Genes that encodes NAGT, MIF1 and MIF2 are not virulence factors for kala-azar caused by Leishmania infantum

Bruno Guedes Alcoforado Aguiar[1], Daniela Lemos Coelho[1], Dorcas Lamounier Costa[1], Betânia Paiva Drumond[2], Luiz Felipe Leomial Coelho[3], Livio Carvalho Figueiredo[4], Danielle Alves Zacarias[1], Jailthon Carlos da Silva[1], Diego Peres Alonso[5], Paulo Eduardo Martins Ribolla[5], Edna Aoba Yassui Ishikawa[6], Samara Belchior Gaido[1] and Carlos Henrique Nery Costa[1]

[1]. Laboratório de Leishmanioses, Instituto de Doenças Tropicais Natan Portella, Universidade Federal do Piauí, Teresina, PI. [2]. Laboratório de Virologia, Universidade Federal de Juiz de Fora, Juiz de Fora, MG, Brasil.. [3]. Laboratório de Vacinas, Universidade Federal de Alfenas, Alfenas, MG. [4]. Departamento de Agricultura, Centro de Ciências Humanas, Sociais e Agrárias, Universidade Federal da Paraíba, Bananeiras, PB. [5]. Departamento de Parasitologia, Instituto de Biotecnologia, Universidade Estadual Paulista, São Paulo, SP. [6]. Laboratório de Biologia Molecular, Núcleo de Medicina Tropical, Universidade Federal do Pará, Belém, PA.

ABSTRACT

Introduction: Kala-azar is a disease resulting from infection by Leishmania donovani and Leishmania infantum. Most patients with the disease exhibit prolonged fever, wasting, anemia and hepatosplenomegaly without complications. However, some patients develop severe disease with hemorrhagic manifestations, bacterial infections, jaundice, and edema dyspnea, among other symptoms, followed by death. Among the parasite molecules that might influence the disease severity are the macrophage migration inhibitory factor-like proteins (MIF1 and MIF2) and N-acetylglucosamine-1-phosphotransferase (NAGT), which act in the first step of protein N-glycosylation. This study aimed to determine whether MIF1, MIF2 and NAGT are virulence factors for severe kala-azar. Methods: To determine the parasite genotype in kala-azar patients from Northeastern Brazil, we sequenced the NAGT genes of L. infantum from 68 patients as well as the MIF1 and MIF2 genes from 76 different subjects with diverse clinical manifestations. After polymerase chain reaction (PCR), the fragments were sequenced, followed by polymorphism identification. Results: The nucleotide sequencing of the 144 amplicons revealed the absence of genetic variability of the NAGT, MIF1 and MIF2 genes between the isolates. The conservation of these genes suggests that the clinical variability of kala-azar does not depend upon these genes. Additionally, this conservation suggests that these genes may be critical for parasite survival. Conclusions: NAGT, MIF1 and MIF2 do not alter the severity of kala-azar. NAGT, MIF1 and MIF2 are highly conserved among different isolates of identical species and exhibit potential for use in phylogenetic inferences or molecular diagnosis.


INTRODUCTION

Leishmaniasis is a group of diseases that compromise skin, mucous membranes and visceral organs. Leishmaniasis diseases are zoonotic, widely distributed globally and are closely associated with poverty. As such, leishmaniasis is an important public health problem. The annual incidence is estimated to be approximately two million people, and approximately 350 million people are at risk of contracting the disease. Among the types of leishmaniasis, kala-azar is well known and is considered by the World Health Organization (WHO) to be one of the six most significant endemic diseases in the World.

The majority of infected individuals are asymptomatic; however, visceral leishmaniasis (VL) exhibits clinical and laboratory presentations that include a prolonged course of fever, pallor, weight loss, hepatosplenomegaly and pancytopenia. With these classic manifestations, clinical presentations including bleeding and bacterial infections are present in the most severe forms of the disease and are associated with mortality.

The identification of genetic variants is critical for the detection of a predisposition to various diseases. High conservation of genes suggests that a small number of species-specific genes are relevant to pathogenicity or virulence. Leishmania, the etiological agent of kala-azar, has 34-36 chromosomes. From a total of approximately 8,300 genes in each Leishmania species, slightly more than 200 are differentially distributed between the genomes of Leishmania infantum, Leishmania major and Leishmania braziliensis. The most divergent species is L. braziliensis, containing 49 different
genes, followed by *L. infantum*, with 27 genes specific to the species; the least divergent is *L. major*; with five genes.

The existence of two distinct patient profiles (classic and severe) might be a consequence of several factors such as the infective strength level, host immunosuppression, and the genetic background of the host, or it might be genetically determined by parasitic virulence factors. The identification of genetic changes in the pathogenic target genes is a starting point in the search for virulence factors. The term ‘virulence factors’ is used to describe proteins that are closely related to the severity of the disease or to genes encoding such peptides. Genes such as A2, which causes a virulent phenotype when introduced in *L. major*, or proteins such as the zinc metalloprotease glycopolypeptide (GP63) (leishmanolysin) which confer protection against lysis by the host.

