1942–1947 Nucleic Acids Research, 1995, Vol. 23, No. 11 © 1995 Oxford University Press

Characterisation of the telomeres at opposite ends of a 3 Mb *Theileria parva* chromosome

Baljinder K. Sohanpal, Subhash P. Morzaria, Elke I. Gobright and Richard P. Bishop*

International Livestock Research Institute, PO Box 30709, Nairobi, Kenya

Received February 13, 1995; Revised and Accepted April 20, 1995

GenBank accession nos L36963 and L36964

ABSTRACT

Bacteriophage λ clones containing *Theileria parva* genomic DNA derived from two different telomeres were isolated and the nucleotide sequences of the telomeric repeats and adjacent telomere-associated (TAS) DNA were determined. The *T.parva* telomeric repeat sequences, a tandem array of TTTTAGGG or TTTAGGG interspersed with a few variant copies, showed a high degree of sequence identity to those of the photosynthetic algae *Chlamydomonas reinhardtii* (97% identity) and *Chlorella vulgaris* (87.7% identity) and the angiosperm *Arabidopsis thaliana* (84.4% identity). Unlike most organisms which have been studied, no significant repetitive sequences were found in the nucleotide sequences of TAS DNA located centromere-proximal to the telomeric repeats. Restriction mapping and hybridisation analysis of λEMBL3 clones containing 16 kilobases of TAS DNA derived from one telomere suggested that they did not contain long regions of repetitive DNA. The cloned TAS DNAs were mapped to *T.parva* Muguga genomic *Sfi*I fragments 8 and 20, which are located at opposite ends of the largest *T.parva* chromosome. A 126 bp sequence located directly centromere-proximal to the telomeric repeats was 94% identical between the two cloned telomeres. The conserved 126 bp sequence was present on all *T.parva* Muguga telomeric *Sfi*I fragments.

INTRODUCTION

Telomeres are the specialised structures at the ends of eukaryotic chromosomes which ensure their stability and allow for the complete replication of their ends (1). Recent experimental data from *Saccharomyces cerevisiae* demonstrates that telomeres are essential for chromosome survival (2). With the exception of some insects (3), chromosomal termini of eukaryotes thus far studied consist of arrays of 6–8 bp tandem repeats which conform to a consensus sequence $5'[T/A_{1-4}G_{1-8}]3'$ (4,5). The terminal repeats are flanked internally by telomere-associated sequences (TASs) which are usually moderately repetitive, often found at more than one telomere and frequently variable in sequence and distribution between individuals within a species (reviewed in 4,5). In the malaria parasite *Plasmodium falciparum* the TASs contain long arrays of repetitive sequence elements with a conserved organization (6,7). Homologous recombination and other rearrangements in *Plasmodium* TASs are responsible for extensive chromosome size polymorphisms (8,9) and novel restriction fragments are generated in the TAS DNA in association with meiosis in *P.falciparum* (10).

Theileria parva is an intracellular protozoan parasite which is classified together with *Plasmodium* in the class Sporozoa. *T.parva* reversibly transforms bovine lymphocytes (reviewed in 11) and causes a usually fatal disease of cattle in sub-Saharan Africa. The availability of an *Sfi*I restriction map for *T.parva* DNA (12) constitutes a powerful tool for genome analysis in this organism.

In this report we describe the isolation of DNA from two *T.parva* telomeres which were located on the *Sfi*I restriction map at opposite ends of the largest *T.parva* chromosome. The *T.parva* telomeric repeat sequences exhibited a very high degree of identity to those of photosynthetic algae. By contrast with *Plasmodium* and most other eukaryotes, the TAS DNA at the cloned telomeres was not internally repetitious, but a short conserved sequence was present centromere-proximal to the telomeric repeats.

MATERIALS AND METHODS

Parasite material and DNA preparation

The *T.parva* stocks used were Uganda (13) and Muguga (14). Purification of piroplasms and preparation of piroplasm DNA was as described (15).

Isolation of genomic clones containing *T.parva* telomeric DNA

A library of sheared *T.parva* Uganda DNA fragments was prepared in λgt11 from purified piroplasm DNA (16). The library was screened with a 28mer oligonucleotide $(CCCTGAACCCTAAA)^2$ corresponding to the telomeric repeats of a cloned *Plasmodium berghei* telomere (17). The 1.74 kb insert from a positive bacteriophage was subcloned into the *Eco*RI site of pBluescript KS (Stratagene). A construct containing only TAS DNA was generated by deleting the telomeric repeats using a unique *Pae*R7I site in the insert and the *Hin*dIII site in the pBluescript polylinker. A construct containing predominantly

* To whom correspondence should be addressed

telomeric repeat DNA was generated by deletion of the TAS DNA using an exonucleaseIII/mung bean nuclease kit (Stratagene). The constructs were verified by nucleotide sequencing. A library of partially *Sau*3AI-digested *T.parva* Muguga piroplasm DNA fragments was constructed in λEMBL3 (18). The λEMBL3 library was screened with the 1.74 kb telomeric insert and two positive clones were selected for analysis. *Sal*I fragments of 2.0, 3.2 and 8.5 kb were subcloned from λTptel 7.11 into pBluescript KS.

