasmodium falciparum*, the most virulent species. Most geographically widespread and prevalent in some regions is *Plasmodium vivax*, which accounts annually for 70–80 million clinical cases across much of Asia, Central and South America, the Middle East, and Africa. Malaria is a reemerging disease in several countries, and imported malaria is becoming a health problem in Western Europe; ≈6,500 cases are reported annually in Germany, France, Italy, and the United Kingdom. *P. falciparum* infections account for the majority of cases (64%), but *P. vivax* infections are responsible for 23% (1–4). Residual anophelinae populations capable of *P. vivax* transmission pose a permanent risk for the (re)emergence of *P. vivax* malaria where it currently does not exist, as recently observed in Central Italy (5).

Advances in malaria research have been made possible by the complete genome sequence of *P. falciparum* (6) and the rodent parasite *Plasmodium yoelii yoelii* (7); the sequence of *P. vivax* is well under way (8). Parasite control strategies depend on understanding the genetic variability and population structure of the parasites (8). Studies of genetic variability, population structure, and evolution of *P. falciparum* have recently accelerated. Less information exists for *P. vivax*, surely because of its lesser virulence but also because it cannot easily be maintained under *in vitro* continuous culture conditions (9); this has handicapped development of molecular tools for fingerprinting *P. vivax* isolates, such as exist for *P. falciparum* (10), and has negatively impacted the investigation of its population genetic structure (11).

Analyses of polymorphism and population diversity of *P. vivax* have focused on parasite molecules that are under selection by host immunity, particularly antigens homologous to those of *P. falciparum*, such as the merozoite surface protein, circumsporozoite surface protein, or erythrocyte-binding antigens (10, 12). However, some population parameters are better investigated by using DNA markers that are neutral or not under strong selection, such as microsatellite loci (13).

We herein present the analysis of 13 microsatellite loci that we have developed for investigating genetic diversity of *P. vivax* in 108 isolates from 8 localities representing different regions of the world and 26 isolates from 8 phylogenetically related *Plasmodium* species parasitic to Old World monkeys from Asia, plus *Plasmodium simium*, parasitic to New World monkeys, which is genetically indistinguishable from *P. vivax* on the basis of two protein-encoding genes (14–16). We have also investigated, for comparative purposes, the polymorphism of eight tandem repeats (TRs) located on a 100-kb DNA contiguous chromosome segment, previously described as highly polymorphic (4).

Materials and Methods

***Plasmodium*-Infected Blood Samples.** The *P. vivax* samples are from eight geographic areas and several continents. *P. vivax*-infected human blood samples were collected from Azerbaijan, Thailand, Turkey, Venezuela, and Ethiopia. Three blood samples were obtained from symptomatic individuals returning from travel in Mozambique, Papua New Guinea, and Sri Lanka. The *P. vivax* Belem strain from Brazil (17) was also analyzed. Nine *Plasmodium* species parasitic to monkeys were included to test for amplification of the different microsatellite and TR loci. Samples provided by the American Type Culture Collection (ATCC) include: *Plasmodium coatneyi* (ATCC 30128), *Plasmodium cynomolgi* (ATCC 30037, 30121, 30129, 30149, 30150, and 30155), *Plasmodium fieldi (*ATCC 30163, 30157, and 30164), *Plasmodium fragile* (ATCC 30075 and 30076), *Plasmodium hylobati* (ATCC 30154), *Plasmodium inui* (ATCC 30122, 30156, 30162, 30195, 30196, 30198, and 30199), *Plasmodium gonderi* (ATCC 30045), *Plasmodium knowlesi* (ATCC 30153, 30158, 30191, and 30192), and *P. simium* (ATCC 30130).

Microsatellite Libraries. *P. vivax* DNA was extracted by using the QIAamp DNA blood kit (Qiagen, Valencia, CA). The *P. vivax* Belem strain (17) adapted to the *Saimiri* monkey was used for developing microsatellite libraries and reference. Purified DNA of *P. vivax* from the Belem strain was kindly provided by Peter David (Institut Pasteur, Paris). Three partial genomic libraries enriched

Abbreviation: TR, tandem repeat.

†M.C.L. and P.D. contributed equally to this work.

‖To whom correspondence should be addressed. E-mail: renaud@mpl.ird.fr.

Table 1. Allele frequencies at 13 microsatellite loci in 8 populations of *P. vivax* and 9 other *Plasmodium* species

