



Universidade de Brasília
Faculdade de Medicina
Programa de Pós-Graduação em Patologia Molecular
Laboratório de Interação Patógeno-Hospedeiro

Análise proteômica de epimastigotas da cepa G de *Trypanosoma cruzi* isolada de um gambá da Amazônia

Raquel Silva de Negreiros

Brasília

2013

Raquel Silva de Negreiros

**Análise proteômica de epimastigotas da
cepa G de *Trypanosoma cruzi* isolada de
um gambá da Amazônia**

**Tese apresentada ao programa de
Pós-Graduação em Patologia
Molecular da Universidade de
Brasília como requisito parcial à
obtenção do título de Doutor em
Patologia Molecular**

**Orientador: Prof. Dr. Jaime
Martins de Santana**

Brasília

2013

Trabalho realizado no Laboratório de Interação Patógeno-Hospedeiro (LIPH) junto ao Programa de Pós-Graduação em Patologia Molecular da Faculdade de Medicina da Universidade de Brasília.

Apoio financeiro: CNPq, CAPES e FAP-DF.

Dedico esta tese aos meus queridos pais, Djaci e Francisca, pelo amor e pela assistência plenos.

AGRADECIMENTOS

Primeiramente, agradeço a Deus e a Meishu-Sama pela vida, pela oportunidade de ser útil à Ciência e por toda a força e a paciência concedidas a mim durante a realização do doutorado.

Tenho profunda gratidão aos meus melhores amigos e pais, Djaci Magalhães e Francisca Negreiros: pelo imenso privilégio de tê-los como pais, pela vida, pelo amor infinito, pelo apoio nos momentos de aprimoramento e em quaisquer escolhas acadêmicas e profissionais que eu faça, pela educação alicerçada no altruísmo, por acreditarem no meu potencial, pela paciência. Determinação e dedicação são características que vocês mostraram ter em abundância. Todos meus anos de estudo envolvendo colégio, graduação e pós-graduação são dedicados a vocês, assim como dedicaram suas vidas em oferecer o que têm de melhor para mim. Meu pai, meu exemplo-mor de otimismo, estará sempre comigo onde quer que eu esteja, pois nosso elo é eterno. Minha mãe é minha verdadeira heroína, uma pessoa que superou com firmeza todos os obstáculos que a vida lhe trouxe. Espero estar retribuindo satisfatoriamente suas expectativas. Vocês tiveram fundamental importância para a concretização deste propósito que é uma conquista nossa.

Minha sincera gratidão ao meu orientador, prof. Jaime Santana, pela oportunidade de aprender, desenvolver e aprimorar meu pensamento científico durante mais de sete anos de convivência, desde a iniciação científica. Profissionalmente, finalizo este curso com muito mais bagagem científica e sabedoria do que quando iniciei, em grande parte, devido aos seus ensinamentos, ao seu incentivo e à confiança que depositou em mim. Pessoalmente, você foi uma pessoa com quem pude contar e que me acolheu durante a fase difícil pela qual passei.

Aos prof. Sebastien Charneau, Magno Junqueira e Wagner Fontes, que me guiaram pelos caminhos da proteômica e colaboraram de modo crucial neste estudo. De modo especial, agradeço ao prof. Sebastien pelas discussões científicas enriquecedoras, por sempre se mostrar tão solícito durante a fase final e crítica do doutorado, enfim, pela intensa e primordial colaboração.

Ao Rayner Myr e ao Jaques Souza, que me deram todo o suporte nos experimentos de proteômica, e ao Agenor Castro pela ajuda na PCR para genotipagem.

Ao Thiago Gastardelo e ao Hugo de Almeida, por me receberem tão gentilmente e me direcionarem durante os meus passos pioneiros pela Ciência.

Ao prof. Bergmann Ribeiro pela acolhida em seu laboratório durante o período inicial da minha iniciação científica.

Ao prof. Carlos André Ricart, coordenador do programa de Pós-Graduação em Patologia Molecular, por compreender as adversidades que enfrentamos no desenrolar do projeto.

Aos amigos e aos professores dos laboratórios da UnB pelos quais passei ou tive contato mais próximo: Laboratório de Interação Patógeno-Hospedeiro, Laboratório de Bioquímica e Química de Proteínas e Laboratório de Virologia.

Especialmente aos amigos e aos professores do LIPH pela convivência profissional, pelo apoio e pela diversão: André Corrêa, Brina Portugal, Bruna Guido, Camila Lasse, Carla Araújo, Clênia Azevedo, David Neves, Débora Torres, Flávia Nader, Graziella Figueiredo, Hugo de Almeida, Isabel Sousa, Izabela Dourado, Luz Irazabal, Marta Rosa, Paula Beatriz, Tatiana Giraldo, Thiago Gastardelo e a todos os estagiários que permaneceram ou não no laboratório e são essenciais para o compartilhamento mútuo do conhecimento; e também à Solange Rego do Laboratório de Toxinologia.

Aos amigos do LIA e do LIM1 com os quais compartilhei o laboratório durante muito tempo, especialmente a: Ana Camila Souza, Isaque Medeiros, Márcio Jerônimo, Pedro Saavedra, Rafael Corrêa, Raffael Castro e Raquel Almeida; e aos vizinhos do Laboratório de Biologia do Gene.

Aos amigos e aos professores da graduação em Biologia, com os quais dividi muitas experiências, dias de estudos e diversão e, principalmente, a paixão pelo estudo da vida.

Aos meus familiares que torcem e me incentivam sempre.

A todos os outros amigos que fazem parte do meu cotidiano ou que estão distantes por desígnios da vida.

"Aprender por aprender é estudo morto, enquanto aprender algo para ser utilizado na sociedade é estudo vivo."

Meishu-Sama

RESUMO

NEGREIROS, R. S. **Análise proteômica de epimastigotas da cepa G de *Trypanosoma cruzi* isolada de um gambá da Amazônia.** 199 f. Tese de doutorado - Faculdade de Medicina, Universidade de Brasília, Brasília, 2013.

Diversos isolados e cepas de *Trypanosoma cruzi* são classificados em seis grupos genéticos ou *discrete typing units* (DTUs), TcI a TcVI. A cepa G estudada no presente trabalho pertence ao DTU TcI e foi isolada de um gambá da região Amazônica brasileira. Parasitos pertencentes a essa cepa apresentam um genoma não híbrido e menor que da cepa CL Brener. Além disso, tripomastigotas metacíclicas da cepa G são qualificados como pouco virulentas e, surpreendentemente, amastigotas axênicas são altamente infectivas. Dentro de um amplo projeto do nosso grupo de pesquisa que envolve o futuro sequenciamento do genoma da cepa G de *T. cruzi*, está inserida a análise proteômica de epimastigotas em larga escala que auxiliará na anotação dos genes. Para isso, foi realizada a análise por LC-MS/MS dos peptídeos gerados por digestão em gel do lisado proteico de epimastigotas. Foram feitas três buscas automatizadas em bancos de dados. Para a primeira busca, utilizando-se do banco de dados de *Trypanosoma* spp. e outros tripanossomatídeos, 1.460 proteínas foram identificadas. Para a segunda busca, no banco de dados apenas da cepa CL Brener, foram obtidas 953 identificações. Para ambas as buscas, foi utilizado o algoritmo do programa Sequest. Já com a utilização do algoritmo do PEAKS, foram detectadas 2.947 proteínas a partir do confronto dos espectros MS/MS com o banco de dados das cepas CL Brener e Sylvio X10/1. Considerando essa última abordagem, este é o proteoma mais abrangente realizado até hoje disponível independentemente do estágio de vida. Em todas as buscas, a via metabólica mais abundante em número de anotações foi metabolismo de purinas, o que está condizente com o alto requerimento energético de formas epimastigotas replicativas. A via glicólise também foi listada entre as dez vias metabólicas mais representativas em todas as buscas, provavelmente devido à presença de glicose no meio de cultivo e à preferência por essa hexose como fonte de carbono por epimastigotas. Todas as buscas resultaram na predominância de proteínas preditas a serem secretadas pela

via não clássica e na identificação de diversas proteínas integrais de membrana potenciais, mostrando que a extração de proteínas por solubilização com SDS foi eficiente para o propósito de um proteoma abrangente. Particularmente, na busca 3, foram identificadas várias proteínas candidatas a alvos de drogas estudadas pelo nosso grupo, como leucil-aminopeptidase (LAP), oligopeptidase B (OPB), prolil-oligopeptidase (POP), metiltioadenosina fosforilase (MTAF) e catepsina B. Em relação às famílias multigênicas de *T. cruzi*, a terceira busca resultou na identificação de uma grande quantidade de membros das famílias de trans-sialidases (TS), *mucin-associated surface proteins* (MASP), mucinas, proteínas *retrotransposon hot spot* (RHS) e *dispersed gene family 1* (DGF-1) e glicoproteínas de 63 kDa (gp63). Portanto, podemos concluir que a busca no banco de dados das cepas CL Brener e Sylvio X10/1 foi a mais satisfatória. Este proteoma representa um avanço significativo para o conhecimento do conjunto de proteínas expresso em *T. cruzi* e, também, um passo importante para a elucidação da sequência genômica da cepa G de *T. cruzi*.

Palavras-chave: doença de Chagas, *Trypanosoma cruzi*, cepa G, epimastigotas, proteoma, espectrometria de massas.

ABSTRACT

NEGREIROS, R. S. **Proteomic analysis of G strain epimastigotes of *Trypanosoma cruzi* isolated from an opossum of Amazonic Region.** 199 f. Thesis (doctorate) - Faculdade de Medicina, Universidade de Brasília, Brasília, 2013.

Many isolates and strains of *Trypanosoma cruzi* are classified into six genetic groups or DTUs (*discrete typing units*), TcI to TcVI. G strain studied in the present work belongs to DTU TcI and was isolated from an opossum of Brazilian Amazon. G strain parasites have a non-hybrid and smaller genome than CL Brener strain. Moreover, G strain metacyclic trypomastigotes forms are characterized as low virulent and surprisingly axenic amastigote forms are highly infective. Within a broad project of our research group, which involves the subsequent genome sequencing of *T. cruzi* G strain, high-throughput proteomic analysis of epimastigotes is inserted to help gene annotation. For this, LC-MS/MS analysis of peptides generated by in-gel digestion of epimastigote protein lysate was accomplished. Three database automated searches were done. The search using the database of *Trypanosoma* spp. and other trypanosomatids allowed identification of 1.460 proteins. Secondly the search against *T. cruzi* CL Brener strain database resulted in 953 identifications. For both searches, Sequest algorithm was employed. With the use of PEAKS algorithm, 2.947 were detected from search against both CL Brener and Sylvio X10/1 strains database. Considering the last approach, this is the most abundant proteome that has been revealed so far regardless the life stage of the parasite. In all searches, the most abundant metabolic pathway in number of annotations was purine metabolism, which is consistent with high energy requirement of replicative epimastigote forms. The glycolysis pathway also was related to the ten most representative metabolic pathways in all searches, probably due to the presence of glucose in culture medium and the preference for this hexose as carbon source by epimastigotes. All searches revealed the predominance of proteins secreted by non classical pathway and in the identification of many potential integral membrane proteins, showing that protein extraction was efficient for the purpose of an in-depth proteome. Notably, with the search against both CL Brener and Sylvio X10/1 strains database, several drug

target candidates studied by our group, as leucyl aminopeptidase (LAP), oligopeptidase B (OPB), prolyl oligopeptidase (POP), methylthioadenosine phosphorylase (MTAF) and cathepsin B were detected. In relation to multigenic families of *T. cruzi*, the third search resulted in the identification of a large amount of members of families of trans-sialidases (TS), mucin-associated surface proteins (MASP), mucins, retrotransposon hot spot (RHS) and dispersed gene family 1 (DGF-1) proteins and 63 kDa glycoproteins (gp63). Therefore, we conclude that the search against both CL Brener and Sylvio X10/1 strains of *T. cruzi* was the most satisfactory. This proteome represent a significant advance in understanding the set of proteins expressed in *T. cruzi* and also an important step for elucidation of genome sequence of *T. cruzi* G strain.

Keywords: Chagas disease, *Trypanosoma cruzi*, G strain, epimastigotes, proteome, mass spectrometry.

LISTA DE FIGURAS

| | |
|--|----|
| Figura 1. Ciclo de vida do <i>T. cruzi</i> | 25 |
| Figura 2. Distribuição geográfica aproximada dos DTUs de <i>T. cruzi</i> no continente americano com relação aos ciclos doméstico e silvestre..... | 36 |
| Figura 3. Princípio do algoritmo do programa Sequest..... | 49 |
| Figura 4. Etapas da análise dos dados gerados por espectrometria de massas pelo programa PEAKS..... | 50 |
| Figura 5. Esquema de um espectrômetro de massas híbrido, um LTQ-Orbitrap..... | 52 |
| Figura 6. Estrutura interna do analisador de massas Orbitrap..... | 52 |
| Figura 7. Curva de crescimento de formas epimastigotas da cepa G de <i>T. cruzi</i> obtida ao longo de 7 dias..... | 61 |
| Figura 8. Análise em gel de agarose 1% da amplificação por PCR dos genes <i>24Sa</i> e <i>SL</i> a partir de DNA genômico das cepas G e Y de <i>T. cruzi</i> | 62 |
| Figura 9. Extrato total de proteínas de formas epimastigotas da cepa G submetidas a SDS-PAGE 12%..... | 63 |
| | |
| I. Busca 1: banco de dados de <i>Trypanosoma</i> spp. e outros tripanossomatídeos | |
| Figura 10. Relação entre o número de peptídeos por identificação e a porcentagem de identificações obtidas..... | 66 |
| Figura 11. Gráficos de setores para cada categoria de GO: (A) componente celular, (B) processo biológico e (C) função molecular de acordo com o programa Blast2GO..... | 70 |
| Figura 12. Diagrama da via de recuperação/salvação das purinas predita em <i>Trypanosoma</i> spp., com base nas enzimas descritas para <i>Leishmania</i> spp..... | 77 |
| Figura 13. Esquema simplificado da via de síntese <i>de novo</i> das pirimidinas..... | 79 |
| Figura 14. Predição de proteínas secretadas pelas vias clássica (SignalP) e não clássica (SecretomeP) e não anotadas como secretadas (nenhum)..... | 84 |
| Figura 15. Diagrama de Venn correlacionado as proteínas com preditas hélices transmembrânicas (TMHMM) e as preditas a serem secretadas pelas vias clássica (SignalP) e não-clássica (SecretomeP)..... | 85 |

II. Busca 2: banco de dados da cepa CL Brener de *T. cruzi*

| | |
|--|-----|
| Figura 16. Relação entre o número de peptídeos por identificação e a porcentagem de identificações obtidas..... | 87 |
| Figura 17. Gráficos de setores para cada categoria de GO: (A) componente celular, (B) processo biológico e (C) função molecular de acordo com o programa Blast2GO..... | 89 |
| Figura 18. Diagrama da via de recuperação/salvação das purinas predita em <i>Trypanosoma</i> spp., com base nas enzimas descritas para <i>Leishmania</i> spp..... | 98 |
| Figura 19. Representação esquemática da possível organização do ciclo do ácido cítrico em <i>T. cruzi</i> | 100 |
| Figura 20. Representação esquemática da via de biossíntese de aminoacil-tRNA gerada pelo programa Blast2GO..... | 102 |
| Figura 21. Esquema representativo da via de metabolismo de piruvato gerado pelo programa Blast2GO..... | 104 |
| Figura 22. Representação esquemática da glicólise/gliconeogênese gerada pelo Blast2GO..... | 107 |
| Figura 23. Predição de proteínas secretadas pelas vias clássica (SignalP) e não clássica (SecretomeP) e não anotadas como secretadas (nenhum)..... | 111 |
| Figura 24. Diagrama de Venn correlacionado as proteínas com preditas hélices transmembrânicas (TMHMM) e as preditas a serem secretadas pelas vias clássica (SignalP) e não-clássica (SecretomeP)..... | 113 |

III. Busca 3: banco de dados das cepas CL Brener e Sylvio X10/1 de *T. cruzi*

| | |
|---|-----|
| Figura 25. Diagrama de Venn correlacionado as proteínas identificadas nas três buscas realizadas..... | 116 |
| Figura 26. Relação entre o número de peptídeos por identificação e a porcentagem de identificações obtidas..... | 117 |
| Figura 27. Gráficos de setores para cada categoria de GO: (A) componente celular, (B) processo biológico e (C) função molecular de acordo com o programa Blast2GO..... | 118 |
| Figura 28. Diagrama da via de recuperação/salvação das purinas predita em <i>Trypanosoma</i> spp., com base nas enzimas descritas para <i>Leishmania</i> spp..... | 126 |

| | |
|--|-----|
| Figura 29. Representação esquemática da possível organização do ciclo do ácido cítrico em <i>T. cruzi</i> | 127 |
| Figura 30. Representação esquemática da via de síntese e de metabolismo de metiltioadenosina (MTA)..... | 129 |
| Figura 31. Representação esquemática da via de biossíntese de aminoacil-tRNA gerada pelo programa Blast2GO..... | 130 |
| Figura 32. Predição de proteínas secretadas pelas vias clássica (SignalP) e não clássica (SecretomeP) e não anotadas como secretadas (nenhum)..... | 135 |
| Figura 33. Diagrama de Venn correlacionado as proteínas com preditas hélices transmembrânicas (TMHMM) e as preditas a serem secretadas pelas vias clássica (SignalP) e não-clássica (SecretomeP)..... | 137 |

LISTA DE TABELAS

I. Busca 1: banco de dados de *Trypanosoma* spp. e outros tripanossomatídeos

| | |
|--|----|
| Tabela 1. Relação entre o organismo e o número de proteínas identificadas através de busca dos espectros de massas MS/MS por meio de busca no banco de dados de <i>Trypanosoma</i> spp. e outros tripanossomatídeos usando o algoritmo do programa Sequest..... | 67 |
| Tabela 2. Abundância de anotações e de atividades enzimáticas para cada uma das seis classes de <i>Enzyme Commission</i> (EC)..... | 72 |
| Tabela 3. Proteínas anotadas para a atividade de peptidase (EC 3.4) distribuídas em cinco sub-subclasses de EC..... | 73 |
| Tabela 4. As dez vias metabólicas mais abundantes em número de anotações e os respectivos números de atividades enzimáticas..... | 74 |
| Tabela 5. Relação entre as vias metabólicas mais representativas em número de proteínas anotadas e as respectivas atividades enzimáticas..... | 81 |

II. Busca 2: banco de dados da cepa CL Brener de *T. cruzi*

| | |
|--|-----|
| Tabela 6. Abundância de anotações e de atividades enzimáticas para cada uma das seis classes de EC..... | 91 |
| Tabela 7. Proteínas anotadas para a atividade de peptidase (EC 3.4) distribuídas em cinco sub-subclasses de EC..... | 94 |
| Tabela 8. As dez vias metabólicas mais abundantes em número de anotações e os respectivos números de atividades enzimáticas..... | 95 |
| Tabela 9. Relação entre as vias metabólicas mais representativas em número de proteínas anotadas e as respectivas atividades enzimáticas..... | 109 |

III. Busca 3: banco de dados das cepas CL Brener e Sylvio X10/1 de *T. cruzi*

| | |
|--|-----|
| Tabela 10. Abundância de anotações e de atividades enzimáticas para cada uma das seis classes de EC..... | 119 |
| Tabela 11. Proteínas anotadas para a atividade de peptidase (EC 3.4) distribuídas em oito sub-subclasses de EC..... | 122 |
| Tabela 12. As dez vias metabólicas mais abundantes em número de anotações e os respectivos números de atividades enzimáticas..... | 124 |

| | |
|---|-----|
| Tabela 13. Relação entre as vias metabólicas mais representativas em número de proteínas anotadas e as respectivas atividades enzimáticas..... | 132 |
|---|-----|

IV. Famílias multigênicas de *T. cruzi* nas buscas realizadas

| | |
|---|-----|
| Tabela 14: Proteínas identificadas das famílias multigênicas TS, MASP, mucinas, proteínas RHS, proteínas DGF-1 e gp63 obtidas na cepa G de <i>T. cruzi</i> pelas três buscas realizadas..... | 139 |
|---|-----|

| | |
|---|-----|
| Tabela 15. Comparação entre o número de proteínas identificadas neste e em outros proteomas totais previamente publicados de formas epimastigotas de <i>T. cruzi</i> | 146 |
|---|-----|

APÊNDICE A. Tabelas suplementares da busca 1

| | |
|---|-----|
| Tabela suplementar 1. Proteínas putativas preditas para secreção por via clássica pelo programa SignalP..... | 172 |
|---|-----|

| | |
|--|-----|
| Tabela suplementar 2. Proteínas putativas preditas para secreção por via não clássica pelo programa SecretomeP..... | 173 |
|--|-----|

APÊNDICE B. Tabelas suplementares da busca 2

| | |
|---|-----|
| Tabela suplementar 3. Proteínas putativas preditas para secreção por via clássica pelo programa SignalP..... | 178 |
|---|-----|

| | |
|--|-----|
| Tabela suplementar 4. Proteínas putativas preditas para secreção por via não clássica pelo programa SecretomeP..... | 179 |
|--|-----|

APÊNDICE C. Tabelas suplementares da busca 3

| | |
|---|-----|
| Tabela suplementar 5. Proteínas putativas preditas para secreção por via clássica pelo programa SignalP..... | 185 |
|---|-----|

| | |
|--|-----|
| Tabela suplementar 6. Proteínas putativas preditas para secreção por via não clássica pelo programa SecretomeP..... | 188 |
|--|-----|

LISTA DE ABREVIATURAS E SIGLAS

| | |
|------------|--|
| 2-DE | eletroforese bidimensional |
| AAA ATPase | <i>extended ATPase associated with various cellular activities</i> |
| ACN | acetonitrila |
| DDA | <i>data-dependent acquisition</i> |
| DGF-1 | <i>dispersed gene family 1</i> |
| DNA | ácido desoxirribonucleico |
| DTU | <i>discrete typing unit</i> |
| EC | <i>enzyme commission</i> |
| EDTA | ácido etilenodiamino tetracético |
| ESI | <i>electrospray ionization</i> |
| FDR | <i>false discovery rate</i> |
| GO | <i>gene ontology</i> |
| gp63 | glicoproteína de 63 kDa |
| GPI | glicosil fosfatidil inositol |
| HGPRT | hipoxantina-guanina fosforribosil transferase |
| HPLC | <i>high-performance liquid chromatography</i> |
| HSP | proteína de choque térmico |
| IDH | isocitrato desidrogenase |
| kDNA | DNA do cinetoplasto |
| KEGG | <i>Kyoto encyclopedia of genes and genomes</i> |
| LAP | leucil-aminopeptidase |
| LC-MS/MS | cromatografia líquida acoplada a MS/MS |
| LIT | <i>liver infusion tryptose</i> |
| LTQ | <i>linear trap quadrupole</i> |
| MALDI-TOF | <i>matrix-assisted laser desorption/ionization-time of flight</i> |
| MASP | <i>mucin-associated surface protein</i> |
| mRNA | RNA mensageiro |
| MS | espectrometria de massas |
| MS/MS | espectrometria de massas <i>in tandem</i> |
| MTAF | metiltioadenosina fosforilase |
| OMS | Organização Mundial de Saúde |

| | |
|----------|---|
| OPB | oligopeptidase B |
| ORF | fase aberta de leitura |
| PCR | reação em cadeia da polimerase |
| PGC | grupo de genes policistrônicos |
| POP | prolil-oligopeptidase |
| PRPP | fosforribosil pirofosfato |
| PTM | modificação pós-traducional |
| RHS | <i>retrotransposon hot spot</i> |
| RNA | ácido ribonucleico |
| rRNA | RNA ribossomal |
| SDS-PAGE | eletroforese em gel de poliacrilamida contendo dodecil sulfato de sódio |
| SL | <i>spliced leader</i> |
| spp. | espécies |
| TcMUC | <i>T. cruzi mucin</i> |
| tRNA | RNA de transferência |
| TS | trans-sialidase |
| VSG | glicoproteína de superfície variante |

SUMÁRIO

| | |
|---|-----------|
| INTRODUÇÃO | 21 |
| 1. Doença de Chagas: aspectos gerais | 21 |
| 1.1. Descoberta da doença de Chagas | 21 |
| 1.2. Aspectos epidemiológicos | 22 |
| 1.3. <i>Trypanosoma cruzi</i> | 23 |
| 1.4. Insetos vetores e hospedeiros vertebrados do <i>T. cruzi</i> | 26 |
| 1.5. Formas de transmissão do <i>T. cruzi</i> para o homem | 29 |
| 1.6. Fases e manifestações clínicas | 30 |
| 1.7. Diagnóstico laboratorial e tratamento | 31 |
| 1.8. Controle e prevenção | 34 |
| 2. Nomenclatura intraespecífica de <i>T. cruzi</i> | 36 |
| 2.1. Cepa G de <i>T. cruzi</i> | 37 |
| 3. Mecanismos moleculares da infecção por via oral pelo <i>T. cruzi</i> | 37 |
| 3.1. gp82 | 38 |
| 3.2. gp30 | 38 |
| 3.3. gp90 | 39 |
| 3.4. gp35/50 | 39 |
| 3.5. Cepa G e suas particularidades na infecção | 39 |
| 4. Organização genômica de <i>T. cruzi</i> | 41 |
| 5. Regulação da expressão gênica em <i>T. cruzi</i> | 43 |
| 6. Proteômica | 43 |
| 6.1. Proteomas totais de <i>T. cruzi</i> | 45 |
| 6.2. Subproteoma da cepa G | 46 |
| 6.3. Espectrometria de massas | 47 |
| 6.3.1. LTQ-Orbitrap | 51 |
| JUSTIFICATIVA | 53 |
| OBJETIVO | 54 |
| MATERIAIS E MÉTODOS | 55 |
| Cultura <i>in vitro</i> de epimastigotas da cepa G de <i>T. cruzi</i> | 55 |
| Curva de crescimento de epimastigotas da cepa G de <i>T. cruzi</i> | 55 |
| Extração de DNA genômico de epimastigotas da cepa G de <i>T. cruzi</i> | 55 |
| Genotipagem da cepa G de <i>T. cruzi</i> : amplificação por PCR do rDNA 24S α e do gene <i>SL</i> | 56 |
| Preparação do extrato de proteínas de epimastigotas da cepa G de <i>T. cruzi</i> ... | 56 |
| Processamento do extrato de proteínas | 57 |
| LC-MS/MS | 58 |
| Análise de dados | 58 |
| RESULTADOS E DISCUSSÃO | 61 |
| 1. Curva de crescimento de epimastigotas da cepa G de <i>T. cruzi</i> | 61 |
| 2. Genotipagem da cepa G de <i>T. cruzi</i> : amplificação por PCR do rDNA 24S α e do gene <i>SL</i> | 61 |
| 3. Preparação do extrato de proteínas expressas por epimastigotas da cepa G de <i>T. cruzi</i> | 62 |
| I. Busca 1: banco de dados de <i>Trypanosoma</i> spp. e outros tripanossomatídeos | 65 |
| 1. Classificação de <i>Gene Ontology</i> (GO) | 67 |

| | |
|---|------------|
| 2. Predição de atividades enzimáticas | 71 |
| 2.1. Peptidases (EC 3.4) putativas | 72 |
| 2.2. Predição de enzimas envolvidas em vias metabólicas..... | 73 |
| 2.2.1. Metabolismo de purinas | 74 |
| 2.2.2. Metabolismo de pirimidinas..... | 78 |
| 2.2.3. Degradação de glicanas | 79 |
| 2.2.4. Metabolismo de esfingolipídeos..... | 79 |
| 2.2.5. Metabolismo de tiaminas (vitamina B ₁)..... | 80 |
| 3. Predição de proteínas secretadas..... | 83 |
| 3.1. Predição de proteínas integrais de membrana..... | 84 |
| II. Busca 2: banco de dados da cepa CL Brener de <i>T. cruzi</i>..... | 86 |
| 1. Classificação de <i>Gene Ontology</i> (GO)..... | 88 |
| 2. Predição de atividades enzimáticas | 90 |
| 2.1 Peptidases (EC 3.4) putativas | 91 |
| 2.2. Predição de enzimas envolvidas em vias metabólicas..... | 95 |
| 2.2.1. Metabolismo de purinas | 96 |
| 2.2.2. Ciclo do ácido cítrico | 98 |
| 2.2.3. Biossíntese de aminoacil-tRNA..... | 100 |
| 2.2.4. Metabolismo de piruvato | 103 |
| 2.2.5. Glicólise/gliconeogênese | 104 |
| 2.2.6. Metabolismo de histidina e prolina..... | 108 |
| 3. Predição de proteínas secretadas..... | 111 |
| 3.1. Predição de proteínas integrais de membrana..... | 112 |
| III. Busca 3: banco de dados das cepas CL Brener e Sylvio X10/1 de <i>T. cruzi</i> | 114 |
| 1. Classificação de <i>Gene Ontology</i> (GO)..... | 117 |
| 2. Predição de atividades enzimáticas | 119 |
| 2.1 Peptidases (EC 3.4) putativas | 119 |
| 2.2. Predição de enzimas envolvidas em vias metabólicas..... | 124 |
| 2.2.1. Metabolismo de purinas | 125 |
| 2.2.2. Ciclo do ácido cítrico | 126 |
| 2.2.3. Metabolismo de cisteína e metionina..... | 128 |
| 2.2.4. Biossíntese de aminoacil-tRNA..... | 129 |
| 2.2.5. Metabolismo de arginina e prolina | 130 |
| 3. Predição de proteínas secretadas..... | 135 |
| 3.1. Predição de proteínas integrais de membrana..... | 136 |
| IV. Famílias multigênicas de <i>T. cruzi</i> nas buscas realizadas | 138 |
| 1. TS..... | 139 |
| 2. MASP | 141 |
| 3. Mucinas | 142 |
| 4. Proteínas RHS | 143 |
| 5. Proteínas DGF-1 | 143 |
| 6. gp63 | 144 |
| CONCLUSÃO | 146 |
| REFERÊNCIAS BIBLIOGRÁFICAS..... | 149 |
| APÊNDICE A. Tabelas suplementares da busca 1 | 172 |
| APÊNDICE B. Tabelas suplementares da busca 2 | 178 |
| APÊNDICE C. Tabelas suplementares da busca 3 | 185 |

INTRODUÇÃO

1. Doença de Chagas: aspectos gerais

1.1. Descoberta da doença de Chagas

A doença de Chagas, também conhecida como Tripanossomíase Americana, é causada pelo protozoário *Trypanosoma cruzi*. Foi descrita em 1909 pelo médico brasileiro Carlos Ribeiro Justiniano Chagas (1878-1934). Em 1907, Chagas foi designado por Oswaldo Cruz para implementar uma campanha contra a malária que estava comprometendo as obras de extensão na Estrada de Ferro Central do Brasil, no estado de Minas Gerais. Na atual cidade de Lassance, Carlos Chagas instalou um pequeno laboratório improvisado em um vagão de trem. Ao mesmo tempo, ele coletava e examinava espécies da fauna brasileira devido ao seu crescente interesse por entomologia e protozoologia. Em 1908, ele identificou, na amostra de sangue de um sagui (*Callithrix penicillata*), um protozoário do gênero *Trypanosoma* nomeado *Trypanosoma minasense*, não patogênico para o animal (Kropf e Sa, 2009).

