Proc. Natl. Acad. Sci. USA
Vol. 88, pp. 3140–3144, April 1991
Evolution

Plasmodium falciparum appears to have arisen as a result of lateral transfer between avian and human hosts

(malaria/phylogeny/small-subunit ribosomal RNA/Farenholz's rule)

A. P. WATERS*, D. G. HIGGINS†, AND T. F. MCCUTCHAN*‡

*Malaria Section, Laboratory of Parasitic Disease, National Institute of Allergy and Infectious Diseases, Building 4, National Institutes of Health, Bethesda, MD 20892; and †European Molecular Biology Organization Data Library, Postfach 10.2209, Meyerhofstrasse 1, Heidelberg 6900, Federal Republic of Germany

Communicated by Louis H. Miller, December 31, 1990 (received for review November 7, 1990)

ABSTRACT **It has been proposed that the acquisition of *Plasmodium falciparum* by man is a relatively recent event and that the sustained presence of this disease in man is unlikely to have been possible prior to the establishment of agriculture. To establish phylogenetic relationships among the *Plasmodium* species and to unravel the mystery of the origin of *P. falciparum*, we have analyzed and compared phylogenetically the small-subunit ribosomal RNA gene sequences of the species of malaria that infect humans as well as a number of those sequences from species that infect animals. Although this comparison confirmed the three established major subgroups, broadly classed as avian, simian, and rodent, we find that the human pathogen *P. falciparum* is monophyletic with the avian subgroup, indicating that *P. falciparum* and avian parasites share a relatively recent avian progenitor. The other important human pathogen, *P. vivax*, is very similar to a representative of the simian group of *Plasmodium*. The relationship between *P. falciparum* and the avian parasites, and the overall phylogeny of the genus, provides evidence of an exception to Farenholz's rule, which propounds synchronous speciation between host and parasite.**

Plasmodium falciparum is a major source of morbidity and death in the tropical and subtropical regions. The proposal that the infection of man by *P. falciparum* is recent in origin (1) stems from a number of factors relating to the virulence and epidemiology of *P. falciparum* infection. One factor central to this proposal is that for the disease to be maintained in an area there must be a high and constant rate of infection of both mosquito and man. In contrast to diseases such as smallpox, where a single infected individual can be the source of a future epidemic, if the level of transmission drops below a critical point then *P. falciparum* malaria is slowly and irrevocably eliminated from the area (2). In this light, it is certain that the establishment of the agricultural community had profound effects on the relationship among man, mosquito, and parasite. Burgeoning human populations altered local environments and served as a vast and stationary blood-meal source for the mosquito. Supportive of this scenario is solid evidence demonstrating a period of rapid evolution of the most anthropophilic and *P. falciparum*-transmitting mosquito species, *Anopheles gambiae*, coincident with the onset of agriculture (3). Only at this point would man have established the level of contact with mosquito populations necessary to have sustained the disease as we know it (4). If indeed *P. falciparum* rapidly evolved as a new species during this period of change in the human and mosquito populations, where did it come from? A clear and precise knowledge of the evolutionary history of *P. falciparum* and other species of *Plasmodium* should provide many insights and may enable a more directed approach to the study of the human pathogens of the genus.

An early classification scheme for *Plasmodium*, based on a variety of biological and morphological criteria, erected nine subgenera, three of which infect mammals, four avians, and two reptiles (5), although, given their use of a different dipteran vector, the reptilian subgenera are not thought to be true representatives of *Plasmodium* (5). This picture was expanded by the observation that the G+C content of the genomic DNA of the parasites suggested that a human parasite, *P. falciparum*, was very similar to both rodent and avian parasites (6). This evidence suggested that there may be a closer relationship between *P. falciparum* and these species than had been previously acknowledged.

We present here an analysis of the phylogenetic relationships within the genus *Plasmodium*. Our conclusions are based upon standard analyses of an alignment of asexually expressed small-subunit ribosomal RNA (SSU rRNA) sequences from a variety of species of *Plasmodium* representative of the major animal models available.§ It is apparent that *P. falciparum* is monophyletic with avian malaria species and that the genesis of this species may have arisen by an event akin to lateral transfer. This is opposed to the many examples of coevolution of host and parasite.

