

***Trypanosoma rangeli* Transcriptome Project: Generation and analysis of expressed sequence tags**

Cristiane Quimelli Snoeijer¹, Gisele Fernanda Picchi²,
Bibiana Paula Dambrós¹, Mário Steindel¹, Samuel Goldenberg²,
Stênio Perdigão Fragoso², Daniel Macedo Lorenzini¹ and
Edmundo Carlos Grisard*¹

Address: ¹Departamento de Microbiologia e Parasitologia, Universidade Federal de Santa Catarina, Caixa postal 476, Santa Catarina, Brazil, 88040-900 and ²Instituto de Biologia Molecular do Paraná, Paraná, Brazil

Email: Cristiane Quimelli Snoeijer - crisquimelli@hotmail.com; Gisele Fernanda Picchi - gpicchi@tepar.br;
Bibiana Paula Dambrós - bibipaula@pop.com.br; Mário Steindel - ccb1mst@ccb.ufsc.br; Samuel Goldenberg - sgoldenb@tepar.br;
Stênio Perdigão Fragoso - sfragoso@tepar.br; Daniel Macedo Lorenzini - dloren@usp.br; Edmundo Carlos Grisard* - grisard@ccb.ufsc.br

* Corresponding author

Published: 13 May 2004

Received: 09 February 2004

Kinetoplastid Biology and Disease 2004, 3:1

Accepted: 13 May 2004

This article is available from: <http://www.kinetoplastids.com/content/3/1/1>

© 2004 Snoeijer et al; licensee BioMed Central Ltd. This is an Open Access article: verbatim copying and redistribution of this article are permitted in all media for any purpose, provided this notice is preserved along with the article's original URL.

Abstract

Trypanosoma rangeli is an important hemoflagellate parasite of several mammalian species in Central and South America, sharing geographical areas, vectors and reservoirs with *T. cruzi*, the causative agent of Chagas disease. Thus, the occurrence of single and/or mixed infections, including in humans, must be expected and are of great importance for specific diagnosis and epidemiology. In comparison to several Trypanosomatidae species, the *T. rangeli* biology and genome are little known, reinforcing the needs of a gene discovery initiative. The *T. rangeli* transcriptome initiative aims to promote gene discovery through the generation of expressed sequence tags (ESTs) and Orestes (ORF ESTs) from both epimastigote and trypomastigote forms of the parasite, allowing further studies of the parasite biology, taxonomy and phylogeny.

The *Trypanosoma rangeli* EST project

Trypanosoma rangeli and *T. cruzi* are hemoflagellate protozoan parasites transmitted by triatomine bugs in a wide, overlapped distribution area in Central and South America [1,2]. *T. cruzi* is the etiological agent of Chagas disease that affects 16 million people [3], whereas *T. rangeli* infection does not seem to be pathogenic for vertebrate hosts [1]. Beyond their morphological similarity, the sharing of several antigenic determinants allows the occurrence of inconclusive or wrong serological diagnosis of human infection, constituting a serious problem for the epidemiology of both etiologies [4,5]. Nowadays, only few methods or specialized techniques can undoubtedly

differentiate *T. cruzi* from *T. rangeli*, being not available for routine diagnosis [2]. Moreover, the controversial taxonomic position of *T. rangeli* has been discussed for several years. Presenting a remarkable pleomorphism and biological characteristics of distinct *Trypanosoma* Subgenus, the *T. rangeli* biology in its vertebrate hosts is still awaiting to be revealed [2]. Thus, it is necessary to develop new strategies in order to allow specific differentiation in a fast, secure, easy and economically viable way, to provide valuable information on genetic markers and to support functional studies in order to clarify the parasite biology and taxonomic position. Comparing to the actual number of *T. cruzi* entries in the GenBank (42,410 nucle-

