Analysis and chromosomal mapping of Leishmania (Leishmania) amazonensis amastigote expressed sequence tags

Luciana Girotto Gentil, Fernanda Lasakovsvitsch, José Franco da Silveira, Márcia Regina Machado dos Santos, Clara Lúcia Barbieri

Departamento de Microbiologia, Imunologia e Parasitologia, Universidade Federal de São Paulo, Escola Paulista de Medicina, Rua Botucatu 862, 6º andar, 04023-062 São Paulo, SP, Brasil

The characterization of expressed sequence tags (ESTs) generated from a cDNA library of Leishmania (Leishmania) amazonensis amastigotes is described. The sequencing of 93 clones generated new L. (L.) amazonensis ESTs from which 32% are not related to any other sequences in database and 68% presented significant similarities to known genes. The chromosome localization of some L. (L.) amazonensis ESTs was also determined in L. (L.) amazonensis and L. (L.) major. The characterization of these ESTs is suitable for the genome physical mapping, as well as for the identification of genes encoding cysteine proteinases implicated with protective immune responses in leishmaniasis.

Key words: Leishmania (Leishmania) amazonensis - amastigotes - cDNA library

Sequencing of cDNA libraries to generate expressed sequence tags (ESTs) has been used for gene identification in Plasmodium falciparum (Chakrabarti et al. 1994), Trypanosoma brucei (El Sayed et al. 1995), Trypanosoma cruzi (Brandão et al. 1997, Verdún et al. 1998), and Leishmania (L.) major (Levick et al. 1996). The L. (L.) major genome project provided the sequence of the 36 chromosomes of the 32.8-megabase haploid genome of the parasite, predicting 911 RNA genes, 39 pseudogenes, and 8272 protein-coding genes (Ivens et al. 2005). In contrast, the genome of L. (L.) amazonensis, a causative agent of human cutaneous leishmaniasis in Brazil, is poorly studied and only 107 of its genes have been reported (http://ncbi.nlm.nih.gov). The availability of a cDNA library from amastigote forms of Leishmania can provide the characterization of genes implicated in parasite survival and host parasite interactions, opening perspectives for the development of new tools for disease control. The aim of the present work was the analysis of ESTs generated from a cDNA library of L. (L.) amazonensis amastigotes and their mapping in chromosomal bands.

In our laboratory we constructed a cDNA library from the amastigote form of L. (L.) amazonensis. It is important to emphasize that the amastigotes used for RNA extraction were derived from hamster foot lesions. This provides a measure of confidence of our cDNA library, since it was demonstrated that there is an increase in the magnitude of the transcript levels in Leishmania amastigotes proceeding from axenic cultures (Holzer et al. 2006). We obtained 40,000 clones and 100 were randomly selected and sequenced generating new L. (L.) amazonensis ESTs with an average insert size of 1.3 kb. These ESTs were sequenced from the 5’ and 3’ ends with T7 primer (forward) and SP6 primer (reverse) originating 147 sequences with an average size of 458 bp. These ESTs were compared to database sequences using the BLAST program. Sequence homologies identified by BLAST programs were considered statistically significant with a Poisson P value of ≤ 10^-4. Among 100 sequences, six matched with rRNA and one with hamster gene were excluded from further analysis. From a total of 93 EST sequences, 63 (68%) showed significant identity to gene sequences available on databases, whereas 30 ESTs (32%) did not have significant matches in databases and therefore were classified as unidentified. Genes from other Leishmania species were matched to 64.5% of the ESTs, 2.1% matched to L. (L.) amazonensis sequences, and 1.1% matched to genes from other organisms. Considering both the identified and the unidentified ESTs we could estimate the level of redundancy by use of SeqManII (DNASTar), a sequence analysis software. The overall diversity was 88.3% against 11.7% of redundancy. The most frequent genes on this analysis were those encoding the L. (L.) mexicana cysteine proteinase Lmcpb2.8 (GenBank accession no. Z49962) and ribosomal phosphoprotein S6 (GenBank accession no. AF045457), which appeared three and two times. This level of redundancy was significantly lower than those observed in a non normalized library of L. (L.) major (Levick et al. 1996) and in other protozoan parasites (Chakrabarti et al. 1994, El Sayed et al. 1995, Brandão et al. 1997) whose redundancy was about 27%.

Further analysis of our data were performed by considering only nonredundant ESTs that showed significant match in the databases. The distribution of the L. (L.) amazonensis ESTs according to their putative cellular functions is shown in the Table. Most clones are related to protein synthesis (15%) that encode predominantly

Nucleotide sequences data reported in this paper are available in the EMBL, GenBank, and DDJB data bases under the accession numbers DY763147-DY763293.

