Leptomonas pyrrhocoris: Genomic insight into Parasite’s Physiology

Anzhelika Butenko¹,², Tamara da Silva Vieira³, Alexander O. Frolov⁴, Fred R. Oppendoes⁵, Rodrigo P. Soares³, Alexei Yu. Kostygov¹,⁴, Julius Lukes³,⁶,⁷ and Vyacheslav Yurchenko¹,²,⁸,⁹,α

¹Life Science Research Centre, Faculty of Science, University of Ostrava, 710 00 Ostrava, Czech Republic; ²Biological Centre, Institute of Parasitology, Czech Academy of Sciences, 370 05 České Budejovice (Budweis), Czech Republic; ³Centro de Pesquisas René Rachou/FIOCRUZ, Belo Horizonte, Minas Gerais, 30190-002, Brazil; ⁴Zoological Institute of the Russian Academy of Sciences, St. Petersburg 199034, Russia; ⁵de Duve Institute, Université Catholique de Louvain, 1200, Brussels, Belgium; ⁶Faculty of Science, University of South Bohemia, 370 05 České Budějovice (Budweis), Czech Republic; ⁷Canadian Institute for Advanced Research, Toronto, ON M5G1Z8, Canada; ⁸Institute of Environmental Technologies, Faculty of Science, University of Ostrava, 710 00 Ostrava, Czech Republic; ⁹Department of Pathology, Albert Einstein College of Medicine, Bronx, NY 10461, USA

Abstract: Background: Leptomonas pyrrhocoris is a parasite of the firebug Pyrrhocoris apterus. This flagellate has been recently proposed as a model species for studying different aspects of the biology of monoxenous trypanosomatids, including host – parasite interactions. During its life cycle L. pyrrhocoris never tightly attaches to the epithelium of the insect gut. In contrast, its dixenous relatives (Leishmania spp.) establish a stable infection via attachment to the intestinal walls of their insect hosts.

Material and Methods: This process is mediated by chemical modifications of the cell surface lipophosphoglycans. In our study we tested whether the inability of L. pyrrhocoris to attach to the firebug’s midgut is associated with the absence of these glycoconjugates. We also analyzed evolution of the proteins involved in proper lipophosphoglycan assembly, cell attachment and establishment of a stable infection in L. pyrrhocoris, L. seymouri, and Leishmania spp. Our comparative analysis demonstrated differences in SCG/L/R repertoire between the two parasite subgenera, Leishmania and Viania, which may be related to distinct life strategies in various Leishmania spp. The genome of L. pyrrhocoris encodes 6 SCG genes, all of which are quite divergent from their orthologs in the genus Leishmania. Using direct probing with an antibody recognizing the β-Gal side chains of lipophosphoglycans, we confirmed that these structures are not synthesized in L. pyrrhocoris.

Conclusion: We conclude that either the SCG enzymes are not active in this species (similarly to SCG5/7 in L. major), or they possess a different biochemical activity.

Keywords: LPG, Insect gut’s attachment, Host-parasite interaction, Monoxenous trypanosomatids, L. pyrrhocoris, SCG enzymes.

1. INTRODUCTION

Trypanosomatids are obligatory parasitic protists of the class Kinetoplastea [1]. Some of them (dixenous = having two hosts in their life cycle: Trypanosoma, Leishmania, and Phytomonas spp.) were intensely studied because they cause diseases in humans, domestic animals, and cultured plants. The vast majority of trypanosomatids is represented by monoxenous (with one host, insect, in their life cycle) species [2, 3]. For decades, these parasites were neglected and only sporadically used in physiological, biochemical, or molecular studies [4-6]. The situation has dramatically improved in recent years with realization that the evolutionary history of dixenous pathogens can be traced back to their monoxenous relatives [7], and with appreciation of the roles these parasites can play in insect communities [8, 9]. Trypanosomatids exhibit numerous unusual molecular and biochemical traits, such as RNA editing, trans-splicing, compartmentalized glycolysis, to name the most prominent ones [10-12].

