Chapter 5 from:

Protozoan Parasitism

From Omics to Prevention and Control

Luis Miguel de Pablos Torró

ISBN: 978-1-910190-83-8 (paperback)
ISBN: 978-1-910190-84-5 (ebook)

Caister Academic Press www.caister.com

Omics Approaches for Understanding Gene Expression in *Leishmania*: Clues for Tackling Leishmaniasis

5

José M Requena[1]*, Pedro J Alcolea[2], Ana Alonso[2] and Vicente Larraga[2]

[1]Centro de Biologia Molecular Severo Ochoa (CSIC-UAM), Universidad Autonoma de Madrid, Madrid, Spain.
[2]Laboratorio de Parasitologia Molecular, Centro de Investigaciones Biologicas (CSIC), Madrid, Spain.

Correspondence: jmrequena@cbm.uam.es

https://doi.org/10.21775/9781910190838.05

Abstract

Leishmaniases, a group of parasitic diseases caused by species of the genus *Leishmania*, afflict millions of people across the globe and cause significant morbidity and mortality. Unfortunately, vaccine and chemotherapy options are limited. The advances in whole genome sequencing have led to renewed impetus in identifying druggable targets for future development of more effective treatments. Hence, the last decade has witnessed a revolution in our understanding of the *Leishmania* genomes through the completion of an increasing number of genome sequencing projects for several species and strains. However, the completion of a genome sequence is not the final product, rather it is just the beginning towards the objective of linking the wealth of data encoded in millions of bases to the biological processes of an organism. Moreover, the genome features (genomics) is only a part of the problem to be solved: the genome yields on transcription, the transcriptome, which in turn yields the proteome on translation, and ultimately the proteins either produce metabolites or are modulated by them. The size of genomics, transcriptomics, proteomics and metabolomics datasets has impelled a new way of analysing data together with the development of potent bioinformatics tools. During their life cycles, *Leishmania* parasites undergo significant changes in their morphology and metabolism. These changes clearly demand a developmental regulation of differential gene expression. Moreover, it is now becoming clear that epigenetic control can also regulate other aspects of the parasitic life cycle, including the control of the switch from proliferative to developmental programs, and the adaptations required for host and cellular tropisms.

Introduction

Protists of the genus *Leishmania* are aetiological agents of a spectrum of clinical diseases, known as leishmaniases, ranging from disfiguring skin lesions to life-threatening visceral infection. Depending on the *Leishmania* species, infection of humans may result in varying forms of leishmaniasis: cutaneous, diffuse cutaneous, mucocutaneous and visceral. The parasite has a worldwide distribution, both in tropical and subtropical regions, and around 1.5 million new cases occur annually, resulting in 50,000 deaths, principally from visceral leishmaniasis (Alvar *et al.*, 2012). Moreover, local socio-economical changes and world climate alterations are contributing to spreading both endemic regions and the number of affected persons (Dujardin *et al.*, 2008; Pigott *et al.*, 2014; Herrador *et al.*, 2015). Canine leishmaniasis is also a serious problem; it is estimated that 2.5 million dogs are infected in the Mediterranean basin alone (Moreno and Alvar, 2002).

Despite the existence of around 20 *Leishmania* species as human pathogens and the wide spectrum of clinical pathologies that they produce, the genus *Leishmania*, regarding morphology and life cycle features, constitutes a homogenous group (Akhoundi *et al.*, 2016). *Leishmania* parasites are transmitted between vertebrate hosts by phlebotomine sandflies of the genera *Phlebotomus* (Old World) and *Lutzomyia* (New World). In the phlebotomine insect vector, the parasite thrives as an extracellular flagellated form (named promastigote), which is attached to the alimentary tract of the insect (Dostálová and Volf, 2012). Upon transmission to the vertebrate host, the parasite is phagocytosed by macrophages, being able to survive and grow inside phagolysosomes, where it adopts a rounded form known as amastigote. These two life cycle stages are adapted to different environmental conditions. In the insect vector, promastigotes grow extracellularly, using sugars as the main nutritional source. In contrast, amastigotes grow intracellularly within the acidic environment of the phagolysosome, in which amino acids are the more abundant nutrients (McConville and Naderer, 2011). Also, drastic changes in temperature occur during transmission of promastigotes from the vector to the mammalian host. A pre-adaptation process takes place within the sandfly vector that involves the development from a noninfective, 'procyclic' stage into an infective, 'metacyclic' stage that is adapted for transmission in the fly and survival in the mammalian host (reviewed in Requena, 2012). In summary, *Leishmania* parasites must cope with many and heterogeneous stressful situations during their life cycle: (i) temperature changes, from ambient temperature in the sandfly vector to 37°C in the mammalian host; (ii) pH changes, after internalization into the macrophage, inside the phagolysosome, the environmental pH decreases to very acidic values; (iii) changes in nutrient and oxygen availability; and (iv) induction of resistance mechanisms to evade both the immune attack by complement components and the harmful effects of antimicrobial molecules produced by phagocytic cells. To cope with these hostile environmental conditions, *Leishmania* has evolved mechanisms of environmental sensing that are able to distinguish between the physiology of vector and mammalian hosts, and controlling accordingly the expression of *Leishmania* stage-specific factors. Nevertheless, how the different *Leishmania* stages are able to sense the environmental conditions, which are typical for either the insect or the mammalian, and to respond accordingly by altering gene expression are two largely unresolved issues.

Leishmania promastigotes are easily cultured in axenic media, and it is generally accepted that these forms are equivalent to the stages thriving in the sandfly midgut. In fact, the axenic promastigotes also undergo a process of programmed differentiation from a relatively

noninfective stage (logarithmic-phase cultures) to a highly infective metacyclic stage, which appears when the cultures reach the stationary phase (Da Silva and Sacks, 1987). Furthermore, highly infective metacyclic promastigotes of some *Leishmania* species can be isolated from stationary-phase cultures on the basis of their failure to be agglutinated by the lectin, peanut agglutinin (Sacks *et al.*, 1985). Similarly, for some strains of particular *Leishmania* species, the amastigote forms can be obtained in axenic conditions. Thus, using conditions mimicking those encountered during insect-to-mammal transmission, i.e. temperature upshifts *in vitro* from 25°C to 37°C combined with acidification of the growth medium, promastigote-to-amastigote differentiation is readily obtained in axenic culture (Castilla *et al.*, 1995; Callahan *et al.*, 1997; Balanco *et al.*, 1998). Nevertheless, there exists some controversy regarding whether or not these axenic amastigotes must be considered *bona fide* amastigotes. On the one hand, several studies have previously shown that axenically grown amastigotes clearly resemble intracellular amastigotes with regards to their ultrastructural, biological, biochemical and immunological properties (Gupta *et al.*, 2001). On the other hand, however, it cannot be forgotten that amastigotes grow inside macrophages, where host molecules presumably modulate parasite cellular processes. In this regards, there are also several studies illustrating that axenic amastigotes (often known as amastigote-like forms) differ from intracellular amastigotes in many molecular and biochemical aspects (Gupta *et al.*, 2001; Pescher *et al.*, 2011). A conclusion emerging from these studies is that axenic growth of the *Leishmania* forms is helpful for both the characterization of parasite molecules and understanding some molecular/cellular processes, but environmental conditions are different to those encountered by the parasite along its natural life cycle, even when considering the promastigote stage. In fact, *L. infantum* promastigotes obtained from the anterior thoracic midgut of *Phlebotomus perniciosus* are considerably more infective than promastigotes in the stationary phase of axenic cultures, and important differences in their transcriptomes were reported (Alcolea *et al.*, 2016a). Moreover, it has been shown that the presence of mammalian serum in the culture media has a dramatic effect on *Leishmania* gene expression (Alcolea *et al.*, 2016d).

The mechanisms controlling gene expression in *Leishmania* substantially differ from those in other eukaryotes (Papadopoulou *et al.*, 2008; Requena, 2011; De Pablos *et al.*, 2016). Thus, regulation of transcription on a per gene basis does not exist for almost the totality of genes; instead, genes are initially transcribed as polycistronic precursor RNAs from poorly characterized promoters located at the 5′ of large directional gene clusters (DGCs). Therefore, we will dedicate the first sections of this chapter to describe our current knowledge about genome structure and the mechanisms of gene expression in *Leishmania*. Then, the readers would be ready to better appreciate the peculiarities of the analysis of Omics data derived from *Leishmania* studies. Finally, after presenting the contributions of Omics approaches to our knowledge of gene expression, we will discuss how -omics are adding new strategies for diagnosis and control of leishmaniasis.

The *Leishmania* genome

It is relevant to briefly summarize the history leading to the uncovering of the genome structure in *Leishmania* parasites. In 1992, P. Bastien and co-workers published an article entitled '*Leishmania*: Sex, Lies and Karyotype' (Bastien *et al.*, 1992); this outstanding article clearly shows the particular difficulties that scientists had to solve, many were hurdles

derived from the peculiarities of the *Leishmania* genome. In fact, nowadays, questions like ploidy and genetic exchange remain unsolved. As *Leishmania* chromosomes do not condense during mitosis (Solari, 1995), their number and sizes could be studied only by pulsed field gel electrophoresis (PFGE), a technique that allows for electrophoretic separation of DNA molecules of very different size (from some kilobases (Kb) to megabase (Mb)-size molecules). The images derived from those analyses were called 'molecular karyotypes'. These images suggested that the karyotype of *Leishmania* should be extremely polymorphic, since strains of the same *Leishmania* species yielded PFGE-pictures very different with chromosomal bands of different size and intensity of staining (Blaineau *et al.*, 1991). However, this puzzling finding did not discourage to the Montpellier team in its objective of determining the *Leishmania* karyotype and chromosomal structures. For this purpose, they decided to establish physical linkage groups by hybridization of specific DNA probes to PFGE-separated chromosomes. After individual hybridization of 244 different probes, they concluded that the *L. infantum* genome comprises 36 chromosomes ranging in size from 0.35 to ≈ 3 Mb (Wincker *et al.*, 1996). Moreover, these authors demonstrated that the linkage groups were consistently conserved in other species, such as *L. major*, *L. tropica* and *L. aethiopica*. This finding contrasted with the size heterogeneity of the chromosomal bands derived from the PFGE analysis, and pointed to a high conservation of the chromosomal scaffolds within the genus *Leishmania* (Wincker *et al.*, 1996). This outstanding work paved the way for deciphering the genome sequence for this genus, first in *L. major* (Ivens *et al.*, 2005), afterwards for *L. infantum* and *L. braziliensis* (Peacock *et al.*, 2007) and, nowadays, for many other species (Cantacessi *et al.*, 2015; www.tritrypdb.org). Nevertheless, as shown in a recent article, the *L. major* (Friedlin) genome, which is the best-assembled genome to date for a *Leishmania* species, cannot be considered as set in stone (Alonso *et al.*, 2016), and users should be aware that a genome sequence is only a hypothesis.

The Old World *Leishmania* spp. have 36 chromosomes, while the New World *Leishmania* spp. have 34 or 35 chromosomes. Thus, *L. mexicana* has linkage groups of chromosomes 8 and 29 as well as of chromosomes 20 and 36, and *L. braziliensis* has a linkage group of chromosomes 20 and 34 (Britto *et al.*, 1998), regarding the *L. major* chromosomal set. Nevertheless, taking into account the evolutionary distance existing between these *Leishmania* species (see below), translocations must be considered extremely rare, and the conservation of the major linkage groups seems to be an essential feature of the genome of this parasite (Smith *et al.*, 2007).

The completion of the genome sequence in *L. major* and those of two related trypanosomatids in 2005, *Trypanosoma cruzi* (El-Sayed *et al.*, 2005) and *T. brucei* (Berriman *et al.*, 2005), revealed an extraordinary feature regarding gene organization. In the genome of these parasites, protein-coding genes are arranged into large unidirectional clusters, in which the genes present have no common nor akin function. These directional gene clusters (DGCs) are separated by short sequences termed strand-switch regions (SSRs), where the clusters converge towards or diverge from. This peculiar gene organization is dictated by another surprising mechanism existing in these parasites: polycistronic transcription (see next section). Also, this unusual way of genome organization would be linked to the very high gene density (around one gene every 3.2 kb) existing in the *Leishmania* genome (and those of related trypanosomatids); this may be an evolutionary advantage for the parasite in order to reduce cell cycle length. In this regards, recent results point to the existence of a large number of potential replication origins in *Leishmania*,

whose activation would be regulated in a concerted manner with transcriptional activity (Lombraña *et al.*, 2016).

These peculiarities in gene organization and transcriptional mechanisms existing in *Leishmania* are related, in part, to the fact that these protists occupy a deep-branched position on the evolutionary tree of eukaryotes (Sogin *et al.*, 1986). The genus *Leishmania* belongs to the order Trypanosomatida (Moreira *et al.*, 2004), in which are also included *T. brucei* and *T. cruzi*, causative agents of other important infectious diseases for humans, sleeping sickness and Chagas disease, respectively. In fact, the genera *Leishmania* and *Trypanosoma* share mechanisms of gene expression, and even a remarkable synteny regarding gene organization.

According to recent estimations, the common ancestor for *Trypanosoma* and *Leishmania* lasted around 400 or 600 millions of years (Stevens *et al.*, 2001). However, in spite of the high level of divergence existing at the level of protein-coding sequences, species like *T. brucei*, *T. cruzi* and *L. major* exhibit a striking conservation in the gene arrangement, suggesting that selection has maintained gene order in this parasitic species over hundreds of millions of years of evolution (Ghedin *et al.*, 2004). Furthermore, when comparisons of genome sequences of *L. major*, *L. infantum* and *L. braziliensis* are done, the synteny across the genus is almost absolute, only a few differences regarding gene content exist (Peacock *et al.*, 2007). This is really surprising taking into account that the divergence of these species took place around 10 million (between *L. infantum* and *L. major*) and 100 million (between *L. infantum* and *L. braziliensis*) years ago, and the different pathologies that each one of these species causes in humans (McGwire and Satoskar, 2014). At present, how those distinct disease pathogeneses emerge from this strong conservation in both gene content and synteny is a question yet to be solved. It is possible that just a few species-specific genes are crucial to disease outcome, but, alternatively, it is also possible that the parasite genome may play only a minor role in determining disease phenotype, which would be mainly determined by differences in the expression levels of key conserved genes.

Vanishing of transposable elements and flourishing of retroposon-derived, repeated sequences in the *Leishmania* genome

Transposable elements (TEs) and related sequences constitute a substantial proportion of eukaryotic genomes, composing, for instance, more than 50% of the human genome (Richard *et al.*, 2008). TEs are classified into two major groups based on the mechanisms used for their transposition: retroelements or retroposons, which transpose via reverse transcription of an RNA intermediate, and DNA transposons that move through a DNA intermediate. Based on the presence (or its absence) of a long terminal repeat (LTR), retrotransposons are subdivided into LTR- and non-LTR-retrotransposons. A particularly abundant group of non-LTR retrotransposons are the long interspersed nuclear elements (LINEs), which in the human genome have a copy number of 850,000, and 660,000 LINEs were found in the mouse genome (Richard *et al.*, 2008). LINEs encode a reverse transcriptase and other enzymes required for retrotransposition, and some of these elements are indeed functional (Belancio *et al.*, 2008). However, the tendency to accumulate LINEs in the genomes has not been a general rule in eukaryotes, and even some organisms, as many species of *Leishmania*, have eliminated them from their genomes (Bringaud *et al.*, 2006).

In contrast, the related trypanosomatids *T. brucei* and *T. cruzi* contain a significant amount of retrotransposons, which are very similar in both species, suggesting a common evolutionary history. Thus, they contain VIPER, an LTR retrotransposon originally characterized

in the *T. cruzi* genome (Vazquez *et al.*, 2000). Also, they contain non-LTR retrotransposons, which are divided into site-specific retroelements (*T. brucei* SLACS, *T. cruzi* CZAR and *Crithidia fasciculata* CRE1/CRE2 elements), which are always inserted at the same relative position in the spliced leader (SL) RNA genes (Aksoy, 1991) and non-site-specific retroposons (Ingi and L1Tc) (Bringaud *et al.*, 2008; Thomas *et al.*, 2010). The genomes of those trypanosomatids also contain small non-autonomous retroposons derived from the autonomous retroelements (Ingi and L1Tc): NARTc in the *T. cruzi* genome and RIME in the *T. brucei* one (Bringaud *et al.*, 2008; Thomas *et al.*, 2010). These retroelements and their derivatives constitute up to 5% of the genomic content of both species of trypanosomes. In contrast to trypanosomes, the genomes of the Old world species *L. major* and *L. infantum* do not contain any complete retroelement. However, a few highly degenerated Ingi/L1Tc-related elements (DIREs) have been identified in the genomes of these species, indicating that the genome of their ancestor contained active retroposons but they have been lost over time to become vestigial retroelements in the genome of the present day *L. major* and *L. infantum* (Bringaud *et al.*, 2006).

Another remarkable feature of Ingi and L1Tc retroelements is the presence at their 5′-end of a conserved stretch of 77–79 bp (Bringaud *et al.*, 2002), which was also present in the short non-autonomous non-LTR retrotransposons identified in *T. brucei* (RIME; Hasan *et al.*, 1984) and *T. cruzi* (NARTc; Bringaud *et al.*, 2002). This sequence, named 'non-LTR retrotransposon signature' (Bringaud *et al.*, 2006) or Pr77 (Heras *et al.*, 2007), was found to be present in the *L. major* genome at the time of NARTc characterization in *T. cruzi* (Bringaud *et al.*, 2002), and afterwards in some of the *L. major* DIREs (Bringaud *et al.*, 2006). Subsequently, using the 79-bp signature (the hallmark of trypanosomatid retroposons) as bait, Bringaud and co-workers identified in the *L. major* genome two large families of small repeated elements that were named LmSIDER1 (785 copies) and LmSIDER2 (1,073 copies); LmSIDER stands for *L. major* Short Interspersed DEgenerated Retroposon (Bringaud *et al.*, 2007). A detailed analysis of LmSIDER2 elements in the *L. major* indicated that most LmSIDER2 would be located within the 3′UTR of mRNAs, suggesting a role of these elements in modulating mRNA stability (Bringaud *et al.*, 2007). Moreover, it was shown that, unlike most eukaryotic transcripts, some of the SIDER2-bearing mRNAs do not undergo poly(A) tail shortening prior to rapid turnover, but instead, they are targeted for degradation by a site-specific endonucleolytic cleavage (Müller *et al.*, 2010a; Müller *et al.*, 2010b). In this regard, the auto-catalytic ribozyme activity, demonstrated in the Pr77 sequence (retroposon-hallmark, see above) of L1Tc (Sánchez-Luque *et al.*, 2011), might be responsible for the endonucleolytic cleavage observed in those SIDER2-bearing mRNAs (Sánchez-Luque *et al.*, 2014). Another salient feature of SIDER2 elements is that they can group, based on sequence similarity, into subfamilies, whose members are restricted to particular regions of a sole chromosome (Requena *et al.*, 2008; Requena *et al.*, 2017). This peculiar pattern of distribution would be related to the genomic plasticity of this parasite (see below). Also, members of the family SIDER1 have been involved in regulatory mechanisms of gene expression as stage-specific translational regulators (McNicoll *et al.*, 2005). In summary, *Leishmania* would be an outstanding example in eukaryotes of domestication and expansion processes in which TEs have evolved to fulfil essential roles related to gene expression, genome plasticity and perhaps transcriptional initiation (Bringaud *et al.*, 2008; Requena *et al.*, 2017).

