Infection, Genetics and Evolution 9 (2009) 290–297

Contents lists available at ScienceDirect

Infection, Genetics and Evolution

journal homepage: www.elsevier.com/locate/meegid

Discussion

Structure, dynamics and function of *Leishmania* genome: Resolving the puzzle of infection, genetics and evolution?

Jean-Claude Dujardin [a,b,*]

[a] Unit of Molecular Parasitology, Institute of Tropical Medicine, Nationalestraat, 155, B-2000 Antwerpen, Belgium
[b] Department of Biomedical Sciences, University of Antwerpen, Universiteitsplein 1, B-2610 Antwerpen, Belgium

ARTICLE INFO

Article history:
Received 28 July 2008
Received in revised form 25 November 2008
Accepted 25 November 2008
Available online 3 December 2008

Keywords:
Leishmania
Molecular karyotyping
Genome sequencing
Evolution
Gene amplification
Molecular epidemiology
Pathogenicity

ABSTRACT

We review and discuss here the specificity and contribution of genome (re-)arrangement studies for the exploration of genetic diversity among *Leishmania*. We show how the early molecular karyotyping studies generated an original perception of the genetics and evolution of these Protozoa, while providing some possible explanations on the parasite diverse phenotypes (drug resistance, pathogenicity...). We compare the results with the enormous amount of data provided by the recent genome sequencing projects, so far focused on one strain per genus/species. We highlight the relevance of parallel sequencing of different strains of a same species, now made possible by the new sequencing technologies. We recommend paying a particular attention to variation in gene copy number, a feature showed by the karyotyping studies to be extremely informative.

1. Introduction

Every day, health professionals are confronted with the clinical and epidemiological diversity of infectious diseases: facing f.i. patients suffering a more severe disease than others or not responding to the treatment. This diversity results from the interaction between several actors (host, pathogen, vector -if any-, environment...), but obviously, the longer evolutionary history of the pathogens points them as a central piece in this puzzle. Perception of their diversity is therefore essential and this is highly dependant of technological advances: for long, only the phenotype of pathogens could be analysed (f.i. morphology, susceptibility to antibiotics, isoenzyme pattern) and later on, the direct access to the genotype was made possible by the explosion of molecular biology. However, the power of molecular techniques should not mask the importance of evolutionary and genetic concepts. In this context, it is interesting to highlight that most genotyping studies focused so far on the nucleotide sequence of one or several genes, giving a major attention to point mutations as genetic characters. Now, genomes are not merely collections of genes (Danchin, 1998), and the map of the cell would be in the chromosome (Danchin and Hénaut, 1997). The karyotype (defined as the chromosomal complement of an organism) would be the product of a selection for gene arrangement, chromosomal size and number and centromere position, hereby constituting a mean to reach an adaptive phenotype (Mc Gregor, 1982). In other words, re-arrangement of genes could either induce or reflect evolutionary changes. This hypothesis proposed for Prokaryotes (Danchin and Hénaut, 1997) is documented in higher Eukaryotes too and, according to Wilson et al. (1974), gene re-arrangements would be more important than point mutations as sources for evolutionary changes. We propose to discuss here the specificity and contribution of genome (re-)arrangement studies for addressing the genetic diversity of lower Eukaryotes, taking as model *Leishmania* (Protozoa, Kinetoplastida, trypanosomatids). Therefore, we will first review the methodological aspects of the two main approaches which successively allowed exploring genome (re-)arrangements: molecular karyotyping and whole genome sequencing. In a second part, we will illustrate the contribution of both approaches to understand (i) the dynamics and evolution of genome, (ii) the genetics of the parasites and (iii) the potential link with parasites' phenotypes. In each application, we will follow an historical approach, explaining first the input of molecular karyotyping and second the contribution of genome sequencing, hereby highlighting the complementarity between both approaches (Table 1).

* Tel.: +32 3 2476355; fax: +32 3 2476359.
E-mail address: jcdujardin@itg.be.

1567-1348/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.meegid.2008.11.007

Table 1
Clinical and genomic features of *Leishmania* species mentioned in the review. CL, cutaneous leishmaniasis, VL, visceral leishmaniasis, PKDL, post kala-azar dermal leishmaniasis, DCL, diffuse cutaneous leishmaniasis, asympt, asymptomatic infections. Number of chromosomes, from Wincker et al. (1996) and Britto et al. (1998); n.d., not determined. Number of genes (pseudo-genes included), species-specific genes (from the 3-leish genome comparison) and genes shared between two species and absent in the third one (from the 3-leish genome comparison), from Smith et al. (2007).

Species	Clinical features	Chromosomes	Genome sequence	Genes	Species-specific genes	Shared genes
Subgenus *Leishmania*						
Old World						
L. major (Lmaj)	CL	36	Completed	8395	5	71 with Linf 5 with Lbra
L. tropica	CL, difficult to treat	36	n.d.	n.d.	n.d.	n.d.
L. infantum (Linf)	VL, CL, asympt	36	Completed	8195	27	71 with Lmaj 33 with Lbra
L. donovani	VL, PKDL, asympt, drug resistance	n.d.	Starting	n.d.	n.d.	n.d.
New World						
L. mexicana	CL, DCL	34 (fused 8–29, 20–36)	In progress	n.d.	n.d.	n.d.
Subgenus *Viannia*						
L. braziliensis (Lbra)	CL (severe), MCL	35 (fused 20–34)	Completed	8314	49	33 with Linf 5 with Lmaj
L. peruviana	CL (benign)	35 (fused 20–34)	n.d.	n.d.	n.d.	n.d.

2. Methods to explore genome structure and its variation

Protists, the primordial Eukaryotes, are characterised by different types of mitoses, classified according to the state of the membrane during mitosis, spindle symmetry...etc: trypanosomatids undergo a closed orthomitosis, as mitosis occurs within an intact nuclear membrane, with a symmetric spindle and a well-defined chromosomal plate (Raikov, 1994). However, study of trypanosomatids' karyotype is not possible with traditional cytogenetic methods, because chromosomal condensation is not detectable by photonic microscopy (Van der Ploeg et al., 1984). Therefore, new approaches were needed to explore genome structure and its variation.

2.1. Molecular karyotyping

The development of pulsed field gradient gel electrophoresis (PFGE, Schwartz et al., 1983) in the early 1980s represented a first and crucial step in the study of Trypanosomatid chromosomes. PFGE and derived methods (like orthogonal field alternated gel electrophoresis, OFAGE; contour clamped homogeneous electric field, CHEF; field inversion gel electrophoresis, FIGE; rotating field gel electrophoresis, RFGE) were applied on high molecular weight DNA (lysed and purified after embedding in agarose plugs) and allowed the resolution of whole chromosomes according to their size (up to a few Mb) or structure (see review by Maule, 1998). Literature analysis shows that OFAGE initially and CHEF later on were the methods of choice for most micro-organisms (Beadle et al., 2003). The resolutive power of these devices was – and is still – extensively used in the molecular epidemiology of bacteria (more than 40 pathogens or groups of pathogens, reviewed by Singh et al., 2006), but through the analysis of large restriction fragments of DNA: a kind of 'super' RFLP method. In contrast, real karyotype analysis seems less popular in molecular epidemiology, and trypanosomatids seem particularly neglected in comparison to fungi, for instance (the latter being reviewed by Beadle et al., 2003), even if a paper on the chromosomal organisation of *Phytomonas* sp. was recently published (Marín et al., 2008). Furthermore, in most of these reports, molecular karyotyping is used for strain identification (e.g. comparing if the general chromosome banding profile of strain A is similar to that of strain B, thus allowing to establish epidemiological links between the two) and papers reporting studies on quantified karyotype diversity (f.i. by phenetic methods) are – to my knowledge – quite rare. This is likely explained by the difficulty encountered by investigators for numerical processing of chromosome size-polymorphism. Indeed, the application of classical algorithms – that consider genetic variation as a discrete variable – would give the same weight to any size variation. However, the link between chromosomal size and copy number of repeats encountered in trypanosomatids suggested that chromosomal size variation was mostly progressive (Dujardin et al., 1995b). Thus, assuming that the more the size of two homologous chromosomes differs, the more they are genomically dissimilar, the absolute chromosomal size-difference index (aCSDI) was developed (Dujardin et al., 1995b). It was demonstrated in *Leishmania* and *T. cruzi* that karyotype variability was much better accounted for by aCSDI than by methods based on the model of discrete size variation (Dujardin et al., 1998; Henriksson et al., 2002). This algorithm (or any other one based on the same concept) is therefore essential to explore chromosome size variation in trypanosomatids.