MATERIALS AND METHODS

Generation of the SSU rRNA Sequence Data. The full-length sequences of the asexually expressed SSU rRNA genes of *P. falciparum* (6), *P. berghei* (7), and *P. lophurae* (8) and a partial sequence of *P. vivax* (9) have been reported elsewhere. The sequences of the SSU rRNA genes for *P. gallinaceum* and *P. fragile* are reported here and aligned with the reported sequence for *Acanthamoeba castellanii* (10). DNA was cloned after mung bean nuclease digestion (*P. gallinaceum*) (11) or restriction endonuclease cleavage (*P. fragile*) of genomic DNA. The DNA was cloned into plasmid pUC9 and positive clones were selected by screening with a panel of antisense, phylogenetically conserved oligonucleotides reactive with SSU rDNA. Individual clones were checked for identity by sequencing and hybridization of predicted, species-specific antisense oligonucleotides to RNA. The clones were sequenced using standard Sequenase (United States Biochemical) protocols and the data were compiled using Staden software (Amersham). Alignments were produced using the CLUSTAL series of programs (12, 13). A partial 1767 base pairs (bp) of a *P. gallinaceum* asexually expressed SSU rRNA gene and 2052 bp of an equivalent gene

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviation: SSU rRNA, small-subunit ribosomal RNA.

§The sequences reported in this paper have been deposited in the GenBank data base [accession nos. M61722 (*P. fragile*) and M61723 (*P. gallinaceum*)].

‡To whom reprint requests should be addressed.

from *P. fragile* were characterized and added to the *Plasmodium* data base for this analysis.

Phylogenetic Analysis of the Plasmodium SSU rRNA Sequence Data Base. The full-length sequences were aligned using the CLUSTAL package (12, 13) and pairwise distances were calculated using Kimura's two-parameter model (14), with 1421 sites from the multiple alignment; only sites where none of the sequences required a gap were included. The tree topology was evaluated using the neighbor-joining method of Saitou and Nei (15). The bootstrap resampling (16) was carried out by taking 1000 random samples of sites, with replacement, from the multiple alignment. For each sample, a rooted tree was constructed: the number of trees that had *P. falciparum*, *P. lophurae*, and *P. gallinaceum* as a monophyletic group was recorded. The standard deviation of the length of the branch leading to this group was calculated by Li's method (17). To be significant at the 5% level, the observed branch length must be at least 1.96 times the standard deviation. Li's method is used for four- or five-taxon trees. To apply this method to the present tree, the following pairs of species were grouped: *P. gallinaceum* with *P. lophurae* and *P. vivax* with *P. fragile*. The sequence of *A. castellanii* was used to determine the root of the tree. The alignments used in these analyses are available on request.

Further analyses were also performed using a parsimony analysis software package, PAUP (18), configured for use on the Macintosh computer. The package allows bootstrapping of the output trees to assess confidence levels.

RESULTS

Sequences of the SSU rRNA Genes of *P. gallinaceum* and *P. fragile*. The sequences of the asexually expressed SSU rRNA genes from *P. gallinaceum* and *P. fragile* are shown in Fig. 1 aligned with the gene from *A. castellanii*, which was used as the reference outgroup for these analyses. The *Plasmodium* genes appear to be typical for members of the genus; for instance, both of them demonstrate the expanded stem–loop 41 (19) that is characteristic of *Plasmodium* (20).

