Proc. Natl. Acad. Sci. USA
Vol. 91, pp. 619–623, January 1994
Microbiology

The polymorphic subtelomeric regions of *Plasmodium falciparum* chromosomes contain arrays of repetitive sequence elements

(malaria/recombination/genome organization)

DERIK DE BRUIN, MICHAEL LANZER, AND JEFFREY V. RAVETCH*

DeWitt Wallace Research Laboratory, Division of Molecular Biology, Sloan-Kettering Institute, 1275 York Avenue, New York, NY 10021

Communicated by Maclyn McCarty, October 7, 1993

ABSTRACT **The human malaria parasite *Plasmodium falciparum* exhibits a high degree of chromosomal polymorphism, which may contribute to its ability to evade host defenses. The analysis of parasite chromosomes has revealed that these polymorphisms are confined to the subtelomeric regions, which are transcriptionally silent and contain repetitive sequence elements. Several subtelomeric repetitive elements have been isolated and mapped by using *P. falciparum* yeast artificial chromosome (YAC) clones. Structural analysis of parasite telomeric and subtelomeric YAC clones demonstrated that these repetitive elements are conserved between *P. falciparum* chromosome ends. We suggest that these subtelomeric elements promote chromosome pairing in *P. falciparum* and facilitate meiotic recombination and gene conversion between telomere-proximal genes.**

The malaria-causing organisms of the genus *Plasmodium* are obligate intracellular protozoan parasites that require two hosts, a mosquito and a vertebrate, for the completion of their life cycle. In humans, the most severe form of malaria is caused by *Plasmodium falciparum*. While in the human host, the parasite is haploid and multiplies asexually within erythrocytes. In the mosquito host, a diploid zygote is formed, in which meiotic recombination occurs. Substantial genetic diversity exists between different parasite isolates. Both antigen genes and individual parasite chromosomes display striking polymorphism. Large size polymorphisms, ranging from 50 to over 300 kb, occur between homologous chromosomes of different geographical isolates and are observed both in culture and in field isolates (1–3). Chromosomal polymorphism can arise either during mitotic growth or through meiotic recombination (4). The mechanisms responsible for generating the mitotically arising polymorphic chromosomes and the role these polymorphisms play in parasite biology are not well understood.

The analysis of the strain-dependent chromosomal polymorphism associated with the *P. falciparum* chromosomes 1, 2, and 8 reveals that polymorphisms are restricted to chromosome ends (5–7). Many polymorphisms result from large subtelomeric deletions and are a consequence of chromosome breakage and telomere-healing events (6–8). However, polymorphisms resulting from the possible insertion or amplification of DNA sequences within subtelomeric regions have also been observed (5).

Previously, it was shown that wild-type chromosome ends contain a conserved *Apa* I restriction endonuclease site located ≈10–15 kb from the telomere (9). This terminal *Apa* I fragment, designated the terminal complex, includes the telomere repeats (9). In addition to the *Apa* I site, several other restriction enzyme sites appear conserved within subtelomeric regions, suggesting an even higher degree of conservation (10). Approximately 10 kb from the chromosome end, telomere-distal to the *Apa* I site, a repetitive DNA sequence present on all wild-type chromosome ends, Pfrep20, has been described (11).

Transcription unit mapping of *P. falciparum* yeast artificial chromosome (YAC) clones corresponding to the entire chromosome 2 suggests that, unlike the transcriptionally active central domains, subtelomeric regions are devoid of intraerythrocytic stage genes (5). *P. falciparum* chromosomes seem to be compartmentalized into transcribed, conserved internal domains and transcriptionally silent, polymorphic ends. Therefore, subtelomeric sequences may have a different structural organization than the central domain of a chromosome.

This study provides a structural analysis of the polymorphic *P. falciparum* subtelomeric regions. These regions were found to contain arrays of repetitive sequence elements, which are common to wild-type chromosomes. This suggests that these regions may be sites of preferential chromosome pairing and formation of synaptonemal complexes and chiasmata during recombination. This preference would help promote crossing-over and gene conversion between subtelomerically localized antigen loci and thus encourage the emergence of genetically diverse recombinant parasites with novel antigen complements. Preliminary analysis of *P. falciparum* genetic crosses supports this hypothesis (12).

MATERIALS AND METHODS

Parasite Culture and DNA Preparation. The *in vitro* culture of *P. falciparum* and the preparation of parasite DNA have previously been described (13, 14).

Analysis of YAC Clones. The cloning and analysis of *P. falciparum* (FCR 3 strain) subtelomeric and telomeric YAC clones have been described (5, 14).