Southern hybridisation

Restriction enzyme digestions were according to the manufacturer's instructions (New England Biolabs). Agarose gel electrophoresis, Southern blotting onto nylon (Hybond N, Amersham) filters and hybridisation of filters used standard procedures (19). DNA probes were labelled using a Prime-it kit (Stratagene). All filters were washed in 2 × SSC/0.1% SDS at 65°C unless indicated otherwise in the figure legends. Labelling, hybridisation and washing conditions for the *P.berghei* telomeric oligonucleotide were as described previously (20).

*Bal*31 exonuclease treatment of genomic DNA

T.parva piroplasm DNA (14 μg) was digested with 10 U *Bal*31 exonuclease (Bethesda Research Laboratories), calibrated to digest 400 bp/min. Serial time points were taken at 0, 1, 2 and 3 min. For each time point 2 μg aliquots of *Bal*31 treated DNA were extracted with phenol–chloroform, ethanol precipitated, digested with *Eco*RI and size fractionated on 0.8% agarose gels for Southern blot analysis.

Nucleotide sequencing

For sequencing, nested deletions were generated using an exonuclease III/mung bean nuclease kit (Stratagene), in combination with the synthesis of oligonucleotides, based on acquired sequence, as primers for regions which were not covered by the deletions. DNA sequencing by the chain termination method (21) used either the Sequenase™ kit (US Biochemicals) or the 'fmol™' DNA sequencing system (Promega). For direct sequencing the 'fmol' system was used according to the manufacturers instructions with 2 μg of uncloned genomic DNA as the template. Both strands of the DNA were sequenced. Sequences were deposited with the GenBank database under accesssion nos L36963 (1747 bp derived from a *T.parva* Uganda λgt11 clone) and L36964 (851 bp derived from *T.parva* Muguga λEMBL3 clone).

Pulsed-field gel electrophoresis (PFGE)

Preparation of high molecular weight *T.parva* piroplasm DNA embedded in low melting point agarose blocks and digestion of agarose-embedded DNA with the restriction enzyme *Sfi*I was performed as described (20). The *T.parva Sfi*I fragments were separated using contour clamped homogeneous electric field (CHEF) electrophoresis (22) and *T.parva* chromosome sized DNA molecules were separated using the pulsaphor system with point electrodes (Pharmacia). For CHEF electrophoresis, 1.5% agarose gels prepared in 0.5 × TBE were used. The *Sfi*I fragments were separated using a pulse frequency of 10 s for 16 h and 40 s for 2.5 h at 200 V. The *T.parva* chromosomes were separated in 1% agarose in 1 × TBE using pulses of 900 s for 24 h, 600 s for 24 h, 480 s for 24 h and 400 s for 24 h at 200 V. Size markers were concatemers of bacteriophage λ DNA CI 857 S7 and *S.cerevisiae* chromosomes (Cambridge Bioscience).

Polymerase chain reaction (PCR) amplification of DNA

Amplification reactions used *Taq* polymerase as specified by the manufacturer (Promega). Cycling conditions were 1 min at 94°C, 1 min at 55°C and 30 s at 72°C for 30 cycles. PCR products were cloned into the pCRtm11 vector (Invitrogen).

RESULTS

Isolation and validation of a *T.parva* telomeric DNA clone

A λgt11 clone containing a 1.74 kb *T.parva* genomic DNA fragment which hybridised strongly to an oligonucleotide derived from a *P.berghei* telomeric repeat sequence (17) was isolated, and the insert subcloned into pBluescript KS. The plasmid clone was designated pTpUtel. To verify that pTpUtel contained telomere-specific sequences, the clone was hybridised to *Eco*RI-digested *T.parva* Muguga piroplasm DNA which had been treated with *Bal*31 exonuclease. The probe hybridised to eight *Eco*RI fragments in untreated *T.parva* genomic DNA (Fig. 1, lane 1) all of which were susceptible to *Bal*31 digestion (Fig. 1A, lanes 2–4) indicating that they were located at or near telomeres. The smallest *Eco*RI fragment recognised by the probe was diffuse (Fig. 1A, lane 1), a feature typical of telomeric restriction fragments due to length heterogeneity in the telomeric repeats (5). The blot was re-hybridised with a *T.parva* ribosomal RNA gene probe, pTmr5.7, which is located in a chromosome internal position (23). The rRNA gene probe recognised a 5.7 kb *Eco*RI fragment which did not reduce in size after *Bal*31 digestion (Fig. 1B). Two derivatives of pTpUtel were generated. The first, pTpUsubtel, contained only TAS DNA centromere-proximal to the single *Pae*R7I site at base pair 987, the second, pTpUtelrep, contained the telomeric repeat sequences and the centromere-proximal 10 bp.