Chagas soube que um inseto hematófago se escondia nas cabanas de pau-a-pique da região, popularmente conhecido como barbeiro devido ao hábito alimentar noturno de picar os rostos das pessoas enquanto elas dormiam. Ele encontrou formas de protozoários flagelados no intestino de alguns insetos e hipotetizou que poderiam ser parasitos naturais do inseto ou de vertebrados como o *T. minasense*. Os insetos foram enviados para o Instituto de Manguinhos, depois denominado Instituto Oswaldo Cruz, onde Oswaldo Cruz fez com que eles se alimentassem em saguis criados em laboratório. Depois, formas de tripanossomas foram encontradas em amostras de sangue de um dos saguis, que adoeceu. Chagas retornou ao Instituto e constatou que não se tratava do *T. minasense* e sim de uma nova espécie nomeada *Trypanosoma cruzi* em homenagem ao mestre Oswaldo Cruz (Kropf e Sa, 2009).

Pesquisando outros hospedeiros vertebrados do parasito, Chagas suspeitou que o homem poderia ser um deles e retornou a Lassance onde identificou o *T. cruzi* na amostra de sangue de um gato, um hospedeiro doméstico. Assim, a possibilidade de que o *T. cruzi* seria também um parasito capaz de infectar humanos era

considerável. Após realizar exames de sangue em moradores da região, Carlos Chagas identificou o *T. cruzi* no sangue de uma criança de dois anos de idade chamada Berenice que apresentava febre, anemia profunda e gânglios inchados. Assim, estava identificada uma nova tripanossomíase humana, a doença de Chagas (Chagas, 1909; Kropf e Sa, 2009).

É considerada uma descoberta inédita e excepcional devido à identificação do inseto transmissor, do agente etiológico e dos hospedeiros, à descrição anatomopatológica e do quadro clínico da doença humana pelo mesmo pesquisador. Além disso, a cronologia desse estudo foi incomum em razão da prévia identificação do vetor e do parasito e da posterior associação deles à doença (Lewinsohn, 1981).

Apesar de a descoberta da doença de Chagas ter ocorrido há pouco mais de cem anos, um estudo de paleoparasitologia sugere que a doença de Chagas é muito provavelmente uma doença antiga. DNA de *T. cruzi* foi detectado em corpos humanos mumificados encontrados no deserto do Atacama, no norte do Chile e sul do Peru, datados por radiocarbono dentro do intervalo de 9000 a 450 anos atrás. A explicação mais plausível para essa constatação sugere que, antes da ocupação dessa região por humanos, o ciclo silvestre da doença de Chagas compreendia um vasto número de hospedeiros mamíferos e insetos vetores. Com a colonização dessa área, os humanos teriam se inserido nesse ciclo silvestre e se tornado hospedeiros do parasito. Assim, o ciclo doméstico teria se desenvolvido gradualmente, atingindo seu estado atual (Aufderheide *et al.*, 2004).

1.2. Aspectos epidemiológicos

A Organização Mundial de Saúde (OMS) define as doenças tropicais negligenciadas como um grupo de doenças que dificultam o desenvolvimento econômico em populações pobres e não têm a atenção de indústrias farmacêuticas. Entre as doenças tropicais negligenciadas, encontram-se: doença de Chagas, dengue, tripanossomíase africana, leishmaniose e esquistossomose, entre outras (Who, 2013). Essas doenças ocorrem principalmente em áreas rurais ou urbanas pobres e geram um grande impacto sobre a saúde pública e, conseqüentemente, sobre as economias nacionais.

Estima-se que aproximadamente 8 milhões de pessoas estejam infectadas pelo *T. cruzi*. A doença de Chagas é principalmente distribuída em áreas endêmicas de 21 países da América Latina: Argentina, Belize, Bolívia, Brasil, Chile, Colômbia,

Costa Rica, El Salvador, Equador, Guatemala, Guiana, Guiana Francesa, Honduras, México, Nicarágua, Panamá, Paraguai, Peru, Suriname, Uruguai e Venezuela. Especificamente no Brasil há a estimativa de 2 a 3 milhões de pessoas infectadas por esse parasito (Martins-Melo *et al.*, 2012). Nas últimas décadas, a detecção de casos da doença de Chagas tem sido cada vez mais frequente nos Estados Unidos, no Canadá, em países da Europa, no Japão e na Austrália devido à migração populacional entre a América Latina e o resto do mundo (Who, 2012; 2013).

Observa-se um declínio do número estimado de pessoas infectadas pelo *T. cruzi*, do número de novos casos por ano e, conseqüentemente, do número de mortes por ano, da década de 80 até os dias atuais: em 1990, 30 milhões de pessoas infectadas, 700.000 novos casos por ano e 45.000 mortes; em 2000, 18 milhões de pessoas infectadas, 200.000 novos casos por ano e 21.000 mortes; em 2006, 15 milhões de pessoas infectadas, 41.200 novos casos por ano e 12.500 mortes (Who, 2007). A aparente redução do número de indivíduos infectados de 15 milhões em 2007 para 8 milhões em 2013 pode ser decorrente de iniciativas de controle da doença de Chagas na América Latina pela OMS e por autoridades regionais e nacionais ou apenas ser resultado de uma subnotificação dos casos dessa doença principalmente devido à existência de uma fase aguda geralmente assintomática, o que dificultaria o diagnóstico. Ainda assim, a estimativa mais recente de aproximadamente 8 milhões de pessoas infectadas em todo o mundo permanece altamente preocupante.

1.3. *Trypanosoma cruzi*

O parasito *T. cruzi* apresenta a seguinte classificação taxonômica: domínio Eukarya, reino Protista, sub-reino Protozoa, filo Sarcomastigophora, subfilo Mastigophora, classe Zoomastigophora, ordem Kinetoplastida, família Trypanosomatidae, gênero *Trypanosoma* (Rassi, Rassi e Marcondes De Rezende, 2012). A família Trypanosomatidae também inclui os protozoários *Trypanosoma brucei* e *Leishmania* spp., que causam a Tripanossomíase Africana e a Leishmaniose, respectivamente, e é caracterizada pela presença de um único flagelo e uma mitocôndria chamada cinetoplasto, que contém DNA mitocondrial com uma organização única (kDNA: *kinetoplast* DNA). Juntos, os parasitos *T. cruzi*, *T. brucei* e *Leishmania major* são denominados TriTryp.

O *T. cruzi* é predominantemente diploide (El-Sayed *et al.*, 2005) e replica-se

por fissão binária, um processo assexuado (Zingales *et al.*, 2012). Tendo a transmissão vetorial como foco inicial neste momento, o ciclo de vida do *T. cruzi* é complexo e digenético, compreendendo insetos hematófagos e hospedeiros vertebrados (Figura 1). Iniciando a descrição do ciclo pela ingestão do *T. cruzi* pelo inseto vetor, quando o barbeiro macho ou fêmea pica o hospedeiro vertebrado infectado, ingere formas tripomastigotas circulantes no sangue. No trato digestivo do triatomíneo, formas tripomastigotas não replicativas diferenciam-se em formas epimastigotas e algumas formas esferomastigotas. Epimastigotas são replicativas, dividem-se no intestino médio, migram para o intestino posterior e o reto do inseto onde se aderem e se transformam em formas tripomastigotas metacíclicas, dando-se o processo chamado de metaciclogênese. Tripomastigotas metacíclicas são liberadas nas fezes e na urina do triatomíneo infectado durante o repasto sanguíneo e podem ser inoculadas no hospedeiro vertebrado por injúria no sítio da picada ou pré-existente, ou pelas mucosas dos olhos e da boca. No vertebrado, tripomastigotas metacíclicas podem potencialmente infectar todos os tipos de células nucleadas; inicialmente, as células do sítio da picada são as células-alvo dos parasitos. Intracelularmente, tripomastigotas são liberadas dos vacúolos parasitóforos, completam a diferenciação em amastigotas (amastigogênese) que multiplicam-se no citoplasma. Amastigotas diferenciam-se em tripomastigotas e as células infectadas são lisadas por causa do grande número de parasitos, podendo invadir tecidos adjacentes e se difundir pelos vasos linfáticos e pelo sangue para tecidos distantes, principalmente células musculares estriadas cardíacas, estriadas esqueléticas e lisas, células ganglionares, onde ocorrem vários ciclos de replicação intracelular. Eventualmente, formas tripomastigotas podem ser ingeridas por um outro triatomíneo e transmitidas para um próximo hospedeiro vertebrado (Rassi, Rassi e Marin-Neto, 2010; Rassi, Rassi e Marcondes De Rezende, 2012).

1.4. Insetos vetores e hospedeiros vertebrados do *T. cruzi*

Os insetos hematófagos ou barbeiros que podem abrigar o *T. cruzi* e transmiti-lo ao homem e a outros hospedeiros vertebrados são também denominados triatomíneos (domínio Eukarya, reino Animalia, filo Arthropoda, classe Insecta, ordem Hemiptera, família Reduviidae, subfamília Triatominae).

A transmissão da doença de Chagas para o homem está associada a dois cenários epidemiológicos: peridoméstico e doméstico. Os barbeiros têm como habitats naturais os ambientes silvestres, tais como ninhos de gambás, tocas de tatu, abrigos em pedras e palmeiras. Porém, com a destruição desses habitats por desmatamento e colonização desenfreados, a disponibilidade de animais é reduzida, o que leva ao desequilíbrio entre as fontes de alimento e as populações de barbeiros. Isso faz com que esses insetos busquem alimento em regiões peridomiciliares que possuem criações de animais e, até mesmo, atinjam as áreas domiciliares. Nas áreas domiciliares, os triatomíneos tipicamente vivem em fendas de cabanas de pau-a-pique, onde se ocultam durante o dia e saem à noite em busca de alimento. Assim, os ciclos de transmissão do *T. cruzi* silvestre, peridoméstico e doméstico apresentam sobreposição e são interrelacionados (Argolo *et al.*, 2008).

O tempo médio de vida dos barbeiros é de dois anos. Como citado na descrição do ciclo de vida do *T. cruzi*, tanto o macho quanto a fêmea alimentam-se de sangue. Após o repasto sanguíneo, a fêmea adulta deposita de uma a duas centenas de ovos no seu abrigo. Cada ovo origina uma ninfa, que apresenta cinco estádios antes de atingir a forma adulta (Argolo *et al.*, 2008).

Atualmente, na subfamília Triatominae, um total de 142 espécies já foram descritas (Gurgel-Gonçalves *et al.*, 2012) e são potenciais hospedeiros do *T. cruzi*, embora relativamente poucas sejam vetores competentes para esse parasito (Lent e Wygodzinsky, 1979; Rassi, Rassi e Marin-Neto, 2010; Rassi, Rassi e Marcondes De Rezende, 2012). As espécies de vetores com maior importância na transmissão da doença de Chagas são *Triatoma infestans*, *Triatoma brasiliensis* e *Panstrongylus megistus* nos países do Cone Sul (Argentina, Bolívia, Brasil, Chile, Paraguai e Uruguai); *Rhodnius prolixus* e *Triatoma dimidiata* nos países do Pacto Andino (Colômbia, Equador, Peru e Venezuela) e em partes da América Central; *T. dimidiata* e *Triatoma barberi* no México. Historicamente, *T. infestans* tem sido o mais importante vetor domiciliar no sul da América do Sul, nas regiões endêmicas subamazônicas (Rassi, Rassi e Marin-Neto, 2010; Rassi, Rassi e Marcondes De

Rezende, 2012; Who, 2012), e teve sua erradicação domiciliar certificada em 2006 no Brasil (Silveira e Dias, 2011). No Brasil, 62 espécies de triatomíneos têm sido encontradas, das quais 39 são exclusivas do país. Os estados brasileiros que apresentam os maiores números de espécies de triatomíneos são Bahia (25 spp.) e Mato Grosso (18 spp.), seguidos por Pará, Tocantins e Minas Gerais (15 spp., cada). No que concerne aos biomas, o Cerrado e a Caatinga são os detentores da maior diversidade de espécies (Gurgel-Gonçalves *et al.*, 2012). Segundo levantamento entomológico de base, em 2007, as principais espécies de vetores do *T. cruzi* na rotina do programa de controle da doença de Chagas no Brasil eram *Triatoma sordida*, *Triatoma pseudomaculata*, *T. brasiliensis* e *P. megistus* (Silveira e Dias, 2011).

Antigamente, conceituava-se um reservatório de um parasito como um organismo selvagem quase inerte que abriga um parasito que causa uma doença humana. O conceito mais recente de reservatório considera o hospedeiro como um sistema vivo complexo, capaz de sofrer e aplicar pressão seletiva em relação aos parasitos, conforme as condições ambientais desses. No caso do *T. cruzi*, que teve o ciclo de vida ampliado a partir do ambiente silvestre para as áreas domiciliares, tanto animais selvagens quanto domésticos podem agir como hospedeiros desse parasito. Assim, um reservatório pode ser definido como uma espécie ou um conjunto de espécies, sejam hospedeiros selvagens ou domésticos, responsáveis pela manutenção a longo prazo de um dado parasito em um dado ambiente (Noireau, Diosque e Jansen, 2009). Além do ser humano, há uma grande variedade de animais que podem ser reservatórios do *T. cruzi*, como marsupiais, primatas, morcegos, quatis, guaxinins, raposas, tatus, roedores, cachorros, gatos, porcos, cabras.

Os marsupiais são considerados os mais antigos hospedeiros selvagens do *T. cruzi*, juntamente com tatus e tamanduás. Gambás do gênero *Didelphis* criados em laboratório e infectados por via subcutânea com *T. cruzi* apresentaram formas epimastigotas e tripomastigotas metacíclicas no lúmen das glândulas anais (Deane, Lenzi e Jansen, 1986), assim como tripomastigotas no sangue periférico (Deane, Lenzi e Jansen, 1984). Foi observado que a metaciclogênese nas glândulas anais ocorre independentemente de adesão às células epiteliais (Carreira *et al.*, 2001), uma característica que difere da metaciclogênese no trato digestório posterior do barbeiro. Em contrapartida, em gambás naturalmente infectados, a prevalência de

infecção nas glândulas anais é baixa (Steindel *et al.*, 1988; Fernandes *et al.*, 1989). Esses resultados sugerem que gambás *Didelphis* spp. podem atuar simultaneamente como reservatório e vetor do *T. cruzi*. Em termos evolutivos, possivelmente a transmissão do *T. cruzi* pelas excreções das glândulas anais dos gambás é um mecanismo anterior à aquisição dos hábitos hematofágicos pelos triatomíneos, o que também garantiria a perpetuação desse parasito na natureza (Noireau, Diosque e Jansen, 2009).

A ligação entre os ciclos silvestre e peridoméstico ou doméstico pode ser feita pelos animais domésticos como cachorros, gatos e de criações peridomésticas como porcos, cabras, por representarem uma alternativa alimentar para os triatomíneos em decorrência da degradação dos habitats silvestres.

Aves, répteis e anfíbios são refratários à infecção pelo *T. cruzi* e, portanto, não considerados reservatórios do parasito. Em aves, isso se deve ao efeito lítico mediado pelo sistema complemento sobre os parasitos, independentemente de anticorpos (Kierszenbaum, Gottlieb e Budzko, 1981). Apesar disso, as aves, principalmente as galinhas, exercem um papel indireto, porém significativo, na manutenção do *T. cruzi* nos ambientes selvagem e peridoméstico por representarem uma fonte alimentar para os triatomíneos. A frequência com que os triatomíneos infectados com *T. cruzi* se alimentam afeta a densidade populacional e a proporção dos diferentes estágios de desenvolvimento na população. Esses insetos hematofagos podem ficar sem se alimentar por vários meses antes de morrer e a perda de nutrientes no trato intestinal influencia na população do *T. cruzi*. Após jejum prolongado, a densidade da população de parasitos no intestino posterior e no reto de triatomíneos é fortemente reduzida, além do aumento da porcentagem de formas esferomastigotas em detrimento de formas epimastigotas e tripomastigotas. Essas formas esferomastigotas parecem ser induzidas em condições de estresse (Kollien e Schaub, 1998).

Portanto, as aves apesar de não abrigarem o *T. cruzi*, participam indiretamente na densidade e na composição populacionais no triatomíneo ao servirem como fonte de sangue. Ao mesmo tempo, galinheiros em áreas peridomésticas ou domésticas são atrativos para os barbeiros e podem agir como uma barreira à invasão domiciliar ou aumentar as chances de contato entre animais domésticos ou humanos e triatomíneos (Noireau, Diosque e Jansen, 2009).

1.5. Formas de transmissão do *T. cruzi* para o homem

As formas de transmissão do *T. cruzi* para o homem são múltiplas. Além da transmissão vetorial, esse parasito também pode ser contraído pelo contato direto com as mucosas, abrasões na pele, ingestão acidental, transfusões sanguíneas, transmissão vertical, transplante de órgãos, acidentes laboratoriais (Rassi, Rassi e Marcondes De Rezende, 2012; Who, 2012) e ainda deve ser considerada a possibilidade de transmissão sexual em humanos (Dias e Amato Neto, 2011) que foi recentemente descrita em camundongos (Silva, 2013).

A transmissão por intermédio de triatomíneos é considerada usual por grande parte dos autores e inclusive pela OMS e os mecanismos de transmissão não vetoriais são tidos como menos frequentes. Entretanto, Coura (2006) considera que a infecção por via oral proveniente da ingestão de alimentos e líquidos contaminados com *T. cruzi* tornou-se o mais importante modo de transmissão no Brasil, originando microepidemias e surtos da fase aguda da doença de Chagas em várias regiões do país. A contaminação de alimentos e bebidas com esse protozoário pode ocorrer durante o preparo e duas possibilidades são sugeridas: contato com excreções das glândulas anais de gambás ou fezes de barbeiros, ou barbeiros infectados inteiros que seriam triturados ou esmagados. A ingestão de carne crua de mamíferos silvestres contendo pseudocistos e formas tripomastigotas circulantes também é considerada uma potencial fonte de ingestão acidental do *T. cruzi*. No Brasil, surtos de fase aguda da doença de Chagas acabam ocorrendo com frequência. Em relação a alguns casos notificados, em 2005, ocorreu um surto em Santa Catarina associado à ingestão de caldo de cana contaminado (Steindel *et al.*, 2008). No Pará, no Amazonas e na Bahia, também ocorreram surtos, em 2006 e 2007, ocasionados pela ingestão de sucos de bacaba, açaí e caldo de cana contaminados e até pela possível contaminação de água inapropriadamente armazenada (Dias *et al.*, 2008; Nobrega *et al.*, 2009; Bastos *et al.*, 2010; Monteiro *et al.*, 2010). Em uma escola de Caracas, Venezuela, ocorreu um surto em 2007 devido à ingestão de suco de goiaba contaminado (Alarcon De Noya *et al.*, 2010). No estado do Amazonas, a ingestão de açaí contaminado levou a ocorrência de casos de fase aguda em 2010 (Souza-Lima Rde *et al.*, 2013). A severidade da doença proveniente da infecção por via oral é indiscutível, com alguns casos resultando em mortes (Coura *et al.*, 2002; Yoshida, 2008). Por isso, o cuidado no armazenamento e no preparo de alimentos e

bebidas em regiões endêmicas pode minimizar a contaminação com *T. cruzi* e a subsequente ocorrência de surtos da fase aguda da doença de Chagas.

A questão de qual modo de transmissão predomina atualmente é bastante controversa. É importante ressaltar que, na maioria dos casos de doença de Chagas, não fica evidente a forma de aquisição do parasito e, por isso, esses dados epidemiológicos podem não corresponder à realidade. Possivelmente, por suas características alarmantes de gerar sintomas severos e microepidemias e surtos em que muitas pessoas são infectadas conjuntamente e algumas delas morrerem rapidamente após o contágio, a infecção por via oral tenha sido destacada por alguns pesquisadores perante as outras formas de aquisição do parasito.

1.6. Fases e manifestações clínicas

A doença de Chagas pode ser caracterizada por duas fases: aguda e crônica. Alguns autores consideram três fases: aguda, indeterminada e crônica. No presente trabalho, a classificação considerada foi a primeira citada, em que as formas determinada e indeterminada são encaixadas dentro da fase crônica. Uma variedade de fatores não exclusivos pode determinar quais manifestações clínicas ocorreram, tais como: características genéticas do parasito e do hospedeiro, infecções mistas (causadas por mais de um clone de *T. cruzi*), fatores culturais e geográficos (Zingales *et al.*, 2012).

A fase inicial aguda é geralmente assintomática e é reconhecida em apenas 1-2% dos indivíduos infectados. Nesses casos, os sintomas são variáveis e incluem: febre prolongada, mal-estar, inchaço do fígado, baço e linfonodos, edema subcutâneo localizado ou generalizado. Na transmissão vetorial, existem sinais que indicam o sítio de entrada do *T. cruzi*: o chagoma, uma área endurecida com eritema e inchaço, e o sinal de Romaña, edema unilateral indolor de ambas as pálpebras e de tecidos periorbitais. Em 90% dos casos, as manifestações clínicas declinam espontaneamente após 4-8 semanas. Ocasionalmente, ocorrem mortes na fase aguda (5-10% dos casos sintomáticos) em decorrência de miocardite aguda e/ou meningoencefalite (Rassi, Rassi e Marin-Neto, 2010; Rassi, Rassi e Marcondes De Rezende, 2012; Who, 2012).

Na fase crônica da doença de Chagas, 60-70% dos indivíduos positivos para anticorpos IgG contra *T. cruzi* no soro são assintomáticos e são agrupados na forma indeterminada da fase crônica, enquanto que os outros 30-40% pertencem à forma

determinada da fase crônica e desenvolvem patologias cardíacas e/ou digestivas (megacólon e megaesôfago) severas e até lesões no sistema nervoso central, depois de 10-30 anos da infecção inicial. 2-3% dos indivíduos da forma indeterminada evoluem para manifestações sintomáticas descritas acima por mecanismos ainda desconhecidos (Rassi, Rassi e Marin-Neto, 2010; Rassi, Rassi e Marcondes De Rezende, 2012; Who, 2012). A progressão direta da fase aguda para a fase crônica determinada ocorre em alguns pacientes (5-10%) (Rassi, Rassi e Marin-Neto, 2010).

1.7. Diagnóstico laboratorial e tratamento

Atualmente, o diagnóstico laboratorial da fase aguda da doença de Chagas é baseado em testes parasitológicos, por observação microscópica do sangue fresco, hemocultura, extensões sanguíneas coradas ou xenodiagnóstico, devido à elevada parasitemia nessa fase. Na fase crônica, a parasitemia persiste, porém é muito reduzida em relação à fase aguda e dificilmente detectada. Há uma elevada produção de anticorpos IgG em resposta ao *T. cruzi* e o diagnóstico é baseado em técnicas sorológicas como ensaios de imunofluorescência, *enzyme-linked immunosorbent assay* (ELISA) e ensaios de hemaglutinação indireta. Dois resultados positivos em quaisquer dessas três técnicas são recomendados para um diagnóstico final. Os testes diagnósticos comercialmente disponíveis empregam extratos crus de *T. cruzi*, frações semipurificadas ou antígenos recombinantes. Outro teste diagnóstico é a detecção do material genético do parasito por PCR ou RT-PCR. Alguns autores consideram que a falta de padronização dos protocolos entre laboratórios e países, a potencial contaminação cruzada com DNA, os polimorfismos de sequência entre diferentes cepas do *T. cruzi* e os resultados falsos negativos poderiam tornar o diagnóstico por PCR incerto em ambas as fases da doença (Britto, 2009; Afonso, Ebell e Tarleton, 2012; Rassi, Rassi e Marcondes De Rezende, 2012; Who, 2012; Murcia *et al.*, 2013). Um estudo envolvendo laboratórios de 16 países foi realizado na tentativa de selecionar os protocolos de PCR de melhor performance para detecção de *T. cruzi* em amostras de sangue humano (Schijman *et al.*, 2011). A PCR pode ser um método auxiliar para o diagnóstico da doença quando a sorologia para *T. cruzi* é inconclusiva e para pacientes submetidos a transplantes ou imunossuprimidos. Deve-se ressaltar que um resultado negativo por testes

sorológicos e/ou por PCR não descarta a infecção, como demonstrado por Gilber *et al.* (2013).

Existem apenas duas drogas utilizadas no tratamento da doença de Chagas, nifurtimox e benzonidazol, introduzidas nos anos 60 e 70, respectivamente. O nifurtimox (3-metil-4(5'-nitrofurfurilideno-amino)tetrahidro-4H-1,4-tiazina-1,1-dióxido) foi introduzido sob o nome Lampit® pelo laboratório Bayer e, subsequentemente, retirado do mercado no Brasil, na Argentina, no Chile e no Uruguai. Hoje, esse medicamento é produzido e utilizado na América Central. O benzonidazol (N-benzil-2-nitroimidazol-acetamida) foi comercializado pela Roche sob os nomes comerciais de Rochagan® e Radanil®, que recentemente transferiu a patente para o Laboratório Farmacêutico do Estado de Pernambuco (LAFEPE). A distribuição do benzonidazol está restrita a clínicas especializadas para acompanhamento médico durante o tratamento (Coura, 2009).

Esses dois compostos nitro-heterocíclicos podem causar inúmeros e graves efeitos adversos clinicamente observados. Os mais frequentes efeitos adversos por administração do nifurtimox são: anorexia, perda de peso, alterações psíquicas, excitabilidade ou sono, manifestações digestivas como náusea, vômito, cólica intestinal, diarreia. Para o benzonidazol, foram observados sintomas de hipersensibilidade, dermatite, erupções cutâneas, edemas, dores articular e muscular, depressão da medula óssea, púrpura trombocitopênica, agranulocitose, polineuropatias (Coura e De Castro, 2002). Segundo a OMS, a cura para a doença de Chagas é caracterizada pela soroconversão (não detecção de anticorpos anti-*T. cruzi* por testes sorológicos) e parasitemia negativa. Em estudos clínicos, esses fármacos mostraram-se efetivos em recém-nascidos com infecção congênita e durante a fase aguda com 99% de cura e até 80% de cura, respectivamente. O tratamento também é indicado em casos de reativação da infecção que pode ocorrer em indivíduos cronicamente infectados e que se tornam imunossuprimidos, como pessoas coinfectadas com HIV ou que estão sendo tratadas com drogas imunossupressoras. Entretanto, durante a fase crônica, a porcentagem de cura é de, aproximadamente, 20% (Braz, Amato Neto e Okay, 2008; Coura, 2009; Who, 2012).

Foi demonstrado em vários estudos que o benzonidazol e o nifurtimox apresentam atividade mutagênica (Ferreira e Ferreira, 1986; Gorla, 1987; Gorla *et al.*, 1988; 1989). Além disso, exibiram efeitos carcinogênicos em coelhos e camundongos (Teixeira *et al.*, 1990; Teixeira, Calixto e Teixeira, 1994). Não há

relatos em humanos de tumorigênese após tratamento com os fármacos em questão.

Dentro da comunidade científica, a questão sobre o emprego amplo desses medicamentos em pacientes adultos enquadrados nas formas indeterminada e determinada da fase crônica gera opiniões divergentes. Por exemplo, Rassi *et al.* (2010; 2012) consideram que crianças e adultos de 18 a 50 anos na fase crônica da doença, com exceção daqueles que apresentam doença cardíaca avançada, insuficiência renal ou hepática, megaesôfago e grávidas, devam ser tratados com esses medicamentos. Além disso, sugerem a adoção de estratégias no uso do benzonidazol relacionadas à posologia e ao tratamento concomitante com corticosteroides no caso de aparecimento de dermatites. Issa e Bochi (2010) discordam da ideia de tratamento como terapia padrão para pacientes na fase crônica devido aos sérios efeitos adversos, aos efeitos mutagênico e carcinogênico, à baixa eficácia, mostrados na literatura. Para fornecer uma base mais sólida sobre a eficácia do tratamento tripanocida em pacientes na fase crônica com patologias cardíacas da América Latina, um largo ensaio clínico internacional, multicêntrico, controlado e randomizado denominado *The BENznidazole Evaluation For Interrupting Trypanosomiasis* (BENEFIT) está sendo realizado (Marin-Neto *et al.*, 2008).

O tratamento prolongado com os nitroderivados que exige a administração de 2-3 doses diárias por 60 dias e a ocorrência de reações adversas podem levar à descontinuação do tratamento sem consentimento médico e ao subsequente agravamento das manifestações clínicas. O sucesso do tratamento também está relacionado à fase da infecção em que o paciente se encontra, à idade do indivíduo e à cepa do *T. cruzi* responsável pela infecção (Coura, 2009; Coura e Borges-Pereira, 2011; Who, 2012).

Em razão de a doença de Chagas acometer cerca de 8 milhões de pessoas em todo o mundo, os nitroderivados apresentarem alta toxicidade e baixa eficácia durante a fase crônica, a resistência de algumas cepas do *T. cruzi* às drogas e não existirem vacinas, esperavam-se grandes investimentos das indústrias farmacêuticas no desenvolvimento de novos medicamentos eficazes, seguros e idealmente ativos para ambas as fases da doença. Lamentavelmente, não é isso que se observa após cem anos da descoberta da doença de Chagas. O investimento para a síntese de novos fármacos pelos centros de pesquisa e pelas

indústrias farmacêuticas seria desproporcionalmente maior do que o retorno financeiro gerado pelo mercado alvo constituído, em grande parte, por pessoas desfavorecidas economicamente.

A iniciativa para pesquisa de novas drogas tem se dado por associações entre laboratórios de pesquisa, órgãos de saúde e pouquíssimos laboratórios farmacêuticos. Por exemplo, *Drugs for Neglected Diseases initiative* (DNDi) lançou, em 2009, a Plataforma de Pesquisa Clínica em Doença de Chagas com mais de 60 organizações, objetivando a disponibilização de novos medicamentos para o tratamento da infecção pelo *T. cruzi*. Estão em andamento os estudos clínicos de fase II para avaliar a segurança e a eficácia das diferentes doses do E1224, uma pró-droga do ravuconazol, em parceria com o laboratório farmacêutico japonês Eisai (Dndi, 2009). Outro composto, o posaconazol, está na fase II de testes clínicos e a pesquisa tem a participação do Hospital Vall d'Hebron de Barcelona (CHAGASAZOL, <http://clinicaltrials.gov/show/NCT01162967>) e do laboratório farmacêutico Merck & Co (STOP CHAGAS, <http://clinicaltrials.gov/show/NCT01377480>). Um inibidor da cruzipaina, uma protease de *T. cruzi*, está na fase de testes pré-clínicos (Who, 2012).

1.8. Controle e prevenção

Associada ao diagnóstico precoce e ao tratamento das pessoas infectadas pelo *T. cruzi*, a execução de estratégias de controle e prevenção da doença de Chagas em relação às formas de transmissão vetorial e não vetoriais é necessária para a redução drástica que se espera no número de novos casos e, conseqüentemente, na quantidade de pessoas infectadas em todo o mundo.

Os triatomíneos podem atingir os habitats doméstico e/ou peridoméstico e formar colônias substanciais ou não. No caso de formação e estabelecimento de colônias, a eliminação do vetor pode ser realizada pela pulverização de inseticidas nas casas e nas áreas adjacentes. Melhorias habitacionais ajudariam a prevenir a instalação dos triatomíneos nos ambientes doméstico e peridoméstico, assim como a utilização de mosquiteiros para impedir a picada pelo inseto hematófago durante a noite. Se a espécie do triatomíneo não coloniza esses habitats, o papel do controle do vetor é limitado e o recurso primário é a detecção adequada e o tratamento de quaisquer novos casos de infecção que possam surgir (Who, 2012; 2013).