Population or species		BEL	AZE	THA	TUR	VEN	PNG	MOZ	ETH	SRL	*sim*	*cyn*	*fie*	*inu*	*hyl*	*kno*	*coa*	*fra*	*gon*
Sample size		1	40	20	19	18	1	1	7	1	1	6	3	7	1	4	1	2	1
Locus	Allele																		
AY391730	135												0.34						—
	136																1		—
	137												0.33*	1*					—
	138	1*	1*	1*	1*	1*	1	1	1	1	1*		0.33*		1				—
	139											1*				1*			—
	142																	0.50*	—
	144																	0.50*	—
AY391732	124													1	—	—		—	—
	127												0.50		—	—		—	—
	134														—	—	1	—	—
	135												0.50*		—	—		—	—
	137											0.67*			—	—		—	—
	138											0.17			—	—		—	—
	141											0.16			—	—		—	—
	150	1*	1	1	1	1	1	1	1	1	1				—	—		—	—
AY391733	128											1	—	—	—	—	—	—	—
	143	1*	1*	1	1	1	1	1	1	1	1*		—	—	—	—	—	—	—
AY391734	99												1*	—	1	—			—
	100											0.33*		—		—			—
	106											0.67*		—		—			—
	117	1*	1	1*	1	0.94	1	1	1	1	1			—		—	1		—
	119					0.06								—		—			—
	124													—		—		0.50*	—
	126													—		—		0.50	—
AY391735	141												0.50		—	—	—	—	—
	145												0.50		—	—	—	—	—
	148											0.67		1*	—	—	—	—	—
	149											0.17*			—	—	—	—	—
	155	1*	1	1	1	1	1	1	1	1	1*				—	—	—	—	—
	159											0.16			—	—	—	—	—
AY391736	128	1*	1*	1	1*	1	1	1	1	1	1*				—				—
	131											0.40	0.66		—	1	1	1	—
	144													1	—				—
	216											0.40	0.33		—				—
	218											0.20*			—				—
AY391737	160	1*	1*	1*	1*	1	1	1	1	1	1*	—	—	—		—		—	—
	165											—	—	—	1	—		—	—
	167											—	—	—		—	1	—	—

for dinucleotide (CA and GA) and tetranucleotide (GATA) repeats were constructed following described protocols (18, 19).

Microsatellite PCR Conditions. Primers of the 13 microsatellite loci were determined by using the PRIMER 3 program (available at www.genome.wi.mit.edu/cgi-bin/primer/primer3_www.cgi). Amplification was performed in a reaction mixture of 20 μl containing 2 μl of genomic DNA, 1× reaction buffer, 2.5 mM $MgCl_2$, 80 μM each dNTP, 6 pmol of each primer, and 1.3 units of *Taq*DNA polymerase (Promega). One denaturation cycle at 94°C for 2 min was followed by 30 cycles at 94°C for 20 sec, annealing temperature (varied according to locus) for 40 sec and extension at 72°C for 30 sec. The microsatellite PCR products were size-genotyped by using a standard-size Genescan 500 LIZ on an ABI Prism 310 Genetic Analyzer (PE Applied Biosystems). We checked for homology by sequencing and comparing alleles of several *Plasmodium* species to DNA sequences from *P. vivax* Belem (Table 1). Primers, core sequences, and annealing temperatures for the 13 loci are given in Table 4, which is published as supporting information on the PNAS web site.

TR PCR Conditions. Amplification was performed in a reaction mixture of 25 μl containing 1 μl (≈20 ng) of total DNA, 0.8 mM each dNTP, 20 pmol of each primer, 2.5 μl of buffer 10× (Promega), 1.5 mM Mg^{2+} (Promega), and 0.25 units of *Taq*DNA polymerase (Promega). Thermocycling with the MJ Research (Cambridge, MA) PTC100 96 well consisted of initial denaturation at 94°C for 2 min, 35 cycles at 94°C for 20 sec, at 55°C for 10 sec, at 50°C for 10 sec, at 65°C for 90 sec, and final extension at 65°C for 5 min. PCR products were size genotyped as for the microsatellites. TR sequences and primers are in Table 4.

Results

Microsatellite Variability in *P. vivax*. We developed 13 microsatellite loci from DNA sequences of the *P. vivax* Belem strain extracted from *Saimiri* monkeys, controlling that neither human nor *Saimiri* DNA amplified with any of these microsatellite loci. Table 1 displays the allele frequencies at 13 loci in 108 *P. vivax* samples. Genetic variability is very limited. Nine loci are monomorphic in every locality and in the whole sample. One locus is polymorphic, with nine alleles in the whole sample, two to six alleles in a given locality. Three loci show, each in only one or two localities, two alleles, but the second allele in low frequency. One of these three loci occurs in Venezuela, where the rare allele appears only once among 18 individual samples, and in the Thai population, where a

Table 1. (continued)

Population or species		BEL	AZE	THA	TUR	VEN	PNG	MOZ	ETH	SRL	*sim*	*cyn*	*fie*	*inu*	*hyl*	*kno*	*coa*	*fra*	*gon*
Sample size		1	40	20	19	18	1	1	7	1	1	6	3	7	1	4	1	2	1
Locus	Allele																		
AY391738	94														1		—		—
	96											0.67	1*	1**		1	—	1	—
	99											0.33*					—		—
	100	1*	1	1	1	1	1	1	1	1	1*						—		—
AY391739	144				0.05*										—		—	—	—
	146	1*	1*	1	0.95	1	1	1	1	1	1*				—		—	—	—
	148														—	1	—	—	—
	164												0.50*		—		—	—	—
	174												0.50		—		—	—	—
	186													1*	—		—	—	—
	188											0.16*			—		—	—	—
	198											0.34*			—		—	—	—
	201											0.34			—		—	—	—
	209											0.16			—		—	—	—
AY391740	152			0.15								—	—	—	—	—	—	—	—
	162					0.05*						—	—	—	—	—	—	—	—
	164	1*	1*	0.85	1*	0.95	1	1	1	1	1*	—	—	—	—	—	—	—	—
AY391742	116												—		—	1			—
	156												—		—			1	—
	157											1	—	1	—				—
	159												—		—		1		—
	164	1*	1	1*	1	1*	1	1	1	1	1*		—		—			—	
AY391743	125	1*	1*	1	1*	1*	1	1	1	1	1*	0.40*			1		—	1	—
	131											0.60*	0.50*	1*		1	—		—
	137												0.50*				—		—
AY391744	98		0.24		0.78*	0.44						—	—	—	—	—	—	—	—
	101		0.60	0.07								—	—	—	—	—	—	—	—
	104			0.36					1			—	—	—	—	—	—	—	—
	108	1*	0.03	0.22		0.39						—	—	—	—	—	—	—	—
	110		0.05	0.21				1			1*	—	—	—	—	—	—	—	—
	112			0.07	0.22*	0.11				1		—	—	—	—	—	—	—	—
	114			0.07		0.06						—	—	—	—	—	—	—	—
	116		0.08*									—	—	—	—	—	—	—	—
	128						1					—	—	—	—	—	—	—	—