The life cycles of dixenous Trypanosoma and Leishmania spp. are characterized in fine detail [13-15], while much less is known about the life cycle and molecular features of Phytomonas spp. [16, 17]. In contrast, the life cycles of monoxenous Trypanosomatidae remain virtually unknown. So far, they have been studied only in a handful of species (Leptomonas pyrrhocoris, Crithidia brevicula, C. fasciculata, Herpetomonas nabulicae, Blastocrithidia triatomae, B. papi, and few others, mainly using light microscopy techniques [18-22]. One outstanding question unifying these studies was to understand how trypanosomatids establish a stable infection, which in absolute majority of cases, seems to be achieved via the attachment of flagel-
lates to the intestinal walls of their insect hosts. Several mechanisms were proposed which included simple physical anchoring of the parasite’s flagellum, or specific interactions of modified flagella with microvilli or microvilli-free epithelial cuticle [23].

Leptomonas pyrrhocoris is a parasite of the cosmopolitan firebug, Pyrrhocoris apterus. The life cycle of this trypanosomatid is well understood and consists of developmental stages in the insect’s midgut, followed by the occasional invasion of the host’s hemolymph and salivary glands [24]. It is transmitted by coprophagy [18] and, supposedly, by necrophagy [25]. Importantly, at any stage of its life cycle, L. pyrrhocoris is never tightly attached to the cells, neither in the firebug’s gut nor in hemocoel [23]. The genome of this flagellate has been recently sequenced and assembled almost to chromosomal level [26]. Therefore, L. pyrrhocoris has been proposed as a model species for monoxenous trypanosomatids. Taken together, knowledge of the life cycle and the available genomic information provide us with a unique opportunity to analyze physiology of this organism and to compare it with its trypanosomatid relatives. The best candidates for this would be Leishmania spp. and Leptomonas seymouri [27]. The latter is an emerging opportunistic pathogen with sequenced genome [28-30] although the details of its life cycle and the specific host remain unknown. Such a comparison between these representatives of two distinct life styles is well justified because: i) the dixenous life cycle of Leishmania apparently evolved from a Leptomonas-like ancestor [7, 31, 32]; ii) mechanisms governing the attachment of Leishmania to the insect’s gut are known in considerable detail [33-35]; iii) genomes of numerous Leishmania spp. were sequenced [27, 36] and a comprehensive comparative analysis is thus feasible.

The development of Leishmania spp. within the alimentary tract of the sand fly’s host is mediated by chemical modifications of the lipophosphoglycan (LPG) on the cell surface [37-40]. This molecular complex consists of an 1-O-alkyl-2-lyso-phosphatidylinositol anchor, a conserved Gal(α1,6)Gal(α1,3)Gal(β1,3)Glc(α1)-PO4[Man(α1,3)Man (α1,4)-GlcN(α1) glycan core region, a conserved repetitive element Gal(β1,4)Man(α1)-PO4, and a terminal cap [41]. However, the repetitive units and caps may incorporate additional stage- and species-specific modifications. For example, in L. major and L. turania, the repetitive units are branched-off with β-1,3-galactosyl side chains (scfβ1,3-Gal) [37, 42]. This modification mediates attachment of the procyclic promastigotes to the midgut of the vector Phlebotomus papatasi [34]. Upon differentiation into infectious metacyclic promastigotes, scGal is further capped with arabinose, enabling parasite’s release and subsequent migration to the foregut [43]. The scfβ1,3-Gal is added by members of the scGal transferase family (SCG/L/R, 14 enzymes in L. major), while the α1,2-Ara attachment to the scβGal is facilitated by two scAra transferases, SCA1.2 [44, 45].

In this work, we tested whether the inability of L. pyrrhocoris to attach to the firebug’s midgut is associated with the absence of surface lipophosphoglycans. In addition, we analyzed evolution of the SCG/L/R and SCA-encoding genes in L. pyrrhocoris, L. seymouri, and Leishmania spp.

