

Research Article

Application of Serial Analysis of Gene Expression to the Study of the Gene Expression Profile of *Leishmania infantum chagasi* Promastigote

Adelino Soares Lima Neto,¹ Osvaldo Pompílio de Melo Neto,²
and Carlos Henrique Nery Costa¹

¹Laboratório de Pesquisas em Leishmanioses, Instituto de Doenças Tropicais Natan Portella (IDTNP), Rua Gov. Artur de Vasconcelos 151, Centro/Sul, 64.001-450 Teresina, PI, Brazil

²Departamento de Microbiologia, Centro de Pesquisas Aggeu Magalhães (CPqAM), Fiocruz, Avenida Moraes Rego s/n. Campus da UFPE, Cidade Universitária, 50.670-420 Recife, PE, Brazil

Correspondence should be addressed to Adelino Soares Lima Neto, adelinolima@yahoo.com.br

Received 18 May 2011; Accepted 30 November 2011

Academic Editor: Zoran Obradovic

Copyright © 2012 Adelino Soares Lima Neto et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This study describes the application of the LongSAGE methodology to study the gene expression profile in promastigotes of *Leishmania infantum chagasi*. A tag library was created using the LongSAGE method and consisted of 14,208 tags of 17 bases. Of these, 8,427 (59.3%) were distinct. BLAST research of the 1,645 most abundant tags showed that 12.8% of them identified the coding sequences of genes, while 82% (1,349/1,645) identified one or more genomic sequences that did not correspond with open reading frames. Only 5.2% (84/1,645) of the tags were not aligned to any position in the *L. infantum* genome. The UTR size of *Leishmania* and the lack of CATG sites in some transcripts were decisive for the generation of tags in these regions. Additional analysis will allow a better understanding of the expression profile and discovering the key genes in this life cycle.

1. Introduction

Visceral leishmaniasis (VL) is a disease caused by the protozoan *Leishmania chagasi* (in the New World) and *L. infantum* or *L. donovani* (in the Old World) transmitted between humans and other mammals through the bite of sand flies of the genera *Lutzomyia* and *Phlebotomus*. It is commonly accepted that *L. infantum* is genetically identical to *L. chagasi* [1], which is currently named *L. infantum chagasi* [2]. VL occurs in 65 countries, with an estimated annual incidence of 500,000 cases. The majority (90%) of these occur in rural areas and suburbs of large urban centers in five countries: Bangladesh, India, Nepal, Sudan, and Brazil [3].

Leishmania parasites have a life cycle characterized by the presence of a promastigote stage, flagellated, inside the gut of the insect vector and an amastigote stage, nonflagellated, and finally within macrophages of the host [4]. Each of these life cycle stages presents with a unique set of biochemical,

genetic, and morphological traits that are either unique to trypanosomatids, such as mitochondrial minicircles, glycosome, and RNA editing, or used to a greater extent than in other organisms, such as membrane proteins anchored by glycosyl phosphatidylinositol (GPI), polycistronic transcription, and transsplicing, among others [5].

The publication of the genome sequence of *L. major* [6] and *L. infantum* [7] has enabled new approaches of profile studies of comparative gene expression of different stages of the *Leishmania* life cycle [8–10] and during differentiation [11] and has fostered a better understanding of various aspects of its biology and pathogenesis.

The serial analysis of gene expression (SAGE) is a powerful methodology that can be used to obtain the complete gene expression profile of a cell or tissue. It is based on the following principles: first, a short nucleotide sequence of 9 or 10 base pairs (or tag) contains sufficient information to identify an mRNA transcript as unique. Second, the

concatenation of dozens of tags enables serial analysis of multiple transcripts, by determining the sequence of multiple tags within a single clone [12].

A major advantage of SAGE is that it can be applied to study the profile of expressed genes in organisms whose genome is not yet sequenced or publicly available. Moreover, in the years since the original description of SAGE, changes have been reported in various stages of the protocol, some of which have enabled a reduction of the quantity of initial RNA [13] or an increase in the size of the tag generated [14–16]; these changes have made the method even more specific.