Locus, GenBank accession number. Allele, allele size in base pair. Sample size, number of individual samples. Capital-letter abbreviations are for geographic populations of *P. vivax*; lower-case abbreviations are for *Plasmodium* species. BEL, Belem strain; AZE, Azerbaijan; THA, Thailand; TUR, Turkey; VEN, Venezuela; PNG, Papua New Guinea; MOZ, Mozambique; ETH, Ethiopia; SRL, Sri Lanka; *sim*, *P. simium*; *cyn*, *P. cynomolgi*; *fie*, *P. fieldi*; *inu*, *P. inui*; *hyl*, *P. hylobati*; *kno*, *P. knowlesi*; *coa*, *P. coatneyi*; *fra*, *P. fragile*; and *gon*, *P. gonderi*. —, no amplification.

*Seventy-seven alleles were sequenced and compared for similarity to DNA sequences of *P. vivax* Belem strain. (Full data available upon request from F.R.)

second allele is present in 3 of 20 individual samples. The two other loci occur in Venezuela and Turkey, where the rare allele appears only once.

The *P. vivax* genome has been partially sequenced (8). We checked by National Center for Biotechnology Information BLAST2 whether our microsatellite sequences are present in the databank European Molecular Biology Laboratory–European Bioinformatics Institute. Two microsatellite sequences (AY391732 and AY391742) appear in preexisting genomic sequences (GenBank accession nos. AZ571837 and AZ572184, respectively). None of the 13 microsatellite loci seems to be located within a known sequence of antigenic surface protein genes, which could be under strong selective pressures by host immune response.

TR Variability in *P. vivax*. Among 33 polymorphic TRs described (4), we selected eight that display strong polymorphism (see figure 3 in ref. 4). Table 2 shows allele frequencies of *P. vivax* at 8 TR loci (a total of 96 samples were analyzed; DNA templates were not available for TR analysis in 12 samples used for microsatellites). Variability ranges from 2 to 10 alleles (Table 2). In localities with 8 or more individuals, the mean number of alleles per locus varies from 2.0 to 2.9. The number of alleles per population varies from 1 to 5, with a prevailing allele in most cases (in 35 of 40 cases where more than 8 *P. vivax* individuals were analyzed, 1 allele displays a frequency ≥ 0.5; Table 2). The number of alleles per TR locus within a sample does not exceed the number observed in the most polymorphic microsatellite locus, indicating that TR polymorphism is not unusually high.

We have calculated genetic differentiation between samples of eight or more individuals. F_{st} estimator values range from 0.25 to 0.60 with strong differentiation between populations; 8 of 10 pairwise comparisons are significant (Table 3).

Genetic Variability in Related *Plasmodium* Species. The limited variability of *P. vivax* might be due to mutational constraints on the microsatellite loci. We tested this possibility by investigating several primate *Plasmodium* species with different degrees of phylogenetic relatedness to *P. vivax*. Fig. 1 displays a consensus phylogenetic tree, derived from three sets of gene sequences encoding circumsporozoite protein (15), cytochrome *b* (16, 20), and small subunit ribosomal RNA (14). The three phylogenies are consistent with one another and with the number of shared microsatellite loci (i.e., the number of loci that could be amplified) between *P. vivax* and the other *Plasmodium* species.

The *Plasmodium* species in Fig. 1 are parasitic to Old World monkeys, except *P. simium*, which parasitizes New World monkeys

EVOLUTION

Table 2. Allele frequencies at eight TR loci in eight populations of *P. vivax* and eight other *Plasmodium* species