Em 1991, um acordo internacional assinado entre Brasil, Argentina, Bolívia, Chile, Paraguai, Uruguai (e mais tarde Peru) teve como objetivo o controle da transmissão vetorial por eliminação do principal vetor domiciliar, *T. infestans*, e também o controle sistemático nos bancos de sangue (Who, 2002). Após 10 anos, tal iniciativa resultou em uma redução média de 94% na incidência da doença nos países do Cone Sul (Who, 2002) principalmente por causa da eliminação das populações domésticas de *T. infestans* em vastas áreas, embora ainda existam populações residuais em áreas peridomésticas como na região do Chaco do norte da Argentina e no sul da Bolívia (Who, 2012), o que tem sido associado à alta resistência do barbeiro aos inseticidas piretroides empregados nessas regiões (Noireau, Diosque e Jansen, 2009). Em 2006, a Comissão Intergovernamental dos Países do Cone Sul e a Organização Pan-Americana da Saúde (OPAS)/OMS certificaram formalmente a interrupção da transmissão vetorial do *T. cruzi* pelo *T. infestans* no Brasil e consideraram essa forma de transmissão como improvável em razão dos baixos índices de infestação e densidade das populações residuais (Silveira e Dias, 2011). Porém foi relatada recentemente a ocorrência dessas populações na Bahia e no Rio Grande do Sul (Gurgel-Gonçalves *et al.*, 2012). Uma vigilância entomológica constante é importante para a prevenção da reinfestação pelo *T. infestans* e da formação de colônias por outras espécies de triatomíneos em ambientes domésticos e peridomésticos.

Ainda não existe vacina efetiva para a doença de Chagas. No caso de infecção por via oral, boas práticas de higiene na preparação, no transporte, na armazenagem e no consumo de comidas ou bebidas podem auxiliar na prevenção da contaminação pelo parasito. Em relação à transmissão transfusional, a triagem rigorosa do sangue doado e hemoderivados e a atuação mais efetiva da vigilância sanitária sobre os serviços de hemoterapia públicos e privados contribuem para o aumento da segurança transfusional e, conseqüentemente, para eliminação da transmissão sanguínea do parasito. A triagem de órgãos, tecidos e células de doadores para transplante e também de neonatos e filhos de mães infectadas pelo *T. cruzi* são outras abordagens recomendadas pela OMS para a prevenção da doença de Chagas (Who, 2013).

2. Nomenclatura intraespecífica de *T. cruzi*

Em 2009, foi estabelecido um novo consenso para nomenclatura intraespecífica de *T. cruzi* no *Second Satellite Meeting* em Búzios (Zingales *et al.*, 2009). Essa nomenclatura foi baseada em genotipagem multilocus e tem o objetivo de favorecer a comunicação entre a comunidade científica envolvida na pesquisa da doença de Chagas, afinal, *T. cruzi* compreende grupos de populações geneticamente distintas em termos de conteúdo de DNA, crescimento, infectividade, perfis isoenzimáticos. Isolados e cepas conhecidas devem ser atribuídos a um dos seis grupos genéticos ou *discrete typing units* (DTUs), TcI a TcVI, e DTUs são definidos como grupos de estoques que são geneticamente mais relacionados uns aos outros do que a qualquer outro estoque e que são identificáveis por marcadores genéticos, moleculares ou imunológicos comuns (Tibayrenc, 1998). A distribuição geográfica aproximada dos seis DTUs de *T. cruzi* nas Américas com relação aos ciclos de transmissão doméstico e silvestre é ilustrada na Figura 2.

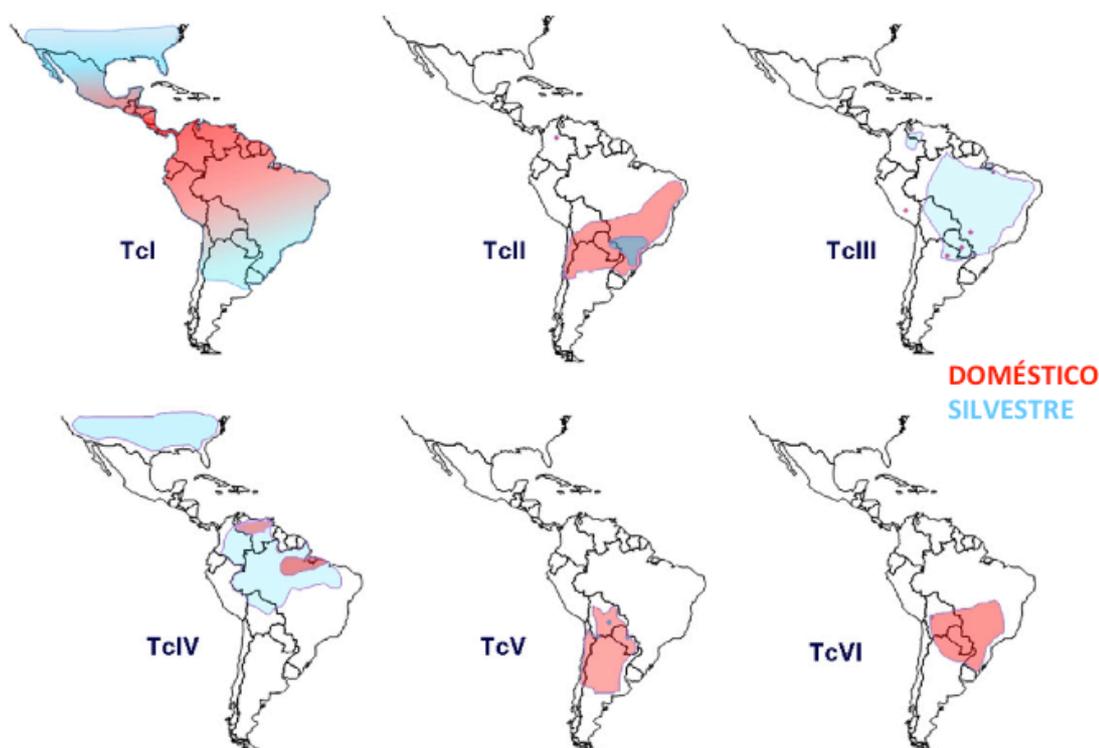


Figura 2. Distribuição geográfica aproximada dos DTUs de *T. cruzi* no continente americano com relação aos ciclos doméstico e silvestre. Modificado de Zingales *et al.* (2012).

O DTU TcI é o mais abundante e amplamente disperso de todos os DTUs de *T. cruzi* no continente americano. Uma distribuição ampla de triatomíneos que abrigam parasitos pertencentes a TcI pode estar associada à ocorrência desses em ambos os tipos de ciclos de transmissão, doméstico e silvestre. Infecção humana por TcI ocorre essencialmente na América Central e no norte da América do Sul. Em relação à diversidade genética, cepas pertencentes a TcI mostram substancial homozigose alélica, assim como TcII, TcIII e TcIV, diferentemente daquelas enquadradas nos DTUs TcV e TcVI, que apresentam uma pronunciada heterozigose e, portanto, são cepas híbridas (Zingales *et al.*, 2012). TcI tem sido fortemente relacionado à infecção em marsupiais (Steindel *et al.*, 2008; Monteiro *et al.*, 2012).

2.1. Cepa G de *T. cruzi*

A cepa G de *T. cruzi*, estudada no presente trabalho, pertence ao DTU TcI e, como Sylvio X10/1, apresenta um genoma menor do que CL Brener (Souza *et al.*, 2011). A cepa G foi isolada de um gambá da Amazônia brasileira (Yoshida, 1983). A potencial transmissão do *T. cruzi* por via oral através da excreção de gambás tornaria a cepa G também potencialmente transmissível para humanos, mas nenhum caso de infecção em humanos foi relatado. Em apenas poucos casos e microepidemias de fase aguda da doença de Chagas associados com transmissão por via oral, o perfil genético de *T. cruzi* foi determinado e muitos DTUs têm sido associados, principalmente TcI (Steindel *et al.*, 2008; Marcili *et al.*, 2009; Valente *et al.*, 2009; Alarcon De Noya *et al.*, 2010; Andrade *et al.*, 2011; Monteiro *et al.*, 2012). Significativamente, a cepa G tem sido amplamente usada para elucidar os mecanismos moleculares da infecção por via oral, entre outras cepas.

3. Mecanismos moleculares da infecção por via oral pelo *T. cruzi*

Experimentos de infecção por via oral em camundongos mostraram que formas tripomastigotas metacíclicas podem invadir e se replicar como formas amastigotas no epitélio da mucosa gástrica, mas não há evidência de invasão da orofaringe ou do esôfago de animais. A mucosa gástrica parece ser a única via de entrada do parasito oralmente transmitido para a infecção sistêmica (Hoft, 1996; Hoft *et al.*, 1996). O ambiente ácido no estômago não diminui a infectividade do parasito, pois proteínas de superfície protegem o parasito da degradação pelo suco gástrico,

principalmente as glicoproteínas de superfície mucinas-símiles do parasito (Mortara *et al.*, 1992). Pelo contrário, existem cepas que se tornam ainda mais infectivas após o contato com o suco gástrico (Cortez *et al.*, 2006; Covarrubias *et al.*, 2007). Para invadir as células epiteliais gástricas, as formas tripomastigotas metacíclicas devem primeiramente se ligar à camada de mucina gástrica, depois atravessá-la e, então, se ligar às células epiteliais gástricas. Nesse contexto, várias moléculas de superfície do *T. cruzi* já tiveram influência demonstrada nos mecanismos moleculares da infecção por via oral e, conjuntamente, são determinantes para a capacidade de invasão de células mamíferas por tripomastigotas metacíclicos das diferentes cepas de *T. cruzi*.

3.1. gp82

Do grupo mais para o menos abrangente, a glicoproteína de superfície gp82 (assim como gp90, TSA-1, SA85, ASP-2) é um membro da superfamília de trans-sialidases, da família de trans-sialidases-símiles, do grupo II composto por glicoproteínas de superfície gp85. A proteína gp82, altamente expressa em tripomastigotas metacíclicas, liga-se ao seu receptor na célula hospedeira e inicia uma cascata de sinalização que leva à ativação sequencial de proteína tirosinoquinase e de fosfolipase C. Consequentemente, ocorre a geração de IP₃ (inositol trifosfato) que promove a mobilização de Ca²⁺ necessária para a internalização do parasito (Yoshida *et al.*, 2000).

No contexto da infecção por via oral, gp82 é resistente à degradação por pepsina em pH ácido (Cortez *et al.*, 2006) e tem um papel na invasão das células epiteliais da mucosa gástrica. gp82 medeia ligação à camada de mucina gástrica (Neira *et al.*, 2003) e posterior ligação às células epiteliais que induz uma resposta de Ca²⁺ bidirecional requerida para a invasão (Ruiz *et al.*, 1998).

3.2. gp30

Algumas cepas de *T. cruzi* são deficientes em gp82 e as formas tripomastigotas metacíclicas pertencentes a elas expressam a glicoproteína de superfície gp30, que também é capaz de induzir a sinalização por Ca²⁺ (Cortez *et al.*, 2003). *In vitro*, a expressão de gp30 apenas não interferiu na capacidade de invasão de células epiteliais por formas tripomastigotas metacíclicas de cepas deficientes em gp82. Entretanto, gp30 tem uma afinidade mais baixa à mucina gástrica do que

gp82, o que resultou em baixos níveis de parasitemia na infecção por via oral de camundongos (Cortez *et al.*, 2003; Cortez *et al.*, 2012). Portanto, a capacidade de invasão reduzida *in vivo* após infecção por via oral pelas cepas deficientes em gp82 deve-se, ao menos em parte, à baixa capacidade de adesão de gp30 à camada de mucina gástrica e à consequente dificuldade desses parasitos em migrar através da camada de muco e atingir as células epiteliais gástricas.

3.3. gp90

A gp90, outra glicoproteína abundantemente expressa em formas tripomastigotas metacíclicas, age como um modulador negativo da invasão de células mamíferas (Malaga e Yoshida, 2001). Após o contato com o suco gástrico, gp90 intacta age como um modulador negativo do processo de invasão (Cortez *et al.*, 2006), por isso, é sugerido que as interações entre gp90 e a célula hospedeira predominam sobre as interações mediadas por gp82 ou gp30 (Yoshida, 2009). Mesmo assim, gp90 induz uma mobilização de Ca^{2+} na célula hospedeira, porém de mais baixa intensidade do que a interação produtiva provocada por gp82, visto que a maioria da atividade de sinalização por Ca^{2+} de extrato de tripomastigotas metacíclicas foi devido à gp82 (Ruiz *et al.*, 1998).

Resumidamente, para parasitos deficientes em gp82, a camada de muco representa uma barreira para a invasão das células epiteliais. Para parasitos que expressam gp82 e gp90 em altos níveis, existem duas particularidades: 1) parasitos que expressam gp90 suscetível à digestão por pepsina apresentam alta infectividade após a exposição ao suco gástrico ou 2) parasitos que expressam gp90 resistente à digestão por pepsina apresentam baixa capacidade infectiva.

3.4. gp35/50

As mucinas gp35/50 são expressas em formas tripomastigotas metacíclicas e epimastigotas (Yoshida *et al.*, 1989). Assim como gp82, moléculas de gp35/50 ligam-se às células alvo e provocam uma mobilização de Ca^{2+} tanto no parasito quanto na célula hospedeira, contudo, menos intensa que a induzida por gp82 (Ruiz *et al.*, 1998).

3.5. Cepa G e suas particularidades na infecção

Mecanismos moleculares da infecção pelo *T. cruzi* por via oral têm sido

elucidados por infecção de camundongos com diversas cepas, inclusive a cepa G, como já citado. Formas tripomastigotas metacíclicas da cepa G mostraram-se pouco infectivas *in vitro* (Ruiz *et al.*, 1998; Yoshida, 2006) e não produziram parasitemia *in vivo* em experimentos com camundongos tanto de infecção por via oral (Yoshida, 2006; Covarrubias *et al.*, 2007) quanto intraperitoneal (Yoshida, 1983; 2006) em comparação a outras cepas. Especificamente em relação à infecção por via oral, esse fenótipo observado foi atribuído à expressão da isoforma de gp90 resistente à digestão por pepsina por formas tripomastigotas metacíclicas da cepa G (Covarrubias *et al.*, 2007) que é uma das principais moléculas da superfície delas (Ruiz *et al.*, 1998). A ligação de gp90 à célula hospedeira induz uma baixa resposta de Ca^{2+} (Ruiz *et al.*, 1998) e, provavelmente por isso, a infectividade de formas tripomastigotas metacíclicas da cepa G é baixa. Mesmo expressando moléculas de gp82 na superfície, a presença de gp90 intacta impede a entrada do parasito nas células epiteliais da mucosa gástrica, o que resulta em uma pobre internalização. Todavia, algumas formas tripomastigotas metacíclicas são internalizadas e isso parece ser dependente de gp35/50 (Yoshida *et al.*, 1989; Ruiz *et al.*, 1998; Yoshida, 2006).

Em contraste às formas metacíclicas, amastigotas axênicas da cepa G apresentaram-se altamente infectivas *in vitro*, mas não sustentaram uma infecção patente *in vivo* após inoculação intraperitoneal. Foi investigado como o hospedeiro estaria respondendo à infecção por amastigotas de forma a abolir completamente a parasitemia. Constatou-se a susceptibilidade de amastigotas da cepa G à produção de IFN- γ por células do hospedeiro que atuam na resposta imune inata, majoritariamente por células NK. É geralmente observado que *T. cruzi* é sensível a diferentes mecanismos da resposta imune, porém também é habilitado em subverter essas respostas, estabelecendo uma fase aguda caracterizada por parasitemia e morte do hospedeiro. Esse foi o primeiro relato de uma resposta imune tradicionalmente bem conhecida baseada na produção de IFN- γ que é capaz de controlar a infecção pela cepa G pouco virulenta (Rodrigues *et al.*, 2012). Amastigotas da cepa G também mostraram-se indutores efetivos de fagocitose clássica em células com atividade fagocítica não profissional. A fagocitose por células com atividade fagocítica não profissional ocorreu em níveis similares ou levemente reduzidos aos de macrófagos, fagócitos profissionais. Foi sugerido que a habilidade em realizar fagocitose é provavelmente uma adaptação específica desses

parasitos em invadir vários tipos celulares e propagar a infecção no hospedeiro mamífero (Fernandes *et al.*, 2013).

O aprofundamento dos estudos de cepas tidas como pouco virulentas como a cepa G é de grande importância para o entendimento das diferentes respostas imunes dos hospedeiros frente à infecção por parasitos de diversos DTUs de *T. cruzi*.

4. Organização genômica de *T. cruzi*

O DNA genômico de *T. cruzi*, assim como o de todos os membros da Ordem *Kinetoplastida*, é composto pelo DNA nuclear e pelo kDNA. A rede do kDNA consiste de milhares de anéis de DNA interligados, de dois tipos, mini e maxicírculos. Em *T. cruzi*, os minicírculos são moléculas de aproximadamente 1,4 kb concatenadas e presas aos maxicírculos; codificam a maioria dos RNAs guias (gRNA: *guide RNA*), que participam da editoração de transcritos dos maxicírculos. Cada célula possui de 40 a 50 moléculas de maxicírculos de aproximadamente 40 kb que codificam rRNAs e algumas subunidades da maquinaria bioenergética mitocondrial (Junqueira, Degrave e Brandao, 2005; Jensen e Englund, 2012).

Os genomas dos tripanossomatídeos são organizados em grupos de genes policistrônicos (PGC: *polycistronic gene cluster*) longos que são dezenas a centenas de genes codificadores de proteínas organizados sequencialmente em um mesma fita de DNA, os quais não são funcionalmente relacionados. A vasta maioria dos genes codificadores perdeu os íntrons e o *cis-splicing* tem sido observado apenas para o gene que codifica poli(A) polimerase de *T. cruzi* e *T. brucei* (Mair *et al.*, 2000). Os genes são densamente agrupados e separados por pequenas regiões intergênicas ricas em pirimidinas envolvidas no processamento do mRNA. Arranjos *in tandem* de genes que codificam rRNAs são encontrados entre PGCs, enquanto a maioria dos genes que codificam tRNAs são localizados nas vizinhanças de PGCs (Martinez-Calvillo *et al.*, 2010; De Gaudenzi *et al.*, 2011).

Duas cepas de *T. cruzi* já tiveram os genomas sequenciados e publicados, CL Brener e Sylvio X10/1. O primeiro genoma a ser sequenciado, da cepa CL Brener (TcVI), revelou que trata-se de um organismo híbrido originado da fusão de dois haplótipos parentais, Esmeraldo (TcII) e não Esmeraldo (TcIII). Foi estimado que o genoma haploide é composto por, aproximadamente, 12.000 genes. O tamanho do

genoma diploide foi estimado entre 106,4 e 110,7 Mb. No mínimo, 50% do genoma dessa cepa é composto por sequências repetitivas, outra característica geral do genoma de tripanossomatídeos, predominantemente formadas por membros de grandes famílias de genes que codificam proteínas de superfície como 63 kDa *glycoprotein* (gp63), trans-sialidasas (TS), mucinas e *mucin-associated surface proteins* (MASP), e de retrotransposons e regiões subteloméricas. Em *L. major* e *T. brucei*, é evidente a expansão dessas grandes famílias de genes em *T. cruzi* (El-Sayed *et al.*, 2005).

A publicação das sequências dos genomas de TriTryp em 2005 (Berriman *et al.*, 2005; El-Sayed *et al.*, 2005; Ivens *et al.*, 2005) foi um importante avanço no estudo desses parasitos. Porém essas publicações não revelaram sequências completas, sendo a de *T. cruzi* a menos abrangente de todas devido à grande quantidade de sequências repetitivas e ao caráter híbrido da cepa CL Brener. Além disso, a ausência de informação em nível de cromossomos do sequenciamento desse genoma dificulta análises sobre a localização de genes e seus alelos e/ou parálogos, o que é importante para estudos de genômica comparativa e estratégias de nocaute de genes.

A cepa Sylvio X10/1 (Tcl) apresenta um genoma não híbrido, com baixos níveis de heterozigose, menor e menos repetitivo do que o genoma de CL Brener. Uma montagem parcial do genoma foi realizada, contendo apenas 49% dos dados gerados. O tamanho do genoma haploide de Sylvio X10/1 foi estimado em 44 Mb, pela extrapolação do tamanho combinado dos *contigs* e de sequências repetitivas que não foram utilizadas na montagem (Franzen *et al.*, 2011).

A heterogeneidade e a presença maciça de sequências repetitivas do genoma de *T. cruzi* foram determinantes para a montagem incompleta ou parcial dos genomas de CL Brener e de Sylvio X10/1. A ausência de informação de alelos e genes parálogos dificulta estudos de genômica comparativa e nocaute gênico. O sequenciamento de um genoma não híbrido, menor, menos repetitivo, com baixos níveis de heterozigose aliado à realização de uma montagem de qualidade superior àquelas executadas é de grande importância para estudos de intensa análise de genômica e até estudos que envolvam apenas clonagem gênica em *T. cruzi*. Mesmo com uma qualidade inferior, o genoma da cepa CL Brener é o melhor modelo existente.

5. Regulação da expressão gênica em *T. cruzi*

Na transcrição de genes em *T. cruzi* e outros tripanossomatídeos, PGCs são transcritos pela RNA polimerase II, originando unidades de pré-mRNAs policistrônicos. Durante a maturação dos mRNAs, essas unidades são processadas em monocistrons por duas reações de clivagem. No mecanismo de *trans-splicing*, é adicionada uma sequência de 39 nucleotídeos, SL (*spliced leader*) ou mini-éxon, na extremidade 5' dos mRNAs. Na poliadenilação, é adicionada uma cauda poli(A) na extremidade 3' dos transcritos. Apesar da transcrição policistrônica, observam-se diferentes níveis de expressão dos monocistrons de genes adjacentes e padrões de expressão que variam de acordo com a forma de vida do parasito. Adicionalmente, não foram evidenciados promotores clássicos para a RNA polimerase II nas regiões a montante da maioria dos genes que codificam proteínas nos tripanossomatídeos. Dessa forma, em *T. cruzi*, a regulação da expressão gênica ocorre principalmente em nível pós-transcricional e, para isso, três mecanismos principais são propostos: processamento do pré-mRNA, degradação do mRNA e repressão traducional (De Gaudenzi *et al.*, 2011; Li *et al.*, 2012).

6. Proteômica

O termo proteoma foi cunhado em meados dos anos 90 para designar a análise da expressão gênica baseada em proteínas. Um PROTEOMA foi definido como todo o complemento de PROTEÍNAS expressas simultaneamente por um genOMA, de uma célula ou de um tipo tecidual em um determinado tempo, considerando-se que o proteoma pode variar sob diferentes condições e em vários tecidos de um único organismo (Wasinger *et al.*, 1995; Wilkins *et al.*, 1996). Conceitualmente, a amostra biológica pode ser derivada de células, tecidos, organelas, fração celular, vírus.

A área científica que estuda proteomas, a proteômica, engloba grande parte da análise funcional dos produtos gênicos ou "genômica funcional", incluindo a identificação em larga-escala. A análise de proteoma também pode fornecer informação sobre: isoformas, modificações pós-traducionais, localização subcelular de proteínas, interações proteína-proteína, composição molecular de estruturas celulares, quantificação de proteínas expressas por diferentes estágios de vida de

um microrganismo ou em diferentes condições experimentais (Pandey e Mann, 2000).

Ao mesmo tempo que as sequências nucleotídicas disponibilizadas pela genômica contribuem para a proteômica, essa área pode ser útil para a identificação e a validação da anotação de genes em organismos que tiveram os genomas sequenciados, pois a existência de uma fase aberta de leitura (ORF: *open reading frame*) não necessariamente implica na existência de um gene funcional. Além disso, a proteômica pode contribuir para: correção de anotações de genes, identificação de variantes de *splicing* alternativo, correção de sítios de iniciação da tradução usando peptídeos acetilados na extremidade N-terminal. A utilização de dados da proteômica é complementar à anotação genômica convencional porque foca nos produtos dos genes. Isso resultou em outro termo, a proteogenômica, que se refere à correlação dos dados da proteômica com os dados da genômica e da transcritômica com o objetivo de aprimorar o entendimento do genoma (Pandey e Mann, 2000; Renuse, Chaerkady e Pandey, 2011).

Nos anos recentes, as tecnologias na proteômica, particularmente a identificação de proteínas baseada em espectrometria de massas (MS - *mass spectrometry*), têm sido amplamente aperfeiçoadas em virtude de avanços na instrumentação, na preparação de amostra e na análise computacional. Há uma década, o sequenciamento e a identificação de uma proteína individual eram um grande desafio, enquanto que, atualmente, a identificação e a quantificação de grande parte das proteínas expressas são alcançáveis em um único experimento (Altelaar, Munoz e Heck, 2013). Esse aprimoramento tecnológico permitiu a utilização de dois tipos de abordagens: proteoma total e subproteoma. O proteoma total refere-se a uma análise global das proteínas expressas por um organismo ou um tipo celular, enquanto que, para um subproteoma, a amostra biológica é submetida a um fracionamento celular ou enriquecimento. Para o subproteoma, a complexidade da amostra é reduzida, o que possibilita a detecção de proteínas pouco expressas e que estariam sub-representadas e poderiam nem ser detectadas no proteoma total. Análises de proteomas de alto rendimento ainda são incompletas, pois mesmo com as tecnologias mais modernas não é possível a identificação e a quantificação de todas as proteínas expressas por um organismo em um determinado tempo. Entretanto, a amplitude de proteínas identificadas após a

submissão de uma amostra a um espectrômetro de massas tem se tornado cada vez maior.

6.1. Proteomas totais de *T. cruzi*

Como uma análise de proteoma total de *T. cruzi* foi realizada neste trabalho, serão abordados a seguir proteomas totais previamente publicados de diferentes cepas desse protozoário.

Antes da publicação do sequenciamento do genoma de *T. cruzi* em 2005, alguns estudos de proteoma total de *T. cruzi* foram realizados mesmo com um número reduzido de sequências gênicas disponíveis nos bancos de dados. No proteoma pioneiro de *T. cruzi*, foi feita a padronização das condições experimentais de eletroforese de proteínas em gel bidimensional (2-DE) para formas epimastigotas, tripomastigotas e amastigotas da cepa Berenice, análises por espectrometria de massas do tipo *matrix-assisted laser desorption/ionization-time of flight* (MALDI-TOF) e de proteínas diferencialmente expressas entre as formas de vida. Isso resultou em 26 identificações que corresponderam a 19 proteínas diferentes (Paba, Santana, *et al.*, 2004).

Com os mesmos detalhes experimentais descritos acima, Parodi-Talice *et al.* (2004) realizaram a análise de proteoma total de formas epimastigotas da cepa Dm28c, com a qual foram obtidas 45 identificações (25 proteínas diferentes), entre elas proteínas de choque térmico (HSP - *heat shock proteins*), proteínas estruturais (α - e β -tubulinas), proteínas envolvidas na síntese proteica e no ciclo celular.

Ainda em 2004, Paba, Ricart, *et al.* (2004) publicaram o primeiro artigo sobre a análise proteômica comparativa quantitativa entre formas tripomastigotas e amastigotas da cepa Berenice usando uma tecnologia livre de separação eletroforética, *isotope-coded affinity tag* (ICAT) associada a LC-MS/MS (*liquid chromatography-in tandem mass spectrometry*), que resultou na identificação de 41 proteínas, das quais 29 mostraram expressão conservativa em ambos os estágios de vida, 9 mais expressas em tripomastigotas e 3 mais expressas em amastigotas.

Nesses três trabalhos, foi realizada busca no banco de dados NCBI (*National Cancer for Biotechnology Information*, Bethesda, USA) não redundante.

Um trabalho publicado em 2005 (Atwood *et al.*, 2005) produziu um forte impacto científico, pois nessa ocasião, já utilizando o banco de dados com as anotações de genes fornecidas pelo consórcio de sequenciamento do genoma de *T.*

cruzi, 2.784 proteínas de formas epimastigotas, tripomastigotas, tripomastigotas metacíclicas e amastigotas da cepa Brazil foram identificadas. Particularmente, de formas epimastigotas, foram detectadas 1.861 proteínas. A análise dos peptídeos fornecidos por digestão em solução de lisados celulares foi realizada por LC-MS/MS. Aproximadamente 30% das proteínas identificadas foram detectadas em todas as quatro formas de vida analisadas. Além disso, mais de 1.000 proteínas identificadas nessa análise serviram para a anotação de genes antes classificados como hipotéticos, validando-os como genes autênticos (*bona fide*) em *T. cruzi*. A identificação direta dos produtos gênicos pode conduzir à anotação de genes em um genoma sequenciado, evidenciando que a cooperação mútua entre a proteômica e a genômica pode ser de grande valia. Ademais, duas enzimas identificadas e envolvidas na via metabólica de conversão de histidina em glutamato, histidina amônia-liase e urocanato hidratase, são abundantes nos estágios de vida do *T. cruzi* em triatomíneos, epimastigotas e tripomastigotas metacíclicas, e quase não detectáveis nos estágios de vida em mamíferos. A histidina é um aminoácido livre dominante nas excreções e na hemolinfa de triatomíneos *R. prolixus*, sugerindo que o parasito parece adaptado para a utilização dessa fonte de energia abundante no intestino do inseto vetor. Em formas amastigotas, a obtenção de energia através dos ácidos graxos foi indicada pela identificação de enzimas putativas do catabolismo deles e pela ausência virtual de transportadores de glicose.

Em virtude da baixa resolução das proteínas básicas de *T. cruzi* em 2-DE de vários trabalhos, foi realizada a otimização do protocolo de 2-DE e, depois, a comparação entre os proteomas alcalinos de formas epimastigotas e tripomastigotas da cepa Berenice. Os resultados desse trabalho corroboraram a expressão diferencial ou estágio-específica de proteínas entre formas epimastigotas e tripomastigotas observada em estudos prévios (Magalhaes *et al.*, 2008).

6.2. Subproteoma da cepa G

Uma análise subproteômica comparativa da cepa G de *T. cruzi* foi realizada a partir de frações ricas em proteínas ancoradas a GPI e outras proteínas de membrana de formas tripomastigotas metacíclicas e epimastigotas submetidas a LC-MS/MS, resultando em 98 e 280 proteínas identificadas, respectivamente (Cordero *et al.*, 2009). Aproximadamente 40% das proteínas identificadas em tripomastigotas metacíclicas como membros da superfamília das TS apresentaram classificação

funcional predita na interação parasito-hospedeiro, adesão e invasão de células hospedeiras. Por outro lado, epimastigotas tiveram mais proteínas putativas do metabolismo celular, em concordância com a alta exigência energética desse estágio proliferativo. Deve-se ressaltar que, no referido trabalho, consideraram-se identificações com um ou mais peptídeos e, em contrapartida, a maioria dos estudos prioriza identificações resultantes de dois ou mais peptídeos. Também foi usado o banco de dados de proteínas não redundante de *T. cruzi* proveniente do GenBank contendo 191.762 sequências. A utilização de um banco de dados portador de uma quantidade tão vasta de sequências de proteínas pode ser desfavorável à obtenção das identificações, pois provavelmente apresenta sequências redundantes, visto que o genoma de referência haploide da cepa CL Brener de *T. cruzi* é formado por 12.000 genes.

