Volume 16 Number 10 1988 Nucleic Acids Research

Chromosome size polymorphism in *Plasmodium falciparum* can involve deletions of the subtelomeric pPFrep20 sequence

Jintana Patarapotikul and Gordon Langsley+

Unit of Experimental Parasitology, Institut Pasteur, 25 rue du Dr Roux, 75724 Paris, France

Received February 26, 1988; Revised and Accepted April 15, 1988

ABSTRACT

The P.falciparum pPFrep20 repetitive element from the Palo Alto Uganda strain has been isolated and sequenced. The Palo Alto pPFrep20 repeat (pPFPArep20) has a clustered subtelomeric location and on chromosome 1 has been deleted from one end. Analysis of chromosome 1 from 5 other strains has revealed that pPFrep20 sequences have been deleted from one end in 3 of them. Thus, deletion of pPFrep20 appears to be a frequent event that could significantly contribute to chromosome size polymorphism in P.falciparum.

INTRODUCTION

One of the first DNA sequences isolated from the Plasmodium falciparum genome was a highly reiterated element termed pPFrep20 (1.2). Other repetitive sequences resembling pPFrep20 have since been isolated from different P.falciparum strains (3-9). The apparent variability, amount and species specifity of pPFrep20 has resulted in this sequence being proposed as a means of strain typing and as a diagnostic probe for P.falciparum malaria (1-10).

The usefulness of pPFrep20 in strain typing is based on the observation that hybridisation of restricted genomic DNA identifies many fragments, the number and size of which are a characteristic of each strain (1,10). However, only a few large fragments were identified when enzymes EcoR1 and DraI were used, indicating that pPFrep20 sequences are clustered in the genome (10,11). DNA sequence analysis revealed that pPFrep20 is composed of a 21 base pair (bp.) repeat (10).

A similar 21bp. sequence has also been described for other P.falciparum strains (4,9). The clustered nature, size and sequence variation of the pPFrep20 sequence are typical of

satelite 1 DNA and it has been proposed that pPFrep20 may be associated with heterochromatin and telomers (10).
The pPFrep20 sequences can indeed be found on all P.falciparum chromosomes and the differing degree of hybridisation has lead to the supposition that the loss of pPFrep20 sequences may be one of the causes of chromosome size variation in P.falciparum (12,13).
Here, we describe the isolation and sequence of the P.falciparum Palo Alto (FUP strain) pPFrep20 repeat element. We show furthermore, that pPFrep20 sequences are highly concentrated in a subtelomeric location and we demonstrate that chromosome size polymorphism in P.falciparum can be due to deletion of the subtelomeric pPFrep20 sequences.

MATERIALS AND METHODS

Isolation of pPFPArep20 and derivation of its DNA sequence.

The pPFPArep20 clone was isolated from an HindIII/EcoRI genomic library constructed in pUC9. A detailed description of the construction of this bank and the isolation of specific clones has been given in Langsley and Ponnudurai (1988). pPFPArep20 was identified in the same manner as the original pPFrep20 clone, namely by screening the library with radiolabeled total genomic DNA (1).The pPFPArep20 sequence was obtained from analysis of double stranded recombinant pUC9 DNA using standard techniques (14).

Restriction analysis of purified P.falciparum chromosomes

Chromosomal blocks were prepared as described in (12) and the pulse field gels as in (13). Briefly, after separation of chromosome 1 (275volts/75sec pulse/0.5XTBE/20hrs.) the gel was stained with ethidium bromide and the chromosome exised from the gel. The agarose blocks containing the individual chromosomes were then equilibriated over night in the relevent restriction buffer. Prior to digestion, the restriction buffer was changed and acetylated bovine serum albumin (BSA) added to a final concentration of 0.1mg./ml. The block was then digested over night at the correct temperature for the corresponding enzyme. The digested chromosomal DNA was reseparated together with phage lamda concatomers and yeast chromosomes as size

markers. A size range of up to 530kb. was chosen using the same conditions as above, but with a pulse of 30 seconds.