Population		AZE	THA	TUR	VEN	ETH	*sim*	*cyn*	*fie*
Sample size		37	15	16	16	8	1	6	2
Locus	Allele								
MN1	243	1	1	1	1	1	1	—	
	273							—	0.50
	284							—	0.50
MN2	133	0.81	0.07	0.63		0.14		—	—
	152	0.19	0.80	0.37	0.94	0.72		—	—
	171		0.13		0.06	0.14		—	—
	189						1	—	—
MN7	192				0.06			—	—
	209	1	1	1	0.81	1	1	—	—
	220				0.13			—	—
MN18	174	0.11	0.93	0.88	1	0.75	1	—	—
	189	0.89	0.07	0.12		0.25		—	—
MN19	217					0.12		—	—
	221	1	1	1	0.94	0.38	1	—	—
	226				0.06	0.38		—	—
	229					0.12		—	—
MN21	228					0.38		—	—
	231					0.12		—	—
	243		0.19	0.06				—	
	254	0.91		0.38	0.76	0.25	1		—
	258							0.25	—
	266	0.03	0.21	0.12		0.25			—
	270							0.75	—
	278	0.06	0.72		0.06				—
	290		0.07		0.06				—
	301			0.31	0.06				—
MN23	190							1	
	198	0.16	0.08			0.12			
	219	0.06	0.50	0.44	0.82	0.12			
	237								1
	240	0.78	0.42	0.12		0.12	1		
	261			0.44	0.06	0.38			
	282				0.06	0.25			
	302				0.06				
MN25	146			0.25				—	
	165		0.13					—	1
	169	0.83	0.13	0.12	0.13	1	1	—	
	191	0.03		0.63	0.13			—	
	212	0.14	0.13		0.06			—	
	234		0.50		0.12			—	
	256		0.12		0.56			—	

TR loci are defined as in ref. 4. Other conventions are as in Table 1. All TR loci failed to amplify in five species (*hylobati*, *knowlesi*, *coatneyi*, *fragile*, and *gonderi*, sample size 1–4) and only MN25 amplified in *inui*, where all five individual samples share one private allele; these six species are not listed in the table. Only one individual sample was amplified for each of four populations of *P. vivax* (BEL, PNG, MOZ, SRL), which also are not listed. The Belem strain has a private allele at two loci (MN1 and MN7); the three other populations show at all eight loci one allele common in other populations, and so does the Belem strain at six loci.

Table 3. F_{st} estimators (θ) for eight TR between *P. vivax* populations, computed with FSTAT (56) (FSTAT, http://www2.unil.ch/izea/softwares/fstat. html)

	ETH	TUR	VEN	THA
AZE	0.483	0.483*	0.609*	0.594*
ETH		0.257*	0.321*	0.294
TUR			0.263*	0.282*
VEN				0.252*

AZE, Azerbaijan; ETH, Ethiopia, TUR, Turkey; VEN, Venezuela; and THA, Thailand.
*Statistically significant ($P < 0.05$) after the Bonferroni correction.

but is genetically indistinguishable from *P. vivax* at two loci (15, 16). The microsatellite data confirm this lack of genetic differentiation. First, all 13 microsatellite loci could be amplified in *P. simium*, the only species for which this is the case. Second, *P. simium* carries the same allele as *P. vivax* at the nine monomorphic loci, the most common *P. vivax* allele at the three slightly polymorphic loci, and one of the *P. vivax* alleles at the polymorphic locus (Table 1). *P. gonderi*, a parasite of African Cercopithecoidea, does not share any microsatellite (or TR) locus with *P. vivax* (Fig. 1 and Table 1). Our results confirm that *P. vivax* belongs to a monophyletic group of *Plasmodium* species that parasitize Asian Cercopithecoids (14–16, 20, 21).

It is significant that eight loci that are monomorphic (or nearly) in *P. vivax* are polymorphic in three of the five *Plasmodium* species with more than one DNA sample (Table 1 and Fig. 1; two other monomorphic, or nearly monomorphic, loci are polymorphic between species, and two loci could not be amplified). Despite the low number of DNA samples, these species are polymorphic. Among three DNA samples of *P. fieldi*, 2 or 3 alleles occur at each of 6 of the 8 amplified loci. In *P. cynomolgi,* 7 of the 10 amplified loci are polymorphic (Table 1). We have confirmed the homology between alleles by sequencing them in different *Plasmodium* species, whenever possible and comparing the DNA sequences to those of *P. vivax* Belem (Table 1). All sequences are similar among the different *Plasmodium* species. The meager genetic polymorphism of *P. vivax* may not be attributed to strong evolutionary conservation of the sequences analyzed.

We also tested the eight TR loci in the other *Plasmodium* species except for *P. hylobati* because of lack of DNA template. A main result is that, once again, *P. simium* presents alleles identical to those of *P. vivax* except at one TR locus, where *P. simium* presents one private allele. This, however, is also the case for the Belem, Venezuelan, Ethiopian, and Turkish samples of *P. vivax*, which each display at least one private allele at one TR locus (Table 2). The number of shared TR loci between the other *Plasmodium* species and *P. vivax* or *P. simium* is markedly lower than for microsatellite loci. This could reflect that the 100-kb region where the TRs are located is more polymorphic than the chromosomal regions where the microsatellites are located. This consideration strengthens genetic indistinguishability of *P. simium* and *P. vivax.* A neighbor-joining tree based on TR genetic distances between populations of *P. vivax* and *P. simium* includes *P. simium* within the *P. vivax* polymorphism. On the basis of TR variability, *P. simium,* which comes from South America, is more closely related to *P. vivax* from Azerbaijan than that from Venezuela (Fig. 2).

Discussion

Two paradoxes arise concerning microsatellite polymorphism in *P. vivax*. The first arises because of the extremely low polymor-

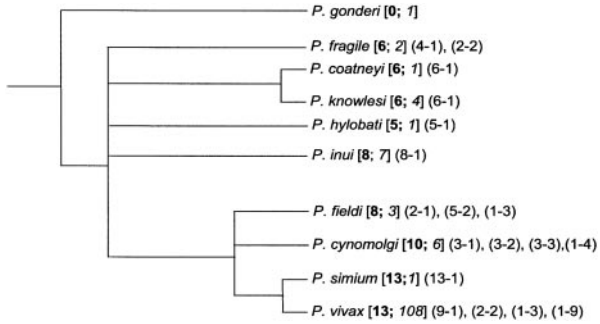

Fig. 1. Consensus phylogenetic tree of 10 *Plasmodium* species. In brackets, number of shared microsatellite loci in bold; sample size in italics. In parentheses, number of loci−number of alleles per locus.