6.3. Espectrometria de massas

Há vinte anos, o sequenciamento de proteínas dava-se principalmente por uma técnica conhecida por degradação de Edman (Edman, 1949) que, fundamentalmente, baseia-se na identificação de aminoácidos que são quimicamente clivados, um por vez, da extremidade N-terminal de um peptídeo proveniente da digestão de uma proteína. Isso requer muita perícia e uma quantidade substancial de proteína purificada, e falha completamente se o peptídeo é acetilado ou, de outra forma, bloqueado na extremidade N-terminal, pois para a reação de Edman é necessária uma extremidade N-terminal livre (revisado por Steen e Mann (2004)).

Atualmente, a espectrometria de massas consolidou-se como uma técnica analítica central para o campo da proteômica e possibilitou a análise simultânea de diversas proteínas numa mesma amostra (*shotgun proteomics*) (Wu e Maccoss, 2002) por LC-MS/MS, em modo DDA (*data-dependent acquisition*) para controlar o processo de aquisição pelo espectrômetro de massas. Simplificadamente, as proteínas são preparadas a partir de uma fonte biológica e podem ser antes fracionadas por SDS-PAGE ou 2-DE ou diretamente convertidas (digestão em solução) em peptídeos por uma protease específica, como a endoprotease Lys-C e/ou tripsina. A tripsina cliva proteínas muito especificamente na extremidade C-terminal de resíduos de arginina e lisina gerando pequenos peptídeos e é amplamente utilizada até mesmo na presença de 1,0 M de ureia. Os peptídeos são

separados por *high-performance liquid chromatography* (HPLC) de fase estacionária reversa, por exemplo, que é diretamente acoplada ao espectrômetro de massas. Depois, os peptídeos podem ser ionizados por MALDI ou ESI (*electrospray ionization*) e adentrar o espectrômetro de massas (Steen e Mann, 2004).

Os espectrômetros de massas diferem em relação aos analisadores da razão massa/carga (m/z) dos peptídeos ionizados. Três tipos de analisadores são frequentemente usados em abordagens proteômicas por MS: *quadrupole*, *time of flight* (TOF) e *quadrupole ion trap*; e também equipamentos híbridos que combinam princípios. Os espectrômetros de massas geram espectros de massas (espectros MS), que correlacionam a intensidade do sinal e a razão m/z de cada íon. Em seguida, os íons podem ser fragmentados isoladamente por colisões com um gás inerte (CID - *collision-induced dissociation* ou HCD - *high energy collision induced dissociation*), resultando em espectros MS/MS ou *tandem* MS (Steen e Mann, 2004).

Muitas ferramentas computacionais têm sido desenvolvidas para a identificação automatizada de peptídeos a partir de LC-MS/MS. Esses programas podem ser divididos em duas categorias: busca em um banco de dados e sequenciamento *de novo* (Zhang *et al.*, 2012).

A busca automatizada em um banco de dados é amplamente aplicada na proteômica, entretanto, requer uma alta especificidade de correlação entre espectro e sequência. Discrepâncias entre eles podem prevenir a identificação de um peptídeo, tais como: substituição de aminoácido, modificação pós-traducional (PTM), via de fragmentação peculiar. Conseqüentemente, um grande número de espectros adquiridos em proteômica de larga escala permanece sem correspondência ou corresponde a peptídeos com baixa pontuação, resultando em identificações errôneas de proteínas que, na verdade, não estão na amostra (falsos positivos) ou ignorando proteínas que efetivamente estão na amostra (falsos negativos) (Johnson 2005). A estratégia mais comum de determinar o chamado *false discovery rate* (FDR) é pesquisar em um banco de dados randomizado (*decoy*) em paralelo ao banco de dados alvo (Elias e Gygi, 2007). FDR é definido como a porcentagem de falsas identificações em relação a todas as identificações acima do limiar de pontuação ou a razão entre o número de peptídeos que corresponderam ao banco de dados *decoy* e a soma dos peptídeos que corresponderam a ambos os bancos de dados, *decoy* e alvo. Essa abordagem assume que as correspondências para

sequências de peptídeos *decoy* e as correspondências denominadas falsos positivos seguem a mesma distribuição (Junqueira, 2009).

Enquadrados na categoria de busca em banco de dados estão os programas Sequest (Eng, McCormack e Yates, 1994), Mascot (Perkins *et al.*, 1999), X!Tandem (Craig e Beavis, 2004), ProteinProspector (Chalkley *et al.*, 2005) e MaxQuant (Cox e Mann, 2008; Cox *et al.*, 2011), que ainda podem experimentar problemas de baixos índices de identificação e altos FDR, resultando em baixas sensibilidade e acuidade, respectivamente (Zhang *et al.*, 2012). Um dos dois programas utilizados no presente trabalho foi o Sequest, que emprega um algoritmo de autocorrelação para matematicamente determinar a sobreposição entre um espectro teórico derivado de um banco de dados e o espectro experimental MS/MS em questão. A sobreposição é dada na forma de uma pontuação (*score*), como ilustrado na Figura 3, e a pontuação para a melhor correspondência (*matching*) é frequentemente fornecida (Steen e Mann, 2004).

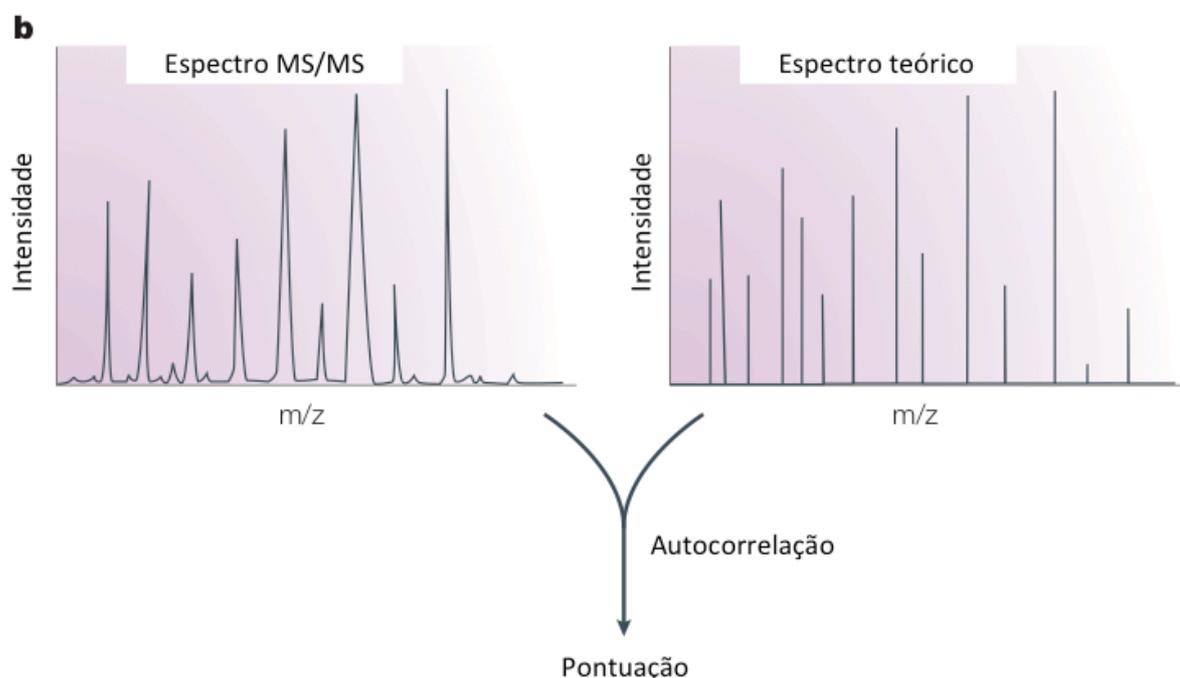


Figura 3. Princípio do algoritmo do programa Sequest descrito por Eng, McCormack e Yates (1994). Utiliza a autocorrelação para determinar a sobreposição entre um espectro teórico derivado de um banco de dados e o espectro experimental MS/MS em questão. O resultado é fornecido sob a forma de pontuação. Modificado de Steen e Mann (2004).

Alguns representantes da categoria de sequenciamento *de novo* são os programas PEAKS (Ma *et al.*, 2003), PepNovo (Frank e Pevzner, 2005) e NovoHMM (Fischer *et al.*, 2005). Por exemplo, PEAKS, o outro programa usado neste trabalho, fornece uma análise diferenciada, pois combina quatro algoritmos complementares para sequenciamento *de novo*, busca em um banco de dados, caracterização de PTM não especificadas e detecção de mutações, respectivamente denominados de: PEAKS *de novo sequencing*, PEAKS DB, PEAKS PTM e SPIDER (Figura 4). Segundo o manual do programa, isso maximiza o número de identificações com altas sensibilidade e acuidade. O sequenciamento *de novo* pode derivar a sequência peptídica diretamente do espectro experimental MS/MS sem a dependência de um banco de dados de proteínas, o que possibilita a identificação de proteínas expressas por organismos sem genoma sequenciado (Ma *et al.*, 2003). Mesmo assim, é possível configurar para a realização de busca dos peptídeos obtidos pelo sequenciamento *de novo* no banco de dados de interesse de modo a aumentar a sensibilidade e a acuidade em relação à busca automatizada padrão (Zhang *et al.*, 2012). A identificação de PTM pelo PEAKS não requer que o usuário especifique as possíveis PTM, pelo contrário, todas aquelas provenientes de banco de dados de PTM (como o banco de dados Unimod) são consideradas (Han *et al.*, 2011). Por fim, o algoritmo SPIDER fornece sequências peptídicas novas e mutações em relação ao banco de dados (Han, Ma e Zhang, 2004).

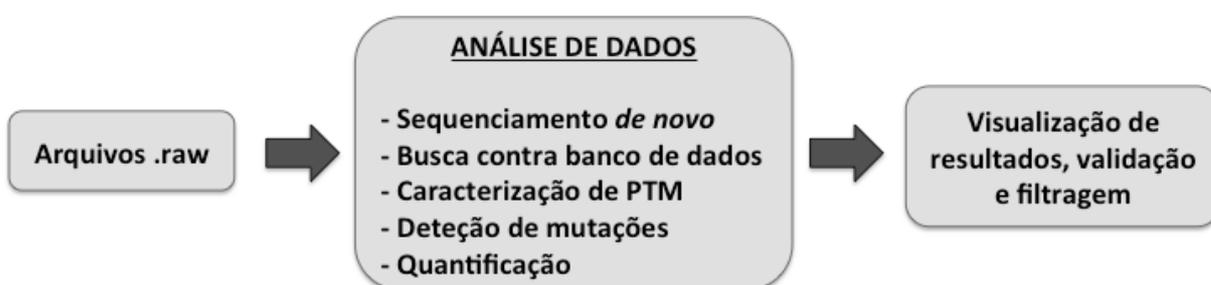


Figura 4. Etapas da análise dos dados gerados por espectrometria de massas pelo programa PEAKS.

6.3.1. LTQ-Orbitrap

Espectrômetros de massas híbridos fornecem uma plataforma instrumental versátil e poderosa para a análise de proteínas. O mais moderno espectrômetro de massas LC-MS/MS, Orbitrap, foi descrito por Makarov (2000) e baseia-se na captura de íons que se movem orbitalmente em um campo eletrostático. Esse tipo de captura dinâmica - *orbital trapping* - já havia sido implementada na década de 20 e Makarov utilizou desse conceito para aplicação em espectrometria de massas.

O espectrômetro de massas híbrido LTQ-Orbitrap (Figura 5) é constituído por duas partes funcionais principais: o *linear trap quadrupole (LTQ)* ou *linear ion trap* e o Orbitrap. Inicialmente os peptídeos são ionizados por ESI, ionização química à pressão atmosférica ou *nanoelectrospray*. O LTQ executa a captura de íons, a seleção de íons, as reações de fragmentação e a detecção de íons com baixa resolução, enquanto que o Orbitrap realiza a detecção de íons com altas acuidade de massas e resolução. E ainda podem funcionar de dois modos: no primeiro conhecido como FT-IT e amplamente utilizado para análises *shotgun*, o Orbitrap e o LTQ adquirem espectros MS e espectros MS/MS, respectivamente; no segundo chamado FT-FT, ambos os tipos de espectros são analisados no Orbitrap. O analisador de massas Orbitrap é formado por dois eletrodos externos e um central (Figura 6). Os íons injetados no Orbitrap são capturados em um campo eletrostático e a atração eletrostática em direção ao eletrodo central é compensada por uma força que resulta da velocidade tangencial dos íons. O campo eletrostático força os íons a se moverem em padrões espirais complexos. O componente axial dessas oscilações é independente de energia inicial, ângulos e posições, e pode ser detectado como uma imagem nos eletrodos externos. A transformação de Fourier é usada para se obter as frequências de oscilação para íons com diferentes massas, resultando em uma leitura precisa dos m/z deles. As frequências são finalmente convertidas em espectros de massas (Scigelova e Makarov, 2006; Makarov e Scigelova, 2010; Kalli *et al.*, 2013).

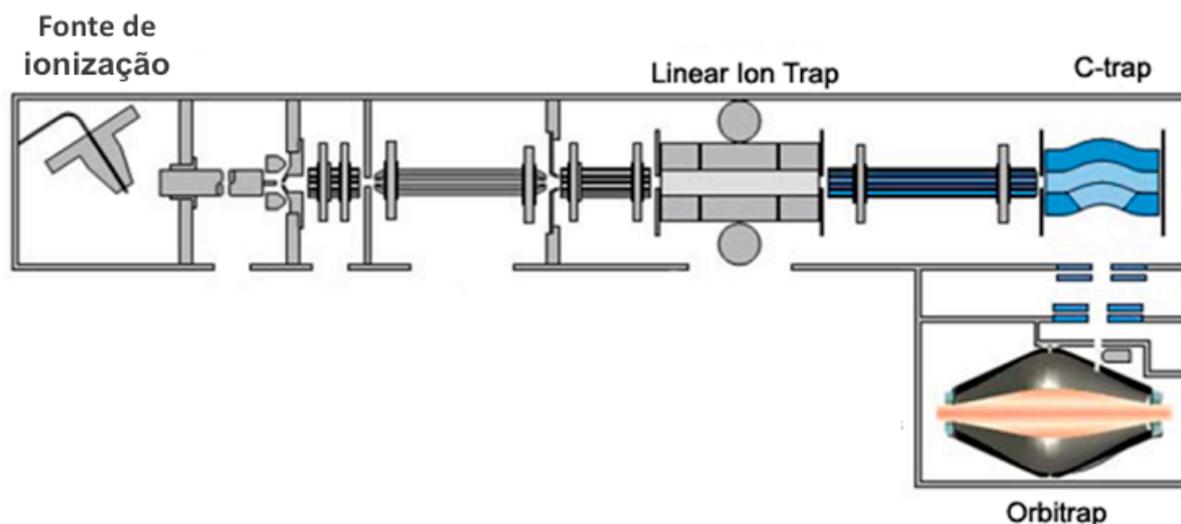


Figura 5. Esquema de um espectrômetro de massas híbrido, um LTQ-Orbitrap. A fonte de ionização pode ser por ESI, ionização química à pressão atmosférica ou *nanoelectrospray*. No *linear ion trap*, os íons são capturados, selecionados, fragmentados e detectados a uma baixa resolução. O *C-trap* estoca uma população significativa de íons e injeta-os no Orbitrap. No analisador de massas Orbitrap, os íons são detectados com altas acuidade e resolução. Modificado de Kalli *et al.* (2013).

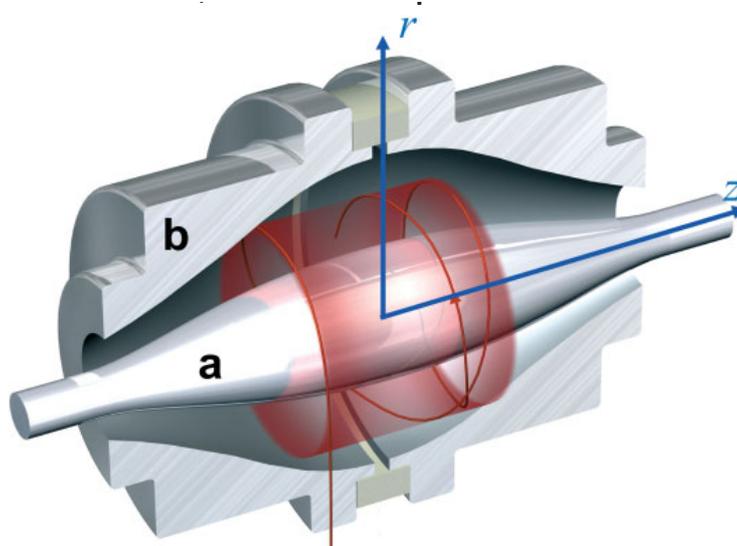


Figura 6. Estrutura interna do analisador de massas Orbitrap. Através do corte transversal, observa-se que o Orbitrap é formado por dois eletrodos externos (b) e um eletrodo interno (a). Os íons oscilam orbitalmente ao eletrodo central devido à ação de campo eletrostático. Sinais são detectados nos eletrodos externos e são convertidos em frequências de oscilação. Os valores m/z dos íons podem ser determinados a partir das respectivas frequências de oscilação depois da transformação de Fourier (Scigelova e Makarov, 2006).

JUSTIFICATIVA

Parasitas da cepa G de *T. cruzi* (DTU TcI) apresentam um genoma não híbrido e menor em relação à cepa CL Brener que teve uma montagem parcial do genoma realizada, porém é o de melhor qualidade disponibilizado até o momento. Tais características genômicas aliadas à utilização de avançadas tecnologias de sequenciamento de DNA ou *next-generation sequencing* (NGS) facilitarão a realização e a disponibilização de uma montagem de genoma de *T. cruzi* com qualidade satisfatória e superior às das cepas CL Brener e Sylvio X10/1. Posteriormente, realizaremos a montagem do genoma da cepa G cujo DNA genômico já foi sequenciado. Dentro desse grande projeto do nosso grupo, a análise proteômica da cepa G seria importante para a anotação dos genes fornecidos pela montagem do genoma devido à identificação dos produtos gênicos.

Além disso, estudos já citados demonstraram que a cepa G tida como pouco virulenta apresenta peculiaridades de infectividade dependendo do estágio de vida e do tipo de infecção experimental analisados: formas tripomastigotas metacíclicas mostraram-se pouco infectivas *in vitro* e não produziram parasitemia *in vivo* em experimentos com camundongos tanto de infecção por via oral quanto intraperitoneal; enquanto que, surpreendentemente, amastigotas axênicas foram altamente infectivas *in vitro*, mas não sustentaram uma infecção patente *in vivo* após inoculação intraperitoneal. Assim como é importante o estudo de cepas virulentas, também é fundamental aprofundar o entendimento de cepas com pouca importância clínica e epidemiológica como a cepa G a fim de se entender quais moléculas e mecanismos estão envolvidos no estabelecimento do fenótipo de baixa infectividade.

A análise proteômica de epimastigotas da cepa G de *T. cruzi* será útil tanto para a anotação dos genes fornecidos pela montagem do genoma quanto para uma posterior análise comparativa das proteínas expressas entre as outras formas de vida do parasito e também de outras cepas. O conhecimento de um conjunto de proteínas expressas pela cepa G em determinadas condições pode ser de grande utilidade para melhorar a compreensão sobre a biologia da interação *T. cruzi*-hospedeiro.

OBJETIVO

Diante do que foi exposto até agora, o objetivo deste trabalho é descrever de forma abrangente o proteoma total de formas epimastigotas em fase exponencial de crescimento da cepa G de *T. cruzi*.

MATERIAIS E MÉTODOS

Cultura *in vitro* de epimastigotas da cepa G de *T. cruzi*

Formas epimastigotas da cepa G de *T. cruzi* (cedidas pela profa. Nobuko Yoshida - UNIFESP) foram crescidas a 28 °C em meio LIT (*liver infusion tryptose*) suplementado com 5% de soro fetal bovino (SFB) (Invitrogen) e 100 µg/mL de gentamicina (Neoquímica).

Curva de crescimento de epimastigotas da cepa G de *T. cruzi*

A curva de crescimento de formas epimastigotas da cepa G foi feita através da contagem das células em câmara de Neubauer ao longo de 7 dias. A concentração inicial da cultura foi de $3,13 \times 10^5$ células/mL.

Extração de DNA genômico de epimastigotas da cepa G de *T. cruzi*

O DNA genômico foi extraído segundo Medina-Acosta e Cross (1993), com algumas modificações. A quantidade de 3×10^9 parasitos foi lavada três vezes com 1x de PBS, ressuspensa em 10 mL de tampão de lise (10 mM de Tris-HCl pH 8,0, 100 mM de EDTA pH 8,0, 0,5% de SDS, 20 µg/mL de RNase A (Invitrogen)) e incubada a 60 °C por 3 h. Depois foi feita a incubação do lisado com 100 µg/mL de proteinase K (Sigma) a 37 °C por 18 h. O lisado foi dividido em vários tubos de 1,5 mL, adicionou-se igual volume de fenol clorofórmico (Sigma) e misturou-se por inversão suavemente por 5 min. As fases foram separadas por centrifugação a 13.000 g por 5 min. As fases aquosas (superior) foram transferidas para tubos novos e adicionou-se igual volume de fenol clorofórmico. Foram feitas, no total, três extrações consecutivas com fenol clorofórmico. Ao final dessas, as fases aquosas foram agrupadas em um único tubo Falcon de 50 mL. Foram adicionados dois volumes de etanol 100%, misturou-se por inversão levemente e o DNA foi pescado e transferido para um tubo novo contendo 1 mL de etanol 100%. O DNA foi centrifugado a 13.000 g por 10 min. Foram feitas duas lavagens com etanol 100%. Ao fim da segunda lavagem, o sobrenadante foi descartado e, após secagem, o DNA foi ressuspensa em 100 µL de água milli-Q.

Genotipagem da cepa G de *T. cruzi*: amplificação por PCR do rDNA 24S α e do gene *SL*

Oligonucleotídeos usados para genotipagem de *T. cruzi* segundo Souto *et al.* (1996) e DNA genômico da cepa Y foram gentilmente cedidos pela profa. Beatriz Dolabela de Lima - UnB. Para o rDNA 24S α do cistron ribossômico foram utilizados os oligonucleotídeos 5'-AAG GTG CGT CGA CAG TGT GG-3' e 5'-TTT TCA GAA TGG CCG AAC AGT-3'. A amplificação por PCR foi realizada com as seguintes etapas: uma desnaturação a 94 °C por 1 min; 30 ciclos de desnaturação a 94 °C por 30 s, anelamento a 64 °C por 30 s e extensão a 72 °C por 35 s; uma extensão final a 72 °C por 1 min. Para o gene *SL*, três oligonucleotídeos foram empregados: 5'-CCC CCC TCC CAG GCC ACA CTG-3', 5'-GTG TCC GCC ACC TCC TTC GGG CC-3' e 5'-CCT GCA GGC ACA CGT GTG TGT G-3'. As etapas da PCR foram: uma desnaturação a 94 °C por 1 min; 27 ciclos de desnaturação a 94 °C por 30 s, anelamento a 55 °C por 30 s e extensão a 72 °C por 30 s; uma extensão final a 72 °C por 1 min. Para ambos os tipos de genes a serem amplificados, as condições dos reagentes usados na PCR, por reação, foram: 100 ng de DNA genômico de *T. cruzi*, 0,33 μ M de cada um dos iniciadores, 1x de tampão da enzima Platinum *Taq* DNA Polimerase (Invitrogen), 0,2 mM de dNTPs, 1,5 mM de MgCl₂, 1,5 U de enzima Platinum *Taq* DNA Polimerase (Invitrogen). Para a amplificação dos genes 24S α e *SL* são gerados produtos de PCR de respectivas massas moleculares aproximadas: para a cepa Y (TcII), 125 pb e 300 pb; para a cepa G (TcI), 110 pb e 350 pb. Os fragmentos de DNA foram analisados em gel de agarose 1%.

Preparação do extrato de proteínas de epimastigotas da cepa G de *T. cruzi*

Os parasitos ($2,5 \times 10^9$) em fase log de crescimento foram coletados por centrifugação a 2.500 rpm por 10 min a 4 °C e lavados cinco vezes com PBS gelado. O sedimento foi diretamente ressuspendido em 5 mL de tampão de amostra Laemmli 1x (18,8 mM de Tris-HCl pH 6,8, 0,6% de SDS, 0,6 mM de EDTA, 2,76% de glicerol, 2,86 mM de 2-mercaptoetanol, 0,02% de azul de bromofenol) contendo *Complete Protease Inhibitor Cocktail* (Roche) e fervido por 5 min. 1 mL do extrato total de proteínas foi centrifugado a 15.000 g por 15 min a 4 °C e o sobrenadante foi separado para a análise seguinte.

Processamento do extrato de proteínas

Extrato total de proteínas (50 μ L relativos a $2,5 \times 10^7$ formas epimastigotas) foi separado por SDS-PAGE 12% de 18 cm x 16 cm no sistema de eletroforese vertical Hoefer SE600 (Hoefer) e a corrida foi parada antes de o tampão de amostra sair do gel para prevenir a perda de proteínas de baixa massa molecular. O gel foi corado com Coomassie Brilliant Blue R-250. Depois de remover o excesso de corante, a linha inteira incluindo o gel concentrador foi excisada em 22 faixas de gel. Essas faixas foram isoladas em tubos Eppendorf cortadas em pedaços ainda menores para posterior digestão em gel com tripsina como descrito previamente por Shevchenko *et al.* (2006) com algumas modificações.

Primeiramente, os pedaços de gel contidos em 22 tubos referentes a cada uma das faixas de gel foram encolhidos com acetonitrila (ACN) pura por 10 min. ACN foi removida para a adição de 10 mM de ditioneína (DTT) diluído em solução de bicarbonato de amônio a 20 mM até cobrir os pedaços de gel e posterior incubação por 30 min a 56 °C, etapa de redução das pontes dissulfeto das proteínas. Os tubos foram resfriados à temperatura ambiente e, então, foi feita incubação com ACN pura por 10 min. Após a remoção da ACN, a etapa de alquilação das proteínas foi realizada com a adição de 55 mM de iodoacetamida diluída em solução de bicarbonato de amônio a 20 mM até cobrir os pedaços de gel e a incubação por 20 min à temperatura ambiente e no escuro. Os pedaços de gel foram encolhidos com ACN que foi depois removida. Após as etapas de redução e alquilação em gel, os pedaços de gel provenientes de gel corado com Coomassie Blue foram descorados com 20 mM de bicarbonato de amônio/ACN (1:1 v/v) e incubados por 30 min com ocasional homogeneização. ACN pura foi adicionada e procedida a incubação à temperatura ambiente. As amostras estavam prontas para a digestão em gel. Solução de tripsina (Promega) diluída a 12,5 ng/ μ L em 10 mM de bicarbonato de amônio e 10% de ACN foi adicionada aos tubos até cobrir os pedaços de gel e a incubação foi feita a 37 °C durante a noite. Dois passos de extração dos peptídeos dos pedaços de gel com 5% de ácido fórmico/ACN (1:2 v/v) foram realizados a 37 °C por 15 min, cada. Por fim, os peptídeos foram submetidos à secagem em centrífuga a vácuo. Durante todo o processo, cuidados especiais foram tomados para prevenir a massiva contaminação com queratina e polímeros, tais como: utilização de luvas de nitrila, touca e máscara, lavagem de vidrarias com metanol-água-metanol, centrifugação das soluções em tubos Falcon sob rotação máxima, filtração das

soluções através de poros de 0,22 µm, água LC-MS Chromasolv (Sigma).

LC-MS/MS

Conforme Junqueira (2009), os peptídeos resultantes foram redissolvidos em 10 µL de ácido trifluoroacético (TFA) 0,1%. 4 µL das misturas de peptídeos foram aplicadas no sistema Easy II nanoLC (Thermo Fisher Scientific) sob pressão constante de 250 bars. Depois do carregamento e da lavagem, peptídeos foram eluídos em 15 cm x 75 µm *nano-column packed with* C18 Reprosil (Dr. Maish) 3 µm *particles* sob o fluxo de 200 nL/min. O gradiente de eluição para análise *in-gel* foi realizado em um gradiente de 5-20% de solvente B em 20 min; 20-50% de solvente B em 16 min; 50-100% de solvente B em 5 min; 100% de solvente B em 10 min e de volta para 5% de solvente B em 5 min. O solvente A era 95:5 H₂O:ACN (v/v) contendo ácido fórmico 0,1% e o solvente B era 20:80 H₂O:ACN (v/v) contendo ácido fórmico 0,1%. O nanoLC foi acoplado *online* ao espectrômetro de massas híbrido LTQ-Orbitrap Velos (Thermo Fisher Scientific) via uma fonte ionização Easy (Thermo Fisher Scientific) operada com a voltagem de ionização de 1,8 kV. MS/MS foi feito em modo DDA sob controle do programa Xcalibur 2.0 (Thermo Fisher Scientific). O ciclo DDA consistiu de um *survey scan* adquirido em um µscan dentro do intervalo de m/z 300-2000 realizado sob a resolução de massa alvo de 60.000 FWHM (*full width at half-maximum*) dentro do analisador Orbitrap e com o valor alvo de 1E06 íons. *Survey scan* foi seguido por fragmentação MS/MS dos dez precursores multiplamente carregados mais abundantes. Espectros foram adquiridos em um µscan sob energia de colisão normalizada de 35% e valor alvo de 1E04 íons no *linear ion trap* (*ion selection threshold 400 counts; precursor ions isolation width 4 amu*). Parâmetro de ativação q = 0,25 e tempo de ativação de 30 ms foram aplicados. Precursores previamente fragmentados foram dinamicamente excluídos para 90 s. Ao fim, foram gerados arquivos .raw com os espectros MS/MS.

Análise de dados

Para as buscas 1, 2 e 3, os espectros MS/MS contidos nos arquivos .raw foram confrontados com os respectivos bancos de dados: (1) *Trypanosoma* spp. e outros tripanossomatídeos e (2) *T. cruzi* CL Brener, ambos pelo algoritmo do programa Sequest, e (3) *T. cruzi* CL Brener e Sylvio X10/1 pelo algoritmo do

programa PEAKS6. Na busca 1, o banco de dados continha 88.279 sequências baixadas em maio de 2013; na busca 2, 19.244 sequências baixadas em julho de 2013; na busca 3, 30.158 sequências baixadas em outubro de 2013. Todos os bancos de dados foram obtidos do UniprotKB/Swiss-Prot e UniprotKB/TrEMBL (<http://www.uniprot.org>). 111 proteínas contaminantes (vários tipos de queratinas, albuminas e tripsinas) foram incluídas no banco de dados da busca 3, apenas. Todas as proteínas contaminantes identificadas foram manualmente removidas da lista de resultados. Os seguintes parâmetros foram ajustados no Sequest para as buscas 1 e 2: tolerâncias de massa do precursor e de massa do fragmento iguais a 10 ppm e 0,8 Da, respectivamente; até 2 clivagens perdidas permitidas. Os resultados foram filtrados para apenas peptídeos *rank* 1 e número mínimo de peptídeos por identificação igual a 2. Para a busca 3, os parâmetros da busca no programa PEAKS6 foram: tolerâncias de massa do precursor e de massa do fragmento iguais a 10 ppm e 0,8 Da, respectivamente; até 2 clivagens perdidas. Os filtros utilizados para a busca 3 foram: FDR (PSM) = 0,1%; $-10\lg P$ (peptídeo) $\geq 28,2$; $-10\lg P$ (proteína) ≥ 20 ; *de novo* TLC ≥ 3 ; *de novo* ALC $\geq 50\%$, apenas peptídeos *rank* 1.

