

A Comparative *In Silico* Study of the Antioxidant Defense Gene Repertoire of Distinct Lifestyle Trypanosomatid Species

Ingrid Thaís Beltrame-Botelho^{1,2}, Carlos Talavera-López³, Björn Andersson³,
Edmundo Carlos Grisard¹ and Patricia Hermes Stoco¹

¹Departamento de Microbiologia, Imunologia e Parasitologia, Universidade Federal de Santa Catarina, Florianópolis, SC, Brazil.

²Universidade do Sul de Santa Catarina, Palhoça, SC, Brazil. ³Department of Cell and Molecular Biology, Karolinska Institutet, Stockholm, Sweden.

ABSTRACT: Kinetoplastids are an ancestral group of protists that contains free-living species and parasites with distinct mechanisms in response to stress. Here, we compared genes involved in antioxidant defense (AD), proposing an evolution model among trypanosomatids. All genes were identified in *Bodo saltans*, suggesting that AD mechanisms have evolved prior to adaptation for parasitic lifestyles. While most of the monoxenous and dixenous parasites revealed minor differences from *B. saltans*, the endosymbiont-bearing species have an increased number of genes. The absence of these genes was mainly observed in the extracellular parasites of the genera *Phytomonas* and *Trypanosoma*. In trypanosomes, a distinction was observed between stercorarian and salivarian parasites, except for *Trypanosoma rangeli*. Our analyses indicate that the variability of AD among trypanosomatids at the genomic level is not solely due to the geographical isolation, being mainly related to specific adaptations of their distinct biological cycles within insect vectors and to a parasitism of a wide range of hosts.

KEYWORDS: Kinetoplastea, antioxidant system, evolution, trypanosomatids

CITATION: Beltrame-Botelho et al. A Comparative *In Silico* Study of the Antioxidant Defense Gene Repertoire of Distinct Lifestyle Trypanosomatid Species. *Evolutionary Bioinformatics* 2016;12:263–275 doi: 10.4137/EBO.S40648.

TYPE: Original Research

RECEIVED: July 29, 2016. **RESUBMITTED:** October 05, 2016. **ACCEPTED FOR PUBLICATION:** October 11, 2016.

ACADEMIC EDITOR: Liuyang Wang, Associate Editor

PEER REVIEW: Five peer reviewers contributed to the peer review report. Reviewers' reports totaled 1867 words, excluding any confidential comments to the academic editor.

FUNDING: This study was supported by grants from CAPES and CNPq, Brazilian Government Agencies, and the STINT (Sweden)/CAPES (Brazil) joint program. ITB-B is a recipient of a CAPES scholarship. The authors confirm that the funder had no influence over the study design, content of the article, or selection of this journal.

COMPETING INTERESTS: Authors disclose no potential conflicts of interest.

CORRESPONDENCE: patricia.stoco@ufsc.br

COPYRIGHT: © the authors, publisher and licensee Libertas Academica Limited. This is an open-access article distributed under the terms of the Creative Commons CC-BY-NC 3.0 License.

Paper subject to independent expert blind peer review. All editorial decisions made by independent academic editor. Upon submission manuscript was subject to anti-plagiarism scanning. Prior to publication all authors have given signed confirmation of agreement to article publication and compliance with all applicable ethical and legal requirements, including the accuracy of author and contributor information, disclosure of competing interests and funding sources, compliance with ethical requirements relating to human and animal study participants, and compliance with any copyright requirements of third parties. This journal is a member of the Committee on Publication Ethics (COPE).

Published by Libertas Academica. Learn more about this journal.

Introduction

The class Kinetoplastea composes an ancestral group of protists that includes free-living species as well as parasitic species of invertebrates, vertebrates, and plants. The subclass Metakinetoplastina comprises the following four orders: Eubodonida, Neobodonida, Parabodonida, and Trypanosomatida. The first three orders are named as bodonids and contain both free-living and parasitic species. The order Trypanosomatida exclusively contains species with a parasitic lifestyle, including (1) dixenous genera of vertebrate parasites such as *Endotrypanum* and important human pathogens such as *Leishmania* and *Trypanosoma*; (2) dixenous parasites of plants such as *Phytomonas*; and (3) monoxenous parasites of insects represented by various genera such as *Strigomonas*, *Critidia*, *Herpetomonas*, *Leptomonas*, and *Angomonas*.^{1–5} Phylogenetic analyses reinforce the late emergence of trypanosomatids as a sister group of the Eubodonida,⁶ which includes *Bodo saltans*, a free-living protist found worldwide in marine and freshwater habitats.

The ability of human-infective trypanosomatids (*Trypanosoma cruzi*, *Trypanosoma brucei*, and *Leishmania* spp.) to infect and multiply within mammalian hosts as well as to avoid the

host's immune system is directly associated with the success of the parasite in establishing the infection.^{7–9} Besides continuous exposure to oxidative environments during their life cycles within mammalian hosts and triatomine vectors, these parasites also produce reactive oxygen species (ROS) and reactive nitrogen species (RNS) via their aerobic metabolism.^{10,11} It is, thus, essential for these parasites to maintain a repertoire of genes coding for proteins involved in antioxidant defense (AD) mechanisms that play a central role for parasite survival. At the same time, these proteins are promising targets for the development of therapeutic drugs, since the majority of the enzymes involved are unique or distinct from the mammalian host orthologs when they exist.^{12–14}

Although the mechanisms responsible for trypanosomatid resistance to reducing environments are well studied, the functional and structural characterization of the components of redox metabolism in the nonpathogenic and mammalian pathogenic species is still incomplete.¹⁵

For most of the eukaryotic organisms, catalase and selenium-dependent glutathione peroxidase are major components of the defense mechanisms against cellular damage induced



by the imbalance between reactive species and detoxification enzymes.¹⁶ However, trypanosomatids lack such enzymes and their antioxidant machinery is based on the reduced form of the thiol trypanothione.^{7,17–19} In response to ROS and/or RNS, the maintenance of a reducing environment by trypanothione is supported by action of other important enzymes such as iron superoxide dismutase (SOD),^{20,21} non-selenium dependent glutathione peroxidase (GPx),^{22,23} trypanredoxin peroxidase (TrPx),^{24–27} ascorbate peroxidase (APx),^{25,28} and trypanothione reductase (TR),^{29,30} which is responsible for the reduced and active form of trypanothione.

Considering the evolutionary and biological differences among the trypanosomatids and the availability of genomic sequence data from diverse parasitic and nonparasitic species, we have performed a comparative *in silico* characterization of the trypanosomatid enzymes involved in the response to oxidative (ROS) and nitrosative (RNS) stress, and we propose a model for the evolution of this mechanism based on lifestyle and transmission way.

Methods

Genes related to the AD machinery. Due to the central role of trypanothione in the AD machinery of trypanosomatids, all genes coding for the enzymes involved in the synthesis of this thiol and in the synthesis of its precursors as well as in enzymatic reactions related to trypanothione activity were targeted in this study. A total of 15 genes were selected, which belong to the following pathways: (1) trypanothione precursor synthesis: cysteine biosynthesis-related enzymes, such as serine acetyltransferase (SAT, EC 2.3.1.30), cysteine synthase (CS, EC 2.5.1.47), cystathionine β synthase (C β S, EC 4.2.1.22), and cystathionine gamma lyase (CGL, EC 4.4.1.1); glutathione (GSH) biosynthesis, such as γ -glutamylcysteine synthetase (γ GcS, EC 6.3.2.2) and GSH synthetase (GS, EC 6.3.2.3); and polyamine biosynthesis or transport, such as ornithine decarboxylase (ODC, EC 4.1.1.17), polyamine transporter 1 (POT1), and spermidine synthase (SpdS, EC 2.5.1.16); (2) Trypanothione biosynthesis: trypanothione synthetase (TryS, EC 1.8.1.12) and glutamylspermidine synthase (GspS, EC 3.5.1.78); and (3) Trypanothione-related cascade of detoxification of ROS and/or RNS: TR (EC 1.8.1.12), SOD (EC 1.15.1.1), APx (EC 1.11.1.11), GPx (EC 1.11.1.9), and TrPx (EC 1.11.1.15).

Comparative analysis procedure was performed using *Trypanosoma rangeli* genomic and transcriptomic sequences available at GenBank, TriTrypDB,³¹ and local datasets. Sequences for SAT, CS, ODC, GspS, APx, and GPx-C were not identified in *T. rangeli* data or were present as pseudogenes. In these cases, sequences from *Leishmania major* and *T. cruzi* were used for the comparisons. Manual curation was performed confirming the ROS/RNS-related proteins by BlastP analysis using the UniProtKB/Swiss-Prot database (cutoff *e*-value 1e–10).

Then, homologous genes from *T. cruzi* CL-Brener (Esmeraldo and non-Esmeraldo haplotypes),³² *Trypanosoma cruzi marinkellei* B7,³³ *T. brucei* TREU927,³⁴ *Trypanosoma grayi* ANR4,³⁵ *Trypanosoma evansi* STIB805,³⁶ *Trypanosoma vivax* Y486,³⁷ *Leishmania infantum* clone JPCM5 (MCAN/ES/98/LLM-877),³⁸ *L. major* Friedlin,³⁹ *Leishmania tarentolae* Parrot-TarII,⁴⁰ *Leishmania braziliensis* MHOM/BR/75/M2903, *Crithidia fasciculata* Cf-C1, *Angomonas deanei* ATCC 30255,⁴¹ *Strigomonas culicis* ATCC 30268,⁴¹ *Endotrypanum monterogeii* LV88, *Leptomonas seymouri* ATCC 30220,⁴² *Phytomonas* spp. EM1,⁴³ and *B. saltans*⁴⁴ were searched online using BlastP (cutoff *e*-value 1e–5 and >35% identity) at TriTrypDB (<http://tritrypdb.org/tritrypdb>) or Genbank (<http://www.ncbi.nlm.nih.gov/genbank/>) (Supplementary File 1). The databases include the available genomic data of the selected species and can be modified after updates. The orthologs and paralogs were mapped online with OrthoMCL-DB v5⁴⁵ (<http://orthomcl.org/orthomcl/home.do>), using the default parameters (BlastP, cutoff of 1e–5 and 50% match) and using the ortholog and paralog tables generated to all datasets included in TriTrypDB. Gene copy numbers were determined after ortholog determination using OrthoMCL-DB. Functional characterization and conserved domains were analyzed in all sequences using the Conserved Domains Search Service (NCBI), InterPro (EMBL-EBI), and Prosite (ExPASy) software.

***In silico* characterization of ROS/RNS-related proteins.** Gene features were predicted using the ExPASy tools (<http://expasy.org>), while molecular mass and theoretical pI of the predicted proteins based on the amino acid sequences were predicted using the ProtParam software.

The putative signal peptides and cleavage sites were predicted using SignalP software,⁴⁶ Phobius software,⁴⁷ PrediSi software,⁴⁸ and iPSORT software,⁴⁹ based on eukaryote models and using default parameters. Only the signal peptide predicted by two programs was considered as valid. Cellular localization was predicted by iPSORT,⁴⁹ TargetP,⁵⁰ MitoProt,⁵¹ and ESLPred2⁵² using default parameters and non-plant or eukaryote prediction models. Subcellular localization equally predicted by at least two methods was considered as valid.