2.2. Genome sequencing

The second crucial step for the exploration of genome structure was the development of strategies and technological platforms allowing the sequencing of whole genomes. Initially, two methods were applied: (i) the whole genome shotgun (WGS), in which the whole genome is sheared into small fragments, which are subsequently cloned, sequenced and assembled, and (ii) the clone-by-clone (CBC) approach, in which an intermediate step is introduced, i.e. the cloning of large inserts of genomic DNA, which serve as material for shearing DNA and are used themselves as units for assembling large pieces of genome (Worthey and Myler, 2005). The Tritryp (*L. major*, *T. brucei* and *T. cruzi*) genome project was initiated with the CBC approach. In the first two species, the approach was progressively replaced by a chromosome shotgun approach (using CHEF-purified chromosomal DNA, Blunt et al., 1997), while in *T. cruzi*, the CBC approach was abandoned and replaced by WGS, because of the high repeat content of the genome (>50%) and the extreme heterozygosity between diploid alleles (the chosen strain, CL Brener, showing to be a hybrid, Brisse et al., 1998). Obviously, the success of these projects did not only rely in the technological performance of large dedicated genome centres able to undertake high-throughput sequencing, but also to a huge bio-informatic support: different algorithms were applied or specifically developed for data mining of parasite genomes and gene annotation (reviewed by Worthey and Myler, 2005; Berri-

man, 2004). Recent developments in enzymology, imaging and microfluidics now offer new approaches to sequencing that could provide a massive increase in capacity while removing the need for the huge infrastructure required today (Hall, 2007). One of the main differences is that these new methods rely on '*in vitro*' amplification (like on special beads) of DNA prior to sequencing, while '*in vivo*' amplification (i.e. through cloning in a vector, with inherent biases against certain stretches of DNA that do not replicate well in the bacteria or code for toxic compounds) was used in previous methods (Hall, 2007). Some of these new methods, like the Solexa sequencing (Bennett et al., 2005) are able to sequence over 1 Gb of DNA in a single run and they should have a major impact in future studies, in particular in comparative genomics.

3. *L. major*, the first reference for *Leishmania* genome

L. major was the first species considered for exploration of the *Leishmania* genome. This choice was not always understood by health professionals, given the lower medical importance of that species vs. 'killers' like *L. donovani*, but it was justified by the fact that *L. major* constituted the best experimental model worldwide for several biological features. Early exploration of the molecular karyotype of *L. major* demonstrated the existence of 36 chromosome pairs with a size ranging between 0.28 and 2.8 Mb (Wincker et al., 1996). The completion of the genome sequencing project estimated the presence of 8311 genes and provided insight on the physical architecture of *L. major* chromosomes. This revealed an unusual pattern of gene distribution on all chromosomes, with clusters of genes present as a contiguous unit on one DNA strand with other similar units on the opposite strand (Myler et al., 1999; Smith et al., 2007). These clusters were shown to be transcribed as a single unit (Martínez-Calvillo et al., 2003) prior to trans-splicing and polyadenylation, being consistent with the polycistronic transcription model of kinetoplastids (Pays et al., 1994). These so called directional gene clusters (DGCs) range in size from a few to hundreds of genes stretching over 1 Mb of DNA (Smith et al., 2007) and they are separated by AT-rich strand-switch regions considered to contain sites for transcription initiation and termination (Martínez-Calvillo et al., 2004). DGCs do not contain clusters of genes of related function like in prokaryotic operons (Smith et al., 2007), but may contain tandem arrays of genes (like the rRNA, Inga et al., 1998) or multigene families (like the gp63 glycoprotein, Victoir et al., 2005). Chromosome ends of *L. major* are characterised by short and highly conserved subtelomeric regions bordering distinctive repeat sequences that precede heterogeneous telomeres (Ivens et al., 2005). These ends lack the large subtelomeric multigene families involved in host immune evasion, common to other parasites like *T. brucei* (Pays et al., 1994) or *Plasmodium* (Hall et al., 2002), but are not totally devoid of genes. Of particular interest are the phosphoglycan scβ-galactosyltransferase (SCB) genes involved in the side chain modification of the lipophosphoglycan (LPG). Their subtelomeric location could favour gene conversion events that may ultimately support the species diversity in LPG galactosylation, a feature playing a major role in parasite-vector specificity (Smith et al., 2007) with non-permissive sandfly species.

4. Dynamics and evolution of genome

The recent sequencing of other genomes brought precise information on the content and architecture of Trypanosomatid genomes, but was so far restricted to one strain of a few species only. In contrast, the early 'rough' analyses by molecular karyotyping were more superficial, but allowed the analysis of multiple strains and their diversity. Both approaches are complementary and might provide an original perception on the evolutionary history of trypanosomatids, and this at various taxonomical levels.

4.1. Genera

4.1.1. Molecular karyotyping

Comparison of the molecular karyotype of *L. major* to that of *T. brucei* and *T. cruzi* showed an overall architecture which was quite similar between *T. cruzi* (35 chromosomes with a size ranging between 0.4 and 3.3 Mb) and *L. major* (see above), while the karyotype of *T. brucei* was very different: presence of 11 large (1–5.2 Mb) diploid chromosomes together with an unspecified number of mini- (30 kb) and intermediate-sized (700 kb) chromosomes (Wickstead et al., 2004). Completion of the Tritryp genome project supported this higher similarity of chromosomal architecture between *L. major* and *T. cruzi* and revealed that the current chromosomal architecture of *T. brucei* seems to have derived from an ancestor with the more fragmented genomic organisation of *L. major* and *T. cruzi*: indeed, several chromosomes of *T. brucei* appeared to result from fusions of two chromosomes of *L. major* or *T. cruzi* (El-Sayed et al., 2005a), or reciprocally.

4.1.2. Genome sequencing

Detailed genome sequence analysis revealed a conserved core proteome of about 6200 genes in large syntenic polycistronic gene clusters (El-Sayed et al., 2005a). This might be considered as a remarkable degree of conservation for organisms that diverged an estimated 200–500 million years ago (Ghedin et al., 2004). Selection against re-arrangements within these conserved regions would be explained by a need of co-directionality for replication and transcription (Ghedin et al., 2004). However, 26% of the *T. brucei* proteome appeared to be species-specific (El-Sayed et al., 2005a), while 20% of the genome would be dedicated to antigenic variation (Kissinger, 2006); noteworthy, the published genome of that species does not include the mini- and intermediate-sized chromosomes and additional *T. brucei*-specific genes might therefore be discovered. A larger proportion of the *T. cruzi* proteome was species-specific (32%), while only 12% was specific to *L. major*. Evolutionary interpretation of Tritryp genome sequence data (El-Sayed et al., 2005a) indicates contrasting results when considering as evolutionary characters the point mutations or the major chromosomal re-arrangements. When considering the amino acid sequence of the genes (point mutations) shared by the three species, *T. cruzi* and *T. brucei* are much closer (57% identity) to each other than to *L. major* (44% identity). In contrast, when counting the number of genes shared by two species and absent in the third one (gene re-arrangement), *T. cruzi* would appear slightly closer to *L. major* (482 shared genes) than to *T. brucei* (458 shared genes); the small number of genes shared between *L. major* and *T. brucei* (74) would, in a simple parsimony evolutionary model, position *T. cruzi* in an intermediate position between the two other species.