2. MATERIALS AND METHODS

2.1. Cultivation of Trypanosomatids and Experimental Infection of Pyrrhocoris apterus with Leptomonas pyrrhocoris

Leptomonas pyrrhocoris isolates PP1 [46] and ATCC 30974 were cultivated in Brain Heart Infusion (BHI) medium (Sigma-Aldrich, St. Louis, USA) supplemented with 10 µg/ml hemin, 100 µg/ml ampicillin, 100 µg/ml chloramphenicol and 50 µg/ml tetracycline at 25 °C as described previously [47]. Due to the low number of passages of the primary culture preserved in the collection, the isolate PP1 was chosen for experimental infection. Non-infected firebugs Pyrrhocoris apterus were starved for 4 days and then provided with a medium containing cultivated trypanosomatids [22]. The smears prepared from experimentally infected insects were prepared on day 10 post-infection.

For extraction and purification of GPI-anchored glycoconjugates, starter cultures represented by promastigotes of Leishmania major strain MHOM/IL/80/Friedlin and Leptomonas pyrrhocoris ATCC 30974 were grown in M199 medium supplemented with 10% heat-inactivated fetal bovine serum, 100 U/ml penicillin, 50 mg/ml streptomycin, 12.5 mM glutamine, 0.1 M adenine, 5 µg/ml hemin, and 40 mM HEPES, pH 7.4 at 25 °C. Two liters of BHI medium supplemented with 27 µg/ml adenosine and 5 µg/ml hemin were seeded with 200 ml of starter cultures. Cells were grown for two days at 25 °C with agitation to a density of 5 - 7 x 10^7 cells/ml, as described previously [39].

2.2. Light and Transmission Electron Microscopy

Smears on slides containing life stages of either species were air-dried, fixed with ethanol and stained with Giemsa. Cells were examined using Leica DM 2500 light microscope (Leica Microsystems, Wetzlar, Germany) equipped with UCMOS14000KPA 14-Mpx camera. For transmission electron microscopy (TEM), the insect gut sections were fixed for 2 days at 25 °C with agitation to a density of 5 - 7 x 10^7 cells/ml, as described previously [16].

2.3. Genomic Analysis and Phylogenetic Inferences

For the identification of the SCG/L/R and SCA gene family members, a BLASTP search was performed using these proteins of Leishmania major Friedlin as a query [44, 48] and annotated proteins of L. infantum JPCM5 (MCAN/ES/98/LLM-877), L. (Viannia) braziliensis M2904 (MHOM/BR/75M2904) [49], Leptomonas pyrrhocoris H10 [26] and L. seymouri ATCC 30220 [28] as a database with E-value cut-off of 10^-20 and the minimal coverage of 60%. Amino acid sequences of all identified proteins were aligned using Muscle v. 3.8.31 with default parameters [50]. Poorly aligned regions of the alignment were removed using Gblocks [51] with the following parameters: -b3 = 10; -b4 = 2; -b5 = h. The SCA and SCA-like proteins were analyzed separately from the SCG/L/R family members due to the notable divergence of their sequences. The average identities within the SCA/SCA-like and SCG/L/R groups were 37.53% and 27.39%, respectively, whereas the same value between them was only 5.09%. Maximum likelihood phylogenetic
trees were inferred using the IQTREE v.1.5.3 web-server with JTTDCMut + G4 model for SCG/L/R and LG + F + G4 model for SCA/SCA-like proteins, as selected by the built-in module [52]. Branch supports were assessed by bootstrapping with 100 standard replicates. Bayesian inference of phylogeny was accomplished in MrBayes 3.2.6 with analysis run for 1 million generations, sampling every 100th of them and other parameters of MCMC left as default [53]. The model of site heterogeneity was set based on IQTREE analyses, while the models of amino acid substitutions were assessed by the MrBayes (mixed amino acid model prior). The resulting models were WAG/Jones + F + G4 for SCG/L/R and Jones + F + G4 for SCA/SCA-like.

2.4. Extraction and Purification of Glycoconjugates

Glycoconjugates (LPG-like structures) from parasites were subjected to organic extraction in solvent E (H2O/ethanol/diethyl ether/pyridine/NH4OH, 15:15:5:1:0.017). The extract was dried by N2 evaporation, dissolved in 0.1 M acetic acid/0.1 M NaCl, and applied to a column of phenyl-Sepharose (2 ml), equilibrated in the same buffer. LPG was eluted using solvent E as described elsewhere [54].