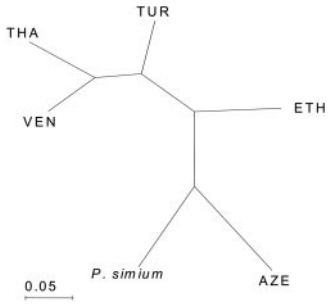

Fig. 2. Unrooted neighbor-joining tree (57) showing genetic relationships of five *P. vivax* populations and *P. simium*. AZE, Azerbaijan; ETH, Ethiopia; TUR, Turkey; VEN, Venezuela; THA, Thailand. Genetic distances are computed by using GENETIX 4.0.1 (GENETIX at www.univ-montp2.fr/~genetix/genetix/genetix.htm). The tree was constructed by using MEGA 2.1 (MEGA at http://megasoftware.net).

phism: total absence of polymorphism in most or all populations at all but one locus. The second refers to the lack of genetic differentiation in microsatellite and TR loci between *P. vivax* and *P. simium*, a New World monkey parasite, a lack of differentiation that has been previously observed for other genetic data (15, 16). The two paradoxes might be related.

Microsatellite polymorphisms arise at high rates, in other organisms as well as in *Plasmodium* (22–25), by replication slippage yielding new alleles with a different number of repeating units. The scarcity of microsatellite polymorphism in *P. vivax*, therefore, is puzzling. In the AT-rich genome of *P. falciparum* (26), $(TA)_n$ microsatellites are abundant (22–25), but they are much less frequent in *P. vivax* (4). The $(CA)_n$ repeats that predominate in many animal and plant genomes are also rare in *P. vivax* (4). [$(TA)_n$ and $(CA)_n$ refer to multiple (n) consecutive copies of the dinucleotides TA and CA.] These observations may account for the low incidence of microsatellite loci in the *P. vivax* DNA library but cannot fully account for the scarcity of polymorphism. The polymorphism at these loci is extensive within and between other *Plasmodium* species (Table 1). Moreover, microsatellites with similar ranges of repeat motif structure show high variability in *P. falciparum* (22–24) and other organisms (27). Even the one locus truly polymorphic in *P. vivax* (GenBank accession no. AY391744) does not display a large number of alleles when compared to *P. falciparum* (22–24). The same observation applies to the polymorphism observed in the TR loci, which contrasts with previous results, indicating highly polymorphic TR sequences in the *P. vivax* genome (4), even though this previous study included only five samples of *P. vivax* from India, El Salvador, Brazil (two samples), and Thailand.

The global polymorphism of *P. vivax* TRs is higher than for microsatellite loci but is relatively low within populations. This is surprising, because the *P. vivax* TR loci are located in a 100-kb region syntenic to a section of chromosome 3 of *P. falciparum*, which contains many putative genes encoding proteins and the circumsporozoite surface protein gene, likely to be under strong selective pressure, both purifying and diversifying (4). The similarity between the amplifications obtained for the primate *Plasmodium* species is much lower for TR than for microsatellite loci, indicating that the region including the TR loci may be subject to diversifying selection. Some TR loci may be located near sequences subject to selection and thus, effectively, would not be neutral.

Considerable polymorphism has been observed in *P. vivax* genes involved in drug resistance or coding for surface proteins (4, 28–33). The variability observed in the corresponding genes of *P. falciparum* was first ascribed to ancient origin (34), but the virtual absence of variation at silent (synonymous) sites in these and other protein-encoding genes is consistent with a recent population expansion of *P. falciparum* (35, 36). This hypothesis has been challenged (37, 38), but additional data (39–43) support a recent bottleneck and world expansion of *P. falciparum* (review in ref. 43). The data for *P. vivax* are less abundant, particularly for within-population polymorphism. The genes encoding the apical membrane antigen 1 (*Pv*AMA1) and the merozoite surface protein 1 (*Pv*MSP1) exhibit considerable synonymous and nonsynonymous polymorphism (32, 33). However, the rate of synonymous substitutions is eight times larger in *PvMSP1* than in *PvAMA1*, suggesting that the synonymous substitutions may not be neutral (32). That synonymous substitutions may not always be neutral is apparent in *P. vivax* (4), where synonymous substitutions are significantly more frequent than nonsynonymous substitutions at the *14070c* locus, whereas they are less frequent or absent at other loci (*14090w* and *14110w*), a result compatible with a recent population bottleneck (or selective sweep; see below) in *P. vivax*. The TR loci we have analyzed are located in the same chromosomal region as *14070c*, *14090w*, and *14110w*; they could be subject to the same evolutionary processes. Consistent with a recent world dispersal of *P. vivax* is Feng *et al.*'s (4) observation that among 37 polyT repeats and 39 polyA repeats >15 bp, only 7 are polymorphic among 5 geographically widely dispersed isolates (4), which contrasts with highly polymorphic polyT, polyA, and (TA) in *P. falciparum* (26).