Phylogenetic analysis of AD machinery. The predicted amino acid sequence alignments of target genes were carried out online using Clustal Omega (<http://www.ebi.ac.uk/Tools/msa/clustalo/>) (Supplementary File 2).⁵³ Evolutionary analyses and phylogenetic trees were constructed using MEGA 6.0.⁵⁴ The alignment of each gene dataset was submitted to the search of best protein model using MEGA 6.0, and the model with the lowest Bayesian Information Criterion score was considered to describe the substitution pattern. The evolutionary history was inferred by using the Maximum Likelihood method based on the Le_Gascuel_2008 model.⁵⁵ For all analyses, the tree with the highest log-likelihood was showed. The bootstrap test (1000 replicates) was applied, and percentage of trees in which the associated taxa clustered

together (above 80%) is shown next to the branches. Initial trees for the heuristic search were obtained automatically by applying Neighbor Joining and BioNJ algorithms to a matrix of pairwise distances estimated using the selected models, and then selecting the topology with superior log-likelihood value. A discrete gamma distribution was used to model evolutionary rate differences among sites (five categories), and all sites were used in analyses. Trees are drawn to scale, with branch lengths measured in the number of substitutions per site.

Results and Discussion

Trypanothione precursors. *Cysteine and GSH biosynthesis.* Besides the involvement of cysteine in several biological processes due to the ability to undergo redox reactions, L-cysteine forms the basic building block of all thiol antioxidants, acting as a direct antioxidant and also as a precursor for the biosynthesis of GSH, trypanothione, or ovothiol. Cysteine is a key molecule in the AD of important protozoan parasite species such as *Entamoeba*, *Giardia*, and *Trichomonas*.¹⁷ In kinetoplastid parasites, cysteine is a precursor of GSH, which ultimately generates the major antioxidant molecule, trypanothione.

Two distinct cysteine biosynthesis pathways are shared among eukaryotes; both composed of the following two-step reactions: the *de novo* or assimilatory pathway and the reverse transsulfuration pathway. Reverse transsulfuration has been demonstrated in fungi and mammals where cysteine is obtained from homocysteine through the action of C β S and CGL. In this study, ortholog sequences to C β S and CGL were found in all analyzed species, except CGL in *T. vivax* (Supplementary File 3). Analysis of the number of genes coding for C β S and CGL that were observed among the studied species is shown in Figure 1. While five C β S genes were found in the *A. deanei* genome and in the non-Esmeraldo haplotype of *T. cruzi*, an extra copy was identified in the *T. cruzi* Esmeraldo haplotype. Four genes were found in *T. cruzi marinkellei*, three

genes were found in *T. rangeli*, and two genes were found in *T. grayi* and *S. culicis*. Among C β S sequences, a number of partial sequences missing either the N-terminal part or the C-terminal part were not included in the phylogenetic analyses (Supplementary File 3). Multiple copies for CGL were only observed in *A. deanei* and *S. culicis*. Phylogenetic analysis using C β S sequences did not allow a separation with high bootstrap value of trypanosomatid-related groups, such as the *Trypanosoma* genus (Fig. 2).

Cysteine can be synthesized *de novo* from serine in plants, bacteria, and some protozoa like *Entamoeba histolytica* and *Entamoeba dispar*, including trypanosomatids.¹⁷ This process depends on the action of SAT and CS. Among the trypanosomatids, *T. cruzi* and *T. grayi* maintained an intact *de novo* pathway, forming a distinct clade with high bootstrap value in the phylogenetic analysis based on CS (Fig. 2). Another clade was represented by all *Leishmania* spp., *E. monterogei*, and the monoxenous parasites *C. fasciculata* and *L. seymouri*. However, sequences coding for both enzymes (SAT and CS) are missing in *Phytomonas* spp., *T. brucei*, *T. evansi*, and *T. vivax* and present in *T. rangeli* as pseudogenes (Fig. 1). Except for the monoxenous *A. deanei* and *S. culicis* that contains many copies of CS and SAT in their genomes, a single CS copy was detected in most of the studied kinetoplastids.

T. brucei is considered auxotrophic for cysteine, taking up exogenous cysteine via specific transporters.⁵⁶ Such transporters represent one of the largest families of permeases in the kinetoplastid genomes.⁵⁷ While cysteine transport is found in *L. major* and *T. cruzi*, transporters of cysteine in *T. brucei* are ~200-fold more efficient.^{58,59} For parasites that are deficient in *de novo* cysteine synthesis, like *T. vivax*, *T. evansi*, and *T. rangeli*, the efficiency of their cysteine transporters is probably related to the availability of this specific amino acid. Considering that *T. brucei*, *T. vivax*, and *T. evansi* are extracellular parasites of mammals and the lifestyle of *T. rangeli* within these hosts is unknown, the deficiency in *de novo* cysteine

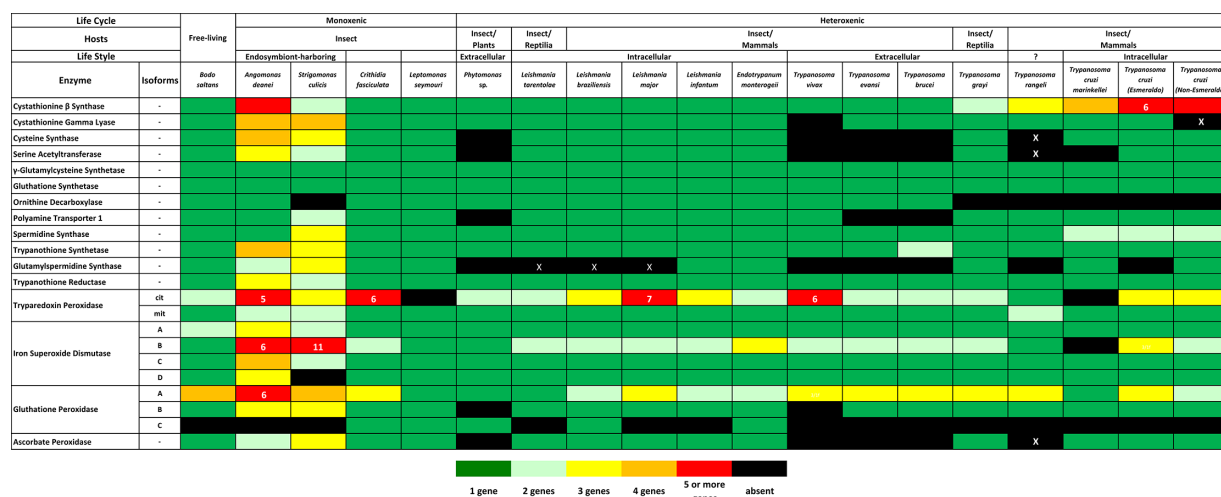


Figure 1. Heat map generated by the analysis of the gene copy numbers of antioxidant enzymes in kinetoplastids. X represents pseudogenes.

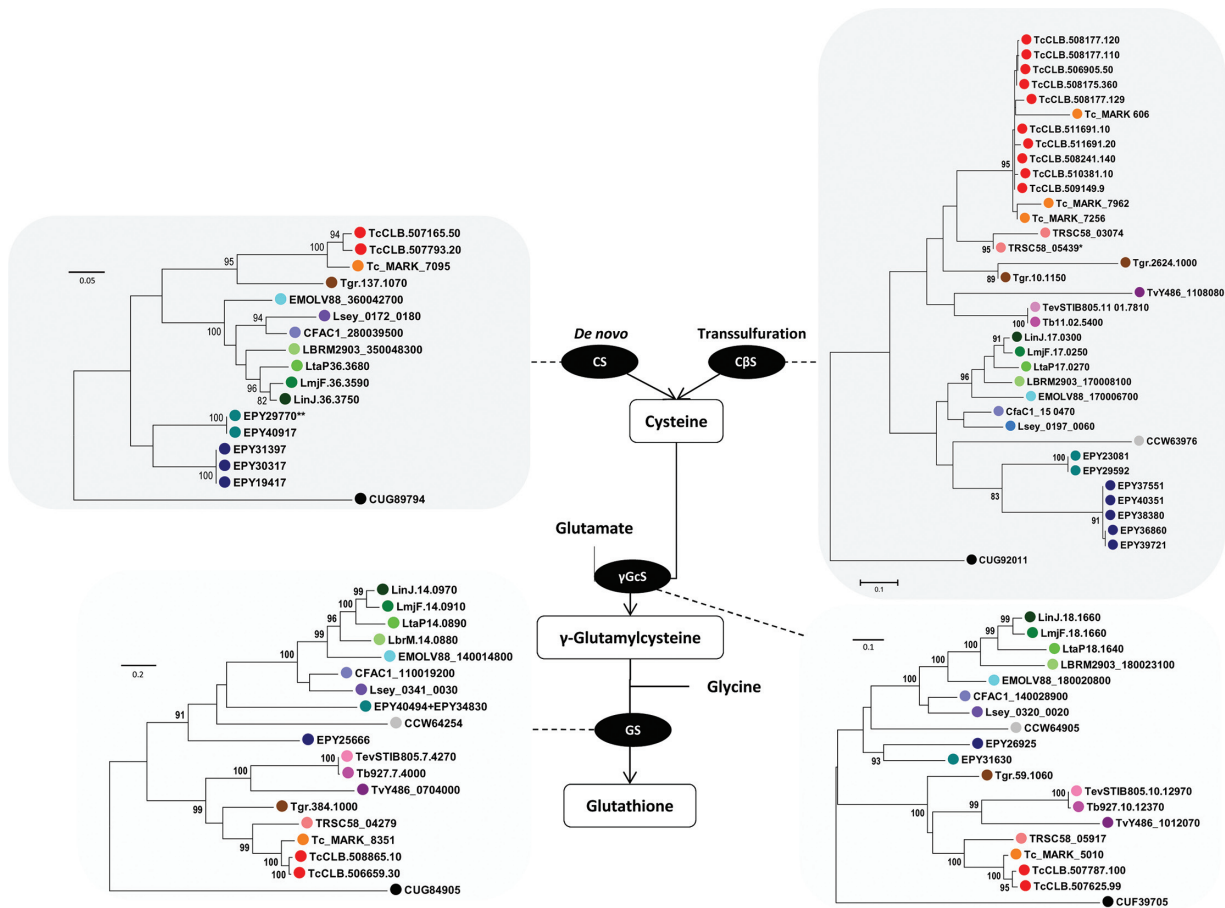


Figure 2. Molecular phylogenetic analyses of amino acid sequences from kinetoplastid enzymes related to cysteine and glutathione biosynthesis pathways. Protein alignments were performed using Clustal Omega, and the evolutionary histories were inferred by using the Maximum Likelihood method of the MEGA software (Version 6.0). Phylogenies of CS, CβS, GS, and γGcS were constructed using the LG + G amino acid substitution model. Trees are drawn to scale, with branch lengths measured in the number of substitutions per site. The analysis involved 36 amino acid sequences to CβS (435 sites), 17 to CS (340 sites), 19 to GS (727 sites) and γGcS (733 sites). Bootstrap values >80 are shown next to the branches, corresponding to the percentage of trees in which the associated taxa clustered together (1000 replicates). Annotated sequences from kinetoplastids were retrieved from TriTrypDB or Genbank and the access numbers are shown. *Identical sequence identified; ● *T. cruzi* (Esmeraldo and non-Esmeraldo); ● *T. cruzi marinkellei*; ● *T. grayi*; ● *T. rangeli*; ● *T. brucei*; ● *T. evansi*; ● *T. vivax*; ● *L. infantum*; ● *L. major*; ● *L. braziliensis*; ● *L. tarantolae*; ● *C. fasciculata*; ● *L. seymouri*; ● *E. monterogei*; ● *A. deanei*; ● *S. culicis*; ● *Phytomonas* spp.; ● *B. saltans*.

synthesis and/or cysteine transport could explain the reduced levels of total thiols observed in *T. rangeli* (eight times lower) when compared with *T. cruzi*.⁶⁰

GSH synthesis is also a two-step process catalyzed by γGcS and GS. The search sequences coding for γGcS and GS on the available genome datasets revealed that these enzymes are conserved (identities of 50%–97%) and present as single copy genes in all studied species, representing the most conserved pathway of the antioxidant metabolism of kinetoplastids. Phylogenetic analyses using γGcS and GS revealed a similar resolution, but rooting the trees on the *B. saltans* sequences formed the following two distinct groups: one including all trypanosomes and the second including all *Leishmania* species and other trypanosomatids (Fig. 2). Except for the endosymbiont-harboring species (*S. culicis* and *A. deanei*) and *Phytomonas* spp., all species were clearly separated at the species level.

