The Small Ribosomal Subunit RNA Isoforms in *Plasmodium cynomolgi*

Vladimir Corredor¹ and Vincenzo Enea

Department of Medical and Molecular Parasitology, New York University Medical Center,
New York, New York 10010

Manuscript received May 8, 1993

Accepted for publication November 3, 1993

ABSTRACT

We report the isolation, characterization and analysis of the small subunit rRNA genes in *Plasmodium cynomolgi* (Ceylon). As in other *Plasmodium* species, these genes are present in low copy number, are unlinked and form two types that are distinct in sequence and are expressed stage specifically. The asexually expressed (type A) genes are present in four copies in the Ceylon⁻ and in five copies in the Berok⁻ strain. Surprisingly, the sexually expressed (type B) gene is present in a single copy. The vast majority of the differences between gene types is confined to the variable regions. The pattern of divergence is different from that observed in *Plasmodium berghei* or in *Plasmodium falciparum*. Analysis of the small subunit rRNA sequences of *P. cynomolgi*, *P. berghei* and *P. falciparum*, indicates that the two gene types do not evolve independently but rather interact (through gene conversion or some form of recombination) to such an extent as to erase whatever stage-specific sequence signatures they may have had in the last common ancestor.

THE organization of the ribosomal RNA genes in *Plasmodium* (the etiological agent of malaria) is unusual on several counts: the genes are few in number (4–8 copies per haploid genome), scattered on different chromosomes, and, most notably, are organized in two distinct subgroups whose expression is developmentally regulated (DAME and McCUTCHAN 1983; LANGSLEY *et al.* 1983; VAN DER PLOEG *et al.* 1985; GUNDERSON *et al.* 1987; WATERS *et al.* 1989). Type A genes are expressed in the asexual stages (which take place in the vertebrate host), and type B genes are expressed in the sexual stages (which initiate in the vertebrate host and continue in the insect vector) (GUNDERSON *et al.* 1987; WATERS *et al.* 1989). The origin and significance of this condition is not understood. The small subunit ribosomal RNA (SSU rRNA) sequence of both gene types has been determined only in *Plasmodium berghei* and *Plasmodium falciparum* (GUNDERSON *et al.* 1987; McCUTCHAN *et al.* 1988). Analysis of these sequences showed that: (1) the isoforms of a given species are more similar to one another than to the corresponding genes in the other species and (2) the pattern of conservation and divergence between isoforms is quite different in the two species (ENEA and CORREDOR 1991). To extend the database we chose to analyze a lineage, *Plasmodium cynomolgi*, which being relatively distantly related to *P. berghei* and to *P. falciparum* (DI GIOVANNI *et al.* 1990), and being composed of a number of closely related but distinct strains (COATNEY *et al.* 1971; GALINSKI *et al.* 1987), would allow for analyses involving large phyletic distances as well as short ones.

Here we report the isolation, characterization and analysis of the SSU rRNA isoforms in *P. cynomolgi* (Ceylon) and the results of some comparative work on other strains and on other species.

MATERIALS AND METHODS

Parasites: *P. cynomolgi* parasites were propagated in rhesus monkeys (*Macaca mulatta*). Infected red blood cells (IRBC) containing parasites of the Berok and Ceylon strains of *P. cynomolgi* used for pulse field gel (PFG) analysis were kindly provided by Dr. J. W. Barnwell.

Nucleic acids and restriction enzyme digestion analysis: DNA preparations from plasmid, phage and parasites as well as restriction enzyme digestions, agarose gel fractionation, blotting and hybridizations were performed as previously described (ENEA *et al.* 1984; SAMBROOK *et al.* 1989). Following hybridization with oligonucleotides, filters were washed sequentially in 2 × SSC/0.2% SDS for 30 min, 1 × SSC/0.2% SDS for 30 min and in 0.2 × SSC/0.2% SDS for 5 min at room temperature. Total RNA was extracted as previously described (ENEA *et al.* 1984) and electrophoresed on agarose/formaldehyde gels (SAMBROOK *et al.* 1989).

A *P. cynomolgi* (Ceylon strain) *EcoRI* genomic library was constructed by standard methods in lambda EMBL4 (SAMBROOK *et al.* 1989) and screened with oligonucleotides spanning various conserved or semiconserved regions of the SSU rRNA segment.

Oligonucleotides: Oligonucleotides were synthesized by standard phosphoramidite chemistry in an Applied Biosystems 381A DNA synthesizer according to the manufacturers instructions and purified on acrylamide gels. Oligonucleotides were 5' end labeled with [γ -³²P]ATP (5000 Ci/mmol) (Amersham) and polynucleotide kinase (SAMBROOK *et al.* 1989). Oligonucleotides representing sequences with different degree of conservation were used for screening genomic libraries, clone characterization and sequencing in both orientations (Table 1).