In this context, it was surprising to detect potentially active retroposons in the genome of *L. braziliensis* (Peacock *et al.*, 2007). Thus, non-LTR retrotransposons, equivalent to *T. brucei* SLACS and *T. cruzi* CZAR, are associated also with the SL gene array in *L. braziliensis*. In addition, the telomeres in the *L. braziliensis* genome accommodate DNA transposable elements, called TATEs (telomeric associated transposable elements), encoding putative reverse transcriptase, phage integrase (site-specific recombinase) and DNA/RNA polymerase domains. Another remarkably distinct feature found in *L. braziliensis*, but absent from the genome of Old World *Leishmania* species, is the existence of an active mechanism of RNA interference (RNAi) (Lye *et al.*, 2010). In trypanosomatids, an RNAi mechanism was initially identified in *T. brucei* (Ngô *et al.*, 1998), but could not be demonstrated in other species including *L. major* and *T. cruzi* (Lye *et al.*, 2010). An open question is whether RNAi machinery has been lost or acquired across the *Leishmania* genus. Examination of the syntenic regions in *L. major* and *L. infantum* identified remnants of Argonaute in both species, an observation that favours the proposal that RNAi was lost in *L. major/L. infantum* after their divergence from the *L. braziliensis* evolutionary branch (Smith *et al.*, 2007). It can be hypothesized that the presence of retrotransposons in *L. braziliensis* parasites forced them to maintain the RNAi pathway, taking into account that it functions as a genome immune defence mechanism to limit the potentially deleterious consequences of transposon/retroposon mobilization (Carthew and Sontheimer, 2009).

Plasticity of the *Leishmania* genome: gene copy number alterations, repeated sequences and aneuploidy

It may result paradoxical to describe the *Leishmania* genome as unstable or plastic, after having indicated in the previous section that gene organization has been maintained static among evolutionarily distant *Leishmania* species, and even relatively few changes have occurred after the separation of the genera *Leishmania* and *Trypanosoma*. However, the *Leishmania* genome, when analysed as a whole, exhibits a remarkable degree of inter- and intraspecific variability. Initial gene mapping studies suggested that *Leishmania* is mostly diploid (Iovannisci and Beverley, 1989). However, it was later shown during gene knock-out experiments that *Leishmania* parasites can undergo changes in ploidy. Currently, the standard approach for deletion of a given gene in *Leishmania* consists of the sequential deletion of both chromosomal alleles by homologous gene replacement using constructs carrying different drug resistance markers flanked with the boundary regions of the gene to be deleted (Cruz *et al.*, 1991). Nevertheless, attempts to delete essential genes in *Leishmania* more often yield parasites bearing the planned allelic replacements but containing additional gene copies that arise via aneuploidy or polyploidy (Cruz *et al.*, 1993; Dumas *et al.*, 1997; Genest *et al.*, 2005). In fact, this observation is considered as a clue about the essentiality of the targeted gene, especially when accompanied by successful replacement in the presence of an ectopic gene (Ilgoutz *et al.*, 1999; Vergnes *et al.*, 2005; Murta *et al.*, 2009).

Other manifestations of the genomic plasticity of *Leishmania* are found in the DNA amplification events that occur after incubation of this parasite with a variety of anti-leishmanial drugs, and even amplicons have been found in unselected laboratory isolates (Beverley, 1991; Grondin *et al.*, 1996). In fact, an idea widely accepted is that *Leishmania* parasites can re-shape their genome rapidly *in vitro* to survive stressful environments (Leprohon *et al.*, 2009). Homologous recombination events between direct repeats (mostly involving SIDER elements) would lead to the generation of the circular amplicons, whereas recombination

between inverted repeats would generate linear or circular, inverted amplicons (Ubeda *et al.*, 2014). Gene amplification through homologous recombination between repeated sequences is a common mechanism of drug resistance in *Leishmania* (Grondin *et al.*, 1993, 1996), and it is probably the only possible strategy for altering gene expressions in organisms like trypanosomatids lacking transcriptional control (see below). Clues about the enzymes and molecular mechanisms responsible for these amplification events are being obtained. Extrachromosomal circular amplicons are dependent on the activity of RAD51 and RAD51–4 (Ubeda *et al.*, 2014; Genois *et al.*, 2015) while linear amplicons depend on MRE11 (Laffitte *et al.*, 2014). In summary, genomic rearrangements in *Leishmania* occur genome-wide (Ubeda *et al.*, 2014), they can lead to the formation of extrachromosomal elements (Leprohon *et al.*, 2009; Downing *et al.*, 2011), to supernumerary chromosomes and to mosaic aneuploidy (Sterkers *et al.*, 2011).

A special comment merits aneuploidy as a genetic feature of *Leishmania* parasites. Thus, different sets of data support the idea that chromosomal dosage in *Leishmania* genomes is aneuploid rather than strictly diploid (Ravel *et al.*, 1998; Sunkin *et al.*, 2000; Dubessay *et al.*, 2002b; Martínez-Calvillo *et al.*, 2005; Rogers *et al.*, 2011; Sterkers *et al.*, 2011). More strikingly, the genome of *Leishmania*, when considered at the population level, exhibits an unusual feature termed 'mosaic aneuploidy' (Sterkers *et al.*, 2011). This concept serves to describe that in a given *Leishmania* population, there are cells that are mono-, di-, trisomic (and so on) in proportions that vary from one chromosome to another, from one strain to another, and even from cell to cell (Lachaud *et al.*, 2014; Seco-Hidalgo *et al.*, 2015). How this remarkable genome plasticity is generated and tolerated is poorly understood at present.

Regulation of gene expression in *Leishmania*

Initial studies dealing with the analysis of the transcriptional activity of particular *Leishmania* genes showed somewhat unexpected results: transcription of individual genes remains constant irrespective of the abundance of the mRNA products. The heat shock protein (HSP) genes of *Leishmania* were among the first to be analysed, because it was easy to find them due to their remarkable sequence conservation along the evolutionary scale and, on the other hand, because they seemed to be ideal candidates for identification of inducible RNA polymerase II promoters (reviewed in Folgueira and Requena, 2007). Such analyses were based on the technically demanding nuclear-run-on procedure, which consists of the isolation of nuclei from promastigotes and their incubation with labelled ribonucleotide triphosphates. Thus, polymerases engaged to DNA in the act of transcriptional elongation continue to transcribe (or 'run-on') for several hundred additional nucleotides. Labelled nuclear RNA is then purified and hybridized to blots containing specific DNA probes. The hybridization signals equate directly to the density of PolII and, consequently, the rate of transcription of a given gene. Run-on analyses of the expression of the *HSP70* and *HSP83/90* genes in several *Leishmania* species concluded that there was not transcriptional activation of these genes when the parasites are exposed to a heat shock, in spite that accumulation of the respective mRNAs occurred (Brandau *et al.*, 1995; Quijada *et al.*, 1997). Subsequent studies analysing the transcriptional regulation of additional genes fostered the idea that most protein-coding genes in *Leishmania*, and related trypanosomatids, do not have their own promoters and are transcribed as part of polycistronic transcripts (Clayton, 2002). In an outstanding article, Martínez-Calvillo and co-workers (2003), by

using strand-specific nuclear run-on assays, showed that most of the transcriptional activity occurs on the sense DNA strand, and that RNA polymerase II-mediated transcription initiates within the strand-switch regions (SSRs) and radiates bidirectionally from them. The idea that transcription initiation is occurring at the SSRs was strengthened by the finding that acetylated histones are concentrated at divergent SSRs (putative sites of transcription initiation) but are absent from convergent SSRs (putative sites of transcription termination) in *T. cruzi* (Respuela *et al.*, 2008). In *T. brucei*, two histone modifications associated with open chromatin (H4K10ac and H3K4me3) and two histone variants (H2AZ and H2BV) concentrate in divergent SSRs, whereas variants of H3 and H4 are found near the probable pol II transcription termination sites of DGCs (Siegel *et al.*, 2009; Wright *et al.*, 2010). In *L. major*, acetylated histones H3 accumulate at all divergent SSRs (as well as other internal locations that are in turn predicted to be transcription start sites) in a manner that is more pronounced in dividing cells than in stationary-phase cells. Moreover, a remarkable enrichment of transcription factors TATA-binding protein (TBP) and SNAP50 has been also detected at these putative transcription start sites (Thomas *et al.*, 2009). In *Leishmania*, a modified thymidine, base J (β-D-glucosyl-hydroxymethyluracil), is located at chromosome-internal RNA polymerase II termination sites and promotes termination of transcription (van Luenen *et al.*, 2012; Reynolds *et al.*, 2016). This DNA modification, which has been found in the nuclear DNA of trypanosomatids and in the genera *Diplonema* and *Euglena* (van Leeuwen *et al.*, 1998; Dooijes *et al.*, 2000), is introduced enzymatically through the hydroxylation of thymidine, forming hydroxymethyluridine, followed by the transfer of a glucose by a specific glucosyltransferase (Bullard *et al.*, 2015).

However, currently, it is not clear where transcription precisely starts or how RNA pol II is recruited to the DNA. Furthermore, the principle that SSRs are the sole sites of transcription initiation is not absolute, since resistance genes introduced into a copy of *L. major* chromosome 1 lacking the divergent SSR can still be transcribed at levels sufficient for selection (Dubessay *et al.*, 2002a).

A polycistronic mode of transcription, as occurs in *Leishmania* and related trypanosomatids, does not leave much scope for regulation at the level of transcription initiation and explains why gene expression is mainly controlled post-transcriptionally in these organisms. The first task after polycistronic transcription is to separate the individual open reading frames into translatable mRNAs. In eukaryotes, crucial are the addition of a 5′-cap structure, which is made of modified bases and is usually put in place by the transcribing polymerase, RNA polymerase II, and the addition at the 3′-end of a poly(A) tail. Both modifications are required in an mRNA to be recognized as properly processed in order to be escorted to the cytosol; otherwise, the RNA may be held back in the nucleus and degraded (Moore, 2005).

In *Leishmania*, and related trypanosomatids, processing of the polycistronic transcripts to monocistronic RNAs and capping are accomplished by transcriptional *trans*-splicing of a capped, 39-nt RNA exon (named spliced leader -SL- or miniexon) to the 5′ untranslated region (5′UTR) and polyadenylation of the 3′UTR (Liang *et al.*, 2003). The mini-exon is transcribed by RNA polymerase II (pol II), yielding a ≈ 110–140-nt RNA donor (SL medRNA) for the *trans*-splicing reaction (Campbell *et al.*, 2000). The SL medRNA is encoded in an array of monocistronic tandem repeats located on chromosome 2; in contrast to the case for pol II-transcribed protein-coding genes, the SL genes are transcribed from bona fide pol II promoters (Gilinger and Bellofatto, 2001).The signal for *trans*-splicing consists of an invariant dinucleotide AG (3′-splice acceptor site) and an upstream polypyrimidine tract of

variable length (Requena *et al.*, 2003). Interestingly, polyadenylation is inextricably coupled to *trans*-splicing, and contrary to most eukaryotes, polyadenylation signals have not been found in trypanosomatids. Instead, poly(A) site selection is conditioned by the position of the downstream splice acceptor site (LeBowitz *et al.*, 1993).

Taking into account that the two principal roles of promoters in other eukaryotes (capping and mRNA start site selection) are instead handled in trypanosomatids by the *trans*-splicing machinery, it is conceivable that promoters may be not necessary and possibly might not exist for pol II-transcribed protein-encoding genes. Nevertheless, the lack of transcriptional regulation in gene expression does not mean that these parasites have renounced to regulatory purposes; quite the contrary, the complexity of their life cycles makes differential gene expression of central importance to complete their biphasic developmental programmes. Therefore, their needs of controlling gene expression must be accomplished through downstream mechanisms that involve RNA and protein regulatory processes.

The requirement for the addition of the SL sequence to the polycistronic RNAs represents a potential point for modulation of mRNA levels between transcripts, since differences in the efficiency of *trans*-splicing would contribute to producing differences in steady state levels of specific mRNAs. Similarly, the polyadenylation process may also contribute to generating differences in expression for transcripts derived from the same polycistronic precursor. In fact, mutagenesis of the polypyrimidine tracts leads to the utilization of cryptic 3′-splice and polyadenylation sites (Matthews *et al.*, 1994). Additionally, for a given transcript, alternative *trans*-splicing events may occur, generating differences in the length of 5′-UTRs, but also in the protein-coding regions. The relevance of these processes for regulation of gene expression is currently unknown, but answers are likely to be forthcoming soon.

The heat shock response in *Leishmania*: a prototypical model for studying gene expression

For all living organisms, a moderate increase of just a few degrees above the optimum growth temperature represents a challenging problem for survival (Richter *et al.*, 2010). Temperature affects many physiological processes in the cell, as protein conformation is highly dependent on this physiochemical factor, and proteins tend to denature in response to temperature increases. As a survival mechanism, in response to heat stress, all the cells and organisms activate the so called heat shock response, which consists of several signalling pathways aimed to increase the expression of the heat shock proteins (HSPs; de Nadal *et al.*, 2011). The main role of HSPs (also known as molecular chaperones) is alleviating the deleterious effects of heat increase on the cellular homeostasis, by facilitating protein folding and preventing (or reversing) protein misfolding. In response to heat shock, gene expression changes quickly; hence, the heat shock response is being used as an amenable system to investigate gene expression mechanisms in different cells and organisms.

Changes in temperature occur regularly in the life cycle of *Leishmania*, when the parasite is transmitted between the insect vector and mammalian hosts, and accordingly temperature shift is considered one of the factors triggering differentiation from the promastigote form to the amastigote stage (Requena, 2012). However, the parasite also experiences abrupt changes in temperature in the poikilothermic vector, for example when it is being exposed directly to the sunlight. Therefore, it is conceivable that heat shock response is being used as a suitable model for analysing gene expression in *Leishmania* (Schwede *et al.*, 2012; Requena, 2012). In most eukaryotes, the regulation of HSP synthesis is mediated by

heat shock transcription factors (HSFs), which act as master regulators of stress-induced gene expression through binding to defined cis-acting heat shock elements (HSE) inside promoters or enhancers of HSP genes (Santoro, 2000). In *Leishmania*, there is no evidence for the existence of either HSE promoter elements or HSFs. This is not so surprising taking into account that, indeed, the entire gene expression in *Leishmania* is regulated mainly at the post-transcriptional level (Requena, 2011). Transcript stabilization and preferential translation of HSP mRNAs are the main mechanisms by which these parasites control HSP synthesis (Quijada *et al.*, 1997, 2000; Zilka *et al.*, 2001; Larreta *et al.*, 2004; Folgueira *et al.*, 2005; David *et al.*, 2010). By way of example, the regulation of HSP70 expression is illustrated in Fig. 5.1, and presented briefly. The genomic organization and expression of *HSP70* genes have extensively been studied in *L. infantum*. The *L. infantum* genome contains a single locus with six *HSP70* genes arranged tandemly in a head-to-tail manner (Quijada *et al.*, 1997). This gene organization was found to be conserved in other *Leishmania* species (Folgueira *et al.*, 2007; Ramírez *et al.*, 2011). All six *HSP70* genes have identical 5′-UTRs and coding sequences, the sole difference is the 3′-UTR of gene 6 (located at the 3′ end of the cluster; Fig. 5.1A) that is absolutely divergent in sequence relative to the 3′-UTR of the other five genes. Hence, for simplicity, genes 1 to 5 are referred as *HSP70-I* genes and gene 6 as *HSP70-II* gene (Folgueira *et al.*, 2005). Regarding transcripts levels, it is remarkable that nearly all the mRNAs come from gene 6, but only expression levels from *HSP70-I* genes are increased by heat shock (Fig. 5.1B; Quijada *et al.*, 1997). Subsequently, it was shown that the temperature-dependent accumulation of *HSP70-I* transcripts is associated with the presence of a *cis*-acting sequence in their 3′-UTR (Quijada *et al.*, 2000). Regarding translation, *HSP70-II* mRNAs are preferentially translated at heat shock temperatures, but not

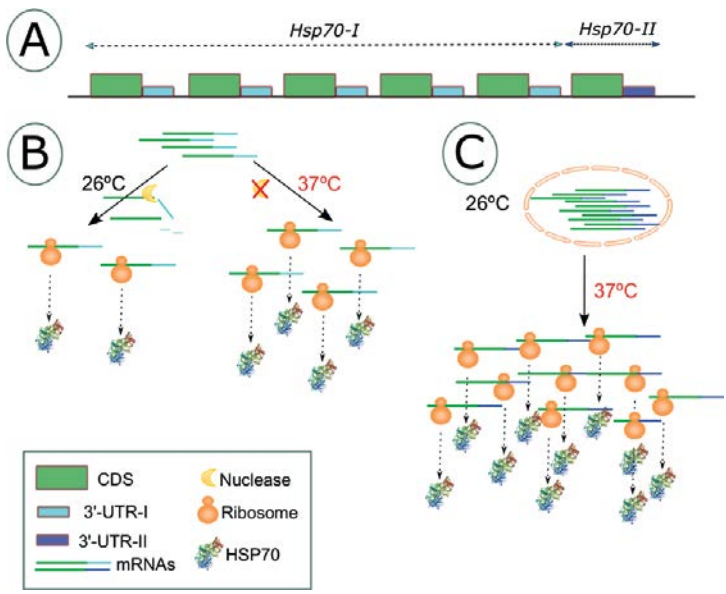

Figure 5.1 Regulation of *HSP70* gene expression in *Leishmania*. (A) Organization of the *HSP70* locus. Expression of (B) *HSP70-I* and (C) *HSP70-II* transcripts at normal (26°C) or heat shock (37°C) temperatures. See text for additional details.

at 26°C, whereas *HSP70-I* mRNAs are bound to polysomes at 26°C and 37°C (Fig. 5.1B; Folgueira *et al.*, 2005). This well-documented gene model illustrates how the expression of a gene may be controlled by variations in mRNA abundance, stability and translatability through differences in sequence elements residing in 3′-UTRs. Differential gene expression controlled by mechanisms affecting mRNA stability and translation has been documented in other *Leishmania* genes (Boucher *et al.*, 2002; Myung *et al.*, 2002; Soto *et al.*, 2004).