Polyamine synthesis and uptake. The biosynthesis of trypanothione depends on GSH synthesis as well as spermidine levels, a polyamine synthesized from putrescine via SpdS.^{61,62} SpdS is a single copy gene, well-conserved enzyme for most of the analyzed kinetoplastids (Fig. 1 and Supplementary File 4). Interestingly, two copies of the SpdS gene were found in each of the *T. cruzi* haplotypes (Esmeraldo and non-Esmeraldo) and *T. cruzi marinkellei* and three copies of the SpdS gene were found in *S. culicis*. Phylogenetic analysis of the SpdS clustered all American trypanosomes, including *T. cruzi* and *T. rangeli*, in a single branch, indicating a duplication of this gene in the *T. cruzi* genome. Although *T. brucei*, *T. evansi*, and *T. vivax* have been clustered together on a separate branch, the African *T. grayi* was not clearly positioned (Fig. 3).

Some kinetoplastids, such as *Leishmania* sp., can either synthesize polyamines de novo or use uptake from the

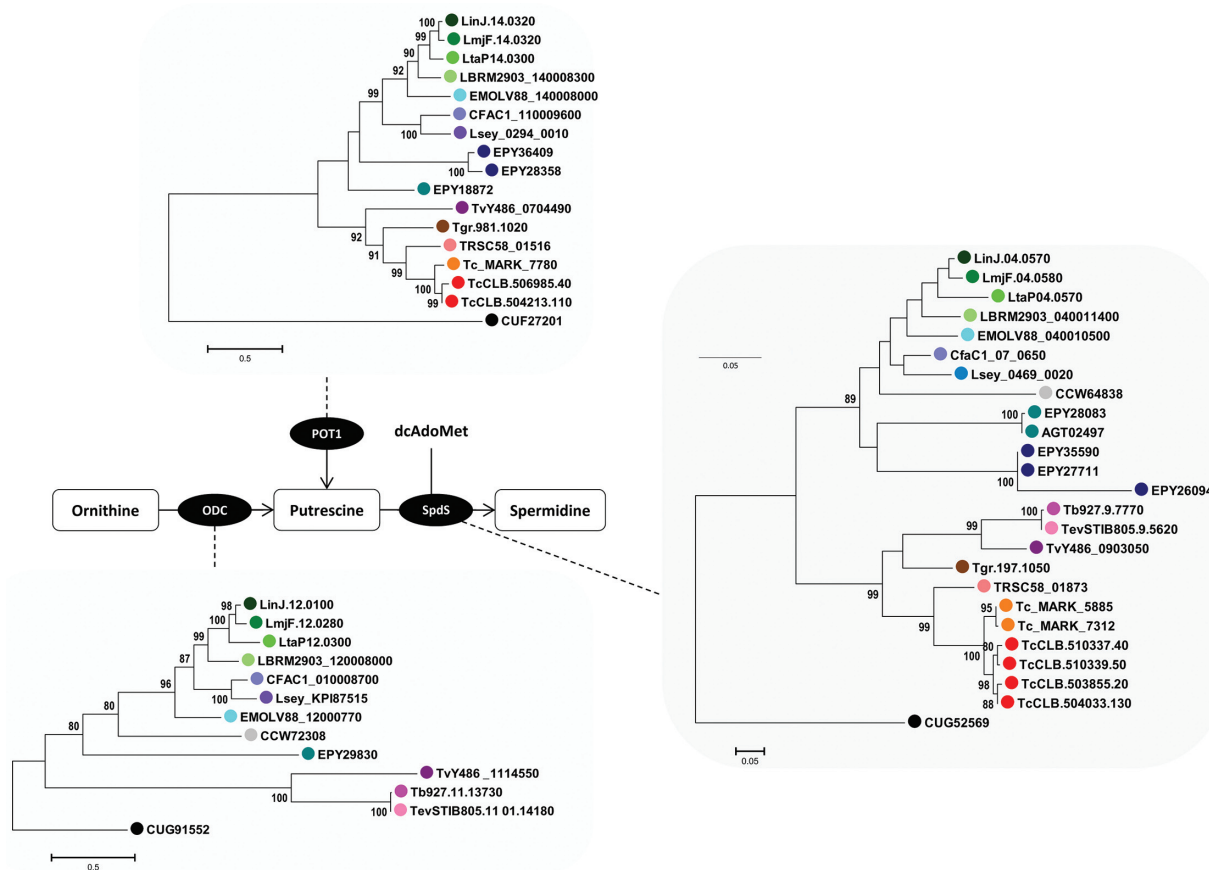


Figure 3. Molecular phylogenetic analyses of amino acid sequences from kinetoplastid enzymes related to spermidine synthesis. Protein alignments were made with Clustal Omega, and the evolutionary histories were inferred by using the Maximum Likelihood method using MEGA version 6.0. Phylogenies of ODC and POT1 were constructed using the JTT + G amino acid substitution model, and WAG + G was used to SpdS. Trees are drawn to scale, with branch lengths measured in the number of substitutions per site. The analysis involved 13 amino acid sequences to ODC (864 sites), 17 to POT1 (857 sites), and 24 to SpdS (319 sites). Bootstrap percentages above 80% are shown next to the branches, corresponding to the percentage of trees in which the associated taxa clustered together (1000 replicates). Kinetoplastid-annotated sequences were retrieved from TriTrypDB or Genbank and the access numbers are included. ● *T. cruzi* (Esmeraldo and non-Esmeraldo); ● *T. cruzi marinkellei*; ● *T. grayi*; ● *T. rangeli*; ● *T. brucei*; ● *T. evansi*; ● *T. vivax*; ● *L. infantum*; ● *L. major*; ● *L. braziliensis*; ● *L. tarentolae*; ● *C. fasciculata*; ● *L. seymouri*; ● *E. monterogeii*; ● *A. deanei*; ● *S. culicis*; ● *Phytomonas* spp.; ● *B. saltans*.

environment,⁶³ but others, such as *T. brucei* and *T. cruzi*, completely rely on polyamine biosynthesis⁶⁴ and on the uptake of putrescine from the host,⁶⁵ respectively. Putrescine synthesis de novo is related to the presence of the enzyme ODC, which is absent in *S. culicis*, *T. cruzi*, *T. rangeli*, and *T. grayi* (Fig. 1). Phylogenetic analyses of ODC (Fig. 3) clustered all *Trypanosoma* spp. and *Leishmania* spp. in separate branches, so that *C. fasciculata* and *L. seymouri* formed a clearly distinct clade with high bootstrap. *Phytomonas* spp., *A. deanei*, and *E. monterogeii* were not clustered with other trypanosomatids, although the latter was clearly closer to the *Leishmania* spp. and *Phytomonas* spp./*A. deanei* branches.

The low identity of the *Odc* gene between trypanosomes and the other kinetoplastids (~35%) reinforces the hypothesis of the loss of this gene in the *Trypanosoma* lineage and a downstream acquisition of an ortholog gene from a vertebrate host as suggested by Steglich and Schaeffer.⁶⁶ However, such an event may not be attributed to all African trypanosomes, since *T. grayi* does not have an *Odc* gene. The absence of ODC in

this crocodylian trypanosome could be due to the horizontal gene transfer occurring from a mammalian source.

The absence of ODC in *S. culicis* was explored by Alves et al.⁶⁷ In this monoxenous parasitic organism, endosymbiont-coded enzymes promote the interconversion of agmatine into putrescine. Agmatine is synthesized by arginine decarboxylase, which is exclusively present in the genomes of *S. culicis* and *A. deanei* endosymbionts but absent in all other studied trypanosomatids.⁶⁷

POTs are present on the cell membranes of *L. major* and *T. cruzi*,^{63,65} and orthologs of these sequences were identified in all studied trypanosomatid species, except for *T. brucei*, *T. evansi*, and *Phytomonas* spp. (Fig. 1). Interestingly, all ODC-negative trypanosomes harbor a POT1 gene that is highly similar to the *T. cruzi* POT1 (TcPOT1). Also, *T. vivax* is the single trypanosome with both ODC and POT1 genes, and the ODC-negative *S. culicis* has two copies of the POT1 gene. Two major clades composed of *Trypanosoma* sp. and *Leishmania* sp./*C. fasciculata*/*L. seymouri* were generated in the phylogenetic analysis, in accordance with the ODC-based analysis discussed



earlier (Fig. 3). It is noteworthy to mention that other sequences annotated as POTs in the *T. brucei* genome showed only 33% identity with TcPOT1 and, thus, were not included in this analysis as their function is likely to be different.

Trypanothione synthesis and reduction. The conjugation of GSH with other molecules is common in eukaryotes,⁶⁸ but the combination of GSH and polyamine occurs only in Enterobacteria and Euglenozoa. The conjugation of GSH-spermidine is Adenosine triphosphate (ATP) dependent and can be solely catalyzed by TryS or by the subsequent action of GspS and TryS, both exhibiting synthetase and amidase activities.^{69,70} The following two distinct domains are present in these proteins and are related to opposite activities: the amidase N-terminal domain with a cysteine, histidine-dependent amidohydrolase/peptidase (CHAP) (pfam05257) and the synthetase C-terminal domain similar to circularly permuted ATP-grasp type 2 (pfam03738).

With the exception of *A. deanei* and *S. culicis*, TryS was found in all trypanosomatids as a single copy gene (Fig. 1). On the other hand, the GspS genes are missing in a number of species, including *T. rangeli*, *T. evansi*, *T. vivax*, *T. brucei*, and

Phytomonas spp. GspS pseudogenes were found in three out of four analyzed *Leishmania* species and as a full-length gene in *L. infantum*. Although the characteristic domains have been found in both genes, reflecting the high similarity among the sequences, the trypanosomes were firmly clustered in a separate branch, clearly distinct from other trypanosomatids (Fig. 4A).

The presence of TryS in all trypanosomatids and the lack of GspS in some species were previously pointed out by Manta et al.⁵⁶, who proposed an initial gene transfer event from the endosymbiont, followed by gene duplication, generating GspS and TryS. During the evolutionary history of these organisms, some species have lost the GspS gene and maintained the TryS gene.