PCR amplification and cloning: The SSU rDNA complement was amplified using 1–2 ng of total parasite genomic DNA and oligonucleotides 0009 and 2134.1. Polymerase chain

¹ Present address: Department of Cell Biology and Anatomy, Box 1007, Mount Sinai Medical Center, One Gustave L. Levy Place, New York, NY 10029-6574.

TABLE 1

Catalog of oligonucleotides used in this study

Primer	Sequence ^a	Polarity	Position ^b
0001	actggcaagatcaaccaggtt	(-)	1-21
0009	tgatcttgccagtagtcatat	(+)	9-29
0096	tataactgttttaatgagccg	(-)	97-117
0382	caggetccctctccggaatcg	(-)	389-409
0401	atgacgggtaacggggaatta	(+)	363-383
0416	ctgctgcttccttagatgt	(-)	423-443
0508	gagaggtagtgacaagaata	(+)	470-490
0651	taactgcaacaattttaatat	(-)	611-631
1037	acctctgacatctgaatacaa	(-)	969-989
1126	gaaagttaagggagtgagac	(+)	1064-1084
6211	gtcttcaactcccttaactttc	(-)	1064-1084
1162	atagtttatggttaagatta	(-)	1099-1119
1171	aagactttgatttctcataag	(-)	1202-1222
1273	gcttaatttgactcaacacgg	(+)	1303-1323
1508	tggttaattccgataacgaac	(+)	1432-1452
1623	tctaacacaaggaagttaag	(+)	1715-1735
1880	gacatcacagacctgtgttg	(-)	1737-1757
1863	cttactaggcattctcgttc	(-)	1944-1964
1900	tgccctttgtacacaccgcc	(+)	1997-2017
2134	tcacctacggaaacctgttta	(-)	2133-2153
2134.1	gtaacaagggtttccgtaggtga	(+)	2133-2153
2401	aagaaagagctattatctgt	(-)	1358-1378
2401.1	tgacagattaatagctctttc	(+)	1356-1376
3553	ctttaaagatagatttacg	(-)	2099-2119
V7-84	ccgcaatttagcaaatcatatc	(-)	1562-1582
V2-114	caagacattcgtcgagagtaa	(-)	185-205
V7-174	acaatcaagcaaatatattct	(-)	1668-1688

^a Primers display different degrees of conservation varying from universally conserved to strain specific. Primers used for screening of libraries only hybridize to Plasmodium DNA as determined by aligning the corresponding region to the available SSU rRNA genes considered as more closely related to the parasite's hosts. These sequences include the human (*Homo sapiens*), and mosquito (*Aedes albopictus*) genes. All primer sequences are oriented 5' to 3' and the polarity given with respect to the rRNA sequence.

^b Primer positions refer to the sequence alignment in Figure 1.

reactions (PCR) were performed in 50- μ l reaction volumes, as described by SAKI (1990) with the following parameters: 90°/1 min denaturation, 55°/1 min annealing and 71°/2 min polymerization, for 30 cycles.

PCR products were extracted with phenol, phenol/chloroform and chloroform followed by two to three cycles of ethanol precipitation in the presence of ammonium acetate. DNA was blunt ended with Klenow fragment (New England Biolabs), phenol/chloroform and chloroform extracted and ligated to pUC 13 and pUC 18 vectors, previously cut with *Hinc*II (New England Biolabs).

Chromosome preparations and pulse field electrophoresis: Chromosomes were prepared essentially as described (SMITH *et al.* 1988). Briefly, infected red blood cells were washed several times in tissue culture medium RPMI 1640 (Sigma), resuspended at a density of 10⁹ parasites/ml in the same medium, mixed with an equal volume of molten 2% low melting agarose (IBI) and poured into block molds. Solidified blocks were mixed with at least 10 volumes of lysis buffer (1% sodium lauryl sarcosinate, 2 mg/ml proteinase K, 0.5 M EDTA pH 9) and incubated at 42-50° for 48 hr. Chromosome blocks were equilibrated with, and stored in, 100 mM Tris-HCl, pH 7.4, 50 mM EDTA at 4°.

Chromosomes were fractionated on an LKB 2015 Pulsaphor apparatus (Pulsaphor 2015; LKB Instruments, Inc.) under the following conditions: 125 sec/24-hr, 250 sec/24-hr, 450 sec/48-hr pulses, in 1 $\times$ TBE buffer, at 10°, using 1% PFG grade agarose (Boehringer) at 10 V/cm, using an hexagonal electrode array.

Southern transfer was performed by standard methods except that the gel was soaked in 0.15 N HCl for 15 min before denaturation and neutralization.

Sequence determination and analysis: The sequence of the V7 variable region (NEEFS *et al.* 1990) of Berok and Ceylon blood stage SSU rRNA was determined by primer extension of the RNA (GHOSH 1980), using a 5' end-labeled oligonucleotide (1880) and M-MLV reverse transcriptase (BRL) and the product sequenced by the MAXAM-GILBERT (1980) method as described.