Transfer of chromosomal gels and hybridisation with various probes

The gels were alkaline transferred to nylon (Zeta-probe) following the Bio-Rad (32nd. & Griffin Ave, Richmond, Calif.94804) protocol with the addition, after the completion of transfer, of a 30 minute neutralisation step (0.2MTris/2XSSC). The filters were hybridised (6X SSC/5% Dextran Sulphate/5X Denhardts/0.1%SDS/50ug./ml. salmon sperm DNA) at 65°C with the following probes.

pPFPArep20- described in the present communication.

Telomere- a kind gift of Clara Frontali and described in (15).

rRNA genes- described in (12,16).

RESA- a recombinant expression clone containing just the 3'-repeats (Jurgen Kuhn, pers.comm. and 17).

The filters were washed in 1XSSC at 65°C and after exposure to film the fiters were dehybridised by boiling 2X 30 minutes in 0.1XSSC/0.5%SDS., re-exposured to confirm the dehybridisation and then reprobed.

RESULTS

The P.falciparum Palo Alto (FUP Uganda) pPFrep20 element

In the course of our studies on chromosome size polymorphism a number of chromosome specific probes were isolated (13). Some of the probes hybridised to multiple fragments on Southern blots and to all Palo Alto chromosomes (data not shown). One of these is described here and for consistency is called pPFPArep20 for the P.falciparum Palo Alto rep20 sequence. It contained a 1.6kb. HindIII/EcoR1 fragment which encodes many examples of a 21bp. repeat. The sequence of 7 repeats is shown in Fig.1 and the Palo Alto consensus is compared to the consensus sequences derived for the 3 other P.falciparum strains (1,4,9).The overall structure of the 21bp. repeat is

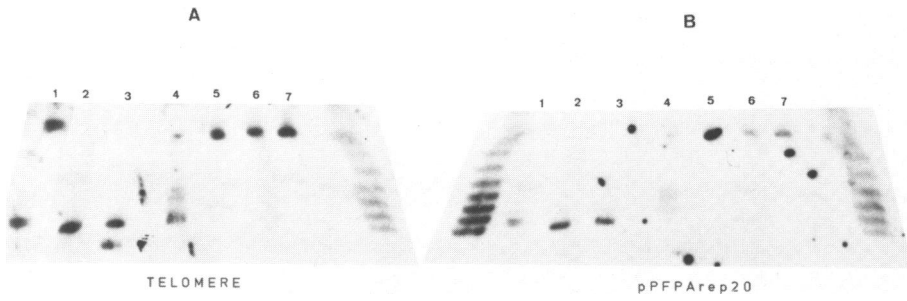
pPFPArep20 Sequence	Consensus Sequence	Isolate Origin & Ref
ACTTATATAAGGACCTATATT	G G C G ACTTATATTAGGACCTATATT T	Uganda (This paper)
ACTTATATTAGGACCTATACT		
AGTTATGTTAGGACCTATAT-	CA TG ACTAATCTAGGTCTTAAGATT	Gambia (10)
ACGTATCTTTGGACCTATATT		
AGTTTATTAAGGACCTATGTT		
	ACTAA--TAGGTCTTA---T-	Tanzania F-32 (4)
AGTACTTTGAAGACCTATATT		
AGTAGTTAGGACTACCTAATT	TT A A A ACTAACATAGGTCTTATTTTC GG C G G C C	Tanzania 1 (9)

Legend to Figure 1.
The DNA sequence of the repetitive element of P.falciparum Palo Alto, FUP Uganda strain (pPFPArep20) and its comparison with pPFrep20 sequences from other strains. The sequence of 7 repeats of pPFPArep20 is given and they cleary identify it as being the Palo Alto rep20 element. Comparison with the previously described elements reveals considerable homology. Interestingly, the variable bases are clustered in 2 regions which are 9 bases apart.

not only conserved within, but also between strains.