The search of the genome databases revealed two other groups of sequences that are annotated as *TryS*, *putative* or *TryS-like, putative*. Due to the exclusive presence of the amidase (CHAP) domain in the predicted sequences, we have named them CHAP1 and CHAP2 in this study (Fig. 4B). The CHAP domain is present at the N-terminal region in all the analyzed sequences, including TryS, GspS, CHAP1, and CHAP2. The differences between CHAP1 and CHAP2 are the total protein size (CHAP1 has ~250 aa and CHAP2 ~650 aa) and

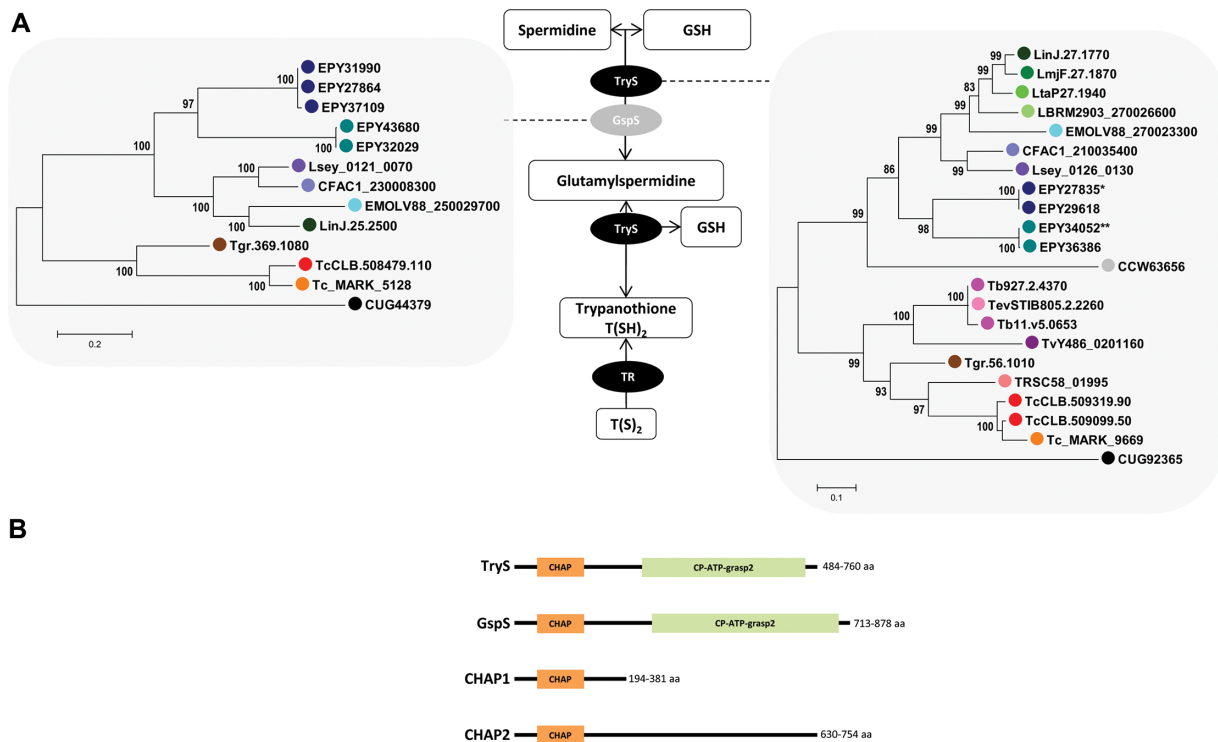


Figure 4. Analyses of enzymes evolved in trypanothione biosynthesis. (A) Molecular phylogenetic analyses of two enzymes directly evolved in trypanothione biosynthesis: trypanothione synthetase (TryS) and glutamylspermidine synthase (GspS). Protein alignments were made with Clustal Omega, and the evolutionary histories were inferred by using the Maximum Likelihood method using MEGA version 6.0. Phylogenies of TryS and GspS were constructed using the LG + G and JTT + G amino acid substitution models, respectively. Trees are drawn to scale, with branch lengths measured in the number of substitutions per site. The analysis involved 22 amino acid sequences to TryS (780 sites) and 13 to GspS (997 sites). Bootstrap percentages above 80% are shown next to the branches, corresponding to the percentage of trees in which the associated taxa clustered together (1000 replicates). Kinetoplastid-annotated sequences were retrieved from TriTrypDB or Genbank, and the access numbers are included. *Identical sequence identified; ●, *T. cruzi* (Esmeraldo and non-Esmeraldo); ●, *T. cruzi marinkellei*; ●, *T. grayi*; ●, *T. rangeli*; ●, *T. brucei*; ●, *T. evansi*; ●, *T. vivax*; ●, *L. infantum*; ●, *L. major*; ●, *L. braziliensis*; ●, *L. tarentolae*; ●, *C. fasciculata*; ●, *L. seymouri*; ●, *E. monterogeei*; ●, *A. deanei*; ●, *S. culicis*; ●, *Phytomonas* spp.; ●, *B. saltans*. (B) Conserved domain analyses of sequences annotated as trypanothione synthetase in kinetoplastid genome databases. Figures are on scale and represent the media size of predicted proteins among the species. Numbers at right side represent the smaller and larger sizes for each gene.

the exclusive prediction to mitochondrial localization in some CHAP1 sequences (Supplementary File 5). Another interesting difference is the presence of CHAP1 genes in all the analyzed species (except *Phytomonas* spp.), in contrast to the lack of CHAP2 in *T. brucei*, *T. evansi*, *T. brucei*, *L. braziliensis*, and *Phytomonas* spp. Like the unclear role of the amidase activity in GspS/TryS of trypanosomatids, the possible function of CHAP1/CHAP2 in trypanothione metabolism still remains to be investigated.

After synthesis, the trypanothione is maintained in a reduced state (T(SH)₂) by an NADPH-dependent flavoenzyme TR. The presence of this enzyme is common to all trypanosomatids (Supplementary File 6), being present as a single copy gene despite the central role in the redox metabolism.²⁹

TR exhibits the following two characteristic domains: Pyr Redox (pyridine nucleotide-disulfide oxidoreductases) related to oxidoreductase action and Pyr Redox dimerization that allows the typical dimeric structure. Although all studied species contain at least one sequence containing both domains, some species, such as *T. cruzi* Esmeraldo haplotype, *A. deanei*, and *S. culicis*, have a second TR gene copy

that contains only one of the domains, possibly coding for a nonfunctional enzyme.

Peroxidases. The following five distinct peroxidases have been described in *T. cruzi*: two peroxiredoxins cytosolic and mitochondrial TrPxs (TrPxcit and TrPxmit); two nonselenium-dependent GPxs (cytosolic, GPxI; endoplasmic reticulum (ER), GPxII), and one ascorbate-dependent hemoperoxidase.^{22,23,26,28} The peroxidases are also present in other kinetoplastids where they show differences in cellular localization and substrate affinity.¹⁹

Tryparedoxin peroxidases. TrPx are considered to be the most important enzymes in peroxide detoxification.⁷¹ Unlike the GSH-dependent peroxiredoxins of mammals, TrPx is trypanothione dependent. Like H₂O₂, TrPx utilizes the trypanothione (TPX) (a thioredoxin homolog) as an electron donor in the reduction of substrates, short-chain hydroperoxides,²⁶ and peroxinitrite.²⁵

Phylogenetic analyses using amino acid sequences from distinct kinetoplastid species clearly separated two major groups containing sequences with predicted mitochondrial and cytosolic localization (TrPxmit or TrPxcit) (Fig. 5A). All sequences within the mitochondrial clade were approximately

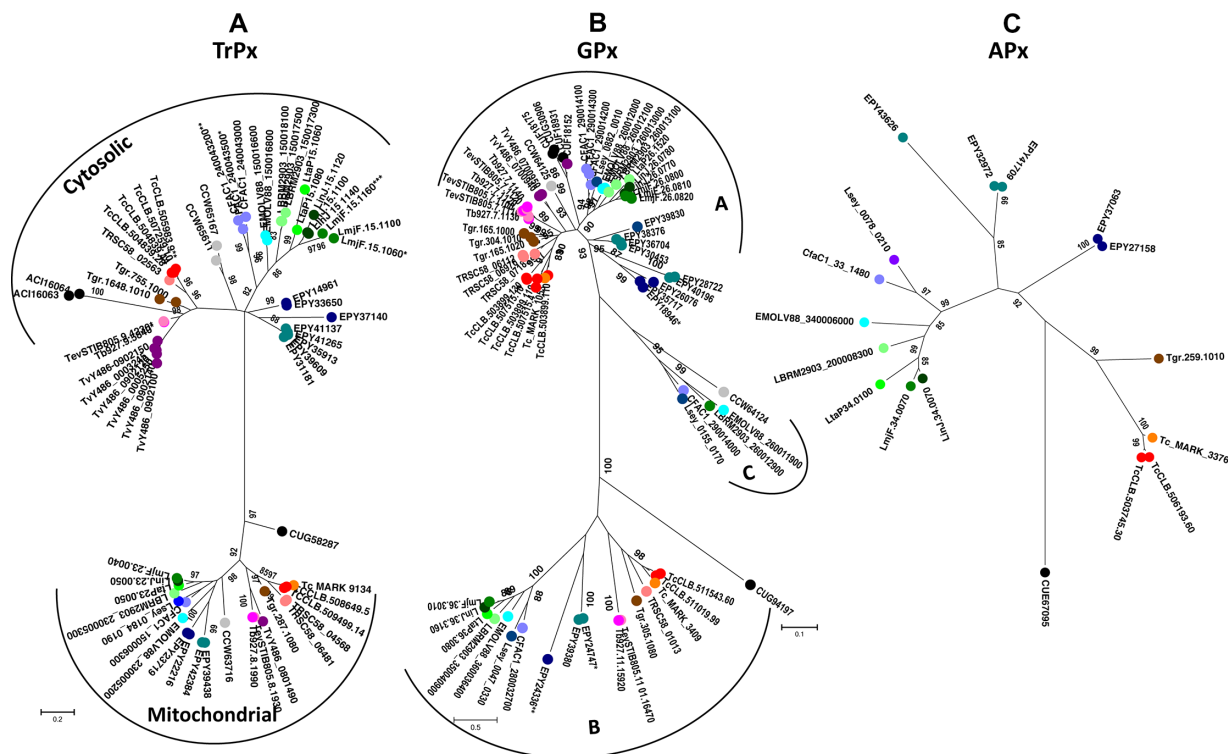


Figure 5. Molecular phylogenetic analyses of kinetoplastid peroxidases: (A) tryparedoxin peroxidase (TrPx); (B) glutathione peroxidase (GPx); and (C) ascorbate peroxidase (APx). Protein alignments were made with Clustal Omega, and the evolutionary histories were inferred by using the Maximum Likelihood method using MEGA version 6.0. Phylogenies of GPx and APx were constructed using the WAG + G amino acid substitution model, and LG + G was used to TrPx. Trees are drawn to scale, with branch lengths measured in the number of substitutions per site. The analysis involved 71 amino acid sequences to GPx (384 sites), 64 to TrPx (337 sites), and 17 to APx (370 sites). Bootstrap percentages above 80% are shown next to the branches, corresponding to the percentage of trees in which the associated taxa clustered together (1000 replicates). Kinetoplastid-annotated sequences were retrieved from TriTrypDB or Genbank, and the access numbers are included. *Identical sequence identified; ● *T. cruzi* (Esmeraldo and non-Esmeraldo); ● *T. cruzi marinkellei*; ● *T. grayi*; ● *T. rangeli*; ● *T. brucei*; ● *T. evansi*; ● *T. vivax*; ● *L. infantum*; ● *L. major*; ● *L. braziliensis*; ● *L. tarentolae*; ● *C. fasciculata*; ● *L. seymouri*; ● *E. monterogeei*; ● *A. deanei*; ● *S. culicis*; ● *Phytomonas* spp.; ● *B. saltans*.



30 amino acids longer in the N-terminal portion than the sequences in the cytosolic clade. These regions probably correspond to the mitochondrial-addressing signal, despite the fact that no *in silico* prediction tool could identify a recognizable signal peptide (Supplementary File 7).