Phylogenetic Analysis of the *Plasmodium* SSU Genes. Fig. 2 shows the phylogeny of the genus *Plasmodium* derived from standard analyses (14, 15) of the aligned, asexually expressed SSU rRNA sequences of six species of *Plasmodium*. One unique factor in comparing phylogenetic relationships among malaria parasites is that they contain two independently expressed rRNA operons (21). When comparisons are made, it may be important to compare types that have the same pattern of expression. We chose to analyze the asexually expressed genes, since our aim was to establish parasite evolutionary relationships in the light of the range of vertebrate host. The pattern of relative evolution suggests that the rodent *Plasmodium* parasites are the most diverged. Two major groups can be distinguished that evolved in addition to the rodent parasite line, and these can be broadly classified as avian and simian. The two human pathogens included in this analysis do not form a monophyletic grouping. Rather they show a phylogenetic relationship within the genus which suggests that they evolved into human parasites by different routes. It is clear that the avian species *P. gallinaceum* and *P. lophurae* are more closely related to the major human pathogen, *P. falciparum*, than representatives of the rodent and simian lines. The level of confidence in the branching order is highly significant. Bootstrapping revealed that the avian parasites and *P. falciparum* were monophyletic in 999 of 1000 sampled trees, which is significant to the 5% level. As a further statistical test on this branching, the standard deviation of the branch length leading to this group was calculated by the method of Li (17). To be significant at the 5% level, the observed branch length must be at least 1.96 times the standard deviation. The length of this particular branch is 0.0202 unit with a standard deviation of ± 0.004793. Other bootstrap values grouped the avian parasites together 1000 times; *P. vivax* and *P. fragile* were monophyletic in every bootstrap test. The separation of *P. berghei* appears distinct but occurs only in 536 of 1000 bootstrap trees and reflects that the branch separating it is the shortest in the tree (0.055 ± 0.048).

Phylogenetic Estimations Based Upon a Species-Specific Region of the SSU rRNA Genes of *Plasmodium*. It would be of considerable use if these relationships could be accurately assessed when a single region of species-specific sequence is used as the data set. We previously established that *Plasmodium* has a characteristic region of expanded sequence complexity (20) (designated stem/loop 41 in ref. 19). This extends over 200 nucleotides in each species characterized, and in each species this region forms an equivalent secondary structure that allows a valid phylogenetic comparison (data not shown). Alignment of the sequences comprising this region shows islands of similarity, suggesting that structurally equivalent regions are maintained (data not shown). Ten species of *Plasmodium* were evaluated (Fig. 3). Although the statistical certainty associated with the tree is greatly reduced (the avian, simian, and rodent groupings are all individually significant to the 95% level, whereas the *P. falciparum*/avian cluster is seen only 75% of the time), the validity of such an analysis is confirmed by the maintenance of the relative topology of the initial tree drawn from the data set of longer sequences. Furthermore, no unexpected phylogenetic groupings are predicted by this tree. The tree is expanded by the addition of four species which serve to emphasize the original conclusions. The two remaining authentic human pathogens, *P. malariae* and *P. ovale*, occupy an intermediate position within the phylogenetic tree. They appear to be equidistant from both the *P. vivax*/simian and the *P. falciparum*/avian groupings.

DISCUSSION

We present here a reevaluation of the phylogeny of the genus *Plasmodium* based on a comparison of the asexually expressed SSU rRNA gene sequences. The analysis is facilitated by standard statistical analytical techniques that are typically applied to such questions. The results of this study suggest that the origin of *P. falciparum* is from avian stock and that an event akin to lateral transfer was responsible for the introduction of this parasite into the human population. It is a formal possibility that the transfer occurred in the opposite direction. This is highly unlikely, given the abundance and diversity of the avian *Plasmodium* species and the apparently unique position of *P. falciparum* among the mammalian parasites (5). This study therefore would be consistent with the commonly held notion that infection by *P. falciparum* is a recent acquisition of man and possibly coincident with the onset of an agriculture-based life style (4).