Sequence of genomic clones, pUC subclones and cloned amplification products was determined by the chain termination dideoxy method (SANGER *et al.* 1977) using either Taq polymerase and 5' end-labeled primers as recommended by the manufacturer (Promega) or T7 polymerase sequencing in the presence of [³⁵S]dCTP using 7-deaza dGTP and 7-deaza dITP analogs as substitutes for dGTP as recommended by the manufacturer (Pharmacia LKB). Oligonucleotides represented in Table 1 were used for sequencing reactions in both orientations.

Sequence alignments were generated with the PILEUP/high road program available in the Genetic Computer Group (DEVEREUX *et al.* 1984) with gap opening penalties ranging from 0.5 to 5.0 and gap extension penalties equal to 0.06 times or 0.1 times the gap opening penalty. Different alignments were collated with the help of the program MALIGNED (CLARK 1991).

Parsimony and distance trees were generated with the PHYLIP package, version 3.4 (FELSENSTEIN 1985). Bootstrappings were performed with the PHYLIP programs DNABOOT for parsimony-based trees, and SEQBOOT for distance-based trees (FELSENSTEIN 1985). Alternative topologies were generated with the program DNAMOVE for parsimony (FELSENSTEIN 1985) and TREE for distance trees (OLSEN 1991).

The three pairs of A and B sequences were analyzed in detail. A number of alignments were generated, including pairwise alignments, separate alignments of the A and B sequences, alignments of the A and B sequences of pairs of species and alignments of the full set. For each of the above, several gap opening and gap extension penalties were used. These were used as inputs to generate trees by distance methods, parsimony and evolutionary parsimony.

RESULTS

Identification of two different SSU rRNA genes: genomic clones from the Ceylon strain of *P. cynomolgi* were isolated by screening a lambda library with oligonucleotide 2134. Hybridization to a number of oligonucleotides spanning the entire SSU rRNA gene, showed that the 2134-positive clones contained complete SSU rRNA genes and that all genes were of Plasmodial origin as judged by hybridization to a Plasmodium specific oligonucleotide (3553). Three clones (λ 586-84, λ 586-114 and λ 586-174) were plaque-purified and analyzed further. Physical mapping suggested that clone 84 and clone 114 were similar while clone 174 exhibited a different restriction pattern. This was confirmed by sequencing short tracts of the inserts by priming with conserved SSU rRNA oligonucleotides, adjacent to variable regions. The sequence information thus generated was used to synthesize clone-specific oligonucleotides (Table 1) with which we tested the other

FIGURE 1.—Nucleotide sequences of the mature portions of the SSU rRNA (positions 1–2171) and of the immediate flanks (–223 to 0 and 2172 to 2213). The top sequence was derived from λ 174, and, as will be shown below, represents an asexually expressed (type A) gene. The bottom sequence was derived from λ 84 and will be shown to represent the sexually expressed gene (type B). The alignment was generated with the GAP program of the GCG package (DEVEREUX *et al.* 1984) using a gap opening penalty of 2 and a gap extension penalty of 0.2. Dots indicate indels. The GenBank accession numbers are L08241 and L08242. Variable regions [*cf.* NEEFs *et al.* (1990)] are indicated above the alignment.

FIGURE 1—Continued

genomic clones. An additional clone was found to hybridize with the 174-specific oligonucleotide V7-174 while the others hybridized with the other clone specific oligonucleotides. None hybridized with both.

Sequence analysis: To facilitate sequencing, pUC 13 subclones were generated from clones 84 and 174. The complete sequence of the two SSU rRNA moieties, as well as parts of the 5' and 3' flanking regions was determined by priming with the oligonucleotides listed in Table 1. Figure 1 shows a Needelman-Wunsch alignment of the two sequences and the location of those regions known to diverge rapidly and referred as variable (V) regions or expansion segments [see NEEFS *et al.* (1990) for nomenclature and secondary structure alignment]. The vast majority of the nucleotide differences are located in V regions. The two sequences differ by 8.2% overall, while on average the variable regions differ by 15.2%. The V7 and V5 regions display the largest number of differences (Table 2). Variation in size is also observed in the variable regions, the largest, in V7, being 75 additional nucleotides in the clone 84 (Figure 1). The sequences of the immediate flanking regions (portions -223 to 0 and 2172 to 2213) show no significant similarity.

Clone specific probes from several V regions were synthesized and used to probe the SSU rRNA gene complement of other *P. cynomolgi* strains. By way of illustration, Figure 2 shows the hybridization pattern obtained with the clone 174-specific V7 probe (V7-174) and the clone

TABLE 2

Similarity between isoforms in individual variable regions

V	berA,berB	cynA,cynB	falA,falB
V1	0.1250	0.1177	0.2121
V2	0.0968	0.1879	0.2081
V3	0.0000	0.1017	0.2581
V4	0.1602	0.1496	0.2234
V5	0.1594	0.2000	0.1972
V7	0.0055	0.2295	0.0622
V8	0.0000	0.1359	0.2323
V9	0.0000	0.0952	0.0971

Variable regions are defined as per NEEFS *et al.* (1990). The V6 region is found only in prokaryotes. Differences were calculated from pairwise alignments using gap opening penalty of 2 and gap extension penalty of 0.2.