The occurrence of cytosolic and mitochondrial TrPx reinforces the needs for cell protection against exogenous and endogenous toxic peroxides. *L. seymouri* and *T. cruzi marinkellei* do not have the cytosolic isoform. All other analyzed kinetoplastids had both isoforms, having multicopy genes encoding for the cytosolic enzyme. On the other hand, TrPxmit is a single copy gene for most of the studied organisms except for *T. rangeli*, *A. deanei*, and *S. culicis* (Fig. 1).

Sequences from the cytosolic clade showed high similarity levels (>62.12%), and no phylogenetic grouping could be seen. A separation in two major clades as reported for many other genes (see above) occurred with the mitochondrial sequences (Fig. 5A).

All TrPx sequences have peroxidatic and resolving cysteines (Cp and Cr) in the conserved domain PRX_Typ2cys (cd0301). The Cp was observed in all kinetoplastids, except for one sequence from *A. deanei* (EPY41265), which, from alignment with the other species, appears to be truncated. The Cr was found in all sequences, and in the cytosolic group, it was contained inside the VCP domain. In the mitochondrial sequences, the Cr was in the VIPC domain, which is characteristic of the mitochondrial TrPx. Both cysteine residues are present in all analyzed TrPx-predicted proteins in the peroxiredoxin 2-Cys family.^{19,72}

GSH peroxidase. The majority of eukaryotes use GSH-dependent peroxidases as key enzymes in peroxide metabolism. Before the characterization of this enzyme in *C. fasciculata*⁷³ they were considered to be absent in trypanosomatids. Upon description in *T. cruzi*, *T. brucei*, and *Leishmania* sp.,^{26,74,75} the peroxidases were renamed as nonselenium GPX-like enzymes,¹⁹ GPX like,⁷² GPX type TrPx,⁷⁶ tryparedoxin-dependent peroxidase (TDPx),⁷⁷ or just GPx.⁷⁸ Distinct isoforms of the GPx enzymes, including isoform A (or I), which can be subdivided into AI, AII, and AIII, and isoform B (or II), have been described. However, the use of these classifications is still not well established, since some authors apply the classification based on sequence similarity and others in accordance with the subcellular localization.^{22,23} Phylogenetic analyses of GPx showed a separation of the sequences into two distinct branches according to the isoforms A (GPxA) and B (GPxB) (Fig. 5B). Even though isoform A sequences have been found in all trypanosomatids, the subgroups AI, AII, and AIII were not identified in our analyses. With the exception of *Phytomonas* spp. and *T. vivax*, isoform B was found in all species (Fig. 1).

Variation in GPx was investigated further by closer examination of gene copy numbers and the genomic organization. One to three genes were identified per species for isoform A (Supplementary File 7), organized in tandem. *Leishmania* in general have this pattern, with the unique exception of *L. tarentolae* (LtaP26.1520), which has only one short

and nonsynthetic sequence. Interestingly, the lizard parasite *L. tarentolae*, which belongs to the Sauroleishmania subgenera, is nonpathogenic to humans and does not multiply intracellularly. The isoform A sequences did not separate into three specific subgroups, and the branch separations were not always at the species level. Ortholog sequences in more related species (*T. brucei* and *T. evansi*/*T. cruzi* and *T. cruzi marinkellei*) grouped closer than paralogous sequences.

GPx classification in *T. cruzi* is based on different cellular localizations, affinity to the reductor agent, and the metabolized target. TcGPxA is a cytosolic enzyme that has TPX as the reductor agent and metabolizes hydroperoxides. TcGPxA shows high similarity to the PHGPX group, which includes plant and mammalian GPx.^{19,73} TcGPxB is located in the ER, uses GSH, and is involved in the protection against lipid peroxidation.

According to Castro and Tomas,¹⁹ the GPxA subdivision (I, II, and III) must be based on the reducing agent and the substrate affinity. Some isoforms can bind to TPX, while others bind to GSH, and some use H₂O₂ and some lipid hydroperoxides as reductors. No indications of these specificities were found at the amino acid sequence level.

Alphey et al.⁷⁷ described the *L. major* isoform AI (TDPX1) as a cytosolic protein and the isoforms AII and AIII (TDPX2 and TDPX3) as mitochondrial proteins due to the presence of N-terminal signal peptides. The TDPX3 also has a glycosomal signal in the C-terminal region.⁷⁵

Most of the GPxB sequences were predicted to be cytosolic. The GPxA sequences include cytosolic and mitochondrial predictions. Based on this criterion, we can define the following two distinct groups of trypanosomatid GPxA: GPxAI with predicted cytosolic sequences and GPxAII with prediction or signal peptides for mitochondria or ER (Supplementary File 7).

We also identified a set of divergent sequences (here named GPxC), which include one sequence each from *E. monterogei*, *L. braziliensis*, *L. seymouri*, *C. fasciculata*, and *Phytomonas* spp. All GPx sequences showed the complete GSH peroxidase domain (pfam00255) with all conserved catalytic residues (C, Q, and W). The divergent sequences are different in size, and there are sequence differences in the N- and C-termini and in the conserved central region (Supplementary File 2).

Ascorbate peroxidase. A third type of peroxidase is related to ascorbate-dependent hemoperoxidases found in plants. *T. cruzi* APx decomposes only H₂O₂ and uses ascorbate as a source of reducing equivalents. Dehydroascorbate is reduced by trypanothione.⁷⁹ The discovery and characterization of this enzyme in *T. cruzi*²⁸ and *Leishmania* sp.⁸⁰ (intracellular parasites) and the absence of the protein in the extracellular parasite *T. brucei* suggested that this kind of antioxidant detoxification could be related to an intracellular lifestyle.⁸¹

We did, however, identify APx coding sequences in other trypanosomatids that are not intracellular parasites, including the free-living kinetoplastid *B. saltans*, monoxenous trypanosomatids, and in the extracellular parasite of crocodiles, *T. grayi* (Fig. 5C). APx orthologs were only absent in

Phytomonas spp., *T. brucei*, and the related species *T. evansi* and *T. vivax*, and only a pseudogene was found in *T. rangeli* (Fig. 1). *B. saltans* APx sequence has 46% identity to *Picea glauca* (seed plant) APx and ~50% identity to the green algae and Oomycetes peroxidases. All these findings suggest that the presence of APx is probably not an acquisition to adapt parasites to an intracellular environment, but instead, a gene acquired early in the kinetoplastid evolution, reinforcing the hypothesis of a common origin of the hemoperoxidases of trypanosomatids and plants.^{28,81,82} The presence of *plant-like* genes had been reported in kinetoplastids, suggesting that plastids were initially acquired by endosymbiosis and, after losing these organelles, have retained some of its genes.⁸³

Superoxide dismutases. The SODs are a group of metalloproteins involved in eliminating the superoxide radical $O_2^{\cdot-}$.

The presence of this enzyme in all the analyzed species is not surprising, since it is present in all aerobic organisms.²¹ SOD exhibits a high degree of sequence and structure conservation,⁸⁴ suggesting a common ancestor.⁸⁵

All kinetoplastid SODs identified thus far have Fe affinity and are therefore named iron SODs (FeSODs).⁸⁴ This group of enzymes is restricted to protozoans, prokaryotes, and chloroplasts.⁸⁶ Distinct isoforms are described in kinetoplastids, and again the nomenclature is highly variable. Most of the time, four isoforms named as A, B (1 and 2), and C are reported. The fifth isoform was found in *T. brucei* (Tb927.6.4030),⁸⁴ and orthologs of this isoform were found in all kinetoplastid species analyzed here, except *S. culicis* (Fig. 1), which was described here as FeSOD-D. All sequences were included in the iron/manganese SOD family according to domain analyses.

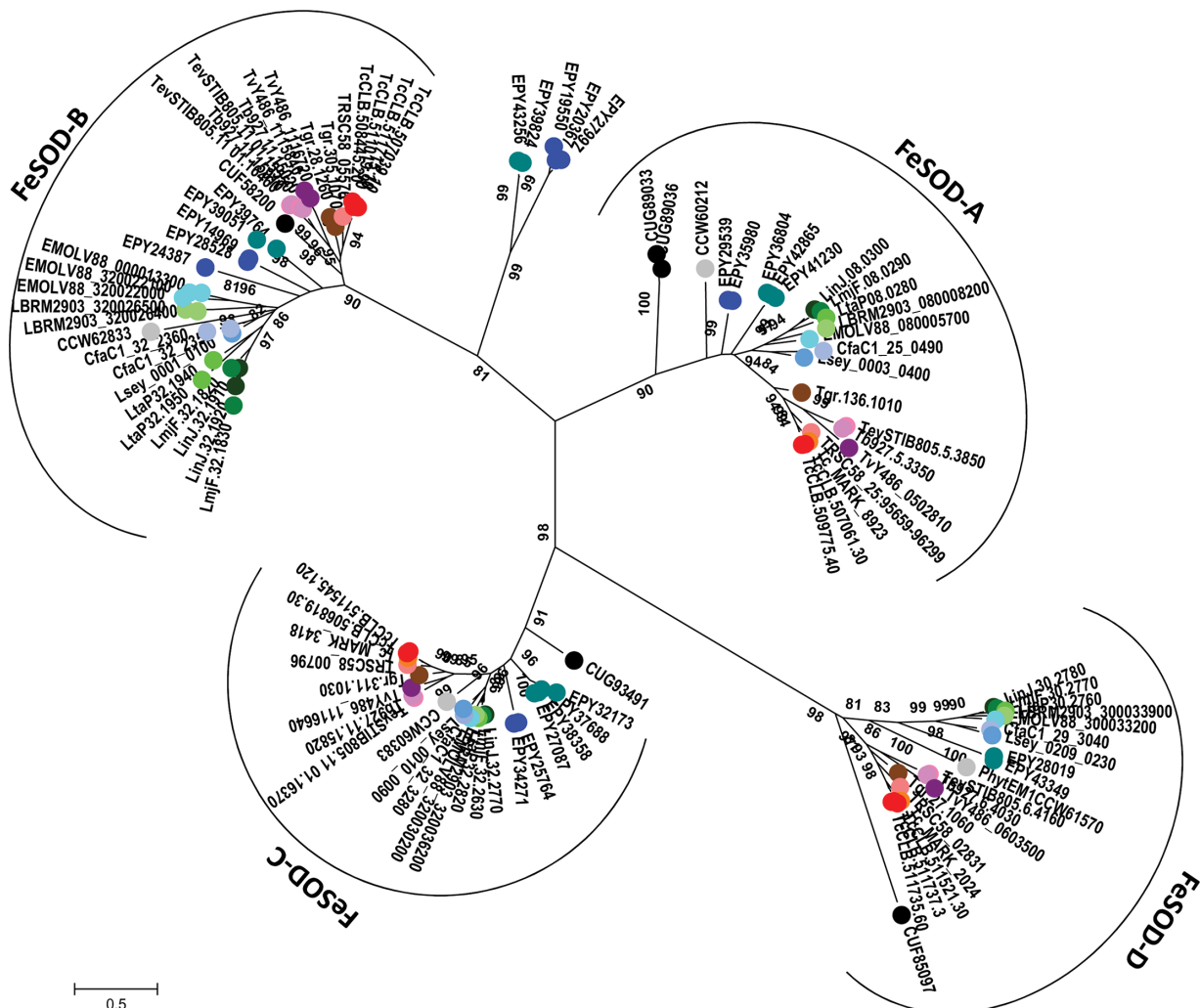


Figure 6. Molecular phylogenetic analysis of kinetoplastid FeSODs. Protein alignment was made with Clustal Omega, and the evolutionary history was inferred by using the Maximum Likelihood method using MEGA version 6.0. Phylogeny of FeSOD was constructed using the LG + G amino acid substitution model. Tree is drawn to scale, with branch lengths measured in the number of substitutions per site. The analysis involved 105 amino acid sequences (402 sites). Bootstrap percentages above 80% are shown next to the branches, corresponding to the percentage of trees in which the associated taxa clustered together (1000 replicates). Kinetoplastid-annotated sequences were retrieved from TriTrypDB or Genbank, and the access numbers are included. *Identical sequence identified; ● *T. cruzi* (Esmeraldo and non-Esmeraldo); ● *T. cruzi marinkellei*; ● *T. grayi*; ● *T. rangeli*; ● *T. brucei*; ● *T. evansi*; ● *T. vivax*; ● *L. infantum*; ● *L. major*; ● *L. braziliensis*; ● *L. tarentolae*; ● *C. fasciculata*; ● *L. seymouri*; ● *E. monterogeii*; ● *A. deanei*; ● *S. culicis*; ● *Phytomonas* spp.; ● *B. saltans*.