The pPFPArep20 sequences are clustered and subtelomeric on chromosome 1

Chromosome 1 from Palo Alto was purified and digested with a number of different restriction enzymes (see Materials and Methods). The PFG separated restriction digests were first hybridised with the P.berghei telomere probe to identify the telomere end fragments (Fig.2a). Many of the enzymes used, with either 8 base pair or G+C rich recognition sites, did not cut chromosome 1 and shown as examples are KpnI and PstI (tracks 6 and 7). Others however, restricted the chromosome and telomere end fragments were identified. In particular, BssH2, PvuII and

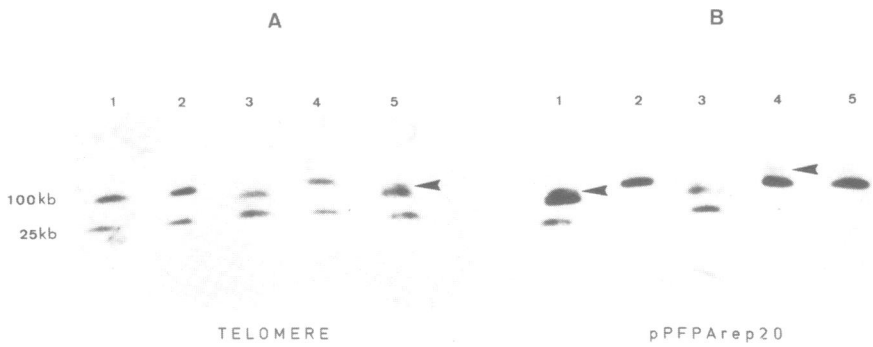


Legend to Figure 2.
The pPFPArep20 element is clustered and subtelomeric. Figure 2A shows those restriction fragments containing telomere sequences and since these are located at each end of the chromosome only 2 fragments are identified. Note that BssH2 (track 1) gives fragments of approximately 100kb. and 550kb. Pvull (track 2) yields 2 fragments of 75kb. which co-migrate. Xhol (track 3) gives fragments of 100 and 25kb. BamHl (track 4) gave only partial digestion and both Kpnl and Pstl (tracks 6 and 7) failed to cut. Track 5 shows uncut chromosome 1. The sizes were estimated by comparison with yeast and lambda markers. Figure 2B shows the location of pPFPArep20. Note that only the 100kb BssH2 (track 1) and Xhol (track 3) fragments have pPFPArep20 sequences. Since the two 75kb. Pvull fragments co-migrate it is impossible to differentiate which one contains pPFPArep20. The 25kb. Xhol fragment not only has lossed pPFPArep20 sequences, but it also hybridised less intensely with the telomere probe (Fig.2A, track 3) indicating that some teomere repeats have also been deleted.

Xhol proved usefull (Fig.2a, tracks 1,2 and3). The filter was then dehybridised and probed with pPFPArep20 (Fig.2b). The only fragments that hybridised to pPFPArep20 were those previously shown to be telomeric. This clearly demonstrates that pPFPArep20 is clustered and subtelomeric. To our surprise, only one BssH2 and Xhol Palo Alto telomeric fragment hybridised to pPFPArep20 indicating that pPFPArep20 sequences are unique to one end on chromosome 1.

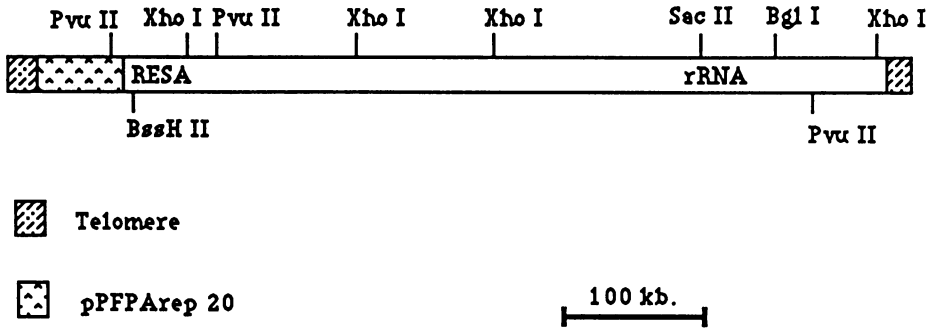
Chromosome 1 size polymorphism can involve deletions of the subtelomeric pPFPArep20 sequence