84-specific V7 probe (V7-84). It can be seen that V7-174 hybridizes to the DNA of all the strains tested while the V7-84 probe hybridizes to the DNA of some strains but yields no signal with others (strains in Figure 2 are: B, Berok; C, Ceylon; G, Gombak; M, Mulligan; N, NIH; L, London).

Clone 174 and clone 84 represent developmentally regulated genes: To study the pattern of expression of these genes, we carried out a primer extension on asexual RNA of the Berok and Ceylon strains, with primer 1880 which is adjacent to the highly variable V7 region (positions 1457 to 1624 in Figure 1). Maxam-Gilbert sequencing of the resulting products of reverse transcription revealed that the Berok SSU rRNA sequence was the same as that derived from clone 174 and that the Ceylon sequence differed at a single position (a C instead of a T at position 1562). Due to the lack of sexual stage material of the Ceylon strain and knowing that both the Mulligan and NIH strains contain the two genes represented by clones 84 and 174 (Figure 2), we hybridized clone-specific oligonucleotides to total RNA derived from the sexual and asexual stages of the NIH and Mulligan strains. As shown in Figure 3 oligonucleotide V7-174 hybridizes exclusively to RNA extracted from blood stage forms and oligonucleotide V7-84 hybridizes exclusively to RNA derived from sporozoites. Similar results were obtained with the V2 region of sporozoite and blood stages RNA and of Ceylon blood stage RNA (data not shown).

These results establish that clone 174 represents a gene that is expressed in the asexual stages (type A) and that clone 84 represents a gene that is expressed in the sexual stages (type B).

Determination of the copy number: Initial Southern blot experiments aimed at determining the organization of SSU rRNA genes in *P. cynomolgi* (Ceylon strain) showed that, as is the case in other *Plasmodium* species, the SSU rRNA genes are present in low copy number and are not tandemly repeated. Additionally, these experiments showed that all the fragments that hybridized to conserved probes, also hybridized with one or the

FIGURE 2.—Autoradiograms of type-specific oligonucleotides hybridized to the SSU rRNA gene complement generated by PCR from the Berok (B), Ceylon (C), Gombak (G), Mulligan (M), NIH (N) and London (L) strains of *P. cynomolgi*. Autoradiograms result from approximately 20 ng of target DNA per lane. Samples were electrophoresed in 1% agarose gels and transferred to nitrocellulose filters. Filters were hybridized first with the clone 174 (type A)-specific oligonucleotide V7-174 at 47° and after removal of the probe, with the clone 84 (type B)-specific oligonucleotide V7-84 at 53°. Removal of radiolabeled oligonucleotides was carried out by incubating the membrane in 0.2 N NaOH/2 mM EDTA for 2–5 min followed by neutralization in 1.5 M NaCl/0.5 M Tris-HCl pH 7.5.

other of the two sets of clone-specific probes, which suggested that the genome does not harbor SSU rRNA types other than those represented by the clone-specific oligonucleotides (not shown).

By way of illustration Figure 4 shows the hybridization pattern of Ceylon DNA digested with *Bst*BI. As can be seen in lane A, five bands of approximately equal intensities are detected with a generalized SSU rRNA probe. The same filter, hybridized with the radiolabeled clone 174-specific oligonucleotide (V7-174) detects all but the top band (lane 3A), which is detected by a clone 84-specific probe (lane 2B). A number of other hybridization experiments with other restriction enzyme digests confirmed that there are four type A genes and one type B gene (not shown). Also shown in Figure 4 (lane 2A and 1B) is the result of a hybridization experiment with a 174-specific probe which is located 5 kb upstream of the SSU rRNA. A comparison with lane 3A shows that this probe hybridizes strongly with the 9.5-kb fragment, less so with the 6.5- and 4.2-kb fragments and not at all with the fourth type A fragment. This and similar results not shown here, indicate that the four type A genes differ with regard to their context. The results obtained with the Berok strain were similar except that five type A genes were detected. Figure 5 shows the chromosomal localization of the SSU rRNA sequences in these strains. It can be seen that in both strains the SSU rRNA sequences are scattered in different chromosomes. In fact, hybridization with type-specific probes showed that only chromosome 6 in both strains contains more than one SSU rRNA sequence (a type A gene and the single type B gene, not shown). Since the type A genes were found

FIGURE 3.—Northern blot of *P. cynomolgi* RNA. Total RNA from sexual (lanes 1 and 3) and asexual (lanes 2 and 4) stages of the NIH (lanes 1 and 2) and Mulligan (lanes 3 and 4) strains was electrophoresed in 1.5% denaturing agarose gels (SAMBROOK *et al.* 1989) and transferred to a nitrocellulose filter. The filter was hybridized with clone 174-specific oligonucleotide V7-174 and, after removal of the labeled oligonucleotide by incubation at 60° for 30 min in the presence of 0.2% SDS/0.2 × SSC, it was rehybridized with clone 84-specific oligonucleotide V7-84.

to differ in their 5' flanking sequences we carried out additional experiments to determine whether and to what extent they differed in sequence. To this end we performed physical mapping of the PCR generated SSU rRNA gene complement from the Ceylon strain. Hybridization of these products with type A specific oligonucleotide V7-174 revealed an *Msp*I polymorphism in the V4 region.