The NAGT gene of *Leishmania* is a single-copy gene with 1,401 base pairs per haploid genome. The translated protein is a transmembrane protein of the endoplasmic reticulum (ER), identified as N-acetylglucosamine-1-phosphotransferase (NAGT), whose enzymatic action is the first step of N-glycosylation. Proteins that are extensively modified post-translationally, including GP63, can become unstable because of mutations affecting N-glycosylation, probably due to the decreased structural stability of leishmanolysin. This instability decreases the response against proteolytic degradation and causes the strain to be less virulent.

Two genes that encode orthologues of human macrophage migration inhibitory factor (MIF) in the *Leishmania* genus have been described. Human MIF is a major mediator of inflammation. This gene encodes a cytokine involved in cell-mediated immunity, immunoregulation, and inflammation. MIF plays a role in the regulation of macrophage function in host defense through the suppression of the anti-inflammatory effects of glucocorticoids. The presence of MIF orthologues in other parasites (*Eimeria* sp., *Trichinella* sp., *Plasmodium* sp. and *Brugia malayi*) is related to immune modulation. Macrophase migration inhibitory factor gene-deficient mice (MIF -/-) are susceptible to infection and develop severe disease, thus, *Leishmania* may be unstable because of mutations affecting N-glycosylation, probably due to the decreased structural stability of leishmanolysin. This instability decreases the response against proteolytic degradation and causes the strain to be less virulent.

Two sets of patients participated in the study. For the NAGT study, 35 complication-free kala-azar patients and 33 kala-azar patients with complications (bleeding, opportunistic infections, sepsis or death) were included as shown in Table 1. For the study of the MIF orthologues, 43 *L. infantum* isolates from other patients with uncomplicated kala-azar and 33 isolates from other patients with complications were analyzed for the presence of genetic polymorphisms (Table 1). The isolates were obtained from patients recruited from a reference hospital in Teresina, Brazil. The patients were referred from endemic areas of *L. infantum* infection in the neighboring States of Piauí and Maranhão. The clinical and laboratory diagnoses were confirmed by typical clinical presentations, such as fever, wasting, anemia and hepatosplenomegaly, in addition to reactive serology and the presence of the parasites in the bone marrow. The species identification was performed using monoclonal antibodies. The parasites were cultured in Novy-MacNeal-Nicolle (NNN) media with Schneider’s insect medium (Sigma, St. Louis, United States of America) supplemented with fetal calf serum (Cultilab, Campinas, Brazil). The cultures were incubated at 28°C for five days before DNA extraction.

**Methods**

Study design

DNA was extracted with Chelex® 100 resin (Bio-Rad, Hemel Hempstead, United Kingdom) and with a QIAamp® DNA mini kit (250) from Qiagen (Qiagen, Hilden, Germany). For the NAGT amplification, four primers were designed (L0, L2B, L3B and L5), based on the GenBank reference AF205934. For the MIF orthologue amplification, the primers were as follows: MIF1 forward 5’-ATGCCGGTCATTCAAACG-3’ and MIF2 reverse 5’-TCAAAAGTTAGCGCCGTTC-3’ and MIF1 reverse 5’-TTAGAAGTTTGTGCCGTTCCA-3’; MIF1 reverse 5’-TCAAAAGTTAGCGCCGTTC-3’ (Table 2), MIF2 reverse 5’-TCAAAAGTTAGCGCCGTTC-3’ (Table 2), and MIF2 forward 5’-ATGCCATTTCTGCAGACGA-3’. For the study of the MIF orthologues, 43 *L. infantum* isolates from other patients with uncomplicated kala-azar and 33 isolates from other patients with complications were analyzed for the presence of genetic polymorphisms (Table 1).

**DNA extraction, PCR and sequencing**

DNA was analyzed with Chelex® 100 resin (Bio-Rad, Hemel Hempstead, United Kingdom) and with a QIAamp® DNA mini kit (250) from Qiagen (Qiagen, Hilden, Germany). For the NAGT amplification, four primers were designed (L0, L2B, L3B and L5), based on the GenBank reference AF205934 (Table 2). For the MIF orthologue amplification, the primers were as follows: MIF1 forward 5’-ATGCCGGTCATTCAAACG-3’ and MIF1 reverse 5’-TCAAAAGTTAGCGCCGTTC-3’; MIF2 reverse 5’-TCAAAAGTTAGCGCCGTTC-3’ (Table 2), and MIF2 forward 5’-TCAAAAGTTAGCGCCGTTC-3’ (Table 2), and MIF2 forward 5’-ATGCCATTTCTGCAGACGA-3’ and MIF2 reverse 5’-TCAAAAGTTAGCGCCGTTC-3’ (Table 2).