To test the possibility that pPFPArep20 sequences have been deleted from one end of chromosome 1 in the Palo Alto strain, several different P.falciparum clones were examined. These clones are derived from a single Thai isolate (J.Patarotikul,



Legend to Figure 3.
Location of pPFrep20 sequences on chromosome 1 of different P.falciparum clones. Figure 3A shows the XhoI telomeric fragments of chromosome 1 from 5 clones derived from a single isolate. Note that the size of the XhoI fragments varies, indicating that the ends of chromosome 1 are polymorphic and that in one clone (track 5) 3 fragments are identified. Figure 3B shows the same filter hybridised with pPFPArep20. The pPFrep20 sequences are present in a subtelomeric location at both ends of chromosome 1 in 2 clones (tracks 1 and 3), but are missing from the smaller XhoI fragment in the others (tracks 2,4 and 5). The varying intensities seen with the pPFPArep20 probe imply that the size polymorphism observed for the XhoI fragments is due to loss of pPFrep20 sequences. This is particularly obvious in panel B track 3.

C.Gentil and P.Druihle, in preparation). The enzyme XhoI was used, because in Palo Alto it yielded 2 clearly distinguishable telomeric fragments. As can be seen in Fig.3, the telomere probe identified 2 XhoI fragments in 4 of the clones, but a 3rd. faint telomere fragment in the 5th. (Fig.3A track 5, arrowed). In this clone the telomere repeats appear to extend over 100kb. to the adjacent XhoI fragment. When the same filter is probed with pPFPArep20 (Fig.3B) it can be seen that pPFrep20 sequences are located at both ends of chromosome 1 in 2 clones (tracks 1 and 3), but they are present at only one end of chromosome 1 in the 3 other clones (tracks 2,4 and 5). Thus, in 4 of the 6 strains examined pPFrep20 sequences have been deleted from one end of chromosome 1. Moreover, using the same filter allowed us to make an interesting observation, namely that although pPFrep20 sequences are subtelomeric they can be over 100kb. from the telomere. This is based on the following



Legend to Figure 4.
Restriction map of chromosome 1 of the P.falciparum Palo Alto FUP Uganda strain. The chromosome has a size of approximately 650kb. and shown together with the restriction sites mapped are the locations of the rRNA and RESA genes. Note that RESA is proximal to the telomere containing pPFPArep20. Note also that the pPFPArep20 sequences have been deleted from one end giving the rRNA genes a more telomeric location. Finally ethidium bromide staining (not shown) indicated that PvuII yields several fragments between 50 and 100kb. and only 3 of these PvuII sites have been mapped.

observations, firstly, in one clone the pPFrep20 sequences are located on the larger XhoI fragment that contains only a few telomeric repeats (Fig.3A, track 5, arrowed). Secondly, in two other clones the pPFPArep20 probe identified 2 approximately 100kb. fragments (Fig.3B, tracks 1 and 4, arrowed); the larger one in both cases being that one which hybridised to the telomere probe (compare A and B).

Physical map of P.falciparum Palo Alto chromosome 1

In the course of our analysis chromosome 1 of Palo Alto has been digested with 9 different restriction enzymes. Two enzymes, BamHI and SmaI only partially digested chromosome 1 and KpnI and PstI did not cut the chromosome at all. Comparison of the restriction patterns with the 5 other enzymes has enable us to construct a physical map for chromosome 1 (see Fig.4). Orientation of the map was facilitated by pPFPArep20 sequences being located exclusively at one end of the chromosome. As both the rRNA and RESA genes have been shown to be on this chromosome (12, 18), we decided to take advantage of our restriction analysis and position the two genes on the physical

map by probing the Southern blots used above with the cloned rRNA and RESA genes (data not shown). It can be noted that for the chromosome 1 of Palo Alto the RESA gene is located near the telomere which contains the pPFPArep20 repeats, whereas the rRNA genes are located at the other end.