The DGCs initially described in *L. major* were also encountered in the two other trypanosomatids. Interestingly, a large proportion of the synteny breakpoints occurred in the proximity of the strand-switch regions separating the DGCs (El-Sayed et al., 2005a). This was explained by a strong selective pressure to maintain gene order and keep the DGCs intact, hence highlighting their functional importance. In contrast, the subtelomeric architecture of *T. brucei* and *T. cruzi* appeared to be quite different from the short regions described in *L. major* and the other species: indeed, in both species, large blocks of non-syntenic genes were encountered at the telomeres of chromosomes. In *T. brucei* these blocks contained essentially the variable surface glycoproteins involved in the immune evasion of the parasite (Berriman et al., 2005), while in *T.*

cruzi, many surface molecules were also encountered in that location, like the transsialidase super-family (El-Sayed et al., 2005a).

4.2. Species

So far, we considered in this review one individual from one species as the representative of a given group of parasites. However, comparison of different species within a group is needed for evolutionary studies within this taxon, but also to revise the evolutionary links between them: the species closer to the common ancestor or *Leishmania* is indeed essential for genomic comparison with *T. cruzi* and *T. brucei*, and it is not sure that *L. major* is the most adequate for this.

4.2.1. Molecular karyotyping

At the end of the 1990s, the first complete physical maps of different *Leishmania* species were published (Wincker et al., 1996; Britto et al., 1998) and revealed that out of 36 chromosomes, 31 linkage groups were preserved across the whole genus, while 5 were the object of fusion/fission events. ACSDI analysis of the 31 «conserved» chromosomes evidenced a structuring into three clusters (Dujardin et al., 2000): (i) Old World representants of subgenus *Leishmania* (OWL), (ii) New World representants of subgenus *Leishmania* (NWL) and (iii) subgenus *Viannia* (NWV). Interestingly, NWL parasites constituted the most remote group and this corresponded to a significantly lower size for 11 chromosomes (accounting for about 1400 kb). This result was also supported by counting the minimal number of fusion/fission events among the five chromosomes not considered by aCSDI analysis. While one event was only separating NWV and OWL (fusion/fission of chromosomes 20 and 34), 2 (fusion/fission of chromosomes 8 and 29 and 20 and 36) and 3 events (sum of previous ones) did separate NWL from OWL and NWV, respectively (Britto et al., 1998).

4.2.2. Genome sequencing

This structuring of the genus *Leishmania* in three major groups according to the karyotype architecture could recently be further explored for two of them (OWL and NWV), after the sequencing of the whole genome of *L. major*, *L. infantum* (both OWL) and *L. braziliensis* (NWV). This revealed a background of conserved gene content, synteny and architecture and a few genes only that were specific to each species (5, 26 and 47, respectively; Peacock et al., 2007). This extent of synteny and similarity is greater than what was observed between human and rodent *Plasmodium* species (more than 1000 different genes; Kooij et al., 2005), despite a similar divergence time among the three *Leishmania* species (estimated between 14.6 and 46 million years; Lukes et al., 2007 and between 20 and 100 millions; Kerr, 2006) and *Plasmodium* ones (Peacock et al., 2007). However, this difference in magnitude of genomic differences could simply reflect the main propagation mode of both organisms: the clonality of trypanosomatids underlying 'reproduction' (in its etymological meaning) and the sexual recombination of *Plasmodium*, base of 'innovation'. Despite the overall genome conservation among the three species, a deeper analysis confirmed the outlier character of *L. (V.) braziliensis* in comparison with the two other *Leishmania* species: (i) lower average amino acid and nucleotide identities, (ii) presence of transposable elements and genes possibly involved in the RNAi pathway (see below), and (iii) differences in size of large tandem gene arrays (Peacock et al., 2007). When counting the number of genes shared by two species and absent in the third one, *L. major* and *L. infantum* are the closest species (71 shared genes), while *L. braziliensis* is closer to *L. infantum* (33 shared genes) than to *L. major* (5 shared genes only); interpreted in a simple parsimony evolutionary model, *L. infantum* would appear to occupy an intermediate position between the two other species. Completion of the sequencing of *L. mexicana* (a representative of NWL, in progress) should bring more light on the evolutionary relationships within genus *Leishmania*. According to the comparative analysis of the linkage maps mentioned above, we hypothesise that more genomic differences might be encountered.

4.3. Strains

4.3.1. Molecular karyotyping

Reports on the application of molecular karyotyping to multiple strains of trypanosomatids are scanty, despite of the particular adequacy of this method for micro-evolutionary studies within species. The first example concerns the study of natural populations of *L. braziliensis* and *L. peruviana*, two Peruvian species belonging to subgenus *Viannia* (NWV), phylogenetically very close, but endemic in different environments, respectively the Amazonian basin in the Eastern side of the Andes and the Western slopes of the Andes (Guerra, 1988). In order to better understand the evolutionary relationships between both parasites, an allopatric sampling was performed in their territory of endemism. Four chromosomes (out of 35 in subgenus *Viannia*, Britto et al., 1998) were selected (2, 10, 11, 27) for their significant size-polymorphism and analysed using the aCSDI method. Results showed a continuum of chromosomal size divergence between both species, centred on a specific geographical point, the only pass across the Andes between the Amazonian region and the Pacific coast, situated in the North of the country. This lead to the hypothesis that *L. peruviana* would descend from *L. braziliensis* and would have acquired its '*peruviana*' character during the southward colonisation and adaptation of the transmission cycle in the Andean valleys. A second example concerns the study of karyotype diversity in French populations of *L. infantum*, the main *Leishmania* species encountered in Southern Europe. Five chromosomes (2, 10, 12, 14, 27) were also studied in parasites originating from two distinct foci, the Pyrénées Orientales (close to the Spanish border) and the Cévennes (about 200 km North-East from the previous one). Like in the Peruvian example, a continuum of chromosomal size variation was observed between the two populations, with one chromosome (number 2) showing a size significantly correlated with geography. Interestingly, this chromosome was also involved in the geographical structuring observed in Peru, and in both models, smaller size variants were observed in the isolated mountain foci (Andes and Pyrénées). In both cases, size variation of chromosome 2 involved dosage of mini-exon genes (Kebede et al., 1999). Considering the key role of mini-exon genes in the translation control, it was hypothesised that a decrease in the amount of mini-exon transcripts (due to gene dosage or to alteration of regulatory inter-genic regions, McCoy et al., 1998), could have an influence on the general metabolism of the parasite and its growth behaviour (Iovannisci and Beverley, 1989). In both cases, environmental factors (climate, altitude, vector) could have selected parasites of different fitness.

4.3.2. Genome sequencing

So far, genome sequencing projects did not yet reach the level of strain analysis in trypanosomatids, but the recent technological developments should make it feasible in a close future, allowing high-throughput sequencing to be more than just a discovery exercise but also a routine assay for hypothesis testing (Hall, 2007). This is illustrated by the initial survey of genetic variation across the *P. falciparum* genome (Volkman et al., 2007): after the generation of high-quality draft genome sequences (7x genome coverage) of two parasite clones, 12 additional lines were sequenced at light genomic coverage (0.25x each) and submitted

to mutation screening. This revealed 46,937 SNPs in the analysed sample, demonstrating rich diversity among *P. falciparum* parasites and opening interesting avenues for understanding the population genetics of this organism. Most of this variation was caused by small regions of extremely high diversity associated with antigen-coding genes. Considering that the extent of synteny and similarity is greater between the genome of *L. major*, *L. infantum* and *L. braziliensis* than what was observed between human and rodent *Plasmodium* species (Peacock et al., 2007; Kooij et al., 2005), it is tempting to predict that the genome of different strains of one *Leishmania* species will be overall even more conserved. This does not exclude the possibility to discover interesting characters by this approach. SNPs constitute obviously a first panel of relevant information, but two additional features should be considered. First, a particular attention should be paid to the tandemly repeated genes, which were shown by the early molecular karyotype studies (i) to play a major role in the 'chromosome respiration' (expansion–contraction, through amplification–deletion among tandem arrays) of strains, (ii) to be evolutionary informative and (iii) to display size variation correlated with phenotypic differences. Repeated genes are not the easiest to sequence and assemble, and specific strategies, like the database of *T. cruzi* repeated genes (Arner et al., 2007) might need to be applied here. Second, we cannot exclude that sequencing of additional strains would not reveal a few additional genes. A study of group B *Streptococcus* strains revealed that, as each new strain was sequenced, an average of 33 new genes was discovered (Tettelin et al., 2005). This has led to the theory of the Pan-genome, which predicts that any bacterial species will be made up of a core set of genes that is found in all individuals and a dispensable set of genes that may or may not be present in any particular individual (Medini et al., 2005; Tettelin et al., 2005). Species delineation in bacteria and trypanosomatids (and even among the latter) is not necessarily comparable and their genetics is different. The amount of strain-specific genes in trypanosomatids (if any) will probably be low, but the Pan-genome theory should be kept in mind, amplified genes or isogenes (Victoir et al., 2005) maybe constituting one source of 'dispensable' genes in trypanosomatids.