2.5. Stains-all and Immunoblotting

In order to confirm purification, purified GPI-anchored glycoconjugates (~20 µg per lane) eluted from the Phenyl-Sepharose column were resolved by SDS-PAGE and subjected to stains-all technique as described previously [55]. Samples were also transferred to nitrocellulose membrane and subjected to immunoblotting with the WIC 79.3 monoclonal antibody (1:1,000), which recognizes terminal β1,3Gal sequences that branches off the Galβ1,4Manα1-PO4 repeat units in the L. major LPG [56].

3. RESULTS

3.1. Leptomonas pyrrhocoris Does not Attach to the Insect’s Gut

It was observed that during its life cycle L. pyrrhocoris never tightly attaches to the epithelium of the insect gut [18, 23]. To test this, we prepared smears and ultrathin sections of the P. apterus gut of specimens experimentally infected with L. pyrrhocoris, and examined them by light and electron microscopy (Fig. 1). Consistent with previous observations, we have never detected L. pyrrhocoris cells attached to the host midgut epithelium.

3.2. Leptomonas pyrrhocoris Encodes Almost Complete Set of Divergent SCG/L/R and SCA Enzymes

In order to analyze genes potentially involved in the biosynthesis of LPG in L. pyrrhocoris, we first identified a set of orthologs to the well-characterized SCA and SCG/L/R families of L. major Friedlin [44].

Thirteen of the SCG/L/R family genes found in L. pyrrhocoris (compared to 14 in L. major Friedlin, 12 in L. infantum JPCM5, and 17 in L. braziliensis M2904) cluster together in the same assembled contig (LpyrH10_27) and span about 90 kb (Fig. 2A). This contrasts with the situation in all Leishmania spp. analyzed so far, in which these genes are located on different chromosomes. Of note, only a single gene of this family was identified in our analysis of L. seymouri genome.

To assess relationships within this group of proteins, we built a phylogenetic tree with all sequences of the SCG/L/R family members of L. major, L. infantum, L. braziliensis, L. pyrrhocoris, and L. seymouri (Fig. 2B). There are 6 orthologs of SCG proteins in L. pyrrhocoris (LpyrH10_27_290 – LpyrH10_27_330, and LpyrH10_27_420) com-

Fig. (1). Leptomonas pyrrhocoris does not attach to the cells of insect’s gut. Light- (A, Giemsa-stained) and transmission electron microscopy (B) of the insect’s gut. Selected trypanosomatid cells are marked by arrowheads. The intestinal epithelium cell is labeled IEC. Scale bars are 20 µm (A) and 2 µm (B).
pared to 7 (SCG1–7) in L. major. These protein sequences form a divergent monophyletic group sister to all Leishmania SCG sequences. Confirming a previous observation [44], SCG5/7 differ from other SCG proteins. Interestingly, all 8 SCG orthologs from L. braziliensis cluster together with SCG1–6, while 4 SCG orthologs in L. infantum belong to the SCG5/7 group (Fig. 2B). It appears that separation of SCG5/7 from the rest of the SCG proteins occurred before radiation of Leishmania spp. The species-specific clusters on the tree suggest that each species has experienced an independent expansion of these genes.

The SCG5/7 enzymes are inactive in L. major [45], implying that L. infantum uses mechanisms of attachment to the insects’ gut that differ from those of L. major and L. braziliensis. Indeed, a structural analysis of the LPG molecules from L. infantum and L. braziliensis demonstrated that they are different from their L. major counterparts [39, 40, 57]. However, whether divergent L. pyrrhocoris orthologs play the same functional role in the biosynthesis of LPG (addition of the scβ1,3-Gal) as they do in L. major cannot be assessed from the genomic analysis.