Hybridization Probes. Probes were radiolabeled by random priming as described (ref. 15, p. 10.13). R-FA3 and R-CG7 were previously designated as STS 1 and STS 31 (5). R-1A2 is the 5.5-kb *Sac* I/*Sac* II fragment recovered from the end rescue of the parasite telomere YAC clone 1A2 (14). PfTEL and REP 20 were subcloned from the *P. falciparum* chromosome 2-specific telomere YAC clone GC6 (5). Briefly, genomic yeast DNA containing the GC6 YAC was treated with BAL-31 exonuclease, and recessed ends were repaired by T4 DNA polymerase (ref. 15, p. F.4). Treated DNA was digested with *Rsa* I and inserted into *Hin*cII-linearized pUC19 plasmid (ref. 15, p. 1.70). A plasmid, PfTEL, containing *P. falciparum* telomere repeat sequences was recovered. In addition, the GC6 subclone library was screened with the Pfrep 20 probe of Patarapotikul and Langsley (11). A REP 20-containing plasmid was isolated.†

Abbreviations: PFG, pulsed-field gel; STS, sequence tag site; YAC, yeast artificial chromosome.
*To whom reprint requests should be addressed.
†The sequences discussed in this paper have been deposited in the GenBank data base (accession nos. L22442–L22446 are, in order, for R-CG7, R-1A2, REP 20, PfTEL, and R-FA3).

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.

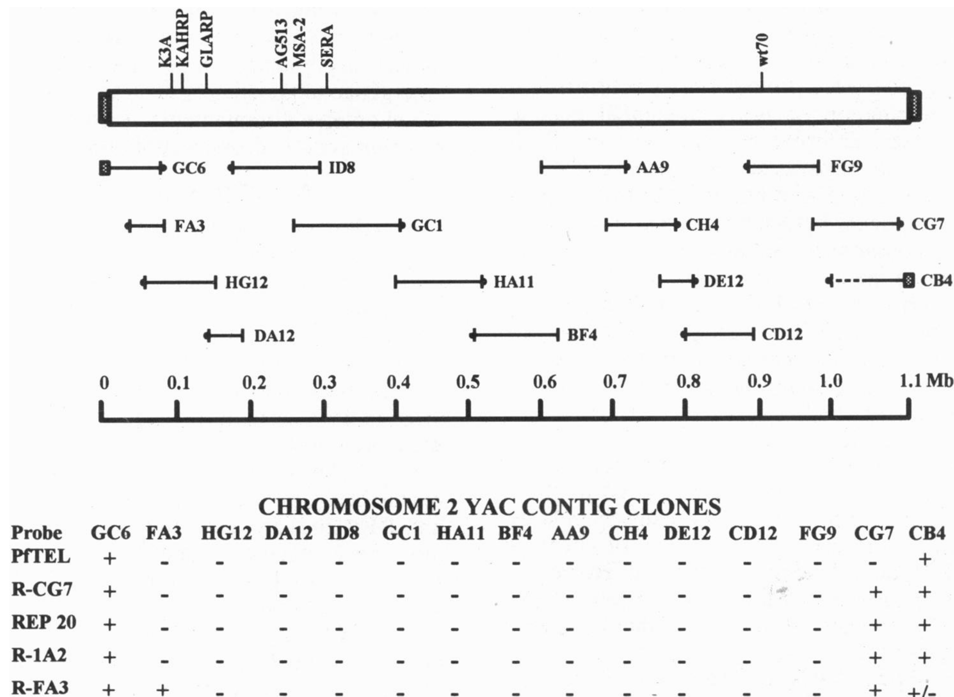
RESULTS

Subsets of Repetitive Elements Are Found in the Subtelomeric Regions of *P. falciparum* Chromosomes. We have described the construction of a stable representative *P. falciparum* YAC library and the cloning of parasite telomeric regions as YACs by complementation in yeast (14). Using these clones, we recovered the 1.1-Mb chromosome 2 as a telomere-to-telomere YAC contig (5). This contig revealed that the conserved internal regions of the chromosome were devoid of major repetitive DNA sequences (5). Of the 31 sequence tag site (STS) probes recovered across the chromosome, only the two final ones, STS 1 (the end-rescued left arm of the FA3 YAC) and STS 31 (the end-rescued left arm of the CG7 YAC), located ≈50 and 10 kb from the telomere, respectively, were repetitive (5) (Fig. 1). Because STS 1 and STS 31 do not refer to single-copy genetic loci, they are renamed here and used in the following studies under the designation STS 1 = R-FA3 and STS 31 = R-CG7. In addition to R-FA3 and R-CG7, another repetitive DNA element, R-1A2, was isolated. R-1A2 was recovered after the end rescue of DNA sequences adjacent to the *Nhe* I cloning site from a parasite telomere YAC clone as previously described (14).