Genes encoding surface antigens targeted by the host's immune response are subject to diversifying selection promoting genetic polymorphism (44). There is evidence for diversifying selection acting on erythrocyte-binding antigens in *P. vivax* as well as *P. falciparum* (12). Directional selection favoring a particular allele will reduce or eliminate genetic polymorphism around the selected site (45–47). If the selective sweep is recent and strong, a large region around the site may become genetically uniform, and the size of the unvaried region may be used to estimate the intensity of the selection and the age of the selective sweep (46, 47). The virtual absence of polymorphism at 12 microsatellite loci of *P. vivax* could be due to each locus being closely linked to a gene (site) recently subject to a selective sweep. There is no convincing direct evidence supporting these multiple recent selection events. The large number of microsatellite loci (but not alleles) shared between *P. vivax* and *P. cynomolgi* (10 loci) or *P. fieldi* (8 loci) (Fig. 1 and Table 1) suggests that the flanking regions of the microsatellite loci may be very similar in sequence and thus evolutionarily conserved.

It seems likely that *P. vivax* may have been recently subject, like *P. falciparum* (35, 36, 43, 48), to either a demographic sweep (i.e., a recent expansion of the world populations from one or few parasites) or a selective sweep, or several sweeps, because if the loci are unlinked, each one would become monomorphic only if it is closely linked to a site that has undergone a selective sweep. The polymorphisms observed at antigenic loci would then have arisen by the diversifying selective pressure of the human host's immune system, which largely occurs locally. This is consistent with the high levels of allelic differentiation between populations at TR loci (Table 3). The genetic differentiation between *P. vivax* populations, moreover, reflects the existence of only weak gene flow between them and strengthens the possibility of particular evolutionary processes that prevail locally.

The mutation rate of microsatellites is orders of magnitude higher than for nucleotide substitutions because of replication slippage that increases or decreases the number of nucleotide repeats. Yet, *P. vivax* has a complete absence (or nearly so) of polymorphism at all but one microsatellite locus, much less polymorphism than *P. falciparum* (22–25). The evidence favors a recent world expansion of *P. falciparum* from a small African propagule within the last 10,000 years (43), likely ≈6,000 years ago, in association with the spread in Africa of slash-and-burn agriculture, world climate change after the last glaciation, and the evolution of highly anthropophilic *Anopheles* mosquito vectors (35, 36, 48). The scarcity of microsatellite polymorphisms in *P. vivax* supports a recent world expansion of this parasite, which by reference to *P.*

falciparum may have occurred not earlier than 10,000 years ago, starting from a single progenitor or small propagule, whether in Africa, Asia, or America. One possibility is that *P. vivax* became a human parasite only in recent historical times, by lateral host transfer from some Old World monkey species. In such case, the ancestral *Plasmodium* species remains undiscovered; genetically, it would be all but indistinguishable from *P. vivax*. The only known *Plasmodium* species meeting this requirement is *P. simium*, a parasite of New World monkeys.

There is abundant evidence that *P. vivax* belongs phylogenetically to a monophyletic group of *Plasmodium* species parasitizing Old World monkeys (Cercopithecoidea), so that the species would have diversified concomitantly with their hosts (14–16, 20, 48, 49). Our microsatellite data are consistent with this phylogenetic position of *P. vivax*. The position of *P. simium* within this phylogenetic group is paradoxical, because *P. simium* parasitizes New World monkeys (Ceboidea), whose phylogeny is separate from that of the Cercopithecoidea. *P. vivax* and *P. simium* are genetically indistinguishable at several gene loci (15, 16, 48), an observation corroborated by the microsatellite and TR results. The *P. simium* allele found at all 12 microsatellite loci is the unique or prevailing allele found in *P. vivax*, and at the 13th locus is one of the alleles found at *P. vivax*'s only polymorphic locus (Table 1). The same observation can be made for the TR loci (Table 2 and Fig. 2). On this evidence, *P. vivax* and *P. simium* would be considered the same species were it not that they parasitize very different host species. Infected monkeys may constitute "reservoir hosts" in the ecosystems where they cohabit with humans.

A parsimonious interpretation of the previous observations is that a recent lateral transfer has occurred from humans to New World monkeys, *P. vivax* → *P. simium*. This would explain the genetic similarity between the two species but not the scarcity of microsatellite polymorphism in *P. vivax*. There are biological and historical considerations other than phylogenetic position that would favor a recent transfer from New World monkeys to humans, that is, after the human colonizations of America (15, 48). This would explain the virtual absence of microsatellite polymorphism in *P. vivax*, but then there is no simple way to account for the phylogenetic position of *P. vivax* and *P. simium* within a monophyletic group of Cercopithecoidean parasites.

Lateral transmission of *Plasmodium* parasites from monkey hosts to humans is known for several species, including *P. simium* (50), *Plasmodium brasilianum* (51), *P. cynomolgi* (52), *P. knowlesi* (53), and perhaps *Plasmodium simiovale* (54). Transmission from humans to monkeys can be accomplished experimentally (53) and may also occur naturally (55).