DISCUSSION

The pPFPArep20 repeat given here is the fourth example of such a sequence described for P.falciparum. All the pPFrep20 sequences have been isloated from different strains. Surprisingly, for a non-expressed (10) highly repetitive sequence the overall structure of the repeat is conserved not only within, but also between strains. This argues for some structural role for the 21bp. repeat. A repeated element with a subtelomeric location has also been described for P.berghei (20) and it has been shown to contain a site of intrinsic bending (21). The 2.3kb. element contains within it a 27bp. repeat and this sequence, unlike pPFrep20, is very closely related to the G+C rich telomere repeats (20).

Since it has been proposed that pPFrep20 may be associated with centomers and telomeres (10), we decided to directly examine the chromosomal location of pPFPArep20 and to map the pPFPArep20 sequence to purified and digested chromosome 1 of the homologous Palo Alto strain.

The pPFPArep20 repeats are clustered and subtelomeric on chromosome 1. If pPFPArep20 sequences are also associated with centromers, then they must be so in only a few copies, as we were unable to detect by hybridisation their presence other than associated with the telomers. Surprisingly, in Palo Alto and 3 of the 5 other strains examined, the pPFrep20 repeats have been deleted from one end of chromosome 1. Thus, deletion of pPFrep20 sequences appears to occur at high frequency and this may be the major basis of chromosome 1 size polymorphism in P.falciparum.

Deletion of pPFrep20 sequences may also significantly contribute to size polymorphism for the other P.falciparum chromosomes. Preliminary analysis of chromosomes 2 and 3 of Palo Alto (not shown) has revealed that these 2 chromosomes

also have pPFPArep20 repeats only at one subtelomeric end. Deletion of pPFrep20 may then be a potential feature of all P.falciparum chromosomes and consistent with this is the earlier Southern blot analysis which showed that chromosomes hybridised to differing extents to pPFrep20 indicating that the amount of this sequence on each chromosome could be variable (12).

In the 3 of the 5 other P.falciparum clones studied the pPFrep20 sequences appear to be deleted from the XhoI fragment on which is located the RESA gene (not shown). Thus pPFrep20 sequences can be deleted from either end of chromosome 1. Recombination resulting in a large deletion at one end of chromosome 1 could result in the loss of the RESA gene, an event perhaps deleterious for the parasite and therefore, rarely seen. Three examples are known where antigen genes have been deleted from chromosomes and they all involve the subtelomerically located histidine rich protein genes and the conclusions drawn from these studies are that the histidine rich protein genes are not essential to the parasite (22,23 and 24). Thus, it follows that a subtelomeric location is a potential hot spot for recombination and antigen genes located there may be deleted following a recombination event involving pPFrep20 sequences. This process could contribute to the antigenic polymorphism of malaria parasites.

ACKNOWLEDGMENTS

We would like to thank Odile Mercereau-Puijalon for her encouragement and Luiz Pereira da Silva and Pierre Druihle for their support. Thanks to Artur Scherf and Jurgen Kuhn for the RESA gene and to Luisa Dandolo and Christine Petit for their advice and the lamda and yeast chromosomal blocks. Finally we are indebted to Hernando del Portillo for his advice on the double standed DNA sequencing. This work was supported by a grant no. M.R.E.S. 87W0043-Eureka, by W.H.O grant 840111 and a French Goverment Tropical Medecine fellowship to J.P.