The repetitive DNA elements are unrelated to each other as revealed by DNA sequence and Southern analysis (data not shown). PfTel contains ≈800 bp of the malaria parasite telomere sequence GGGTT(T/C)A (16). The REP 20 clone shows the 21-bp degenerate tandem repeat expected in a Pfrep 20 hybridizing sequence (11). Analysis of the R-FA3 clone revealed that 234 of the first 240 nucleotides were identical to the *P. falciparum* a95 repetitive DNA sequence, a random clone isolated from the Malaysian CAMP strain (17). The R-CG7 and R-1A2 DNA sequences are unrelated to previously described malaria sequences, and they do not show either complex or simple sequence repeat patterns (data not shown).

To determine the localization and distribution of the repetitive elements on a chromosome, members of an ordered series of the chromosome 2 YAC contig clones were separated by PFG electrophoresis and then transferred to nitrocellulose filters. Radiolabeled PfTEL, R-CG7, R-FA3, R-1A2, and REP 20 probes were sequentially incubated with the filters. The patterns of hybridization across chromosome 2 are summarized in Fig. 1. No hybridization to the 800-kb internal conserved domain of the chromosome is seen, and each repetitive probe hybridizes only to the YAC clones containing the polymorphic 100-kb subtelomeric regions (Fig. 1).

The sequence elements R-CG7, R-FA3, R-1A2, and REP 20 are not chromosome specific. Instead, they hybridize to chromosomes of all sizes as determined by Southern analysis of PFG-separated chromosomes (Fig. 2). However, the intensity of hybridization for an individual probe can vary between chromosomes of the same strain and between chromosomes of unrelated strains (Fig. 2, R-CG7 probe, compare strains D10 and FVO⁻). The variation in hybridization intensity could result from differences either in the DNA element copy number or in the genetic drifting of the DNA sequences. In addition, some parasite chromosomes, such as chromosome 1 of strain D10 (Fig. 2, leftmost, ethidium bromide-stained, gel, D10 lane, first band from the bottom of gel), fail to hybridize to any of the sequences characterized.



CHROMOSOME 2 YAC CONTIG CLONES

Probe	GC6	FA3	HG12	DA12	ID8	GC1	HA11	BF4	AA9	CH4	DE12	CD12	FG9	CG7	CB4
PfTEL	+	-	-	-	-	-	-	-	-	-	-	-	-	-	+
R-CG7	+	-	-	-	-	-	-	-	-	-	-	-	-	+	+
REP 20	+	-	-	-	-	-	-	-	-	-	-	-	-	+	+
R-1A2	+	-	-	-	-	-	-	-	-	-	-	-	-	+	+
R-FA3	+	+	-	-	-	-	-	-	-	-	-	-	-	+	+/-

FIG. 1. Repetitive DNA sequences are localized to the subtelomeric regions of *P. falciparum* chromosome 2. The minimal number of YAC clones representing *P. falciparum* chromosome 2 were separated by pulsed-field gel (PFG) electrophoresis, transferred to a nitrocellulose filter, and then analyzed by sequential hybridizations with the indicated DNA probes (PfTEL, R-CG7, REP 20, R-1A2, and R-FA3). The locations of chromosome 2-specific genes (*KAHRP*, knob-associated histidine-rich protein; *MSA-2*, merozoite surface antigen 2; *SERA*, serine-repeat antigen) and one anonymous STS marker (λwt70) are indicated (5). The *K3A* and *GLARP* genes will be described elsewhere. −, No hybridization; +, strong hybridization; +/−, weak hybridization. The left (yeast centromere-containing) arm of the YAC clone is indicated by the filled circle. The differences in the intensity of hybridization of the R-FA3 probe between the CB4 and CG7 YAC clones were unpredicted, since preliminary gross restriction enzyme mapping data indicated that the clones were 90% overlapping (5). End-rescued probes from these YAC clones hybridize back to chromosome 2; however, the CB4 left end probe (indicated by the broken line) overlaps with the chromosome 2-specific AA9 YAC clone. Further analysis revealed that the CB4 YAC most likely represents a distinct subpopulation within FCR3 parasites with a unique chromosome 2 that are present in the laboratory cultures (data not shown).