Recently, tag libraries that use SAGE (ShortSAGE) were constructed with the aim of identifying differentially expressed genes between the promastigote and axenic amastigote of *L. donovani* [9]. The comparison of these libraries showed that about 90% of genes were expressed in both stages. A total of 968 genes revealed statistically different transcript levels between promastigotes and axenic amastigotes. Of these, most (642) were derived from amastigotes [9]. Another application of SAGE in *Leishmania* was described by Guerfali et al. [10] to investigate changes in the transcriptome of *L. major* and human macrophages infected with *L. major*. The comparison between libraries of promastigotes and cultured macrophages that were infected or uninfected by the parasite led to the identification of human genes whose expression profile may be relevant to the pathogenesis of the disease. Regarding the parasite, genes showing differential expression in the intracellular stage were also identified, including amastins, oxidative stress proteins, and several ribosomal proteins [10].

One of the main problems regarding the application of SAGE in *Leishmania* sp. is the lack of a database that maps tags to a gene. The absence of such data could be a result of difficulties in the mapping process. It is believed that characteristics inherent to the *Leishmania* sp. genome (such as a percentage of GC bases, size of the coding regions (*open reading frame*—ORFs) and 5' and 3' untranslated regions—UTRs) can negatively influence the process of mapping the tag to a gene. In addition, one characteristic of the SAGE methodology is that tags can be generated from any transcript region where there is an enzyme-anchoring site *Nla*III [12]. Thus, depending on the size, it is possible that many tags are generated from 3'UTR, without corresponding ORFs. Additionally, depending on the size and composition of the bases, it is possible that ORFs do not present the site for the anchoring enzyme, generating tags that are exclusive to UTRs.

This study described the construction of a tag library of *L. i. chagasi* promastigotes using LongSAGE. It analyzed the role of genomic and methodological characteristics that may influence the process of generation and tag to gene mapping and evaluated the applicability of the methodology for studying gene expression in *Leishmania*.

2. Materials and Methods

2.1. *Leishmania* Cultivation. Strain 2230 promastigotes of *L. i. chagasi* were used. The sample was isolated from a patient

diagnosed with VL at the Natan Portella Institute for Tropical Diseases (IDTNP), Teresina, Piauí, Brazil. The promastigotes were grown in Schneider medium (Sigma) supplemented with 10% (v/v) fetal bovine serum (FBS), 2% human urine (v/v) at pH 7.2 and maintained at 25°C.

2.2. RNA Extraction and Assessment. Total RNA was extracted from approximately 8×10^7 promastigotes using TRIzol LS Reagent (Invitrogen, Frederick MD, USA) according to the manufacturer's protocol. The integrity of RNA was assessed on 1% agarose gel in a denaturing condition (MOPS/formamide) and quantification was carried out using a spectrophotometer at 260 nm. About 30 µg of total RNA was used in the procedure.

2.3. SAGE Procedure *L. i. chagasi* Promastigotes. The I-SAGE Long kit (Invitrogen) was used to construct the *L. i. chagasi* library according to the manufacturer's recommendations. The following were included among the main steps: polyadenylated mRNAs were captured by oligo (dT) linked to magnetic beads and used for cDNA synthesis. The cDNA was digested with 60 U of *Nla*III (anchoring enzyme), and the 3' ends of the cDNAs were isolated using the beads. The resulting cDNA 3' was divided into two equal portions and connected to two LongSAGE adapters, A and B. The long tags were released by the enzyme *Mme*I and linked to form ~130 bp ditags. Dilutions of 1:40 of the product were amplified in 27 PCR cycles (300 reactions in total). The precipitated PCR products were separated on 12% polyacrylamide gel, the 130 bp bands were cut out, and the precipitated DNA was digested with *Nla*III. The digestion products were separated on 12% polyacrylamide gel, and ~34 bp ditags were cut out and linked to form concatemers. The concatemers were separated on 6% polyacrylamide gel, and the fractions of 250–500 and 500–800 pb were isolated and cloned into pZER0-1 vector digested with *Sph*I. Vectors with concatemers were cloned in *E. coli* DH10b by electroporation, and the isolated recombinant vectors were used as template for sequencing in the ABI Prism 3100 (Applied Biosystems).