5. Genetics of micro-organisms

5.1. Molecular karyotyping

Karyotype and gene (re-)arrangement studies also provided most relevant information on the genetics of trypanosomatids. A first original feature was observed in the early PFGE studies, especially thanks to the imperfect design of the home-made electrophoretic systems. Indeed, these set-ups often generated conditions of important inhomogeneity of electrical fields, in which some DNA molecules were found to migrate outside the track of linear chromosomes (Gajendran et al., 1987). These elements were shown to be circular DNAs and corresponded to episomal amplifications of short chromosome segments containing several genes. Further analysis revealed that the same regions could also be amplified in a linear form, leading to the appearance of multiple mini-chromosomes in some strains (Hamers et al., 1987). This type of chromosomal amplification might constitute an emergency solution allowing a higher expression of specific genes (see below). We remind here that gene expression is not regulated at initiation in trypanosomatids, but at post-transcription level (McDonagh et al., 2000). Accordingly, the only way for a parasite to up-regulate the expression of a given gene is to amplify it at genomic level (Victoir and Dujardin, 2002).

A second contribution of karyotype analysis to the comprehension of the genetics of trypanosomatids was the detection of hybrids in these organisms considered to reproduce essentially in a clonal way (Tibayrenc et al., 1990). Indeed, despite slight size fluctuation, the average size of some chromosomes is a feature specific of certain taxonomic groups. In that case, the finding of strains which after cloning show to possess sister chromosomes specific of each of these taxonomic groups, may suggest a hybridisation event; this hypothesis is further strengthened if these features may be observed for different chromosomes. This was observed for instance in the Eastern Andean valley of Huanuco (Peru), where at the junction of areas respectively endemic for *L. braziliensis* and *L. peruviana*, hybrids were encountered in 1991 for three chromosomes (Dujardin et al., 1995a). Interestingly, years later these hybrids were still abundant in that region (Nolder et al., 2007), demonstrating that such parasites may have a high fitness. Another well documented case of hybridisation concerns *T. cruzi*, in which lineages IId and IIe (the latter containing CL-Brener, the strain used for the *T. cruzi* genome project) showed for two and three chromosomes respectively hybrid features of other lineages (Brisse et al., 2003).

5.2. Genome sequencing

A third interesting feature of Trypanosomatid genetics was recently discovered through the Tritryp genome sequencing project and concerns the presence of mobile genetic elements, considered to be important in shaping the course of genome evolution. The haploid genome of *T. cruzi* and *T. brucei* was shown to possess 5% and 2%, respectively, of retroelements (transposons that require reverse transcription from an RNA intermediate), while no one was present in *L. major* (El-Sayed et al., 2005a). Surprisingly, potentially active retrotransposons were encountered in *L. braziliensis*, as well as a family of novel DNA transposable elements located in the telomeres of that species (Peacock et al., 2007). The specific role of these mobile elements and the reason behind their absence in *L. major* are still unknown.

Last but not least, the discovery of genes involved in RNAi deserves a particular attention. This mechanism, which was identified in *T. brucei* (Ngô et al., 1998) was not demonstrated in *L. major* and *T. cruzi* (Robinson and Beverley, 2003; El-Sayed et al., 2005b). Interestingly, one of the most unexpected differences between the genomes of the three *Leishmania* species was the identification of genes implicated in the RNAi pathway in *L. braziliensis*. This species-specificity of RNAi genes should be further explored in other lineages of *T. cruzi* and *T. brucei*. Could these genes belong to the Pan-genome of trypanosomatids, being retained by a few lineages only which would require them? It was hypothesised that *L. braziliensis* could have retained RNAi as an antiviral defense mechanism (Smith et al., 2007): indeed RNA viruses were often reported in that species as well as in other species of subgenus *Viannia* (Widmer and Dooley, 1995).

6. Infection

6.1. Molecular karyotyping

Karyotype studies and genome sequencing projects may also provide a better understanding of some phenotypic features of the parasites, like drug resistance or virulence. Several studies were made on the experimental induction of drug resistance in *Leishmania*, and this phenomenon was commonly reported to be associated with the appearance of circular episomes (Beverley, 1991) or linear molecules with telomeric sequences (Olmo et al., 1995). As mentioned earlier, these extra-chromosomal molecules originate from chromosome themselves, and one locus found to be frequently amplified in experimentally induced drug resistant strains of *Leishmania* is known as the H locus, a stretch of about 40 kb that is part of an 800-kb chromosome. 'H circles' were

generated de novo after induction with several unrelated drugs (Ouellette and Borst, 1991) and specific genes they contained were shown to be directly implicated in resistance: a P-glycoprotein-related genes is involved in arsenite resistance (Papadopoulou et al., 1994a) and a short-chain dehydrogenase reductase gene confers resistance to methotrexate, by reducing pterins (Papadopoulou et al., 1994b). This H locus seems to be quite conserved in different *Leishmania* species, but they do not necessarily favour extra-chromosomal gene amplification in response to drug resistance induction, as shown by a recent experimental study comparing *L. major* and *L. braziliensis* (Dias et al., 2007). This was confirmed in the few published reports on natural resistance among clinical isolates. H circles were reported among natural antimony-resistant Indian isolates of *L. donovani* (Mukherjee et al., 2007), while no episomal amplification was encountered in antimony-resistant Iranian isolates of *L. tropica* (Hadighi et al., 2007). Karyotype studies may thus be of (limited) use for the study of natural resistance, but considering the pleomorphic adaptive response of *Leishmania* to the drug (Decuypere et al., 2005), complementary avenues need to be explored. Obviously, whole genome comparison might play a major role in these new explorations, and this is well illustrated by the studies on the genetic diversity of *P. falciparum*. Regional variation in nucleotide diversity can be used to identify recent selective sweeps in a parasite population (Wootton et al., 2002) and using this approach, regions showing lower diversity in populations under selective pressure were identified, allowing the identification of genes specifically associated with drug resistance (Volkman et al., 2007).

With respect to the virulence and disease presentation, the molecular karyotype study of the *L. braziliensis/L. peruviana* and *L. infantum* models mentioned above also permitted to make interesting observations on a possible link between chromosomal size-polymorphism and these phenotypes. Indeed, *L. peruviana* is characterised by a lower pathogenicity than *L. braziliensis*, and the French *L. infantum* sample was characterised by strains isolates from visceral and cutaneous forms, respectively. Interestingly, in both models, clinical structuring was observed for some chromosomes and in both cases, the less pathogenic parasites presented less copies of rDNA genes (Inga et al., 1998; Guerbouj et al., 2001) associated to a lower copy number of genes encoding surface antigens, PSA-2 in *L. infantum* (Guerbouj et al., 2001) and gp63 in *L. (V.) peruviana* (Victoir et al., 1995). Like in the case of mini-exon genes mentioned above, a lower copy number of rDNA genes might have a general effect on the metabolism. This was observed in yeast (Maleszka and Clark-Walker, 1993) and *Drosophila* (Ashburner, 1989). With respect to PSA-2 and gp63, a more specific role would be expected. PSA-2 is a family of glycosil phospholipid anchored glycoproteins, which is expressed on the surface of both amastigote and promastigote forms of the parasite. It would contribute to macrophage attachment and survival to the non-specific complement attack (Murray and Spithill, 1991), and would induce T-cell (Sjölander et al., 1998) and B-cell (Jiménez-Ruiz et al., 1998) mediated immunity. Gp63 is a surface-expressed zinc metalloproteinase, involved in host cell binding and parasite protection from complement-mediated lysis (Yao et al., 2003).