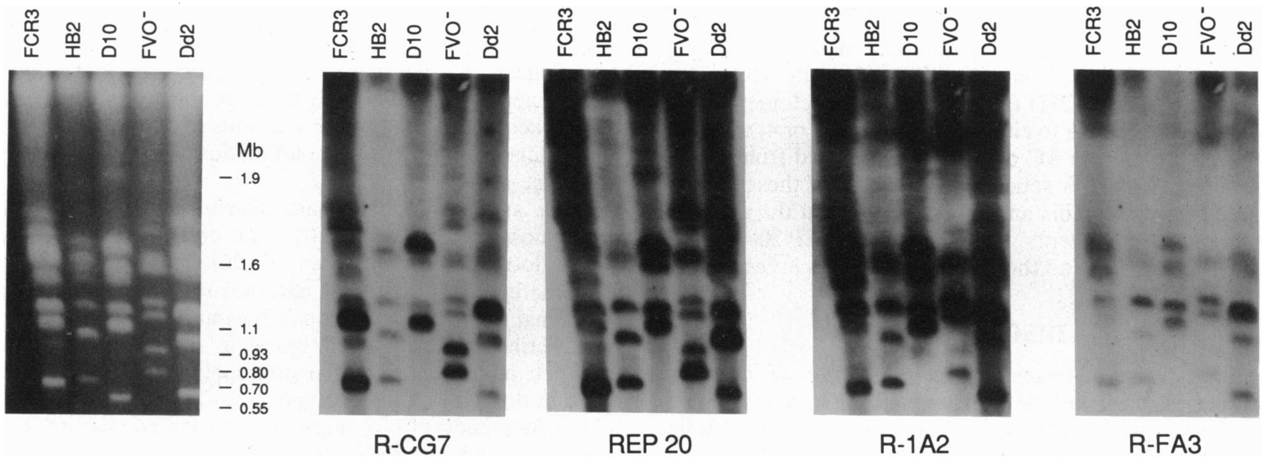


FIG. 2. Southern analysis of PFG-separated parasite chromosomes identifies repetitive DNA elements. Parasite chromosomal DNA embedded in agarose blocks was size fractionated by PFG electrophoresis [Bio-Rad CHEF-DR II, 0.8% SeaKem LE agarose (FMC) in 45 mM Tris borate, pH 8.3/1 mM EDTA, 14°C, pulse time from 60 to 300 sec over 48 hr at 180 V]. The ethidium bromide-stained gel is shown on the left. The DNA was capillary transferred to a nitrocellulose filter, and the filter was hybridized, under stringent conditions, sequentially, with the indicated radiolabeled DNA probes (R-CG7, REP 20, R-1A2, and R-FA3). The geographical origins of the parasites are as follows: FCR3, Gambia; HB2, Honduras; D10, New Guinea; FVO⁻, Vietnam; Dd2, Indochina. Size markers are in megabases.

Cross-species homologies to *P. falciparum* repetitive DNA were examined by Southern analysis. All probes failed to hybridize to DNA from humans, mice, *Giardia lamblia, Trypanosoma brucei, Crithidia fasciculata, Candida albicans, Schizosaccharomyces pombe, Plasmodium berghei*, or *Plasmodium yoelii*. However, R-CG7 showed faint hybridization to *Saccharomyces cerevisiae* chromosomes, and the *P. falciparum* telomere clone (PfTEL) hybridizes to the chromosomes of both *P. berghei* and *P. yoelii* (data not shown).

The Organization of *P. falciparum* Subtelomeric Regions Is Conserved. Restriction enzyme mapping and Southern anal-