2.4. Sequence Analysis and Database. Crude sequences were first analyzed with the EditSeq tool (Lasergene, v4.01, DNASTAR Inc., Madison, WI), to confirm the presence of tags, identifiable due to the occurrence of the CATG motif (*Nla* III site) at 34 bp intervals. SAGE tags were then directly extracted from the sequencing files using the program SAGE2000 (Version 4.5), which also looks for duplicates and quantifies the individual tags.

To identify the corresponding gene to *Leishmania* tag, conventional BLAST (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>) was carried out with the 1,645 tags that had two or more copies in the library against the *L. infantum* nucleotides collection. When the tag presented 100% sequence identity with a genomic sequence, the distance in bp was noted, as well as the orientation and the closest ORF (gene ID and name).

TABLE 1: Summary of the serial analysis of gene expression in *L. i. chagasi* promastigotes.

Copy number	Number of sequenced tags (%)	Number of unique identified tags (%)
1	6,782 (47.7)	6,782 (80.5)
2~4	3,242 (22.8)	1,331 (15.8)
5~19	2,288 (16.1)	272 (3.2)
20~99	1,331 (9.4)	39 (0.46)
≥100	565 (4.0)	3 (0.04)
Total	14,208 (100.0)	8,427 (100.0)

TABLE 2: Summary of BLAST search of 50 most expressed tags of *L. i. chagasi* as Long or Short tag.

	Number of distinct genes matching to tag				
	1	2	3	4	6
Long tag (21 bp)	49	1	0	0	0
Short tag (14 bp)	30	8	4	6	2

3. Results

3.1. Global Data. To analyze the gene expression profile in *L. i. chagasi* promastigotes a tag library was constructed using the LongSAGE method. The library generated was composed of 14,208 tags of 17 bases, here called long tags (LTs). Of the total number of LTs, 8,427 (59.3%) were distinct. Of these 8,427 distinct tags, 1,645 (19.5%) had two or more copies, while the vast majority, 6,782 (80.5%), had only one copy (Table 1). The data discussed in this paper have been deposited in NCBI's Gene Expression Omnibus and are accessible through GEO Series accession number GSE29369 (<http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE29369>).

3.2. Long Tag Specificity. To demonstrate the increased specificity of long tag in relation to short tag, we did a BLAST search with the 50 tags of the most expressed genes of *L. i. chagasi* against the *L. infantum* genome database. For each 21 bp tag, its size was set to 14 bp (short tag) and used as query for BLAST search. The number of distinct *L. infantum* genes showing perfect matching to the given tag was counted for each tag. The 21 bp tag has matching to only one gene in 98% of cases, compared with just 60% of 14 bp tags (Table 2).

3.3. Gene Identification by BLAST. The BLAST research showed that 82% (1,349/1,645) of the tags align to one or more positions of the *L. infantum* genome, which do not correspond to ORFs. Of these, 42% (690/1,349) were aligned more closely to 5' end of a gene, while the remainder (659) were more closely to 3' end. The mean distance between the tag and the closest gene was 718 bp for both regions. About 80% of tags were less than 1 Kb away from the closest gene (Table 3).

About seventy-five percent of these tags aligned up to just one region in the genome, whereas 8.5% (141/1,645) of them identified more than one region that, in most cases, flanked gene copies or genes belonging to the same family. Only

TABLE 3: Results of conventional BLAST alignments of 1,645 tags with two or more copies in the *L. i. chagasi* library per alignment region and distance from the nearest ORF.

Tags/genes	5' UTR (%)	3' UTR (%)	Total (%)
Different tags aligned at	690 (42.0)	659 (40.0)	1,349 (82.0)
Genes identified at			
Up to 1Kb	624 (38.0)	632 (38.4)	1,256 (76.4)
Between 1 to 2 Kb	120 (7.2)	116 (7.0)	236 (14.3)
More than 2 Kb	41 (2.5)	44 (2.7)	85 (5.2)
Total	785 (47.8)	792 (48.1)	1,577 (95.9)
Mean distance from ORF	718	718	—
Reach (bp)	5 to 14,647	1 to 8,161	—

5.2% (84/1,645) of the tags did not have 100% alignment to any position in the *L. infantum* genome. These may have originated from *L. i. chagasi* genome regions presenting polymorphisms in relation to *L. infantum*, from *L. i. chagasi* exclusive transcripts or simply due to sequencing errors.