It is of interest that the conclusions of an earlier study (22), which were based on an estimation of the G+C content of parasite genomic DNA, are not completely supported by this work. Such measurements have been standardly applied to estimations of phylogeny for bacteria. With *Plasmodium*, this had implied a close relationship between rodent and avian parasites and *P. falciparum*. It is impossible to explain this aberration without an understanding of the precise molecular mechanisms underlying the maintenance of the G+C content of a genome. However, in this instance the similarity of their G+C content values should be viewed either as a convergent evolutionary trait or as a property of the progenitor. This would in turn imply that the simian branch of the genus has diverged most sharply from the progenitor. Our study shows that the rodent branch of the genus is in fact the most sharply diverged, supporting the notion of convergent evolution for

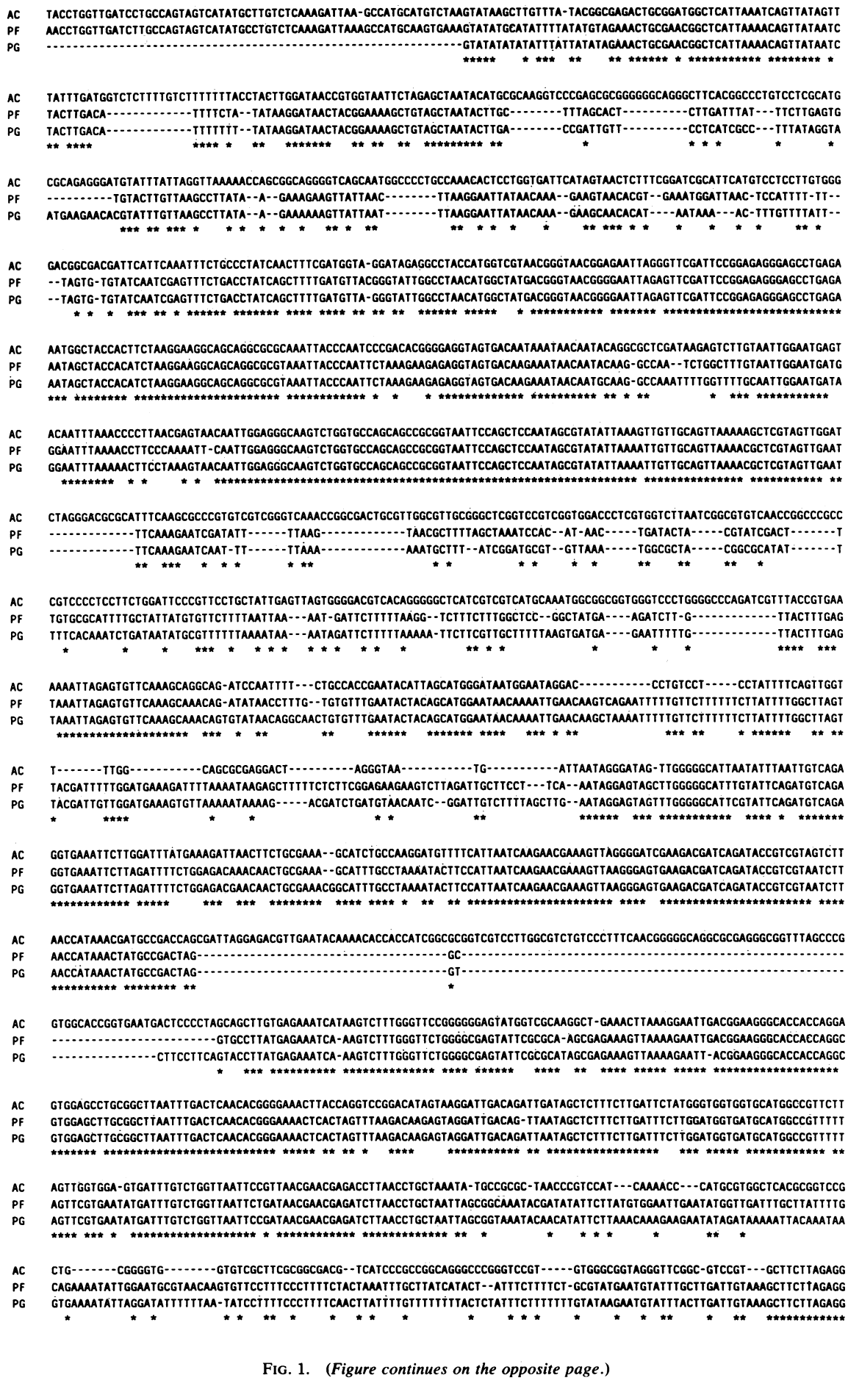

FIG. 1. (*Figure continues on the opposite page.*)

```
AC   GACTGCTGCGCGCCTAGCCAGCGGAAGTTTGAGGCAATAACAGGTCTGTGATGCCCTTAGATGTTCTGGGCCGCACGCGCGCTACACTGATTAATCCAACGAGTCCGCTTCAATCGAGGC
PF   AAC-GATGTGTGTCTAACACAAGGAAGTTTAAGGCAACAACAGGTCTGTGATGTCCTTAGATGAACTAGGCTGCACGCGTGCTACACTGATATGTATAACGAGTAC--TTAAA--AAT-T
PG   GAC-ATTGTGTGTCTAACACAAGGAAGTTTAAGGCAACAACAGGTCTGTGATGTCCTTAGATGAACTAGGCTGCACGCGTGCTACACTGATATATATAACGAGTTA--TTAAA--AATAT
      **  ** * * *** *   ******** ****** ********************* *********  ** *** ******* ***********  *  *******   ** **  *

AC   GCGATGCCGTTGGGGTCAAACCCAACTGTGTCGCTGTCCTCGATCGCGCCTGGGCCGATAGGTCCGGGTAATCTTTGCAAATTTA-ATCGTGCTGGGGATAGATCATTGTAATTATTGAT