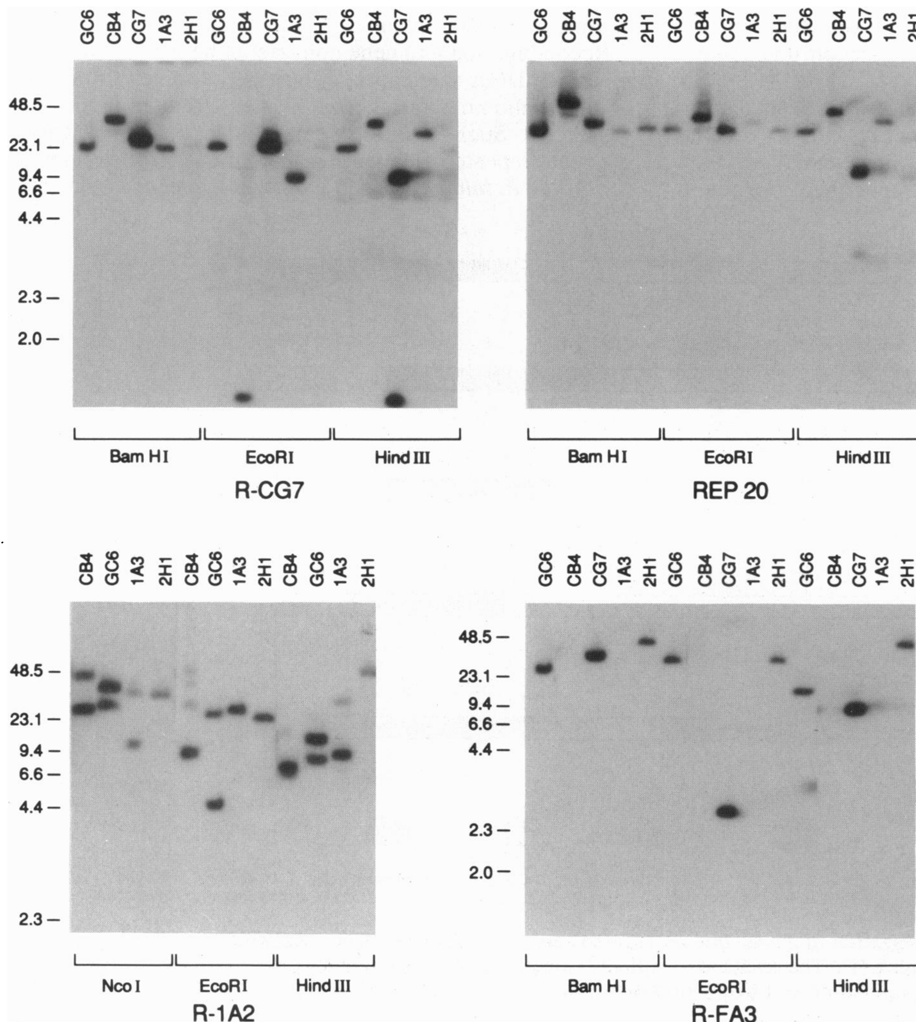


FIG. 3. Mapping of repetitive DNA elements on *Plasmodium* telomeric YAC clones. Agarose plugs containing chromosomal DNA from the GC6 (chromosome 2 "left arm"), CG7 (chromosome 2 "right arm"), CB4 (chromosome 2 "right arm"), 1A3 (chromosome 3), and 2H1 (chromosome 9) plasmodial telomere YAC clones were prepared and digested with the indicated restriction enzymes. The digested YAC clone DNA was size fractionated by PFG electrophoresis and then analyzed for hybridization to radiolabeled R-CG7, REP 20, R-1A2, and R-FA3. A DNA size standard is indicated in kb. Hybridization of R-FA3 to the telomere YAC clone 1A3 is evident only after prolonged exposure. The presence of the R-FA3 element on this YAC clone has been independently confirmed by PCR analysis (data not shown). Further hybridization analysis was performed with PfTel and YAC4 vector left and right arm probes (data not shown). These data are summarized in Fig. 4.

ysis were carried out to localize the repetitive DNA elements on the chromosome 2 telomere and subtelomere YAC clones, GC6, CB4, and CG7, and on two additional, random telomere YAC clones, 1A3 and 2H1 (Fig. 3). The 60-kb telomere YAC clone 1A3 corresponds to either chromosome 3 or 4, whereas the 100-kb telomere YAC clone 2H1 is derived from either chromosome 8 or 9. A schematic summary of these data is presented in Fig. 4. This analysis revealed that the order of the subtelomeric elements, PfTEL–R-CG7–REP 20–R-1A2–R-FA3, is common to all the chromosome ends investigated.

DISCUSSION

Previous data have suggested that *P. falciparum* chromosomes are functionally compartmentalized into conserved transcribed central domains and polymorphic transcriptionally silent subtelomeric ends (5) (Fig. 5). A preliminary analysis suggested that the subtelomeric regions may have a different structural organization than the conserved central regions (5). Here, we have examined the structural organization of the subtelomeric regions from *P. falciparum* chromosomes, using parasite telomere YAC clones from four different chromosome ends. Subtelomeric regions were found to contain repetitive sequence elements, extending over 60 kb from the telomere end. In addition to the previously characterized telomere (PfTEL) and REP 20 sequences, three additional repetitive elements, R-CG7, R-FA3, and R-1A2, were isolated from parasite telomere YAC clones. The order of these repetitive elements within a subtelomeric region, telomere–PfTEL–R-CG7–REP 20–R-1A2–R-FA3–centromere, was found to be common to several chromosome ends. We postulate that this organization is likely to be common to all parasite wild-type chromosome ends (Fig. 5).