+To whom correspondence should be addressed

REFERENCES

1. Goman,M., Langsley,G., Hyde,J.E., Yankovsky,N.K., Zolg,J.W.andScaife,J.G.(1982).Molec.Biochem. Parasit. 5 391-400.
2. Bone,N., Gibson,T., Goman,M., Hyde,J.E., Langsley,G.W., Scaife,J.G., Walliker,D., Yankofky,N.K. and Zolg,J.W. (1982). In: The Molecular Biology of Parasites (Luzzato,L. and Guardia,J. eds.) Raven Press, New York.
3. Franzen,L., Shabo,R., Perlmann,H., Wigzell,H., Westin,G., Aslund,L., Persson,T. and Pettersson,U. (1984). Lancet. 1 525-528.
4. Aslund,L., Franzen,L., Westin,G., Persson,T., Wigzell,H. and Pettersson,U. (1985). J.Mol.Biol. 185 509-516.
5. Bhasin,V.K., Clayton,C., Trager,T. and Cross,G. (1985). Molec. Biochem. Parasit. 15 149-158.
6. McLaughlin,G.L., Edlind;T.D., Campbell,G.H., Eller,R.F. and Ihler,G.M. (1985). Am.J.Med.Hyg. 34(5) 837-840.
7. Enea,V. (1986). Molec. and Cell.Biol. Jan. 321-324.
8. Barker,Jr.R.H., Suebsaeng,L., Rooney,W., Alecrim;G.C., Dourado,H.V. and Wirth,D.F. (1986). Science 231 1434-1436.
9. Zolg,W.J., Andrade,L.E. and Scott,E.D.(1987). Molec.Biochem.Parasit. 22 145-151.
10. Oquendo,P., Goman,M., Mackay,M., Langsley,G., Walliker,D. and Scaife,J.G. (1986). Molec.Biochem. Parasit. 18 89-101.
11. Langsley,G. (1983). Doctoral Thesis, Univ. of Edinburgh, Scotland.
12. Van der Ploeg,L.H.T., Smits,M., Ponnundurai,Th., Vermeulen,A., Meuwissen,J.H.E.T. and Langsley,G. (1985). Science 229 658-661.
13. Langsley,G. and Ponnundurai,TH. (1988). Experimental Parasit. (in press).
14. Chen,E.J. and Seeburg,P.H. (1985). DNA 4 165-170.
15. Ponzi,M., Pace,T., Dore,E. and Frontali,C. (1985). EMBO.J. 4 2991-2995.
16. Langsley,G., Hyde,J.E., Goman,M. and Scaife,J.G. (1983). Nucl. Acid. Res. 11 8703-8717.
17. Coppel,R.L., Cowman,A.F., Anders,R.F., Bianco,A.E., Saint,R.B., Lingelbach,K.R., Kemp,D.J. and Brown,G.V. (1984). Nature 310 789-791.
18. Kemp,D.J., Coppel,R.L., Stahl,H.D., Bianco,A.E., Corcoran,L.M., McIntyre,P., Langford,C.J., Falvalaro,J.M., Crewther,P.E., Brown,G.V., Mitchel,G.F. and Anders,R.F. (1986). Parasitology 91 S83-S108.
19. Brutlag,D.L. (1980). Ann. Rev. Genet. 14 121-144.
20. Pace,T. Ponzi,M., Dore,E. and Frontali,C. (1987). Molec. Biochem. Parasit. 24 193-202.
21. Dore,E., Pace,T., Ponzi,M., Scotti,R. and Frontali,C. (1988). Molec. Biochem. Parasit. 27 201-206.
22. Pologe,L. and Ravetch,J. (1986). Nature 322 474-477.
23. Wellems,T., Walliker,D., Smith,C.L., Rosario,V.E.do., Maloy,W., Howard,R.J., Carter,R. and McCutchan,T. (1987). 633-642.
24. Kemp,D.J., Thompson,J.K., Walliker,D. and Corcoran,L.M. (1987). Proc. Natl. Acad. Sci. USA. 84 7672-7676.