3.4. Most Expressed Genes in *L. i. chagasi*. The BLAST search showed that distinct tags can identify the same transcript as well as a particular tag can identify distinct genes. At least 294 genes were identified by 823 distinct tags. Each tag identified 2.8 genes average. This may be the result that tags can be generated from any transcript region where there is an anchoring-enzyme site or different 5' or 3' sequence generated by alternative splicing or cleavage. Thus, the most expressed genes correspond to the sum of the number of different tags that identifies it. This concept was used to identify the genes most expressed by *L. i. chagasi*. This shows that the number of times that a tag appears in the library is not the most important factor, but rather the total number of times that the gene that it identifies appears in the library, which justifies the need for mapping all the tags in the library. With this strategy, the most expressed genes for *L. i. chagasi* were those listed in Table 4, which includes translation elongation factors, several members of the histone family (H1, H3, H4), several members of the ribosomal proteins family (S and L), tubulins, heat shock proteins (hsp 70 and hsp 83), and hypothetical proteins.

3.5. Comparison between Libraries. Compared to *L. donovani* and *L. major* promastigotes SAGE libraries [9, 10], the results

TABLE 4: Most abundant and annotated transcript and corresponding tags identified in the library of *L. i. chagasi* promastigotes.

Ord.	Long tag	Frequ.	Gene/protein	Gene ID
1	CATGCGCGGATGGTGCCCC	125	elongation factor 1 alpha	LinJ.17.0100
2	CATGCGACTTAGACCGTGAGG	91	histone H1	LinJ.27.1070
3	CATGGGCGCACGGCGCGCCG	68	heat shock protein 83-1	LinJ.33.0350
4	CATGGGCTTCCTCGAGTCCGC	60	histone H3	LinJ.10.1050
5	CATGGAGGAGGGCGAGTTCTC	57	Leishmania infantum JPCM5 alpha tubulin	LinJ.13.0330
6	CATGCTCCATAAAGGAGAGAC	53	heat-shock protein hsp70, putative	LinJ.28.3000
7	CATGCGTACTGCGTTCCAGA	47	ribosomal protein L23, putative	LinJ.35.3840
8	CATGGCGGCCGAAAAGCAGG	43	40S ribosomal protein S17, putative	LinJ.28.2750
9	CATGACTGGACATTCAAAGA	41	histone H1, putative	LinJ.27.1070
10	CATGCGAAAATGTGGTTTCGC	39	60S ribosomal protein L7a, putative	LinJ.07.0550
11	CATGACGCGCCACCCGTTGT	38	ubiquitin-fusion protein	LinJ.31.1930
12	CATGCGTGGACGATTTCAAAG	36	ribosomal protein S29, putative	LinJ.28.2360
13	CATGCCGGCAGCACACCAACA	35	histone h4	LinJ.06.0010
14	CATGCCCTTCATTTCTCCC	32	RNA binding protein, putative	LinJ.32.0790
15	CATGCGCCCTTTCATTTAAC	29	hypothetical protein, conserved	LinJ.33.0970
16	CATGACGCTCTTGCCAACCTG	28	ribosomal protein s26, putative	LinJ.30.3240
17	CATGCGAAAAACAGCACAGCA	27	60S ribosomal protein L6, putative	LinJ.15.1060
18	CATGGCTGCCGCTACCGCAGC	26	ribosomal protein s11 homolog	LinJ.20.1620
19	CATGATCAGATTTGTTTTGT	26	ribosomal protein L3, putative	LinJ.34.2730
20	CATGTGCGTGCCGGTGCCGGT	25	hypothetical protein, conserved	LinJ.32.2840
21	CATGCGGCTTGTTTATCCTTT	25	60S ribosomal protein L12, putative	LinJ.35.2230
22	CATGCGGGCGCGACTTTGCG	24	ribosomal protein L24, putative	LinJ.36.1130
23	CATGCGGAGGGCTGTGAAGC	24	hypothetical protein, conserved	LinJ.27.0130
24	CATGTGCAAGACTCACGTCCA	23	ribosomal protein S25	LinJ.34.0460
25	CATGGTCGTTTTGCGGGCAGC	22	hypothetical protein, conserved	LinJ.13.0270
26	CATGAAATGAAAAGGAAAGGC	22	ribosomal protein L15, putative	LinJ.30.3710
27	CATGTGGATCGCGGGTGCCA	22	60S ribosomal protein L2, putative	LinJ.32.4050
28	CATGGTGGACAGTGGCGAGCG	22	60S acidic ribosomal subunit protein,	LinJ.27.1300
29	CATGGCAGCTGCCGCTGCGTC	21	60S ribosomal protein L34, putative	LinJ.36.3930
30	CATGGGATGGAAGGCATCGCA	20	ATP-dependent zinc metalloproteinase,	LinJ.18.0620
31	CATGCAGAGAGAAGCAGTGCA	20	heat shock protein 83-1	LinJ.33.0350
32	CATGTGCTGTGATCCGTGTGT	20	hypothetical protein, conserved	LinJ.24.2290
33	CATGTATGCCTAAAGTACCAC	20	hypothetical protein, conserved	LinJ.31.0920
34	CATGGGAAACACTCTGCGCCC	19	60S ribosomal protein L19, putative	LinJ.06.0430
35	CATGACGGTGGGTGCGCTAAT	18	hypothetical protein, conserved	LinJ.30.3590