In other organisms, subtelomeric regions are also organized as an array of repetitive elements (18–20). In yeast, subtelomeric regions function as origins of replication and as buffers between the telomeres and transcribed regions, thereby preventing the silencing of transcriptional activity (21–24). While the effect of *P. falciparum* subtelomeric sequences on either chromosome maintenance or transcriptional activity remains unclear, we propose a model whereby these sequences are preferential sites of chromosome pairing and thus help promote recombination and gene conversion events.

The analysis of chiasmata distribution between human chromosomes has shown that the chiasmata are preferentially located near telomeres (25, 26). This suggests that the telomeres are the sites of chromosome pairing and synaptonemal complex formation. Chiasma formation in the untranscribed repetitive subtelomeric regions could lead to branch migrations into the nonrepetitive, transcriptionally active domains of the chromosomes. This could result in a high frequency of crossing-over or gene conversion events near chromosome ends. Of the antigen genes mapped at low resolution on parasite chromosomes a majority appear to cluster within 100–200 kb of the telomeres (5, 27–29). Thus, if recombination and gene conversion are preferentially initiated within subtelomeric regions, this may enhance the emergence of parasites with new antigenic combinations.

The limited existing genetic data on *P. falciparum* chromosomal recombinants suggest that there is indeed some bias towards either crossing-over or gene conversion events involving chromosome ends (12, 29). When two different parasite strains were crossed, Vernick *et al.* (12) observed that the cross-fertilized progeny displayed nonparental restriction fragment patterns in their subtelomeric DNA, a likely result of either gene conversion or recombination events. However, sequences internal to these telomeric regions remained stable.

Recombination and gene conversion between short (2-kb) repetitive DNA sequences, either on the same chromosome or on nonhomologous chromosomes, frequently occur in yeast (30). Such ectopic events can lead to homogeneity between repeated sequences, and they may explain the conserved *P. falciparum* subtelomeric sequence organization

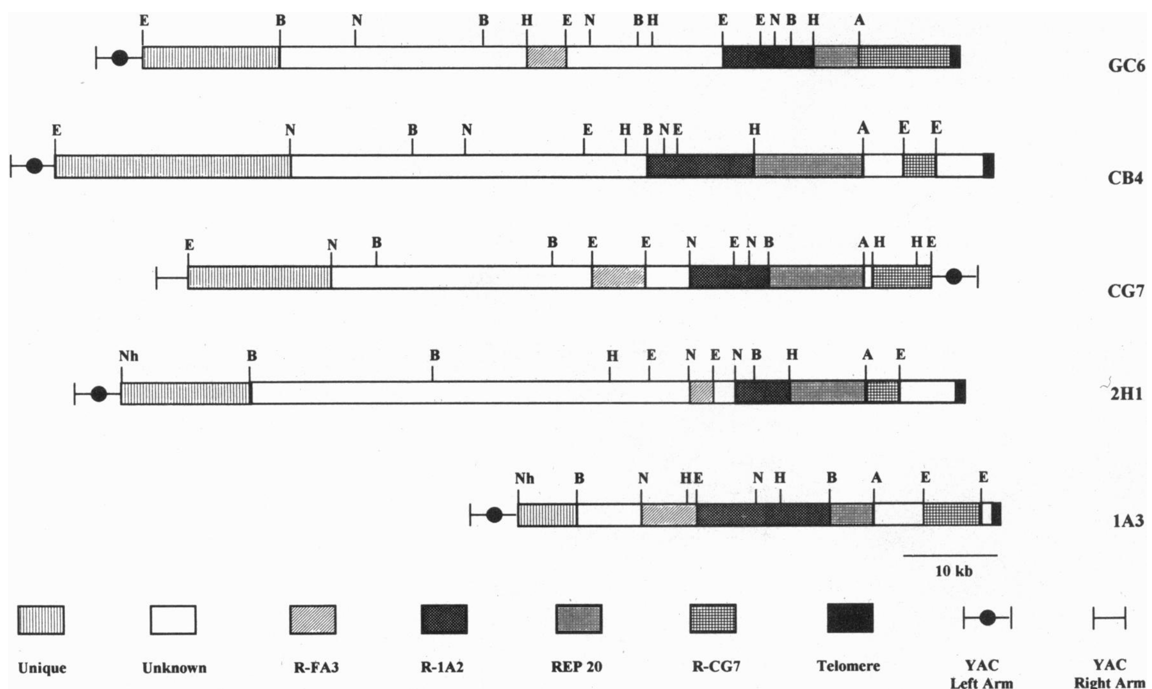


FIG. 4. Schematic summary of the organization of *P. falciparum* chromosome ends. The restriction enzyme maps are aligned at the conserved *Apa I* site of the telomere complex (1). The smallest identifiable repetitive sequence domains are illustrated. CG7 and CB4 represent YACs from major and minor subpopulations of FCR3 chromosome 2, respectively (Fig. 1). Nh, *Nhe* I; B, *Bam*HI; A, *Apa* I; E, *Eco*RI; H, *Hin*dIII; N, *Nco* I.