PF   ACG------TTTGGCC--GCTT-----GCGGGCATTTCGTAC--TTTTCCTCCGCTGAAAAGTGTAGGTAATCTTTATCAATACATATCGTGATGGGGATAGATTATTGCAATTATTA--
PG   ACGAA--TGTATATAT--ATTTTATATATATACATTTTGTATAATTTTCCTCCACTGAAAAGT----------------------------------------------------------
      **     *                      * *  *       ***   * ** * **

AC   CTTCAACGAGGAATTCCTAGTAAGCGCGAGTCATCAGCTCGCGTTGATTACGTCCCTGCCCTTTGTACACACCGCCCGTCGCTCCTACCGATTGAATGGTCCGGTGAAATCCTCGGAGCC
PF   --TGAACGAGGAATGCCTAGTAAGCATGATTCATCAGATTGTGCTGACTACGTCCCTGCCCTTTGTACACACCGCCCGTCGCTCCTACCGATTGAAAGATATGATGAATTGTTTGGA---
PG   ----------------------------------------------------------------------------------------------------------------------

AC   GTGGCCTCTACGCAATCCGGGCAACCGGGCTTGTGAGGTCTCCCCTTTTGGCGGCGAAGTCGATTGAACCTTACCATTTAGAGGAAGGAGAAGTCGTAACAAGGTCTCCGTAGGTGAACC
PF   -----------CAA---GAAGAACGGGGATT---ATATCTT-TGTTTTTCTGGAAAAACCG--TAAATCCTATCTTTTAAAGGAAGGAGAAGTCGTAACAAGGTTTCCGCAGGTTCACC
PG   ----------------------------------------------------------------------------------------------------------------------

AC   TGCGGAGGGATCATTA 2305bp
PF   TACGG---------- 2052bp
PG   ---------------- 1767bp
```

FIG. 1. Alignment of the SSU rRNA genes of *P. gallinaceum* and *P. fragile*. The complete sequence of an asexually expressed gene for *P. fragile* (PF, 2052 bp) and 1767 bp of the equivalent gene from *P. gallinaceum* (PG) are aligned with the SSU rRNA gene from *A. castellanii* (AC, 2305 bp) (10). The alignments were produced using the CLUSTAL software package (12, 13). Dashed lines are introduced by the program at positions where there are no equivalent bases to facilitate alignment with the reference sequence.