As etapas seguintes são relativas às três abordagens supracitadas. Após a obtenção da lista de identificações, as sequências em FASTA foram submetidas à análise pelo programa Blast2GO (Conesa *et al.*, 2005) versão 2.6.6 com parâmetros padrão. A anotação funcional das proteínas identificadas foram referentes a: *Gene Ontology* (GO) em relação a componente celular, processo biológico e função molecular; *Enzyme Commission* (EC); vias metabólicas (KEGG *pathways*). Os programas online SignalP Server versão 4.1 (<http://www.cbs.dtu.dk/services/SignalP/>; (Petersen *et al.*, 2011)) e SecretomeP Server versão 2.0 (<http://www.cbs.dtu.dk/services/SecretomeP/>; (Bendtsen *et al.*, 2004)) foram usados para a predição das proteínas secretadas pelas vias clássica e não clássica, respectivamente. Por não haver uma seleção mais específica para *T. cruzi*, no SignalP, foi utilizada a predição de peptídeo-sinal para todos os tipos de organismos disponibilizados (eucariotos, bactérias gram-negativas e gram-positivas), assim como no SecretomeP (mamíferos, bactérias gram-negativas e gram-positivas). As identificações que pertenceram a ambos os tipos de predição foram excluídas da lista do SecretomeP e, assim, consideradas apenas na lista do SignalP. Também foi utilizado o programa TMHMM Server versão 2.0

(<http://www.cbs.dtu.dk/services/TMHMM/>) para a predição de domínios transmembrânicos nas proteínas identificadas. Os diagramas de Venn foram construídos pelo programa online Venny (<http://bioinfogp.cnb.csic.es/tools/venny/index.html>). Os gráficos foram gerados com a utilização do programa GraphPad Prism versão 6.0.

RESULTADOS E DISCUSSÃO

1. Curva de crescimento de epimastigotas da cepa G de *T. cruzi*

A curva de crescimento de formas epimastigotas da cepa G de *T. cruzi*, após 24 h, mostrou que a população de células mais que dobrou (de $3,13 \times 10^5$ para $7,30 \times 10^5$ parasitos). No 6º dia, a contagem foi de $2,31 \times 10^7$ parasitos e, no 7º dia, $3,33 \times 10^7$ parasitos, portanto, a população aumentou aproximadamente 1,5 vezes. O período analisado de 7 dias englobou a fase de crescimento exponencial ou log da população, havendo a necessidade de mais dias para se observar a entrada da população na fase estacionária de crescimento (Figura 7).

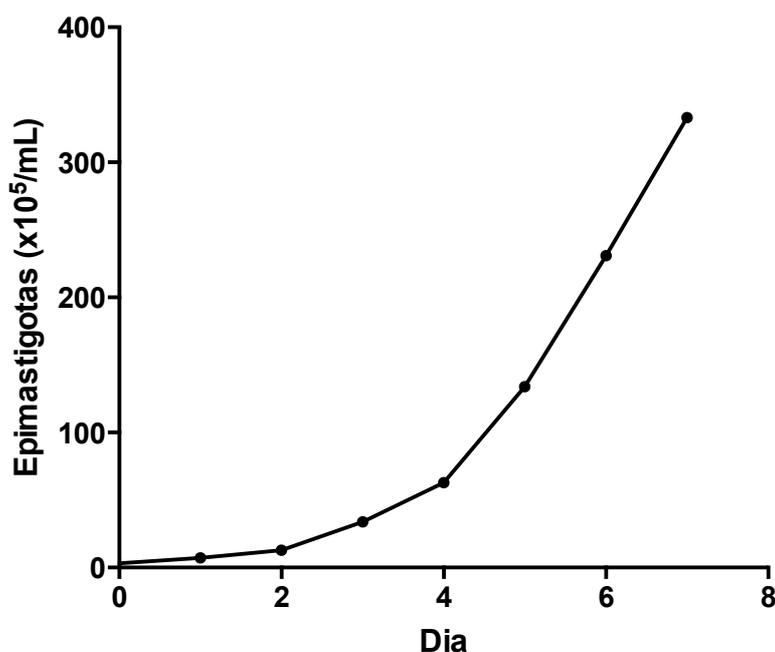


Figura 7. Curva de crescimento de formas epimastigotas da cepa G de *T. cruzi* obtida ao longo de 7 dias por contagem de células em câmara de Neubauer.

2. Genotipagem da cepa G de *T. cruzi*: amplificação por PCR do rDNA 24S α e do gene *SL*

Para confirmar se os parasitos estudados neste trabalho são pertencentes ao DTU TcI, duas amplificações por PCR a partir do DNA genômico foram realizadas, do rDNA 24S α (integrante do cistron ribossômico) e do gene *SL* ou *mini-éxon* (Figura

8). Em conformidade com a literatura sobre a genotipagem de *T. cruzi* que foi revisada por Zingales *et al.* (2012), as PCRs do gene ribossomal *24S α* da cepa G e da cepa Y (TcII) resultaram na amplificação de fragmentos de DNA de tamanho esperado de aproximadamente 110 e 125 pb, respectivamente, assim como as PCRs do gene *SL* da cepa G e da cepa Y que resultaram em fragmentos de tamanho esperado de aproximadamente 350 e 300 pb, respectivamente (Souto e Zingales, 1993; Souto *et al.*, 1996). Portanto, os parasitos *T. cruzi* utilizados nos experimentos subsequentes são membros do DTU TcI.

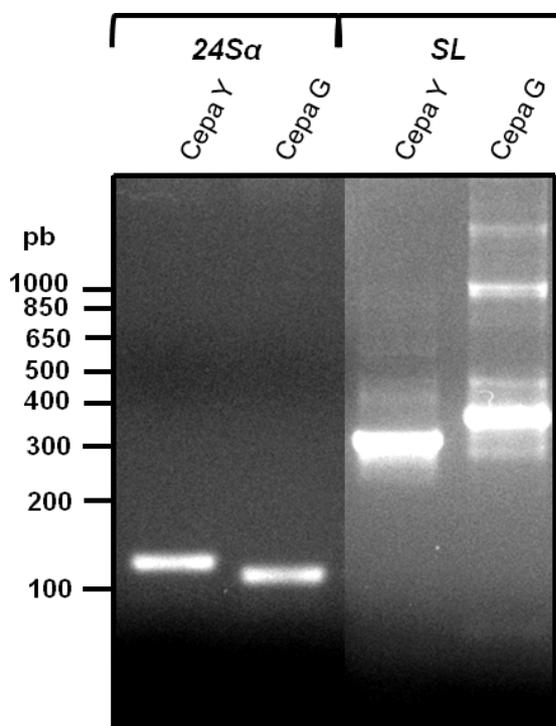


Figura 8. Análise comparativa em gel de agarose 1% da amplificação por PCR dos genes *24S α* e *SL* a partir de DNA genômico das cepas G e Y de *T. cruzi*. Os tamanhos esperados dos produtos de PCR relativos aos genes *24S α* e *SL* são: para a cepa Y (TcII), 125 pb e 300 pb, respectivamente; para a cepa G (TcI), 110 pb e 350 pb, respectivamente.

3. Preparação do extrato de proteínas expressas por epimastigotas da cepa G de *T. cruzi*

O extrato total de proteínas de formas epimastigotas da cepa G de *T. cruzi* obtido por solubilização com SDS foi submetido a SDS-PAGE 12% a diferentes quantidades, $1,25 \times 10^7$ parasitos (linha 1) e $2,5 \times 10^7$ parasitos (linha 2), e o gel foi

corado com Coomassie Blue (Figura 9A). Somente a linha 2, incluindo a parte do gel concentrador, foi utilizada para os experimentos posteriores. As faixas foram excisadas em um padrão aproximado ao ilustrado na Figura 9B. Os peptídeos gerados por digestão em gel foram analisados por LC-MS/MS com o espectrômetro de massas LTQ-Orbitrap.

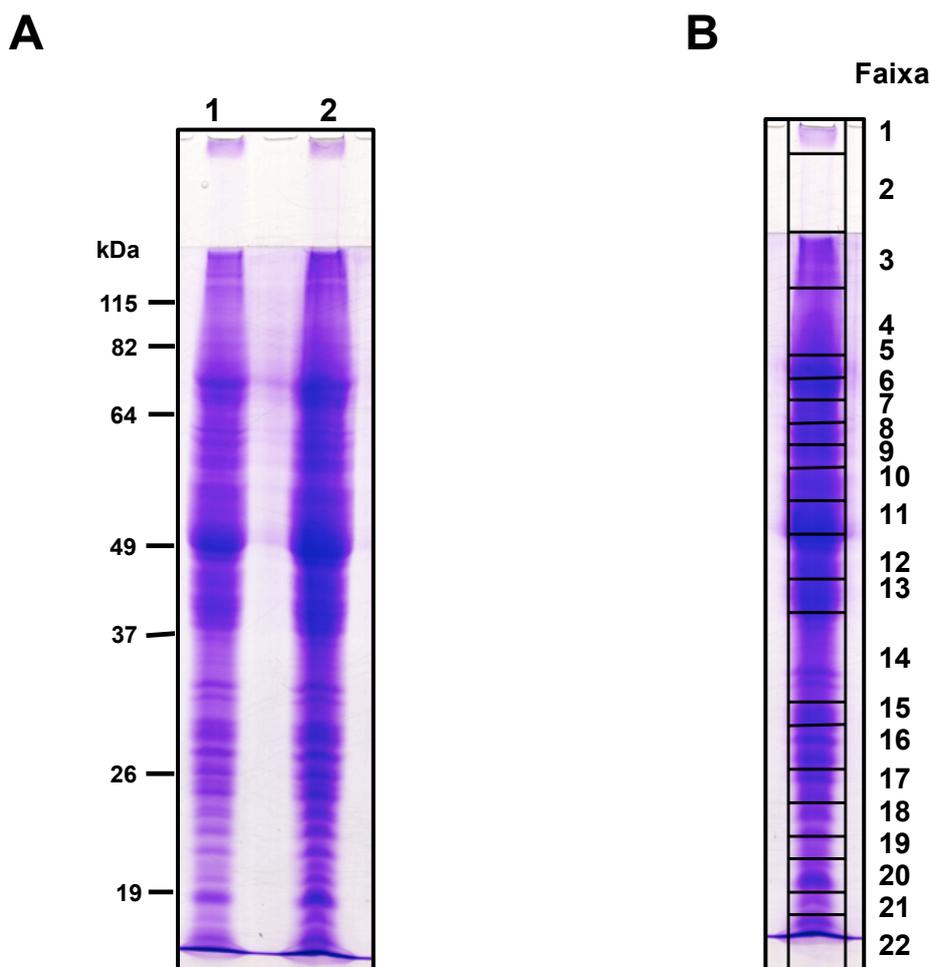


Figura 9. Perfil eletroforético do lisado proteico de formas epimastigotas da cepa G de *T. cruzi* submetidas a SDS-PAGE 12% e corado com azul de Coomassie. (A) Linha 1: $1,25 \times 10^7$ parasitos, linha 2: $2,5 \times 10^7$ parasitos. (B) Padrão de excisão horizontal aproximado das faixas de gel da linha 2 incluindo o gel concentrador, resultando em 22 faixas.

A seguir, Resultados e Discussão foi subdividido em quatro capítulos. Os três primeiros são referentes às buscas em cada um dos três bancos de dados de proteínas: I. Busca 1: banco de dados de *Trypanosoma* spp. e outros tripanossomatídeos; II. Busca 2: banco de dados da cepa CL Brener de *T. cruzi*; III. Busca 3: banco de dados das cepas CL Brener e Sylvio X10/1 de *T. cruzi*. O quarto

capítulo é relativo às seis famílias multigênicas de *T. cruzi* mais abundantes em número de genes e os respectivos membros identificados nas buscas.

**I. Busca 1: banco de dados de
Trypanosoma spp. e outros
triptanosomatídeos**

A primeira abordagem é generalista e baseou-se na busca dos espectros de massas MS/MS no banco de dados de *Trypanosoma* spp. e outros tripanossomatídeos com 88.279 sequências de proteínas. Em relação a sequências de *T. cruzi*, esse banco de dados continha 19.244 sequências da cepa CL Brener e 10.794 sequências da cepa Sylvio X10/1. Utilizando o algoritmo do programa Sequest, 1.460 proteínas foram identificadas e, dessas, 1.079 proteínas são anotadas no banco de dados como não caracterizadas ou hipotéticas. Um dos parâmetros de filtro para os resultados foi o número mínimo de 2 peptídeos para cada proteína identificada. Considerando o número de peptídeos que resultaram em cada uma das identificações, o maior percentual das identificações (23,77%) foi proveniente de 2 peptídeos e 7,40% das proteínas foram identificadas com 10 ou mais peptídeos (Figura 10). O número máximo de peptídeos por identificação foi 31, que resultou na identificação da proteína não caracterizada de *Trypanosoma vivax* (Nome de entrada ou ID: G0U6P0_TRYVY).

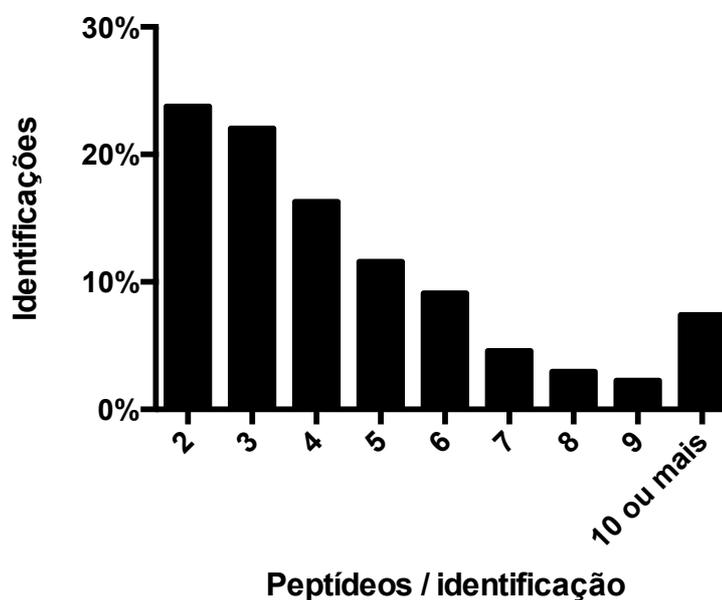


Figura 10. Relação entre o número de peptídeos por identificação e a porcentagem de identificações obtidas por meio de busca no banco de dados de *Trypanosoma* spp. e outros tripanossomatídeos.

Como o banco de dados para organismos do gênero *Trypanosoma* foi usado, identificações de múltiplos organismos foram detectadas (Tabela 1). As maiores

quantidades de identificações foram provenientes de *T. vivax* (38,90%) e *Trypanosoma congolense* (28,70%), seguidos por *T. cruzi* (12,87%). Erros de deposição no banco de dados poderiam ser responsáveis por três identificações oriundas de *L. major* e do oomiceto *Phytophthora infestans*.

Tabela 1. Relação entre o organismo, o número total de proteínas e de proteínas hipotéticas identificadas por meio de busca no banco de dados de *Trypanosoma* spp. e outros tripanossomatídeos usando o algoritmo do programa Sequest.

| Organismo | Total de proteínas identificadas | Proteínas hipotéticas |
|--------------------------------------|----------------------------------|-----------------------|
| <i>Trypanosoma vivax</i> | 568 | 437 |
| <i>Trypanosoma congolense</i> | 419 | 344 |
| <i>Trypanosoma cruzi</i> | 188 | 112 |
| <i>Trypanosoma cruzi marinkellei</i> | 144 | 105 |
| <i>Trypanosoma brucei</i> | 135 | 81 |
| <i>Trypanosoma rangeli</i> | 2 | 0 |
| <i>Trypanosoma evansi</i> | 1 | 0 |
| <i>Leishmania major</i> | 2 | 0 |
| <i>Phytophthora infestans</i> | 1 | 0 |
| Total | 1.460 | 1.079 |

1. Classificação de Gene Ontology (GO)

As proteínas identificadas foram classificadas com o emprego do programa Blast2GO por GO e agrupadas nas seguintes categorias: componente celular (Figura 11A), processo biológico (Figura 11B) e função molecular (Figura 11C). Somando essas três categorias, 1.648 anotações foram geradas e distribuídas em 579 proteínas identificadas diferentes (39,66% do total de identificações). Pelo algoritmo do Blast2GO, uma única identificação pode receber mais de uma anotação de GO.

Para componente celular, 25 termos GO foram fornecidos. Por meio de análise estatística, o termo GO célula foi o mais abundante com 26,69% do total das anotações, seguido por intracelular (19,93%), complexo proteico (11,74%), citoplasma (7,83%), núcleo (4,98%) e assim por diante (Figura 11A). Essa

classificação de GO é generalista, pois se não existe o conhecimento prévio e a informação anotada nos bancos de dados sobre a localização subcelular de proteínas identificadas, são observadas anotações pouco específicas como "célula" e, assim, a classificação aparenta ambiguidade.

Quarenta e três termos GO foram sugeridos em processo biológico para as identificações desta busca. Os termos GO mais abrangentes foram: processo metabólico (8,16%) e simbiose, abrangendo mutualismo até parasitismo (8,16%), etc. (Figura 11B). Consideram-se integrantes do grupo simbiose as proteínas relacionadas com a interação simbiótica entre dois organismos e, entre as várias formas de simbiose, estão o mutualismo, o parasitismo e o comensalismo, segundo a classificação de GO (<http://amigo.geneontology.org/cgi-bin/amigo/term-etails.cgi?term=GO:0044403>). Especificamente no parasitismo, a associação é desvantajosa ou destrutiva para um dos organismos envolvidos. As proteínas de formas epimastigotas da cepa G identificadas e classificadas nesse grupo foram principalmente TS putativas de *T. cruzi* e *T. cruzi marinkellei* (parasito que infecta morcegos) e *variant surface glycoproteins* (VSG) putativas de *T. brucei*, *T. congolense* e *T. vivax*.

A sialilação de proteínas da superfície do *T. cruzi* é devido à atividade de sialidases modificadas, as TS, pois tripanossomatídeos não são capazes de sintetizar ácido siálico. Essas enzimas, ao invés de hidrolisarem ácido siálico, transferem preferencialmente ácido siálico $\alpha(2-3)$ -ligado a partir de sialoglicoconjugados e proteínas do hospedeiro para principalmente mucinas ancoradas a GPI na superfície do *T. cruzi*. Isso confere uma carga negativa que protege os tripomastigotas extracelulares da ação de anticorpos produzidos pelo hospedeiro (Pereira-Chioccola *et al.*, 2000). Em adição ao favorecimento da sobrevivência de tripomastigotas, TS têm importância na infectividade do parasito, no reconhecimento da célula hospedeira, na modulação e escape da resposta imune, no escape dos vacúolos parasitóforos, na adesão às células epiteliais do intestino do triatomíneo, na invasão de células fagocíticas e não fagocíticas do hospedeiro (revisado por Dc-Rubin e Schenkman (2012) e Freire-De-Lima *et al.* (2012)).

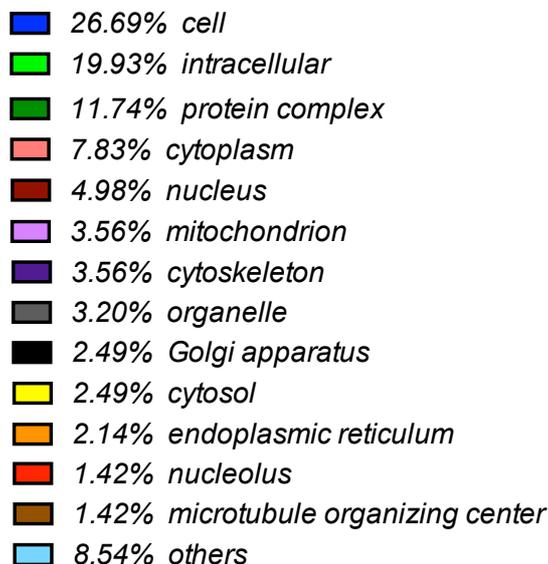
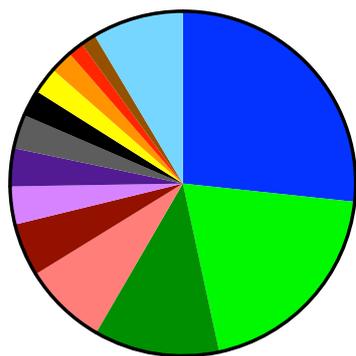
A grande representatividade das famílias de TS e mucinas no genoma de *T. cruzi* pode refletir na habilidade do parasito em evadir ou subverter o sistema imune do hospedeiro. A superfamília polimórfica TcS contém a maior quantidade de genes

entre todas as famílias multigênicas da cepa CL Brener, com um total de 1.430 membros, dos quais 693 são prováveis pseudogenes (Atwood *et al.*, 2005). Ela inclui membros das famílias de trans-sialidases e trans-sialidases-símiles. Apesar de o fato de que quatro grupos da superfamília TcS estavam bem caracterizados e apenas um dos grupos contém trans-sialidases ativas, a maioria dos genes da superfamília possui a descrição "trans-sialidases" no banco de dados do genoma de *T. cruzi* e todas as proteínas identificadas são descritas pelo Uniprot como trans-sialidases e, neste trabalho, estão referenciadas como TS. Assim, 35 TS putativas foram identificadas, todas provenientes de *T. cruzi marinkellei* e *T. cruzi*, mesmo existindo ortólogos em *T. brucei* e esses constarem no banco de dados utilizado. Recentemente, quatro grupos adicionais foram identificados após análise de alinhamento de sequências, sugerindo que essa superfamília é muito mais heterogênea do que previamente pensado (Freitas *et al.*, 2011).

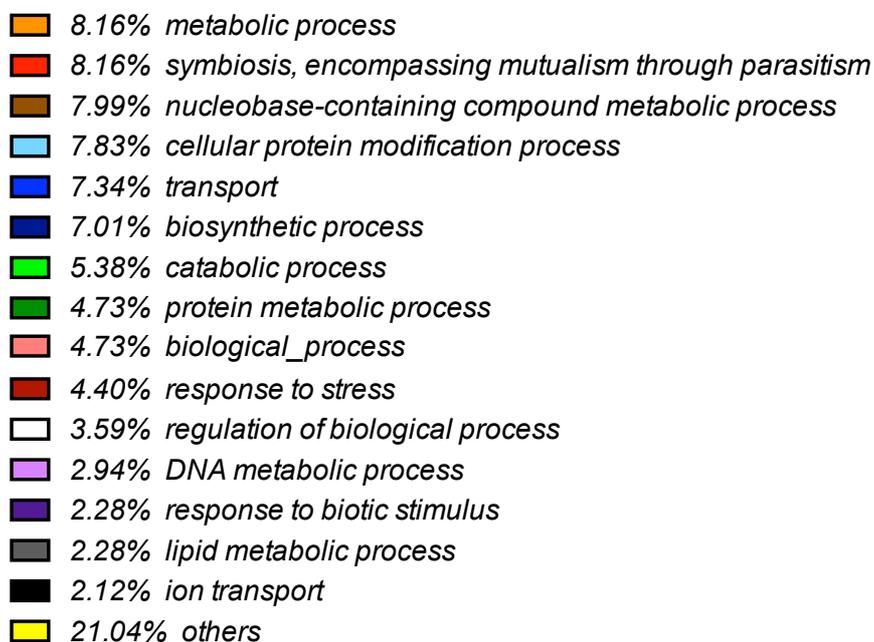
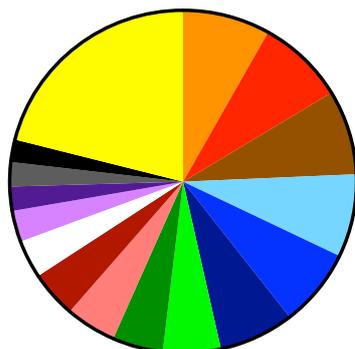
É importante destacar que, por meio de busca no banco de dados geral de *Trypanosoma* spp. e outros tripanossomatídeos, foram obtidas 52 identificações de VSG de *T. brucei*, *T. congolense* e *T. vivax*, porém a família VSG não foi descrita em *T. cruzi*. *T. brucei* tem a superfície coberta por uma capa densa de VSG nas formas metacíclicas e sanguíneas (formas *slender* e *stumpy*), o que gera variedade antigênica e favorece a evasão do sistema imune do hospedeiro mamífero (revisado por Borst e Ulbert (2001) e Matthews (2005)). Talvez possam ter ocorrido erros na análise pelo algoritmo do Sequest ou prevaleceu a correspondência entre espectros MS/MS e VSG em vez de outras proteínas devido à existência de domínios comuns. A análise da sequência genômica da cepa G poderá elucidar tal ponto.

Para função molecular, 30 termos GO foram resultantes da classificação das proteínas identificadas. O termo GO ligação a proteínas apresentou a maior porcentagem de anotações (19,10%), depois aparecem ligação a nucleotídeos (14,32%), atividade de hidrolase (12,20%), etc. (Figura 11C). O Blast2GO classifica proteínas em atividades de hidrolase e peptidase, sendo que essa última é uma atividade de hidrolase. Portanto, para as três categorias geradas pelo Blast2GO, pode existir um termo GO que é integrante de um outro termo GO mais abrangente e ambos serem considerados classificações distintas.

A



B



C

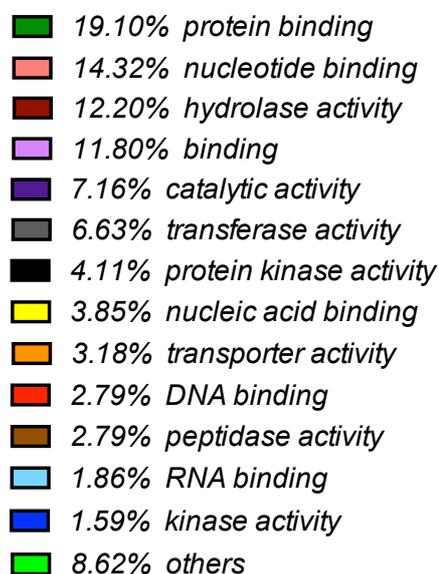
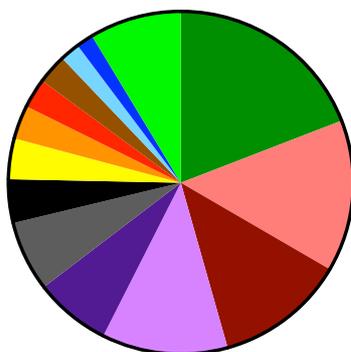


Figura 11. Gráficos de setores para cada categoria de GO: (A) componente celular, (B) processo biológico e (C) função molecular, de acordo com o programa Blast2GO.

2. Predição de atividades enzimáticas

As proteínas identificadas foram pesquisadas para atividades enzimáticas segundo a classificação da Nomenclatura de Enzimas definida pelo *Nomenclature Committee of the International Union of Biochemistry and Molecular Biology* (NC-IUBMB). As enzimas são classificadas em seis classes de acordo com a reação que catalisam (EC 1 a 6) e são atribuídos números de *Enzyme Commission* (EC) a elas. Assim como para anotações de GO, pela classificação do Blast2GO, uma mesma identificação pode receber mais de uma anotação de EC. Foram anotadas 153 proteínas para EC e, de um total de 163 anotações, as classes mais representativas foram EC 2 (transferases) e EC 3 (hidrolases) com 64 e 63 anotações, respectivamente. Essa grande quantidade de anotações de hidrolases está condizente com a alta porcentagem de anotações para o termo GO de função molecular atividade de hidrolase. EC 3 também foi a classe mais diversificada, com 31 atividades enzimáticas diferentes (Tabela 2).

De forma global, as atividades enzimáticas mais numerosas em relação a anotações de todos os seis ECs foram: proteína serino/treonino-quinase (EC 2.7.11.0) com 26 anotações e exo- α -sialidase (EC 3.2.1.18) com 22 anotações. As proteínas quinases são mediadores em vias de transdução de sinal, transmitindo sinais ambientais (extracelulares) e coordenando processos intracelulares. As proteínas quinases eucarióticas são classificadas em duas superfamílias: proteínas Ser/Thr-quinases e proteínas Tyr-quinases. As proteínas Ser/Thr-quinases são enzimas que catalisam a fosforilação de resíduos de Ser ou Thr de proteínas alvo com a utilização de ATP (revisado por Edelman, Blumenthal e Krebs (1987)), o que induz uma mudança conformacional na proteína alvo. Em *T. cruzi* e outros tripanossomatídeos, a fosforilação de proteínas provavelmente está envolvida em mecanismos regulatórios de processos biológicos como ciclo celular, diferenciação e resposta ao estresse durante os ciclos de vida complexos deles (revisado por Flawia, Tellez-Inon e Torres (1997) e Naula, Parsons e Mottram (2005)). A grande representatividade do grupo de atividade de exo- α -sialidases é devido a TS putativas identificadas (17 identificações atribuídas a esse grupo), citadas anteriormente.

Tabela 2. Abundância de anotações e de atividades enzimáticas para cada uma das seis classes de *Enzyme Commission* (EC).

| Classe de EC | Anotações ^a | Atividades enzimáticas ^b |
|---------------------|------------------------|-------------------------------------|
| 1 - Óxido-redutases | 12 | 9 |
| 2 - Transferases | 64 | 31 |
| 3 - Hidrolases | 63 | 25 |
| 4 - Liases | 9 | 6 |
| 5 - Isomerases | 7 | 5 |
| 6 - Ligases | 8 | 4 |
| Total | 163 | 80 |

^a Número de anotações de EC pelo Blast2GO

^b Número de *Enzyme* ID distintos do sistema EC

2.1. Peptidases (EC 3.4) putativas

As enzimas que catalisam a hidrólise de ligações peptídicas de polipeptídeos ou proteínas são denominadas peptidases ou proteases. Tem se tornado claro que as peptidases de tripanossomatídeos patogênicos desempenham um papel essencial em várias etapas da infecção do hospedeiro: absorção, penetração, sobrevivência intracelular, replicação, diferenciação, infectividade, evasão do sistema imune, nutrição, autofagia (revisado por Vermelho *et al.* (2007), Alvarez, Niemirowicz e Cazzulo (2012) e Bastos *et al.* (2013)).

Dois sistemas de classificação e nomenclatura são utilizados para peptidases, EC já citado e MEROPS. MEROPS (<http://merops.sanger.ac.uk>) é um banco de dados sobre as peptidases, os inibidores e os substratos delas e usa uma classificação hierárquica baseada na estrutura das peptidases (Rawlings, Barrett e Bateman, 2012). Cada peptidase é atribuída a uma família com base nas similaridades de sequência de aminoácidos estatisticamente significativas, e as famílias em que se acredita serem homólogas são agrupadas em um mesmo clã. Na versão atual do MEROPS (*release* 9.9), as peptidases são classificadas em nove famílias de acordo com o tipo de mecanismo catalítico (aspártico (A), cisteíno (C), glutâmico (G), metalo (M), asparagino (N), serino (S), treonino (T), misto (P), desconhecido (U)).