This post-transcriptional regulation relies on the interaction between trans-acting RNA-binding proteins and cis-acting RNA elements (Kramer and Carrington, 2011). In this regard, a putative master regulator of HSP mRNA stability has been described in *T. brucei*. Thus, the zinc finger protein ZC3H11 was identified as a specific binding factor for a common AUU repeat motif, which is found inside the 3′UTR of HSP70 and other genes coding for stress proteins (Droll *et al.*, 2013).

Another regulatory level of the stress response in *Leishmania* may operate through phosphorylation/dephosphorylation of HSPs and other proteins (Morales *et al.*, 2010). Thus, during the promastigote-to-amastigote differentiation, an increase of phosphorylated sites in HSP70 and HSP90 has been demonstrated (Morales *et al.*, 2008). These findings suggest that stress-response regulation in *Leishmania* would be done in part by protein kinases that regulate chaperone function (Späth *et al.*, 2015). Among the candidate kinases for regulating the stress response, it has been suggested the *Leishmania* PKR-homologue (named PERK), based on the observation that exposure of promastigotes to a combination of elevated temperature and acidic pH, key signals triggering amastigote differentiation, leads to a marked decrease in global translation initiation and a concomitant increase in the phosphorylation of the translation factor $eIF_{2\alpha}$ (Cloutier *et al.*, 2012). Moreover, impairing PERK leads to a defect in intracellular amastigote differentiation (Chow *et al.*, 2011), which provides a further link between *Leishmania* stress signalling and parasite differentiation. Nevertheless, other *Leishmania* kinases have also been postulated as putative master regulators of the heat shock response (Späth *et al.*, 2015).

Although important advances in the characterization of the response to stress in *Leishmania* have been achieved, little is known about the molecular details of the switch that triggers the heat shock response. In this way, a relevant finding was the observation that calcium uptake and calcineurin signalling are required for the development of parasite thermotolerance at 34–37°C (Naderer *et al.*, 2011).

Genomics and transcriptomics approaches

The Omics technologies are tools designed to gather information on the vast majority of constituents of a cell, a tissue, an organ or even the whole organism. They are aimed primarily at the universal detection of genes (genomics), mRNA (transcriptomics), proteins (proteomics) and metabolites (metabolomics) in a specific biological sample and in a non-biased manner. In this section and the following ones, brief descriptions on the foundations of the main Omics technologies, together with their contribution to increasing our knowledge on *Leishmania* biology, are provided.

Genomics is the study of the sequence and structure of the genome, paying a particular attention to gene organization. Transcriptomics deals with genome-wide identification and quantification of RNA species such as mRNAs, non-coding RNAs and small RNAs. Another Omic, closely related to genomics, is epigenomics, which study modifications and

structural alterations of the chromatin, and their relationships with genome replication and transcription dynamics.

When genome-sequencing projects began to generate a wealth of information about gene content in *Leishmania*, and other organisms, DNA microarrays became a major high-throughput platform by which genome-wide analysis of gene expression could be explored at that time (Boothroyd *et al.*, 2003; Duncan, 2004). Serial Analysis of Gene Expression (SAGE) was another sequencing-based methodology used for transcriptomics studies (Velculescu *et al.*, 1995), but currently it has fallen into disuse.

Microarrays consist of ordered collections of thousands of nucleotide sequences (genomic DNA, cDNA, or oligonucleotides) set as microscopic spots on a glass slide (Boothroyd *et al.*, 2003; Duncan, 2004). With DNA microarrays, gene expression levels are measured as the amount of RNA in the sample that matches the set of probes fixed on the array. Either RNA molecules or more often their cDNA copies are fluorescently labelled and hybridized onto the array, where the intensity of the signal, measured for a given probe, is assumed to be proportional to the quantity of RNA present in the sample. The simultaneously use of two probes (labelled separately with Cy3- or Cy5-fluorescent dyes) allows for direct comparison between two environmental/physiological conditions, and the relative abundance of a given sequence under the two conditions can be deduced from the ratio of fluorescence intensity between both dyes. According to the nature of the probes, which are mainly derived from genomic DNA or cDNAs, synthesized from the RNA samples of interest, the application of this technology serves to address genomic or transcriptomic purposes, respectively.

The quality of the microarrays relies largely on the quality of the libraries, and the construction of them is the most cumbersome step. Thus, cDNA libraries have the problem that some genes are over-represented (those with a high level of expression) whereas others, corresponding to low expressed transcripts, may be completely absent. To avoid this problem of complexity, genomic libraries are a valid alternative in *Leishmania*, as the vast majority of genes lack introns and its genome has a remarkable gene density. However, a clear drawback of genomic libraries is that some segments can span two or more genes. For *Leishmania* species, whose genomes have been sequenced and annotated, these limitations were solved by using oligonucleotides that, furthermore, can be prepared and spotted directly onto the microarray slides.

In the last 10 years, an extraordinary progress has been made in sequencing technology. Currently, various sequencing platforms can generate simultaneously several millions of sequences with very affordable costs. These methodologies are often referred as next-generation sequencing (NGS); for readers interested in the technical details about the main types of NGS approaches, we suggest consulting an outstanding review published recently (Goodwin *et al.*, 2016). NGS platforms can be classified into two categories regarding the size of the sequences generated: short-reads (35–700 nucleotides) and long-reads (several hundreds of nucleotides). Within the category of short-read sequencing, NGS research is increasingly being conducted with Illumina instruments, which provide vast quantities of data; whereas the most widely used long-read platform is the single-molecule real-time (SMRT) sequencing approach (Eid *et al.*, 2009), developed by Pacific Biosciences (PacBio). Nevertheless, this is a field in continuous evolution and surely, in the next few years, novel sequencing solutions will emerge.

NGS can indeed be used for DNA or RNA sequence analyses. In particular,

high-throughput sequencing of cDNA copies of mRNA (RNA-seq), because millions of sequences are generated, provides a digital read-out of mRNA levels over several orders of magnitude and allows for mapping of transcripts to the nucleotide level. Hence, RNA-seq has become particularly useful for quantitation of relative abundance of transcripts and for detection of polymorphisms, minor RNA species, and RNA editing. Currently, RNA-seq is replacing the use of microarrays for analysis of gene expression, because it has several advantages over microarrays: array design is not required, it has better resolution, and very low amounts of input material are used. Furthermore, microarrays can be limited by probe cross-reactivity, high background, and signal saturation. But depending on the array design, chips may provide rapid and high-quality data at lower cost. In fact, microarrays remain widely used in genomic research, such as allelic polymorphism determinations, pathogen identification and measuring of expression levels for selected genes. For example, DNA microarrays have been developed for simultaneous identification of 18 species of common blood protozoa affecting humans, among them *Leishmania* (Chen *et al.*, 2016).

Translation of the vast amount of genetic data into biological contexts is currently a bottleneck; mapping the vast amount of short read sequences (Martin and Wang, 2011) and assessing the significance of differences in reads coverage are two relevant questions to address. In addition, well-validated experimental samples would be a requisite, before proceeding with NGS data collection. Finally, powerful databases containing this plethora of data, continuously organized and updated, are crucial. Currently, TriTrypDB (http://tritrypdb.org) is the most relevant database aimed to provide access to genome-scale datasets for many kinetoplastid parasites (Aslett *et al.*, 2010). This database forms part of the platform Eukaryotic Pathogen Bioinformatics Resource Centre (EuPathDB; Aurrecoechea *et al.*, 2013).

Leishmania transcriptomics

Just before the completion of the genome sequence for a *Leishmania* species, DNA microarrays emerged as a powerful method to study global gene expression in terms of quantitation of mRNA levels (see previous section for details on microarrays foundations). The initial studies dealt with the use of whole-genome shotgun DNA microarrays (Saxena *et al.*, 2003; Akopyants *et al.*, 2004), but the completion of the genome sequencing of *L. major* (Ivens *et al.*, 2005), *L. infantum* and *L. braziliensis* (Peacock *et al.*, 2007) allowed the generation of high-density oligonucleotide microarrays suitable for genome-wide expression profiling analyses (Rochette *et al.*, 2008). Microarrays monitoring gene expression have been used for the study of various aspects of *Leishmania* biology: metacyclogenesis (Saxena *et al.*, 2003; Alcolea *et al.*, 2009), promastigote-to-amastigote differentiation (Holzer *et al.*, 2006; McNicoll *et al.*, 2006; Leifso *et al.*, 2007; Srividya *et al.*, 2007; Rochette *et al.*, 2009; Alcolea *et al.*, 2010; Lahav *et al.*, 2011), sandfly-derived promastigotes (Alcolea *et al.*, 2014, 2016a), species-specific gene expression (Depledge *et al.*, 2009), intraspecific heterogenicity of clinical isolates (Adaui *et al.*, 2011), resistance to antileishmanial drugs (Guimond *et al.*, 2003; Leprohon *et al.*, 2009; Singh *et al.*, 2010), and response to stress conditions imposed by heavy metals (Alcolea *et al.*, 2011a). On the other hand, microarrays have also been used for monitoring the expression of host genes being altered by *Leishmania* infection (Novais *et al.*, 2015).

The application of high-throughput RNA sequencing (RNA-seq) to the study of *Leishmania* parasites has represented a milestone in the knowledge of their transcriptome, both

regarding the real number of genes and their primary structure. The first transcriptome for a species of the genus *Leishmania* was reported by Rastrojo *et al* (2013). These authors documented the existence of 10,285 different transcripts, of which 1884 were considered novel, as they did not match previously annotated genes. In addition, the annotations for around 500 ORFs were corrected based on the location of SL-addition sites. In 2015, Fiebig *et al* (2015), using also RNA-seq, defined the transcriptome for *L. mexicana*; a total of 9169 protein-coding transcripts, including 936 corresponding to previously non-annotated genes, were identified.

Apart from its usefulness for establishing the gene models, RNA-seq methodology is an excellent tool for quantitation of relative expression levels for each one of the transcripts composing the *Leishmania* genome (Rastrojo *et al*., 2013). Also, RNA-seq has begun to be applied to the measurement of mRNA transcript abundance in some developmental stages of *Leishmania*. Thus, Dillon *et al.* (2015b) used RNA-seq to identify global changes in gene expression occurring as *L. major* undergoes metacyclogenesis, a developmental progression that is well mimicked *in vitro* using reliable axenic cultivation methods (Sacks and Perkins, 1984). Remarkably, this analysis identified 3138 genes that were expressed at significantly different levels between procyclic and metacyclic promastigotes (Dillon *et al*., 2015b). The transcriptomes of *L. mexicana* promastigotes, axenic amastigotes and intracellular amastigotes have also been compared by RNA-seq (Fiebig *et al*., 2015). More than one-third of all genes were found to be differentially expressed between promastigotes and intracellular amastigotes.

RNA-seq methodologies have also been used for analysing the complete repertoire of *L. major* small nucleolar RNAs (snoRNAs), which are the molecules guiding modifications by methylation and pseudouridylation on rRNAs (Eliaz *et al*., 2015). rRNA processing events in trypanosomatids are unique, as the large subunit rRNA undergoes trypanosome-specific cleavages during rRNA maturation, yielding two large rRNA molecules and four small ones (Hernández and Cevallos, 2014).

RNA-seq has also been used to simultaneously identify global transcriptional changes in host cells and *L. major* parasites during the entry into and persistence within murine macrophages (Dillon *et al*., 2015a). Similar studies were carried out using human macrophages, and the infection with two different *Leishmania* species, *L. amazonensis* and *L. major* (Fernandes *et al*., 2016). No significant differences were observed between the two *Leishmania* species transcriptomes or in the transcriptional response of human macrophages infected with each species. Nevertheless, depending on the host cells (human or mouse macrophages), remarkable differences in the *L. major* transcriptome were observed (Fernandes *et al*., 2016). More recently, host and parasite gene expression in skin biopsies from *L. braziliensis*-infected patients were simultaneously analysed using RNA-seq (Christensen *et al*., 2016). It was evidenced a clear uniformity in the parasite gene expression in all the patient biopsies analysed. Remarkably, 4579 host genes were differentially expressed in patient's skin lesions relative to the expression occurring in normal skin. Again, the most abundant parasite transcripts expressed in lesions were distinct from transcripts expressed *in vitro* in human macrophage cultures infected with *L. amazonensis* or *L. major* (Christensen *et al*., 2016). Additionally, RNA-seq approaches have been used to uncover microRNAs (short single-stranded ribonucleic acids that act blocking mRNA translation) that are specifically expressed in human monocyte-derived dendritic cells and macrophages after infection with *Leishmania* parasites (Geraci *et al*., 2015).

Determination of the mRNA levels existing in a cell at a given moment does not allow for directly inferring the abundance of the encoded protein, as the latter also depends on other factors like nuclear-cytoplasmic transport, transcript-specific silencing mechanisms and mRNA translational efficiency. However, technologies for assessing the relative abundance of all proteins existing in a cell is not currently an achievable goal. A recently developed technique, ribosome profiling, has emerged as a valuable approach to filling this gap, as it allows inferring global protein production via sequencing small mRNA fragments protected by the assembled ribosomes (Ingolia *et al.*, 2012). This approach has only recently been applied to the identification of actively translated transcripts in protozoan parasites *T. brucei*, *T. cruzi*, and *Plasmodium falciparum* (reviewed in Parsons and Myler, 2016), and in *L. major* (Dillon *et al.*, 2015a).

Application of whole-genome sequencing to *Leishmania* studies

Currently, whole genome sequencing (WGS) is becoming one of the most widely used applications of NGS methodologies. Through these platforms, researchers can obtain the most comprehensive view of genomic information. Using these methodologies, an increasing number of genomic sequences for *Leishmania* species is being reported (Downing *et al.*, 2011; Rogers *et al.*, 2011; Raymond *et al.*, 2012; Real *et al.*, 2013; Gupta *et al.*, 2015; Llanes *et al.*, 2015; see also the TriTryp database at www.tritryp.org).

The relevance of WGS for epidemiological purposes has been highlighted recently in an article in which genome sequences from 204 clinical isolates have been determined and used to track the evolution and epidemiology of *L. donovani* in the Indian subcontinent (Imamura *et al.*, 2016). In addition, this approach results crucial to understand how drug resistance emerges and spreads among parasite populations. These studies have furtherly demonstrated the plasticity of *Leishmania* genomes that allows these parasites to develop resistance to drugs by increasing the copy number of particular genes.

Leishmania proteomics

Overview of proteomics methodologies

The term proteome was coined to refer to the protein complement encoded in a given genome, and the process of studying the proteome is known as proteomics. The experimental objective of proteomics approaches is to detect and identify in a reproducible and robust quantitative manner as many proteins in a single sample as possible. The development of mass spectrometry (MS), which was initially an analytical technique to determine mass to charge (m/z) ratios of chemical compounds (reviewed in Aebersold and Mann, 2003), resulted pivotal for proteomics analyses. Mass spectrometry is widely used to study proteins (proteomics) and also metabolites (metabolomics, see next section). Several methods of sample analysis and types of spectrometers exist, but their description falls out of the scope of this chapter. Detailed information can be obtained elsewhere (Dunn, 2011).

Proteomics aims at defining all of the proteins present in a cell, a tissue, or an organism (or any other biological compartment) and employs large-scale, high-throughput techniques dealing with protein content, modifications, subcellular localization, and protein interactions. Initial proteomic studies were based on a combination of protein fractionation

using either one- or two-dimensional gel electrophoresis (2-DE) and MS determination of peptides derived from trypsin digestion of selected protein spots. Firstly, total protein extracts and/or subcellular fractions are separated by isoelectric focusing and afterwards in a second dimension by SDS-PAGE. For comparative purposes, images of stained gels for the different protein extracts are superimposed, spots are matched between the different gels and their intensity quantitated using specialized software. This critical and error-prone analytical step must be meticulously conducted. Many of these difficulties disappeared after the introduction of the difference gel electrophoresis (DIGE) technique (Tannu and Hemby, 2006). Although DIGE technology also relies on 2-DE electrophoresis, up to three different protein samples can be analysed in the same gel. Prior to 2-DE, protein samples are separately labelled with size-matched, charge-matched spectrally resolvable fluorescent dyes (cyanine derivatives Cy2, Cy3 and Cy5). Once spots of interest are selected and excised from 2-DE gels, proteins are fragmented into peptides for subsequent MS identification. Trypsin is the most common protease used to produce peptides, but other proteases can be also used. Peptides in the digested samples can be separated by HPLC prior to identification by tandem MS. In the first MS round, peptides are ionized and separated producing mass spectra with peaks corresponding to peptides (peptide fingerprint). Then, peptides are sequenced in the second round (fragmentation spectra).

Quantitative proteomics approaches based on 2DE-MS/MS are being replaced by shotgun proteomics, which relies exclusively on MS analysis because the former are cumbersome and expensive, especially DIGE. Currently, complex protein samples are separated by means of liquid chromatography (LC) or liquid-phase isoelectric focusing (L-IEF). These separation methods precede qualitative and quantitative analysis by MS. Liquid chromatographers are coupled to mass spectrometers in many cases, but large samples may require additional separation by L-IEF or an alternative method prior to LC-MS. This will depend on the genome size of the organism of interest and the desired coverage. Once samples have been run, a complex bioinformatics analysis is required for peptide identification, assembly of polypeptides and protein database search. Additionally, for quantification purposes, differential isotopic labelling is used; hence, different samples may be mixed and analysed together in the same run. Among the labelling methods is the Stable Isotope-Labelling by Amino acids in Culture (SILAC) approach (Ong *et al.*, 2002), which is based on the growth of cell cultures in the presence of ^{13}C6-arginine, ^{15}N4-arginine, ^{13}C6-lysine or another labelled amino acid. Then, the samples are mixed, digested and analysed by means of LC-MS/MS. Another labelling method is Isotope-coded Affinity Tag (ICAT); this technique only allows comparison of two samples because proteins are labelled on their cysteine sulfhydryl groups after protein extraction and before digestion (Smolka *et al.*, 2001). The ICAT reagent consists of an iodoacetylamide residue coupled to a linker region (light or heavy) containing a biotin residue for avidin-affinity LC coupled to MS. In this regard, Isobaric Tag for Relative and Absolute Quantitation (iTRAQ) is more advantageous because samples are labelled after digestion and up to eight samples can be pooled and simultaneously compared (Phanstiel *et al.*, 2009).