otides and 1,602 proteins), a small number of *T. rangeli* nucleotide sequences or genes (68) and proteins (42) have been characterized and the overall picture of the genome is still absent. Due the absence of such information, little is known about important biological features such as the host-parasite relationships at the molecular level and the taxonomic position of the parasite [5,6]. Considering the ongoing initiatives over the *T. cruzi* genome [7] and the usefulness of cDNA partial sequencing to generate expressed sequence tags (ESTs) or Orestes (ORF ESTs) as a rapid and efficient approach to establish a detailed profile of the genes and the gene expression as well [8-14], we have started the "Trypanosoma rangeli EST Project" under the auspices of CNPq, a Brazilian Government Agency. Considering the importance of this parasite and its under representation on the current genomic/proteomic scenario, the project aims to generate an overview of the *T. rangeli* transcriptome through sequencing ESTs and ORESTES from both epimastigotes and trypomastigotes forms of the parasite. Initially, three distinct normalized and non-directional *T. rangeli* cDNA libraries were constructed using epimastigote forms of the Choachi strain [15]. Total RNA was obtained using standard protocols [16] followed by phenol-chloroform extraction and polyA⁺ RNA was prepared using the Oligotex RNA purification kit (Qiagen, Valencia). The cDNA libraries were constructed by cloning size-purified fragments in pGEM T-Easy plasmid (Promega, Madison) using the Clontech PCR-SelectTM cDNA Subtraction kit (BD Biosciences, Franklin Lakes). A total of ~3,360 clones were obtained and the validation of the libraries was performed by sequencing a single end of the inserts in a MegaBace 1000[®] equipment (Molecular Dynamics/Amersham Biosciences, Piscataway) using the DYEnamic^{ET} dye terminator kit (Amersham Biosciences). The mean size of the inserts for all cDNA libraries was of ~750 bp as revealed by *EcoRI* restriction analysis. The obtained sequences were edited and analyzed using the new ESTonSQL system under development in collaboration of D.M. Lorenzini with the UFSC Bioinformatics laboratory <http://www.bioinformatica.ufsc.br>. ESTs with 150 bp and phred quality > 10 <http://www.phrap.org> were considered valid by ESTonSQL and were analyzed by comparison with the EMBL-EBI, GenBank, Swissprot and Interpro databases. Sequences < 150 bp or with Phred value < 10 were labeled as invalid ESTs and stored. Up to now, a total of 656 valid ESTs were generated among which, 386 showed similarity with trypanosomatid sequences (Table 1). Several matches were observed with distinct organisms, but considering only the trypanosomatids database, 168 sequences were similar to *T. cruzi* sequences and only 20 showed similarity with *T. rangeli* sequences (Table 2). Despite the preliminary analysis, several ESTs revealed similarity to sequences not previously reported for *T. rangeli* such as a human topoisomerase III (GenBank accession

AAH02432) and a putative *Leishmania major* quinone oxidoreductase (NP_859482). Over 37% of the valid ESTs returned no hits and may represent *T. rangeli* specific genes, unknown genes or even 5' or 3' end untranslated regions (UTRs), proving the usefulness of this type of strategy to gather new information and the needs of genomic studies of this parasite. Along with the continuing EST generation by sequencing both ends of all obtained clones, ORESTES (ORF ESTs) profiles [17] will be obtained from the same libraries used for EST generation. While ESTs are generated by sequencing the 5' and/or 3' ends of cDNA clones, the ORESTES technique allows the generation of sequences belonging to the middle portion of the genes [17]. Through the combination of both EST and ORESTES approaches to study the transcriptome of a target organism, a significant number of coding regions can be rapidly obtained. Having described a method to induce *T. rangeli* metacyclogenesis *in vitro* [18], our group is currently working on a cDNA library of this parasite form in order to perform an intra and inter-specific comparative study of both epimastigotes and trypomastigote ESTs and ORESTES.

Conclusions

The number of studies involving *T. rangeli* is increasing due several aspects such as i) the importance on Chagas disease diagnosis and epidemiology, ii) the unknown life cycle on vertebrate hosts and iii) the contradictory taxonomic position of the parasite. Several molecular approaches have been used during the last few years to address these points. However, in order to perform comparative transcriptomics to address taxonomic studies, to bring up new markers for specific diagnosis, to discover, understand and/or compare some biological processes such as the involvement of the 3' end UTR on the post-transcriptional regulation of gene expression or even to assess intra-specific variability, a genomic panorama of the parasite must be revealed. Upon the release of the *T. cruzi*, *T. brucei* and *L. major* genome databases, not so far from now, the existence of well annotated ESTs and ORESTES from distinct *T. rangeli* forms will allow comparative transcriptomic studies with the state of the art technologies such as microarrays. Upon the end of the project, free accession to the *T. rangeli* transcriptome data, as well as tools for comparative genomics/transcriptomics under development, will be available through the UFSC Bioinformatics Laboratory website <http://www.bioinformatica.ufsc.br>.