O Blast2GO utiliza o sistema EC e, por isso, as identificações da cepa G de *T. cruzi* foram pesquisadas e classificadas segundo essa nomenclatura. As peptidases (EC 3.4) que atuam na extremidade N-terminal ou C-terminal são as exopeptidases (EC 3.4.11-19) e as que clivam dentro da cadeia polipeptídica são as

endopeptidases (EC 3.4.21-25 e EC 3.4.99). Em relação à subclasse de peptidases, um total de dez identificações foram atribuídas a cinco sub-subclasses: aminopeptidases (EC 3.4.11.0), serino-carboxipeptidases (EC 3.4.16.0), metalo-carboxipeptidases (EC 3.4.17.0), serino-endopeptidases (EC 3.4.21.0) e metalo-endopeptidases (EC 3.4.24.0). As anotações corresponderam às seguintes proteínas identificadas: uma leucil-aminopeptidase (LAP) citosólica putativa de *T. congolense*, uma aminopeptidase putativa de *T. vivax*, uma serino-carboxipeptidase putativa (CBP1) de *T. vivax*, duas carboxipeptidases dependentes de zinco putativas de *T. vivax*, duas proteínas de superfície putativas de *T. vivax*, uma gp63 putativa de *T. cruzi*, uma proteína não caracterizada de *T. congolense* e uma proteína não caracterizada de *T. vivax* (Tabela 3).

Tabela 3. Proteínas anotadas para a atividade de peptidase (EC 3.4) distribuídas em cinco sub-subclasses de EC.

| Sub-subclasse de EC | Descrição ^a | Código identificador ^b |
|--|--|-----------------------------------|
| Aminopeptidases (EC 3.4.11) | <i>Aminopeptidase</i> | G0TTF8_TRYVY |
| | <i>Cytosolic leucyl aminopeptidase</i> | G0URS4_TRYCI |
| Serino-carboxipeptidases (EC 3.4.16) | <i>Serine carboxypeptidase (CBP1)</i> | G0U597_TRYVY |
| Metallo-carboxipeptidases (EC 3.4.17) | <i>Zinc carboxypeptidase</i> | G0TU56_TRYVY |
| | <i>Zinc carboxypeptidase</i> | G0UCS1_TRYVY |
| Serino-endopeptidases (EC 3.4.21) | <i>Uncharacterized protein</i> | F9W6B7_TRYCI |
| | <i>Uncharacterized protein</i> | G0TU13_TRYVY |
| Metallo-endopeptidases (EC 3.4.24) | <i>Major surface protease A</i> | G0UBW9_TRYVY |
| | <i>p63-1 surface protease homolog</i> | G0UC84_TRYVY |
| | <i>Surface protease gp63</i> | K4DJ96_TRYCR |

^a Descrição da proteína anotada proveniente do Uniprot

^b Nome de entrada ou ID do Uniprot

2.2. Predição de enzimas envolvidas em vias metabólicas

As proteínas identificadas foram anotadas por KEGG *pathways* - Blast2GO

em 45 vias metabólicas, das quais as dez mais representativas em relação ao número de anotações por via estão listadas na Tabela 4. A via metabólica de metabolismo de purinas apresentou a maior variedade de atividades enzimáticas (*Enzyme ID* do sistema EC) com 12 atividades enzimáticas distintas, seguida por metabolismo de pirimidinas com 6. Formas epimastigotas de *T. cruzi* são replicativas e, por isso, realizam uma alta taxa de síntese de nucleotídeos para a propagação dos ácidos nucleicos para as células filhas após a fissão binária. Assim, a alta frequência de anotações nas vias metabólicas de síntese de nucleotídeos seria uma consequência disso. Na Tabela 5, estão detalhadas as atividades enzimáticas e os nomes de entrada relativos às cinco vias metabólicas mais abundantes em número de anotações.

Tabela 4. As dez vias metabólicas mais abundantes em número de anotações e os respectivos números de atividades enzimáticas.

| Via metabólica ^a | Anotações ^b | Atividades enzimáticas ^c |
|---|------------------------|-------------------------------------|
| Metabolismo de purinas | 25 | 12 |
| Degradação de glicanas | 22 | 1 |
| Metabolismo de esfingolipídeos | 22 | 1 |
| Metabolismo de pirimidinas | 10 | 6 |
| Metabolismo de tiaminas | 6 | 1 |
| Metabolismo de glicerofosfolipídeos | 5 | 3 |
| Metabolismo de drogas - outras enzimas | 3 | 3 |
| Vias de sinalização de receptor de célula T | 3 | 1 |
| Glicólise/Gliconeogênese | 2 | 2 |
| Metabolismo de ácidos graxos | 2 | 2 |

^aVia metabólica sugerida por KEGG *pathways* - Blast2GO

^bNúmero de anotações (pode haver uma proteína anotada mais de uma vez na mesma via metabólica)

^cNúmero de *Enzyme ID* distintos do sistema EC

2.2.1. Metabolismo de purinas

As purinas são essenciais para a síntese de nucleotídeos. Distintamente dos hospedeiros vertebrados, os protozoários parasitos que têm sido examinados são auxotróficos para purino-nucleotídeos, ou seja, são incapazes de realizar a síntese *de novo* das purinas. Assim, são dependentes da via de recuperação das purinas

pré-formadas do hospedeiro ou do meio (Gutteridge e Gaborak, 1979; Ullman e Carter, 1997) ilustrada na Figura 12.

A PRPP sintetase (também denominada ribose fosfato pirofosfoquinase) participa no início da via de síntese *de novo* das purinas, que conduz à síntese de AMP e GMP. Essa enzima catalisa a transformação de ribose-5-fosfato em PRPP (fosforribosil pirofosfato) na presença de ATP. Em razão da necessidade de PRPP como doador de ribose-fosfato para a recuperação de adenina, guanina e hipoxantina, essa enzima também está envolvida nesta via. A PRPP sintetase de *Leishmania donovani* foi caracterizada molecularmente (Hendrickson, Allen e Ullman, 1993). A identificação de uma PRPP sintetase putativa (Q57ZX2_TRYB2) foi constatada na cepa G de *T. cruzi* por busca no banco de dados *Trypanosoma* spp. e outros tripanossomatídeos.

Uma enzima hipoxantina-guanina fosforribosil transferase putativa (HGPRT, G0U5D5_TRYVY) foi identificada. Nos gêneros *Trypanosoma* e *Leishmania*, HGPRT catalisa as reações de recuperação de guanina e de hipoxantina livres na presença de PRPP, produzindo GMP e IMP, respectivamente (Ullman e Carter, 1997). Foi publicado um estudo envolvendo mutagênese de dois resíduos de aminoácidos do *loop* do sítio ativo da HGPRT de *T. cruzi* que são conservados entre as HGPRT, Leu67 e Gly69 (Butterworth *et al.*, 2004).

Na via de recuperação das purinas, há enzimas que convertem IMP e XMP, os respectivos produtos da HGPRT e da XPRT (xantina fosforribosil transferase), em AMP e GMP. Também existem enzimas que catalisam a reação inversa de conversão de AMP e GMP em IMP. Entre as referidas enzimas de interconversão estão: adenilsuccinato liase, adenilsuccinato sintetase, IMP desidrogenase, GMP sintetase, AMP desaminase, GMP redutase.

Na via biossintética *de novo* das purinas, a enzima bifuncional adenilsuccinato liase catalisa a conversão de succinil aminoimidazol carboxamida ribonucleotídeo (SAICAR) em aminoimidazol carboxamida ribonucleotídeo (AICAR) e fumarato e a conversão de adenilsuccinato em AMP e fumarato (Toth e Yeates, 2000). Em *T. cruzi* e outros protozoários parasitos, a adenilosuccinato liase catalisa apenas a última reação supracitada porque somente a via de recuperação de purinas está presente. A adenilsuccinato liase de *T. cruzi* foi purificada e bioquimicamente caracterizada (Spector, Berens e Marr, 1982). Nesse contexto, uma adenilsuccinato liase putativa (G0U2G5_TRYVY) foi identificada na cepa G.

Enzimas globais nos processos de replicação do DNA e de transcrição como as respectivas DNA polimerases e RNA polimerases foram classificadas no grupo de metabolismo de purinas, assim como no grupo de metabolismo de pirimidinas que será exposto no próximo tópico. DNA polimerases são atualmente classificadas em sete famílias, A, B, C, D, X, Y e RT. Em *T. cruzi*, são preditas sequências gênicas que codificam membros pertencentes a quatro famílias de DNA polimerases, A, B, X e Y (El-Sayed *et al.*, 2005). No proteoma total de formas epimastigotas da cepa Brazil, foram identificadas duas DNA polimerases β mitocondriais, membros da família X (Atwood *et al.*, 2005). No presente proteoma de formas epimastigotas da cepa G, duas DNA polimerases η putativas (G0UVH1_TRYCI e G0U5H0_TRYVY) foram detectadas. As DNA polimerases η e κ , integrantes da família Y, atuam em uma via de reparo alternativa do DNA chamada de síntese translesão (TLS - *translesion synthesis*), na qual são capazes de processar lesões do DNA permitindo que a replicação da fita de DNA prossiga. As DNA polimerases η e κ de *T. cruzi* já foram estudadas (De Moura *et al.*, 2009; Rajao *et al.*, 2009). Parasitos superexpressando a TcPol η foram mais resistentes a tratamento com peróxido de hidrogênio, que causa danos oxidativos ao DNA, comparados a células não transfectadas. Porém não restaurou lesões em parasitos transfectados após tratamento com os agentes genotóxicos luz ultravioleta e cisplatina (De Moura *et al.*, 2009). Além disso, duas identificações relativas a RNA polimerases putativas também foram constatadas, uma RNA polimerase II dirigida a DNA (G0UY95_TRYCI) e uma subunidade RPB8 da RNA polimerase B (G0UC34_TRYVY).

Adenilato ciclases e o produto de sua catálise, o mensageiro secundário cAMP, têm sido envolvidos na regulação da metaciclologênese de *T. cruzi* em resposta a ligantes peptídicos exógenos (Gonzales-Perdomo, Romero e Goldenberg, 1988; Fraidenaich *et al.*, 1993). Nesse protozoário, essas proteínas são codificadas por uma grande família gênica polimórfica e parecem possuir um grande domínio extracelular N-terminal variável, uma única hélice transmembrânica e um domínio catalítico citosólico altamente conservado (Taylor *et al.*, 1999). Uma adenilato ciclase de *T. cruzi* estudada é ativada por Ca^{2+} e interage com a proteína paraflagelar *rod*, um componente da estrutura flagelar de formas epimastigotas e tripomastigotas (D'angelo *et al.*, 2002). Quatro adenilato ciclases foram relacionadas na cepa G (Q57WZ8_TRYB2, Q4D322_TRYCC, G0U4Z9_TRYVY e

G0TUP8_TRYVY).

Assim, purinas são de vital importância para a síntese de ácidos nucleicos, moléculas regulatórias, co-fatores enzimáticos, assim como podem atuar como transportadores de energia. Uma grande quantidade de "motores" moleculares utilizam a energia química do ATP para exercer trabalho mecânico e nesse grupo estão incluídas a família de AAA+ ATPases (*extended ATPases associated with various cellular activities*), as superfamílias de dineína, miosina e cinesina. As AAA+ ATPases são responsáveis por uma ampla variedade de processos celulares e tem como característica um módulo de ATPase de 230 aminoácidos, composto pelos subdomínios de ligação a nucleotídeo ASCE, α -helicoidal e C-terminal (Wendler *et al.*, 2012). Dineínas são proteínas motoras dos microtúbulos e apresentam várias cadeias pesadas, intermediárias e leves (Cho e Vale, 2012). Na cepa G de *T. cruzi*, identificamos duas AAA ATPases putativas (G0V0C1_TRYCI e G0TX46_TRYVY) e uma cadeia pesada de dineína (Q4CKA4_TRYCC).

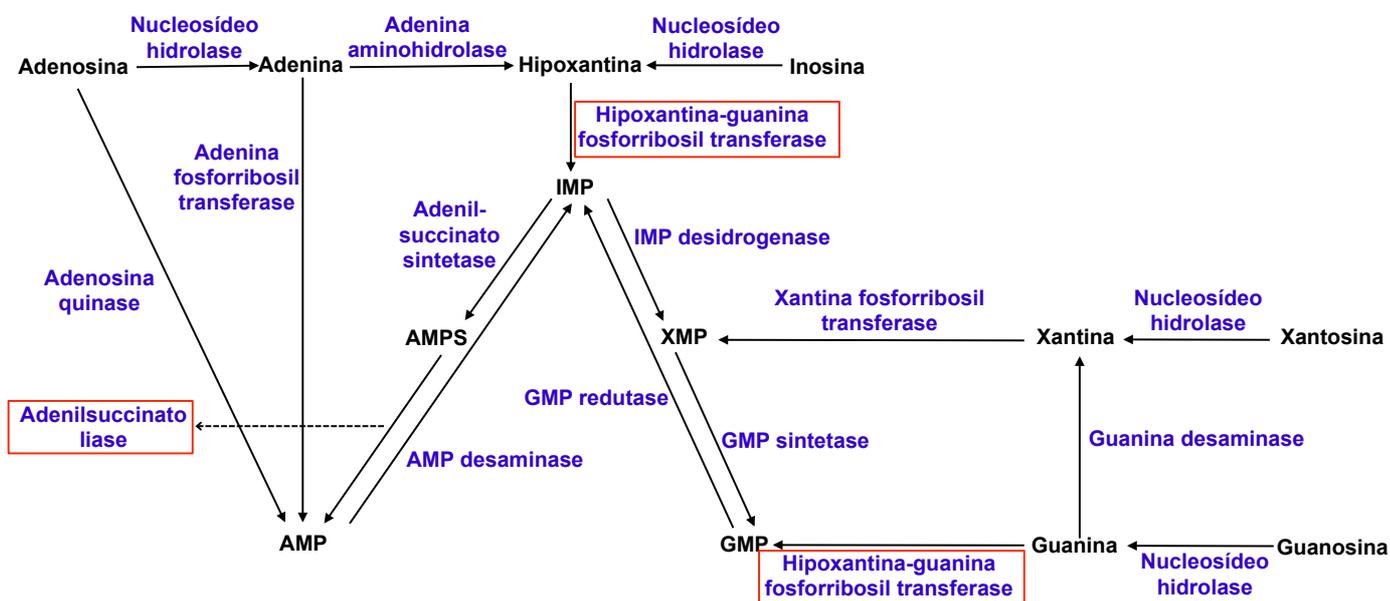


Figura 12. Diagrama da via de recuperação/salvação das purinas predita em *Trypanosoma spp.*, com base nas enzimas descritas para *Leishmania spp.* As enzimas marcadas em vermelho correspondem às proteínas identificadas na cepa G de *T. cruzi* pela busca 1. Modificado de Boitz *et al.* (2013). AMP: adenilato, AMPS: adenilsuccinato, GMP: guanilato, IMP: inosinato, XMP: xantinilato.

2.2.2. Metabolismo de pirimidinas

Há vias *de novo* e de recuperação de nucleotídeos pirimidínicos em *T. cruzi*, a partir das quais UTP, CTP e TTP são sintetizados. Na síntese *de novo* das pirimidinas UTP e CTP, inicialmente, a biossíntese de carbamil fosfato é feita pela enzima carbamil fosfato sintetase II e depois o anel piridímico é formado a partir de carbamil fosfato e aspartato pela enzima aspartato transcarbamilase (Figura 13). Apenas uma aspartato transcarbamilase putativa G0TW09_TRYVY foi identificada e atribuída a esta via metabólica. A forma recombinante dessa enzima de *T. cruzi* foi cristalizada (Matoba *et al.*, 2009). As três enzimas iniciais envolvidas na síntese *de novo* de pirimidinas em *T. cruzi*, carbamil fosfato sintetase II, aspartato transcarbamilase e diidroorotase, foram expressas em *Escherichia coli* e interagem molecularmente entre si, formando um complexo terciário (Nara *et al.*, 2012).

Por estarem relacionadas aos processos de replicação do DNA e transcrição, as duas DNA polimerases η putativas (G0UVH1_TRYCI e G0U5H0_TRYVY), a RNA polimerase II dirigida a DNA putativa (G0UY95_TRYCI) e a subunidade RPB8 da RNA polimerase B putativa (G0UC34_TRYVY) que foram encaixadas na via de metabolismo de purinas também foram atribuídas à via de metabolismo de pirimidinas.

A timidina quinase putativa (F9WAY4_TRYCI), que participa da síntese de DNA e, na presença de ATP, fosforila desoxitimidina, foi atribuída à via de metabolismo de pirimidinas.

Os RNAs de transferência (tRNAs) sofrem processamento pós-transcricional e, na etapa final, ocorrem modificações de algumas bases como a isomerização de uridina em pseudo-uridina. Identificamos uma tRNA pseudo-uridina sintase A putativa (G0U6Q1_TRYVY) que catalisa essa reação e também foi anotada nesta via.

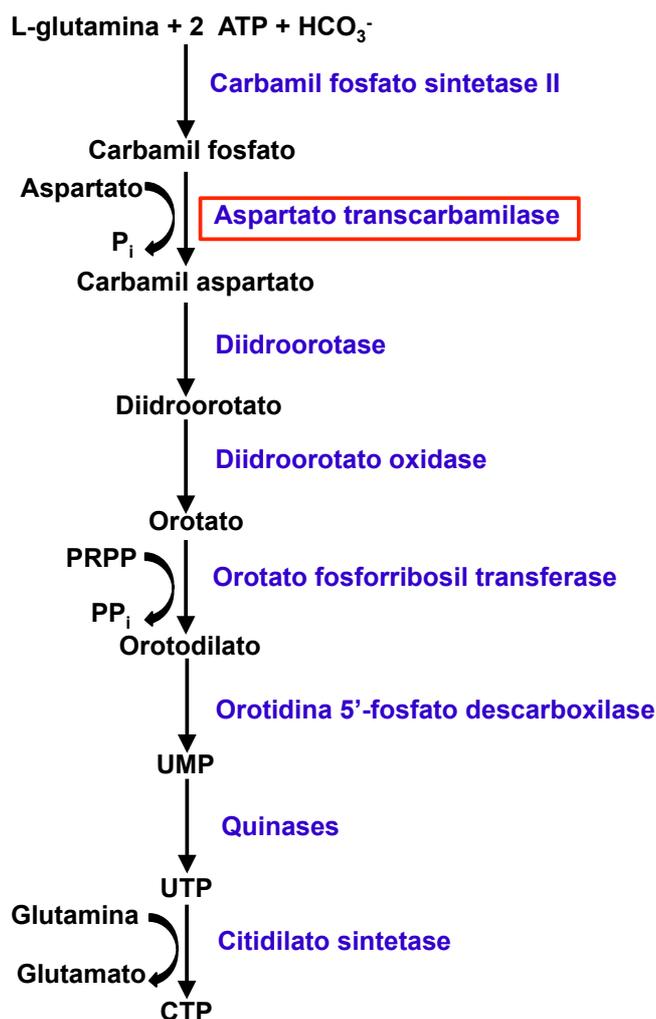


Figura 13. Esquema simplificado da via de síntese *de novo* das pirimidinas. A enzima marcada em vermelho corresponde à proteína identificada na cepa G de *T. cruzi*. HCO_3^- : bicarbonato, PRPP: fosforibosilpirofosfato, P_i : fosfato inorgânico, PP_i : pirofosfato.

2.2.3. Degradação de glicanas

Relacionadas à via metabólica de degradação de glicanas estão 17 identificações de TS putativas que promovem a remoção de ácido siálico de glicoconjugados do hospedeiro. A atividade enzimática associada a essa via é neuraminidase ou exo- α -sialidase (EC 3.2.1.18), já citada anteriormente como o segundo grupo de atividade enzimática mais numeroso em relação a anotações.

2.2.4. Metabolismo de esfingolipídeos

As mesmas anotações com atividade de neuraminidase para a via de degradação de glicanas estão associadas à via de metabolismo de esfingolipídeos.

2.2.5. Metabolismo de tiaminas (vitamina B₁)

Seis enzimas com atividade predita de fosfatase foram atribuídas à via de metabolismo de tiaminas, das quais duas são AAA ATPases putativas (G0V0C1_TRYCI e G0TX46_TRYVY). Vitamina B₁, na forma de tiamina pirofosfato (TPP), é necessária para as vias metabólicas de fosforilação oxidativa e das pentoses-fosfato, pois age como cofator para α -cetoácido desidrogenases (Hohmann e Meacock, 1998). O envolvimento de ATPases no metabolismo de tiaminas seria no transporte dessas moléculas para o interior dos parasitos.

Tabela 5. Relação entre as vias metabólicas mais representativas em número de anotações e as respectivas atividades enzimáticas. (continua)

| Via metabólica ^a | Atividade enzimática ^b | Código identificador ^c |
|-----------------------------------|---|--|
| | Fosforribosil transferase (EC 2.4.2.8) | G0U5D5_TRYVY ^d |
| | Fosforribosil transferase (EC 2.4.2.22) | G0U5D5_TRYVY ^d |
| | Quinase (EC 2.7.4.6) | G0U704_TRYVY |
| | Difosfoquinase (EC 2.7.6.1) | Q57ZX2_TRYB2 |
| | RNA polimerase (EC 2.7.7.6) | K2MYR2_TRYCR, G0UY95_TRYCI, G0UC34_TRYVY |
| 1) Metabolismo de purinas | DNA polimerase (EC 2.7.7.7) | F9W6J4_TRYCI, G0UVH1_TRYCI, G0U5H0_TRYVY |
| | Fosfodiesterase (EC 3.1.4.17) | F9W8G7_TRYCI |
| | Nucleosidase (EC 3.2.2.2) | G0UQA6_TRYCI |
| | Adenil pirofosfatase (EC 3.6.1.3) | Q4CKA4_TRYCC, F9WEX2_TRYCI |
| | Fosfatase (EC 3.6.1.15) | G0UJN2_TRYCI, F9W3N5_TRYCI, G0V2J6_TRYCI, G0V0C1_TRYCI, G0TX46_TRYVY, G0TYA2_TRYVY |
| | Liase (EC 4.3.2.2) | G0U2G5_TRYVY |
| | Ciclase (EC 4.6.1.1) | Q57WZ8_TRYB2, Q4D322_TRYCC, G0U4Z9_TRYVY, G0TUP8_TRYVY |
| 2) Degradação de glicanas | | K2NWD0_TRYCR, K2MLX8_TRYCR, K2M785_TRYCR, K2MWI6_TRYCR, K2N4I7_TRYCR, Q4DR51_TRYCC, Q4DSY7_TRYCC, Q4E222_TRYCC, Q4D349_TRYCC, Q4DAK1_TRYCC, Q4DGS3_TRYCC, Q4CKL4_TRYCC, |
| 3) Metabolismo de esfingolipídeos | Neuraminidase (EC 3.2.1.18) | Q4DYT7_TRYCC, Q4W4T2_TRYRA, F9W5R1_TRYCI, F9WJ72_TRYCI, K4DQI5_TRYCR, K4EAY0_TRYCR, K4DSK4_TRYCR, K4DJC1_TRYCR, K4DWC7_TRYCR, K4DM44_TRYCR |
| 4) Metabolismo de pirimidinas | Carbamil transferase (EC 2.1.3.2) | G0TW09_TRYVY |
| | Quinase (EC 2.7.1.21) | F9WAY4_TRYCI |
| | Quinase (EC 2.7.4.6) | G0U704_TRYVY |
| | RNA polimerase (EC 2.7.7.6) | K2MYR2_TRYCR, G0UY95_TRYCI, G0UC34_TRYVY |

Tabela 5. Relação entre as vias metabólicas mais representativas em número de anotações e as respectivas atividades enzimáticas (conclusão)

| Via metabólica ^a | Atividade enzimática ^b | Código identificador ^c |
|--|--|--|
| 4) Metabolismo de pirimidinas (continuação) | DNA polimerase (EC 2.7.7.7) Sintase (EC 4.2.1.70) | F9W6J4_TRYCI, G0UVH1_TRYCI, G0U5H0_TRYVY G0U6Q1_TRYVY |
| 5) Metabolismo de tiaminas | Fosfatase (EC 3.6.1.15) | G0UJN2_TRYCI, F9W3N5_TRYCI, G0V2J6_TRYCI, G0V0C1_TRYCI, G0TX46_TRYVY, G0TYA2_TRYVY |

^a Via metabólica sugerida por KEGG - Blast2GO

^b *Enzyme* ID do sistema EC anotado na via metabólica

^c Nome de entrada ou ID do Uniprot

^d Identificação anotada com dois *Enzyme* ID na mesma via metabólica

3. Predição de proteínas secretadas

Na via de secreção clássica, as proteínas que contêm peptídeo-sinal, localizado na extremidade N-terminal ou internamente na proteína, são dirigidas à membrana plasmática ou ao meio extracelular por um sistema de endomembranas a partir do retículo endoplasmático (RE). Essa via de secreção convencional compreende a seguinte rota intracelular: RE - aparato de Golgi - TGN (*trans-Golgi network*) - membrana plasmática e é altamente conservada entre eucariotos (revisado por Lee *et al.* (2004)).

Embora o sistema RE-Golgi seja um processo de exportação de proteínas extremamente eficiente, existe o transporte por via não clássica. Proteínas que não contêm peptídeo-sinal podem ser diretamente direcionadas à membrana plasmática de modo independente do RE-Golgi e até mesmo algumas proteínas portadoras de sequência sinalizadora podem ser secretadas independentemente do aparato de Golgi. Quatro mecanismos têm sido propostos para vias secretórias não convencionais: (1) translocação direta a partir do citoplasma através da membrana plasmática para o espaço extracelular, (2) secreção lisossomal, (3) secreção por exossomos derivados de corpos multivesiculares e (4) secreção por "bolhas" e por brotamento de vesículas, diretamente da membrana plasmática (revisado por Nickel e Seedorf (2008) e Nickel e Rabouille (2009)).

Neste trabalho, para secreção de proteínas por via clássica preditas pelo programa SignalP, entre proteínas putativas e não caracterizadas, foram anotadas 149 proteínas (10,21%), e para secreção de proteínas por via não clássica preditas pelo SecretomeP, 978 proteínas (66,99%) (Figura 14). As proteínas de cada uma das vias são mutuamente exclusivas, ou seja, as proteínas que são atribuídas à via de secreção clássica não são anotadas na via não clássica e vice-versa. Essa predominância de proteínas preditas que seguem pela via de secreção não clássica está de acordo com trabalhos de outros tripanossomatídeos, como promastigotas de *L. braziliensis* (Cuervo *et al.*, 2009), formas sanguíneas de *T. congolense* (Grebaut *et al.*, 2009) e tripomastigotas da cepa Y de *T. cruzi* (Queiroz, 2013).

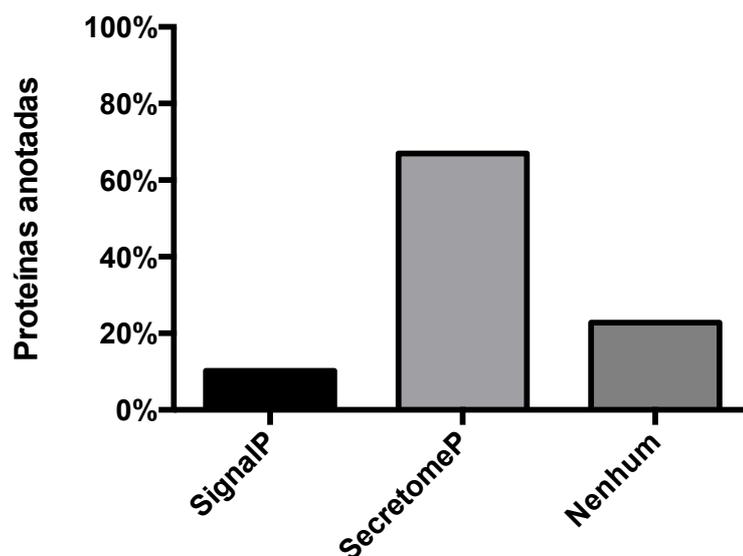


Figura 14. Predição de proteínas secretadas pelas vias clássica (SignalP) e não clássica (SecretomeP) e não anotadas como secretadas (nenhum).

O SignalP predisse peptídeo sinal para 19 tipos de proteínas putativas, de acordo com a descrição do Uniprot, listadas na Tabela suplementar 1, dos quais os mais representativos foram: VSG (29 membros), MASP (11 membros), mucinas TcMUCII (5 membros) e TS (4 membros). Para proteínas secretadas por via não clássica pelo SecretomeP, 135 tipos de proteínas putativas foram preditas e os mais abundantes foram: TS (31 membros), VSG (23 membros) e *dispersed gene family protein 1* (DGF-1) (6 membros) (Tabela suplementar 2). Considerando que as identificações dos dois programas de predição são mutuamente exclusivas, os seguintes tipos de proteínas foram enquadrados em ambas as vias de secreção de proteínas: MASP, mucina TcMUCII, mucina-símile, receptor adenilato ciclase, proteína *retrotransposon hot spot* (RHS), proteína específica de *T. brucei* spp., TS e VSG.

3.1. Predição de proteínas integrais de membrana

Como citado, as proteínas secretadas podem seguir para a membrana plasmática ou para o meio extracelular. Na membrana, as proteínas integrais deslocam-se entre os lipídeos, mantidas pelas interações hidrofóbicas com as cadeias não-polares dos aminoácidos delas. Com a utilização do programa TMHMM, 302 proteínas identificadas (20,68%) foram consideradas candidatas a possuírem

domínios transmembrânicos, variando de 1 a 27 hélices transmembrânicas por identificação.

Estabelecendo uma comparação entre a lista de proteínas preditas com domínios transmembrânicos e a lista de proteínas preditas serem secretadas por via clássica, 49 proteínas pertencem a ambas as listas. Das 302 proteínas preditas com domínios transmembrânicos, 227 proteínas são preditas serem secretadas por via não clássica. Portanto, 75,17% das proteínas integrais de membrana preditas seriam secretadas por via não clássica e, apenas 16,23%, por via clássica (Figura 15).

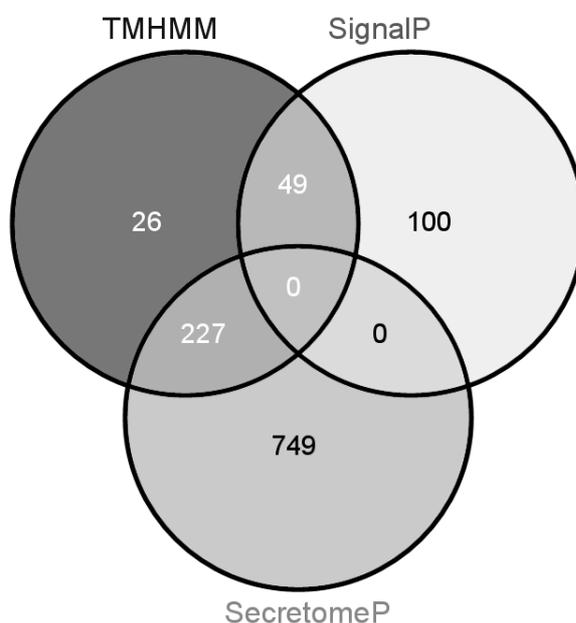


Figura 15. Diagrama de Venn correlacionado as proteínas com preditas hélices transmembrânicas (TMHMM) e as preditas a serem secretadas pelas vias clássica (SignalP) e não-clássica (SecretomeP).