Finally, protein microarrays based on parasite proteins can be also considered a proteomic approach as they offer the potential for a high-throughput screen to determining either protein interactomes or host antibody responses when using sera from patients suffering parasitic infections (Davies *et al.*, 2005).

Proteomics studies in *Leishmania*

Although the protein-coding genes in *Leishmania* are regulated at the post-transcriptional level, mRNA abundance may not necessarily reflect protein levels, and methodologies to globally analyse gene expression at the protein level are particularly relevant for understanding the molecular dynamics in this organism. Thus, proteomics analyses are being conducted to understand different aspects of *Leishmania* biology (Cuervo *et al.*, 2010; Paape and Aebischer, 2011; Tsigankov *et al.*, 2012), such as stage differentiation, but also to study aspects of the host that result altered during the infection and disease processes (Veras and Bezerra de Menezes, 2016). Initial proteomics approaches were based on visual comparison of two-dimensional gel electrophoresis (2-DE) of proteins, followed by liquid chromatography – tandem mass spectrometry (LC-MS/MS) of the selected protein spots (Acestor *et al.*, 2002; El Fakhry *et al.*, 2002; Bente *et al.*, 2003; Drummelsmith *et al.*, 2003). When genome annotations for *Leishmania* species became available, larger numbers of proteins could be identified; conversely, proteomic data can be used as a complementary approach (referred as proteogenomics) to conventional informatics annotation (Pawar *et al.*, 2014). Additionally, technological improvements like protein fractionation by isoelectric point using free-flow electrophoresis were introduced for better resolving in the 2-DE gels of the basic proteins, which tend to be poorly represented on classical 2-DE gels compared to the acidic ones (Brotherton *et al.*, 2010). In many of these studies, it was often observed that some particular proteins were present in multiple spots, suggesting, therefore, that post-translational modifications might be extensive in *Leishmania*. Remarkably, for several proteins, different isoforms appeared to be stage-specific (McNicoll *et al.*, 2006; Dea-Ayuela *et al.*, 2006; Gupta *et al.*, 2007; Vergnes *et al.*, 2007; Alcolea *et al.*, 2011b, 2016c; Coelho *et al.*, 2012).

The use of conventional 2-DE electrophoresis for analysing relative levels of proteins has a number of limitations, mainly derived from the need of densitometry analysis of the stained gels and superimposition of images from different 2-DE runs. Some methodological approaches are being used for solving these drawbacks (see the previous section). Thus, quantitative proteomic approaches based on Isobaric Tags for Relative and Absolute quantification (iTRAQ) coupled to high-resolution mass spectrometry are being used to identify differentially expressed proteins. This method is based on the tagging of primary amines present in tryptic-digested proteins; by using different tags, it is possible multiplexing up to eight distinct samples in a single experiment. Samples are then combined at an equal ratio and subjected to LC-MS/MS and, finally, relative abundance values of all peptides attributed to a specific protein are averaged to represent the relative abundance of the entire protein. This methodology has been used for quantification of protein changes during the promastigote-to-amastigote differentiation (Rosenzweig *et al.*, 2008b; Biyani and Madhubala, 2012; Lynn *et al.*, 2013), and also for detecting protein modifications (phosphorylation, methylation, acetylation, and glycosylation sites) throughout the differentiation process (Rosenzweig *et al.*, 2008a; Tsigankov *et al.*, 2014). Also, Sardar *et al.* (2013) used iTRAQ proteome analyses to assess the effect of oxidative and nitrosative stresses on protein abundance in *L. donovani* promastigotes.

Quantitative proteomics, using the isotope-coded affinity tag (ICAT) technology and mass spectrometry (MS) were used to identify *Leishmania* stage-specific, differentially expressed proteins (Leifso *et al.*, 2007). Nevertheless, new technical improvements like label-free quantification of proteins are greatly contributing to improve comparative proteomics (Wong and Cagney, 2010). This technique is relatively simple as it relies on

computational procedures for recovering the quantitative information derived from the mass spectrometry data of peptides, obtained after trypsin enzymatic cleavage of protein extracts and subsequent separation through LC. This methodology has been recently used to identify differentially modulated proteins after hypericin treatment, an inhibitor of spermidine synthase (Singh and Dubey, 2016).

Additionally, studies aimed to identify phosphorylated proteins (phosphoproteomics) in *Leishmania* have been done by immobilized metal affinity chromatography (IMAC), separation by 2-DE, fluorescent staining and peptide identification by MS (Morales *et al.*, 2008). Analysis of the phosphoproteome has revealed that the overall phosphorylation pattern in *Leishmania* substantially changes during the differentiation process (Tsigankov *et al.*, 2014). In this regards, it should be noticed that *Leishmania*, and related trypanosomatids, possess a large set of protein kinases, comprising approximately 2% of the proteome (Parsons *et al.*, 2005). Overall, these findings suggest a key role of phosphorylation in regulating protein function and signalling pathways in these parasites. In fact, a few number of studies have been published, in which the relevance of some protein kinases begins to be envisaged. For example, temperature increase and pH decrease trigger activation of three mitogen-activated protein kinases (MAP4, 7 and 10), suggesting their involvement in signalling events during the *Leishmania* life cycle (Morales *et al.*, 2007). In another study, *L. donovani* promastigotes were engineered to express a mammalian protein tyrosine phosphatase, or were treated with inhibitors of protein tyrosine kinases. As a result, following both approaches, a partial differentiation from promastigotes to amastigotes was elicited (Nascimento *et al.*, 2003). However, we are still at the initial stages to understand the regulatory pathways involved. In fact, regulation of phosphorylation patterns and intracellular signalling cascades have not been characterized in these parasites to date, and there is no guarantee that any of the regulation pathways described in model organisms are present in these protists.

Other relevant phenotypic features of *Leishmania* that have been studied through proteomic approaches are: virulence factors (Magalhães *et al.*, 2014), antigenic proteins (Forgber *et al.*, 2006; Costa *et al.*, 2011; Coelho *et al.*, 2012), resistance to oxidative and nitrosative stresses (Alcolea *et al.*, 2016b,e). Also, proteomics techniques have been used to analyse composition of exosomes (Hassani and Olivier, 2013), glycosomes (Jamdhade *et al.*, 2015) and plasma membrane proteins (Yao *et al.*, 2010).

Metabolomics applied to the study of *Leishmania*

In order to fully understand the phenotypic characteristics of an organism, it is essential to complement genomic, transcriptomic and proteomic approaches with metabolomics studies. Metabolomics is the high-throughput characterization of the mixture of all metabolites in a biological system, both endogenous and exogenous small molecules. Metabolites are lipids, peptides, and amino, nucleic and organic acids. Metabolite levels are not easily predicted by protein or transcript levels due to the continuous flux and interconversions of metabolites occurring between the different metabolic pathways. Furthermore, particular aspects of the cellular physiology are likely impacting on both the steady state levels of metabolites and the metabolic dynamics, and consequently, measurements of the metabolome can be also used to detect subtle changes in a biological system. Metabolomics is now widely used in microbiology, nutrition, agriculture and environmental sciences, and in the clinical and pharmaceutical fields. Also, metabolomics has emerged as a key area in the study

of *Leishmania* and related trypanosomatids (t'Kindt *et al.*, 2010b; Creek *et al.*, 2012; Saunders *et al.*, 2014).

The analysis of cellular metabolomes is complicated by the chemical diversity and wide range of concentrations of different metabolite classes and by the fact that there is no single analytical platform that provides complete coverage of all cellular metabolites. However, recent advances in both NMR and hyphenated mass spectrometry, in which liquid chromatography, gas chromatography and capillary electrophoresis (CE) are interfaced with mass spectrometry, are allowing to detect and quantitate a significant proportion of a given cellular metabolome in broad untargeted analyses, at least in relative terms (Scheltema *et al.*, 2010; Berg *et al.*, 2013b). The LTQ-orbitrap mass spectrometry is also used for metabolomics purposes (t'Kindt *et al.*, 2010a). The CE-ESI-TOF-MS metabolic fingerprinting has also been used as a complementary technique to LC in order to separate metabolites (Canuto *et al.*, 2012). Additionally, stable-isotope labelling approaches result appropriate for identification of metabolite fluxes and mapping metabolic pathways. This method, for instance, has been used to characterize the changes in metabolism occurring during the promastigote-to-amastigote differentiation in *L. mexicana*, and the results showed an increased dependence of the amastigote stage on hexose and mitochondrial metabolism (Saunders *et al.*, 2014).

Metabolic pathways databases began to be established years ago, and they are an inextricable complement to metabolomics analyses. Among the many metabolic pathways databases available, several have emerged as front runners, i.e. LeishCyc (Doyle *et al.*, 2009), KEGG (Kanehisa *et al.*, 2010; www.genome.jp/kegg/), and BioCyc (Caspi *et al.*, 2010; http://metacyc.org/). Metabolic reconstructions of *L. major*, *L. infantum* and *L. braziliensis* are all available at KEGG. Network theory applied to metabolism has been the subject of a comprehensive recent review (Lacroix *et al.*, 2008). Various web servers, such as MetExplore (Cottret *et al.*, 2010) and MetaboAnalyst 2.0 (www.Metaboanalyst.ca; Xia and Wishart, 2016), are available for metabolomic data analysis and interpretation.

To date, only a few metabolomics studies of *Leishmania* have been published. These have been directed towards the analysis of changes in metabolites associated with the differentiation process in some *Leishmania* species (Silva *et al.*, 2011; Arjmand *et al.*, 2016). Interestingly, using heavy water labelling of infected BALB/c mice and metabolomics determinations, it has been possible to estimate *Leishmania* metabolic activity and growth rate *in vivo* (Kloehn *et al.*, 2015). Also, metabolomics studies have been used for the detection of metabolic differences that may correlate to drug resistance in parasite strains from clinical isolates (Scheltema *et al.*, 2010; Shaw *et al.*, 2016), for analysing the mode of drug action (Berg *et al.*, 2013a, 2015; Vincent *et al.*, 2014; Rojo *et al.*, 2015), and for the detection of possible new drug targets using *in vitro* studies (Cunha-Júnior *et al.*, 2017). Additionally, metabolomics approaches have been found useful for identification of *Leishmania* species (Westrop *et al.*, 2015). Complementary analyses based on the identification of genes coding for metabolism enzymes serve for drawing up particular biosynthetic routes (Opperdoes *et al.*, 2016). Additionally, integration of proteomics and metabolomics data has been done in order to elucidate metabolic adaptation in *Leishmania* mutant parasites (Akpunarlieva *et al.*, 2017), and integration of genomics and metabolomics analysis for elucidating differences in virulence (Alves-Ferreira *et al.*, 2015). Obviously, metabolomics approaches are also well suited for analysing relevant pathways, such as polyamine metabolism (Castilho-Martins *et al.*, 2015).

Future trends

Whole-genome sequencing is becoming every day more accessible and affordable, and soon it will be an essential tool for investigating the evolution and epidemiology of *Leishmania* and all other human infectious agents. Moreover, parallel sequencing of different strains, representative of the clinical and phenotypic diversity of a given *Leishmania* species, is revealing genome-wide sequence differences in pathogen populations (Imamura *et al.*, 2016) that might explain the phenotypic variation of the infection outcomes. Insights derived from those genomic data should aid to leishmaniasis control in disease-burdened regions as well as facilitate drug resistance diagnosis and vaccine development.

A recurrent theme emerging from some people working in the field of gene expression in *Leishmania* and related trypanosomatids is that transcriptomics data may have a marginal value for quantitation of protein expression as mRNA levels are often poorly correlated with protein or enzymatic activity. This represents a simplistic view, since every gene-encoded protein is expressed through the translation of the corresponding mRNA. Therefore, the mRNAs are central players for gene expression, which is achieved by regulation one or more of the following aspects: successful maturation of pre-mRNA (*trans*-splicing and polyadenylation, in *Leishmania*), export from the nucleus, cytoplasmic mRNA half-life, the fraction of mRNA in translation and the efficiency of that translation. In this regards, transcripts might be sequestered into ribonucleoprotein particles (or RNA granules) that ultimately control their availability to the translation machinery, storage or degradation in response to different stresses (i.e. temperature, starvation, oxidative stress) as shown in *T. brucei* (Fernández-Moya *et al.*, 2012; Kramer *et al.*, 2012, 2013). The regulation of these aspects is mediated by *cis*-elements mainly located at 5′ and 3′ untranslated region (UTR) sequences in concert with *trans*-acting protein factors. Thus, the 5′- and 3′-UTR should be considered, together with the coding regions, when defining gene models (Rastrojo *et al.*, 2013). Additionally, RNA-binding proteins (RBPs), which are responsible for the rapid gene expression remodelling that occurs during cell differentiation in *Leishmania*, and the identification of the cohorts of mRNAs they bind, must be the objects of extensive studies in the forthcoming years (Alves and Goldenberg, 2016; Kramer, 2014).

Another of the most exciting advances to look forward to is the ability to integrate data across genomics, transcriptomics, proteomics, metabolomics and perhaps other broad approaches. Pioneering studies, using this integrative approach, are being produced (Vacchina *et al.*, 2016). Current high-throughput technologies produce very large data sets and have shifted the bottleneck from data production to data analysis. Thus, great efforts have to be invested in the training of a new generation of scientists in the fields of bioinformatics and systems biology.

Although we are still faced with inadequate chemotherapy and no effective vaccines, the information that is emerging from the -omics techniques promise to alleviate this problem in the near future. For *Leishmania*, this means identifying the genes that are expressed in forms of the parasite that invade and survive in the vertebrate host. This is the case of the recently described DNA polymerase theta from *L. infantum*, which increases the infection ability of the parasite by repairing the DNA damaged by the oxidative stress that the parasite experience inside the phagolysosomal vacuole of macrophages (Fernández-Orgiler *et al.*, 2016). Similarly, metabolite profiling of both the parasite and the mammalian host during the course of infection will lead to the identification of novel diagnostic biomarkers, as well as the development of new strategies for controlling leishmaniasis.

Currently, genetic manipulation of *Leishmania* is a difficult and challenging task, mainly if essential genes are targeted (Späth and Clos, 2016). Fortunately, into the two last years, novel and very useful tools for the phenotypic analysis of essential genes have been developed. Firstly, the CRISPR/Cas9 system, which has been shown to be a powerful tool for genome editing in *Leishmania* (Sollelis *et al.*, 2015; Zhang and Matlashewski, 2015). More recently, it has been implemented an inducible gene deletion system based on a dimerized Cre recombinase (diCre) that allows the targeting of essential genes (Duncan *et al.*, 2016). Thus, these new methods will serve to directly assess whether a gene is essential to parasite viability, and therefore they will become useful tools for obtaining novel insights into the function of essential genes in *Leishmania*.

Acknowledgements

This field is already too expansive for a comprehensive, single review; thus we apologize in advance for the many omissions. Research activity is supported by grants from Proyecto del Ministerio de Economía y Competitividad (SAF2013-47556-R, co-financed with FEDER funds), and the Fondo de Investigaciones Sanitarias (ISCIII-RETIC RD12/0018/0009-FEDER). An institutional grant from the Fundación Ramón Areces is also acknowledged.

References

Acestor, N., Masina, S., Walker, J., Saravia, N.G., Fasel, N., and Quadroni, M. (2002). Establishing two-dimensional gels for the analysis of *Leishmania* proteomes. Proteomics *2*, 877–879. https://doi.org/10.1002/1615-9861(200207)2:7<877::AID-PROT877>3.0.CO;2-D.

Adaui, V., Castillo, D., Zimic, M., Gutierrez, A., Decuypere, S., Vanaerschot, M., De Doncker, S., Schnorbusch, K., Maes, I., Van der Auwera, G., *et al.* (2011). Comparative gene expression analysis throughout the life cycle of *Leishmania braziliensis*: diversity of expression profiles among clinical isolates. PLOS Negl. Trop. Dis. *5*, e1021. https://doi.org/10.1371/journal.pntd.0001021.

Aebersold, R., and Mann, M. (2003). Mass spectrometry-based proteomics. Nature *422*, 198–207. https://doi.org/10.1038/nature01511.

Akhoundi, M., Kuhls, K., Cannet, A., Votýpka, J., Marty, P., Delaunay, P., and Sereno, D. (2016). A Historical Overview of the Classification, Evolution, and Dispersion of *Leishmania* Parasites and Sandflies. PLOS Negl. Trop. Dis. *10*, e0004349. https://doi.org/10.1371/journal.pntd.0004349.

Akopyants, N.S., Matlib, R.S., Bukanova, E.N., Smeds, M.R., Brownstein, B.H., Stormo, G.D., and Beverley, S.M. (2004). Expression profiling using random genomic DNA microarrays identifies differentially expressed genes associated with three major developmental stages of the protozoan parasite *Leishmania major*. Mol. Biochem. Parasitol. *136*, 71–86. https://doi.org/10.1016/j.molbiopara.2004.03.002.

Akpunarlieva, S., Weidt, S., Lamasudin, D., Naula, C., Henderson, D., Barrett, M., Burgess, K., and Burchmore, R. (2017). Integration of proteomics and metabolomics to elucidate metabolic adaptation in *Leishmania*. J Proteomics *155*, 85–98. https://10.1016/j.jprot.2016.12.009.

Aksoy, S. (1991). Site-specific retrotransposons of the trypanosomatid protozoa. Parasitol. Today *7*, 281–285.

Alcolea, P.J., Alonso, A., Sánchez-Gorostiaga, A., Moreno-Paz, M., Gómez, M.J., Ramos, I., Parro, V., and Larraga, V. (2009). Genome-wide analysis reveals increased levels of transcripts related with infectivity in peanut lectin non-agglutinated promastigotes of *Leishmania infantum*. Genomics *93*, 551–564. https://doi.org/10.1016/j.ygeno.2009.01.007.