This table shows the 35 most abundant and annotated tags aligning to only one gene, their occurrences in the *L. i. chagasi* library, their corresponding name and gene ID in GeneDB.

were highly consistent. In all libraries the most expressed genes were those related to cellular metabolism as histone, tubulin, ribosomal protein, and elongation factors. However, in the *L. i. chagasi* library we found tags matching to genes encoding proteins not found in the other libraries, such as heat shock proteins, an ATP-dependent zinc metalloproteinase, and more genes encoding hypothetical proteins. These proteins may play an important and possibly not yet identified role for the parasite. It is important to verify the sequence, expression, and cellular location to access its cell function.

4. Discussion

SAGE is a technique for studying global gene expression based on a sequence that enables researchers to obtain the complete gene expression profile of a cell or tissue, even if their genes are not previously known [12]. At least two variants of traditional SAGE (14 bp tag) have been developed to improve the specificity of the technique with respect to the mapping tags to genes: LongSAGE, which generates tags of 21 bp (with CATG) [14], and SuperSAGE, which generates tags of 26 bp (with CATG) [16].

In this study, a tag library of *L. i. chagasi* promastigote constructed with LongSAGE was presented. A total of 14,208 long tags were generated, of which 8,427 are distinct. This is the first description of SAGE tags library in *L. i. chagasi* and is the only one for LongSAGE in *Leishmania*. The advantage of LongSAGE over the conventional SAGE is that the information content of a 21 bp LongSAGE tag is appreciably higher than a 14 bp tag of conventional SAGE [12].

As shown, more than 80% of the tags originated from noncoding sequences. This characteristic seems to be more pronounced in *Leishmania* sp. than in other organisms. In a SuperSAGE library (tags of 26 bp) from human bone marrow, about 90% of the tags identified the corresponding gene by coding sequence. The high percentage of tags aligning to non-coding sequence in *Leishmania* raises an important question: why the vast majority of the tags align up perfectly to the genomic sequence and not to the coding sequence? An attempt was made to answer this question by analyzing the following factors: (1) it is possible that not all ORFs have the enzyme *NlaIII* (anchoring enzyme) site that is used to delimit the tags, while others may have more than one site; (2) characteristics such as base composition and relative size of the UTRs and ORF of each transcript can determine the tag position.