Financial support: Fapesp, CNPq

*Corresponding author: barbiericl@ecm.epm.br

Received 26 February 2007
Accepted 14 May 2007
ribozymal proteins, corroborating data previously obtained in other protozoa (El Sayed et al. 1995, Verdún et al. 1998, Levick et al. 1996, Dijkeng et al. 1998). Other clones are related to transport facilitation (10%), cellular biogenesis (8.3%), metabolism (6.7%), protein destination (5%), including two Leishmania cysteine proteinase genes, intracellular transport (3.3%), signal transduction (3.3%), and transcription (1.7%), whereas 41.7% were not identified to known proteins. This identification was based on the annotation present in the database GeneDB of L. (L.) major (www.gene-db.org/genedb/leish). Only two ESTs, LlaAm0013T7 and LlaAm0088 (GenBank accession nos. DY763238 and DY763159), showed identity to L. (L.) amazonensis genes deposited in database, translation elongation factor and Ca\(^{2+}\)-ATPase, respectively. This could be due to the scarcity of data relative to this Leishmania species. However, we identified several genes not yet described in L. (L.) amazonensis and one unidentified in trypanosomatids, the LlaAm0088SP6, a F1 ATP synthase (GenBank accession no. DY763167). We have also compared our results to previous studies employing mass spectrometry and microarray experiments for proteomic analysis of L. (L.) amazonensis differentiation (Dijkeng et al. 1998). Five ESTs identified in our studies were detected by these authors and two are exclusively expressed in amastigotes, the translation elongation factor-1 and the 40S ribosomal protein S12 (GenBank accession nos. DY763238 and DY763240, respectively). The other three ESTs were alpha-tubulin, ATPase beta subunit, and cysteine proteinase (GenBank accession nos. DY763245, DY763167, and DY763163, respectively). Recent data from microarray analysis of L. (L.) major and L. (L.) mexicana transcription are also in agreement with our results. Genes encoding cysteine proteinase and cytochrome b were predominantly transcribed in amastigote forms. Three ESTs identified in L. (L.) amazonensis (GenBank accession nos. DY763245, DY763151, and DY763244, respectively) correspond to alpha-tubulin, histone H2B and 60S ribosomal protein L13 transcripts detected almost exclusively in promastigotes, whereas one L. (L.) amazonensis EST (GenBank accession no. DY763267) corresponds to amino acid permease only detected in metacyclic promastigotes (Saxena et al. 2003, Akopyants et al. 2004, Almeida et al. 2004, Holzer et al. 2006), confirming that transcription is constitutive among trypanosomatids. The significant number of L. (L.) amazonensis ESTs belonging to a hypothetical protein category described in the present work corroborates data obtained by microarray analysis of L. (L.) mexicana and demonstrates the importance to characterize genes related to specific functions in Leishmania (Saxena et al. 2003).

The identified ESTs and genes have been often used as chromosomal markers for determination of molecular karyotype as well as for construction of physical and genetic maps in trypanosomatids (Ivens et al. 1998, Melville et al. 2000, Porcile et al. 2003). The sequence analysis of our clones showed that some ESTs had identities with sequences previously annotated in L. (L.) major chromosomes (www.ncbi.nlm.nih.gov). In order to confirm the chromosomal polymorphism often observed in trypanosomatids (Henriksson et al. 1990) as well as to identify new chromosomal markers, we mapped the L. (L.) amazonensis ESTs in pulsed field gel electrophoresis (PFGE) karyotype of L. (L.) amazonensis and L. (L.) major. For this analysis, eight L. (L.) amazonensis ESTs were chosen: two (LlaAm0053 and LlaAm0108) presented identity with chromosome 14, two (LlaAm0044 and LlaAm0047) with chromosome 19, three (LlaAm0134, LlaAm0048, and LlaAm0087) with chromosomes 4, 21, and 27, respectively, and one (LlaAm0088) with chromosomes 21 and 35 from L. (L.) major. The L. (L.) amazonensis cysteine proteinase genes Llacys1 and Llacys2 previously cloned in our laboratory (Lasakovitsch et al. 2003) and the 24S\(\alpha\)-rRNA were also used as probes and presented identity with L. (L.) major chromosomes 19, 8, and 27, respectively. A total of 25 chromosomal bands of L. (L.) amazonensis, ranging in size from 0.2 to \(\geq\) 2.2 Mb, were resolved by PFGE, corroborating previous results obtained by Conte and Cano (2005). Some of bands showed variable ethidium bromide-staining intensities possibly due to co-migration of chromosomes of similar size (Figure). In control experiments using total DNA of L. (L.) amazonensis or L. (L.) major as a probe, all chromosomal bands were labeled by the probe, confirming their integrity and identity as chromosomes (data not shown). The hybridization of ESTs, cysteine proteinase genes, and 24S\(\alpha\)-rRNA with L. (L.) amazonensis chromosomal bands showed that nine of them mapped in a single band, whereas two EST (LlaAm0088) and the 24S\(\alpha\)-rRNA hybridized to two chromosomal bands with different sizes (Figure). The results obtained with seven markers (LlaAm0044, LlaAm0047, LlaAm0048, LlaAm0087, LlaAm0088, Llacys1 gene, and 24S\(\alpha\)-rRNA) were in agreement with data available in the L. (L.) major GeneDB. These genes are localized in chromosomes with different size, suggesting that they may represent homologue chromosomes with different size, as it has been observed in trypanosomatids (Henriksson et al. 1990, Santos et al. 1999).

