 Nucleic Acids Research, 1999, Vol. 27, No. 12 **2473–2477**

High conservation of the fine-scale organisation of chromosome 5 between two pathogenic *Leishmania* species

C. Ravel, P. Dubessay, C. Britto, C. Blaineau, P. Bastien and M. Pagès*

CNRS EP 613 'Génome et Biologie Moléculaire des Protozoaires Parasites', Laboratoire de Parasitologie, Faculté de Médecine, 163 Rue A. Broussonet, 34090 Montpellier, France

Received March 11, 1999; Revised and Accepted April 23, 1999

ABSTRACT

In a previous work we showed a remarkable conservation of the general structure of the genome (chromosome number and synteny) among different pathogenic species of Old World *Leishmania*, indicating the absence of major interchromosomal rearrangements during evolution. In the present study, we have compared the fine structure of chromosome 5 among two of these divergent species (*Leishmania major* and *Leishmania infantum*) by means of physical mapping. Remarkably, the 42 markers jointly mapped on these two chromosomes were found in an identical order along the chromosome. This perfect colinearity of the markers is in striking contrast to the large genetic distance that separates these species and suggests that conservation of the fine-scale organisation of chromosomes may be critical in *Leishmania*. If this colinearity is confirmed on the whole of the chromosome set, the current systematic sequencing programme of the genome of *L.major* should greatly help in the development of comparative genetics between different species of *Leishmania*.

INTRODUCTION

The protozoan parasite *Leishmania* is the causative agent of a wide spectrum of human diseases of which some can be fatal in the absence of drug treatment. Endemic foci are distributed in 97 countries on four continents. These diseases constitute a serious public health problem affecting around 15 million people around the world (1). The genus *Leishmania* comprises about 15 species complexes infecting mammals; in man, they can cause widely polymorphic clinical pictures, from fatal visceral disease to benign cutaneous lesions, depending on the species as well as the host immune status. In 1994, the World Health Organisation (WHO) launched a specific programme for the mapping and sequencing of the genome of *Leishmania* which was coordinated through the *Leishmania* Genome Network (LGN) (http://www.ebi.ac.uk/parasites/leish.html); the reference organism selected for this programme was the *Leishmania major* 'Friedlin' strain (2).

The general organisation of the genome is now well documented for several species of *Leishmania*. It is a small genome (35 Mb) subdivided in a relatively high number of chromosomes (3). Recently, by the use of 244 chromosome-specific markers as probes, it was possible to define synteny groups corresponding to whole chromosomes and thus to identify all the chromosomes in the genomes of 11 different species (3,4). A chromosome nomenclature was then proposed for this protozoan. Thus, the complete chromosome set (for the haploid genome) was shown to comprise 36 chromosomes for all Old World species (3,5), 35 for the New World complex *Leishmania braziliensis* and 34 for the New World complex *Leishmania mexicana* (4). Remarkably, these synteny groups were found to be highly conserved among species, including all species with a lower number of chromosomes. This contrasts with the extensive size polymorphisms exhibited by homologous chromosomes among strains (3,6,7). Medium range restriction maps of full-length homologous chromosomes of different sizes permitted us to show the role of the instability of subtelomeric regions in these size variations, as opposed to less variable central regions (8,9). Subtelomeric regions are known to contain non-coding repetitive sequences (10–12), while the central regions seem to consist of long stretches of densely packed coding sequences (12). Nevertheless, the conservation of the order of the genes/markers along the chromosomes (i.e. of physical linkage as opposed to synteny) among species had yet to be determined. Here, we have physically mapped and compared the fine-scale organisation of chromosome 5 in two divergent species, *Leishmania infantum* and *L.major*, for which no restriction sites are conserved. Forty-two DNA markers identifying 21 loci previously mapped along this chromosome in *L.infantum* were found to be in the same order in *L.major*. This perfect map colinearity between two divergent species should promote comparative genomics studies for this protozoan parasite.