6.2. Genome sequencing

Outputs of the 3-*Leishmania* genome project were systematically scrutinised for searching species-specific genes that might contribute to differences in disease presentation, immune response and pathogenicity (Peacock et al., 2007). Despite the difficulty of deriving conclusions from one strain per species (given the clinical pleomorphism existing within a species), this study is a première and it illustrates the power of comparative genomic analysis. As mentioned above, a few genes or tandem duplications, followed by diversification were found to be specific of each species. Also, despite of the difficulty for correct assembling of highly repetitive regions from randomly sequenced, gene dosage could be somehow assessed: f.i. the amastin gene array (largest family of surface-expressed protein genes in *Leishmania*) was twice larger in *L. major* than in *L. braziliensis*, while the gp63 gene array is 4 times larger in *L. braziliensis* than in the two other species. The authors concluded that a few species-specific parasite gene may be important in pathogenesis, that parasite gene expression levels differ considerably between species (perhaps as a consequence of variation in gene copy number) or that, contrary to the expectation, the parasite genome plays only a small part in determining clinical presentation. Comparison of these results with those observed by molecular karyotyping suggests that the genome probably plays a role, but that a particular attention should be paid on the consequences of gene dosage, specifically those encoding for surface proteins.

7. General conclusion

The examples shown in this review illustrate the power of karyotype and gene (re-)arrangement studies for resolving the puzzle of infection, genetics and evolution of trypanosomatids. Several observations were confirmed by the large genome projects but are restricted so far to a few individuals. A single genome may give a very poor representation of the genetic potential of the species (Hall, 2007), hence the current technological revolution allowing massive parallel DNA sequencing is of extreme interest, as it provides for the first time the possibility to study the genome architecture of a species instead of 1 reference genome of an individual. Parallel sequencing of different strains representative of the clinical and phenotypic diversity and subsequent comparison to reference genomes (just initiated for *L. donovani*) will reveal genome-wide sequence differences in pathogen populations that may explain phenotypic variation. In this context, molecular karyotyping might seem obsolete, but should not be consigned too rapidly in the oblivion. It is still a very democratic tool for fine molecular epidemiology studies (if using the adequate algorithms for numerical processing) and most of all, it might still be extremely useful for guiding comparative genome projects and target their efforts or adapt their strategies, e.g. to tandemly repeated genes.

Acknowledgements

Studies of the Molecular Parasitology Unit here mentioned were funded by EU (contracts TS2-CT90-0315, TS3-CT92-0129, IC18-CT98-0256, ICA4-CT-2001-10076); the parallel genome sequencing of *L. donovani* strains is funded by EU (contract 222895) and ITM (project GeMInI).

References

Arner, E., Kindlund, E., Nilsson, D., Farzana, F., Ferella, M., Tammi, M.T., Andersson, B., 2007. Database of *Trypanosoma cruzi* repeated genes: 20,000 additional gene variants. BMC Genomics 8, 391.

Ashburner, M., 1989. Drosophila. A laboratory handbook. Cold Spring Harbor Laboratory Press.

Beadle, J., Wright, M., McNeely, L., Bennett, J.W., 2003. Electrophoretic karyotype analysis in fungi. Adv. Appl. Microbiol. 53, 243–270.

Bennett, S.T., Barnes, C., Cox, A., Davies, L., Brown, C., 2005. Toward the 1,000 dollars human genome. Pharmacogenomics 6, 373–382.

Berriman, M., 2004. Data mining parasite genomes. Parasitology 128 (Suppl. 1), S23–31.

Berriman, M., Ghedin, E., Hertz-Fowler, C., Blandin, G., Renauld, H., Bartholomeu, D.C., Lennard, N.J., Caler, E., Hamlin, N.E., Haas, B., Böhme, U., Hannick, L., Aslett, M.A., Shallom, J., Marcello, L., Hou, L., Wickstead, B., Alsmark, U.C., Arrowsmith, C., Atkin, R.J., Barron, A.J., Bringaud, F., Brooks, K., Carrington, M., Cherevach, I., Chillingworth, T.J., Churcher, C., Clark, L.N., Corton, C.H., Cronin, A., Davies, R.M., Doggett, J., Djikeng, A., Feldblyum, T., Field, M.C., Fraser, A., Goodhead, I., Hance,

Z., Harper, D., Harris, B.R., Hauser, H., Hostetler, J., Ivens, A., Jagels, K., Johnson, D., Johnson, J., Jones, K., Kerhornou, A.X., Koo, H., Larke, N., Landfear, S., Larkin, C., Leech, V., Line, A., Lord, A., Macleod, A., Mooney, P.J., Moule, S., Martin, D.M., Morgan, G.W., Mungall, K., Norbertczak, H., Ormond, D., Pai, G., Peacock, C.S., Peterson, J., Quail, M.A., Rabbinowitsch, E., Rajandream, M.A., Reitter, C., Salzberg, S.L., Sanders, M., Schobel, S., Sharp, S., Simmonds, M., Simpson, A.J., Tallon, L., Turner, C.M., Tait, A., Tivey, A.R., Van Aken, S., Walker, D., Wanless, D., Wang, S., White, B., White, O., Whitehead, S., Woodward, J., Wortman, J., Adams, M.D., Embley, T.M., Gull, K., Ullu, E., Barry, J.D., Fairlamb, A.H., Opperdoes, F., Barrell, B.G., Donelson, J.E., Hall, N., Fraser, C.M., Melville, S.E., El-Sayed, N.M., 2005. The genome of the African trypanosome *Trypanosoma brucei*. Science 309, 416–422.

Beverley, S.M., 1991. Gene amplification in *Leishmania*. Ann. Rev. Microbiol. 45, 417–444.

Blunt, D.S., Khramtsov, N.V., Upton, S.J., Montelone, B.A., 1997. Molecular karyotype analysis of *Cryptosporidium parvum*: evidence for eight chromosomes and a low-molecular-size molecule. Clin. Diagn. Lab. Immunol. 4, 11–13.

Brisse, S., Barnabé, C., Bañuls, A.L., Sidibé, I., Noël, S., Tibayrenc, M., 1998. A phylogenetic analysis of the *Trypanosoma cruzi* genome project CL Brener reference strain by multilocus enzyme electrophoresis and multiprimer random amplified polymorphic DNA fingerprinting. Mol. Biochem. Parasitol. 92, 253–263.

Brisse, S., Henriksson, J., Barnabé, C., Douzery, E.J.P., Berkvens, D., Serrano, M., De Carvalho, M.R.C., Buck, G.A., Dujardin, J.C., Tibayrenc, M., 2003. Evidence for genetic exchange and hybridization in *Trypanosoma cruzi*. Inf. Gen. Evol. 2, 173–183.

Britto, C., Ravel, C., Bastien, P., Blaineau, C., Pagès, M., Dedet, J.P., Wincker, P., 1998. Conserved linkage groups associated with large-scale chromosomal rearrangements between Old World and New World *Leishmania* genomes. Gene 222, 107–117.

Danchin, A., 1998. The Delphic boat or what the genomic text tell us. Bioinformatics 14, 383.

Danchin, A., Hénaut, A., 1997. The map of the cell is in the chromosome. Curr. Opin. Genet. Dev. 7, 852–854.

Decuypere, S., Yardley, V., De Doncker, S., Laurent, T., Rijal, S., Chappuis, F., Dujardin, J.C., 2005. A study of the mechanism of natural Sb(V) resistance in Nepalese *Leishmania donovani* isolates based on gene expression analysis. Antimicrob. Agents Chemother. 49, 4616–4621.