<table>
<thead>
<tr>
<th>TABLE 1 - Rates of complications among selected patients* in the NAGT and MIF orthologues study.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>NAGT</strong></td>
</tr>
<tr>
<td><strong>n</strong></td>
</tr>
<tr>
<td>No complications</td>
</tr>
<tr>
<td>Complications**</td>
</tr>
<tr>
<td>Total</td>
</tr>
</tbody>
</table>

**NAGT**: N-acetylglucosamine-1-phosphate transferase; **MIF**: macrophage migration inhibitory factor-like protein. *For each gene target, NAGT and MIF orthologues, were used different patients. **Complications: bleeding, opportunistic infections, sepsis or death.
Mega 535. All 68 NAGT sequences were selected, with the maximum-likelihood estimation method, and the tree topologies were evaluated using the nucleotide and amino acid sequences of MIF1 (HLc1), MIF2 (HLc2), Lin33.2090, Lin33.2100, Lmj1740, Lmj1750, MIF-Label and MIF human, the nucleotide and amino acid percentages of identity among the Leishmania MIF orthologues were calculated using LALIGN software.36

Ethical considerations

This work is part of a research project, Influence of the Leishmania chagasi genotype over the visceral leishmaniasis pathology, which is registered by the National Council of Research Ethics (CONEP) under 0116.045.203-05 and was approved by the National Council of Ethics (CEP - Conselho Nacional de Ética) at the Federal University of Piauí in 12-14-2005. Since then, Leishmania strains have been isolated from infected patients at the Natan Portella Tropical Diseases Institute (IDTPN - Instituto de Doenças Tropicais Natan Portella). At the time, the patients were advised regarding the research and signed the clarified free-consent form. This work has not presented harm to any patient because the Leishmania data used in the project were previously collected and stocked at -70°C in liquid nitrogen. The patients were identified by numeric codes, which shielded their identities.

RESULTS

The nucleotide sequencing of the 144 amplicons (76 to the MIF orthologues and 68 to NAGT) revealed the absence of genetic variability of the NAGT, MIF1 and MIF2 genes among these analyzed isolates of L. infantum. The complete identity of all the MIF1 and MIF2 nucleotide sample sequences was observed with the Lin33.2090 and Lin33.2100 sequences of the L. infantum MIF orthologues (Figure 1).

In Figure 1, the cluster in which the MIF gene orthologues were located was divided into two distinct sub-clusters. The first cluster, located at the top of the tree, comprises the MIF1 (HLc1), Lin33.2090, Lmj1740 and LbrM33.V2.2010 genes. The second sub-cluster (lower) comprises the MIF2 (HLc2), Lin33.2100 and Lmj1750 genes. The LbrM33.V2.2030 gene from L. braziliensis appears to be more similar than the other MIF gene of Leishmania to the human MIF orthologues. Similarly, the phylogenetic tree of NAGT indicates strong gene similarity independence of the strain or species (Figure 2). The tree with a representative sequence of all the sequences generated in this work (3928 contig 1) and other sequences from NCBI reveals that all of the NAGT sequences from the L. donovani complex are clustered together; however, L. mexicana, L. major and L. tropica each appear in different clusters.

DISCUSSION

The data indicate that the clinical variability and severity of kala-azar could not be genetically attributed to the studied genes. Factors such as other parasitic genes and host determinants including host genetics, immune status, and nutrition

<table>
<thead>
<tr>
<th>Name</th>
<th>Target</th>
<th>Primer a</th>
<th>Primer b</th>
</tr>
</thead>
<tbody>
<tr>
<td>L0</td>
<td>NAGT</td>
<td>CACAGAAGCTCTTCTCTCAGCTC</td>
<td>ATGCCGGTCATTCAAAACG</td>
</tr>
<tr>
<td>L5</td>
<td>NAGT</td>
<td>GTGAGACGGTCCTCCTTACAAC</td>
<td>ATGCCATTCTGCAGACGA</td>
</tr>
<tr>
<td>L2B</td>
<td>NAGT</td>
<td>CGACTGCGATCACAATGCTC</td>
<td>TCAAAAGTTAGCGCCGTTC</td>
</tr>
<tr>
<td>L3B</td>
<td>NAGT</td>
<td>GGCACACGGTGATCTACCT</td>
<td>TGAGAGACGGTCCTCCTTACAAC</td>
</tr>
<tr>
<td>MIF1_f</td>
<td>MIF1</td>
<td>ATGCCGGTCATTCAAAACG</td>
<td>ATGCCATTCTGCAGACGA</td>
</tr>
<tr>
<td>MIF1_r</td>
<td>MIF1</td>
<td>TTGAGGATTTGTGCGCGTTCCA</td>
<td>TCAAAAGTTAGCGCCGTTC</td>
</tr>
<tr>
<td>MIF2_f</td>
<td>MIF2</td>
<td>ATGCCATTTCTGCAGACGA</td>
<td>ATGCCGGTCATTCAAAACG</td>
</tr>
<tr>
<td>MIF2_r</td>
<td>MIF2</td>
<td>TCAAAAGTTAGCGCCGTTC</td>
<td>ATGCCGGTCATTCAAAACG</td>
</tr>
</tbody>
</table>