MATERIALS AND METHODS

Parasite chromosomal DNA preparation

The two cloned strains of *L.major* and *L.infantum* ('Friedlin' and LEM1317, respectively) examined in this study have been

*To whom correspondence should be addressed. Tel: +33 4 67 63 55 13; Fax: +33 4 67 63 00 49; Email: gpp@sc.univ-montpl.fr

described (2,9). The *in vitro* cultivation conditions were as in Blaineau *et al.* (7).

The chromosome nomenclature used here is that proposed by Wincker *et al*. (3) and used by the LGN (2). The preparation of chromosomal DNA in agarose blocks for PFGE and the restriction analysis of the PFGE gel-purified chromosome 5 were as described in Ravel *et al*. (9).

Screening of the cosmid library

The reference resource for the sequencing programme of the *L.major* genome is a cosmid library of 9216 cosmid clones constructed in the shuttle vector cLHyg and theoretically covering nine times the genome (13). High density grids on which the whole cosmid library was arrayed (14) were generously provided by Alasdair Ivens (Imperial College of Science, Technology and Medicine, London) through the LGN. They were hybridised with 81 chromosome 5-specific anonymous DNA probes (termed st… followed by a figure) previously described (9), as well as, occasionally, cosmid end probes which were purified from agarose gels after enzyme restriction. The probes were ^{32}P-labelled using random-primed synthesis as described previously (10). All hybridisations were performed in 5× SSPE (1× SSPE = 180 mM NaCl, 10 mM NaH_2PO_4, 1 mM EDTA, pH 7.5), 1% SDS at 65°C overnight, followed by washes at high stringency (0.1× SSPE, 0.1% SDS). Only cosmids showing clear hybridisation signals were retained for further analysis. The selected positive cosmids were then subcloned and the subclones analysed by restriction fingerprinting.

Cosmid restriction analysis

For restriction analysis, purified cosmid DNA obtained by alkaline lysis was restricted with *Not*I and *Pst*I, separated in a 0.7% agarose gel, Southern blotted onto nylon filters (Hybond N+; Amersham) and hybridised with the ^{32}P-labelled oligonucleotides T7 (5'-ATAATACGACTCACTATAGG-3') and T3cLHyg (5'-TCCTTGCCCTGCAGTTG-3') present on each side of the polycloning site. Filters were hybridised in 5× SSPE, 1% SDS at 45°C, then washed at the same temperature with 3× SSPE, 0.1% SDS.

RESULTS

Construction of a cosmid contig for chromosome 5 in *L.major*

In a previous study, 81 DNA probes isolated from chromosome 5 of *L.infantum* were located on 31 restriction fragments 5–45 kb in size constituting the full-length physical map of this chromosome (9; unpublished data). These 81 probes were hybridised onto high density arrays of a genomic library of 9216 cosmid clones theoretically covering nine times the genome of the reference strain of the LGN, *L.major* 'Friedlin' (13). Although these were cross-species hybridisations, high stringencies were used so as to lower the high rate of false-positive signals (13). First, 39 probes were removed from the analysis: 19 revealing a high number of clones (>60) and 20 giving no signal at all. The 42 remaining probes revealed a mean of 17 (from 2 to 52) clones per probe. They allowed the identification of 21 loci out of the 31 loci previously mapped on chromosome 5 of *L.infantum* (Fig. 1, lines C and D). They were then used for the construction of cosmid contigs. In the first place, 191 clones which hybridised with a minimum of two probes were retained for primary analysis. These hybridisation data allowed the integration of these clones into five (sub)contigs. Contig joining and gap closure were achieved by the use of contig end probes (CEPs) obtained by gel purification and radiolabelling of restriction fragments from the cosmid clones located at the ends of the subcontigs. These CEPs were hybridised onto Southern blotted restriction patterns of the terminal cosmids of each subcontig, allowing the assembly of a single chromosomal contig. Moreover, one of the terminal restriction fragments of cosmid L8138 used as a CEP on high density grids of the library allowed the identification of a more distal clone, L9501. Based on the whole of these hybridisation data, a number of redundant clones could be removed, and 62 cosmid clones were selected for further analysis. The physical overlap between cosmids was then assessed by RFLP analysis: the double restriction by *Not*I and *Pst*I, of which some of the sites are close to the polycloning site, allowed the release of the inserts and their fragmentation into specific patterns comprising relatively few bands (Fig. 2). This revealed a high overlap between certain cosmids (Fig. 2A). The hybridisation of radiolabelled DNA from each of the selected cosmids onto cosmid restriction patterns (Fig. 2B) allowed confirmation of the overlap, as well as detection of other cosmids showing less overlap, with few or no fragments of the same size.