To assess the degree of variability among type A genes more accurately, we sequenced two of the largest and more variable regions (V4 and V7) of a type A clone that contained the additional *Msp*I site. The sequence revealed five differences between the two type A genes in the V4 region (C/T, G/T, T/C, G/A at positions 722, 730, 844, 940 respectively and an indel involving a C at position 822. The change at position 844 generates the additional *Msp*I site). Two transitions were detected in the V7 region (T/C and A/G at positions 1690 and 1700, respectively).

Evolutionary analysis of the SSU rRNA isoforms: Figure 6 lists the number of steps required to generate the most parsimonious tree and the symmetrical trees that group together the sequences expressed in a given stage. It can be noted that in all cases the most parsimonious tree requires far fewer steps than any of the alternative trees. Even if the regions whose alignment is more open

FIGURE 4.—Hybridization of clone-specific probes to *P. cynomolgi* (Ceylon strain) genomic DNA digested with *Bst*BI. DNA was fractionated by PFG electrophoresis under the following conditions: 2 sec/20-hr pulses (panel A) and a 20-hr run of 1-, 2-, 3-, 4- and 5-sec pulses of 4 hr each (panel B), in $1 \times$ TBE at 10° in 1% PFG grade agarose at 10 V/cm using an hexagonal electrode array and transferred to nitrocellulose. Probes used for hybridizations were as follows. (1) The Ceylon SSU rRNA gene complement (2 kb in size) generated by PCR amplification with oligonucleotides 0009 and 2134 as described in methods (panel A, lane 1). (2) A 1.1-kb *Sma*I fragment representing a clone 174 specific sequence, 5.5 kb upstream of the SSU rRNA (panels A and B, lanes 2 and 1, respectively). (3) A clone 84 specific fragment representing 4.5 kb upstream of the SSU rRNA (panel B, lane 2). (4) The clone 174-specific oligonucleotide V7-174 (panel A, lane 3). Removal of radiolabeled probes was carried out by incubating the membrane in 0.2 N NaOH/2 mM EDTA for 2–5 min followed by neutralization in 1.5 M NaCl/0.5 M Tris-HCl, pH 7.5.

to questions are removed, such that the best tree requires less than half as many steps, the alternative trees still require 15–21% more steps.

The Fitch-Margoliash tree corresponding to the alignment made with gap opening penalty of 2 is shown in Figure 7. The same topology as with parsimony is obtained and with a high degree of confirmation (100/100 bootstrappings). The same results are obtained with the other two alignments. The trees obtained with the Neighbor-Joining method (SAITOU and NEI 1987) closely resemble the Fitch-Margoliash tree (not shown).

Given the high degree of similarity between *berA* and *berB*, we analyzed separately the *cyn* plus *fal* sequences. Again the isoforms grouped together. The topology was confirmed in 1000/1000 bootstrappings by parsimony and in 100/100 by distance. It scored 24 by evolutionary parsimony (LAKE 1987) *vs.* 7 for the topology that partitions the A and B genes and 0 for the third topology.

Inspection of the sequences showed the following additional points. (1) the near totality of the differences is

FIGURE 5.—Pulse field gel electrophoresis of *P. cynomolgi* Berok (B) and Ceylon (C) chromosomes. Chromosomes were fractionated, stained with ethidium bromide, transferred to a nitrocellulose filter and hybridized to nick-translated Ceylon SSU rRNA gene complement (2 kb in size) generated by PCR amplification with oligonucleotides 0009 and 2134. Bands are numbered consecutively from bottom to top. Hybridizing bands are marked with arrowheads.

found in the variable regions [*cf.* NEEFS *et al.* (1990)]; there are only 35 nucleotide substitutions outside the V regions, of which 22 are autoapomorphs, 4 support the partition [(*berA*, *berB*, *falA*, *falB*) (*cynA*, *cynB*)], 2 support [(*berA*, *berB*, *cynA*, *cynB*) (*falA*, *falB*)] and the other 7 appear only once. None support a partition of the A and B genes. (2) To determine if sequence motifs exist that are gene type specific and conserved across taxa we analyzed the 609 positions that exhibit identical morphs in at least one group of genes (*i.e.*, *berA/berB*, *cynA/cynB*, *falA/falB*, *berA/cynA/falA* and *berB/cynB/falB*). This occurs at 72%, 57% and 45% of the positions in the *ber*, *cyn* and *fal* pairs, respectively. The A genes are identical at 35% of the positions and the B genes at only 6% of the positions. Of these, only two, non consecutive residues are present exclusively in the B genes. Hence, there are no B-type specific sequence signatures. Similarly, out of 220 consensual positions of the type A genes, only six single nucleotide positions are unique to them.