Nucleotide sequences of telomeric DNAs

The nucleotide sequence of the 1747 bp insert in pTpUtel was determined (GenBank accession no. L36963). 490 bp at the 3′ end of the cloned sequence consisted of interspersed tandem repeats of TTTTAGGGG (43 copies), TTTAGGG (13 copies), TTTTAGG (six copies) and TTTAGG (three copies). There was no obvious order in the arrangement of the repeats. These sequences conform to the typical consensus for eukaryotic telomeric repeats (5). A search of the nucleotide sequence databases revealed very high levels of similarity to the telomeric repeat sequences of the photosynthetic algae, *Chlamydomonas reinhardtii* (97% identity in a 299 bp overlap; reference 24), *Chlorella vulgaris* (87.7% identity in a 261 bp overlap; GenBank accession no. D26374) and the angiosperm *Arabidopsis thaliana* (84.8% identity in a 222 bp overlap; reference 25). The telomeric repeats of *P.falciparum* were less similar to *T.parva* (75% identity in a 411 bp overlap; reference 26). The 1250 bp of sequence centromere-proximal to the telomeric repeats was A/T (70%) and T (44%) rich. The Ts in the sequence were often non-randomly arranged repetitions of two to six, as indicated by a high positive autocorrelation function for this base. A 22 bp sequence TAGATTAGAGTGTTTTTAGAGT (two copies separated by

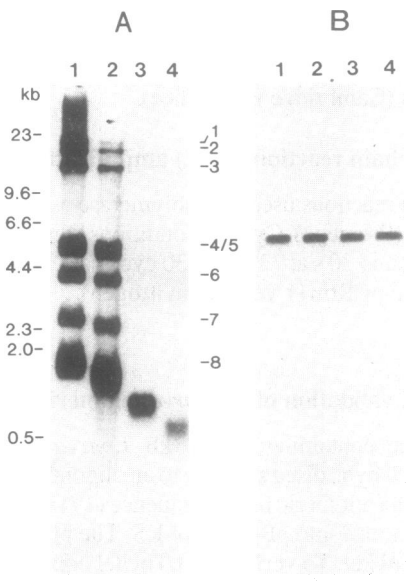

Figure 1. pTpUtel recognises *Eco*RI fragments in *T.parva* genomic DNA which are susceptible to digestion with exonuclease *Bal*31. *T.parva* Muguga genomic DNA was incubated for 1–3 min with *Bal*31 exonuclease, digested with *Eco*RI and size fractionated. Following Southern transfer the filter was hybridised with pTpUtel (**A**) and re-hybridised, after stripping, with pTpmr5.7, a ribosomal RNA gene probe (**B**). Lanes 1–4 show *T.parva* DNA treated with *Bal*31 for 0, 1, 2 and 3 min, respectively.

354 bp) was the only perfect repeat of 20 bp or more present in the TAS DNA. No perfect repeats were detected in 610 bp of TAS DNA from a second cloned *T.parva* telomere isolated from a λEMBL3 recombinant (GenBank accession no. L36964). No significant identities with sequences in the databases were detected in the TAS DNA.

Analysis of λEMBL3 clones containing TAS DNA

To isolate more extensive TAS DNA sequences, a λEMBL3 library was screened with the cloned *T.parva* DNA contained in pTpUtel. Two positively hybridising bacteriophages λTptel 4.12 and λTptel 7.11, containing insert sizes of ~16.5 and 13.5 kb, respectively, were isolated. The restriction maps of the *T.parva* sequences contained in λTptel 4.12 and λTptel 7.11, which are shown in Figure 2, revealed frequent 6 bp recognition restriction enzyme sites in the TAS DNA. The similarity of the restriction maps, together with the cross-hybridisation of corresponding restriction fragments, suggested that the two λ clones were derived from the same telomere. The 2.0 kb (pTpMtel2.0), 3.2 kb (pTpMsubtel 3.2) and 8.5 kb (pTpMsubtel8.5) *Sal*I fragments from λTptel 7.11 (see Fig. 2) were subcloned, and the nucleotide sequences of the termini were determined. The analysis revealed 231 bp of telomeric repeat sequence at the 3′ end of the 2.0 kb fragment (GenBank accession no. L36964). This was consistent with the cloning of an intact telomere. Genomic restriction fragments recognised by the cloned 2.0, 3.2 and 8.5 kb *Sal*I fragments were sensitive to *Bal*31 digestion (not shown), confirming that they were located at or near telomeres and ruling out the possibility that they might represent internally located telomere-similar sequences, such as those observed in *S.cerevisiae* and *A.thaliana* (5,27). Direct nucleotide sequencing of *T.parva* genomic DNA also demonstrated that the 2.0, 3.2 and 8.5 kb fragments were adjacent in the *T.parva* genome. The 2.0, 3.2 and 8.5 kb *Sal*I restriction fragments from λTptel 7.11 (see Fig. 2) predominantly recognised single *Sal*I fragments when hybridised to *T.parva* genomic DNA (Fig. 3). This result, together with the lack of cross-hybridisation between the 2.0, 3.2 and 8.5 kb fragments (data not shown), indicated the absence of extensive repeat sequences in the DNA. The diffuse nature of the genomic *Sal*I fragment recognised by the 2.0 kb fragment provided further confirmation of a telomeric location (Fig. 3, lane 1).