Alcolea, P.J., Alonso, A., Gómez, M.J., Moreno, I., Domínguez, M., Parro, V., and Larraga, V. (2010). Transcriptomics throughout the life cycle of *Leishmania infantum*: high down-regulation rate in the amastigote stage. Int. J. Parasitol. *40*, 1497–1516. https://doi.org/10.1016/j.ijpara.2010.05.013.

Alcolea, P.J., Alonso, A., and Larraga, V. (2011a). Genome-wide gene expression profile induced by exposure to cadmium acetate in *Leishmania infantum* promastigotes. Int. Microbiol. *14*, 1–11.

Alcolea, P.J., Alonso, A., and Larraga, V. (2011b). Proteome profiling of *Leishmania infantum* promastigotes. J. Eukaryot. Microbiol. *58*, 352–358. https://doi.org/10.1111/j.1550-7408.2011.00549.x.

Alcolea, P.J., Alonso, A., Gómez, M.J., Postigo, M., Molina, R., Jiménez, M., and Larraga, V. (2014). Stage-specific differential gene expression in *Leishmania infantum*: from the foregut of *Phlebotomus perniciosus* to the human phagocyte. BMC Genomics *15*, 849. https://doi.org/10.1186/1471-2164-15-849.

Alcolea, P.J., Alonso, A., Domínguez, M., Parro, V., Jiménez, M., Molina, R., and Larraga, V. (2016a). Influence of the Microenvironment in the Transcriptome of *Leishmania infantum* Promastigotes: Sand Fly versus Culture. PLOS Negl. Trop. Dis. *10*, e0004693. https://doi.org/10.1371/journal.pntd.0004693.

Alcolea, P.J., Alonso, A., García-Tabares, F., Mena, M.C., Ciordia, S., and Larraga, V. (2016b). Increased Abundance of Proteins Involved in Resistance to Oxidative and Nitrosative Stress at the Last Stages of Growth and Development of *Leishmania* amazonensis Promastigotes Revealed by Proteome Analysis. PLOS ONE *11*, e0164344. https://doi.org/10.1371/journal.pone.0164344.

Alcolea, P.J., Alonso, A., García-Tabares, F., Mena, M.C., Ciordia, S., and Larraga, V. (2016c). Proteome profiling of the growth phases of *Leishmania* pifanoi promastigotes in axenic culture reveals differential abundance of immunostimulatory proteins. Acta Trop. *158*, 240–247. https://doi.org/10.1016/j.actatropica.2016.03.015.

Alcolea, P.J., Alonso, A., Moreno-Izquierdo, M.A., Degayón, M.A., Moreno, I., and Larraga, V. (2016d). Serum Removal from Culture Induces Growth Arrest, Ploidy Alteration, Decrease in Infectivity and Differential Expression of Crucial Genes in *Leishmania infantum* Promastigotes. PLOS ONE *11*, e0150172. https://doi.org/10.1371/journal.pone.0150172.

Alcolea, P.J., Tuñón, G.I., Alonso, A., García-Tabares, F., Ciordia, S., Mena, M.C., Campos, R.N., Almeida, R.P., and Larraga, V. (2016e). Differential protein abundance in promastigotes of nitric oxide-sensitive and resistant *Leishmania chagasi* strains. Proteomics Clin Appl *10*, 1132–1146. https://doi.org/10.1002/prca.201600054.

Alonso, G., Rastrojo, A., López-Pérez, S., Requena, J.M., and Aguado, B. (2016). Resequencing and assembly of seven complex loci to improve the *Leishmania major* (Friedlin strain) reference genome. Parasit Vectors *9*, 74. https://doi.org/10.1186/s13071-016-1329-4.

Alvar, J., Vélez, I.D., Bern, C., Herrero, M., Desjeux, P., Cano, J., Jannin, J., and den Boer, M. (2012). Leishmaniasis worldwide and global estimates of its incidence. PLOS ONE *7*, e35671. https://doi.org/10.1371/journal.pone.0035671.

Alves-Ferreira, E.V., Toledo, J.S., De Oliveira, A.H., Ferreira, T.R., Ruy, P.C., Pinzan, C.F., Santos, R.F., Boaventura, V., Rojo, D., López-Gonzálvez, Á., *et al.* (2015). Differential Gene Expression and Infection Profiles of Cutaneous and Mucosal *Leishmania braziliensis* Isolates from the Same Patient. PLOS Negl. Trop. Dis. *9*, e0004018. https://doi.org/10.1371/journal.pntd.0004018.

Alves, L.R., and Goldenberg, S. (2016). RNA-binding proteins related to stress response and differentiation in protozoa. World J. Biol. Chem. *7*, 78–87. https://doi.org/10.4331/wjbc.v7.i1.78.

Arjmand, M., Madrakian, A., Khalili, G., Najafi Dastnaee, A., Zamani, Z., and Akbari, Z. (2016). Metabolomics-Based Study of Logarithmic and Stationary Phases of Promastigotes in *Leishmania major* by 1H NMR Spectroscopy. Iran. Biomed. J. *20*, 77–83.

Aslett, M., Aurrecoechea, C., Berriman, M., Brestelli, J., Brunk, B.P., Carrington, M., Depledge, D.P., Fischer, S., Gajria, B., Gao, X., *et al.* (2010). TriTrypDB: a functional genomic resource for the Trypanosomatidae. Nucleic Acids Res. *38*, D457–462. https://doi.org/10.1093/nar/gkp851.

Aurrecoechea, C., Barreto, A., Brestelli, J., Brunk, B.P., Cade, S., Doherty, R., Fischer, S., Gajria, B., Gao, X., Gingle, A., *et al.* (2013). EuPathDB: the eukaryotic pathogen database. Nucleic Acids Res. *41*, D684–691. https://doi.org/10.1093/nar/gks1113.

Balanco, J.M., Pral, E.M., da Silva, S., Bijovsky, A.T., Mortara, R.A., and Alfieri, S.C. (1998). Axenic cultivation and partial characterization of *Leishmania braziliensis* amastigote-like stages. Parasitology *116*, 103–113.

Bastien, P., Blaineau, C., and Pages, M. (1992). *Leishmania*: sex, lies and karyotype. Parasitol. Today *8*, 174–177.https://doi.org/10.1016/0169-4758(92)90016-U.

Belancio, V.P., Hedges, D.J., and Deininger, P. (2008). Mammalian non-LTR retrotransposons: for better or worse, in sickness and in health. Genome Res. *18*, 343–358. https://doi.org/10.1101/gr.5558208.

Bente, M., Harder, S., Wiesgigl, M., Heukeshoven, J., Gelhaus, C., Krause, E., Clos, J., and Bruchhaus, I. (2003). Developmentally induced changes of the proteome in the protozoan parasite *Leishmania donovani*. Proteomics *3*, 1811–1829. https://doi.org/10.1002/pmic.200300462.

Berg, M., García-Hernández, R., Cuypers, B., Vanaerschot, M., Manzano, J.I., Poveda, J.A., Ferragut, J.A., Castanys, S., Dujardin, J.C., and Gamarro, F. (2015). Experimental resistance to drug combinations in *Leishmania donovani*: metabolic and phenotypic adaptations. Antimicrob. Agents Chemother. *59*, 2242–2255. https://doi.org/10.1128/AAC.04231-14.

Berg, M., Mannaert, A., Vanaerschot, M., Van Der Auwera, G., and Dujardin, J.C. (2013a). (Post-) Genomic approaches to tackle drug resistance in *Leishmania*. Parasitology *140*, 1492–1505. https://doi.org/10.1017/S0031182013000140.

Berg, M., Vanaerschot, M., Jankevics, A., Cuypers, B., Maes, I., Mukherjee, S., Khanal, B., Rijal, S., Roy, S., Opperdoes, F., *et al.* (2013b). Metabolic adaptations of *Leishmania donovani* in relation to differentiation, drug resistance, and drug pressure. Mol. Microbiol. *90*, 428–442. https://doi.org/10.1111/mmi.12374.

Berriman, M., Ghedin, E., Hertz-Fowler, C., Blandin, G., Renauld, H., Bartholomeu, D.C., Lennard, N.J., Caler, E., Hamlin, N.E., Haas, B., *et al.* (2005). The genome of the African trypanosome *Trypanosoma brucei*. Science *309*, 416–422.

Beverley, S.M. (1991). Gene amplification in *Leishmania*. Annu. Rev. Microbiol. *45*, 417–444. https://doi.org/10.1146/annurev.mi.45.100191.002221.

Biyani, N., and Madhubala, R. (2012). Quantitative proteomic profiling of the promastigotes and the intracellular amastigotes of *Leishmania donovani* isolates identifies novel proteins having a role in *Leishmania* differentiation and intracellular survival. Biochim. Biophys. Acta *1824*, 1342–1350. https://doi.org/10.1016/j.bbapap.2012.07.010.

Blaineau, C., Bastien, P., Rioux, J.A., Roizès, G., and Pagès, M. (1991). Long-range restriction maps of size-variable homologous chromosomes in *Leishmania infantum*. Mol. Biochem. Parasitol. *46*, 292–302.

Boothroyd, J.C., Blader, I., Cleary, M., and Singh, U. (2003). DNA microarrays in parasitology: strengths and limitations. Trends Parasitol. *19*, 470–476.

Boucher, N., Wu, Y., Dumas, C., Dube, M., Sereno, D., Breton, M., and Papadopoulou, B. (2002). A common mechanism of stage-regulated gene expression in *Leishmania* mediated by a conserved 3'-untranslated region element. J. Biol. Chem. *277*, 19511–19520. https://doi.org/10.1074/jbc.M200500200.

Brandau, S., Dresel, A., and Clos, J. (1995). High constitutive levels of heat-shock proteins in human-pathogenic parasites of the genus *Leishmania*. Biochem. J. *310*, 225–232.

Bringaud, F., García-Pérez, J.L., Heras, S.R., Ghedin, E., El-Sayed, N.M., Andersson, B., Baltz, T., and Lopez, M.C. (2002). Identification of non-autonomous non-LTR retrotransposons in the genome of *Trypanosoma cruzi*. Mol. Biochem. Parasitol. *124*, 73–78.

Bringaud, F., Ghedin, E., Blandin, G., Bartholomeu, D.C., Caler, E., Levin, M.J., Baltz, T., and El-Sayed, N.M. (2006). Evolution of non-LTR retrotransposons in the trypanosomatid genomes: *Leishmania major* has lost the active elements. Mol. Biochem. Parasitol. *145*, 158–170.

Bringaud, F., Ghedin, E., El-Sayed, N.M., and Papadopoulou, B. (2008). Role of transposable elements in trypanosomatids. Microbes Infect. *10*, 575–581. https://doi.org/10.1016/j.micinf.2008.02.009.

Bringaud, F., Müller, M., Cerqueira, G.C., Smith, M., Rochette, A., El-Sayed, N.M., Papadopoulou, B., and Ghedin, E. (2007). Members of a large retroposon family are determinants of post-transcriptional gene expression in *Leishmania*. PLOS Pathog. *3*, 1291–1307.

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

Brotherton, M.C., Racine, G., Foucher, A.L., Drummelsmith, J., Papadopoulou, B., and Ouellette, M. (2010). Analysis of stage-specific expression of basic proteins in *Leishmania infantum*. J. Proteome Res. *9*, 3842–3853. https://doi.org/10.1021/pr100048m.

Bullard, W., Cliffe, L., Wang, P., Wang, Y., and Sabatini, R. (2015). Base J glucosyltransferase does not regulate the sequence specificity of J synthesis in trypanosomatid telomeric DNA. Mol. Biochem. Parasitol. *204*, 77–80. https://doi.org/10.1016/j.molbiopara.2016.01.005.

Callahan, H.L., Portal, A.C., Devereaux, R., and Grogl, M. (1997). An axenic amastigote system for drug screening. Antimicrob. Agents Chemother. *41*, 818–822.

Campbell, D.A., Sturm, N.R., and Yu, M.C. (2000). Transcription of the kinetoplastid spliced leader RNA gene. Parasitol. Today *16*, 78–82.

Cantacessi, C., Dantas-Torres, F., Nolan, M.J., and Otranto, D. (2015). The past, present, and future of *Leishmania* genomics and transcriptomics. Trends Parasitol. *31*, 100–108. https://doi.org/10.1016/j.pt.2014.12.012.

Canuto, G.A., Castilho-Martins, E.A., Tavares, M., López-Gonzálvez, A., Rivas, L., and Barbas, C. (2012). CE-ESI-MS metabolic fingerprinting of *Leishmania* resistance to antimony treatment. Electrophoresis *33*, 1901–1910. https://doi.org/10.1002/elps.201200007.

Carthew, R.W., and Sontheimer, E.J. (2009). Origins and Mechanisms of miRNAs and siRNAs. Cell *136*, 642–655. https://doi.org/10.1016/j.cell.2009.01.035.

Caspi, R., Altman, T., Dale, J.M., Dreher, K., Fulcher, C.A., Gilham, F., Kaipa, P., Karthikeyan, A.S., Kothari, A., Krummenacker, M., *et al.* (2010). The MetaCyc database of metabolic pathways and enzymes and

the BioCyc collection of pathway/genome databases. Nucleic Acids Res. *38*, D473–479. https://doi.org/10.1093/nar/gkp875.

Castilho-Martins, E.A., Canuto, G.A., Muxel, S.M., da Silva, M.F., Floeter-Winter, L.M., Del Aguila, C., Lopez-Gonzalvez, A., and Barbas, C. (2015). Capillary electrophoresis reveals polyamine metabolism modulation in *Leishmania* (*Leishmania*) *amazonensis* wild type and arginase knockout mutants under arginine starvation. Electrophoresis 36, 2314–2323. https://doi.org/10.1002/elps.201500114.

Castilla, J.J., Sanchez-Moreno, M., Mesa, C., and Osuna, A. (1995). *Leishmania donovani*: in vitro culture and [1H] NMR characterization of amastigote-like forms. Mol. Cell. Biochem. *142*, 89–97.

Chen, M.X., Ai, L., Chen, J.H., Feng, X.Y., Chen, S.H., Cai, Y.C., Lu, Y., Zhou, X.N., Chen, J.X., and Hu, W. (2016). DNA Microarray Detection of 18 Important Human Blood Protozoan Species. PLOS Negl. Trop. Dis. *10*, e0005160. https://doi.org/10.1371/journal.pntd.0005160.

Chow, C., Cloutier, S., Dumas, C., Chou, M.N., and Papadopoulou, B. (2011). Promastigote to amastigote differentiation of *Leishmania* is markedly delayed in the absence of PERK eIF$_{2\alpha}$lpha kinase-dependent eIF$_{2\alpha}$lpha phosphorylation. Cell. Microbiol. *13*, 1059–1077. https://doi.org/10.1111/j.1462-5822.2011.01602.x.

Christensen, S.M., Dillon, L.A., Carvalho, L.P., Passos, S., Novais, F.O., Hughitt, V.K., Beiting, D.P., Carvalho, E.M., Scott, P., El-Sayed, N.M., *et al.* (2016). Meta-transcriptome Profiling of the Human-*Leishmania braziliensis* Cutaneous Lesion. PLOS Negl. Trop. Dis. *10*, e0004992. https://doi.org/10.1371/journal.pntd.0004992.

Clayton, C.E. (2002). Life without transcriptional control? From fly to man and back again. EMBO J. *21*, 1881–1888. https://doi.org/10.1093/emboj/21.8.1881.

Cloutier, S., Laverdière, M., Chou, M.N., Boilard, N., Chow, C., and Papadopoulou, B. (2012). Translational control through eIF2alpha phosphorylation during the *Leishmania* differentiation process. PLOS ONE *7*, e35085. https://doi.org/10.1371/journal.pone.0035085.

Coelho, V.T., Oliveira, J.S., Valadares, D.G., Chávez-Fumagalli, M.A., Duarte, M.C., Lage, P.S., Soto, M., Santoro, M.M., Tavares, C.A., Fernandes, A.P., *et al.* (2012). Identification of proteins in promastigote and amastigote-like *Leishmania* using an immunoproteomic approach. PLOS Negl. Trop. Dis. *6*, e1430. https://doi.org/10.1371/journal.pntd.0001430.

Costa, M.M., Andrade, H.M., Bartholomeu, D.C., Freitas, L.M., Pires, S.F., Chapeaurouge, A.D., Perales, J., Ferreira, A.T., Giusta, M.S., Melo, M.N., *et al.* (2011). Analysis of *Leishmania chagasi* by 2-D difference gel electrophoresis (2-D DIGE) and immunoproteomic: identification of novel candidate antigens for diagnostic tests and vaccine. J. Proteome Res. *10*, 2172–2184. https://doi.org/10.1021/pr101286y.

Cottret, L., Wildridge, D., Vinson, F., Barrett, M.P., Charles, H., Sagot, M.F., and Jourdan, F. (2010). MetExplore: a web server to link metabolomic experiments and genome-scale metabolic networks. Nucleic Acids Res. *38*, W132–137. https://doi.org/10.1093/nar/gkq312.

Creek, D.J., Anderson, J., McConville, M.J., and Barrett, M.P. (2012). Metabolomic analysis of trypanosomatid protozoa. Mol. Biochem. Parasitol. *181*, 73–84. https://doi.org/10.1016/j.molbiopara.2011.10.003.

Cruz, A., Coburn, C.M., and Beverley, S.M. (1991). Double targeted gene replacement for creating null mutants. Proc. Natl. Acad. Sci. U.S.A. *88*, 7170–7174.

Cruz, A.K., Titus, R., and Beverley, S.M. (1993). Plasticity in chromosome number and testing of essential genes in *Leishmania* by targeting. Proc. Natl. Acad. Sci. U.S.A. *90*, 1599–1603.

Cuervo, P., Domont, G.B., and De Jesus, J.B. (2010). Proteomics of trypanosomatids of human medical importance. J Proteomics *73*, 845–867. https://doi.org/10.1016/j.jprot.2009.12.012.

Cunha-Júnior, E.F., Andrade-Neto, V.V., Lima, M.L., da Costa-Silva, T.A., Galisteo Junior, A.J., Abengózar, M.A., Barbas, C., Rivas, L., Almeida-Amaral, E.E., Tempone, A.G., *et al.* (2017). Cyclobenzaprine Raises ROS Levels in *Leishmania infantum* and Reduces Parasite Burden in Infected Mice. PLOS Negl. Trop. Dis. *11*, e0005281. https://doi.org/10.1371/journal.pntd.0005281.

da Silva, R., and Sacks, D.L. (1987). Metacyclogenesis is a major determinant of *Leishmania* promastigote virulence and attenuation. Infect. Immun. *55*, 2802–2806.