DISCUSSION

The SSU rRNA isoforms in *P. cynomolgi*: As is known to be the case for the other Plasmodia species thus far

FIGURE 6.—Number of steps required to obtain the most parsimonious unrooted tree and the alternative, symmetrical trees that partition the A and B genes from three alignments. A and B refer to the corresponding gene types; ber, cyn and fal to the corresponding organisms (*P. berghei*, *P. cynomolgi* and *P. falciparum*). The number of additional steps that are required to generate the alternative topologies are listed in parentheses. Alignment P16g2 was made using gap opening penalty of 2 and gap extension penalty of 0.2. The resulting topology was confirmed in 1000/1000 bootstrappings (FELSENSTEIN 1985). P16g5 was made with gap opening and gap extension penalties of 5 and 0.5. The ber, cyn and fal groupings were confirmed 1000, 878 and 634 times, respectively, out of 1000 bootstrappings. P16g5.ED was generated from P16g5 by removal of the most size-variable regions. Bootstrap values for this alignment were 1000, 895 and 780 for the ber, cyn and fal groupings, respectively.

studied (GUNDERSON *et al.* 1987; WATERS *et al.* 1989) the SSU rRNA genes in *P. cynomolgi* are few in number, unlinked and organized in two distinct, developmentally controlled subsets. In both strains analyzed, Berok and Ceylon, the type B gene is present in a single copy. The obvious possibility of gene amplification has thus far remained unsubstantiated (CORNELISSEN 1985; MCCUTCHAN 1986; our unpublished data). In this context it may be significant that *P. cynomolgi* also harbors in single copy a gene, the hsp70, which as a rule, is present in several to many copies (ECKERT *et al.* 1992).

Our results indicate that the 4/5 copies of the A genes are nearly identical and that a V7 type A specific probe hybridizes to all the strains tested. In contrast, the B gene diverges more rapidly; for instance, the V2 and V7 type B probes fail to hybridize to some of the strains tested (Figure 2).

This difference need not reflect different degrees of constraints in the two gene types and can be understood in terms of the organization and copy number of these genes as follows. The 4/5 copies of the type A gene evolve in near complete concert. Since these genes are unlinked, their concerted evolution requires some form of sequence interaction (conversion) *other than* unequal crossing over. But in contrast to unequal crossing

over, conversion is bound to become biased, as the viable variants that are more apt at converting, as opposed to being converted, spread. Since conversion between the quasi-identical A genes is bound to be more efficient than conversion between type A and type B genes (which will be addressed below) (LISKAY *et al.* 1987), the net result will be that, even if individually the 4/5 A genes diverge as rapidly as the single B gene, the newly arisen A variants will be increasingly more likely to be converted back than to spread.

Evolutionary profile of the plasmodial SSU rRNA: The origin and significance of developmentally regulated rRNA genes are unknown. Three considerations suggest that this trait is ancestral (plesiomorphic): (1) it is shared by all plasmodial species thus far studied, (2) it is utterly rare and (3) the only other known instance of it seems to occur in *Babesia bigemina* (REDDY *et al.* 1991), which is also a member of the class Sporozoa (LEVINE 1985). Thus, the trait may actually be, a shared ancestral character (symplesiomorphic). Furthermore, in as much as the trait involves directly the ontogenetic plan, it should be evolutionarily stable [*e.g.*, GRANT (1991, p. 361)]. From this it follows that the SSU rRNA genes of a given stage specificity should be orthologous (and hence more closely related) and the two

FIGURE 7.—Distance-based cladogram. Differences were computed from an alignment using gap opening penalty of 2 and gap extension penalty of 0.2 (see methods for details) and converted to evolutionary distances using the JUKES-CANTOR (1969) model. Trees were generated by using the Fitch-Margoliash (FITCH and MARGOLIASH 1967) method. Branch lengths are measured as average fixation/100 residues. The numbers at the forks refer to the outcome of 100 bootstrappings (FELSENSTEIN 1985).

gene types of a given taxon should be paralogous. The evolutionary analysis of the sequences, however, shows beyond doubt that the isoforms of each taxon group together (Figure 7).

Thus one must assume either that the establishment of developmentally regulated rRNA genes took place separately in the three lineages or that sequence homogenization between paralogues is extensive enough to exceed the phylogenetic signal. In their most contrasted form the two hypotheses can be formulated as follows: (1) the cladogram reflects the genealogy of the sequences, which evolve independently and (2) the cladogram reflects the largely concerted evolution of the isoforms.

If the isoforms evolve independently, the less constrained regions of the genes should exhibit a higher extent of divergence in all the taxa. Contrariwise, if the isoforms exchange sequences, we should expect the less constrained regions of the genes to undergo divergent evolution and extensive sequence homogenization in essentially a stochastic manner.