**II. Busca 2: banco de dados da cepa CL
Brener de *T. cruzi***

A segunda abordagem para busca dos espectros de massas MS/MS fornecidos pelo LTQ-Orbitrap é mais restrita, pois se fez uso do banco de dados de *T. cruzi* CL Brener com 19.244 sequências de proteínas, visto que a montagem de genoma de *T. cruzi* com a melhor qualidade até hoje é a da cepa CL Brener, mesmo se tratando de uma cepa híbrida. Com a utilização do programa Sequest, 953 proteínas foram identificadas, das quais 385 são anotadas no banco de dados como proteínas hipotéticas. Assim como para a busca no banco de dados de *Trypanosoma* spp. e outros tripanosomatídeos, o número mínimo de peptídeos para cada identificação considerado foi 2. A maioria das identificações (38,41%) foi resultante de 2 peptídeos, seguido por 21,09% provenientes de 3 peptídeos, 10,28% de 4 peptídeos e assim por diante. Apenas 8,50% das proteínas foram identificadas com 10 ou mais peptídeos, percentual próximo ao da primeira busca (7,40%) (Figura 16). Contrastando com as identificações obtidas a partir do banco de dados de *Trypanosoma* spp. e outros tripanosomatídeos, uma porcentagem maior de proteínas foram identificadas com apenas 2 peptídeos a partir do banco de dados de *T. cruzi* CL Brener. As proteínas putativas *heat shock protein 85* (Q4CQS6) e *heat shock protein 70 kDa* (Q4CVR9) apresentaram correspondência a 43 peptídeos, o maior número de peptídeos por identificação obtido por esta busca, que foi superior ao da busca no banco de dados de *Trypanosoma* spp. e outros tripanosomatídeos (31 peptídeos). Todos os nomes de entrada (ID) dessa busca terminam em _TRYCC, mas a seguir serão apresentadas como números de acesso (forma simplificada).

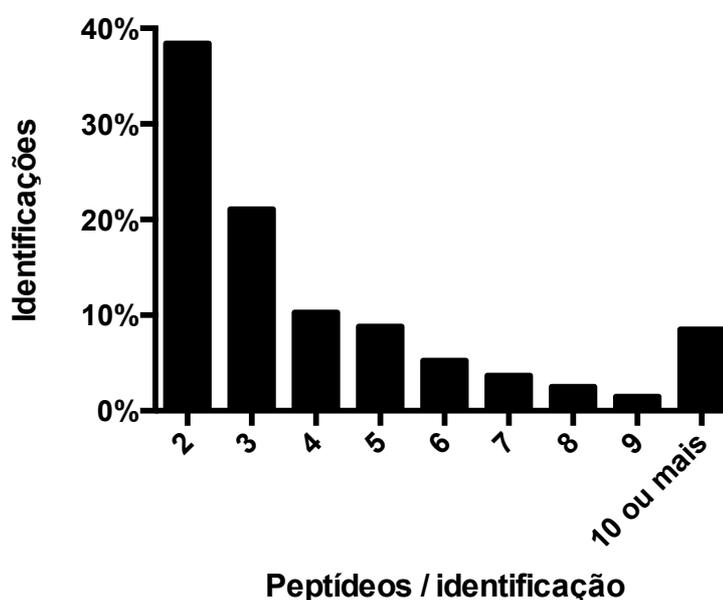


Figura 16. Relação entre o número de peptídeos por identificação e a porcentagem de identificações obtidas por meio de busca no banco de dados de *T. cruzi* CL Brener.

1. Classificação de *Gene Ontology* (GO)

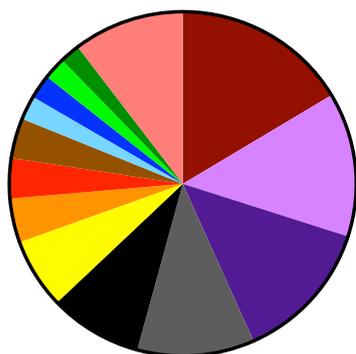
Com o emprego do Blast2GO, 2.564 anotações foram usadas para a classificação por GO de 665 proteínas identificadas distintas (69,78%). Em comparação com a busca no banco de dados de *Trypanosoma* spp. e outros tripanossomatídeos, na qual 39,66% das identificações foram anotadas, o uso do banco de dados de *T. cruzi* CL Brener resultou em uma maior porcentagem de anotações de GO.

Para a categoria componente celular, 28 termos GO foram fornecidos pela classificação das proteínas identificadas nesta busca. O termo GO complexo proteico apresentou o maior número de anotações (16,38%), depois apareceram citoplasma (13,56%), intracelular (13,37%), celular (10,92%), ribossomo (8,66%), etc. (Figura 17A).

Em relação a processo biológico, as identificações foram distribuídas a 40 termos GO. O termo GO processo catabólico (12,04%) foi o mais abundante, seguido por processo metabólico (10,77%), processo metabólico contendo nucleobase (8,87%), etc. (Figura 17B). O termo GO simbiose, abrangendo mutualismo até parasitismo, um dos mais representativos da busca no banco de dados de *Trypanosoma* spp. e outros tripanossomatídeos, surgiu, nesta busca, no fim da lista decrescente em número de anotações, com apenas uma anotação que foi relativa à proteína putativa peptidil-prolil cis-trans isomerase (Q4D932).

Trinta e um termos GO foram resultantes da classificação de função molecular predita para as proteínas identificadas. O termo GO ligação a nucleotídeos apresentou a maior porcentagem de anotações (19,43%), em seguida constaram atividade catalítica (14,36%), ligação a proteínas (11,23%), etc. (Figura 17C).

A



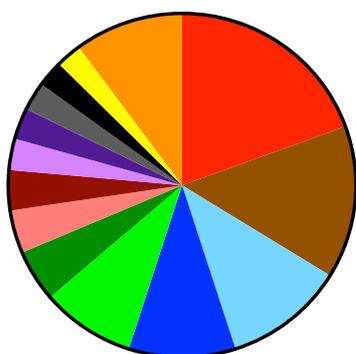
- 16.38% *protein complex*
- 13.56% *cytoplasm*
- 13.37% *intracellular*
- 10.92% *cell*
- 8.66% *ribosome*
- 6.59% *cytoskeleton*
- 4.14% *mitochondrion*
- 3.77% *nucleus*
- 3.77% *cytosol*
- 2.26% *organelle*
- 2.26% *Golgi apparatus*
- 2.07% *plasma membrane*
- 1.88% *peroxisome*
- 10.36% *others*

B



- 12.04% *catabolic process*
- 10.77% *metabolic process*
- 8.87% *nucleobase-containing compound metabolic process*
- 7.81% *protein metabolic process*
- 7.71% *translation*
- 6.86% *biosynthetic process*
- 5.81% *biological_process*
- 4.54% *transport*
- 4.01% *cellular protein modification process*
- 3.70% *carbohydrate metabolic process*
- 3.17% *response to stress*
- 2.96% *protein transport*
- 2.96% *generation of precursor metabolites and energy*
- 2.64% *regulation of biological process*
- 2.01% *signal transduction*
- 14.15% *others*

C



- 19.43% *nucleotide binding*
- 14.36% *catalytic activity*
- 11.23% *protein binding*
- 9.94% *binding*
- 8.75% *hydrolase activity*
- 4.88% *structural molecule activity*
- 4.05% *transferase activity*
- 3.78% *peptidase activity*
- 2.95% *motor activity*
- 2.95% *RNA binding*
- 2.67% *nucleic acid binding*
- 2.49% *DNA binding*
- 2.39% *transporter activity*
- 10.13% *others*

Figura 17. Gráficos de setores para cada categoria de GO: (A) componente celular, (B) processo biológico e (C) função molecular, de acordo com o programa Blast2GO.

2. Predição de atividades enzimáticas

O Blast2GO gerou 263 anotações de EC para um total de 225 proteínas anotadas. A classe das hidrolases (EC 3) é a que apresentou o maior número de anotações (n = 74), e a classe das transferases (EC 2) foi a segunda mais representativa (n = 66). Nessa análise, EC 2 apresentou 39 atividades enzimáticas distintas e, assim, foi o grupo com maior número de membros nesse quesito (Tabela 6).

As atividades enzimáticas mais abundantes de todos os seis ECs foram: nucleosídeo-trifosfato fosfatase (EC 3.6.1.15) e proteína tirosino-quinase (EC 2.7.10.0), ambas com 11 anotações. A reação catalisada por membros do EC 3.6.1.15 é a hidrólise de nucleosídeos trifosfatos, produzindo nucleosídeos difosfatos e fosfato inorgânico. Nesse grupo, foram enquadradas ATPases, metalopeptidase dependente de zinco e de ATP, HSP78, HSP100 e também uma subunidade β de ATP sintase putativa (Q4DTX7). Em relação a essa última anotação, é possível que a anotação do gene como um candidato a ATP sintase proveniente do Uniprot esteja equivocada. Sabe-se que a atividade proteína Tyr-quinase é essencial para a invasão da célula hospedeira pelo *T. cruzi*, pois a inibição de Tyr-quinases pelo inibidor genisteína reduz a infectividade de tripomastigotas metacíclicas e tripomastigotas derivados de cultura tecidual *in vitro*. A ativação de Tyr-quinases induz a fosforilação de p175 em tripomastigotas, fosforilação não detectada em epimastigotas (Favoreto, Dorta e Yoshida, 1998). Surpreendentemente, detectamos Ser/Thr-quinases putativas (Q4E2L0, Q4DBG2, Q4DDT6, Q4C5P0 e Q4CQE6) anotadas pelo Blast2GO como Tyr-quinases. Novamente, isso pode ser um erro de anotação dos genes.

Tabela 6. Abundância de anotações e de atividades enzimáticas para cada uma das seis classes de EC.

| Classe de EC | Anotações ^a | Atividades enzimáticas ^b |
|---------------------|------------------------|-------------------------------------|
| 1 - Óxido-redutases | 55 | 33 |
| 2 - Transferases | 66 | 39 |
| 3 - Hidrolases | 74 | 27 |
| 4 - Liases | 9 | 8 |
| 5 - Isomerases | 16 | 11 |
| 6 - Ligases | 43 | 32 |
| Total | 263 | 150 |

^a Número de anotações de EC pelo Blast2GO

^b Número de *Enzyme* ID distintos do sistema EC

2.1 Peptidases (EC 3.4) putativas

Com base nos resultados da busca no banco de dados de *T. cruzi* CL Brener, 25 proteínas identificadas foram atribuídas à subclasse EC 3.4 (Tabela 7).

Na sub-subclasse das aminopeptidases, foram inseridas 6 identificações, incluindo uma LAP (Q4D4Y0). LAPs são exopeptidases que hidrolisam ligações peptídicas tendo resíduos de leucina na porção N-terminal de proteínas ou substratos peptídicos (Rawlings e Barret, 2004). A partir da montagem do genoma da cepa CL Brener de *T. cruzi*, quatro genes foram anotados como aminopeptidases da família M17 putativas, sendo dois deles como leucil-aminopeptidases (El-Sayed *et al.*, 2005). Nosso grupo de pesquisa realizou um trabalho pioneiro sobre uma LAP de *T. cruzi* (LAPTc). A enzima nativa foi isolada de formas epimastigotas da cepa Berenice e a enzima recombinante foi expressa em *E. coli*. As formas nativa e recombinante apresentaram características de atividade enzimática distintas. LAPTc pertence à família M17 ou família das leucil-aminopeptidases e é expressa nas formas epimastigotas, tripomastigotas e amastigotas de *T. cruzi*. A forma recombinante é constituída de um homohexâmero de aproximadamente 330 kDa (Cadavid-Restrepo *et al.*, 2011). Com a caracterização da LAPTc (Q4DZJ3) como uma autêntica leucil-aminopeptidase, a anotação dessa proteína no banco de dados do Uniprot está desatualizada, pois ainda consta como aminopeptidase. Uma dipeptidil-peptidase, que catalisa a hidrólise de dipeptídeos N-terminais de uma cadeia polipeptídica, e uma peptidase T, que hidrolisa tripeptídeos, também foram classificadas na sub-subclasse das aminopeptidases.

Uma ubiquitina C-terminal hidrolase putativa foi atribuída à sub-suculose das ômega-peptidases.

Entre as serino-endopeptidases anotadas, estão a oligopeptidase B (OPB), a prolil-oligopeptidase (POP) e uma proteína não caracterizada. Com base no sequenciamento do genoma da cepa CL Brener de *T. cruzi*, um gene é candidato a codificar POP, dois genes para OPB e dois genes para OPB-símiles (El-Sayed *et al.*, 2005). Essas enzimas de *T. cruzi* também vêm sendo bastante estudadas pelo nosso grupo. A oligopeptidase B de *T. cruzi* (OPBTc) foi originalmente descrita por Santana *et al.* (1992). Está relacionada à geração de um agonista que ativa fosfolipase C da célula hospedeira, o que resulta na formação de inositol trifosfato (IP₃) e na consequente liberação de Ca²⁺ intracelular. Essa cascata de sinalização induz a fusão de lisossomos no sítio de ligação do parasito à célula hospedeira (Rodriguez *et al.*, 1995; Burleigh *et al.*, 1997). Recentemente, nosso grupo demonstrou que a molécula OPBTc é um dímero estável, ativo e independente de ligações dissulfeto intermoleculares (Motta *et al.*, 2012). A prolil-oligopeptidase de *T. cruzi* (POP Tc80), que também foi descrita pelo nosso grupo de pesquisa, pode estar envolvida na degradação da matriz extracelular, processo necessário para atingir a célula hospedeira, pois foi demonstrado que essa protease é capaz de hidrolisar colágenos do tipo I e IV (Santana *et al.*, 1997) e fibronectina (Grellier *et al.*, 2001). Um outro trabalho do nosso grupo mostrou que inibidores específicos da POP Tc80 impediram a entrada de tripomastigotas nas células mamíferas, mas não a ancoragem à membrana dessas, e também sugeriu um modelo tridimensional da POP Tc80 (Bastos *et al.*, 2005).

Como metalo-endopeptidases, uma thimet oligopeptidase, uma *mitochondrial processing peptidase* (MPP), uma pitrilisina-símile, uma metalopeptidase dependente de zinco e ATP e uma peptidase putativas foram anotadas.

E, por fim, 10 proteínas do proteassoma putativas receberam a classificação de treonino-endopeptidases. O proteassoma 26S pode ser dividido em dois grupos funcionais: a porção central catalítica 20S e a porção regulatória 19S. A porção 20S é formada por dois anéis β internos idênticos, cada um formado por sete diferentes subunidades β, e dois anéis α externos idênticos, cada um formado por sete diferentes subunidades α. A porção 20S é fechada em cada uma das extremidades por uma 19S (Kim, Yu e Cheng, 2011). Várias subunidades α e β putativas da porção 20S foram anotadas. *T. cruzi* expressa tanto o proteassoma 26S (descrito

inicialmente em eucariotos) quanto o complexo HsIVU (descrito originalmente em eubactérias) (Gonzalez *et al.*, 1996; De Diego *et al.*, 2001; Barboza *et al.*, 2012), assim como outros tripanossomatídeos. Uma subunidade do complexo HsIVU putativa também foi atribuída à sub-subclasse das treonino-endopeptidases.

Tabela 7. Proteínas anotadas para a atividade de peptidase (EC 3.4) distribuídas em cinco sub-subclasses de EC.

| Sub-subclasse de EC | Descrição ^a | Código identificador ^b |
|---|---|-----------------------------------|
| Amino-peptidases (EC 3.4.11) | <i>Cytosolic leucyl aminopeptidase</i> | Q4D4Y0 |
| | <i>Aminopeptidase</i> | Q4D5P1 |
| | <i>Uncharacterized protein</i> | Q4D7E5 |
| | <i>Dipeptidyl-peptidase</i> | Q4DE42 |
| | <i>Peptidase T</i> | Q4DPI6 |
| Ômega-peptidases (EC 3.4.19) | <i>Amino-peptidase</i> | Q4E686 |
| | <i>Ubiquitin carboxyl-terminal hydrolase</i> | Q4E680 |
| Serino-endo-peptidases (EC 3.4.21) | <i>Oligopeptidase B</i> | Q4CW30 |
| | <i>Uncharacterized protein</i> | Q4DQZ3 |
| | <i>Prolyl oligopeptidase</i> | Q4E132 |
| Metallo-endo-peptidases (EC 3.4.24) | <i>Thimet oligopeptidase</i> | Q4D7D2 |
| | <i>Peptidase</i> | Q4DFP8 |
| | <i>ATP-dependent zinc metallopeptidase</i> | Q4DJU4 |
| | <i>Mitochondrial processing peptidase</i> | Q4DPR3 |
| | <i>Pitrilysin-like metalloprotease</i> | Q4DU56 |
| | <i>Proteasome alpha 2 subunit</i> | Q4CKY8 |
| Treonino-endo-peptidases (EC 3.4.25) | <i>Proteasome alpha 7 subunit</i> | Q4CTU6 |
| | <i>Proteasome subunit alpha type</i> | Q4D144 |
| | <i>Proteasome subunit beta type</i> | Q4D6T6 |
| | <i>Hslvu complex proteolytic subunit-like</i> | Q4D729 |
| | <i>Proteasome subunit alpha type</i> | Q4DAW0 |
| | <i>Proteasome beta 3 subunit</i> | Q4DHA9 |
| | <i>Proteasome subunit beta type</i> | Q4DND1 |
| | <i>Proteasome subunit alpha type</i> | Q4DWX8 |
| <i>Proteasome subunit beta type</i> | Q4E0L9 | |

^a Descrição da proteína anotada proveniente do Uniprot

^b Número de acesso do Uniprot

2.2. Predição de enzimas envolvidas em vias metabólicas

Anotações de KEGG *pathways* - Blast2GO geraram uma lista de 69 vias metabólicas, das quais as dez mais numerosas em relação a anotações por via estão destacadas na Tabela 8. Metabolismo de purinas foi a via mais representativa com 36 anotações, assim como para a busca no banco de dados de *Trypanosoma* e outros tripanossomatídeos, resultado concorde com o fato de que formas epimastigotas são replicativas. Depois surgiu ciclo do ácido cítrico com 23 anotações.

A via metabólica com a maior quantidade de atividades enzimáticas foi biossíntese de aminoacil-tRNA com 14 atividades enzimáticas distintas (todas de ligase, porém EC diferentes), seguida por metabolismo de purinas e ciclo do ácido cítrico, ambas com 13. Entre as dez vias metabólicas mais representativas, também estão vias relacionadas ao metabolismo energético, como ciclo do ácido cítrico, glicólise/gliconeogênese, fosforilação oxidativa, em conformidade com os altos requerimentos energéticos desse estágio proliferativo do parasito. As atividades enzimáticas e os números de acesso das cinco vias metabólicas mais representativas em número de proteínas anotadas estão especificados na Tabela 9.

Tabela 8. As dez vias metabólicas mais abundantes em número de anotações e os respectivos números de atividades enzimáticas.

| Via metabólica ^a | Anotações ^b | Atividades enzimáticas ^c |
|---|------------------------|-------------------------------------|
| Metabolismo de purinas | 36 | 13 |
| Ciclo do ácido cítrico | 23 | 13 |
| Biossíntese de aminoacil-tRNA | 17 | 14 |
| Metabolismo de piruvato | 16 | 10 |
| Glicólise/Gliconeogênese | 14 | 12 |
| Metabolismo de arginina e prolina | 12 | 10 |
| Metabolismo de butanoato | 11 | 5 |
| Fosforilação oxidativa | 11 | 5 |
| Metabolismo de tiaminas | 11 | 1 |
| Metabolismo de alanina, aspartato e glutamato | 10 | 9 |

^a Sugerida por KEGG *pathways* - Blast2GO

^b Número de anotações (pode haver uma proteína anotada mais de uma vez na mesma via metabólica)

^c Número de *Enzyme ID* distintos do sistema EC

2.2.1. Metabolismo de purinas

Recapitulando a via de recuperação das purinas na Figura 18, uma das primeiras reações consiste na transformação direta de adenina e PRPP em AMP e pirofosfato catalisada pela adenina fosforribosil transferase (APRT). Duas APRTs putativas (Q4DNZ3 e Q4DNZ4) foram identificadas. Assim como pela busca no banco de dados de *Trypanosoma* spp. e outros tripanossomatídeos, uma PRPP sintetase putativa (Q4DIR4) que produz PRPP a partir de ribose-5-fosfato foi identificada.

Algumas enzimas de interconversão de nucleotídeos foram identificadas: uma adenilsuccinato sintetase putativa (Q4CWX1) que catalisa a conversão de IMP em adenilsuccinato na presença de aspartato e GTP; uma guanina desaminase putativa (Q4D711) que converte guanina em xantina; uma GMP sintetase putativa (Q4E1D4) que está envolvida na conversão de xantina monofosfato em GMP. A caracterização bioquímica da adenilsuccinato sintetase de *T. cruzi* foi publicada junto com a da adenilsuccinato liase (Spector, Berens e Marr, 1982).

Geralmente os nucleotídeos são utilizados nas reações de biossíntese na forma de nucleosídeos trifosfato. Na presença de ATP, a adenosina quinase catalisa a conversão de adenosina em AMP e, no presente estudo, uma adenosina quinase putativa (Q4CVK9) foi identificada. Kidder (1982) detectou atividade de adenosina quinase em *T. cruzi* e desde então nenhum outro trabalho especificamente sobre tal enzima desse parasito foi publicado. A fosforilação de AMP em ADP é realizada pela adenilato quinase com gasto de ATP e duas candidatas a essa enzima (Q4D6Z4 e Q4D818) foram atribuídas à via de metabolismo de purinas com atividade de quinase. *T. cruzi* apresenta sete isoformas putativas de adenilato quinases, das quais uma foi citolocalizada em organelas peroxissomos-símile denominadas glicossomos também presentes em outros cinetoplastídeos estudados e, outra, no núcleo (Bouvier *et al.*, 2006; Milagros Camara Mde *et al.*, 2012; Camara Mde *et al.*, 2013). O ADP produzido pelas adenilato quinases é então fosforilado a ATP pelas enzimas glicolíticas ou por meio da fosforilação oxidativa.

Foram identificadas duas DNA polimerases I mitocondriais putativas (Q4D527 e Q4DSW2) pertencentes à família A de DNA polimerases e duas RNA polimerases dirigidas a DNA putativas (Q4CRQ2 e Q4DDX5).

Pela outra busca, adenilato ciclases putativas haviam sido identificadas. Uma vez que as adenilato ciclases formam cAMP, as fosfodiesterases cAMP-específicas

regulam os níveis intracelulares desse mensageiro secundário. De forma geral, os cinetoplastídeos, incluindo o *T. cruzi*, codificam quatro famílias de fosfodiesterases denominadas TcrPDEA, B, C e D, das quais três já foram bem estudadas (Kunz *et al.*, 2006). Em relação ao uso exclusivo de cAMP como substrato, quatro fosfodiesterases de *T. cruzi* foram descritas: TcrPDEA1 mostrou-se resistente aos inibidores clássicos de fosfodiesterases (Alonso *et al.*, 2007), TcrPDEB1 (Diaz-Benjumea *et al.*, 2006) e TcrPDEB2 (D'angelo *et al.*, 2004) e TcrPDEC2 (Alonso *et al.*, 2006) foram caracterizadas bioquimicamente, das quais TcrPDEB2 foi associada fortemente ao flagelo e, TcrPDEC2, à fração de membranas. Recentemente, a associação entre TcrPDEC2 e o complexo do vácuolo contrátil em epimastigotas foi descrita, sugerindo um papel dessa enzima no processo de osmorregulação (Schoijet *et al.*, 2011). Nesse contexto, uma fosfodiesterase cAMP-específica putativa (Q4E0Y6) foi identificada na cepa G de *T. cruzi*.

Diversas proteínas putativas foram classificadas como fosfatases na via de metabolismo de purinas: seis cadeias pesadas de dineína (Q4CQ22, Q4DX41, Q4CPG2, Q4CTR4, Q4CPP9 e Q4CX46); uma subunidade de ATPase do proteassoma (Q4DW08); um transportador ABC (Q4DWA3); uma AAA ATPase (Q4DWK0); uma ATPase que se funde a vesículas (Q4D5R2); uma metalopeptidase dependente de zinco e de ATP (Q4DJU4) já citada no tópico 2.1 desta busca; uma katanina (Q4DLN7); duas HSP100 (Q4CLP1, Q4CNM5); duas HSP78 (Q4CRW0, Q4DTH5); uma ATPase de retículo endoplasmático transitória (Q4DWB5); uma subunidade β de ATP sintase (Q4DTX7).

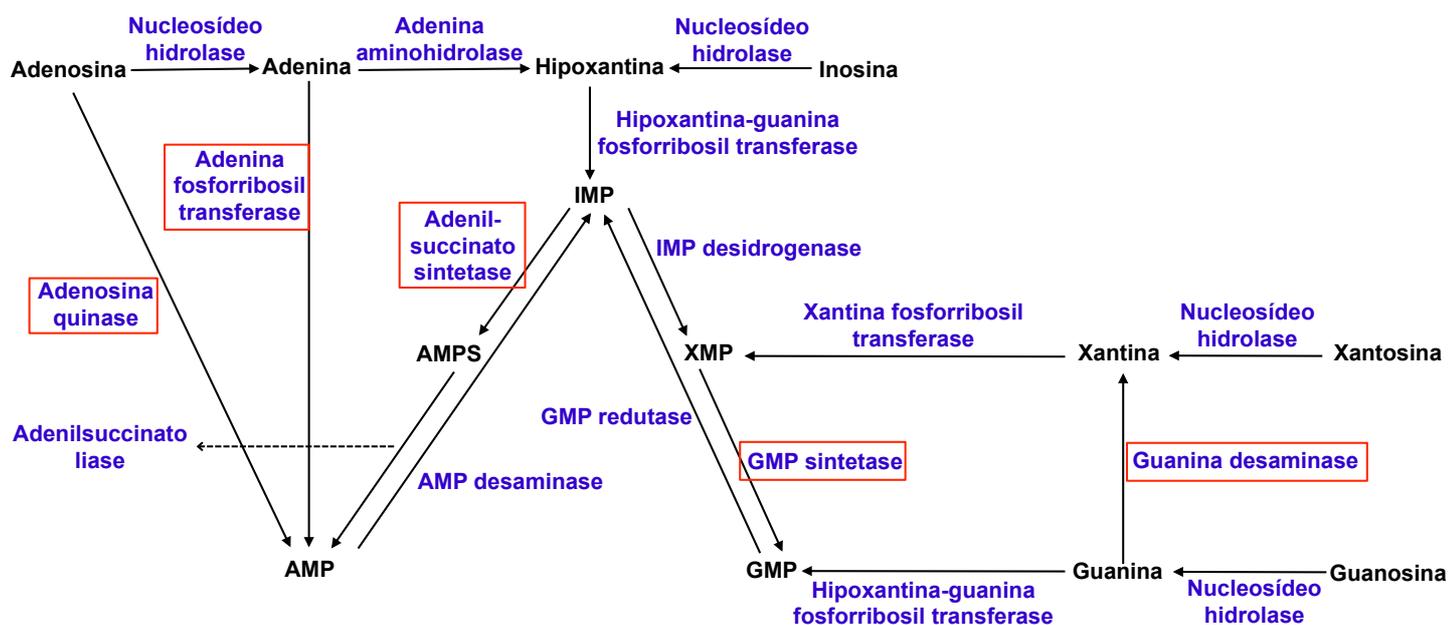


Figura 18. Diagrama da via de recuperação/salvação das purinas predita em *Trypanosoma spp.*, com base nas enzimas descritas para *Leishmania spp.* As enzimas marcadas em vermelho correspondem às proteínas identificadas na cepa G de *T. cruzi* pela busca 2. Modificado de Boitz *et al.* (2013). AMP: adenilato, AMPS: adenilsuccinato, GMP: guanilato, IMP: inosinato, XMP: xantinilato.

2.2.2. Ciclo do ácido cítrico

Nos organismos aeróbicos, o ciclo do ácido cítrico, também chamado de ciclo dos ácidos tricarbóxicos ou ciclo de Krebs, é uma via anfibólica, ou seja, participa de processos catabólicos e anabólicos, e ocorre na matriz mitocondrial. Além de servir ao catabolismo oxidativo de carboidratos, ácidos graxos e aminoácidos, essa via metabólica também produz precursores para muitas outras vias de biossíntese.

Vinte proteínas que podem participar da etapa inicial de conversão de piruvato em acetil-CoA e do ciclo do ácido cítrico foram identificadas (Figura 19). Em relação à descarboxilação oxidativa do piruvato formado na via glicolítica a acetil-CoA, a totalidade do complexo da piruvato desidrogenase descrito foi identificada: subunidade E1 da piruvato desidrogenase (Q4DB91), diidrolipoil desidrogenase (Q4DD33) e diidrolipoil transacetilases (Q4DYI5 e Q4DZT8) putativas. Quanto às identificações atribuídas ao ciclo do ácido cítrico estão: uma aconitase putativa (Q4E5G5), uma isocitrato desidrogenase dependente de NADP^+ (Q4E4L7), quatro subunidades putativas do complexo da 2-oxoglutarato desidrogenase (Q4CTE8, Q4CYI3, Q4DKY2 e Q4D1R9), uma subunidade α da succinil-CoA sintetase putativa

(Q4E2I6), uma flavoproteína succinato desidrogenase putativa (Q4D4X1), cinco malato desidrogenases putativas (Q4DDK8 (glicossomal), Q4DXL5, Q4CTR7, Q4DRD8 (glicossomal) e Q4D123), três fumarato redutases dependentes de NADH putativas (Q4CMU8, Q4CYX1 e Q4D779).

A partir de busca no GeneDB (<http://www.genedb.org>), dois genes de *T. cruzi* foram anotados como isocitrato desidrogenases (IDH). A isocitrato desidrogenase supracitada (Q4E4L7) corresponde à Tc_IDH2, que está principalmente localizada no citosol, enquanto Tc_IDH1 (Q4DG65) foi descrita como uma enzima mitocondrial, em formas epimastigotas (Leroux, Maugeri, Cazzulo, *et al.*, 2011). O fato de ter sido descrita uma isocitrato desidrogenase citoplasmática em *T. cruzi* poderia representar uma evidência de que o ciclo do ácido cítrico tem duas ramificações principais, uma relacionada à condensação de oxaloacetato e acetil-CoA pelo complexo da piruvato desidrogenase produzindo citrato, e a segunda associada à conversão de 2-oxoglutarato em succinato provavelmente resultante da oxidação de aminoácidos (setas grossas na Figura 19), como já descrito para *T. brucei* (Van Weelden *et al.*, 2005). Assim, o citrato produzido na mitocôndria pode ser substrato para a IDH mitocondrial e a IDH citosólica.

Além dessas, constatamos uma fosfoenolpiruvato carboxiquinase putativa (Q4DB81). Nos glicossomos, a reação catalisada pela fosfoenolpiruvato carboxiquinase converte fosfoenolpiruvato, CO₂ e ADP em oxaloacetato e ATP. Depois a malato desidrogenase glicossomal (putativas: Q4DDK8 e Q4DRD8) produz malato a partir de oxalacetato que é repostado no ciclo do ácido cítrico quando está em baixa concentração (Cannata e Cazzulo, 1984; Cazzulo, 1992).

isoleucil- (Q4D2F2), alanil- (Q4D747), valil- (Q4DQQ7), metionil- (Q4D6H2), glicil- (Q4D449), cisteinil- (Q4CQ34), glutaminil- (Q4E0Y0), arginil- (Q4E397), fenilalanil- (Q4CXE3) e asparaginil- (Q4E595) tRNA sintetases. As estruturas cristalográficas das histidil-tRNA sintetases de *T. brucei* e *T. cruzi* já foram determinadas (Merritt *et al.*, 2010).