Localisation of cloned telomeric DNA sequences on the *Sfi*I restriction map

The existence of an *Sfi*I restriction map of the *T.parva* Muguga genome (12) allows cloned *T.parva* sequences to be mapped by hybridisation to blots of *Sfi*I-digested DNA fractionated by PFGE. The plasmid construct containing only telomeric repeat sequences, pTpUtelrep, hybridised to a minimum of six distinct *Sfi*I fragments in the DNA of the *T.parva* Muguga stock (Fig. 4A and B, lane 2), which had previously been identified as telomeric (12). Since *T.parva* Muguga has four chromosomes (12), and several of the telomeric *Sfi*I fragments co-migrate on CHEF gels (20), the data was consistent with the presence of this repeat at all telomeres. The probe hybridised to at least six distinguishable *Sfi*I fragments in the *T.parva* Uganda stock (Fig. 4A and B, lane 1). Several of the telomeric *Sfi*I fragments were size polymorphic between the two *T.parva* stocks. The construct, pTpUsubtel, which contained only sequences located centromere-proximal to the telomeric repeats, hybridised strongly to two size-polymorphic variants of *Sfi*I fragment 8 in *T.parva* Muguga DNA and to a single similar sized fragment in *T.parva* Uganda DNA. The

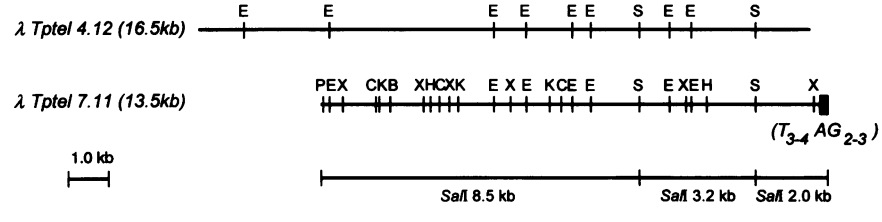

Figure 2. Physical maps of recombinant λphage containing telomeric repeat and TAS DNA. A selective restriction map is shown of λEMBL3 recombinant clones λTptel 7.11 (13.5 kb) and λTptel 4.12 (16.5 kb). Restriction enzyme sites are designated as follows: *Bam*HI = B; *Cla*I = C; *Eco*RI = E; *Hin*dIII = H; *Kpn*I = K; *Pst*I = P; *Sal*I = S; *Xba*I = X. λTptel 4.12 was analysed only with enzymes *Eco*RI and *Sal*I. The extent of the telomeric repeat sequences present in λTptel 7.11 is indicated. The location of 2.0, 3.2 and 8.5 kb *Sal*I fragments in λTptel 7.11 which were subcloned and used as probes is also shown.

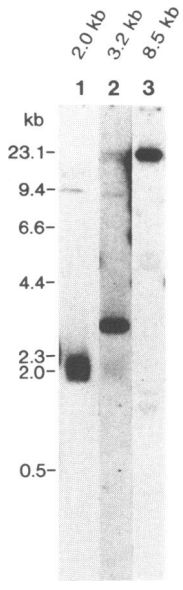

Figure 3. Genomic organisation of *SalI* fragments containing telomeric repeat and TAS DNA. Southern blots of *SalI*-digested *T.parva* piroplasm DNA are shown hybridised with the 2.0 kb (lane 1), 3.2 kb (lane 2) and 8.5 kb (lane 3) *SalI* fragments subcloned from λTptel 7.11.