David, M., Gabdank, I., Ben-David, M., Zilka, A., Orr, I., Barash, D., and Shapira, M. (2010). Preferential translation of Hsp83 in *Leishmania* requires a thermosensitive polypyrimidine-rich element in the 3' UTR and involves scanning of the 5' UTR. RNA *16*, 364–374. https://doi.org/10.1261/rna.1874710.

Davies, D.H., Liang, X., Hernandez, J.E., Randall, A., Hirst, S., Mu, Y., Romero, K.M., Nguyen, T.T., Kalantari-Dehaghi, M., Crotty, S., *et al.* (2005). Profiling the humoral immune response to infection by using proteome microarrays: high-throughput vaccine and diagnostic antigen discovery. Proc. Natl. Acad. Sci. U.S.A. *102*, 547–552.

de Nadal, E., Ammerer, G., and Posas, F. (2011). Controlling gene expression in response to stress. Nat. Rev. Genet. *12*, 833–845. https://doi.org/10.1038/nrg3055.

De Pablos, L.M., Ferreira, T.R., and Walrad, P.B. (2016). Developmental differentiation in *Leishmania* lifecycle progression: post-transcriptional control conducts the orchestra. Curr. Opin. Microbiol. *34*, 82–89.

Dea-Ayuela, M.A., Rama-Iñiguez, S., and Bolás-Fernández, F. (2006). Proteomic analysis of antigens from *Leishmania infantum* promastigotes. Proteomics *6*, 4187–4194. https://doi.org/10.1002/pmic.200600101.

Depledge, D.P., Evans, K.J., Ivens, A.C., Aziz, N., Maroof, A., Kaye, P.M., and Smith, D.F. (2009). Comparative expression profiling of *Leishmania*: modulation in gene expression between species and in different host genetic backgrounds. PLOS Negl. Trop. Dis. *3*, e476. https://doi.org/10.1371/journal.pntd.0000476.

Dillon, L.A., Suresh, R., Okrah, K., Corrada Bravo, H., Mosser, D.M., and El-Sayed, N.M. (2015a). Simultaneous transcriptional profiling of *Leishmania major* and its murine macrophage host cell reveals insights into host-pathogen interactions. BMC Genomics *16*, 1108. https://doi.org/10.1186/s12864-015-2237-2.

Dillon, L.A., Okrah, K., Hughitt, V.K., Suresh, R., Li, Y., Fernandes, M.C., Belew, A.T., Corrada Bravo, H., Mosser, D.M., and El-Sayed, N.M. (2015b). Transcriptomic profiling of gene expression and RNA processing during *Leishmania major* differentiation. Nucleic Acids Res. *43*, 6799–6813. https://doi.org/10.1093/nar/gkv656.

Dooijes, D., Chaves, I., Kieft, R., Dirks-Mulder, A., Martin, W., and Borst, P. (2000). Base J originally found in kinetoplastida is also a minor constituent of nuclear DNA of *Euglena gracilis*. Nucleic Acids Res. *28*, 3017–3021.

Dostálová, A., and Volf, P. (2012). *Leishmania* development in sand flies: parasite-vector interactions overview. Parasit. Vectors *5*, 276. https://doi.org/10.1186/1756-3305-5-276.

Downing, T., Imamura, H., Decuypere, S., Clark, T.G., Coombs, G.H., Cotton, J.A., Hilley, J.D., de Doncker, S., Maes, I., Mottram, J.C., *et al.* (2011). Whole genome sequencing of multiple *Leishmania donovani* clinical isolates provides insights into population structure and mechanisms of drug resistance. Genome Res. *21*, 2143–2156. https://doi.org/10.1101/gr.123430.111.

Doyle, M.A., MacRae, J.I., De Souza, D.P., Saunders, E.C., McConville, M.J., and Likić, V.A. (2009). LeishCyc: a biochemical pathways database for *Leishmania major*. BMC Syst. Biol. *3*, 57. https://doi.org/10.1186/1752-0509-3-57.

Droll, D., Minia, I., Fadda, A., Singh, A., Stewart, M., Queiroz, R., and Clayton, C. (2013). Post-transcriptional regulation of the trypanosome heat shock response by a zinc finger protein. PLOS Pathog. *9*, e1003286. https://doi.org/10.1371/journal.ppat.1003286.

Drummelsmith, J., Brochu, V., Girard, I., Messier, N., and Ouellette, M. (2003). Proteome mapping of the protozoan parasite *Leishmania* and application to the study of drug targets and resistance mechanisms. Mol. Cell Proteomics *2*, 146–155. https://doi.org/10.1074/mcp.M200085-MCP200.

Dubessay, P., Ravel, C., Bastien, P., Crobu, L., Dedet, J.P., Pagès, M., and Blaineau, C. (2002a). The switch region on *Leishmania major* chromosome 1 is not required for mitotic stability or gene expression, but appears to be essential. Nucleic Acids Res. *30*, 3692–3697.

Dubessay, P., Ravel, C., Bastien, P., Stuart, K., Dedet, J.P., Blaineau, C., and Pagès, M. (2002b). Mitotic stability of a coding DNA sequence-free version of *Leishmania major* chromosome 1 generated by targeted chromosome fragmentation. Gene *289*, 151–159.

Dujardin, J.C., Campino, L., Cañavate, C., Dedet, J.P., Gradoni, L., Soteriadou, K., Mazeris, A., Ozbel, Y., and Boelaert, M. (2008). Spread of vector-borne diseases and neglect of Leishmaniasis, Europe. Emerging Infect. Dis. *14*, 1013–1018. https://doi.org/10.3201/eid1407.071589.

Dumas, C., Ouellette, M., Tovar, J., Cunningham, M.L., Fairlamb, A.H., Tamar, S., Olivier, M., and Papadopoulou, B. (1997). Disruption of the trypanothione reductase gene of *Leishmania* decreases its ability to survive oxidative stress in macrophages. EMBO J. *16*, 2590–2598. https://doi.org/10.1093/emboj/16.10.2590.

Duncan, R. (2004). DNA microarray analysis of protozoan parasite gene expression: outcomes correlate with mechanisms of regulation. Trends Parasitol. *20*, 211–215. https://doi.org/10.1016/j.pt.2004.02.008.

Duncan, S.M., Myburgh, E., Philipon, C., Brown, E., Meissner, M., Brewer, J., and Mottram, J.C. (2016). Conditional gene deletion with DiCre demonstrates an essential role for CRK3 in *Leishmania mexicana* cell cycle regulation. Mol. Microbiol. *100*, 931–944. https://doi.org/10.1111/mmi.13375.

Dunn, W.B. (2011). Mass spectrometry in systems biology an introduction. Meth. Enzymol. *500*, 15–35. https://doi.org/10.1016/B978-0-12-385118-5.00002-5.

Eid, J., Fehr, A., Gray, J., Luong, K., Lyle, J., Otto, G., Peluso, P., Rank, D., Baybayan, P., Bettman, B., *et al.* (2009). Real-time DNA sequencing from single polymerase molecules. Science *323*, 133–138. https://doi.org/10.1126/science.1162986.

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., *et al.* (2005). The genome sequence of *Trypanosoma cruzi*, etiologic agent of Chagas disease. Science *309*, 409–415.

El Fakhry, Y., Ouellette, M., and Papadopoulou, B. (2002). A proteomic approach to identify developmentally regulated proteins in *Leishmania infantum*. Proteomics *2*, 1007–1017. https://doi.org/10.1002/1615-9861(200208)2:8<1007::AID-PROT1007>3.0.CO;2-G.

Eliaz, D., Doniger, T., Tkacz, I.D., Biswas, V.K., Gupta, S.K., Kolev, N.G., Unger, R., Ullu, E., Tschudi, C., and Michaeli, S. (2015). Genome-wide analysis of small nucleolar RNAs of *Leishmania major* reveals a rich repertoire of RNAs involved in modification and processing of rRNA. RNA Biol *12*, 1222–1255. https://doi.org/10.1080/15476286.2015.1038019.

Fernandes, M.C., Dillon, L.A., Belew, A.T., Bravo, H.C., Mosser, D.M., and El-Sayed, N.M. (2016). Dual Transcriptome Profiling of *Leishmania*-Infected Human Macrophages Reveals Distinct Reprogramming Signatures. MBio *7*, e00027–16. https://doi.org/10.1128/mBio.00027-16.

Fernández-Moya, S.M., García-Pérez, A., Kramer, S., Carrington, M., and Estévez, A.M. (2012). Alterations in DRBD3 ribonucleoprotein complexes in response to stress in *Trypanosoma brucei*. PLOS ONE *7*, e48870. https://doi.org/10.1371/journal.pone.0048870.

Fernández-Orgiler, A., Martínez-Jiménez, M.I., Alonso, A., Alcolea, P.J., Requena, J.M., Thomas, M.C., Blanco, L., and Larraga, V. (2016). A putative *Leishmania* DNA polymerase theta protects the parasite against oxidative damage. Nucleic Acids Res. *44*, 4855–4870. https://doi.org/10.1093/nar/gkw346.

Fiebig, M., Kelly, S., and Gluenz, E. (2015). Comparative Life Cycle Transcriptomics Revises *Leishmania mexicana* Genome Annotation and Links a Chromosome Duplication with Parasitism of Vertebrates. PLOS Pathog. *11*, e1005186. https://doi.org/10.1371/journal.ppat.1005186.

Folgueira, C., Cañavate, C., Chicharro, C., and Requena, J.M. (2007). Genomic organization and expression of the HSP70 locus in New and Old World *Leishmania* species. Parasitology *134*, 369–377. https://doi.org/10.1017/S0031182006001570.

Folgueira, C., Quijada, L., Soto, M., Abanades, D.R., Alonso, C., and Requena, J.M. (2005). The translational efficiencies of the two *Leishmania infantum* HSP70 mRNAs, differing in their 3'-untranslated regions, are affected by shifts in the temperature of growth through different mechanisms. J. Biol. Chem. *280*, 35172–35183.

Folgueira, C., and Requena, J.M. (2007). A postgenomic view of the heat shock proteins in kinetoplastids. FEMS Microbiol. Rev. *31*, 359–377.

Forgber, M., Basu, R., Roychoudhury, K., Theinert, S., Roy, S., Sundar, S., and Walden, P. (2006). Mapping the antigenicity of the parasites in *Leishmania donovani* infection by proteome serology. PLOS ONE *1*, e40. https://doi.org/10.1371/journal.pone.0000040.

Genest, P.A., ter Riet, B., Dumas, C., Papadopoulou, B., van Luenen, H.G., and Borst, P. (2005). Formation of linear inverted repeat amplicons following targeting of an essential gene in *Leishmania*. Nucleic Acids Res. *33*, 1699–1709.

Genois, M.M., Plourde, M., Éthier, C., Roy, G., Poirier, G.G., Ouellette, M., and Masson, J.Y. (2015). Roles of Rad51 paralogs for promoting homologous recombination in *Leishmania infantum*. Nucleic Acids Res. *43*, 2701–2715. https://doi.org/10.1093/nar/gkv118.

Geraci, N.S., Tan, J.C., and McDowell, M.A. (2015). Characterization of microRNA expression profiles in *Leishmania*-infected human phagocytes. Parasite Immunol. *37*, 43–51. https://doi.org/10.1111/pim.12156.

Ghedin, E., Bringaud, F., Peterson, J., Myler, P., Berriman, M., Ivens, A., Andersson, B., Bontempi, E., Eisen, J., Angiuoli, S., *et al.* (2004). Gene synteny and evolution of genome architecture in trypanosomatids. Mol. Biochem. Parasitol. *134*, 183–191. https://doi.org/10.1016/j.molbiopara.2003.11.012.

Gilinger, G., and Bellofatto, V. (2001). Trypanosome spliced leader RNA genes contain the first identified RNA polymerase II gene promoter in these organisms. Nucleic Acids Res. *29*, 1556–1564.

Goodwin, S., McPherson, J.D., and McCombie, W.R. (2016). Coming of age: ten years of next-generation sequencing technologies. Nat. Rev. Genet. *17*, 333–351. https://doi.org/10.1038/nrg.2016.49.

Grondin, K., Papadopoulou, B., and Ouellette, M. (1993). Homologous recombination between direct repeat sequences yields P-glycoprotein containing amplicons in arsenite resistant *Leishmania*. Nucleic Acids Res. *21*, 1895–1901.

Grondin, K., Roy, G., and Ouellette, M. (1996). Formation of extrachromosomal circular amplicons with direct or inverted duplications in drug-resistant *Leishmania* tarentolae. Mol. Cell. Biol. *16*, 3587–3595.

Guimond, C., Trudel, N., Brochu, C., Marquis, N., El Fadili, A., Peytavi, R., Briand, G., Richard, D., Messier, N., Papadopoulou, B., *et al.* (2003). Modulation of gene expression in *Leishmania* drug resistant mutants as determined by targeted DNA microarrays. Nucleic Acids Res. *31*, 5886–5896.

Gupta, A.K., Srivastava, S., Singh, A., and Singh, S. (2015). De Novo Whole-Genome Sequence and Annotation of a *Leishmania* Strain Isolated from a Case of Post-Kala-Azar Dermal Leishmaniasis. Genome Announc. *3*, e00809–15. https://doi.org/10.1128/genomeA.00809-15.

Gupta, N., Goyal, N., and Rastogi, A.K. (2001). In vitro cultivation and characterization of axenic amastigotes of *Leishmania*. Trends Parasitol. *17*, 150–153.

Gupta, S.K., Sisodia, B.S., Sinha, S., Hajela, K., Naik, S., Shasany, A.K., and Dube, A. (2007). Proteomic approach for identification and characterization of novel immunostimulatory proteins from soluble antigens of *Leishmania donovani* promastigotes. Proteomics *7*, 816–823. https://doi.org/10.1002/pmic.200600725.

Hasan, G., Turner, M.J., and Cordingley, J.S. (1984). Complete nucleotide sequence of an unusual mobile element from *Trypanosoma brucei*. Cell *37*, 333–341.

Hassani, K., and Olivier, M. (2013). Immunomodulatory impact of leishmania-induced macrophage exosomes: a comparative proteomic and functional analysis. PLOS Negl. Trop. Dis. *7*, e2185. https://doi.org/10.1371/journal.pntd.0002185.

Heras, S.R., López, M.C., Olivares, M., and Thomas, M.C. (2007). The L1Tc non-LTR retrotransposon of *Trypanosoma cruzi* contains an internal RNA-pol II-dependent promoter that strongly activates gene transcription and generates unspliced transcripts. Nucleic Acids Res. *35*, 2199–2214.

Hernández, R., and Cevallos, A.M. (2014). Ribosomal RNA gene transcription in trypanosomes. Parasitol. Res. *113*, 2415–2424. https://doi.org/10.1007/s00436-014-3940-7.

Herrador, Z., Gherasim, A., Jimenez, B.C., Granados, M., Granados, M., San Martín, J.V., and Aparicio, P. (2015). Epidemiological changes in leishmaniasis in Spain according to hospitalization-based records, 1997-2011: raising awareness towards leishmaniasis in non-HIV patients. PLOS Negl. Trop. Dis. *9*, e0003594. https://doi.org/10.1371/journal.pntd.0003594.

Holzer, T.R., McMaster, W.R., and Forney, J.D. (2006). Expression profiling by whole-genome interspecies microarray hybridization reveals differential gene expression in procyclic promastigotes, lesion-derived amastigotes, and axenic amastigotes in *Leishmania mexicana*. Mol. Biochem. Parasitol. *146*, 198–218.

Ilgoutz, S.C., Zawadzki, J.L., Ralton, J.E., and McConville, M.J. (1999). Evidence that free GPI glycolipids are essential for growth of *Leishmania mexicana*. EMBO J. *18*, 2746–2755. https://doi.org/10.1093/emboj/18.10.2746.

Imamura, H., Downing, T., Van den Broeck, F., Sanders, M.J., Rijal, S., Sundar, S., Mannaert, A., Vanaerschot, M., Berg, M., De Muylder, G., *et al.* (2016). Evolutionary genomics of epidemic visceral leishmaniasis in the Indian subcontinent. Elife *5*, e12613. https://doi.org/10.7554/eLife.12613.

Ingolia, N.T., Brar, G.A., Rouskin, S., McGeachy, A.M., and Weissman, J.S. (2012). The ribosome profiling strategy for monitoring translation in vivo by deep sequencing of ribosome-protected mRNA fragments. Nat Protoc *7*, 1534–1550. https://doi.org/10.1038/nprot.2012.086.

Iovannisci, D.M., and 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., *et al.* (2005). The genome of the kinetoplastid parasite, *Leishmania major*. Science *309*, 436–442.

Jamdhade, M.D., Pawar, H., Chavan, S., Sathe, G., Umasankar, P.K., Mahale, K.N., Dixit, T., Madugundu, A.K., Prasad, T.S., Gowda, H., *et al.* (2015). Comprehensive proteomics analysis of glycosomes from *Leishmania donovani*. OMICS *19*, 157–170. https://doi.org/10.1089/omi.2014.0163.

Kanehisa, M., Goto, S., Furumichi, M., Tanabe, M., and Hirakawa, M. (2010). KEGG for representation and analysis of molecular networks involving diseases and drugs. Nucleic Acids Res. *38*, D355–360. https://doi.org/10.1093/nar/gkp896.

Kloehn, J., Saunders, E.C., O'Callaghan, S., Dagley, M.J., and McConville, M.J. (2015). Characterization of metabolically quiescent *Leishmania* parasites in murine lesions using heavy water labeling. PLOS Pathog. *11*, e1004683. https://doi.org/10.1371/journal.ppat.1004683.

Kramer, S. (2014). RNA in development: how ribonucleoprotein granules regulate the life cycles of pathogenic protozoa. Wiley Interdiscip Rev RNA *5*, 263–284. https://doi.org/10.1002/wrna.1207.

Kramer, S., Bannerman-Chukualim, B., Ellis, L., Boulden, E.A., Kelly, S., Field, M.C., and Carrington, M. (2013). Differential localization of the two T. brucei poly(A) binding proteins to the nucleus and RNP granules suggests binding to distinct mRNA pools. PLOS ONE *8*, e54004. https://doi.org/10.1371/journal.pone.0054004.

Kramer, S., and Carrington, M. (2011). Trans-acting proteins regulating mRNA maturation, stability and translation in trypanosomatids. Trends Parasitol. *27*, 23–30. https://doi.org/10.1016/j.pt.2010.06.011.