The following four clades are clearly observed in the phylogenetic analyses: A, B, C, and D (Fig. 6). In each clade, there was a clear separation of the *Trypanosoma* species. At least one gene copy from each isoform was identified in the analyzed species, except *S. culicis* (FeSOD-D) and *T. cruzi marinkellei* (FeSOD-B). For FeSOD-B, two copies were present in the majority of the species, probably representing the isoforms B1 and B2. A set of distinct sequences, exclusively from *S. culicis* and *A. deanei*, were grouped near the FeSOD-B sequences.

Most of the FeSOD-A (78%), FeSOD-C (61%), and FeSOD-D (75%) sequences were predicted to be mitochondrial (Supplementary File 8). In *T. brucei*, the FeSOD-B1 and B2 differentiations were based on the identification of distinct residues in the C-terminal region: LKS (TbSOD-B1) and SDL (TbSOD-B2).⁸⁴ The SDL peptide is similar to the SKL signal related to peroxisome location. Thus, TbSOD-B1 is mainly found in the cytoplasm, while TbSOD-B2 is mainly located in glycosomes. We did not identify the peroxisome signal peptide in all FeSOD-B sequences, since the C-terminal region is frequently degenerated.⁸⁷

The presence of this class of enzymes in different cell compartments reinforces the importance of the detoxification of superoxide anions produced in different biochemical processes.⁸⁷ Based on our results, we suggest that the trypanosomatid FeSOD enzymes should be classified as follows: FeSOD-A, FeSOD-C, and FeSOD-D for predicted mitochondrial proteins, FeSOD-B1 for exclusively cytosolic enzymes, and FeSOD-B2 for sequences that contain signal peptides for targeting peroxisome/glycosome (SQL or SDL).

The evolution of the AD. Aerobic organisms are exposed to oxidative challenges generated by their metabolism. All kinetoplastid protozoans have to eliminate their endogenous metabolites, and the parasite group also has to defend themselves against the oxidative stress generated by the host immune system.¹⁷ Thus, we expected that the antioxidant machinery components could differ between parasitic trypanosomatids, which are correlated with differences in the oxidative challenges. The identification of all the analyzed gene sequences in the free-living protozoan *B. saltans*, which is also considered to be an ancestor of parasitic trypanosomatids, suggests that the AD mechanisms are an anterior adaptation of this group of protozoans.

Most of the monoxenous parasites and all the analyzed *Leishmania* sp. had only a few differences in gene copy numbers compared to *B. saltans* (Fig. 1). However, the endosymbiont-bearing *A. deanei* and *S. culicis* showed a higher number of gene copies for the majority of the analyzed genes. This characteristic was also observed by Motta et al.⁴¹, and it is consistent with the high number of CDS predicted in the partial genome sequences of these parasites, 16,957 and 12,157 CDS, respectively.⁴¹ These CDS numbers are smaller than that of *B. saltans* (18,936)⁵ but higher than that of all other analyzed

parasites. In addition, basically in all analyses, genes from *A. deanei* and *S. culicis* did not group with other monoxenous parasites included in the analyses, such as *C. fasciculata* and *L. seymouri*. The latter parasites were more closely related to the heteroxenous parasites *E. monterogeii* and *Leishmania* sp.

Genomic reduction has been observed in a considerable number of parasites, and trypanosomatids appear to have a reduced size of their genomes and the gene density compared to the free-living *B. saltans*.^{3,5} Among the trypanosomatid species included in our analyses, loss of certain genes involved in AD are mainly observed in *Phytomonas* spp. and in the *Trypanosoma* genus, indicating a continued reduction in pathways related to the AD during trypanosomatid diversification. The lack of SAT, CS, POT1, GspS, and APx appears to be related to extracellular parasitic lifestyles, including *Phytomonas* spp., *T. brucei*, *T. vivax*, and *T. evansi* (Fig. 7). The exception was *T. grayi*, which is an extracellular African trypanosome of crocodiles closely related to other parasites of crocodiles in South America.³⁵ *T. grayi* is a stercorarian trypanosome (like *T. cruzi*), transmitted by tsetse flies via oral contamination with infective metacyclics present in the insect feces. Our results thus indicate differences in the antioxidant system between the stercorarian and salivarian trypanosomes. As mentioned earlier, the extracellular Stercoraria section trypanosomes lacked certain genes, and in salivarian parasites, we observed high copy numbers of CBS, the absence of ODC, and the presence of POT1. Interestingly, *T. rangeli*, whose life cycle in the mammalian host is unknown and is mainly transmitted by triatomine bites, lacks certain enzymes (SAT, CS, GspS, and APx) like Salivaria parasites and others (ODC) like Stercoraria parasites. Interestingly, the SAT, CS, and APx genes related to the AD are found as pseudogenes in *T. rangeli*, reinforcing the undefined taxonomic position between Salivaria and Stercoraria of this taxon.⁸⁸ These genes have premature stop codons before the specific domains, and CS has been experimentally demonstrated to be inactive in *T. rangeli*.⁶⁰ A large amount of pseudogenes was described in the *T. rangeli* genome, including genes that compose the RNA interference machinery.³¹ As well observed to some antioxidant-related genes, RNA interference machinery genes are observed in Salivaria parasites, pseudogenes are found in *T. rangeli*, and these same genes are absent in *T. cruzi*.⁸⁹

All studied enzymes identified in *T. rangeli* showed higher similarity with *T. cruzi* orthologs than *T. grayi*. However, when we compared the antioxidant system overall, including the presence or absence of enzymes, *T. cruzi* is more closely related to the African *T. grayi* than the South American *T. rangeli*. This situation reinforces that the presence of specific pathways in each parasite was not a direct result of geographical isolation but is based on differences in the parasitic lifecycle, mainly within their invertebrate hosts. The presence of parasite forms in other insect tissues/organs (hemolymph, salivary glands, or proboscides) than in the alimentary tract seems to be a common feature among the trypanosomatid species that lack more genes related to the AD.

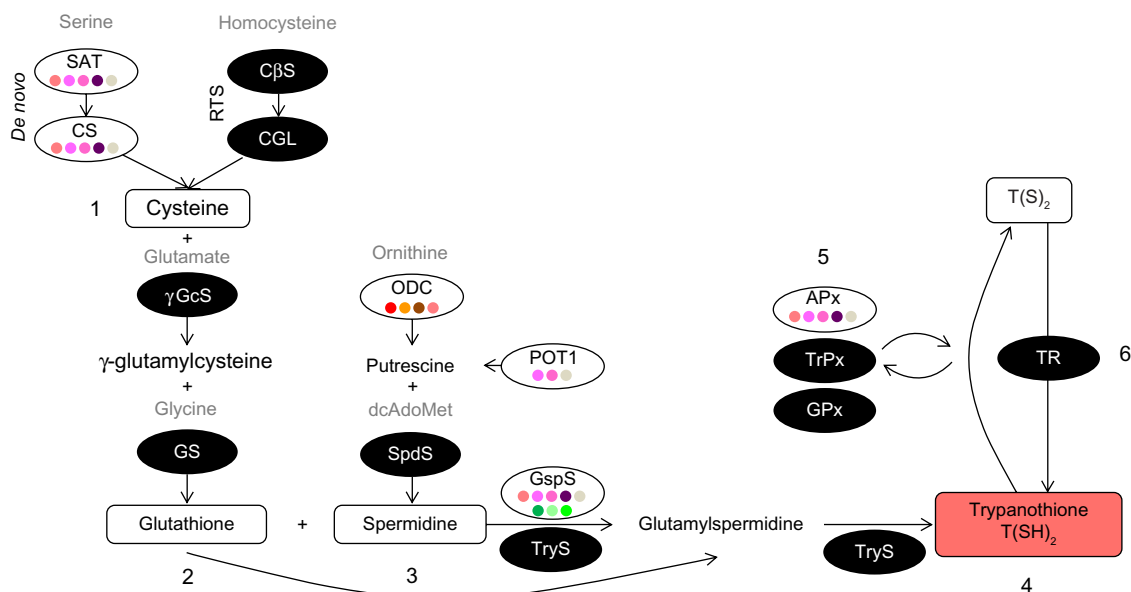


Figure 7. Adaptation of trypanothione ($T(SH)_2$) metabolism in trypanosomatids. The synthesis of $T(SH)_2$ depends on the biosynthesis and/or uptake of cysteine (1), glutathione (2), and polyamine (3) that ultimately allows the conjugation of two glutathione molecules to spermidine (4). Oxidation of $T(SH)_2$ is required to several cellular functions, including the neutralization of radicals and the reduction of biomolecules performed by peroxidases (5). The regeneration of reduced $T(SH)_2$ is catalyzed by trypanothione reductase (TR) (6). Black-filled ovals are fully conserved enzymes involved in trypanothione metabolism. Open ovals represent enzymes that are absent in some species. These species are represented by colored circles as follows: ● *T. cruzi*; ● *T. cruzi marinkellei*; ● *T. grayi*; ● *T. rangeli*; ● *T. brucei*; ● *T. evansi*; ● *T. vivax*; ● *L. major*; ● *L. braziliensis*; ● *L. tarentolae*; ● *Phytomonas* spp.

Conclusion

Most of the studies focusing on the adaptation of trypanosomatids to parasitism were carried out using species pathogenic to humans. However, novel insights concerning such mechanisms have emerged from comparative genomics using *B. saltans*, pathogenic trypanosomes to wild animals and livestock and other nonpathogenic trypanosomatid species. In this sense, partial or complete loss of genes or even entire metabolic pathways related to catabolism of some amino acids has been described.^{3,5,44} Specialization on some metabolic pathways such as the glucose metabolism allows clustering trypanosomatids into two distinct clades composed by the African trypanosomes and by *Phytomonas* spp., suggesting a convergent evolution. Despite the geographical isolation, species from both clades revealed a similar set of absent genes.³ In this study, we assessed the plasticity of the AD machinery at the genomic level using distinct clades of parasitic trypanosomatids. The “African trypanosomes” and the “*Phytomonas* spp.” clades have basically revealed the loss of the same antioxidant gene repertoire, reinforcing the convergent evolution hypothesis.