To investigate the impact of ORFs without *NlaIII* site on the results, 7,993 ORFs of *L. infantum* deposited in NCBI (www.ncbi.nlm.gov/mapview) were downloaded and analyzed for the presence (and quantity) or absence of the CATG site. The mapping of CATG sites was restricted to coding sequences and according to the results identified 251 ORFs (3.2% of coding sequences in the genome) without the enzyme site. This and the results obtained from other genomes (such as *L. major* and *Plasmodium falciparum*) showed that, as expected, the smaller the ORF and GC content, the lower the occurrence of the enzyme site. Certainly, the small percentage of ORFs with no CATG site does not justify the large number of tags (80%) generated from non-coding sequence, but identifies transcripts that can generate tags exclusively on their 3' or 5' UTR sequences.

It is known that UTRs play crucial roles in regulating posttranscriptional gene expression, including regulation of the transportation to outside the nucleus efficiency, subcellular localization, mRNA stability, and translation efficiency [17]. In higher eukaryotes, the mean size of 5' UTR varies from 100 to 200 nucleotides, while for 3' UTR the average size can be as high as 800 nucleotides [17, 18]. Compared with higher eukaryotes, the sizes of the 5' and 3' UTR of trypanosomatids are larger. The 5' UTR can vary from 100 to 300 bp, while the 3' UTR may have a size of 1.5 to 2 Kb [6]. The size differences between the 3' UTRs of humans and *Leishmania* sp. may partly explain the fact that over 80% of *L. i. chagasi* tags aligned with sequences other non-ORFs.

As shown, the observed mean distance between tags and their nearest ORF was the same (718 bp) for both 5' and 3' regions, indicating that (1) it is possible that many tags mapped closer to the 5' end of the genes are, in fact, located within the 3'UTRs of upstream ORFs or (2) the 5'UTR sequences of *Leishmania* sp. are greater than initially anticipated [6]. In *Trypanosoma brucei*, the average

5' UTR length (based in splice-acceptor site) was 184 bp and the average predominant 3' UTR length was 604 bp [19]. Another possibility is that these (especially the largest) can represent unpredicted ORFs.

Since most tags identified 5' or 3' UTR, it is needed to confirm the corresponding genes. The gene annotation or confirmation using the tags that did not match to *L. infantum* ORFs as initiator in the RAGE [20] or GLGI [21] and the building of a database of accurately annotated tags will be the focus of the next steps in this study.

There are many studies that have looked at transcriptomics in *Leishmania* using DNA microarray [22, 23]. However, unlike SAGE, DNA microarray is a closed platform. Recently, new software for analysis and sequence alignment have been described [24, 25], but none permit the comparison of SAGE libraries. In addition, the digital format of SAGE data enables the truncation of 17-base to 10-base tags. This makes it possible to compare libraries built at different times and places—a double benefit of LongSAGE regarding DNA microarrays.

As with the library of *L. donovani* or *L. major* [9, 10], the most expressed genes in the *L. i. chagasi* library were those with constitutive expression, such as those related to DNA and protein related to metabolism: histone, translation elongation factors, ribosomal proteins, ubiquitin, and tubulin. However, at least two genes were more expressed in *L. i. chagasi* when compared to the other libraries: genes encoding heat shock proteins (HSP-83 (LinJ.33.0350) and HSP-70 (LinJ.28.3000)) and an ATP-dependent zinc metalloproteinase (LinJ.18.0620). The significance of high expression of these genes is still not fully understood. It is known that HSP-70 and HSP-83 are constitutively expressed, but the accumulation of mRNA and the increased rate of translation are only found at high temperatures [26], such as the one that the parasite finds in the vertebrate host. Moreover, zinc metalloproteinases are numerous in *Leishmania* [6]. The main one, GP63, is more expressed in metacyclic promastigotes and plays an important role in the resistance to lysis mediated by the complement system in the early stages of infection [27]. It is possible that the metalloproteinase gene identified in the *L. i. chagasi* library plays a similar role to that of GP63. In addition, in the *L. i. chagasi* library we found tags matching to genes responsible for encoding hypothetical proteins with no ortholog found in the other libraries, except one (LinJ.13.0270), whose ortholog (LmjF.13.0450) was abundantly expressed in *L. major* promastigote. These proteins may play an important and possibly not yet identified role for these parasites. It is important to verify the sequence, expression, and cellular location to access its cell function.