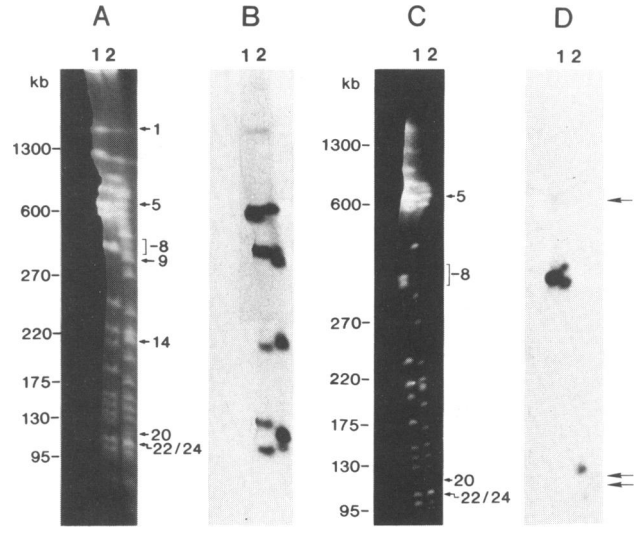

Figure 4. Localisation of telomeric repeat and TAS DNA contained in pTpUtel and pTpUsubtel to *SfiI* restriction fragments. (**A** and **C**) Ethidium bromide stained 1.5% agarose gels of *SfiI*-digested *T.parva* Uganda (lane 1) and *T.parva* Muguga (lane 2) DNA separated by CHEF electrophoresis. (**B** and **D**) Blots of the gels in (A) and (C). The blots were hybridised as follows: (B) with pTpUtelrep (telomeric repeats) and (D) with pTpUsubtel (TAS DNA) The numbered arrows indicate *T.parva* Muguga *SfiI* fragments previously determined as being telomerically located (12). The heterogeneous fragment 8 is marked by a bracket. The weak hybridisation to *SfiI* fragment 1 in (B) is due to inefficient transfer of this 1800 kb fragment.

probe hybridised weakly to three additional telomeric *SfiI* fragments in both stocks (Fig. 4C and D; the weakly hybridising fragments are indicated by arrows). *SfiI* fragment 8 appears as two bands in *T.parva* Muguga due to parasite heterogeneity within the uncloned stock (20). Sequences with some homology to the cloned TAS DNA contained in pTpUsubtel were therefore present at or near telomeres on three of the four *T.parva* chromosomes (12). After high stringency washing (0.2 × SSC/0.1% SDS), hybridisation was observed only to *SfiI* fragment 8 (not shown), which is located on the genome map at the right hand telomere of chromosome 1 (12). The chromosome 1 origin of the TAS DNA was confirmed by hybridisation to a blot of an intact chromosome separation (data not shown). When hybridised to blots of *SfiI* fragments and intact chromosomes, the EMBL3 clone λTptel 4.12, hybridised exclusively to *T.parva* Muguga *SfiI* fragment 20 (Fig. 5A and B), which is a telomeric fragment derived from the left hand end of chromosome 1 (12), and to chromosome 1 (Fig. 5C and D). Similar results were obtained with the 8.5 kb *SalI* fragment from λTptel 7.11 (not shown). This suggested that there was no substantial homology of the cloned DNA in the λEMBL3 clones to sequences at other telomeres. The results indicated that telomeric DNAs derived from opposite ends of chromosome 1 had been isolated.

A conserved sequence located centromere-proximal to the telomeric repeats

To examine whether any sequence identity existed between different telomeres in the TAS DNA centromere-proximal to the telomeric repeats, 600 bp of sequence directly adjacent to the telomeric repeats was compared with the DNA cloned in pTpUtel and that cloned in pTpMtel2.0. The DNA sequences were aligned using the CLUSTAL programme (28). The sequences in pTpUtel were derived from one telomere of *T.parva* chromosome 1 (Fig. 4), and the sequences in pTpMtel2.0 were derived from the opposite telomere of *T.parva* chromosome 1 (Fig. 5). The only significant conserved region was a 126 bp sequence (shown in the lower panel of Fig. 6) directly centromere-proximal to the telomeric repeats which was 94% identical between the two telomeres. The conserved sequence contained a single copy of the minor repeat TTTAGG and one partial copy, TTTTAG, of the major repeat, but was otherwise unrelated to the telomeric repeats. To test whether the 126 bp sequence was present at other telomeres, it was amplified from pTpUtel using primers ILO 2299 (5′-ACT TCG GGG TTT TCA AAT TG-3′) and ILO 2359 (5′-CTT AAT GTC CTA AAG CT-3′) and cloned. Nucleotide sequencing confirmed that the correct sequence had been amplified. The PCR product hybridised at moderate stringency to all eight *T.parva* Muguga telomeric *SfiI* fragments (Fig. 6B and D, upper panel). The apparent hybridisation to an additional non-telomeric *SfiI* fragment in Figure 6B is not evident in Figure 6D and is probably attributable to persistent partial digestion, which is often observed with *SfiI* digestion of *T.parva* DNA (12). The 126 bp PCR product hybridised only to the same eight *Bal*31-sensitive *Eco*RI fragments in *T.parva* Muguga DNA as the pTpUtel probe (data not shown). The results were consistent with the presence of a sequence with a high degree of identity to the conserved 126 bp sequence adjacent or close to the telomeric repeats at all telomeres.