Competing interests

C.Q. Snoeijer and B.P. Dambrós are Research Scholars from Capes and CNPq (Brazilian Government Agencies), respectively. The *T. rangeli* EST project is funded by a MCT/CNPq grant. The ESTonSQL system is under devel-

Table 1: Current status of the *Trypanosoma rangeli* transcriptome project.

EST matching with:	Number of matches*	(%)*
Trypanosomatids	386	58.84
Other organisms	25	3.81
Unknown	245	37.35
Total	656	100

*Updated values can be obtained at <http://www.bioinformatica.ufsc.br/trangeli>

Table 2: Actual number of expressed sequence tags (ESTs) generated by the *Trypanosoma rangeli* transcriptome project distributed according to the similarity with sequences of distinct species.

Species	Number of ESTs*
<i>Trypanosoma cruzi</i>	168
<i>T. brucei brucei</i>	77
<i>T. brucei</i>	25
<i>T. rangeli</i>	20
<i>T. carassii</i>	3
<i>T. brucei rhodesiense</i>	2
<i>T. congolense</i>	1
<i>Leishmania major</i>	71
<i>L. infantum</i>	15
<i>L. tarentolae</i>	1
<i>Neurospora crassa</i>	8
<i>Homo sapiens</i>	5
<i>Vibrio parahaemolyticus</i>	3
<i>Ictalurus punctatus</i>	3
<i>Plasmodium yoelii yoelii</i>	2
<i>Dictyostelium discoideum</i>	1
<i>Emericella nidulans</i>	1
<i>Mycine glutinosa</i>	1
<i>Oryza sativa</i>	1
<i>Shigella flexneri</i>	1
<i>Trypanoplasma borreli</i>	1
<i>Zea mays</i>	1
Total	656

*Updated values can be obtained at <http://www.bioinformatica.ufsc.br/trangeli>

opment by D.M. Lorenzini, to whom inquiries must be sent.

Author's contributions

The present work is part of C.Q. Snoeijer Masters thesis. All other authors have equally contributed to this work.

Acknowledgments

To the IBMP staff for their technical support on the cDNA libraries construction and to Dr. Ana Tereza R. de Vasconcellos from the Laboratório Nacional de Computação Científica (LNCC) for helping on bioinformatics analysis.

References

1. D'Alessandro A, Saravia NG: ***Trypanosoma rangeli***. In: *Protozoal Diseases* Edited by: Herbert M Gilles. London: Arnold Press; 1999:398-412.
2. Grisard EC, Steindel M, Guarnieri AA, Eger-Mangrich I, Campbell DA, Romanha AJ: **Characterization of *Trypanosoma rangeli* strains isolated in Central and South America: overview**. *Mem Inst Oswaldo Cruz* 1999, **94**:203-209.
3. Guhl F, Jaramillo C, Carranza JC, Vallejo GA: **Molecular characterization and diagnosis of *Trypanosoma cruzi* and *T. rangeli***. *Arch Med Res* 2002, **33**:362-370.
4. Afchain D, LeRay D, Fruit J, Capron A: **Antigenic make-up of *Trypanosoma cruzi* culture forms: Identification of a specific component**. *J Parasitol* 1979, **65**:507-514.
5. Grisard EC: **Salivaria or Stercoraria? The *Trypanosoma rangeli* dilemma**. *Kinet Biol Dis* 2002, **1**:5.
6. Stevens JR, Teixeira MM, Bingle LE, Gibson WC: **The taxonomic position and evolutionary relationships of *Trypanosoma rangeli***. *Int J Parasitol* 1999, **29**:749-757.