A weakness of our study is that it presents results from a library of *L. i. chagasi* promastigotes. Although it does not reveal the gene expression profile of the form found in the vertebrate host (amastigotes) and thus makes little contribution to understanding the pathogenesis of VL, it is of great value for understanding the biology of *Leishmania* in the insect vector. From the methodological point of view, LongSAGE is a powerful tool for studying gene expression in *Leishmania*.

5. Conclusions

In conclusion, the description of a tag library using the LongSAGE method in promastigotes of the parasite *L. i. chagasi* was presented. It was shown that characteristics of the method and of the actual genome of *Leishmania* sp. (such as UTRs and ORF sizes) may have negative impacts on the process of mapping tag to gene.

LongSAGE revealed the gene expression profile in promastigotes of *L. i. chagasi* in culture. The most expressed genes were related to cellular metabolism, whereas preferential expression genes in the amastigote stage were also represented in the promastigote form. Additional analysis will allow a better understanding of the expression profile and discovering the key genes in this life cycle.

Acknowledgments

The authors would like to thank Dr. L. Figueirêdo for the critical revision of the paper. Dra C. Docena, V. Carvalho, and the PDTIS Program-FIOCRUZ are acknowledged for the technological development in tools for health, for access and help with automated sequencing. They would also like to thank the Aggeu Magalhães Research Center-FIOCRUZ/PE, the National Council for Scientific and Technological Development (CNPq), and the Research Support Foundation of the State of Piauí (FAPEPI) for financial support.