Construction of the restriction map of chromosome 5 of *L.major*

The restriction sites of the rare-cutting enzymes *Ase*I, *Dra*I, *Spe*I and *Ssp*I were mapped in independent experiments, on the one hand, on the cosmid clones selected and, on the other hand, directly on chromosome 5 of *L.major* 'Friedlin', according to mapping strategies previously described (9). These data allowed the precise positioning of the cosmids in relation to one another as well as along the chromosome. They were also useful in detecting chimeric clones, which were discarded. In the end, the restriction sites mapped on the different cosmids were concordant for all overlaps and in agreement with the purified chromosome analysis data. Finally, this medium range restriction mapping made possible calculation of the position of the contig relative to the ends of the chromosome. By the use of hybridisation of cosmid DNA onto Southern blots of restricted gel-purified chromosome 5, the terminal cosmids L9501 and L1356 (Fig. 1, line E) could clearly be positioned each at ~20 kb from the telomeric ends of the chromosome (Fig. 3).

This lengthy process of screening and elimination finally led to selection of a minimum tile set of 24 cosmid clones, covering most of chromosome 5 (Fig. 1, line E).

Comparative physical maps of chromosome 5 between *L.major* and *L.infantum*

The compared restriction maps of chromosome 5 from *L.infantum* LEM1317 and *L.major* 'Friedlin' are shown in Figure 1 (lines A and F). As expected, the locations of the restriction sites of the four enzymes used in mapping are not conserved between these two species. A most remarkable finding is the perfect colinearity of the DNA markers mapped on both chromosomes: the 21 loci were all found in exactly the same order along both chromosomes,