7. Verdun RE, Di Paolo N, Urmenyi TP, Rondinelli E, Frasch AC, Sanchez DO: **Gene discovery through expressed sequence tag sequencing in *Trypanosoma cruzi***. *Infect Immun* 1998, **66**:5393-5398.
8. Adams MD: **Expressed sequence tags as tools for physiology and genomics**. In: *Automated DNA sequencing and analysis* Edited by: Adams MD, Fields C, Venter JC. London: Academic Press; 1994:71-76.
9. Ajioka JW, John CB, Brunk BP, Hehl A, Hillier L, Manger ID, Marra M, Overton GC, Roos DS, Wan KL, Waterston R, Sibley LD: **Gene discovery by EST sequencing in *Toxoplasma gondii* reveals sequences restricted to the Apicomplexa**. *Genome Res* 1998, **8**:18-28.
10. Dijkeng A, Aguda C, Donelson JE, Majiwa PAO: **Generation of expressed sequence tags as physical landmarks in the genome of *Trypanosoma brucei***. *Gene* 1998, **221**:93-106.
11. Fan J, Minchella DJ, Day SR, McManus DP, Tiu WU, Brindley PJ: **Generation, identification and evaluation of expressed sequence tags from different developmental stages of the asian blood fluke *Schistosoma japonicum***. *Biochem Biophys Res Comm* 1998, **252**:348-356.
12. Levick MP, Blackwell JM, Connor V, Coulson RMR, Miles A, Smith HE, Wan KL, Ajioka JW: **An expressed sequence tag analysis of a full-length, spliced-leader cDNA library from *Leishmania major* promastigotes**. *Mol Biochem Parasitol* 1996, **76**:345-348.
13. Porcel BM, Tran AN, Tammi M, Nyarady Z, Rydåker M, Urmenyi TP, Rondinelli E, Pettersson U, Andersson B, Åslund I L: **Gene survey of the pathogenic protozoan *Trypanosoma cruzi***. *Genome Res* 2000, **10**:1103-1107.
14. Verjovski-Almeida S, Demarco R, Martins EAL, Guimarães PEM, Ojopi EPB, Paquola ACM, Piazza JP, Nishiyama MY, Kitajima JP, Adamson RE, Ashton PD, Bonaldo MF, Coulson PS, Dillon GP, Farias LP, Gregorio SP, Ho PL, Leite RA, Malaquias LCC, Marques RCP, Miyasato PA, Nascimento Alto, Ohlweiler FP, Reis EM, Ribeiro MA, Sá RG, Stukart GC, Soares MB, Gargioni C, Kawano T, Rodrigues V, Madeira AMBN, Wilson RA, Menck CFM, Setubal JC, Leite LCC, Dias-Neto E: **Transcriptome analysis of the acelomate human parasite *Schistosoma mansoni***. *Nat Genet* 2003, **35**:148-157.
15. Schotelius J: **Neuraminidase fluorescent test for differentiation of *Trypanosoma cruzi* and *Trypanosoma rangeli***. *Trop Med Parasitol* 1987, **38**:323-327.
16. Chomczynski P, Sacchi N: **Single-Step Method of RNA Isolation by Acid Guanidinium Thiocyanate-Phenol-Chloroform Extracion**. *Anal Biochem* 1987, **62**:156-159.
17. Dias Neto E, Correa RG, Verjovski-Almeida S, Briones MR, Nagai MA, Da Silva W JR, Zago MA, Bordin S, Costa FF, Goldman GH, Carvalho AF, Matsukuma A, Baia GS, Simpson DH, Brunstein A, De Oliveira PS, Bucher P, Jongeneel CV, O'Hare MJ, Soares F, Brentani RR, Reis LF, De Souza SJ, Simpson AJ: **Shotgun sequencing of the human transcriptome with ORF expressed sequence tags**. *Proc Natl Acad Sci USA* 2000, **97**:3491-3496.
18. Koerich LB, Emmanuelle-Machado P, Santos K, Grisard EC, Steindel M: **Differentiation of *Trypanosoma rangeli*: High production of infective trypomastigote forms in vitro**. *Parasitol Res* 2002, **88**:21-25.

Publish with **BioMed Central** and every scientist can read your work free of charge

"BioMed Central will be the most significant development for disseminating the results of biomedical research in our lifetime."

Sir Paul Nurse, Cancer Research UK

Your research papers will be:

- available free of charge to the entire biomedical community
- peer reviewed and published immediately upon acceptance
- cited in PubMed and archived on PubMed Central
- yours — you keep the copyright

Submit your manuscript here:
http://www.biomedcentral.com/info/publishing_adv.asp