NAGT: N-acetylglucosamine-1-phosphate transferase; MIF: macrophage migration inhibitory factor-like protein; PCR: polymerase chain reaction.

The DNA strands were sequenced using ABI 3730XLs for the NAGT gene and MegaBACE 1,000 for the MIF orthologues at least six times in both orientations. The nucleotide sequences were assembled using the CAP3 Sequence Assembling Program to generate a consensus sequence. The nucleotide sequences were aligned using the Clustal W software.

Phylogenetic analyses

Multiple sequence alignments were performed using Clustal W in Molecular Evolutionary Genetics Analysis - Mega 5. To calculate the evolutionary distances, phylogenetic trees were constructed by the maximum-likelihood estimation method, and the tree topologies were evaluated using Mega 5. All 68 NAGT sequences were selected, with Leishmania infantum NAGT [GenBank: AF205934], L. infantum NAGT var. 2 [GenBank: DQ836147], L. infantum var. 4 [GenBank: DQ836148], L. infantum NAGT var 7 [GenBank: DQ836149], Leishmania donovani [GenBank: DQ836150], Leishmania major [GenBank: AF205930.1], Leishmania tropica [GenBank: AF291678.1], Leishmania mexicana [GenBank: M96635.1], Trypanosoma cruzi [GenBank: XM_807416.1], and Leishmania infantum JPCM5 chromosome 36: 1589082bp – 1590482bp. Another phylogenetic tree was constructed with the identical settings in the presence of one representative sequence of the sequences generated in this work.
should be investigated\(^9\). Because of a likely polygenic influence on the phenotype of the disease, as suggested by the multiple molecular mechanisms associated with *Leishmania* survival in vertebrate hosts\(^{38,39}\), and the crucial involvement of the host response, additional kala-azar patients and other genes should be observed before a definitive conclusion is reached on the role of virulence factors in the pathogenesis of severe kala-azar. This type of information could be critical for drug and vaccine development because the virulence factors might be targeted by specific drugs and may be related to the quality and magnitude of the host immune response\(^40\).

The data support the hypothesis that these genes are conserved in the genome of *Leishmania*, particularly those belonging to the *L. donovani* complex; all of the species that we used from this complex were arranged in a single cluster (Figure 2). We hypothesize the importance of NAGT for the visceralization of *Leishmania* as a predisposing factor for VL. After the genome sequencing of *L. major, L. infantum* and *L. braziliensis*, it was observed that only a minority of genes are species-specific genes\(^{10,41}\). Thus, the difference in the clinical presentation of visceral, cutaneous and mucocutaneous leishmaniasis could not be the result of significant genetic differences throughout the genome. The clinical presentations are likely a consequence of critical point mutations in a few genes or of post-transcriptional or post-translational mechanisms\(^{41}\). We could not exclude the possibility that the studied isolates were differed with respect to gene or protein expression levels, which might be related to the different clinical outcomes of kala-azar.

The high degree of conservation of these genes between the isolates and with the published sequence of *L. infantum* suggests that these genes might be critical for the survival of the parasites in the vertebrate host. This conclusion is drawn from the lack of genetic variability that could impair the
MIF orthologue functions involved with parasite survival inside macrophages or the structural stability of leishmanolysin from N-glycosylation, in the case of NAGT.

The NAGT, MIF1 and MIF2 genes have a high sequence similarity to the *L. infantum* JPCM5 (NCBI) genes used for comparison and are not involved with the severity of kala-azar at the coding level, which suggests a critical role of these genes in the biological cycle of the parasite. Considering their conservation level in identical species, the products of these genes might be used in phylogenetic inferences or, eventually, for diagnostic purposes or second-generation vaccine development.

**ACKNOWLEDGMENTS**

We thank Dr. Guilherme Corrêa de Oliveira for providing the sequencing platform as well as Elisa Cupolillo and Carlos Henrique M. da Silva from FIOCRUZ for providing the standard strain of *L. infantum* L0579.

**CONFLICT OF INTEREST**

The authors declare that there is no conflict of interest.

**REFERENCES**


10. Deborah F. Smith, Christopher S. Peacock and AKC: Comparative genomics: From genotype to disease phenotype in the leishmaniasis.