References

- [1] J. Lukeš, I. L. Mauricio, G. Schönian et al., "Evolutionary and geographical history of the *Leishmania donovani* complex with a revision of current taxonomy," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 104, no. 22, pp. 9375–9380, 2007.
- [2] J. J. Shaw, "Further thoughts on the use of the name *Leishmania (Leishmania) infantum chagasi* for the aetiological agent of American visceral leishmaniasis," *Memorias do Instituto Oswaldo Cruz*, vol. 101, no. 5, pp. 577–579, 2006.
- [3] P. Desjeux, "Leishmaniasis: current situation and new perspectives," *Comparative Immunology, Microbiology and Infectious Diseases*, vol. 27, no. 5, pp. 305–318, 2004.
- [4] G. Milon, "Perpetuation of leishmania: Some novel insight into elegant developmental programs," *Veterinary Research*, vol. 40, no. 2, article 38, 2009.
- [5] D. A. Campbell, S. Thomas, and N. R. Sturm, "Transcription in kinetoplastid protozoa: why be normal?" *Microbes and Infection*, vol. 5, no. 13, pp. 1231–1240, 2003.
- [6] A. C. Ivens, C. S. Peacock, E. A. Worthey et al., "The genome of the kinetoplastid parasite, *Leishmania major*," *Science*, vol. 309, no. 5733, pp. 436–442, 2005.
- [7] C. S. Peacock, K. Seeger, D. Harris et al., "Comparative genomic analysis of three *Leishmania* species that cause diverse human disease," *Nature Genetics*, vol. 39, no. 7, pp. 839–847, 2007.
- [8] K. Leifso, G. Cohen-Freue, N. Dogra, A. Murray, and W. R. McMaster, "Genomic and proteomic expression analysis of *Leishmania* promastigote and amastigote life stages: the *Leishmania* genome is constitutively expressed," *Molecular and Biochemical Parasitology*, vol. 152, no. 1, pp. 35–46, 2007.
- [9] Q. Li, Y. Zhao, B. Ni et al., "Comparison of the expression profiles of promastigotes and axenic amastigotes in *Leishmania donovani* using serial analysis of gene expression," *Parasitology Research*, vol. 103, no. 4, pp. 821–828, 2008.
- [10] F. Z. Guerfali, D. Laouini, L. Guizani-Tabbane et al., "Simultaneous gene expression profiling in human macrophages infected with *Leishmania major* parasites using SAGE," *BMC Genomics*, vol. 9, article 238, 2008.
- [11] A. Saxena, T. Lahav, N. Holland et al., "Analysis of the *Leishmania donovani* transcriptome reveals an ordered progression of transient and permanent changes in gene expression during differentiation," *Molecular and Biochemical Parasitology*, vol. 152, no. 1, pp. 53–65, 2007.
- [12] V. E. Velculescu, L. Zhang, B. Vogelstein, and K. W. Kinzler, "Serial analysis of gene expression," *Science*, vol. 270, no. 5235, pp. 484–487, 1995.
- [13] C. Vilain, F. Libert, D. Venet, S. Costagliola, and G. Vassart, "Small amplified RNA-SAGE: an alternative approach to study transcriptome from limiting amount of mRNA," *Nucleic Acids Research*, vol. 31, no. 6, p. e24, 2003.
- [14] S. Saha, A. B. Sparks, C. Rago et al., "Using the transcriptome to annotate the genome," *Nature Biotechnology*, vol. 20, no. 5, pp. 508–512, 2002.
- [15] H. Matsumura, S. Reich, A. Ito et al., "Gene expression analysis of plant host-pathogen interactions by SuperSAGE," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 100, no. 26, pp. 15718–15723, 2003.
- [16] H. Matsumura, A. Ito, H. Saitoh et al., "SuperSAGE," *Cellular Microbiology*, vol. 7, no. 1, pp. 11–18, 2005.
- [17] F. Mignone, C. Gissi, S. Liuni, and G. Pesole, "Untranslated regions of mRNAs," *Genome Biology*, vol. 3, no. 3, article 0004, 2002.
- [18] G. Pesole, S. Liuni, G. Grillo et al., "UTRdb and UTRsite: Specialized databases of sequences and functional elements of 5' and 3' untranslated regions of eukaryotic mRNAs. Update 2002," *Nucleic Acids Research*, vol. 30, no. 1, pp. 335–340, 2002.
- [19] T. N. Siegel, D. R. Hekstra, X. Wang, S. Dewell, and G. A. M. Cross, "Genome-wide analysis of mRNA abundance in two life-cycle stages of *Trypanosoma brucei* and identification of splicing and polyadenylation sites," *Nucleic Acids Research*, vol. 38, no. 15, pp. 4946–4957, 2010.
- [20] A. Wang, A. Pierce, K. Judson-Kremer et al., "Rapid analysis of gene expression (RAGE) facilitates universal expression profiling," *Nucleic Acids Research*, vol. 27, no. 23, pp. 4609–4618, 1999.
- [21] J. J. Chen, J. D. Rowley, and S. M. Wang, "Generation of longer cDNA fragments from serial analysis of gene expression tags for gene identification," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 97, no. 1, pp. 349–353, 2000.
- [22] A. Rochette, F. Raymond, J. M. Ubeda et al., "Genome-wide gene expression profiling analysis of *Leishmania major* and *Leishmania infantum* developmental stages reveals substantial differences between the two species," *BMC Genomics*, vol. 9, article 255, 2008.
- [23] P. J. Alcolea, A. Alonso, M. J. Gómez et al., "Transcriptomics throughout the life cycle of *Leishmania infantum*: high down-regulation rate in the amastigote stage," *International Journal for Parasitology*, vol. 40, no. 13, pp. 1497–1516, 2010.
- [24] B. Langmead, C. Trapnell, M. Pop, and S. L. Salzberg, "Ultrafast and memory-efficient alignment of short DNA sequences to the human genome," *Genome Biology*, vol. 10, no. 3, article R25, 2009.

- [25] H. Li and R. Durbin, "Fast and accurate short read alignment with Burrows-Wheeler transform," *Bioinformatics*, vol. 25, no. 14, pp. 1754–1760, 2009.
- [26] R. Larreta, M. Soto, L. Quijada et al., "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 Molecular Biology*, vol. 5, article 3, 2004.
- [27] C. Yao, J. E. Donelson, and M. E. Wilson, "The major surface protease (MSP or GP63) of *Leishmania* sp. Biosynthesis, regulation of expression, and function," *Molecular and Biochemical Parasitology*, vol. 132, no. 1, pp. 1–16, 2003.