Table 2 lists the evolutionary distances between isoforms for each variable region. It is evident that the extent of divergence between isoforms varies erratically. Thus, V7 is the most divergent of the regions in the cynA/cynB pair and the least in the falA/falB pair. Conversely, the most divergent region in falA/falB (V3) is unvaried in berA/berB. V8 is also unvaried in berA/berB and quite diverged in cyn and fal.

These data are plainly consistent with the second but not with first hypothesis. Why, then, do plasmodia exhibit this rare trait? This may be a reflection of how the ontogenetic plan is organized. A large number of plasmodial genes, including house keeping genes, are expressed in a developmentally controlled fashion (*e.g.*,

WESSELING *et al.* 1989; DELVES *et al.* 1990; MEIER *et al.* 1992). In this light the present result can be taken as yet another manifestation of the tendency in *Plasmodium* to organize the ontogenetic plan in two stage-specifically expressed subgenomes.

Finally we point out that insofar as these sequences evolve in a semi-concerted manner (as opposed to fully concertedly or fully independently), a gene tree based on them has a high probability of not corresponding to the species tree (SANDERSON and DOYLE 1992). Hence, the generally reliable SSU rRNA sequences may be unsuitable for assessing phyletic relationships *within* the *Plasmodium* genus (CORREDOR and ENEA 1993).

We gratefully acknowledge Dr. John E. Hill's friendly and skillful assistance with the mainframe computer. We thank Ms. Carmen Diaz for her generous help during various stages of this work, Dr. Mauricio Calvo for kind help with various computer problems and Dr. Ruth S. Nussenzweig for support. Computing was supported by National Science Foundation grant DIR 8908095. This work was supported by the McArthur Foundation.

LITERATURE CITED

- CLARK, S., 1991 Using MALIGNED to display multiple sequence alignments. Emphasizing regions of conservation and divergence. Maligned version 1.64 Manual.
- COATNEY, G. R., W. E. COLLINS, M. WARREN and P. G. CONTACOS, 1971 *The Primate Malarias*. National Institute of Allergy and Infectious Diseases, Bethesda, Md.
- CORNELISSEN, A. W. C. A., 1985 Gametocytogenesis and ribosomal rRNA gene organization in the rodent malarial *Plasmodium chabaudi* and *Plasmodium falciparum*. *Mol. Biochem. Parasitol.* **14**: 165–174.
- CORREDOR, V., and V. ENEA, 1993 Plasmodial ribosomal RNA as phylogenetic probe: a cautionary note. *Mol. Biol. Evol.* **10**: 924–926.
- DAME, J. B., and T. F. MCCUTCHAN, 1983 The four ribosomal DNA units of the malaria parasite *Plasmodium berghei*: identification, restriction map, and copy number analysis. *J. Biol. Chem.* **258**: 6984–6990.
- DELVES, C. J., P. ALANO, R. G. RIDLEY, M. GOMAN, S. P. HOLLOWAY *et al.*, 1990 Expression of α and β tubulin genes during the asexual and sexual blood stages of *Plasmodium falciparum*. *Mol. Biochem. Parasitol.* **43**: 271–278.
- DEVEREUX, J., M. HAEBERLI and O. SMITHIES, 1984 A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Res.* **12**: 387–395.
- DI GIOVANNI, L., A. H. COCHRANE and V. ENEA, 1990 On the evolutionary history of the circumsporozoite protein in Plasmodia. *Exp. Parasitol.* **70**: 373–381.
- ECKERT, V., L. SANCHEZ, A. COCHRANE and V. ENEA, 1992 *Plasmodium cynomolgi*: the hsp 70 gene. *Exp. Parasitol.* **75**: 323–328.
- ENEA, V., and V. CORREDOR, 1991 The evolution of plasmodial stage-specific rRNA genes is dominated by gene conversion. *J. Mol. Evol.* **32**: 183–186.
- ENEA, V., D. E. ARNOT, E. C. SCHMIDT, A. COCHRANE, R. GWADZ *et al.*, 1984 Circumsporozoite gene of *Plasmodium cynomolgi* (Gombak): cDNA cloning and expression of the repetitive circumsporozoite epitope. *Proc. Natl. Acad. Sci. USA.* **81**: 7520–7524.
- FELSENSTEIN, J., 1985 Confidence limits on phylogenies: an approach using the bootstrap method. *Evolution* **39**: 783–791.
- FITCH, W. M., and E. MARGOLIASH, 1967 Construction of phylogenetic trees. *Science* **155**: 279–284.
- GALINSKI, M. R., D. E. ARNOT, A. H. COCHRANE, J. W. BARNWELL, R. S. NUSSENZWEIG, *et al.*, 1987 The circumsporozoite gene of the *Plasmodium cynomolgi* complex. *Cell* **48**: 311–319.
- GHOSH, P. K., V. B. REDDY, M. PIATAK, P. LEBOWITZ and S. M. WEISSMAN, 1980 Determination of RNA sequences by primer directed synthesis and sequencing of their cDNA transcripts. *Methods Enzymol.* **65**: 580–595.