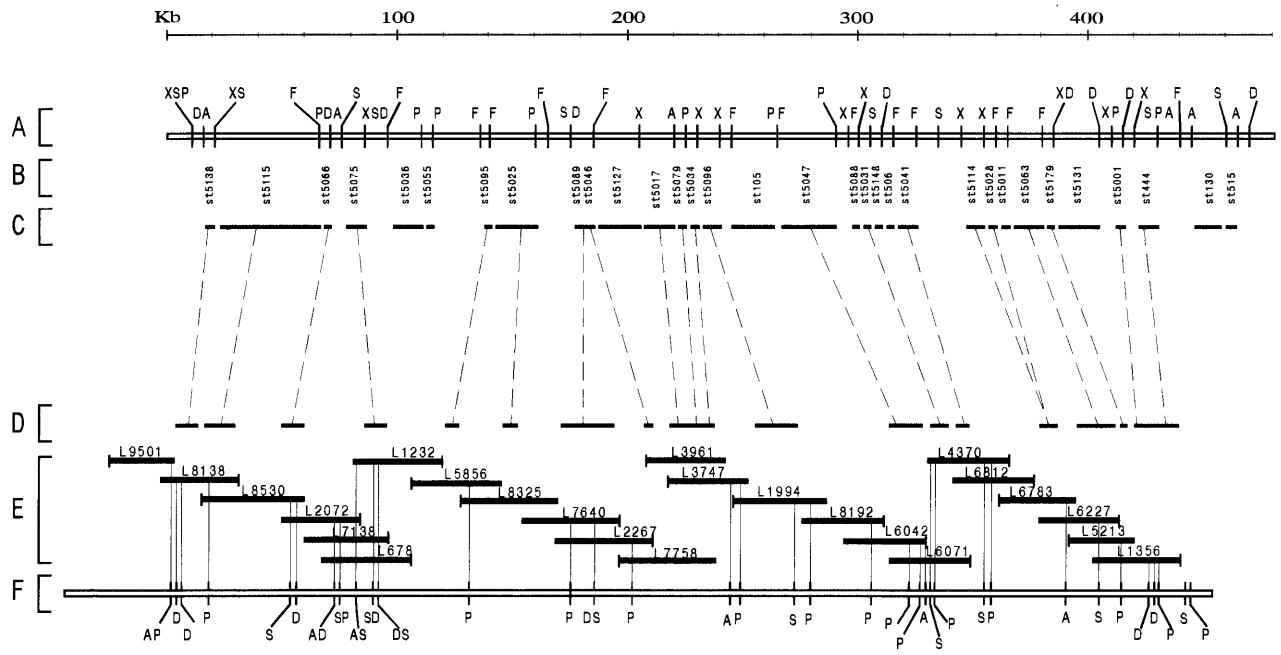

Figure 1. Comparison of the conserved linkage group on chromosome 5 for *L.infantum* LEM1317 (top, line A) and *L.major* 'Friedlin' (bottom, line F). Restriction sites are noted: A, *Ase*I; D, *Dra*I; F, *Sfi*NI; P, *Spe*I; S, *Ssp*I; X, *Xba*I. The locations of 32 out of the 42 (st…) DNA markers retained for the study are shown in line B below the restriction map of *L.infantum* in line A. The 31 loci identified by these probes are shown as bold horizontal lines (line C). Hybridisation of the 42 DNA markers allowed the identification of 21 loci on chromosome 5 of *L.major* (shown likewise in line D). The correspondence between loci of the two strains/species is represented as thin oblique lines: the order of the loci along the chromosome is conserved, although inter-loci distances may vary. The minimal tile set of 24 cosmid clones (horizontal lines) selected for DNA sequencing is shown in line E. The T7 end of each cosmid is marked as a small vertical line. The restriction map of chromosome 5 of *L.major* 'Friedlin' is shown in line F, with the corresponding restriction sites on the cosmid clones projected above on line E.

although the inter-loci distances sometimes appeared different. Another interesting finding was the size difference of both subtelomeric ends between the two strains/species: based on the most distant DNA markers shared by the extremities of both chromosomes (st5138 and st444 for the left and right ends, respectively), the left end of chromosome 5 in 'Friedlin' was found to be 35 kb larger than that in LEM1317, while its right end was 30 kb smaller (Fig. 1, lines A and F). These size differences may be due to the instability of subtelomeric sequences (see Discussion).

DISCUSSION

The genome of Old World *Leishmania* pathogenic species comprises 36 chromosomal synteny groups which are conserved among highly divergent species (3,5). Here, we have shown that the chromosomal organisation is also conserved on the low intrachromosomal scale, with a remarkable conservation of the order of the markers along chromosome 5 between the two species, *L.major* and *L.infantum*. The 21 loci jointly mapped on both chromosomes are distributed on average every 25 kb along the chromosome; at this scale, no rearrangements of the inversion or translocation type could be detected. This fine-scale conservation of the chromosomal organisation between two species phenotypically widely distinct is surprising, and the following points can be discussed.

Some of our data are attributable to the nucleotide divergence rate between the two species. The restriction sites located in the maps of both chromosomes are not conserved, which gives a minimum divergence rate of >8%. Moreover, a quarter of the *L.infantum* DNA anonymous probes did not cross-hybridise to high density filters of the *L.major* cosmid library. These data are in agreement with the estimates of 13–25% calculated by Beverley *et al.* (15). Based upon this figure, the latter authors estimated the separation between the two species at ≥10 million years. This great genetic divergence contrasts with the apparent stability of at least the central regions of the chromosomes (8,9,12; this study). The structural conservation of genomes varies widely from one group of living organisms to another, even down to the genus level (16). In protozoan parasites, similar studies to ours have been carried out on the genomes of different species of *Plasmodium* (17,18). Out of 42 homologous genes analysed among *Plasmodium chabaudi* and *Plasmodium falciparum* (rodent- and man-specific parasites, respectively), 26 appeared conserved within 10 synteny groups, but these were dispersed on different chromosomes (18). The authors estimated that >60% of the total number of housekeeping genes shared between these two species were likely to be within a conserved synteny. A better conservation of the synteny groups was observed in an earlier study on four rodent-specific species (17). Our results concerning different species of *Leishmania* having divergent hosts shows a much stricter conservation of the genome, at the level of synteny groups (3,4) as well as of the intrachromosomal organisation, where