In this evolutionary context and aiming to provide further evidence of the convergent evolution, the following two intriguing species were further included in our study: the American and mammalian infective *T. rangeli* and the African parasite of crocodiles, *T. grayi*. *T. rangeli* shares the majority of the lost genes (including some pseudogenes) of the antioxidant machinery with *T. brucei*, *T. evansi*, *T. vivax*, and *Phytomonas* spp., maintaining the majority of the developmental stages outside the insect gut and being transmitted to its hosts by bite (anterior

transmission) of infected vectors. On the other hand, *T. grayi* revealed an antioxidant machinery similar to the American and mammalian infective *T. cruzi*. Both *T. grayi* and *T. cruzi* develop within their invertebrate host gut (posterior station) and are transmitted to their vertebrate hosts via contaminative method.

The inclusion of *T. rangeli* and *T. grayi* in our analyses indicates the contradictory use of the term *African trypanosomes*, present in most of the studies on trypanosomatid evolution, leading to the misinterpretation that evolution of these species is due geographical isolation. Considering that our results indicate that adaptative evolution of the AD mechanisms in trypanosomatids is related to their biology within the vectors, we propose the use of proper taxonomic nomenclature (genomic and phenotypic) instead of clade denominations.

In conclusion, evolution of the AD system in trypanosomatids includes recent evolutionary events mostly related to adaptations of each species to distinct environments within their insect vectors instead to geographic isolation. Although a genomic core related to AD is preserved from the most ancestral organisms to the trypanosomatids, the loss, acquisition, and/or duplication of genes seem to be specifically related to adaptative mechanisms to the parasite life cycle within their invertebrate hosts, influencing their ability to survive, multiply, and/or generate infective forms.

Acknowledgments

The authors thank Dr Glauber Wagner for his bioinformatics support and Dr Guilherme de Toledo e Silva for critical reading of the article.



Author Contributions

Conceived and designed the experiments: ITB-B, ECG, PHS. Analyzed the data: ITB-B, CT-L, BA, ECG, PHS. Wrote the first draft of the article: ITB-B, PHS. Contributed to the writing of the article: ITB-B, CT-L, BA, ECG, PHS. Agreed with the article results and conclusions: ITB-B, CT-L, BA, ECG, PHS. Jointly developed the structure and arguments for the article: ITB-B, CT-L, BA, ECG, PHS. Made critical revisions and approved the final version: ITB-B, CT-L, BA, ECG, PHS. All authors reviewed and approved the final article.

Supplementary Material

Supplementary File 1. General characteristics of kinetoplastid species used in the antioxidant defence genes analyses.

Supplementary File 2. Multiple sequence alignments of antioxidant defence related protein sequences from kinetoplastids.

Supplementary File 3. General characteristics of kinetoplastid genes related to the cysteine and glutathione biosynthesis.

Supplementary File 4. General characteristics of kinetoplastid genes related to the polyamine synthesis and uptake.

Supplementary File 5. General characteristics of kinetoplastid genes related to the trypanothione synthesis and reduction.

Supplementary File 6. Molecular phylogenetic analyses of kinetoplastid Trypanothione reductase (TR). Protein alignment was made with Clustal Omega and the evolutionary history was inferred by using the Maximum Likelihood method using MEGA version 6.0. Phylogeny of TR was constructed using the LG + G amino-acid substitution model. Tree is drawn to scale, with branch lengths measured in the number of substitutions per site. The analysis involved 22 amino acid sequences (492 sites). Bootstrap percentages above 80% are shown next to the branches, corresponding to the percentage of trees in which the associated taxa clustered together (1,000 replicates). Kinetoplastids annotated sequences were retrieved from TriTrypDB or NCBI and the access numbers are included.

Supplementary File 7. General characteristics of kinetoplastid peroxidases genes.

Supplementary File 8. General characteristics of kinetoplastid Iron Superoxide Dismutase genes.

REFERENCES

- Maslov DA, Votycka J, Yurchenko V, Lukes J. Diversity and phylogeny of insect trypanosomatids: all that is hidden shall be revealed. *Trends Parasitol.* 2013;29(1):43–52.
- Simpson AG, Stevens JR, Lukes J. The evolution and diversity of kinetoplastid flagellates. *Trends Parasitol.* 2006;22(4):168–74.
- Opperdoes FR, Butenko A, Flegontov P, Yurchenko V, Lukes J. Comparative metabolism of free-living *Bodo saltans* and parasitic trypanosomatids. *J Eukaryot Microbiol.* 2016;63(5):657–78.
- Lukes J, Skalicky T, Tyc J, Votycka J, Yurchenko V. Evolution of parasitism in kinetoplastid flagellates. *Mol Biochem Parasitol.* 2014;195(2):115–22.
- Jackson AP, Otto TD, Aslett M, et al. Kinetoplastid phylogenomics reveals the evolutionary innovations associated with the origins of parasitism. *Curr Biol.* 2016;26(2):161–72.
- Deschamps P, Lara E, Marande W, Lopez-Garcia P, Ekelund F, Moreira D. Phylogenomic analysis of kinetoplastids supports that trypanosomatids arose from within bodonids. *Mol Biol Evol.* 2011;28(1):53–8.
- Piacenza L, Peluffo G, Alvarez MN, Martinez A, Radi R. *Trypanosoma cruzi* antioxidant enzymes as virulence factors in Chagas disease. *Antioxid Redox Signal.* 2013;19(7):723–34.
- Alvarez MN, Peluffo G, Piacenza L, Radi R. Intraphagosomal peroxynitrite as a macrophage-derived cytotoxin against internalized *Trypanosoma cruzi*: consequences for oxidative killing and role of microbial peroxiredoxins in infectivity. *J Biol Chem.* 2011;286(8):6627–40.
- Machado-Silva A, Cerqueira PG, Grazielle-Silva V, et al. How *Trypanosoma cruzi* deals with oxidative stress: antioxidant defence and DNA repair pathways. *Mutat Res Rev Mutat Res.* 2016;767:8–22.
- Cosentino-Gomes D, Rocco-Machado N, Meyer-Fernandes JR. Rhodnius prolixus: modulation of antioxidant defenses by *Trypanosoma rangeli*. *Exp Parasitol.* 2014;145:118–24.
- Finzi JK, Chiavegatto CW, Corat KF, et al. *Trypanosoma cruzi* response to the oxidative stress generated by hydrogen peroxide. *Mol Biochem Parasitol.* 2004;133(1):37–43.
- Flohe L. The trypanothione system and its implications in the therapy of trypanosomatid diseases. *Int J Med Microbiol.* 2012;302(4–5):216–20.
- Peloso EF, Goncalves CC, Silva TM, et al. Tryparedoxin peroxidases and superoxide dismutases expression as well as ROS release are related to *Trypanosoma cruzi* epimastigotes growth phases. *Arch Biochem Biophys.* 2012;520(2):117–22.
- Peloso EF, Vitor SC, Ribeiro LH, Pineyro MD, Robello C, Gadelha FR. Role of *Trypanosoma cruzi* peroxiredoxins in mitochondrial bioenergetics. *J Bioenerg Biomembr.* 2011;43(4):419–24.
- Arias DG, Marquez VE, Chiribao ML, et al. Redox metabolism in *Trypanosoma cruzi*: functional characterization of tryparedoxins revisited. *Free Radic Biol Med.* 2013;63:65–77.
- Weydert CJ, Cullen JJ. Measurement of superoxide dismutase, catalase and glutathione peroxidase in cultured cells and tissue. *Nat Protoc.* 2010;5(1):51–66.
- Muller S, Liebau E, Walter RD, Krauth-Siegel RL. Thiol-based redox metabolism of protozoan parasites. *Trends Parasitol.* 2003;19(7):320–8.
- Fairlamb AH, Cerami A. Metabolism and functions of trypanothione in the Kinetoplastida. *Annu Rev Microbiol.* 1992;46:695–729.
- Castro H, Tomas AM. Peroxidases of trypanosomatids. *Antioxid Redox Signal.* 2008;10(9):1593–606.
- Ismail SO, Paramchuk W, Skeiky YA, Reed SG, Bhatia A, Gedamu L. Molecular cloning and characterization of two iron superoxide dismutase cDNAs from *Trypanosoma cruzi*. *Mol Biochem Parasitol.* 1997;86(2):187–97.
- Mateo H, Marin C, Perez-Cordon G, Sanchez-Moreno M. Purification and biochemical characterization of four iron superoxide dismutases in *Trypanosoma cruzi*. *Mem Inst Oswaldo Cruz.* 2008;103(3):271–6.
- Wilkinson SR, Taylor MC, Touitha S, Mauricio IL, Meyer DJ, Kelly JM. TcGPXII, a glutathione-dependent *Trypanosoma cruzi* peroxidase with substrate specificity restricted to fatty acid and phospholipid hydroperoxides, is localized to the endoplasmic reticulum. *Biochem J.* 2002;364(pt 3):787–94.
- Wilkinson SR, Meyer DJ, Taylor MC, Bromley EV, Miles MA, Kelly JM. The *Trypanosoma cruzi* enzyme TcGPXI is a glycosomal peroxidase and can be linked to trypanothione reduction by glutathione or tryparedoxin. *J Biol Chem.* 2002;277(19):17062–71.
- Pineyro MD, Pizarro JC, Lema F, et al. Crystal structure of the tryparedoxin peroxidase from the human parasite *Trypanosoma cruzi*. *J Struct Biol.* 2005;150(1):11–22.
- Piacenza L, Peluffo G, Alvarez MN, Kelly JM, Wilkinson SR, Radi R. Peroxiredoxins play a major role in protecting *Trypanosoma cruzi* against macrophage- and endogenously-derived peroxynitrite. *Biochem J.* 2008;410(2):359–68.
- Wilkinson SR, Temperton NJ, Mondragon A, Kelly JM. Distinct mitochondrial and cytosolic enzymes mediate trypanothione-dependent peroxide metabolism in *Trypanosoma cruzi*. *J Biol Chem.* 2000;275(11):8220–5.
- Trujillo M, Budde H, Pineyro MD, et al. *Trypanosoma brucei* and *Trypanosoma cruzi* tryparedoxin peroxidases catalytically detoxify peroxynitrite via oxidation of fast reacting thiols. *J Biol Chem.* 2004;279(33):34175–82.
- Wilkinson SR, Obado SO, Mauricio IL, Kelly JM. *Trypanosoma cruzi* expresses a plant-like ascorbate-dependent hemoperoxidase localized to the endoplasmic reticulum. *Proc Natl Acad Sci U S A.* 2002;99(21):13453–8.
- Krieger S, Schwarz W, Ariyanayagam MR, Fairlamb AH, Krauth-Siegel RL, Clayton C. Trypanosomes lacking trypanothione reductase are avirulent and show increased sensitivity to oxidative stress. *Mol Microbiol.* 2000;35(3):542–52.
- Krauth-Siegel RL, Enders B, Henderson GB, Fairlamb AH, Schirmer RH. Trypanothione reductase from *Trypanosoma cruzi*. Purification and characterization of the crystalline enzyme. *Eur J Biochem.* 1987;164(1):123–8.
- Stoco PH, Wagner G, Talavera-Lopez C, et al. Genome of the avirulent human-infective trypanosome – *Trypanosoma rangeli*. *PLoS Negl Trop Dis.* 2014; 8(9):e3176.
- El-Sayed NM, Myler PJ, Bartholomeu DC, et al. The genome sequence of *Trypanosoma cruzi*, etiologic agent of Chagas disease. *Science.* 2005;309(5733):409–15.
- Franzen O, Talavera-Lopez C, Ochaya S, et al. Comparative genomic analysis of human infective *Trypanosoma cruzi* lineages with the bat-restricted subspecies *T. cruzi marinkellei*. *BMC Genomics.* 2012;13:531.