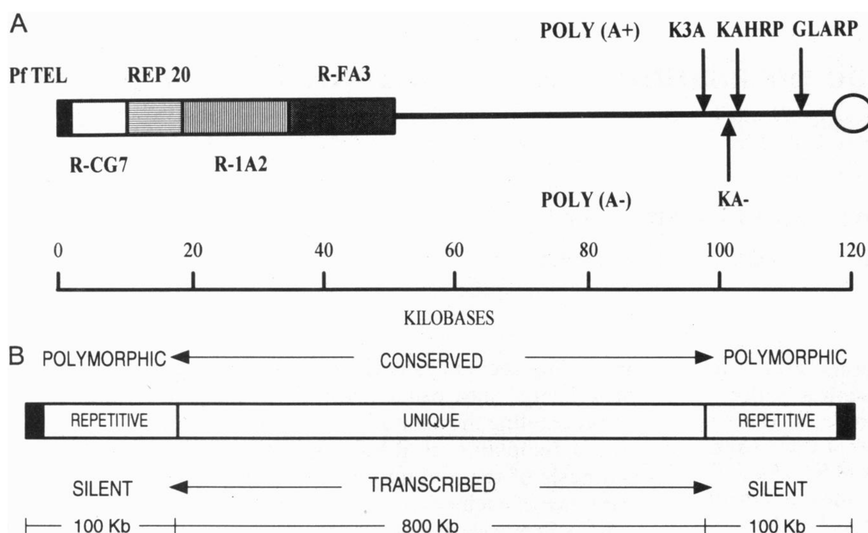


FIG. 5. (A) Map of chromosome 2 subtelomeric region. The left end of chromosome 2 contains ≈60–80 kb of repetitive DNA sequences (R-CG7, REP 20, R-1A2, and R-FA3). These sequences display a conserved order between all chromosome ends examined. Transcription units flanking the KAHRP locus [the poly(A)$^+$ transcribed *K3A* and *GLARP* and the poly(A)$^-$ transcribed *KA*$^-$] are indicated (5). The approximate size range of the repetitive domains is illustrated. (B) Organization of *P. falciparum* chromosome 2. The chromosomal polymorphisms are restricted to within 100 kb at either end of the chromosome. The polymorphic ends are composed of conserved arrays of repetitive DNA sequences that are not transcribed during intraerythrocytic stages of the life cycle. In contrast, the 800-kb central domain of the chromosome contains transcribed unique-sequence DNA that is conserved between parasite isolates. This model is applicable to other *P. falciparum* chromosomes.

(31). The recombination between dispersed repeats can also generate such chromosomal aberrations as translocations, inversions, deletions, and duplications (30). Some of the additional *P. falciparum* chromosomal polymorphism not associated with chromosome breakage and healing type deletions can be explained by intrachromosomal or heterochromosomal crossing-over and gene conversion events between subtelomeric sequences (5, 9). Ectopic recombination events resulting from pairing of chromosomes through subtelomeric repetitive sequences may explain the emergence of several *P. falciparum* gene families such as *HRP II* and *HRP III*, *RESA-1* and *RESA-2*, and *GBP130* and *GBPH* (27, 32, 33). Some of these genes, such as *HRP-II* and *HRP-III*, located on chromosomes 8 and 13, respectively, are found near chromosome ends and are thought to have arisen from the duplication and translocation of an ancestral gene (27). Hence, preferential chromosome pairing between dispersed subtelomeric repeats could provide the parasite with an effective means of generating novel phenotypes by promoting recombination and gene conversion events. Such events would subsequently lead to increased diversity amongst a subset of genes located near chromosome ends; however, those genes in the conserved internal domains of the chromosomes would be more stable (12).

We thank Arthur Lustig and Sam Wertheimer for their helpful discussions and comments. Mark Samuels and J. Carter Ralphe provided expert technical assistance. Secretarial assistance was provided by Cynthia Ritter. This work was supported by U.S. Army Grant DAMD 17-89-Z-9003 and the World Health Organization. D.d.B. is supported by the Cornell University Graduate School of Medical Sciences, and M.L. is supported by fellowships from the Deutscher Akademischer Austauschdienst and the World Health Organization.