Dias, F.C., Ruiz, J.C., Lopes, W.C., Squina, F.M., Renzi, A., Cruz, A.K., Tosi, L.R., 2007. Organization of H locus conserved repeats in *Leishmania (Viannia) braziliensis* correlates with lack of gene amplification and drug resistance. Parasitol. Res. 101, 667–676.

Dujardin, J.C., Bañuls, A.L., Llanos-Cuentas, A., Alvarez, E., De Doncker, S., Jacquet, D., Le Ray, D., Arevalo, J., Tibayrenc, M., 1995a. Putative *Leishmania* hybrids in the Eastern Andean valley of Huanuco, Peru. Acta Trop. 59, 293–307.

Dujardin, J.C., Dujardin, J.P., Tibayrenc, M., Timperman, G., De Doncker, S., Jacquet, D., Arevalo, J., Llanos-Cuentas, A., Guerra, H., Bermudez, H., Hamers, R., Le Ray, D., 1995b. Karyotype plasticity in Neotropical *Leishmania*: an index for measuring genomic distance among *L.(V.)peruviana* and *L.(V.)braziliensis* populations. Parasitology 110, 21–30.

Dujardin, J.C., Henriksson, J., Victoir, K., Brisse, S., Gamboa, D., Arevalo, J., Le Ray, D., 2000. Genomic rearrangements in trypanosomatids: an alternative to the "one gene" evolutionary hypotheses? Mem. Instit. Oswaldo Cruz 95, 527–534.

Dujardin, J.C., Bañuls, A.L., Arevalo, J., Tibayrenc, M., Le Ray, D., 1998. Comparison of chromosomal and isoenzymatic variation in eco-geographical populations of *Leishmania (Viannia) peruviana*. Parasitology 117, 547–554.

El-Sayed, N.M., Myler, P.J., Blandin, G., Berriman, M., Crabtree, J., Aggarwal, G., Caler, E., Renauld, H., Worthey, E.A., Hertz-Fowler, C., Ghedin, E., Peacock, C., Bartholomeu, D.C., Haas, B.J., Tran, A.N., Wortman, J.R., Alsmark, U.C., Angiuoli, S., Anupama, A., Badger, J., Bringaud, F., Cadag, E., Carlton, J.M., Cerqueira, G.C., Creasy, T., Delcher, A.L., Djikeng, A., Embley, T.M., Hauser, C., Ivens, A.C., Kummerfeld, S.K., Pereira-Leal, J.B., Nilsson, D., Peterson, J., Salzberg, S.L., Shallom, J., Silva, J.C., Sundaram, J., Westenberger, S., White, O., Melville, S.E., Donelson, J.E., Andersson, B., Stuart, K.D., Hall, N., 2005a. Comparative genomics of trypanosomatid parasitic protozoa. Science 309, 404–409.

El-Sayed, N.M., Myler, P.J., Bartholomeu, D.C., Nilsson, D., Aggarwal, G., Tran, A.N., Ghedin, E., Worthey, E.A., Delcher, A.L., Blandin, G., Westenberger, S.J., Caler, E., Cerqueira, G.C., Branche, C., Haas, B., Anupama, A., Arner, E., Aslund, L., Attipoe, P., Bontempi, E., Bringaud, F., Burton, P., Cadag, E., Campbell, D.A., Carrington, M., Crabtree, J., Darban, H., da Silveira, J.F., de Jong, P., Edwards, K., Englund, P.T., Fazelina, G., Feldblyum, T., Ferella, M., Frasch, A.C., Gull, K., Horn, D., Hou, L., Huang, Y., Kindlund, E., Klingbeil, M., Kluge, S., Koo, H., Lacerda, D., Levin, M.J., Lorenzi, H., Louie, T., Machado, C.R., McCulloch, R., McKenna, A., Mizuno, Y., Mottram, J.C., Nelson, S., Ochaya, S., Osoegawa, K., Pai, G., Parsons, M., Pentony, M., Pettersson, U., Pop, M., Ramirez, J.L., Rinta, J., Robertson, L., Salzberg, S.L., Sanchez, D.O., Seyler, A., Sharma, R., Shetty, J., Simpson, A.J., Sisk, E., Tammi, M.T., Tarleton, R., Teixeira, S., Van Aken, S., Vogt, C., Ward, P.N., Wickstead, B., Wortman, J., White, O., Fraser, C.M., Stuart, K.D., Andersson, B., 2005b. The genome sequence of *Trypanosoma cruzi*, etiologic agent of Chagas disease. Science 309, 409–415.

Gajendran, N., Dujardin, J.C., Le Ray, D., Matthyssens, G., Muyldermans, S., Hamers, R., 1987. Abnormally migrating chromosomes identify *Leishmania donovani* strains. In: Hart, D. (Ed.), Leishmaniasis: The First Centenary (1885–1985). New strategies for Control NATO ASI Series A, 163. Plenum Press, New York, pp. 539–547.

Ghedin, E., Bringaud, F., Peterson, J., Myler, P., Berriman, M., Ivens, A., Andersson, B., Bontempi, E., Eisen, J., Angiuoli, S., Wanless, D., Von Arx, A., Murphy, L., Lennard, N., Salzberg, S., Adams, M.D., White, O., Hall, N., Stuart, K., Fraser, C.M., El-Sayed, N.M., 2004. Gene synteny and evolution of genome architecture in trypanosomatids. Mol. Biochem. Parasitol. 134, 183–191.

Guerbouj, S., Guizani, I., Speybroek, N., Le Ray, D., Dujardin, J.C., 2001. Genomic polymorphism of *Leishmania infantum*: a relationship with clinical pleomorphism? Infect. Gen. Evol. 1, 49–59.

Guerra, H., 1988. Distribution of *Leishmania* in Peru. In: Walton, Wijeyaratne, Modabber, (Eds.), Research on Control Strategies for the Leishmaniases. IDRC–MR 184e, Ottawa, pp. 135–145.

Hadighi, R., Mohebali, M., Boucher, P., Hajjaran, H., Khamesipour, A., Ouellette, M., 2007. Unresponsiveness to glucantime treatment in Iranian cutaneous leishmaniasis due to drug-resistant *Leishmania tropica* parasites. PLOS Med. 3 (5), e162.

Hall, N., Pain, A., Berriman, M., Churcher, C., Harris, B., Harris, D., Mungall, K., Bowman, S., Atkin, R., Baker, S., Barron, A., Brooks, K., Buckee, C.O., Burrows, C., Cherevach, I., Chillingworth, C., Chillingworth, T., Christodoulou, Z., Clark, L., Clark, R., Corton, C., Cronin, A., Davies, R., Davis, P., Dear, P., Dearden, F., Doggett, J., Feltwell, T., Goble, A., Goodhead, I., Gwilliam, R., Hamlin, N., Hance, Z., Harper, D., Hauser, H., Hornsby, T., Holroyd, S., Horrocks, P., Humphray, S., Jagels, K., James, K.D., Johnson, D., Kerhornou, A., Knights, A., Konfortov, B., Kyes, S., Larke, N., Lawson, D., Lennard, N., Line, A., Maddison, M., McLean, J., Mooney, P., Moule, S., Murphy, L., Oliver, K., Ormond, D., Price, C., Quail, M.A., Rabbinowitsch, E., Rajandream, M.A., Rutter, S., Rutherford, K.M., Sanders, M., Simmonds, M., Seeger, K., Sharp, S., Smith, R., Squares, R., Squares, S., Stevens, K., Taylor, K., Tivey, A., Unwin, L., Whitehead, S., Woodward, J., Sulston, J.E., Craig, A., Newbold, C., Barrell, B.G., 2002. Sequence of *Plasmodium falciparum* chromosomes 1, 3–9 and 13. Nature 419, 527–531.

Hall, N., 2007. Advanced sequencing technologies and their wider impact in microbiology. J. Exp. Biol. 209, 1518–1525.

Hamers, R., Gajendran, N., Dujardin, J.C., Stuart, K., 1987. Circular and linear forms of small nucleic acids in *Leishmania*. In: Hart, D. (Ed.), Leishmaniasis: The First Centenary (1885–1985). New strategies for Control, NATO ASI Series A, vol. 163. Plenum Press, New York, pp. 985–988.