The situation with other members of the SCG/L/R family is more complicated. LpyrH10_27_280 is sister to the SCGL+ SCGR2/3/5/6 cluster and therefore it cannot be assigned to any of the known groups of these proteins. Meanwhile, LpyrH10_27_220, LpyrH10_27_240, LpyrH10_27_260, and LpyrH10_27_270 are phylogenetically related to SCGR2/3/5/6, while LpyrH10_27_430 and LpyrH10_27_510 cluster together with SGR1/4 and SCGL, respectively (Fig. 2B). Again, we observed evidence of independent multiplication of these genes, similar to the case of SCGs. A single gene of the whole SCG/L/R family present in L. seymouri (Lsey_0379_0040) belongs to the SCGL group along with LpyrH10_27_510 of L. pyrrhocoris, LinJ.14.1500 of L. infantum, and LmjF.14.1400 of L. major (Fig. 2B).

Three orthologs of the SCA proteins identified in L. pyrrhocoris are phylogenetically related to the SCA1/2 (LpyrH10_27_440 and LpyrH10_27_450) and SCA-like protein (LpyrH10_05_780) of L. major (Fig. 2C). While orthologs of SCA or SCA-like proteins are present in L. seymouri (Lsey_0703_010 and Lsey_0144_150, respectively) and L. infantum (LinJ.02.190 and LinJ.34.530, respectively), they are absent in Leishmania braziliensis.

### 3.3. Leptomonas pyrrhocoris Does not Express LPG with β-Gal Side Chains

GPI-anchored LPG-like structures were purified from L. major and L. pyrrhocoris as confirmed by stains-all technique (Fig. 3A). After probing the membrane with the WIC 79.3 antibody only the LPG of L. major (positive control) was recognized (Fig. 3B). These data suggest that LPG-like structures from L. pyrrhocoris are devoid of the β-Gal side chains.

**DISCUSSION AND CONCLUSION**

The attachment of Leishmania spp. to the insects’ gut is governed by lipophosphoglycans on their cell surface [41, 58]. Thus, enzymes involved in biosynthesis of LPG are important for understanding the physiology and behavior of the parasitic protists in their hosts. It is a family of side chain galactosyltransferases (SCGs) that in Leishmania determines proper LPG assembly, cell attachment and establishment of a stable infection [45]. These side chains are further modified with the arabinose moiety ensuring detachment and subsequent transmission of the Leishmania species to the vertebrate host [34, 37, 43]. The enzymes involved in this process are side chain arabinosyltransferases (SCAs) [59]. The extended SCG/L/R family in L. major Friedlin contains 14 members [44]. Our comparative analysis demonstrated that there are 12 and 17 SCG/L/R enzymes in L. infantum JPCM5 and L. braziliensis M2904, respectively. The expansion of this family in L. braziliensis (8 SCG orthologs compared to 6 in L. major and 7 in L. infantum; 5 SCG/L/R orthologs compared to 2 in L. major and L. infantum), combined with the absence of SCAs or SCA-like proteins in this species, may be related to the differences in various Leishmania spp. life strategies. The New World leishmanias of the subgenus Viannia, such as L. braziliensis, are peripylarian parasites that parasitize in the hindgut before migrating forward into the midgut. In contrast, most Leishmania species (subgenus Leishmania) are suprapylarian parasites and they develop in the midgut of their insect host [60]. Of note, enzymes of this family can be detected only within Leishmaninae and not in other trypanosomatids with sequenced and annotated genomes. This can be explained either by high divergence of these proteins or by their absence in other species.

Another interesting aspect of our comparative analysis is clustering of SCG orthologs of L. infantum. They appear different from the related enzymes in L. major or L. braziliensis and group together with presumably inactive SCG5/7 sequences. The molecular mechanisms of LPG synthesis and modification in L. infantum remain to be investigated further. In most cases, they are governed by β-glucosyl transferases [39, 61, 62]. The tree topology demonstrating the independent expansion of SCGs implies that these proteins may be responsible for species-specific functions.

L. pyrrhocoris encodes 6 SCG enzymes (Fig. 2B), all of which are only distantly related to their SCGs orthologs in the genus Leishmania. Direct probing with an antibody recognizing the β-Gal side chains of LPG confirmed that these structures are not synthesized (Fig. 3B). There are two alternative explanations for these observations. Either the SCG enzymes are not active in L. pyrrhocoris (similarly to SCG5–7 in L. major), or they possess a different biochemical activity. We favor the second scenario because these genes in L. pyrrhocoris are quite divergent from those of Leishmania and in many cases expanded independently.