G+C richness. In addition, it should be emphasized that similar secondary physical characteristics can be acquired through convergent evolution rather than through maintenance from a common progenitor. For example, the formation of knobs facilitating the deep vascular cytoadherence of schizonts in *P. falciparum* and *P. fragile* is clearly a convergence and not indicative that *P. fragile* is a suitable model for the study of *P. falciparum*.

While it is clear that the rodent parasites are quite distant from the rest of the genus, the exact position of the branch point was not possible to determine with any accuracy. The relatively extreme divergence of the rodent line may be explained by the ecology of the host/vector combination. The natural host for *P. berghei* is the predominantly terrestrial thicket rat *Thamnomys*, which occupies gallery forest in the Katanga, where primate infection is rare (5). All rodent strains in current laboratory use come from West/Central Africa and are thus quite isolated geographically (5). In addition, the vector, *Anopheles dureni*, prefers temperatures at the lower end of the tolerable range for *Anopheles* (19°C). In fact, the more usual temperatures (25–30°C) inhibit transmission (5). These factors, combined with a tendency for strains of mosquito to exhibit altitude preference, probably served to effectively isolate these rodent species early in the

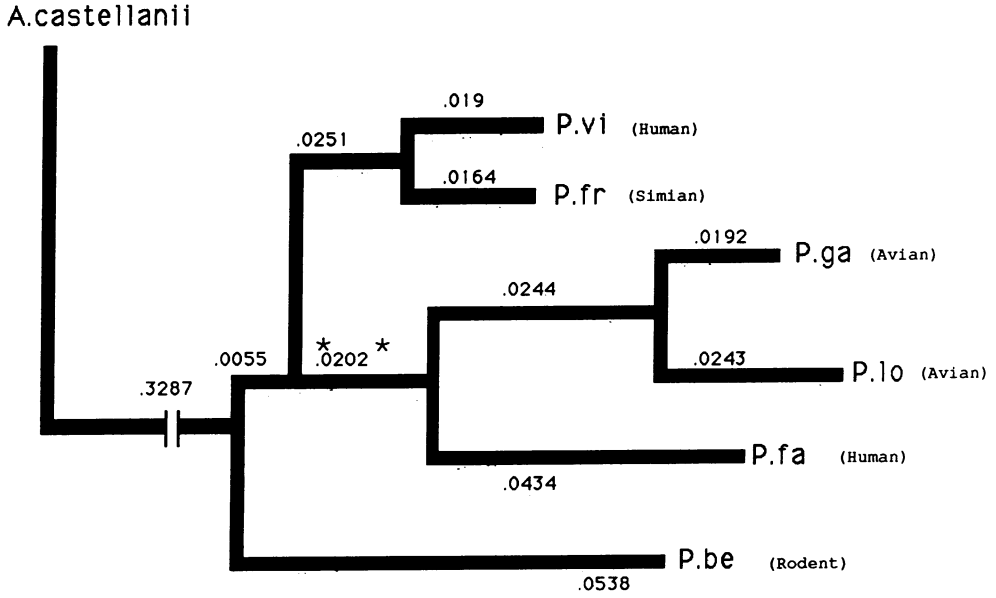

FIG. 2. Phylogenetic tree of the genus *Plasmodium* based upon asexually expressed SSU rRNA gene sequence data. Numbers above each branch reflect the branch length and are the number of nucleotide substitutions per site along that branch from node to endpoint. The grouping of *P. falciparum* with *P. gallinaceum* and *P. lophurae* (marked with asterisks) is statistically significant at the 5% level as determined by a bootstrap technique (16) also using Li's method for calculating the standard deviation of branch lengths (17). This particular branch length is 0.0202 unit with a standard deviation of ±0.004793. The same tree topology was also found using a maximum parsimony program [PAUP (18)] with a similar degree of confidence associated with the bootstrapping (99.5%). The size differences between the genes result mainly from gaps clustered in four regions (numbered approximately in *P. falciparum* as positions 175–205, 650–850, 1120–1160, and 1430–1620); these are known to be areas of distinct variability (expansion regions) within all eukaryotic SSU rRNAs. The vertebrate host for each species of *Plasmodium* is given. P.vi, *P. vivax*; P.fr, *P. fragile*; P.ga, *P. gallinaceum*; P.lo, *P. lophurae*; P.fa, *P. falciparum*; P.be, *P. berghei*.