1. Foote, S. J. & Kemp, D. J. (1989) *Trends Genet.* **5**, 337–342.
2. Corcoran, L. M., Forsyth, K. P., Bianco, A. E., Brown, G. V. & Kemp, D. J. (1986) *Cell* **44**, 87–95.
3. Biggs, B. A., Kemp, D. J. & Brown, G. V. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 2428–2432.
4. Janse, C. J. (1993) *Parisitol. Today* **9**, 19–22.
5. Lanzer, M., de Bruin, D. & Ravetch, J. V. (1993) *Nature (London)* **361**, 654–657.
6. Pologe, L. G., de Bruin, D. & Ravetch, J. V. (1990) *Mol. Cell. Biol.* **10**, 3243–3246.
7. Pologe, L. G. & Ravetch, J. V. (1988) *Cell* **55**, 869–874.
8. Pologe, L. G. & Ravetch, J. V. (1986) *Nature (London)* **322**, 474–477.
9. Corcoran, L. M., Thompson, J. K., Walliker, D. & Kemp, D. J. (1988) *Cell* **54**, 807–813.
10. Scotti, R., Pace, T. & Ponzi, M. (1993) *Mol. Biochem. Parasitol.* **58**, 1–6.
11. Patarapotikul, J. & Langsley, G. (1988) *Nucleic Acids Res.* **16**, 4331–4340.
12. Vernick, K. D., Walliker, D. & McCutchan, T. F. (1988) *Nucleic Acids Res.* **16**, 6973–6985.
13. Trager, W. & Jensen, J. B. (1976) *Science* **193**, 673–675.
14. de Bruin, D., Lanzer, M. & Ravetch, J. V. (1992) *Genomics* **14**, 332–339.
15. Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab. Press, Plainview, NY), 2nd Ed.
16. Vernick, K. D. & McCutchan, T. F. (1988) *Mol. Biochem. Parasitol.* **28**, 85–94.
17. Weber, J. L. (1988) *Mol. Biochem. Parasitol.* **29**, 117–124.
18. Biessmann, H. & Mason, J. M. (1992) *Adv. Genet.* **30**, 185–249.
19. Brown, W. R. A., MacKinnon, P. J., Villasante, A., Spurr, N., Buckle, V. J. & Dobson, M. J. (1990) *Cell* **63**, 119–132.
20. Zakian, V. A. & Blanton, H. M. (1988) *Mol. Cell. Biol.* **8**, 2257–2260.
21. Gottschling, D. E., Aparico, O. M., Billington, B. L. & Zakian, V. A. (1990) *Cell* **63**, 751–762.
22. Weiner, A. M., Deininger, P. L. & Efstratiadis, A. (1986) *Annu. Rev. Biochem.* **55**, 631–661.
23. Carlson, M., Celenza, J. L. & Eng, F. J. (1985) *Mol. Cell. Biol.* **5**, 3584–3587.
24. Horwitz, H. & Haber, J. E. (1985) *Mol. Cell. Biol.* **5**, 2369–2380.
25. Laurie, D. A. & Hulten, M. A. (1985) *Ann. Hum. Genet.* **49**, 189–205.
26. Zickler, D., Moreau, P. J. F., Huynh, A. D. & Slezec, A.-M. (1992) *Genetics* **132**, 135–148.
27. Wellems, T. E. & Howard, R. J. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 6065–6069.
28. Walker-Jonah, A., Dolan, S. A., Gwadz, R. W., Panton, L. J. & Wellems, T. E. (1992) *Mol. Biochem. Parasitol.* **51**, 313–320.
29. Scherf, A., Carter, R., Petersen, C., Alano, P., Nelson, R., Aikawa, M., Mattei, D., Pereiar da Silva, L. & Leech, J. (1992) *EMBO J.* **11**, 2293–2301.
30. Petes, T. D. & Hill, C. W. (1988) *Annu. Rev. Genet.* **22**, 147–168.
31. Louis, E. J. & Haber, J. E. (1992) *Genetics* **133**, 559–574.
32. Cappai, R., Kaslow, D. C., Peterson, M. G., Cowman, A. F., Anders, R. F. & Kemp, D. J. (1992) *Mol. Biochem. Parasitol.* **54**, 213–222.
33. Nolte, D., Hundt, E., Langsley, G. & Knapp, B. (1991) *Mol. Biochem. Parasitol.* **49**, 253–264.