The direction of host transfer between *P. vivax* and *P. simium* can be settled genetically. If the transfer has occurred from Ceboidea to humans, *P. simium* is expected to have greater nucleotide polymorphism at neutral sites than *P. vivax*. Ascertaining this will require the investigation of numerous independent samples of *P. simium*, which are not currently available. Moreover, the issue may not be resolved simply, because the genetic impoverishment of *P. vivax* may be due to a recent demographic sweep, which for now we favor as the likely explanation for this impoverishment.

We thank J. L. Boyer for support during the microsatellite library's construction, F. Rousset for helpful suggestions, G. Reynier for assistance in organizing data, and J. Patarapotikul (Department of Microbiology and Immunology, Faculty of Tropical Medicine, Mahidol University, Bangkok, Thailand) and L. Urdaneta (Centro de Investigaciones Biomédicas, Universidad de Carabobo, Maracay, Venezuela) for providing the *P. vivax* Thai and Venezuelan samples, respectively. This research was supported by a grant from the European Commission, by Copernicus-2 RTD project contract ICA2-CT-2000-10046 (to C.S.), and by National Institutes of Health Grant GM42397 (to F.J.A.).

1. World Health Organization (2000) *Tech. Rep. Ser.* **892,** 1.
2. Mendis, K., Sina, B., Marchesini, P. & Carter, R. (2001) *Am. J. Trop. Med. Hyg.* **64,** 97–106.
3. Feighner, B. H., Pak, S. I., Novakoski, W. L., Kelsey, L. L. & Strickman, D. (1998) *Emerg. Infect. Dis.* **4,** 295–297.
4. Feng, X., Carlton, J. M., Joy, D. A., Mu, J., Furuya, T., Suh, B. B., Wang, Y., Barnwell, J. W. & Su, X.-Z. (2003) *Proc. Natl. Acad. Sci. USA* **100,** 8502–8507.
5. Severini, C., Menegon, M., Gradoni, L. & Majori, G. (2002) *Acta Tropica* **84,** 511–517.
6. Gardner, M. J., Hall, N., Fung, E., White, O., Berriman, M., Hyman, R. W., Carlton, M. J., Pain, A., Nelson, K. E., Bowman, S., *et al.* (2002) *Nature* **419,** 498–511.
7. Carlton, J. M., Angiuoli, S. V., Suh, B. B., Kooij, T. W., Pertea, M., Silva, J. C., Ermolaeva, M. D., Allen, J. E., Selengut, J. D., Koo, H. L., *et al.* (2002) *Nature* **419,** 512–519.
8. Carlton, J. (2003) *Trends Parasitol.* **19,** 227–231.
9. Golenda, C. F., Li, J. & Rosenberg, R. (1997) *Proc. Natl. Acad. Sci. USA* **94,** 6786–6791.
10. Cui, L., Escalante, A. A., Imwong, M. & Snounou, G. (2003) *Trends Parasitol.* **19,** 220–226.
11. Collins, W. E. & Jeffery, G. M. (1996) *Am. J. Trop. Med. Hyg.* **55,** 243–249.
12. Baum, J., Thomas, A. W. & Conway, D. (2003) *Genetics* **163,** 1327–1336.
13. Jarne, P. & Lagoda, P. J. (1996) *Trends Ecol. Evol.* **11,** 424–429.
14. Escalante, A. A. & Ayala, F. J. (1994) *Proc. Natl. Acad. Sci. USA* **91,** 11373–11377.
15. Escalante, A. A., Barrio, E. & Ayala, F. J. (1995) *Mol. Biol. Evol.* **12,** 616–626.
16. Escalante, A. A., Freeland, D. E., Collins, W. E. & Lal, A. A. (1998) *Proc. Natl. Acad. Sci. USA.* **95,** 8124–8129.
17. del Portillo, H. A., Gysin, J., Mattei, D. M., Khouri, E., Udagama, P. V., Mendis, K. N. & David, P. H. (1988) *Exp. Parasitol.* **67,** 346–353.
18. Billotte, N., Lagoda, P. J. L., Risterucci, A.-M. & Baurens, F.-C. (1999) *Fruits* **54,** 277–288.
19. Edwards, K. J., Barker, J. H. A., Daly, A., Jones, C. & Karp, A. (1996) *Biotechniques* **20,** 758–760.
20. Perkins, S. & Schall, J. (2002) *J. Parasitol.* **88,** 972–978.
21. Rathore, D., Wahl, A. M., Sullivan, M. & McCutchan, T. F. (2001) *Mol. Biochem. Parasitol.* **114,** 89–94.
22. Anderson, T. J. C., Haubold, B., Williams, J. T., Estrada-Franco, J. G., Richardson, L., Mollinedo, R., Bockarie, M., Mokili, J., Mharakurwa, S., French, N., *et al.* (2000) *Mol. Biol. Evol.* **17,** 1467–1482.
23. Durand, P., Michalakis, Y., Cestier, S., Leclerc, M. C., Tibayrenc, M. & Renaud, F. (2003) *Am. J. Trop. Med. Hyg.* **68,** 345–349.