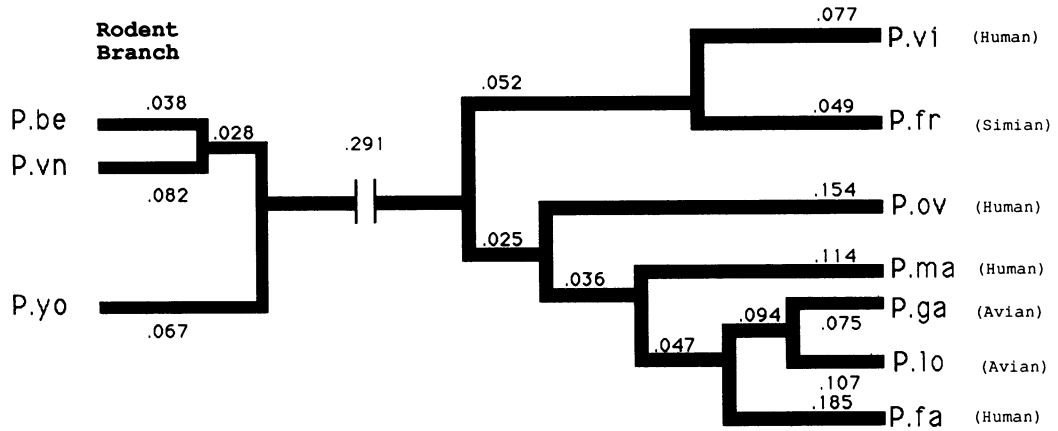

FIG. 3. Phylogenetic tree for the genus *Plasmodium* based upon a species-specific region of the asexually expressed SSU rRNA gene. The phylogenetic tree was constructed as in Fig. 2 but is not drawn to scale for reasons of clarity. The data set was subjected to the same analyses as the previous data and produced essentially the same conclusions. The same tree topology was also found using a maximum parsimony program [PAUP (18)]. The vertebrate host species are indicated. Species are labeled as in Fig. 2; in addition: P.vn, *P. vinckei*; P.yo, *P. yoelii*; P.ov, *P. ovale*; P.ma, *P. malariae*.

evolution of the genus, a possibility recognized by Garnham (5).

This study also permits an evaluation of the phylogenetic status of the other human malarial parasites. *P. vivax* is closely related to the simian malaria, *P. fragile*. Such a relationship suggests a long association of these species with their appropriate vertebrate host and possible coevolution (23). The addition of further species within this branch will elucidate the complex of parasite–host interactions and perhaps provide insight to the most appropriate animal model for the study of *P. vivax*. Further, the capacity for zoonoses could be assessed. It is apparent that both *P. ovale* and *P. malariae* seem dissimilar to any of the other species evaluated in this study. Both these parasites exhibit benign, long-term infections of man and are considered to be the most ancient human *Plasmodium* species (5). Thus their isolated position in this phylogeny may reflect the necessary extent of adaptive change. More detailed conclusions will gain validity when the full-length sequences of their SSU rRNA genes become available.

It appears from the topology of this phylogeny that *Plasmodium* has generated its extensive vertebrate host range in part by lateral transfer. It is likely that such a transfer would be mediated by cross-species challenges by sporozoites that occur frequently but that do not normally initiate active infections. The capacity of the sporozoites of avian *Plasmodium* species to invade a variety of tissues may enable the parasite to make such lateral transfers (5). As a precedent, lateral transfer of *P. lophurae* to mice has been established in the laboratory (24), as has its promiscuous ability to invade erythrocytes from various mammalian sources (25).