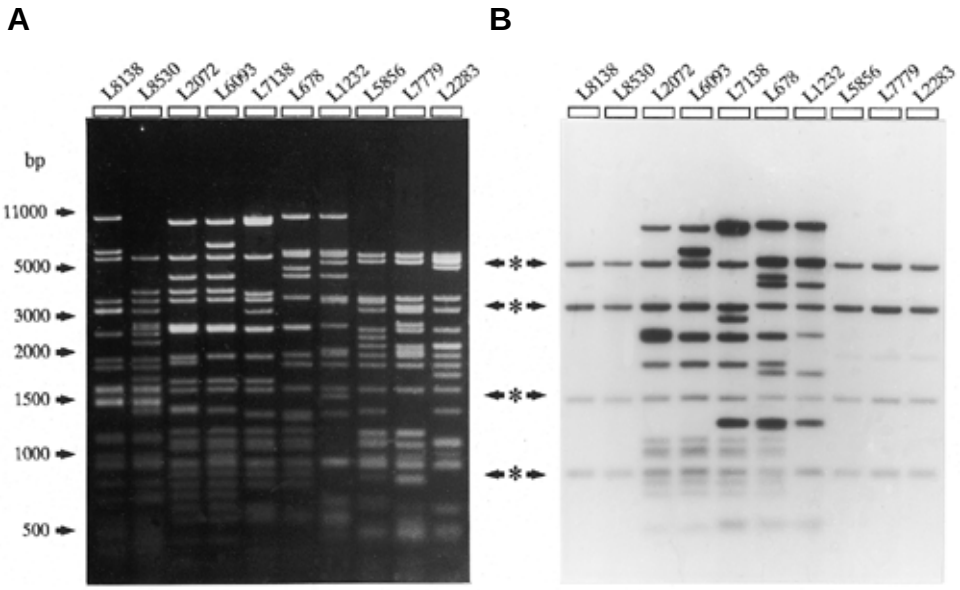

Figure 2. Assessment of the cosmid clone overlaps by restriction analysis (**A**) and Southern blot analysis (**B**). Cosmid DNA was restricted with *Not*I and *Pst*I and electrophoresed in an agarose gel (A). The four fragments corresponding to restriction of the cosmid vector cLHyg are indicated by arrowheads and asterisks. Clones L2072, L6093 and L7138 exhibit a large overlap with >50% of the bands in common. In (B), the same gel was Southern blotted and hybridised with the total DNA of cosmid L678: four cosmid clones (L2072, L6093, L7138 and L1232) were confirmed as overlapping with L678. The clones which present a large overlap (e.g. L6093) are not included in the minimum tile path in Figure 1.

linkage appears completely conserved (this study). The complete sequence of chromosome 1 of *L.major* has shown an unusual distribution of protein coding genes: the first 29 genes are all encoded on one strand, while the remaining 50 genes are encoded on the opposite strand (12). Such an organisation might act as a drastic functional constraint which would restrict intrachromosomal rearrangement eventualities.

We have shown here that the size of both terminal regions of chromosome 5 is different among the two strains/species analysed: the left end in 'Friedlin' appears to be 35 kb larger than that in LEM1317, while its right end is 30 kb smaller (Fig. 1). We have previously demonstrated that the right end of chromosome 5 in *L.infantum* exhibits an intraspecific size instability which might be linked to the presence of tandemly repeated sequences of the minisatellite type (LiSTIR1 and Lmet2) (9,10). These repeat sequences may be present in *L.major* as either divergent or less abundant sequences on chromosome 5 (our unpublished data) and chromosome 1 (12; our unpublished data). More classes of tandem repeats have been shown to be present in *L.major*, but they do not cross-hybridise with *Leishmania donovani* (11,12). This suggests that the organisation of these regions might be species specific. A chromosomal compartmentalisation has been observed in several species of protozoan parasites (19–21). It consists of unstable subtelomeric regions flanking stable 'core' central regions. A functional role for these unstable repeated regions is suspected. In *Plasmodium* the subtelomeric regions contain tandem repeats coding for antigens and membrane proteins (22) and in *Trypanosoma brucei* they comprise the telomeric variant surface glycoprotein gene expression sites (23). The tandemly repeated sequences, whether coding or non-coding, present in these regions in *Leishmania* might be involved in