24. Leclerc, M. C., Durand, P., de Meeüs, T., Robert, V. & Renaud, F. (2002) *Microb. Infect.* **4,** 685–692.
25. Su, X.-Z, Ferdig, M. T., Huang, Y., Huynh, C. Q., Liu, A., You, J., Wootton, J. C. & Wellems, T. E. (1999) *Science* **286,** 1351–1353.
26. Su, X. & Wellems, T. E. (1996) *Genomics* **33,** 430–444.
27. Innan, H., Terauchi, R. & Miyashita, N. (1997) *Genetics* **146,** 1441–1452.
28. Hughes, M. K. & Hughes, A. L. (1995) *Mol. Biochem. Parasitol.* **71,** 99–113.
29. Rosenberg, R., Wirtz, R. A., Lanar, D. E., Sattabongkot, J., Waters, A. P. & Prasittisuk, C. (1989) *Science* **245,** 973–976.
30. Putaporntip, C., Jongtuwiwes, S., Tanabe, K. & Thaithong, S. (1997) *Mol. Biochem. Parasitol.* **84,** 49–56.
31. Coleman, R. E., Sithiprasasna, R., Kankaew, P., Kiattibut, C., Ratanawong, S., Khuntirat, B. & Sattabongkot, J. (2002) *J. Med. Entomol.* **39,** 556–559.
32. Figtree, M., Pasay, C. J., Slade, R., Cheng, Q., Cloonan, N., Walker, J. & Saul, A. (2000) *Mol. Biochem. Parasitol.* **108,** 53–66.
33. Putaporntip, C., Jongwutiwes, S., Sakihama, N., Ferreira, M. U., Kho, W-G., Kaneko, A., Kanbara, H., Hattori, T. & Tanabe, K. (2002) *Proc. Natl. Acad. Sci. USA* **99,** 16348–16353.
34. Hughes, A. L. (1991) *Genetics* **127,** 345–353.
35. Rich, S., Licht, M. C., Hudson, R. R. & Ayala, F. J. (1998) *Proc. Natl. Acad. Sci. USA* **95,** 4425–4430.
36. Rich, R. M. & Ayala, F. J. (2000) *Proc. Natl. Acad. Sci. USA* **97,** 6994–7001.
37. Forsdyke, D. R. (2002) *Trends Parasitol.* **18,** 411–417.
38. Hughes, A. L. & Verra, F. (1998) *Genetics* **150,** 511–513.
39. Volkman, S. K., Barry, A. E., Lyons, E. J., Nielsen, K. M., Thomas, S. M., Choi, M., Thakore, S. S., Day, K. P., Wirth, D. F. & Hartl, D. L. (2001) *Science* **293,** 482–484.
40. Joy, D. A., Feng, X., Mu, J., Furuya, T., Chotivanich, K., Krettli, A. U., Ho, M., Wang, A., White, N. J., Suh, E., Beerli, P. & Su, X. Z. (2003) *Science* **300,** 318–321.
41. Hartl, D. L., Volkman, S. K., Neilsen, K. M., Barry, A. E., Day, K. P., Wirth, D. F. & Winzeler, E. A. (2002) *Trends Parasitol.* **18,** 266–272.
42. Wootton, J. C., Feng, X., Ferdig, M. T., Cooper, R. A., Mu, J., Baruch, D. I., Magill, A. J. & Su, X.-H. (2002) *Nature* **418,** 320–323.
43. Hartl, D. L. (2004) *Nat. Rev. Microbiol.* **2,** 15–22.
44. Riley, E. M., Allen, S. J., Wheeler, J. G., Blackman, M. J., Bennett, S., Takacs, B., Schonfeld, H. J., Holder, A. A. & Greenwood, B. M. (1992) *Parasite Immunol.* **14,** 50–54.
45. Hudson, R. R., Bailey, K., Skarecky, D., Kwiatowski, J. & Ayala, F. J. (1994) *Genetics* **136,** 1329–1340.
46. Hudson, R. R., Sáez, A. G. & Ayala, F. J. (1997) *Proc. Natl. Acad. Sci. USA* **94,** 7725–7729.
47. Sáez, A. G., Tatarenkov, A., Barrio, E., Becerra, N. H. & Ayala, F. J. (2003) *Proc. Natl. Acad. Sci. USA* **100,** 1793–1798.
48. Ayala, F. J., Escalante, A. A. & Rich, S. M. (1999) *Parassitologia* **41,** 55–68.
49. Qari, S. H., Shi, Y. P., Pieniazek, N. J., Collins, W. E. & Lal, A. A. (1996) *Mol. Phylogenet. Evol.* **6,** 157–165.
50. Deane, L. M., Deane, M. P. & Ferreira, N. J. (1966) *Trans. R. Soc. Trop. Med. Hyg.* **60,** 563–564.
51. Contacos, P. G., Lunn, J. S., Coatney, G. R., Kilpatrick, J. W. & Jones, F. E. (1963) *Science* **142,** 676.
52. Eyles, D. E., Coatney, G. R. & Getz, M. E. (1960) *Science* **131,** 1812–1813.
53. Chin, W., Contacos, P. G., Coatney, G. R. & Kimball, H. R. (1965) *Science* **149,** 865.
54. Qari, S. H., Shi, Y.-P., Povoa, M. M., Alpers, M. P., Deloron, P., Murphy, G. S., Harjosuwarno, S. & Lal, A. A. (1993) *J. Infect. Dis.* **168,** 1485–1489.
55. Collins, W. E. (1974) *Ad. Vet. Sci. Comp. Med.* **18,** 1–23.
56. Goudet, J. (1995) *J. Hered.* **8,** 485–486.
57. Nei, M. (1978) *Genetics* **89,** 583–590.