Kramer, S., Marnef, A., Standart, N., and Carrington, M. (2012). Inhibition of mRNA maturation in trypanosomes causes the formation of novel foci at the nuclear periphery containing cytoplasmic regulators of mRNA fate. J. Cell. Sci. *125*, 2896–2909. https://doi.org/10.1242/jcs.099275.

Lacroix, V., Cottret, L., Thébault, P., and Sagot, M.F. (2008). An introduction to metabolic networks and their structural analysis. IEEE/ACM Trans Comput Biol Bioinform *5*, 594–617. https://doi.org/10.1109/TCBB.2008.79.

Lachaud, L., Bourgeois, N., Kuk, N., Morelle, C., Crobu, L., Merlin, G., Bastien, P., Pagès, M., and Sterkers, Y. (2014). Constitutive mosaic aneuploidy is a unique genetic feature widespread in the *Leishmania* genus. Microbes Infect. *16*, 61–66. https://doi.org/10.1016/j.micinf.2013.09.005.

Laffitte, M.C., Genois, M.M., Mukherjee, A., Légaré, D., Masson, J.Y., and Ouellette, M. (2014). Formation of linear amplicons with inverted duplications in *Leishmania* requires the MRE11 nuclease. PLOS Genet. *10*, e1004805. https://doi.org/10.1371/journal.pgen.1004805.

Lahav, T., Sivam, D., Volpin, H., Ronen, M., Tsigankov, P., Green, A., Holland, N., Kuzyk, M., Borchers, C., Zilberstein, D., *et al.* (2011). Multiple levels of gene regulation mediate differentiation of the intracellular pathogen *Leishmania*. FASEB J. *25*, 515–525. https://doi.org/10.1096/fj.10-157529.

Larreta, R., Soto, M., Quijada, L., Folgueira, C., Abanades, D.R., Alonso, C., and Requena, J.M. (2004). The expression of HSP83 genes in *Leishmania infantum* is affected by temperature and by stage-differentiation and is regulated at the levels of mRNA stability and translation. BMC Mol. Biol. *5*, 3. https://doi.org/10.1186/1471-2199-5-3.

LeBowitz, J.H., Smith, H.Q., Rusche, L., and Beverley, S.M. (1993). Coupling of poly(A) site selection and trans-splicing in *Leishmania*. Genes Dev. *7*, 996–1007.

Leifso, K., Cohen-Freue, G., Dogra, N., Murray, A., and McMaster, W.R. (2007). Genomic and proteomic expression analysis of *Leishmania* promastigote and amastigote life stages: the *Leishmania* genome is constitutively expressed. Mol. Biochem. Parasitol. *152*, 35–46.

Leprohon, P., Légaré, D., Raymond, F., Madore, E., Hardiman, G., Corbeil, J., and Ouellette, M. (2009). Gene expression modulation is associated with gene amplification, supernumerary chromosomes and chromosome loss in antimony-resistant *Leishmania infantum*. Nucleic Acids Res. *37*, 1387–1399. https://doi.org/10.1093/nar/gkn1069.

Liang, X.H., Haritan, A., Uliel, S., and Michaeli, S. (2003). trans and cis splicing in trypanosomatids: mechanism, factors, and regulation. Eukaryotic Cell *2*, 830–840.

Llanes, A., Restrepo, C.M., Del Vecchio, G., Anguizola, F.J., and Lleonart, R. (2015). The genome of *Leishmania* panamensis: insights into genomics of the L. (Viannia) subgenus. Sci Rep *5*, 8550. https://doi.org/10.1038/srep08550.

Lombraña, R., Álvarez, A., Fernández-Justel, J.M., Almeida, R., Poza-Carrión, C., Gomes, F., Calzada, A., Requena, J.M., and Gómez, M. (2016). Transcriptionally Driven DNA Replication Program of the Human Parasite *Leishmania major*. Cell Rep *16*, 1774–1786.

Lye, L.F., Owens, K., Shi, H., Murta, S.M., Vieira, A.C., Turco, S.J., Tschudi, C., Ullu, E., and Beverley, S.M. (2010). Retention and loss of RNA interference pathways in trypanosomatid protozoans. PLOS Pathog. *6*, e1001161. https://doi.org/10.1371/journal.ppat.1001161.

Lynn, M.A., Marr, A.K., and McMaster, W.R. (2013). Differential quantitative proteomic profiling of *Leishmania infantum* and *Leishmania mexicana* density gradient separated membranous fractions. J Proteomics *82*, 179–192. https://doi.org/10.1016/j.jprot.2013.02.010.

Magalhães, R.D., Duarte, M.C., Mattos, E.C., Martins, V.T., Lage, P.S., Chávez-Fumagalli, M.A., Lage, D.P., Menezes-Souza, D., Régis, W.C., Manso Alves, M.J., *et al.* (2014). Identification of differentially expressed proteins from *Leishmania* amazonensis associated with the loss of virulence of the parasites. PLOS Negl. Trop. Dis. *8*, e2764. https://doi.org/10.1371/journal.pntd.0002764.

Martin, J.A., and Wang, Z. (2011). Next-generation transcriptome assembly. Nat. Rev. Genet. *12*, 671–682. https://doi.org/10.1038/nrg3068.

Martínez-Calvillo, S., Stuart, K., and Myler, P.J. (2005). Ploidy changes associated with disruption of two adjacent genes on *Leishmania major* chromosome 1. Int. J. Parasitol. *35*, 419–429.

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

Matthews, K.R., Tschudi, C., and Ullu, E. (1994). A common pyrimidine-rich motif governs trans-splicing and polyadenylation of tubulin polycistronic pre-mRNA in trypanosomes. Genes Dev. *8*, 491–501.

McConville, M.J., and Naderer, T. (2011). Metabolic pathways required for the intracellular survival of *Leishmania*. Annu. Rev. Microbiol. *65*, 543–561. https://doi.org/10.1146/annurev-micro-090110-102913.

McGwire, B.S., and Satoskar, A.R. (2014). Leishmaniasis: clinical syndromes and treatment. QJM *107*, 7–14. https://doi.org/10.1093/qjmed/hct116.

McNicoll, F., Drummelsmith, J., Müller, M., Madore, E., Boilard, N., Ouellette, M., and Papadopoulou, B. (2006). A combined proteomic and transcriptomic approach to the study of stage differentiation in *Leishmania infantum*. Proteomics *6*, 3567–3581. https://doi.org/10.1002/pmic.200500853.

McNicoll, F., Müller, M., Cloutier, S., Boilard, N., Rochette, A., Dubé, M., and Papadopoulou, B. (2005). Distinct 3'-untranslated region elements regulate stage-specific mRNA accumulation and translation in *Leishmania*. J. Biol. Chem. *280*, 35238–35246.

Moore, M.J. (2005). From birth to death: the complex lives of eukaryotic mRNAs. Science *309*, 1514–1518.

Morales, M.A., Renaud, O., Faigle, W., Shorte, S.L., and Späth, G.F. (2007). Over-expression of *Leishmania major* MAP kinases reveals stage-specific induction of phosphotransferase activity. Int. J. Parasitol. *37*, 1187–1199.

Morales, M.A., Watanabe, R., Dacher, M., Chafey, P., Osorio y Fortéa, J., Scott, D.A., Beverley, S.M., Ommen, G., Clos, J., Hem, S., *et al.* (2010). Phosphoproteome dynamics reveal heat-shock protein complexes specific to the *Leishmania donovani* infectious stage. Proc. Natl. Acad. Sci. U.S.A. *107*, 8381–8386. https://doi.org/10.1073/pnas.0914768107.

Morales, M.A., Watanabe, R., Laurent, C., Lenormand, P., Rousselle, J.C., Namane, A., and Späth, G.F. (2008). Phosphoproteomic analysis of *Leishmania donovani* pro- and amastigote stages. Proteomics *8*, 350–363. https://doi.org/10.1002/pmic.200700697.

Moreira, D., Lopez-Garcia, P., and Vickerman, K. (2004). An updated view of kinetoplastid phylogeny using environmental sequences and a closer outgroup: proposal for a new classification of the class Kinetoplastea. Int. J. Syst. Evol. Microbiol. *54*, 1861–1875. https://doi.org/10.1099/ijs.0.63081-0.

Moreno, J., and Alvar, J. (2002). Canine leishmaniasis: epidemiological risk and the experimental model. Trends Parasitol. *18*, 399–405.

Müller, M., Padmanabhan, P.K., and Papadopoulou, B. (2010a). Selective inactivation of SIDER2 retroposon-mediated mRNA decay contributes to stage- and species-specific gene expression in *Leishmania*. Mol. Microbiol. *77*, 471–491. https://doi.org/10.1111/j.1365-2958.2010.07226.x.

Müller, M., Padmanabhan, P.K., Rochette, A., Mukherjee, D., Smith, M., Dumas, C., and Papadopoulou, B. (2010b). Rapid decay of unstable *Leishmania* mRNAs bearing a conserved retroposon signature 3′-UTR motif is initiated by a site-specific endonucleolytic cleavage without prior deadenylation. Nucleic Acids Res. *38*, 5867–5883. https://doi.org/10.1093/nar/gkq349.

Murta, S.M., Vickers, T.J., Scott, D.A., and Beverley, S.M. (2009). Methylene tetrahydrofolate dehydrogenase/cyclohydrolase and the synthesis of 10-CHO-THF are essential in *Leishmania major*. Mol. Microbiol. *71*, 1386–1401. https://doi.org/10.1111/j.1365-2958.2009.06610.x.

Myung, K.S., Beetham, J.K., Wilson, M.E., and Donelson, J.E. (2002). Comparison of the post-transcriptional regulation of the mRNAs for the surface proteins PSA (GP46) and MSP (GP63) of *Leishmania chagasi*. J. Biol. Chem. *277*, 16489–16497. https://doi.org/10.1074/jbc.M200174200.

Naderer, T., Dandash, O., and McConville, M.J. (2011). Calcineurin is required for *Leishmania major* stress response pathways and for virulence in the mammalian host. Mol. Microbiol. *80*, 471–480. https://doi.org/10.1111/j.1365-2958.2011.07584.x.

Nascimento, M., Abourjeily, N., Ghosh, A., Zhang, W.W., and Matlashewski, G. (2003). Heterologous expression of a mammalian protein tyrosine phosphatase gene in *Leishmania*: effect on differentiation. Mol. Microbiol. *50*, 1517–1526.

Ngô, H., Tschudi, C., Gull, K., and Ullu, E. (1998). Double-stranded RNA induces mRNA degradation in *Trypanosoma brucei*. Proc. Natl. Acad. Sci. U.S.A. *95*, 14687–14692.

Novais, F.O., Carvalho, L.P., Passos, S., Roos, D.S., Carvalho, E.M., Scott, P., and Beiting, D.P. (2015). Genomic profiling of human *Leishmania braziliensis* lesions identifies transcriptional modules associated with cutaneous immunopathology. J. Invest. Dermatol. *135*, 94–101.

Ong, S.E., Blagoev, B., Kratchmarova, I., Kristensen, D.B., Steen, H., Pandey, A., and Mann, M. (2002). Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. Mol. Cell Proteomics *1*, 376–386.

Opperdoes, F.R., Butenko, A., Flegontov, P., Yurchenko, V., and Lukeš, J. (2016). Comparative Metabolism of Free-living *Bodo saltans* and Parasitic Trypanosomatids. J. Eukaryot. Microbiol. *63*, 657–678. https://doi.org/10.1111/jeu.12315.

Paape, D., and Aebischer, T. (2011). Contribution of proteomics of *Leishmania* spp. to the understanding of differentiation, drug resistance mechanisms, vaccine and drug development. J Proteomics *74*, 1614–1624. https://doi.org/10.1016/j.jprot.2011.05.005.

Papadopoulou, B., Muller, M., Rochette, A., McNicoll, F., Dumas, C., and Chow, C. (2008). Regulation of gene expression in *Leishmania* throughout a complex digenetic life cycle. In: *Leishmania*, after the genome, P.J. Myler and N. Fasel, eds. (Caister Academic Press, Norfolk), pp. 29–54.

Parsons, M., and Myler, P.J. (2016). Illuminating Parasite Protein Production by Ribosome Profiling. Trends Parasitol. *32*, 446–457. https://doi.org/10.1016/j.pt.2016.03.005.

Parsons, M., Worthey, E.A., Ward, P.N., and Mottram, J.C. (2005). Comparative analysis of the kinomes of three pathogenic trypanosomatids: *Leishmania major*, *Trypanosoma brucei* and *Trypanosoma cruzi*. BMC Genomics *6*, 127.

Pawar, H., Renuse, S., Khobragade, S.N., Chavan, S., Sathe, G., Kumar, P., Mahale, K.N., Gore, K., Kulkarni, A., Dixit, T., *et al.* (2014). Neglected tropical diseases and omics science: proteogenomics analysis of the promastigote stage of *Leishmania major* parasite. OMICS *18*, 499–512. https://doi.org/10.1089/omi.2013.0159.

Peacock, C.S., Seeger, K., Harris, D., Murphy, L., Ruiz, J.C., Quail, M.A., Peters, N., Adlem, E., Tivey, A., Aslett, M., *et al.* (2007). Comparative genomic analysis of three *Leishmania* species that cause diverse human disease. Nat. Genet. *39*, 839–847.

Pescher, P., Blisnick, T., Bastin, P., and Späth, G.F. (2011). Quantitative proteome profiling informs on phenotypic traits that adapt *Leishmania donovani* for axenic and intracellular proliferation. Cell. Microbiol. *13*, 978–991. https://doi.org/10.1111/j.1462-5822.2011.01593.x.

Phanstiel, D., Unwin, R., McAlister, G.C., and Coon, J.J. (2009). Peptide quantification using 8-plex isobaric tags and electron transfer dissociation tandem mass spectrometry. Anal. Chem. *81*, 1693–1698. https://doi.org/10.1021/ac8019202.

Pigott, D.M., Bhatt, S., Golding, N., Duda, K.A., Battle, K.E., Brady, O.J., Messina, J.P., Balard, Y., Bastien, P., Pratlong, F., *et al.* (2014). Global distribution maps of the leishmaniases. Elife *3*. https://doi.org/10.7554/eLife.02851.

Quijada, L., Soto, M., Alonso, C., and Requena, J.M. (1997). Analysis of post-transcriptional regulation operating on transcription products of the tandemly linked *Leishmania infantum* hsp70 genes. J. Biol. Chem. *272*, 4493–4499.

Quijada, L., Soto, M., Alonso, C., and Requena, J.M. (2000). Identification of a putative regulatory element in the 3'-untranslated region that controls expression of HSP70 in *Leishmania infantum*. Mol. Biochem. Parasitol. *110*, 79–91.

Ramírez, C.A., Requena, J.M., and Puerta, C.J. (2011). Identification of the HSP70-II gene in *Leishmania braziliensis* HSP70 locus: genomic organization and UTRs characterization. Parasit. Vectors *4*, 166. https://doi.org/10.1186/1756-3305-4-166.

Rastrojo, A., Carrasco-Ramiro, F., Martín, D., Crespillo, A., Reguera, R.M., Aguado, B., and Requena, J.M. (2013). The transcriptome of *Leishmania major* in the axenic promastigote stage: transcript annotation and relative expression levels by RNA-seq. BMC Genomics *14*, 223. https://doi.org/10.1186/1471-2164-14-223.

Ravel, C., Dubessay, P., Bastien, P., Blackwell, J.M., and Ivens, A.C. (1998). The complete chromosomal organization of the reference strain of the *Leishmania* Genome Project, *L. major* 'Friedlin'. Parasitol. Today *14*, 301–303.

Raymond, F., Boisvert, S., Roy, G., Ritt, J.F., Légaré, D., Isnard, A., Stanke, M., Olivier, M., Tremblay, M.J., Papadopoulou, B., *et al.* (2012). Genome sequencing of the lizard parasite *Leishmania* tarentolae reveals loss of genes associated to the intracellular stage of human pathogenic species. Nucleic Acids Res. *40*, 1131–1147. https://doi.org/10.1093/nar/gkr834.

Real, F., Vidal, R.O., Carazzolle, M.F., Mondego, J.M., Costa, G.G., Herai, R.H., Würtele, M., de Carvalho, L.M., Carmona e Ferreira, R., Mortara, R.A., *et al.* (2013). The genome sequence of *Leishmania* (*Leishmania*) amazonensis: functional annotation and extended analysis of gene models. DNA Res. *20*, 567–581. https://doi.org/10.1093/dnares/dst031.

Requena, J.M. (2011). Lights and shadows on gene organization and regulation of gene expression in *Leishmania*. Front. Biosci. *16*, 2069–2085.

Requena, J.M. (2012). The Stressful Life of pathogenic *Leishmania* species. In: Stress Response in Microbiology, J.M. Requena, ed. (Caister Academic Press, Norfolk), pp. 323–346.

Requena, J.M., Folgueira, C., López, M.C., and Thomas, M.C. (2008). The SIDER2 elements, interspersed repeated sequences that populate the *Leishmania* genomes, constitute subfamilies showing chromosomal proximity relationship. BMC Genomics *9*, 263. https://doi.org/10.1186/1471-2164-9-263.

Requena, J.M., Quijada, L., Soto, M., and Alonso, C. (2003). Conserved nucleotides surrounding the trans-splicing acceptor site and the translation initiation codon in *Leishmania* genes. Exp. Parasitol. *103*, 78–81.

Requena, J.M., Rastrojo, A., Garde, E., López, M.C., Thomas, M.C., and Aguado, B. (2017). Genomic cartography and proposal of nomenclature for the repeated, interspersed elements of the *Leishmania major* SIDER2 family and identification of SIDER2-containing transcripts. Mol. Biochem. Parasitol. *212*, 9–15.

Respuela, P., Ferella, M., Rada-Iglesias, A., and Aslund, L. (2008). Histone acetylation and methylation at sites initiating divergent polycistronic transcription in *Trypanosoma cruzi*. J. Biol. Chem. *283*, 15884–15892. https://doi.org/10.1074/jbc.M802081200.

Reynolds, D.L., Hofmeister, B.T., Cliffe, L., Siegel, T.N., Anderson, B.A., Beverley, S.M., Schmitz, R.J., and Sabatini, R. (2016). Base J represses genes at the end of polycistronic gene clusters in *Leishmania major* by promoting RNAP II termination. Mol. Microbiol. *101*, 559–574. https://doi.org/10.1111/mmi.13408.