Henriksson, J., Dujardin, J.C., Barnabe, C., Timperman, G., Venegas, J., Pettersson, U., Tibayrenc, M., Solari, A., 2002. Chromosomal size variation in *Trypanosoma cruzi* is mainly progressive and is evolutionary informative. Parasitology 124, 277–286.

Inga, R., De Doncker, S., Gomez, J., Lopez, M., Garcia, R., Le Ray, D., Arevalo, J., Dujardin, J.C., 1998. Relation between variation in copy number of ribosomal RNA encoding genes and size of harbouring chromosomes in *Leishmania* of subgenus *Viannia*. Mol. Biochem. Parasitol. 92, 219–228.

Iovannisci, D.M., Beverley, S.M., 1989. Structural alterations of chromosome 2 in *Leishmania major* as evidence for diploidy, including spontaneous amplification of the mini-exon array. Mol. Biochem. Parasitol. 34, 177–188.

Ivens, A.C., Peacock, C.S., Worthey, E.A., Murphy, L., Aggarwal, G., Berriman, M., Sisk, E., Rajandream, M.A., Adlem, E., Aert, R., Anupama, A., Apostolou, Z., Attipoe, P., Bason, N., Bauser, C., Beck, A., Beverley, S.M., Bianchettin, G., Borzym, K., Bothe, G., Bruschi, C.V., Collins, M., Cadag, E., Ciarloni, L., Clayton, C., Coulson, R.M., Cronin, A., Cruz, A.K., Davies, R.M., De Gaudenzi, J., Dobson, D.E., Duesterhoeft, A., Fazelina, G., Fosker, N., Frasch, A.C., Fraser, A., Fuchs, M., Gabel, C., Goble, A., Goffeau, A., Harris, D., Hertz-Fowler, C., Hilbert, H., Horn, D., Huang, Y., Klages, S., Knights, A., Kube, M., Larke, N., Litvin, L., Lord, A., Louie, T., Marra, M., Masuy, D., Matthews, K., Michaeli, S., Mottram, J.C., Müller-Auer, S., Munden, H., Nelson, S., Norbertczak, H., Oliver, K., O'neil, S., Pentony, M., Pohl, T.M., Price, C., Purnelle, B., Quail, M.A., Rabbinowitsch, E., Reinhardt, R., Rieger, M., Rinta, J., Robben, J., Robertson, L., Ruiz, J.C., Rutter, S., Saunders, D., Schäfer, M., Schein, J., Schwartz, D.C., Seeger, K., Seyler, A., Sharp, S., Shin, H., Sivam, D., Squares, R., Squares, S., Tosato, V., Vogt, C., Volckaert, G., Wambutt, R., Warren, T., Wedler, H., Woodward, J., Zhou, S., Zimmermann, W., Smith, D.F., Blackwell, J.M., Stuart, K.D., Barrell, B., Myler, P.J., 2005. The genome of the kinetoplastid parasite, *Leishmania major*. Science 309, 436–442.

Jiménez-Ruiz, A., Boceta, C., Bonay, P., Requena, J.M., Alonso, C., 1998. Cloning, sequencing, and expression of the PSA genes from *Leishmania infantum*. Eur. J. Biochem. 251, 389–397.

Kebede, A., De Doncker, S., Arevalo, J., Le Ray, D., Dujardin, J.C., 1999. *Leishmania* of subgenus *Viannia*: size polymorphism of chromosomes bearing mini-exon genes in natural populations is due to rearrangement of the mini-exon gene array. Int. J. Parasitol. 29, 549–557.

Kerr, S.F., 2006. Molecular trees of trypanosomes incongruent with fossil records of hosts. Mem. Inst. Oswaldo Cruz 101, 25–30.

Kissinger, J.C., 2006. A tale of three genomes: the kinetoplastids have arrived. Trends Parasitol. 22, 240–243.

Kooij, T.W., Carlton, J.M., Bidwell, S.L., Hall, N., Ramesar, J., Janse, C.J., Waters, A.P., 2005. A *Plasmodium* whole-genome synteny map: indels and synteny breakpoints as foci for species-specific genes. PLoS Pathog. 1, e44.

Lukes, J., Mauricio, I., Schonian, G., Dujardin, J.C., Soteriadou, K., Dedet, J.P., Kuhls, K., Quispe Tintaya, W., Jirků, M., Zemanová, E., Haralambous, C., Oborník, M., Horák, A., Ayala, F., Miles, M.A., 2007. Evolutionary and geographical history of the *Leishmania donovani* complex with a revision of current taxonomy. PNAS USA 104, 9375–9380.

Maleszka, R., Clark-Walker, G.D., 1993. Yeasts have a four-fold variation in ribosomal DNA copy number. Yeast 9, 53–58.

Marín, C., Alberge, B., Dollet, M., Pagès, M., Bastien, P., 2008. First complete chromosomal organization of a protozoan plant parasite (Phytomonas spp.). Genomics 91, 88–93.

Martínez-Calvillo, S., Yan, S., Nguyen, D., Fox, M., Stuart, K., Myler, P.J., 2003. Transcription of *Leishmania major* Friedlin chromosome 1 initiates in both directions within a single region. Mol. Cell. 11, 1291–1299.