DISCUSSION

Telomeric repeat sequences

The telomeric repeat sequences of *T.parva* conform to the consensus described for most eukaryotes (5). The TTTTAGGG repeat unit of *T.parva* and *C.reinhardtii*, is the least G/C rich

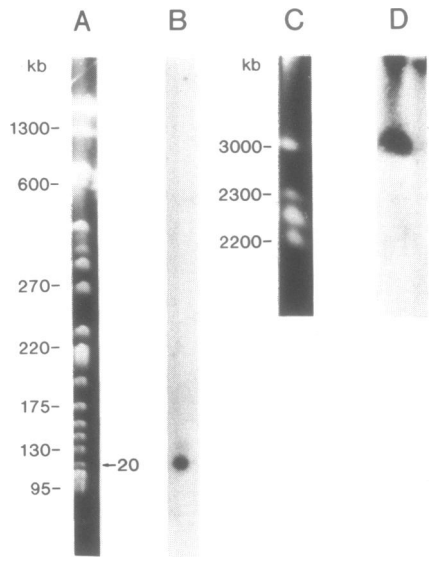

Figure 5. Localisation of TAS DNA contained in λTptel4.12 to *T.parva* Muguga chromosome 1 and *Sfi*I restriction fragment 20. (**A**) Ethidium bromide stained 1.5% agarose gel of *Sfi*I-digested *T.parva* Muguga DNA separated by CHEF electrophoresis. (**B**) Blot of the gel in (A) hybridised with λTptel 4.12. (**C**) Ethidium bromide stained 1% agarose gel of *T.parva* Muguga chromosome sized DNA molecules separated using the Pulsaphor system. (**D**) A blot of the gel in (C) hybridised with λTptel 4.12. The blots were washed in 0.1 × SSC/0.5% SDS.

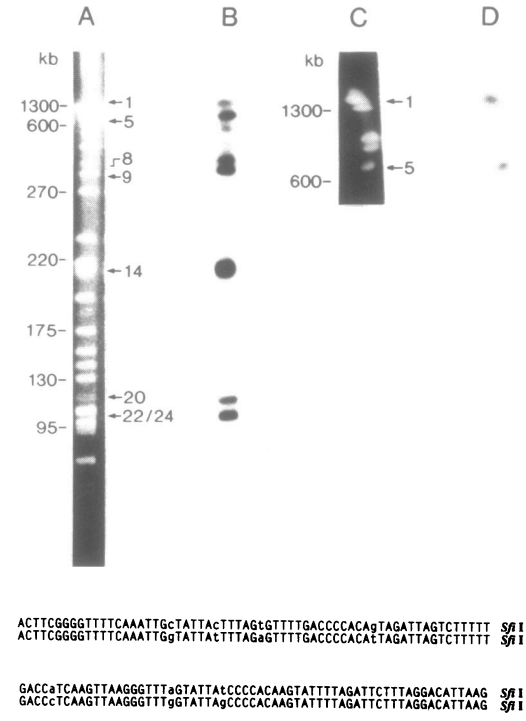

Figure 6. A 126 bp conserved sequence located centromere-proximal to the telomeric repeats is present on all eight *T.parva* Muguga telomeric *Sfi*I fragments. The lower panel shows a CLUSTAL alignment (28) of 126 bp of DNA, adjacent to the telomeric repeats, derived from the telomeres contained on *Sfi*I fragments 20 and 8, which are at opposite ends of the largest *T.parva* chromosome (12). Mismatched bases are in lower case letters. The upper panel shows CHEF gel separations of *Sfi*I-digested *T. parva* Muguga DNA in (**A**) and (**C**), hybridised with a PCR product of the 126 bp sequence in sections (**B**) and (**D**). The blots were washed in 0.5 × SSC/0.1% SDS.

telomeric repeat known (24), but the inverse correlation between genomic DNA and telomeric DNA G/C content observed in most organisms (24) does not hold for *T.parva,* in which the telomeric repeat G/C content is 37.5% but the genomic DNA G/C content is only 30% (29). The *T.parva* telomeric repeats are strikingly similar to those of two photosynthetic algae and an angiosperm. The sequence identity between the invariant TTTTAGGG array of the alga *C.reinhardtii* and the *T.parva* telomeric repeats is 97%, the differences arising due to the occurrence of variants of the predominant repeat in *T.parva*. The similarity between the *T.parva* and plant telomeric repeat sequences may be an example of convergence in molecular evolution. However, evidence exists suggesting that sporozoan protozoa have evolved from photosynthetic algae (reviewed in 30). It is therefore tempting to speculate that a common ancestral telomeric repeat has been conserved in *Theileria* and plants. A hypothesis which is supported by the similarity of the telomeric repeats of the algae *C.reinhardtii* and *C.vulgaris* and the angiosperm *A.thaliana*. Extensive sequence conservation over a wide evolutionary distance contrasts with the divergence observed in the telomeres of some other organisms such as relatively closely related species of budding yeasts (31). The extent of the telomeric repeats can vary from 20 bp to 20 kb in different organisms (5). In *T.parva*, as judged by the size of the smallest genomic restriction fragments recognised by the telomeric probes, the telomeric repeats are between 0.5 and 1 kb in length, which is similar to other lower eukaryotes (1,32).