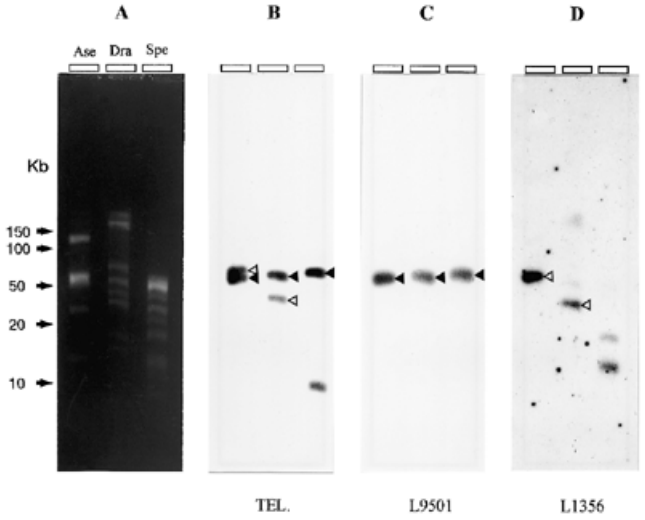

Figure 3. Positioning of the cosmid contig ends in relation to both telomeric ends of chromosome 5 in *L.major* 'Friedlin'. PFGE gel-purified chromosome 5 DNA was restricted with *Ase*I, *Dra*I and *Spe*I and the resulting fragments separated by PFGE (**A**). The same gel was Southern blotted and hybridised with a telomeric probe (**B**) and with radiolabelled DNA from the subtelomeric cosmid clones L9501 (**C**) and L1356 (**D**). Cosmid L9501 hybridised onto the ~50 kb *Ase*I, *Dra*I and *Spe*I fragments and was also recognised by the telomeric probe (closed arrowheads). These three sites are present on or close to the overlap of L9501 with L8138 (see Fig. 1). Cosmid L1356 hybridised onto the telomeric 60 kb *Ase*I and 25 kb *Dra*I fragments (open arrowheads), as well as onto two subtelomeric *Spe*I fragments of 12 and 18 kb, but not onto the telomeric 10 kb *Spe*I fragment. The two clones are therefore each located at ~20 kb from the telomeric ends of the chromosome (see Fig. 1).

their instability (10), as well as in possible chromatin structure modifications, as suggested by recent studies on *T.brucei* (24,25). The association of these two mechanisms might confer on the parasite a rapid adaptative response in gene expression to changes in external constraints.

With the constant increase in DNA sequence output foreseen in the frame of the *L.major* genome sequencing project, the results presented here on the conservation of the chromosomal organisation between two divergent pathogenic species of *Leishmania* should greatly help in the development of positional cloning. They should promote more studies in comparative genetics for this parasite with the objective of better understanding the species specificity of gene composition and expression.

ACKNOWLEDGEMENTS

We wish to thank Alasdair I. Ivens and Jenefer M. Blackwell for generously providing the complete *L.major* 'Friedlin' cosmid library as well as the gridded arrays on high density filters. This study received financial support from GDR 1077 of the CNRS/DGA-DSP and from the European Community INCO-DC Programme (contract no. 970256). P.D. was a grant recipient from the Ministère de l'Enseignement Supérieur et de la Recherche.