- GRANT, V., 1991 *The Evolutionary Process*. Columbia University Press, New York.
- GUNDERSON, J. H., M. L. SOGIN, G. WOLLETT, M. HOLLINGDALE, V. F. DE LA CRUZ *et al.*, 1987 Structurally distinct, stage specific ribosomes occur in *Plasmodium*. *Science* **238**: 933–937.
- JUKES, T. H., and C. R. CANTOR, 1969 Evolution of protein molecules, pp. 21–132 *Mammalian Protein Metabolism*, edited by H. R. Munro. Academic Press, New York.
- LAKE, J. A., 1987 A rate-independent technique for analysis of nucleic acid sequences: evolutionary parsimony. *Mol. Biol. Evol.* **26**: 59–73.
- LANGSLEY, G., J. E. HYDE, M. GOMAN and J. SCAIFE, 1983 Cloning and characterisation of the rRNA genes from the human malaria parasite *Plasmodium falciparum*. *Nucleic Acids Res.* **11**: 8703–8717.
- LEVINE, N. D., 1985 Phylum Apicomplexa, Levine, 1970, pp. 322–374 in *An Illustrated Guide to the Protozoa*, edited by J. J. LEE, S. H. HUTNER and E. C. BOVEE. Society of Protozoologists, Allen Press, Lawrence, Kan.
- LISKAY, M. R., A. LETSOU and J. L. STACHELEK, 1987 Homology requirement for efficient gene conversion between duplicated chromosomal sequences in mammalian cell. *Genetics* **115**: 161–167.
- MAXAM, A., and W. GILBERT, 1980 A new method of sequencing DNA. *Proc. Natl. Acad. Sci. USA* **74**: 560–564.
- MCCUTCHAN, T. F., 1986 The ribosomal gene of *Plasmodium*. *Int. Rev. Cytol.* **99**: 295–309.
- MCCUTCHAN, T. F., V. F. DE LA CRUZ, A. A. LAL, J. H. GUNDERSON, H. J. ELWOOD *et al.*, 1988 Primary sequence of two small subunit ribosomal RNA genes from *Plasmodium falciparum*. *Mol. Biochem. Parasitol.* **28**: 63–68.
- MEIER, B., H. DOBELI and U. CERTA, 1992 Stage-specific expression of aldolase isoenzymes in the rodent malaria parasite *Plasmodium berghei*. *Mol. Biochem. Parasitol.* **52**: 15–28.
- NEEFS, J. M., Y. VAN DE PEER, L. HENDRIKS and R. DE WACHTER, 1990 Compilation of small ribosomal subunit RNA sequences. *Nucleic Acids Res.* **18**: 2237–2318.
- OLSEN G. J., 1991 Phylogenetic tree inference program. Version 3.0.4 manual.
- REDDY, G. R., D. CHAKRABARTI, C. A. YOWELL and J. B. DAME, 1991 Sequence microheterogeneity of the three small subunit ribosomal RNA genes of *Babesia bigemina*: expression in erythrocyte culture. *Nucleic Acids Res.* **19**: 3641–3645.
- SAIKI, R. K. 1990 Amplification of genomic DNA, in *PCR Protocols. A Guide to Methods and Applications*, edited by M. A. INNIS, D. H. GELFAND, J. J. SNINSKY and T. J. WHITE. Academic Press, San Diego, Calif.
- SAITOU, N., and M. NEI, 1987 The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **4**: 406–425.
- SAMBROOK, J., E. F. FRITSCH and T. MANIATIS, 1989 *Molecular Cloning: A Laboratory Manual*, Ed. 2. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
- SANDERSON, M. J., and J. J. DOYLE, 1992 Reconstruction of organismal and gene phylogenies from data on multigene families: concerted evolution, homoplasy, and confidence. *Syst. Biol.* **41**: 4–17.
- SANGER, F., S. NICKLEN, and A. R. COULSON, 1977 DNA sequencing with chain terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74**: 5463–5467.
- SMITH, C. L., S. R. KLCO and C. R. CANTOR, 1988 Pulsed-field gel electrophoresis and the technology of large DNA molecules, in *Genome Analysis. A Practical Approach*, edited by K. E. Davies IRL Press, Oxford.
- VAN DER PLOEG, L. H. T., M. SMITS, T. PONNUDURAI, A. VERMEULEN, J. H. E. MEUWISSEN *et al.*, 1985 Chromosome-sized DNA molecules of *Plasmodium falciparum*. *Science* **229**: 658–661.
- WATERS, A. P., C. SYN and T. F. MCCUTCHAN, 1989 Developmental regulation of stage-specific ribosome populations in *Plasmodium*. *Nature* **342**: 438–441.
- WESSELING, J. G., P. J. F. SNYDERS, P. VAN SOMEREN, J. JANSSEN, M. A. SMITS *et al.*, 1989 Stage specific expression and genomic organization of the actin genes of the malaria parasite *Plasmodium falciparum*. *Mol. Biochem. Parasitol.* **35**: 167–176.

Communicating editor: S. L. ALLEN