Telomere-associated sequences

The organisation of the TASs in *T.parva* is dissimilar to that reported in most organisms. The TASs of yeast, protozoa, insects and vertebrates typically consist of middle repetitive sequences, which are often the location of genomic rearrangements and are polymorphic in their chromosomal distribution (reviewed in 5,33). These can be either short tandemly arranged 'minisatellites' (26,34–37) or moderately repetitive sequences ranging from 350 bp to several kb in length (38–41). Nucleotide sequence analysis did not reveal the presence of similar repeated sequences in *T.parva* TAS DNA at the two telomeres examined. The hybridisation of the TAS DNA contained in the λTptel 4.12 recombinant primarily to a single *Sfi*I restriction fragment (Fig. 5), the lack of cross-hybridisation between *Sal*I restriction fragments from the λEMBL3 clones and the recognition of single genomic *Sal*I fragments by the sequences subcloned from λTptel 7.11, was also consistent with the absence of extensive internal repetition in the DNA centromere-proximal to the telomeric repeat sequences.

In *Plasmodium* species chromosome size polymorphisms between clones can be detected by PFGE and the major source of this polymorphism is DNA rearrangements in TAS regions (reviewed in 9). In *T.parva,* chromosome size polymorphisms are difficult to detect due to the existence of only four chromosomes which are relatively large (2.1–3.0 Mb) and similar in size (12), but this study shows that some telomeric *Sfi*I fragments, including fragment 20 from which the *T.parva* λEMBL3 clones were derived, are size polymorphic between stocks (Fig. 4A and B). Blocks of tandemly repeated sequences located between 10 and 45 kb from most telomeres have been shown to be involved in the

generation of chromosome size polymorphisms in *P.falciparum* (8). Preliminary experiments in which *T.parva* Muguga telomeric *Sfi*I fragments 20 (110 kb) and 9 (294 kb) were gel-purified and hybridised to blots of *Sfi*I-digested *T.parva* DNA revealed significant hybridisation only to the homologous *Sfi*I fragment (Baljinder Sohanpal and Richard Bishop, unpublished), suggesting the absence of major tracts of cross-hybridising repetitive sequences upstream of these telomeres.

It has been reported recently that repetitive DNA sequences with a conserved organisation extending over many kilobases are common to most *P.falciparum* telomeres (6,7). In *T.parva* Muguga we have shown that a conserved 126 bp sequence element is present centromere-proximal to the telomeric repeats. The hybridization of the pTpUsubtel probe (which does not contain the 126 bp sequence) with a subset of telomeric *Sfi*I fragments indicates that additional TASs are shared between different telomeres. It is speculated that such homologies in the TASs may facilitate recombination between the telomeres of non-homologous chromosomes (5,6,33). However, the hybridisation of the λTptel 4.12 clone only to *Sfi*I fragment 20 suggests that, by contrast with *P.falciparum*, such homologies are of limited extent in *T.parva*.

The scarcity of repeat sequences in the TASs suggests the possibility that genes may be located close to telomeres in *T.parva*, whose genome is only 10^7 bp in size (20,29). The data also imply that the size polymorphism observed in *T.parva* telomeric *Sfi*I fragments (Fig. 4; reference 20) is not caused only by rearrangements in repetitive DNA sequences. The cloned sequences described here will allow studies on the molecular basis of telomeric polymorphism in *T.parva* and investigation of whether protein coding genes are located in the TASs.

ACKNOWLEDGEMENTS

We thank Mr J. Kiarie and Mr T. Ndolo for technical assistance. We are indepted to Drs V. Nene and N. Murphy for their critical reading of and suggestions on this manuscript. We are grateful to Mr M. Kibe for the use of the λEMBL3 *T.parva* genomic DNA library. This is ILRAD publication no. 1384