Martínez-Calvillo, S., Nguyen, D., Stuart, K., Myler, P.J., 2004. Transcription initiation and termination on *Leishmania major* chromosome 3. Eukaryot. Cell 3, 506–517.
Maule, J., 1998. Pulsed-field gel electrophoresis. Mol. Biotechnol. 9, 107–126.
McCoy, J.J., Beetham, J.K., Ochs, D.E., Donelson, J.E., Wilson, M.E., 1998. Regulatory sequences and a novel gene in the msp (GP63) gene cluster of *Leishmania chagasi*. Mol. Biochem. Parasitol. 95, 251–265.
McDonagh, P.D., Myler, P.J., Stuart, K., 2000. The unusual gene organization of *Leishmania major* chromosome 1 may reflect novel transcription processes. Nucleic Acids Res. 28, 2800–2803.
Mc Gregor, H.C., 1982. Big chromosomes and speciation among Amphibia. In: Dover, G.A., Flavell, R.B. (Eds.), Genome Evolution. Academic Press, pp. 325–341.
Medini, D., Donati, C., Tettelin, H., Masignani, V., Rappuoli, R., 2005. The microbial pan-genome. Curr. Opin. Genet. Dev. 15, 589–594.
Mukherjee, A., Padmanabhan, P.K., Singh, S., Roy, G., Girard, I., Chatterjee, M., Ouellette, M., Madhubala, R., 2007. Role of ABC transporter MRPA, γ-glutamylcysteine synthetase and ornithine decarboxylase in natural antimony-resistant isolates of *Leishmania donovani*. J. Antimicrob. Chemother. 59, 204–211.
Murray, P.J., Spithill, T.W., 1991. Variants of a *Leishmania* surface antigen derived from a multigenic family. J. Biol. Chem. 266, 24477–24484.
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., Stuart, K., 1999. *Leishmania major* Friedlin chromosome 1 has an unusual distribution of protein-coding genes. Proc. Natl. Acad. Sci. U.S.A. 96, 2902–2906.
Ngô, H., Tschudi, C., Gull, K., Ullu, E., 1998. Double-stranded RNA induces mRNA degradation in *Trypanosoma brucei*. Proc. Natl. Acad. Sci. U.S.A. 95, 14687–14692.
Nolder, D., Roncal, N., Davies, C.R., Llanos-Cuentas, A., Miles, M.A., 2007. Multiple hybrid genotypes of *Leishmania (viannia)* in a focus of mucocutaneous Leishmaniasis. Am. J. Trop. Med. Hyg. 76, 573–578.
Olmo, A., Arrebola, R., Bernier, V., González-Pacanowska, D., Ruiz-Pérez, L.M., 1995. Co-existence of circular and multiple linear amplicons in methotrexate-resistant *Leishmania*. Nucleic Acids Res. 23, 2856–2864.
Ouellette, M., Borst, P., 1991. Drug resistance and P-glycoprotein gene amplification in the protozoan parasite *Leishmania*. Res. Microbiol. 142, 737–746.
Pays, E., Vanhamme, L., Berberof, M., 1994. Genetic controls for the expression of surface antigens in African trypanosomes. Annu. Rev. Microbiol. 48, 25–52.
Papadopoulou, B., Roy, G., Dey, S., Rosen, B.P., Ouellette, M., 1994a. Contribution of the *Leishmania* P-glycoprotein-related gene ltpgpA to oxyanion resistance. J. Biol. Chem. 269, 11980–11986.
Papadopoulou, B., Roy, G., Mourad, W., Leblanc, E., Ouellette, M., 1994b. Changes in folate and pterin metabolism after disruption of the *Leishmania* H locus short chain dehydrogenase gene. J. Biol. Chem. 269, 7310–7315.
Peacock, C.S., Seeger, K., Harris, D., Murphy, L., Ruiz, J.C., Quail, M.A., Peters, N., Adlem, E., Tivey, A., Aslett, M., Kerhornou, A., Ivens, A., Fraser, A., Rajandream, M.A., Carver, T., Norbertczak, H., Chillingworth, T., Hance, Z., Jagels, K., Moule, S., Ormond, D., Rutter, S., Squares, R., Whitehead, S., Rabbinowitsch, E., Arrowsmith, C., White, B., Thurston, S., Bringaud, F., Baldauf, S.L., Faulconbridge, A., Jeffares, D., Depledge, D.P., Oyola, S.O., Hilley, J.D., Brito, L.O., Tosi, L.R., Barrell, B., Cruz, A.K., Mottram, J.C., Smith, D.F., Berriman, M., 2007. Comparative genomic analysis of three *Leishmania* species that cause diverse human disease. Nat. Genet. 39, 839–847.
Raikov, I.B., 1994. The diversity of forms of mitosis in Protozoa: a comparative review. Eur. J. Protistol. 30, 253–269.
Robinson, K.A., Beverley, S.M., 2003. Improvements in transfection efficiency and tests of RNA interference (RNAi) approaches in the protozoan parasite *Leishmania*. Mol. Biochem. Parasitol. 128, 217–228.
Schwartz, D.C., Safran, N., Welsh, J., Haas, R., Goldenberg, M., Cantor, C.R., 1983. New techniques for purifying large DNAs and studying their properties and packaging. Cold Spring Harb. Symp. Quant. Biol. 47, 189–195.
Singh, A., Goering, R.V., Simjee, S., Foley, S.L., Zervos, M.J., 2006. Application of molecular techniques to the study of hospital infection. Clin. Microbiol. Rev. 19, 512–530.
Sjölander, A., Baldwin, T.M., Curtis, J.M., Bengtsson, K.L., Handman, E., 1998. Vaccination with recombinant Parasite Surface Antigen 2 from *Leishmania major* induces a Th1 type of immune response but does not protect against infection. Vaccine 16, 2077–2084.
Smith, D.F., Peacock, C.S., Cruz, A.K., 2007. Comparative genomics: from genotype to disease phenotype in the leishmaniases. Int. J. Parasitol. 37, 1173–1186.
Tettelin, H., Masignani, V., Cieslewicz, M.J., Donati, C., Medini, D., Ward, N.L., Angiuoli, S.V., Crabtree, J., Jones, A.L., Durkin, A.S., Deboy, R.T., Davidsen, T.M., Mora, M., Scarselli, M., Margarit y Ros, I., Peterson, J.D., Hauser, C.R., Sundaram, J.P., Nelson, W.C., Madupu, R., Brinkac, L.M., Dodson, R.J., Rosovitz, M.J., Sullivan, S.A., Daugherty, S.C., Haft, D.H., Selengut, J., Gwinn, M.L., Zhou, L., Zafar, N., Khouri, H., Radune, D., Dimitrov, G., Watkins, K., O'Connor, K.J., Smith, S., Utterback, T.R., White, O., Rubens, C.E., Grandi, G., Madoff, L.C., Kasper, D.L., Telford, J.L., Wessels, M.R., Rappuoli, R., Fraser, C.M., 2005. Genome analysis of multiple pathogenic isolates of *Streptococcus agalactiae*: implications for the microbial "pan-genome". Proc. Natl. Acad. Sci. U.S.A. 102, 13950–13955.
Tibayrenc, M., Kjelleberg, F., Ayala, F.J., 1990. A clonal theory of parasitic protozoa: the population structures of *Entamoeba, Giardia, Leishmania, Naegleria, Plasmodium, Trichomonas* and *Trypanosoma* and their medical and taxonomical consequences. Proc. Natl. Acad. Sci. U.S.A. 87, 2414–2418.
Van der Ploeg, L.H.T., Schwartz, D.C., Cantor, C.R., Borst, P., 1984. Antigenic variation in *Trypanosoma brucei* analyzed by electrophoretic separation of chromosome-sized molecules. Cell 37, 77–84.
Victoir, K., Dujardin, J.C., De Doncker, S., Barker, D.C., Arevalo, J., Hamers, R., Le Ray, D., 1995. Plasticity of gp63 gene organization in *Leishmania (Viannia) braziliensis* and *Leishmania (Viannia) peruviana*. Parasitology 111, 265–273.
Victoir, K., Arevalo, J., De Doncker, S., Barker, D.C., Laurent, T., Godfroid, E., Bollen, A., Le Ray, D., Dujardin, J.C., 2005. Complexity of the major surface protease (msp) gene organisation in *Leishmania (Viannia) braziliensis*: evolutionary and functional implications. Parasitology 131, 207–214.
Victoir, K., Dujardin, J.C., 2002. How to succeed in parasitic life without sex? Asking *Leishmania*. Trends Parasitol. 18, 81–85.
Volkman, S.K., Sabeti, P.C., DeCaprio, D., Neafsey, D.E., Schaffner, S.F., Milner, D.A., Daily, J.P., Sarr, O., Ndiaye, D., Ndir, O., Mboup, S., Duraisingh, M.T., Lukens, A., Derr, A., Stange-Thomann, N., Waggoner, S., Onofrio, R., Ziaugra, L., Mauceli, E., Gnerre, S., Jaffe, D.B., Zainoun, J., Wiegand, R.C., Birren, B.W., Hartl, D.L., Galagan, J.E., Lander, E.S., Wirth, D.F., 2007. A genome-wide map of diversity in *Plasmodium falciparum*. Nature genetics 39, 113–119.
Wickstead, B., Ersfeld, K., Gull, K., 2004. The small chromosomes of *Trypanosoma brucei* involved in antigenic variation are constructed around repetitive palindromes. Genome Res. 14, 1014–1024.
Widmer, G., Dooley, S., 1995. Phylogenetic analysis of *Leishmania* RNA virus and Leishmania suggests ancient virus-parasite association. Nucleic Acids Res. 23, 2300–2304.
Wilson, A.C., Sarich, V.M., Maxson, L.R., 1974. The importance of gene rearrangement in evolution: evidence from studies on rates of chromosomal, protein and anatomical evolution. Proc. Natl. Acad. Sci. U.S.A. 71, 3028–3030.
Wincker, P., Ravel, C., Blaineau, C., Pages, M., Jauffret, Y., Dedet, J.P., Bastien, P., 1996. The *Leishmania* genome comprises 36 chromosomes conserved across widely divergent human pathogenic species. Nucleic Acids Res. 24, 1688–1694.
Wootton, J.C., Feng, X., Ferdig, M.T., Cooper, R.A., Mu, J., Baruch, D.I., Magill, A.J., Su, X.Z., 2002. Genetic diversity and chloroquine selective sweeps in *Plasmodium falciparum*. Nature 418, 320–323.
Worthey, E.A., Myler, P.J., 2005. Protozoan genomes: gene identification and annotation. Int. J. Parasitol. 35, 495–512.
Yao, C., Donelson, J., Wilson, M., 2003. The major surface protease (MSP or gp63) of *Leishmania* sp. Biosynthesis, regulation of expression, and function. Mol. Biochem. Parasitol. 132, 1–16.