34. Berriman M, Ghedin E, Hertz-Fowler C, et al. The genome of the African trypanosome *Trypanosoma brucei*. *Science*. 2005;309(5733):416–22.
35. Kelly S, Ivens A, Manna PT, Gibson W, Field MC. A draft genome for the African crocodilian trypanosome *Trypanosoma grayi*. *Sci Data*. 2014;1:140024.
36. Carnes J, Anupama A, Balmer O, et al. Genome and phylogenetic analyses of *Trypanosoma evansi* reveal extensive similarity to *T. brucei* and multiple independent origins for dyskinetoplasty. *PLoS Negl Trop Dis*. 2015;9(1):e3404.
37. Jackson AP, Berry A, Aslett M, et al. Antigenic diversity is generated by distinct evolutionary mechanisms in African trypanosome species. *Proc Natl Acad Sci U S A*. 2012;109(9):3416–21.
38. Peacock CS, Seeger K, Harris D, et al. Comparative genomic analysis of three *Leishmania* species that cause diverse human disease. *Nat Genet*. 2007;39(7):839–47.
39. Ivens AC, Peacock CS, Worthey EA, et al. The genome of the kinetoplastid parasite, *Leishmania major*. *Science*. 2005;309(5733):436–42.
40. Raymond F, Boisvert S, Roy G, et al. Genome sequencing of the lizard parasite *Leishmania tarentolae* reveals loss of genes associated to the intracellular stage of human pathogenic species. *Nucleic Acids Res*. 2012;40(3):1131–47.
41. Motta MC, Martins AC, de Souza SS, et al. Predicting the proteins of *Angomonas deanei*, *Strigomonas culicis* and their respective endosymbionts reveals new aspects of the trypanosomatidae family. *PLoS One*. 2013;8(4):e60209.
42. Kraeva N, Butenko A, Hlavacova J, et al. *Leptomonas seymouri*: adaptations to the dixenous life cycle analyzed by genome sequencing, transcriptome profiling and co-infection with *Leishmania donovani*. *PLoS Pathog*. 2015;11(8):e1005127.
43. Porcel BM, Denoed F, Opperdoes F, et al. The streamlined genome of *Phytomonas* spp. relative to human pathogenic kinetoplastids reveals a parasite tailored for plants. *PLoS Genet*. 2014;10(2):e1004007.
44. Jackson AP. Genome evolution in trypanosomatid parasites. *Parasitology*. 2015;142(suppl 1):S40–56.
45. Li L, Stoeckert CJ Jr, Roos DS. OrthoMCL: identification of ortholog groups for eukaryotic genomes. *Genome Res*. 2003;13(9):2178–89.
46. Petersen TN, Brunak S, von Heijne G, Nielsen H. SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat Methods*. 2011;8(10):785–6.
47. Kall L, Krogh A, Sonnhammer EL. A combined transmembrane topology and signal peptide prediction method. *J Mol Biol*. 2004;338(5):1027–36.
48. Hiller K, Grote A, Scheer M, Munch R, Jahn D. PrediSi: prediction of signal peptides and their cleavage positions. *Nucleic Acids Res*. 2004;32(Web Server issue):W375–9.
49. Bannai H, Tamada Y, Maruyama O, Nakai K, Miyano S. Extensive feature detection of N-terminal protein sorting signals. *Bioinformatics*. 2002;18(2):298–305.
50. Emanuelsson O, Nielsen H, Brunak S, von Heijne G. Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. *J Mol Biol*. 2000;300(4):1005–16.
51. Claros MG, Vincens P. Computational method to predict mitochondrially imported proteins and their targeting sequences. *Eur J Biochem*. 1996;241(3):779–86.
52. Garg A, Raghava GP. ESLpred2: improved method for predicting subcellular localization of eukaryotic proteins. *BMC Bioinformatics*. 2008;9:503.
53. Sievers F, Higgins DG. Clustal omega. *Curr Protoc Bioinformatics*. 2014;48:3.13.1–3.13.16.
54. Tamura K, Stecher G, Peterson D, Filipinski A, Kumar S. MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol Biol Evol*. 2013;30(12):2725–9.
55. Le SQ, Gascuel O. An improved general amino acid replacement matrix. *Mol Biol Evol*. 2008;25(7):1307–20.
56. Manta B, Comini M, Medeiros A, Hugo M, Trujillo M, Radi R. Trypanothione: a unique bis-glutathionyl derivative in trypanosomatids. *Biochim Biophys Acta*. 2013;1830(5):3199–216.
57. Landfear SM. Nutrient transport and pathogenesis in selected parasitic protozoa. *Eukaryot Cell*. 2011;10(4):483–93.
58. Marciano D, Santana M, Nowicki C. Functional characterization of enzymes involved in cysteine biosynthesis and H(2)S production in *Trypanosoma cruzi*. *Mol Biochem Parasitol*. 2012;185(2):114–20.
59. Canepa GE, Bouvier LA, Miranda MR, Uttaro AD, Pereira CA. Characterization of *Trypanosoma cruzi* L-cysteine transport mechanisms and their adaptive regulation. *FEMS Microbiol Lett*. 2009;292(1):27–32.
60. Romero I, Tellez J, Yamanaka LE, Steindel M, Romanha AJ, Grisard EC. Transsulfuration is an active pathway for cysteine biosynthesis in *Trypanosoma rangeli*. *Parasit Vectors*. 2014;7:197.
61. Heby O, Persson L, Rentala M. Targeting the polyamine biosynthetic enzymes: a promising approach to therapy of African sleeping sickness, Chagas' disease, and leishmaniasis. *Amino Acids*. 2007;33(2):359–66.
62. Gonzalez NS, Huber A, Algranati ID. Spermidine is essential for normal proliferation of trypanosomatid protozoa. *FEBS Lett*. 2001;508(3):323–6.
63. Hasne MP, Ullman B. Identification and characterization of a polyamine permease from the protozoan parasite *Leishmania major*. *J Biol Chem*. 2005;280(15):15188–94.
64. Heby O, Roberts SC, Ullman B. Polyamine biosynthetic enzymes as drug targets in parasitic protozoa. *Biochem Soc Trans*. 2003;31(2):415–9.
65. Hasne MP, Coppens I, Soysa R, Ullman B. A high-affinity putrescine-cadaverine transporter from *Trypanosoma cruzi*. *Mol Microbiol*. 2010;76(1):78–91.
66. Steglich C, Schaeffer SW. The ornithine decarboxylase gene of *Trypanosoma brucei*: evidence for horizontal gene transfer from a vertebrate source. *Infect Genet Evol*. 2006;6(3):205–19.
67. Alves JM, Klein CC, da Silva FM, et al. Endosymbiosis in trypanosomatids: the genomic cooperation between bacterium and host in the synthesis of essential amino acids is heavily influenced by multiple horizontal gene transfers. *BMC Evol Biol*. 2013;13:190.
68. Wang W, Ballatori N. Endogenous glutathione conjugates: occurrence and biological functions. *Pharmacol Rev*. 1998;50(3):335–56.
69. Oza SL, Ariyanayagam MR, Fairlamb AH. Characterization of recombinant glutathionylspermidine synthetase/amidase from *Crithidia fasciculata*. *Biochem J*. 2002;364(pt 3):679–86.
70. Smith K, Nadeau K, Bradley M, Walsh C, Fairlamb AH. Purification of glutathionylspermidine and trypanothione synthetases from *Crithidia fasciculata*. *Protein Sci*. 1992;1(7):874–83.
71. Turrens JF. Oxidative stress and antioxidant defenses: a target for the treatment of diseases caused by parasitic protozoa. *Mol Aspects Med*. 2004;25(1–2):211–20.
72. Patel S, Hussain S, Harris R, et al. Structural insights into the catalytic mechanism of *Trypanosoma cruzi* GPXI (glutathione peroxidase-like enzyme I). *Biochem J*. 2010;425(3):513–22.
73. Flohe L, Hecht HJ, Steinert P. Glutathione and trypanothione in parasitic hydroperoxide metabolism. *Free Radic Biol Med*. 1999;27(9–10):966–84.
74. Tetaud E, Fairlamb AH. Cloning, expression and reconstitution of the trypanothione-dependent peroxidase system of *Crithidia fasciculata*. *Mol Biochem Parasitol*. 1998;96(1–2):111–23.
75. König J, Fairlamb AH. A comparative study of type I and type II trypanedoxin peroxidases in *Leishmania major*. *FEBS J*. 2007;274(21):5643–58.
76. Schlecker T, Schmidt A, Dirdjaja N, Voncken F, Clayton C, Krauth-Siegel RL. Substrate specificity, localization, and essential role of the glutathione peroxidase-type trypanedoxin peroxidases in *Trypanosoma brucei*. *J Biol Chem*. 2005;280(15):14385–94.
77. Alphey MS, König J, Fairlamb AH. Structural and mechanistic insights into type II trypanosomatid trypanedoxin-dependent peroxidases. *Biochem J*. 2008;414(3):375–81.
78. Navrot N, Collin V, Gualberto J, et al. Plant glutathione peroxidases are functional peroxidoxins distributed in several subcellular compartments and regulated during biotic and abiotic stresses. *Plant Physiol*. 2006;142(4):1364–79.
79. Krauth-Siegel RL, Ludemann H. Reduction of dehydroascorbate by trypanothione. *Mol Biochem Parasitol*. 1996;80(2):203–8.
80. Adak S, Datta AK. *Leishmania major* encodes an unusual peroxidase that is a close homologue of plant ascorbate peroxidase: a novel role of the transmembrane domain. *Biochem J*. 2005;390(pt 2):465–74.
81. Wilkinson SR, Prathalingam SR, Taylor MC, Horn D, Kelly JM. Vitamin C biosynthesis in trypanosomes: a role for the glycosome. *Proc Natl Acad Sci U S A*. 2005;102(33):11645–50.
82. Nogueira FB, Rodrigues JF, Correa MM, Ruiz JC, Romanha AJ, Murta SM. The level of ascorbate peroxidase is enhanced in benznidazole-resistant populations of *Trypanosoma cruzi* and its expression is modulated by stress generated by hydrogen peroxide. *Mem Inst Oswaldo Cruz*. 2012;107(4):494–502.
83. Hannaert V, Saavedra E, Duffieux F, et al. Plant-like traits associated with metabolism of *Trypanosoma* parasites. *Proc Natl Acad Sci U S A*. 2003;100(3):1067–71.
84. Dufernez F, Yernaux C, Gerbod D, et al. The presence of four iron-containing superoxide dismutase isozymes in trypanosomatidae: characterization, subcellular localization, and phylogenetic origin in *Trypanosoma brucei*. *Free Radic Biol Med*. 2006;40(2):210–25.
85. Wintjens R, Noel C, May AC, et al. Specificity and phenetic relationships of iron- and manganese-containing superoxide dismutases on the basis of structure and sequence comparisons. *J Biol Chem*. 2004;279(10):9248–54.
86. Irigoien F, Cibils L, Comini MA, Wilkinson SR, Flohe L, Radi R. Insights into the redox biology of *Trypanosoma cruzi*: trypanothione metabolism and oxidant detoxification. *Free Radic Biol Med*. 2008;45(6):733–42.
87. Sommer JM, Wang CC. Targeting proteins to the glycosomes of African trypanosomes. *Annu Rev Microbiol*. 1994;48:105–38.
88. Grisard EC. Salivaria or stercoraria? The *Trypanosoma rangeli* dilemma. *Kinetoplastid Biol Dis*. 2002;1(1):5.
89. Kolev NG, Tschudi C, Ullu E. RNA interference in protozoan parasites: achievements and challenges. *Eukaryot Cell*. 2011;10(9):1156–63.