Three genetic markers (LlaAm0088, Llacys2 gene, and 24S\(\alpha\)-rRNA) showed some polymorphism when we
compared L. (L.) amazonensis and L. (L.) major karyotypes (Figure). The LlaAm0088 hybridized to two bands, 760 kb and 2000 kb, of L. (L.) amazonensis, suggesting a deletion of 400 kb in the homologue in L. (L.) major. An unequivocal evidence of chromosomal rearrangement could also be observed with the 24Sα-rRNA which mapped on a chromosomal band of 1200 kb in both species. In addition, a band of 1500 kb contains the 24Sα-rRNA in L. (L.) amazonensis. It is also possible to observe that the Llacys2 gene mapped on a 565 kb band in L. (L.) major and on a single band of 2000 kb in L. (L.) amazonensis. Among the nine markers that hybridized to a single band in L. (L.) amazonensis, five ESTs did not show any known identity, two were identified as a ribosomal protein (LlaAm0048) and ATP/ADP translocase (LlaAm0044), and two as cysteine proteinase genes (Llacys1 and Llacys2). These clones will be useful as chromosome-specific markers to identify polymorphism among L. (L.) amazonensis and other Leishmania species. The ESTs LlaAm0044, LlaAm0047, and Llacys1 hybridized in a same chromosomal band of 720 kb in L. (L.) amazonensis and L. (L.) major. Analysis of the chromosome 19 sequence from L. (L.) major (www.genedb.org/genedb/leish) showed that the clones LlaAm0044, LlaAm0047, and Llacys1 are located on 50 kb, 190 kb, and 620 kb of this chromosome, respectively. These results suggest that these genes are located in the same chromosome in L. (L.) amazonensis, indicating that they belong to a linkage group and can be used as markers for a specific chromosome. Similarly, the LlaAm0053 and LlaAm00108 ESTs mapped on a chromosomal band of the same size (630 kb) in both species and they could also be used as specific chromosomal markers. In contrast, the markers LlaAm0088, Llacys2 gene, and 24Sα-rRNA showed different hybridization patterns in the two Leishmania species, strengthening the chromosomal polymorphism previously reported among Leishmania species from the Old and New World (Britto et al. 1998).
Cloning of cysteine proteinase genes represents another interesting approach for use of the cDNA library described in the present work. Cysteine proteinases have been considered virulence factors in *Leishmania* (Denise et al. 2003) and targets for the design of vaccines (Rafati et al. 2001). Previous studies from our laboratory demonstrated that a cysteine proteinase of 30 kDa, predominantly expressed in amastigote forms of *L. (L.) amazonensis*, can confer protective immune responses against homologous infection in BALB/c mice (Beyrodt et al. 1997). Cloning of *L. (L.) amazonensis* Llacs1 and Llacs2 genes has led to the development of new schedules of immunization using both recombinant cysteine proteinases as well as Llacs1 and Llacs2 genes.

**ACKNOWLEDGMENTS**

To Simone Katz, Renato A Migliano Lopes, and Paulo B Paiva for technical assistance.

**REFERENCES**


Rafati S, Salmanian AH, Taheri T, Vafa M, Fasel N 2001. A protective cocktail vaccine against murine cutaneous leishma-
niasis with DNA encoding cysteine proteinases of *Leishmania major*. *Vaccine* 19: 3369-3375.