REFERENCES

1. Ashford,R.W., Desjeux,P. and de Raadt,P. (1992) *Parasitol. Today*, **8**, 104–105.
2. Ravel,C., Dubessay,P., Ivens,A., Blackwell,J.M., Bastien,P. and The *Leishmania* Genome Network (1998) *Parasitol. Today*, **14**, 301–303.
3. Wincker,P., Ravel,C., Blaineau,C., Pagès,M., Jauffret,Y., Dedet,J.P. and Bastien P. (1996) *Nucleic Acids Res.*, **24**, 1688–1694.
4. Britto,C., Ravel,C., Bastien,P., Blaineau,C., Pagès,M., Dedet,J.P. and Winker,P. (1998) *Gene*, **222**, 107–117.
5. Ravel,C., Macari,F., Bastien,P., Pagès,M. and Blaineau,C. (1995) *Mol. Biochem. Parasitol.*, **69**, 1–8.
6. Pagès,M., Bastien,P., Veas,F., Rossi,V., Bellis,M., Winker,P., Rioux,J.A. and Roizès,G. (1989) *Mol. Biochem. Parasitol.*, **36**, 161–168.
7. Blaineau,C., Bastien,P. and Pagès,M. (1992) *Mol. Biochem. Parasitol.*, **50**, 197–204.
8. Blaineau,C., Bastien,P., Rioux,J.A., Roizès,G. and Pagès,M. (1991) *Mol. Biochem. Parasitol.*, **46**, 292–302.
9. Ravel,C., Wincker,P., Blaineau,C., Britto,C., Bastien,P. and Pagès,M. (1996) *Genomics*, **35**, 509–516.
10. Ravel,C., Wincker,P., Bastien,P., Blaineau,C. and Pagès,M. (1995) *Mol. Biochem. Parasitol.*, **74**, 31–41.
11. Fu,G. and Barker,D.C. (1998) *Nucleic Acids Res.*, **26**, 2161–2167.
12. Myler,P.J., Audleman,L., deVos,T., Hixson,G., Kiser,P., Lemley,C., Magness,C., Rickel,E., Sisk,E., Sunkin,S., Swartzell,S., Westlake,T., Bastien,P., Fu,G., Ivens,A. and Stuart,K. *Proc. Natl Acad. Sci. USA*, **6**, 2902–2906.
13. Ivens,A.C., Lewis,S.M., Bagherzadeh,A., Zhang,L., Chan,H.M. and Smith,D.F. (1998) *Genome Res.*, **8**, 135–145.
14. Ivens,A.C. and Smith,D.F. (1997) *Trans. R. Soc. Trop. Med. Hyg.*, **91**, 111–115.
15. Beverley,S.M., Ismach,R.B. and McMahon-Pratt,D. (1987) *Proc. Natl Acad. Sci. USA*, **84**, 484–488.
16. Li,W.-H. (1993) *Curr. Opin. Genet. Dev.*, **3**, 896–901.
17. Janse,C.J., Carlton,J.M.R., Walliker,D. and Waters,A.P. (1994) *Mol. Biochem. Parasitol.*, **68**, 285–296.
18. Carlton,J.M., Vinkenoog,R., Waters,A.P. and Walliker,D. (1998) *Mol. Biochem. Parasitol.*, **93**, 285–294.
19. Lanzer,M., de Bruin,D. and Ravetch,J.V. (1993) *Nature*, **361**, 654–657.
20. Hernandez-Rivas,R., Hinterberg,K. and Scherf,A. (1996) *Mol. Biochem. Parasitol.*, **78**, 137–148.
21. Melville,S. (1998) *Parasitol. Today*, **14**, 129–130.
22. Hernandez-Rivas,R., Mattei,D., Sterkers,Y., Peterson,D.S., Wellems,T.E. and Scherf,A. (1997) *Mol. Cell. Biol.*, **17**, 604–611.
23. Pays,E. and Nolan,D.P. (1998) *Mol. Biochem. Parasitol.*, **91**, 3–36.
24. Rudenko,G., Blundell,P.A., Dirks-Mulder,A., Kieft,R. and Borst,P. (1995) *Cell*, **83**, 547–543.
25. Horn,D. and Cross,G.A. (1995) *Cell*, **83**, 555–561.