The overall pattern of evolution within the genus *Plasmodium* contradicts the rule established by Farenholz (26), which proposes the coevolution of parasites with their hosts. Although coevolution would certainly explain the genesis of the *P. vivax*/simian branch, the phylogeny of the *P. falciparum*/avian grouping is incompatible with the relative evolution of the avian, primate, and rodent hosts.

We thank Paul Sharp, Glenn McConkey, David Kaslow, and Louis Miller for critical reading of the manuscript, Manolo Gouy for his advice and the program for the significance testing, and Mario Coluzzi and Sam Telford for discussions. D.G.H. was supported by the European Community Biotechnology Action Programme.

1. Boyd, M. F. (1949) in *Malariology*, ed. Boyd, M. F. (Saunders, Philadelphia), Vol. 1, pp. 3–25.
2. MacDonald, G. (1957) in *The Epidemiology and Control of Malaria* (Oxford Univ. Press, London), p. 18.
3. Coluzzi, M., Petrarca, V. & Di Deco, M. A. (1985) *Boll. Zool.* **52,** 45–63.
4. Livingstone, F. B. (1958) *Am. Anthropol.* **60,** 533–560.
5. Garnham, P. C C. (1966) *Malaria Parasites and Other Haemosporidia* (Blackwell, Oxford).
6. McCutchan, T. F., de la Cruz, V. F., Lal, A. A., Gunderson, J. H., Elwood, H. J. & Sogin, M. L. (1988) *Mol. Biochem. Parasitol.* **28,** 63–68.
7. Gunderson, J. H., McCutchan, T. F. & Sogin, M. L. (1986) *Protozoology* **33,** 525–529.
8. Waters, A. P., Unnasch, T. R., Wirth, D. F. & McCutchan, T. F. (1989) *Nucleic Acids Res.* **17,** 1763.
9. Waters, A. P. & McCutchan, T. F. (1989) *Nucleic Acids Res.* **17,** 2135.
10. Gunderson, J. H. & Sogin, M. L. (1986) *Gene* **44,** 63–70.
11. McCutchan, T. F., Hansen, J. L., Dame, J. B. & Mullins, J. A. (1984) *Science* **225,** 625–628.
12. Higgins, D. G. & Sharp, P. M. (1988) *Gene* **73,** 237–244.
13. Higgins, D. G. & Sharp, P. M. (1989) *Computer Appl. Biosci.* **5,** 151–153.
14. Kimura, M. (1980) *J. Mol. Evol.* **16,** 111–120.
15. Saitou, N. & Nei, M. (1987) *Mol. Biol. Evol.* **4,** 406–425.
16. Felsenstein, J. (1985) *Evolution* **39,** 783–791.
17. Li, W.-H. (1989) *Mol. Biol. Evol.* **6,** 424–435.
18. Swofford, D. L. (1990) PAUP: *Phylogenetic Analysis Using Parsimony* (Illinois Natural History Survey, Champaign, IL), Version 3.0.
19. Neefs, J.-M., Van de Peer, Y., Hendriks, L. & De Wachter, R. (1990) *Nucleic Acids Res.* **18,** 2237–2317.
20. Waters, A. P. & McCutchan, T. F. (1989) *Lancet* **i,** 1343–1346.
21. Gunderson, J. H., Sogin, M. L., Wollett, G., Hollingdale, M., de la Cruz, V. F., Waters, A. P. & McCutchan, T. F. (1987) *Science* **238,** 933–937.
22. McCutchan, T. F., Dame, J. B., Miller, L. H. & Barnwell, J. W. (1984) *Science* **225,** 808–811.
23. Brooks, D. R. (1986) in *Parasitology—quo vadit?*, ed. Howell, M. J. (Aust. Acad. Sci., Canberra, Australia), pp. 291–297.
24. McGhee, R. B. (1951) *J. Infect. Dis.* **88,** 86–97.
25. McGhee, R. B. (1950) *Am. J. Hyg.* **52,** 42–47.
26. Eichler, W. (1948) *Ann. Mag. Nat. Hist.* **12,** 588–598.