In conclusion, trypanosomatids exhibit a wide range of adaptations to the different hosts and environments that they circulate in. The survival of these parasites depends on glycoconjugates enabling them to circumvent adverse conditions they face in their hosts. Genomic and functional studies of the enzymes responsible for the assembly of those molecules may illuminate the mechanisms used by trypanosomatids for interactions with their hosts.
Fig. (2). SCG/L/R orthologs in *Leptomonas pyrrhocoris*. (A) Schematic organization of the genomic locus containing identified SCG/L/R genes (black arrows). Grey arrows and rectangles are genes not involved in LPG biosynthesis. Numbering within LpyrH10_27 contig and predicted SCG/L/R affiliation are shown below the scheme. (B and C) Midpoint-rooted phylogenetic trees of *Leishmania major* SCG/L/R (B) and SCA (C) proteins along with their orthologs from *Leptomonas pyrrhocoris*, *Leptomonas seymouri*, *Leishmania braziliensis*, and *Leishmania infantum*. Numbers at nodes indicate bootstrap percentage/posterior probability. Dots mark branches with maximal statistical support. Dashes (-) indicate bootstrap support below 50%, posterior probability below 0.5, or different topology. The scale bar denotes the number of substitutions per site.

Fig. (3). Stains-all (A) and Western blotting (B) analysis of purified GPI-anchored glycoconjugates from *Leishmania major* (positive control) and *Leptomonas pyrrhocoris* parasites. Glycoconjugates (20 µg per lane) from late-log stage cultures were probed with the WIC 79.3 monoclonal antibody (1:1000), specific for β1,3Gal residues of LPG. Protein ladder sizes on the left are in kDa.

**ETHICS APPROVAL AND CONSENT TO PARTICIPATE**

Not applicable.

**HUMAN AND ANIMAL RIGHTS**

No Animals/Humans were used for studies that are base of this research.

**CONSENT FOR PUBLICATION**

Not applicable.

**CONFLICT OF INTEREST**

The authors declare no conflict of interest, financial or otherwise.

**ACKNOWLEDGEMENTS**

We thank the members of our laboratories for stimulating discussions and Dr. d'Avila-Levy (Fundação Oswaldo Cruz,
Leptomonas pyrrhocoris: Genomic insight into Parasite’s Physiology

Current Genomics, 2018, Vol. 19, No. 2 155

Rio de Janeiro, Brazil) for providing the Leptomonas pyrrhocoris ATCC 30974 strain (COLPROT-026). Support from the Czech Grant Agency awards 17-10656S to V.Y., 16-18699S to J.L. and V.Y., and 15-16406S to A.B. and Moravskoslezský kraj research initiative 01211/2016/RRC to V.Y. and A.Y.K., and the EU COST Action CM1307 to F.R.O. and J.L. is kindly acknowledged. The contribution of A.O.F. (experimental infection of insects, light and electron microscopical studies) was supported by the Russian Foundation for Basic Research grant 15-29-02734. R.P.S. and T.S.V. were supported by CNPq (305065/2016-5) and FAPEMIG (PPM-00102-16). This work was also financially supported by the Ministry of Education, Youth and Sports of the Czech Republic in the “National Feasibility Program I”, project LO1208 “TEWEP” and the EU structural funding Operational Programme “Research and Development for Innovation”, project CZ.1.05/2.1.00/19.0388.

REFERENCES


[4] Babchi, J.; Lambros, C.; Goldberg, B.; Hutner, S.H.; de Carvalho, M.; An unprecedented non-canonical nuclear genetic code with all 46 codons encoding for Basic Research A.O.F. (experimental infection of insects, light and electron microscopic studies) was supported by the Russian Foundation for Basic Research grant 15-29-02734. R.P.S. and T.S.V. were supported by CNPq (305065/2016-5) and FAPEMIG (PPM-00102-16). This work was also financially supported by the Ministry of Education, Youth and Sports of the Czech Republic in the “National Feasibility Program I”, project LO1208 “TEWEP” and the EU structural funding Operational Programme “Research and Development for Innovation”, project CZ.1.05/2.1.00/19.0388.