Richard, G.F., Kerrest, A., and Dujon, B. (2008). Comparative genomics and molecular dynamics of DNA repeats in eukaryotes. Microbiol. Mol. Biol. Rev. *72*, 686–727. https://doi.org/10.1128/MMBR.00011-08.

Richter, K., Haslbeck, M., and Buchner, J. (2010). The heat shock response: life on the verge of death. Mol. Cell *40*, 253–266. https://doi.org/10.1016/j.molcel.2010.10.006.

Rochette, A., Raymond, F., Corbeil, J., Ouellette, M., and Papadopoulou, B. (2009). Whole-genome comparative RNA expression profiling of axenic and intracellular amastigote forms of *Leishmania infantum*. Mol. Biochem. Parasitol. *165*, 32–47. https://doi.org/10.1016/j.molbiopara.2008.12.012.

Rochette, A., Raymond, F., Ubeda, J.M., Smith, M., Messier, N., Boisvert, S., Rigault, P., Corbeil, J., Ouellette, M., and Papadopoulou, B. (2008). Genome-wide gene expression profiling analysis of *Leishmania major* and *Leishmania infantum* developmental stages reveals substantial differences between the two species. BMC Genomics *9*, 255. https://doi.org/10.1186/1471-2164-9-255.

Rogers, M.B., Hilley, J.D., Dickens, N.J., Wilkes, J., Bates, P.A., Depledge, D.P., Harris, D., Her, Y., Herzyk, P., Imamura, H., *et al.* (2011). Chromosome and gene copy number variation allow *major* structural change between species and strains of *Leishmania*. Genome Res. *21*, 2129–2142. https://doi.org/10.1101/gr.122945.111.

Rojo, D., Canuto, G.A., Castilho-Martins, E.A., Tavares, M.F., Barbas, C., López-Gonzálvez, Á., and Rivas, L. (2015). A Multiplatform Metabolomic Approach to the Basis of Antimonial Action and Resistance in *Leishmania infantum*. PLOS ONE *10*, e0130675. https://doi.org/10.1371/journal.pone.0130675.

Rosenzweig, D., Smith, D., Myler, P.J., Olafson, R.W., and Zilberstein, D. (2008a). Post-translational modification of cellular proteins during *Leishmania donovani* differentiation. Proteomics *8*, 1843–1850. https://doi.org/10.1002/pmic.200701043.

Rosenzweig, D., Smith, D., Opperdoes, F., Stern, S., Olafson, R.W., and Zilberstein, D. (2008b). Retooling *Leishmania* metabolism: from sand fly gut to human macrophage. FASEB J. *22*, 590–602.

Sacks, D.L., Hieny, S., and Sher, A. (1985). Identification of cell surface carbohydrate and antigenic changes between noninfective and infective developmental stages of *Leishmania major* promastigotes. J. Immunol. *135*, 564–569.

Sacks, D.L., and Perkins, P.V. (1984). Identification of an infective stage of *Leishmania* promastigotes. Science *223*, 1417–1419.

Sánchez-Luque, F.J., López, M.C., Carreira, P.E., Alonso, C., and Thomas, M.C. (2014). The wide expansion of hepatitis delta virus-like ribozymes throughout trypanosomatid genomes is linked to the spreading of L1Tc/ingi clade mobile elements. BMC Genomics *15*, 340. https://doi.org/10.1186/1471-2164-15-340.

Sánchez-Luque, F.J., López, M.C., Macias, F., Alonso, C., and Thomas, M.C. (2011). Identification of an hepatitis delta virus-like ribozyme at the mRNA 5'-end of the L1Tc retrotransposon from *Trypanosoma cruzi*. Nucleic Acids Res. *39*, 8065–8077. https://doi.org/10.1093/nar/gkr478.

Santoro, M.G. (2000). Heat shock factors and the control of the stress response. Biochem. Pharmacol. *59*, 55–63.

Sardar, A.H., Kumar, S., Kumar, A., Purkait, B., Das, S., Sen, A., Kumar, M., Sinha, K.K., Singh, D., Equbal, A., *et al.* (2013). Proteome changes associated with *Leishmania donovani* promastigote adaptation to oxidative and nitrosative stresses. J Proteomics *81*, 185–199. https://doi.org/10.1016/j.jprot.2013.01.011.

Saunders, E.C., Ng, W.W., Kloehn, J., Chambers, J.M., Ng, M., and McConville, M.J. (2014). Induction of a stringent metabolic response in intracellular stages of *Leishmania mexicana* leads to increased dependence on mitochondrial metabolism. PLOS Pathog. *10*, e1003888. https://doi.org/10.1371/journal.ppat.1003888.

Saxena, A., Worthey, E.A., Yan, S., Leland, A., Stuart, K.D., and Myler, P.J. (2003). Evaluation of differential gene expression in *Leishmania major* Friedlin procyclics and metacyclics using DNA microarray analysis. Mol. Biochem. Parasitol. *129*, 103–114.

Scheltema, R.A., Decuypere, S., T'kindt, R., Dujardin, J.C., Coombs, G.H., and Breitling, R. (2010). The potential of metabolomics for *Leishmania* research in the post-genomics era. Parasitology *137*, 1291–1302. https://doi.org/10.1017/S0031182009992022.

Schwede, A., Kramer, S., and Carrington, M. (2012). How do trypanosomes change gene expression in response to the environment? Protoplasma *249*, 223–238. https://doi.org/10.1007/s00709-011-0282-5.

Seco-Hidalgo, V., Osuna, A., and de Pablos, L.M. (2015). To bet or not to bet: deciphering cell to cell variation in protozoan infections. Trends Parasitol. *31*, 350–356. https://doi.org/10.1016/j.pt.2015.05.004.

Shaw, C.D., Lonchamp, J., Downing, T., Imamura, H., Freeman, T.M., Cotton, J.A., Sanders, M., Blackburn, G., Dujardin, J.C., Rijal, S., *et al.* (2016). In vitro selection of miltefosine resistance in promastigotes of *Leishmania donovani* from Nepal: genomic and metabolomic characterization. Mol. Microbiol. *99*, 1134–1148. https://doi.org/10.1111/mmi.13291.

Siegel, T.N., Hekstra, D.R., Kemp, L.E., Figueiredo, L.M., Lowell, J.E., Fenyo, D., Wang, X., Dewell, S., and Cross, G.A. (2009). Four histone variants mark the boundaries of polycistronic transcription units in *Trypanosoma brucei*. Genes Dev. *23*, 1063–1076. https://doi.org/10.1101/gad.1790409.

Silva, A.M., Cordeiro-da-Silva, A., and Coombs, G.H. (2011). Metabolic variation during development in culture of *Leishmania donovani* promastigotes. PLOS Negl. Trop. Dis. *5*, e1451. https://doi.org/10.1371/journal.pntd.0001451.

Singh, R., Kumar, D., Duncan, R.C., Nakhasi, H.L., and Salotra, P. (2010). Overexpression of histone H2A modulates drug susceptibility in *Leishmania* parasites. Int. J. Antimicrob. Agents *36*, 50–57. https://doi.org/10.1016/j.ijantimicag.2010.03.012.

Singh, S., and Dubey, V.K. (2016). Quantitative Proteome Analysis of *Leishmania donovani* under Spermidine Starvation. PLOS ONE *11*, e0154262. https://doi.org/10.1371/journal.pone.0154262.

Smith, D.F., Peacock, C.S., and Cruz, A.K. (2007). Comparative genomics: from genotype to disease phenotype in the leishmaniases. Int. J. Parasitol. *37*, 1173–1186.

Smolka, M.B., Zhou, H., Purkayastha, S., and Aebersold, R. (2001). Optimization of the isotope-coded affinity tag-labeling procedure for quantitative proteome analysis. Anal. Biochem. *297*, 25–31. https://doi.org/10.1006/abio.2001.5318.

Sogin, M.L., Elwood, H.J., and Gunderson, J.H. (1986). Evolutionary diversity of eukaryotic small-subunit rRNA genes. Proc. Natl. Acad. Sci. U.S.A. *83*, 1383–1387.

Solari, A.J. (1995). Mitosis and genome partition in trypanosomes. Biocell *19*, 65–84.

Sollelis, L., Ghorbal, M., MacPherson, C.R., Martins, R.M., Kuk, N., Crobu, L., Bastien, P., Scherf, A., Lopez-Rubio, J.J., and Sterkers, Y. (2015). First efficient CRISPR-Cas9-mediated genome editing in *Leishmania* parasites. Cell. Microbiol. *17*, 1405–1412. https://doi.org/10.1111/cmi.12456.

Soto, M., Iborra, S., Quijada, L., Folgueira, C., Alonso, C., and Requena, J.M. (2004). Cell-cycle-dependent translation of histone mRNAs is the key control point for regulation of histone biosynthesis in *Leishmania infantum*. Biochem. J. *379*, 617–625. https://doi.org/10.1042/BJ20031522.

Späth, G.F., and Clos, J. (2016). Joining forces: first application of a rapamycin-induced dimerizable Cre system for conditional null mutant analysis in *Leishmania*. Mol. Microbiol. *100*, 923–927. https://doi.org/10.1111/mmi.13374.

Späth, G.F., Drini, S., and Rachidi, N. (2015). A touch of Zen: post-translational regulation of the *Leishmania* stress response. Cell. Microbiol. *17*, 632–638. https://doi.org/10.1111/cmi.12440.

Srividya, G., Duncan, R., Sharma, P., Raju, B.V.S., Nakhasi, H.L., and Salotra, P. (2007). Transcriptome analysis during the process of in vitro differentiation of *Leishmania donovani* using genomic microarrays. Parasitology *134*, 1527–1539. https://doi.org/10.1017/S003118200700296X.

Sterkers, Y., Lachaud, L., Crobu, L., Bastien, P., and Pagès, M. (2011). FISH analysis reveals aneuploidy and continual generation of chromosomal mosaicism in *Leishmania major*. Cell. Microbiol. *13*, 274–283. https://doi.org/10.1111/j.1462-5822.2010.01534.x.

Stevens, J.R., Noyes, H.A., Schofield, C.J., and Gibson, W. (2001). The molecular evolution of Trypanosomatidae. Adv. Parasitol. *48*, 1–56.

Sunkin, S.M., Kiser, P., Myler, P.J., and Stuart, K. (2000). The size difference between leishmania *major* friedlin chromosome one homologues is localized to sub-telomeric repeats at one chromosomal end. Mol. Biochem. Parasitol. *109*, 1–15.

t'Kindt, R., Jankevics, A., Scheltema, R.A., Zheng, L., Watson, D.G., Dujardin, J.C., Breitling, R., Coombs, G.H., and Decuypere, S. (2010a). Towards an unbiased metabolic profiling of protozoan parasites: optimisation of a *Leishmania* sampling protocol for HILIC-orbitrap analysis. Anal. Bioanal. Chem. *398*, 2059–2069. https://doi.org/10.1007/s00216-010-4139-0.

t'Kindt, R., Scheltema, R.A., Jankevics, A., Brunker, K., Rijal, S., Dujardin, J.C., Breitling, R., Watson, D.G., Coombs, G.H., and Decuypere, S. (2010b). Metabolomics to unveil and understand phenotypic diversity between pathogen populations. PLOS Negl. Trop. Dis. *4*, e904. https://doi.org/10.1371/journal.pntd.0000904.

Tannu, N.S., and Hemby, S.E. (2006). Two-dimensional fluorescence difference gel electrophoresis for comparative proteomics profiling. Nat. Protoc. *1*, 1732–1742.

Thomas, M.C., Macias, F., Alonso, C., and López, M.C. (2010). The biology and evolution of transposable elements in parasites. Trends Parasitol. *26*, 350–362. https://doi.org/10.1016/j.pt.2010.04.001.

Thomas, S., Green, A., Sturm, N.R., Campbell, D.A., and Myler, P.J. (2009). Histone acetylations mark origins of polycistronic transcription in *Leishmania major*. BMC Genomics *10*, 152. https://doi.org/10.1186/1471-2164-10-152.

Tsigankov, P., Gherardini, P.F., Helmer-Citterich, M., Späth, G.F., Myler, P.J., and Zilberstein, D. (2014). Regulation dynamics of *Leishmania* differentiation: deconvoluting signals and identifying phosphorylation trends. Mol. Cell Proteomics *13*, 1787–1799. https://doi.org/10.1074/mcp.M114.037705.

Tsigankov, P., Gherardini, P.F., Helmer-Citterich, M., and Zilberstein, D. (2012). What has proteomics taught us about *Leishmania* development? Parasitology *139*, 1146–1157. https://doi.org/10.1017/S0031182012000157.

Ubeda, J.M., Raymond, F., Mukherjee, A., Plourde, M., Gingras, H., Roy, G., Lapointe, A., Leprohon, P., Papadopoulou, B., Corbeil, J., *et al.* (2014). Genome-wide stochastic adaptive DNA amplification at direct and inverted DNA repeats in the parasite *Leishmania*. PLOS Biol. *12*, e1001868. https://doi.org/10.1371/journal.pbio.1001868.

Vacchina, P., Norris-Mullins, B., Abengózar, M.A., Viamontes, C.G., Sarro, J., Stephens, M.T., Pfrender, M.E., Rivas, L., and Morales, M.A. (2016). Genomic Appraisal of the Multifactorial Basis for In Vitro Acquisition of Miltefosine Resistance in *Leishmania donovani*. Antimicrob. Agents Chemother. *60*, 4089–4100. https://doi.org/10.1128/AAC.00478-16.

van Leeuwen, F., Taylor, M.C., Mondragon, A., Moreau, H., Gibson, W., Kieft, R., and Borst, P. (1998). beta-D-glucosyl-hydroxymethyluracil is a conserved DNA modification in kinetoplastid protozoans and is abundant in their telomeres. Proc. Natl. Acad. Sci. U.S.A. *95*, 2366–2371.

van Luenen, H.G., Farris, C., Jan, S., Genest, P.A., Tripathi, P., Velds, A., Kerkhoven, R.M., Nieuwland, M., Haydock, A., Ramasamy, G., *et al.* (2012). Glucosylated hydroxymethyluracil, DNA base J, prevents transcriptional readthrough in *Leishmania*. Cell *150*, 909–921. https://doi.org/10.1016/j.cell.2012.07.030.

Vazquez, M., Ben-Dov, C., Lorenzi, H., Moore, T., Schijman, A., and Levin, M.J. (2000). The short interspersed repetitive element of *Trypanosoma cruzi*, SIRE, is part of VIPER, an unusual retroelement related to long terminal repeat retrotransposons. Proc. Natl. Acad. Sci. U.S.A. *97*, 2128–2133. https://doi.org/10.1073/pnas.050578397.

Velculescu, V.E., Zhang, L., Vogelstein, B., and Kinzler, K.W. (1995). Serial analysis of gene expression. Science *270*, 484–487. https://doi.org/10.1126/science.270.5235.484.

Veras, P.S., and Bezerra de Menezes, J.P. (2016). Using Proteomics to Understand How *Leishmania* Parasites Survive inside the Host and Establish Infection. Int J Mol Sci *17*, E1270. https://doi.org/10.3390/ijms17081270.

Vergnes, B., Gourbal, B., Girard, I., Sundar, S., Drummelsmith, J., and Ouellette, M. (2007). A proteomics screen implicates HSP83 and a small kinetoplastid calpain-related protein in drug resistance in *Leishmania donovani* clinical field isolates by modulating drug-induced programmed cell death. Mol. Cell Proteomics *6*, 88–101.

Vergnes, B., Sereno, D., Tavares, J., Cordeiro-da-Silva, A., Vanhille, L., Madjidian-Sereno, N., Depoix, D., Monte-Alegre, A., and Ouaissi, A. (2005). Targeted disruption of cytosolic SIR2 deacetylase discloses its essential role in *Leishmania* survival and proliferation. Gene *363*, 85–96.

Vincent, I.M., Weidt, S., Rivas, L., Burgess, K., Smith, T.K., and Ouellette, M. (2014). Untargeted metabolomic analysis of miltefosine action in *Leishmania infantum* reveals changes to the internal lipid metabolism. Int J Parasitol Drugs Drug Resist *4*, 20–27. https://doi.org/10.1016/j.ijpddr.2013.11.002.

Westrop, G.D., Williams, R.A., Wang, L., Zhang, T., Watson, D.G., Silva, A.M., and Coombs, G.H. (2015). Metabolomic Analyses of *Leishmania* Reveal Multiple Species Differences and Large Differences in Amino Acid Metabolism. PLOS ONE *10*, e0136891. https://doi.org/10.1371/journal.pone.0136891.

Wincker, P., Ravel, C., Blaineau, C., Pages, M., Jauffret, Y., Dedet, J.P., and Bastien, P. (1996). The *Leishmania* genome comprises 36 chromosomes conserved across widely divergent human pathogenic species. Nucleic Acids Res. *24*, 1688–1694.

Wong, J.W., and Cagney, G. (2010). An overview of label-free quantitation methods in proteomics by mass spectrometry. Methods Mol. Biol. *604*, 273–283. https://doi.org/10.1007/978-1-60761-444-9_18.

Wright, J.R., Siegel, T.N., and Cross, G.A. (2010). Histone H3 trimethylated at lysine 4 is enriched at probable transcription start sites in *Trypanosoma brucei*. Mol. Biochem. Parasitol. *172*, 141–144. https://doi.org/10.1016/j.molbiopara.2010.03.013.

Xia, J., and Wishart, D.S. (2016). Using MetaboAnalyst 3.0 for comprehensive metabolomics data analysis. Curr. Protoc. Bioinformatics *55*, 14.10.11–14.10.91.

Yao, C., Li, Y., Donelson, J.E., and Wilson, M.E. (2010). Proteomic examination of *Leishmania chagasi* plasma membrane proteins: Contrast between avirulent and virulent (metacyclic) parasite forms. Proteomics Clin Appl *4*, 4–16. https://doi.org/10.1002/prca.200900050.

Zhang, W.W., and Matlashewski, G. (2015). CRISPR-Cas9-Mediated Genome Editing in *Leishmania donovani*. MBio *6*, e00861. https://doi.org/10.1128/mBio.00861-15.

Zilka, A., Garlapati, S., Dahan, E., Yaolsky, V., and Shapira, M. (2001). Developmental regulation of heat shock protein 83 in *Leishmania*. 3' processing and mRNA stability control transcript abundance, and translation id directed by a determinant in the 3'-untranslated region. J. Biol. Chem. *276*, 47922–47929. https://doi.org/10.1074/jbc.M108271200