REFERENCES

1 Blackburn, E. H. (1991) *Nature* **350**, 569–573.
2 Sandell. S. L. and Zakian, V. A. (1993) *Cell* **75**, 729–739.
3 Levis, R. W., Ganesan, R., Houtchens, K., Tolar, L.A. and Sheen, F. (1993) *Cell* **75**, 1083–1093.
4 Blackburn, E.H. and Szostak, J. W. (1984) *Annu. Rev. Biochem.* **53**, 163–194.
5 Zakian, V.A. (1989) *Annu. Rev. Genet.* **23**, 579–604.
6 De Bruin, D., Lanzer, M., and Ravetch, J.A. (1994) *Proc. Natl. Acad. Sci. USA* **91**, 619–623.
7 Scotti, R., Pace, T. and Ponzi, M. (1993) *Mol. Biochem. Parasitol.* **58**, 1–6.
8 Corcoran, L.M., Thompson, J.K., Walliker, D. and Kemp, D.J. (1988) *Cell* **53**, 807–813.
9 Janse, C. J. (1993) *Parasitol. Today* **9**, 19–22.
10 Vernick, K.D., Walliker, D. and McCutchan, T.F. (1988) *Nucleic Acids Res.* **16**, 6973–6985.
11 ole-Moi Yoi, O.K. (1989) *Exp. Parasitol.* **69**, 204–210.
12 Morzaria, S.P. and Young, J.R. (1992) *Proc. Natl. Acad. Sci. USA* **89**, 5241–5245.
13 Minami, T., Spooner, P. R., Irvin, A. D., Ocama, J. G. R., Dobbelaere, D. A. E. and Fujinaga, T. (1983) *Res. Vet. Sci.* **335**, 334–340.
14 Brocklesby, D. W., Barnett, S. F. and Scott, G. R. (1961) *British Vet. J.* **117**, 529–531.
15 Conrad, P.A., Iams, K., Brown, W.C., Sohanpal, B. and Ole-Moi, O. (1987) *Mol. Biochem. Parasitol.* **25**, 213–226.
16 Young, R.A., Mehra, V., Sweetser, D., Buchanan, T., Clark-Curtis, J., Davis, R. W. and Bloom, B.R. (1985) *Nature* **316**, 450–452.
17 Ponzi, M., Pace, T., Dore, E. and Frontali, C. (1985) *EMBO J.* **4**, 2991–2995.
18 Frischauf, A., Lehrach, H., Poustka, A. and Murray, N. (1983) *J. Mol. Biol.* **170**, 827–842.
19 Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) *Molecular Cloning: a Laboratory Manual*. Cold Spring Harbor University Press, Cold Spring Harbor, NY.
20 Morzaria, S.P., Spooner, P.R., Bishop, R.P., Musoke, A.J. and Young, J.R. (1990) *Mol. Biochem. Parasitol.* **40**, 203–212.
21 Sanger, F., Nicklen, S. and Coulsen, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467.
22 Chu, G., Vollrath, D. and Davis, R.W. (1986) *Science* **234**, 1582–1585.
23 Kibe, M.K., ole-Moi Yoi, O.K., Nene, V., Khan, B.K., Allsopp, B.A., Collins, N.E., Morzaria, S.P., Gobright, E.I and Bishop, R.P. (1994) *Mol. Biochem. Parasitol.* **66**, 249–259.
24 Petracek, M.E., Lefebrve, P.A., Silflow, C.D. and Berman J. (1990) *Proc. Natl. Acad. Sci.* USA **87**, 8222–8226.
25 Richards, E. J., Chao, S., Vongs, A. and Yang, J. (1992) *Nucleic Acids Res.* **20**, 4039–4046.
26 Vernick, K.D. and McCutchan, T.F. (1988) *Mol. Biochem. Parasitol.* **28**, 85–94.
27 Richards E.J., Goodman, H.M. and Ausubel, F.M. (1991) *Nucleic Acids Res.* **19**, 3351–3357.
28 Higgins, D.G. and Sharp, P.M. (1988) *Gene* **73**, 237–244.
29 Allsopp, B.A. and Allsopp, M.T.E.P. (1988) *Mol. Biochem. Parasitol.* **28**, 77–84.
30 Wilson, R. J., Gardner, M.J., Feagin, J.E., Williamson, D. H. (1991) *Parasitol. Today* **7**, 134–136.
31 Mceachern, M.J. and Blackburn, E.H. (1994) *Proc. Natl. Acad. Sci. USA* **91**, 3453–3457.
32 Coleman, M.J., McHale, M.T., Arnau, J., Watson, A. and Oliver, R.P. (1993) *Gene* **132**, 67–73.
33 Biessmann, H., and Mason, J.M. (1992) *Adv. Genet.* **30**, 185–249.
34 Bachmann, L., Raab, M. and Sperlich, D. (1990) *Genetica* **83**, 9–16.
35 de Lange, T., Shiue, L., Myers, R.M., Cox, D.R., Naylor, S.L., Killery, A.M. and Varmus, H.E. (1990) *Mol. Cell. Biol.* **10**, 518–527.
36 Brown, W.R.A., Mackinnon, P.J., Villesante, A., Spurr, N., Buckle, V.J. and Dobson, M.J. (1990) *Cell* **63**, 119–132.
37 Chan, C.S. and Tye, B.K. (1983) *J. Mol. Biol.* **168**, 505–523.
38 Chan, C.S. and Tye, B.K. (1983) *Cell* **33**, 563–573.
39 Pace, T., Ponzi, M., Dore, F. and Frontali, C. (1987) *Mol. Biochem. Parasitol.* **24**, 193–202.
40 Dore, E., Pace, T., Ponzi, M., Picci, L. and Frontali, C. (1990) *Mol. Cell. Biol.* **10**, 2423–2427.
41 Horowitz, H. and Haber, J.E. (1984) *Nucleic Acids Res.* **12**, 7105–7121.