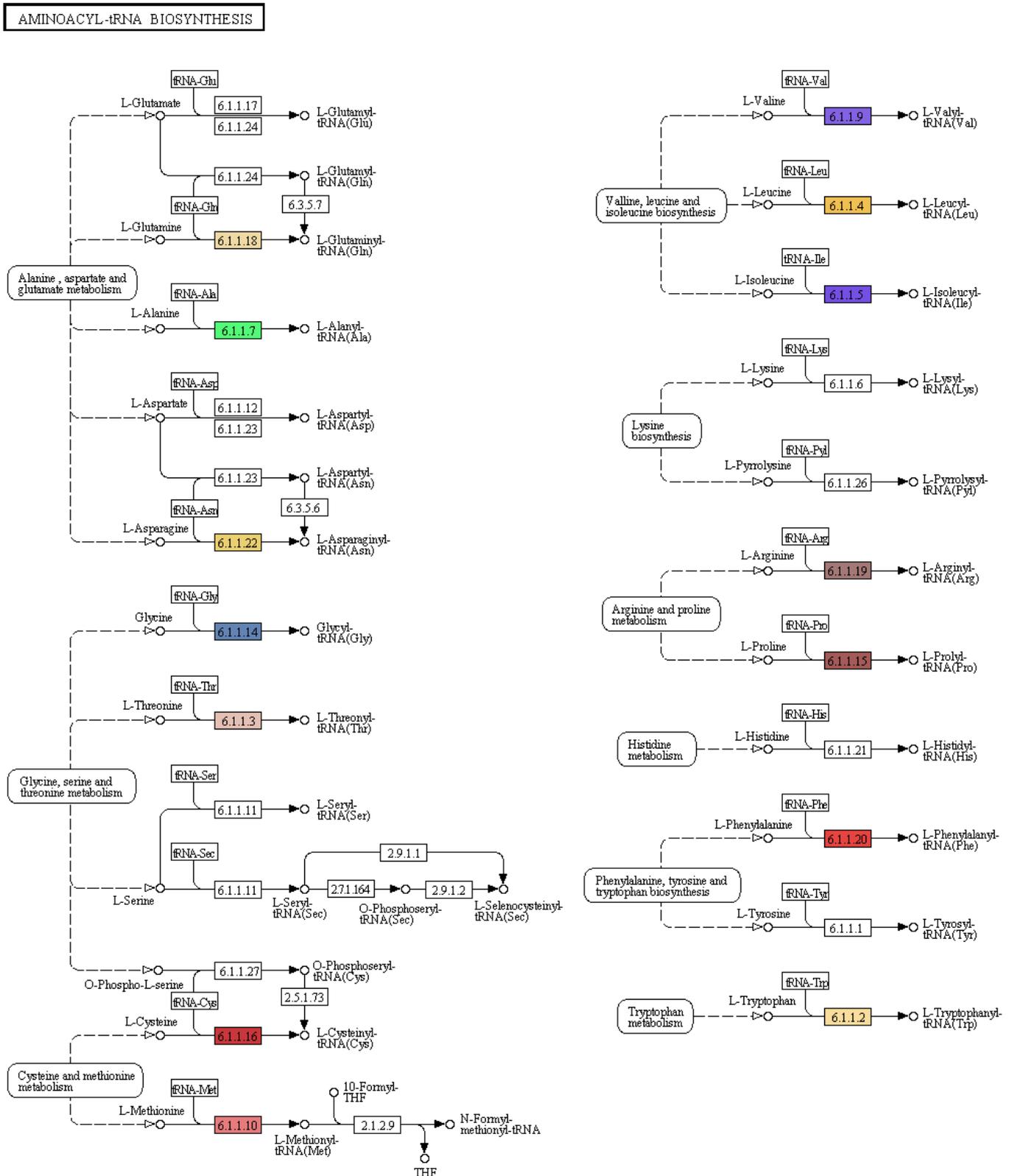


Figura 20. Representação esquemática da via de biossíntese de aminoacil-tRNA gerada pelo programa Blast2GO. Os retângulos em destaque indicam as proteínas identificadas na cepa G de *T. cruzi*.

2.2.4. Metabolismo de piruvato

A Figura 21 representa a via de metabolismo do piruvato com as proteínas identificadas. As cinco malato desidrogenases putativas (Q4DDK8 (glicossomal), Q4DXL5, Q4CTR7, Q4DRD8 (glicossomal) e Q4D123) e as proteínas putativas piruvato desidrogenase (Q4DB91), diidrolipoil desidrogenase (Q4DD33) e diidrolipoil transacetilases (Q4DYI5 e Q4DZT8) possivelmente envolvidas na descarboxilação oxidativa do piruvato a acetil-CoA citadas no tópico do ciclo do ácido cítrico também foram anotadas na via de metabolismo de piruvato.

No genoma de *T. cruzi*, existem quatro ORF que possivelmente codificam enzimas málicas, responsáveis pela conversão de malato em piruvato e fortes candidatas a contribuírem para a produção de NADPH. Como previamente publicado, muito provavelmente as duas enzimas málicas identificadas (Q4DV36 e Q4DJ68) são produtos de cópias de genes alelos localizados em cromossomos homólogos devido ao grau de identidade de 97%. Foi mostrado que a enzima málica TcME2 (Q4DV36) que é específica para NADP⁺ localiza-se no citosol do parasito e é expressa em epimastigotas, tripomastigotas, amastigotas e tripomastigotas metacíclicas, sendo mais abundante nesse último estágio de vida (Leroux, Maugeri, Opperdoes, *et al.*, 2011).

Nós também identificamos uma piruvato fosfato diquinase (Q4E0Q0). A piruvato fosfato diquinase catalisa a reação reversível de conversão de piruvato em fosfoenolpiruvato. Foi sugerido que a termodinâmica dessa enzima de *T. cruzi* é na verdade muito similar a de uma pirofosfatase, produzindo piruvato a partir de fosfoenolpiruvato. Em epimastigotas, localiza-se na matriz glicossomal e essa enzima seria um elo essencial entre a glicólise, a β -oxidação de ácidos graxos e muitas vias biossintéticas produtoras de pirofosfato. Uma vez que nenhuma pirofosfatase foi identificada na fração glicossomal, a piruvato fosfato diquinase seria encarregada da hidrólise de pirofosfato que é tóxico a altas concentrações (Acosta *et al.*, 2004).

Uma fosfoenolpiruvato carboxiquinase putativa (Q4DB81) já citada, uma hidroxiaçilglutaciona hidrolase putativa (Q4E0K0), uma acetil-CoA sintetase (Q4DYN2) e uma acetil-CoA carboxilase putativa (Q4CL50) também estão presentes nesta via.

PYRUVATE METABOLISM

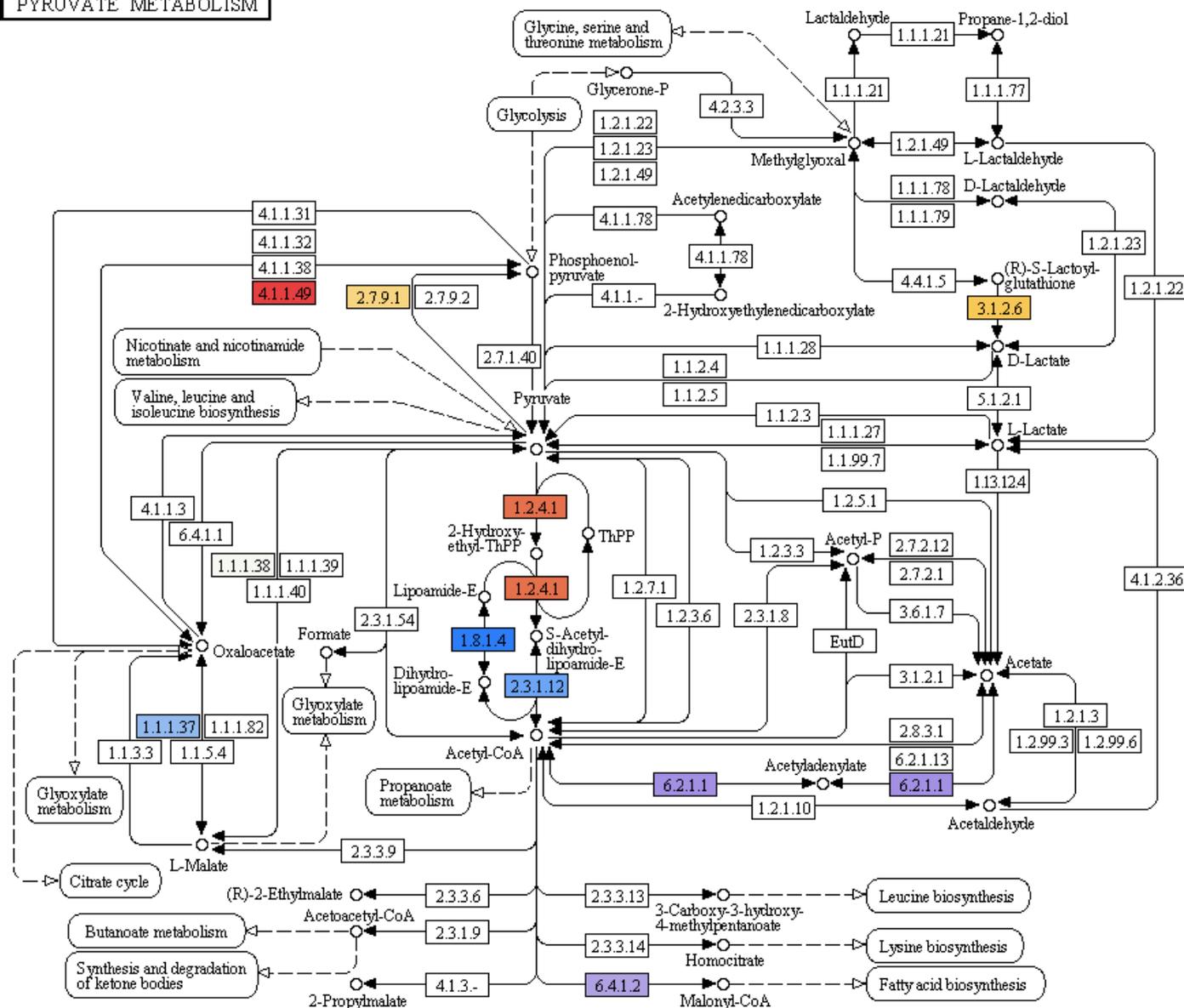


Figura 21. Esquema representativo da via de metabolismo de piruvato gerado pelo programa Blast2GO. Os retângulos em destaque indicam as proteínas identificadas na cepa G de *T. cruzi*.

2.2.5. Glicólise/gliconeogênese

É importante ressaltar que insetos hematófagos obtêm energia de L-prolina e/ou L-glutamina, aminoácidos exuberantes na hemolinfa e nos fluidos teciduais. Conseqüentemente, estágios de vida de *T. cruzi* no triatomíneo como epimastigotas utilizam aminoácidos como fonte de energia, como L-prolina, L-histidina e D-prolina, enquanto tripomastigotas circulantes usam glicose que é abundante nos fluidos do hospedeiro vertebrado. Porém em meio rico em glicose e aminoácidos, essa hexose

é utilizada preferencialmente por epimastigotas (revisado por Bringaud, Riviere e Coustou (2006). No meio LIT para cultura de epimastigotas, havia 0,2% de glicose e, portanto, a presença da via de glicólise entre as mais representativas está de acordo com a disponibilidade da fonte preferencial de energia no meio de cultura.

A glicólise em *T. cruzi* é distinta de outros eucariotos devido à ocorrência de taxas similares do consumo de glicose tanto em anaerobiose quanto em condições aeróbicas. Isso é característico de um metabolismo do tipo fermentativo e, assim, foi proposto que esse parasito realiza fermentação aeróbica. Não ocorre o efeito Pasteur que significa a transição da anaerobiose para a aerobiose acompanhada por uma rápida diminuição na taxa de utilização de glicose, ou seja, a inibição do fluxo glicolítico na presença do oxigênio (revisado por Cazzulo (1992)).

Em relação ao metabolismo de glicose, a Figura 22 ilustra as proteínas identificadas. Como sete reações da via glicolítica são livremente reversíveis e as enzimas envolvidas atuam também na gliconeogênese, talvez por isso o Blast2GO categorize essas duas vias metabólicas em uma só, via glicólise/gliconeogênese.

As seis primeiras reações enzimáticas da glicólise em *T. cruzi* são compartimentalizadas nos glicossomos e as quatro últimas reações ocorrem no citosol. Em formas tripomastigotas sanguíneas de *T. brucei*, a sétima reação da glicólise ocorre nos glicossomos (revisado por Cazzulo (1992) e Hannaert *et al.* (2003)).

A primeira reação, que envolve a fosforilação de glicose produzindo glicose-6-fosfato, é catalisada pela hexoquinase e nós identificamos uma hexoquinase putativa (Q4DQ27). Foi identificada uma candidata a participar da segunda reação, uma glicose-6-fosfato isomerase ou fosfoexose isomerase putativa (Q4E2M9).

Outras quatro proteínas identificadas possivelmente relacionadas a essa via pertencem a três reações reversíveis que são subsequentes. A quarta enzima, a frutose-bifosfato aldolase (putativa: Q4D4R9), catalisa a reação que produz diidroxiacetona fosfato e gliceraldeído-3-fosfato a partir de frutose-1,6-bifosfato. A diidroxiacetona fosfato é rapidamente convertida em gliceraldeído-3-fosfato pela quinta enzima da via glicolítica, a triose fosfato isomerase (putativa: Q4DV43). Por sua vez, gliceraldeído-3-fosfato e fosfato inorgânico são transformados em 1,3-bifosfoglicerato pela gliceraldeído-3-fosfato desidrogenase (putativas: Q4DCN9 e Q4DZT1).

Uma enolase putativa (Q4DZ98) foi atribuída pelo Blast2GO à via

glicólise/gliconeogênese. Na glicólise, a enolase é responsável pela desidratação de 2-fosfoglicerato em fosfoenolpiruvato (nona reação da via) e, na gliconeogênese, pela reação inversa. Em um dos trabalhos de proteomas totais de *T. cruzi*, foi demonstrado que a expressão da enolase 2 foi mais alta em formas tripomastigotas e amastigotas comparadas a epimastigotas da cepa Berenice (Paba, Santana, *et al.*, 2004).

Uma fosfomanomutase putativa (Q4E4A3) foi identificada e, na Figura 22, está associada à reação inversa à terceira reação da glicólise que é irreversivelmente catalisada pela fosfofrutoquinase-1, ou seja, trata-se de uma reação da gliconeogênese.

Por causa da interligação de vias metabólicas por seus produtos intermediários, as proteínas putativas candidatas a membros do complexo da piruvato desidrogenase (Q4DB91, Q4DD33, Q4DYI5 e Q4DZT8), uma acetil-CoA sintetase (Q4DYN2) e uma fosfoenolpiruvato carboxiquinase (Q4DB81) putativas foram enquadradas na glicólise/gliconeogênese.

GLYCOLYSIS / GLUCONEOGENESIS

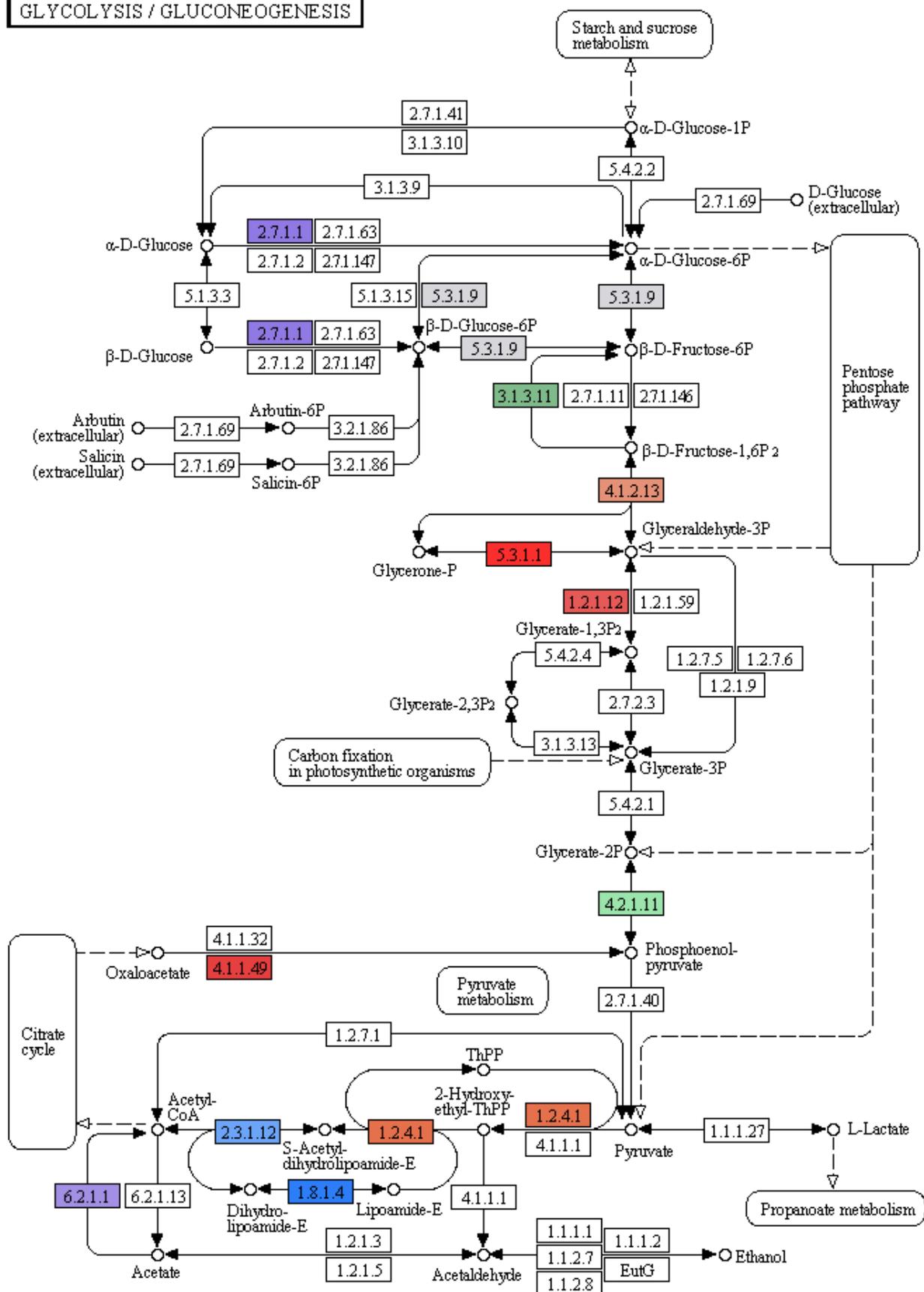


Figura 22. Representação esquemática da glicólise/gliconeogênese gerada pelo Blast2GO. Os retângulos em destaque indicam as proteínas identificadas na cepa G de *T. cruzi*.

2.2.6. Metabolismo de histidina e prolina

Como citado no tópico anterior, epimastigotas utilizam aminoácidos como fonte de energia, como L-prolina, L-histidina e D-prolina. Devido à importância dessas vias metabólicas em epimastigotas, objeto deste estudo, foi criado esse tópico além das cinco vias metabólicas mais representativas em número de atividades enzimáticas relacionadas a esta busca.

Diferentemente de *T. brucei* e *L. major*, *T. cruzi* codifica enzimas responsáveis pela conversão de histidina em glutamato. Assim como para a cepa Brazil, foram identificadas, em epimastigotas da cepa G, uma histidina amônia-liase (Q4E133) e uma urocanato hidratase (Q4D9S6). Essa última foi predita pelo SignalP ser secretada por via clássica que será abordada no tópico seguinte (Tabela suplementar 3).

A L-prolina é o constituinte predominante da hemolinfa e dos fluidos teciduais do vetor e tem se tornado evidente que esse aminoácido serve como uma fonte de carbono chave para formas de tripanossomatídeos presentes no triatomíneo (revisado por Bringaud, Barrett e Zilberstein (2012)). Relacionadas ao metabolismo de prolina, uma pirrolina-5-carboxilato redutase putativa (Q4DH60) e a isoforma 2 da prolina racemase (Q4DA80-2) foram detectadas. Apenas duas espécies possuem prolina racemases descritas, *Clostridium sticklandii* (Cardinale e Abeles, 1968; Rudnick e Abeles, 1975) e *T. cruzi*. O gene *TcPRACA* codifica duas isoformas de prolina racemase, que catalisa a interconversão dos enantiômeros L-prolina e D-prolina: a isoforma 2 de epimastigotas da cepa CL Brener apresenta uma marcação citoplasmática, enquanto que a isoforma 1 é secretada por formas tripomastigotas metacíclicas (Reina-San-Martin *et al.*, 2000). A isoforma 2 é resultante de um mecanismo de *trans-splicing* alternativo, perdendo o sinal para secreção (Chamond *et al.*, 2005). De modo a reforçar esse resultado, a isoforma 2 da prolina racemase identificada foi atribuída à predição por via secretória não clássica (Tabela suplementar 4).

Tabela 9. Relação entre as vias metabólicas mais representativas em número de anotações e as respectivas atividades enzimáticas. (continua)

| Via metabólica ^a | Atividade enzimática ^b | Código identificador ^c |
|----------------------------------|---------------------------------------|--|
| 1) Metabolismo de purinas | Fosforribosiltransferase (EC 2.4.2.7) | Q4DNZ3, Q4DNZ4 |
| | Quinase (EC 2.7.1.20) | Q4CVK9 |
| | Quinase (EC 2.7.4.3) | Q4D6Z4, Q4D818 |
| | Quinase (EC 2.7.4.6) | Q4DBS8 |
| | Difosfoquinase (EC 2.7.6.1) | Q4DIR4 |
| | RNA polimerase (EC 2.7.7.6) | Q4CRQ2, Q4DDX5 |
| | DNA polimerase (EC 2.7.7.7) | Q4D527, Q4DSW2 |
| | Fosfodiesterase (EC 3.1.4.17) | Q4DJT7, Q4E0Y6 |
| | Desaminase (EC 3.5.4.3) | Q4D711 |
| | Adenil pirofosfatase (EC 3.6.1.3) | Q4CQ22, Q4DX41, Q4CY33, Q4CPG2, Q4CTR4, Q4DW08, Q4CPP9, Q4CX46, Q4DWA3 |
| | Fosfatase (EC 3.6.1.15) | Q4DWK0, Q4D5R2, Q4DJU4, Q4DLN7, Q4DSH8, Q4CLP1, Q4CRW0, Q4DTH5, Q4CNM5, Q4DWB5, Q4DTX7 |
| | Sintetase (EC 6.3.4.4) | Q4CWX1 |
| | Sintetase (EC 6.3.5.2) | Q4E1D4 |
| 2) Ciclo do ácido cítrico | Desidrogenase (EC 1.1.1.37) | Q4DDK8, Q4DXL5, Q4CTR7, Q4DRD8, Q4D123 |
| | Desidrogenase (EC 1.1.1.42) | Q4E4L7 |
| | Desidrogenase (EC 1.2.4.1) | Q4DB91 |
| | Desidrogenase (EC 1.2.4.2) | Q4CTE8, Q4CYI3, Q4DKY2 |
| | Desidrogenase (EC 1.3.99.1) | Q4D4X1, Q4CMU8, Q4CYX1, Q4D779 |
| | Desidrogenase (EC 1.8.1.4) | Q4DD33 |
| | Acetiltransferase (EC 2.3.1.12) | Q4DYI5, Q4DZT8 |
| | Succinil transferase (EC 2.3.1.61) | Q4D1R9 |
| | Citrato sintetase (EC 2.3.3.8) | Q4E2I6 ^d |
| | Carboxiquinase (EC 4.1.1.49) | Q4DB81 |
| | Hidratase (EC 4.2.1.3) | Q4E5G5 |
| | Ligase (EC 6.2.1.4) | Q4E2I6 ^d |
| | Ligase (EC 6.2.1.5) | Q4E2I6 ^d |
| 3) Biossíntese de aminoacil-tRNA | Ligase (EC 6.1.1.2) | Q4CTS7 |
| | Ligase (EC 6.1.1.3) | Q4DPR0, Q4D9G6 |

Tabela 9. Relação entre as vias metabólicas mais representativas em número de anotações e as respectivas atividades enzimáticas. (conclusão)

| Via metabólica ^a | Atividade enzimática ^b | Código identificador ^c |
|--|-----------------------------------|--|
| | Ligase (EC 6.1.1.4) | Q4CTR0 |
| | Ligase (EC 6.1.1.5) | Q4DLK5 ^e , Q4D2F2 |
| | Ligase (EC 6.1.1.7) | Q4DLK5 ^e , Q4D747 |
| | Ligase (EC 6.1.1.9) | Q4DQQ7 |
| 3) Biossíntese de aminoacil-tRNA (continuação) | Ligase (EC 6.1.1.10) | Q4D6H2 |
| | Ligase (EC 6.1.1.14) | Q4D449 |
| | Ligase (EC 6.1.1.15) | Q4DLK5 ^e |
| | Ligase (EC 6.1.1.16) | Q4CQ34 |
| | Ligase (EC 6.1.1.18) | Q4E0Y0 |
| | Ligase (EC 6.1.1.19) | Q4E397 |
| | Ligase (EC 6.1.1.20) | Q4CXE3 |
| | Ligase (EC 6.1.1.22) | Q4E595 |
| | Desidrogenase (EC 1.1.1.37) | Q4DDK8, Q4DXL5, Q4CTR7, Q4DRD8, Q4D123 |
| | Desidrogenase (EC 1.1.1.38) | Q4DV36, Q4DJ68 |
| | Desidrogenase (EC 1.2.4.1) | Q4DB91 |
| | Desidrogenase (EC 1.8.1.4) | Q4DD33 |
| 4) Metabolismo de piruvato | Acetiltransferase (EC 2.3.1.12) | Q4DYI5, Q4DZT8 |
| | Fosfato diquinase (EC 2.7.9.1) | Q4E0Q0 |
| | Hidrolase (EC 3.1.2.6) | Q4E0K0 |
| | Carboxiquinase (EC 4.1.1.49) | Q4DB81 |
| | Ligase (EC 6.2.1.1) | Q4DYN2 |
| | Carboxilase (EC 6.4.1.2) | Q4CL50 |
| | Desidrogenase (EC 1.2.1.12) | Q4DCN9, Q4DZT1 |
| | Desidrogenase (EC 1.2.4.1) | Q4DB91 |
| | Desidrogenase (EC 1.8.1.4) | Q4DD33 |
| | Acetiltransferase (EC 2.3.1.12) | Q4DYI5, Q4DZT8 |
| | Hexoquinase e (EC 2.7.1.1) | Q4DQ27 |
| 5) Glicólise/ gliconeogênese | Hexose difosfatase (EC 3.1.3.11) | Q4E4A3 |
| | Carboxiquinase (EC 4.1.1.49) | Q4DB81 |
| | Aldolase (EC 4.1.2.13) | Q4D4R9 |
| | Hidratase (EC 4.2.1.11) | Q4DZ98 |
| | Isomerase (EC 5.3.1.1) | Q4DV43 |
| | Isomerase (EC 5.3.1.9) | Q4E2M9 |
| | Ligase (EC 6.2.1.1) | Q4DYN2 |

^a Via metabólica sugerida por KEGG - Blast2GO

^b *Enzyme* ID do sistema EC anotado na via metabólica

^c Número de acesso do Uniprot

^{d,e} Identificações anotadas com três *Enzyme* ID na mesma via metabólica

3. Predição de proteínas secretadas

Das 953 proteínas identificadas que compreendem proteínas putativas e não caracterizadas, o SignalP anotou, 87 proteínas (9,13%) e o SecretomeP, 477 proteínas (50,05%) (Figura 23). Novamente, as proteínas de cada uma das vias são mutuamente exclusivas. As proteínas preditas a serem secretadas por via não clássica prevaleceram em relação às preditas a serem secretadas por via clássica, assim como pela busca no banco de dados de *Trypanosoma* spp. e outros tripanossomatídeos, embora não se possa excluir que o algoritmo do SignalP é incapaz de identificar todas as sequências sinais de *T. cruzi*. Essa predominância de proteínas secretadas por via não clássica também foi encontrada a partir de proteínas excretadas/secretadas de três frações de epimastigotas e tripomastigotas metacíclicas da cepa Dm28c analisadas conjuntamente: ectossomos que brotam diretamente da membrana plasmática, exossomos resultantes da fusão exocítica de corpos multivesiculares supostamente com a bolsa flagelar e proteínas solúveis que não foram associadas com vesículas (Bayer-Santos *et al.*, 2013).

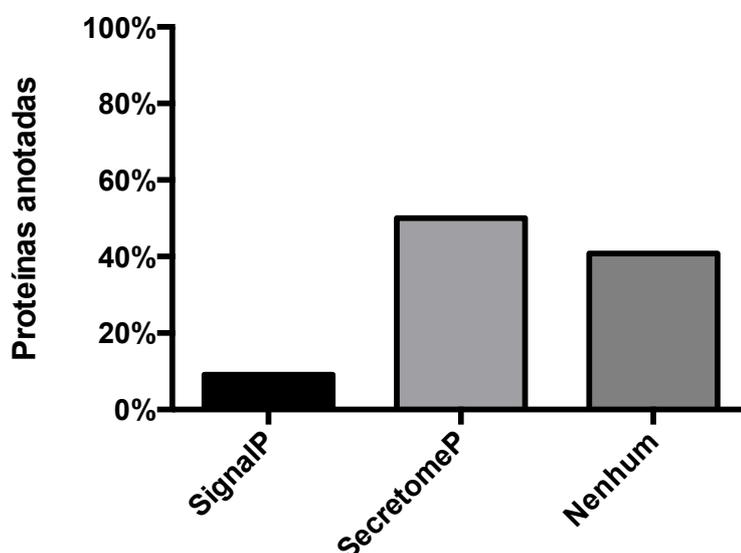


Figura 23. Predição de proteínas secretadas pelas vias clássica (SignalP) e não clássica (SecretomeP) e não anotadas como secretadas (nenhum).

Trinta e oito grupos de proteínas putativas, de acordo com a descrição do Uniprot, poderiam ser secretados por via clássica de acordo com o SignalP, dos quais os mais representativos foram: proteína RHS (5 membros) e DGF-1 (3

membros) (Tabela suplementar 3). Para secreção por via não clássica pelo SecretomeP, 215 grupos de proteínas putativas foram preditos e os mais numerosos foram: *60S ribosomal protein* (17 membros) e *40S ribosomal protein* (12 membros) (Tabela suplementar 4). 17 tipos de proteínas foram atribuídos a ambas as vias de secreção. A secreção de proteínas ribossomais poderia estar envolvida no carregamento de mRNA para fora do parasito. O tráfego de mRNAs e microRNAs para o meio extracelular por meio de exossomos foi demonstrado em células mamíferas por Valadi *et al.* (2007) e muitos desses mRNAs não foram detectados no citosol da célula doadora. mRNAs exossomais foram transferidos e traduzidos após a entrada nas células receptoras. Hipoteticamente, proteínas ribossomais carregariam mRNAs e seriam secretadas por vesículas pelo parasito.

3.1. Predição de proteínas integrais de membrana

Para investigar as possíveis proteínas localizadas na superfície celular, o programa TMHMM fez a predição de 119 proteínas identificadas (12,49%) com domínios transmembrânicos, variando de 1 a 20 hélices transmembrânicas por identificação.

Na intersecção entre proteínas preditas com domínios transmembrânicos e secretadas por via clássica e na intersecção entre proteínas preditas com domínios transmembrânicos e secretadas por via não clássica, estão 23 e 83 proteínas, respectivamente. Logo, 69,75% das proteínas integrais de membrana preditas seriam, teoricamente, secretadas por via não clássica e, apenas 19,33%, por via clássica (Figura 24).

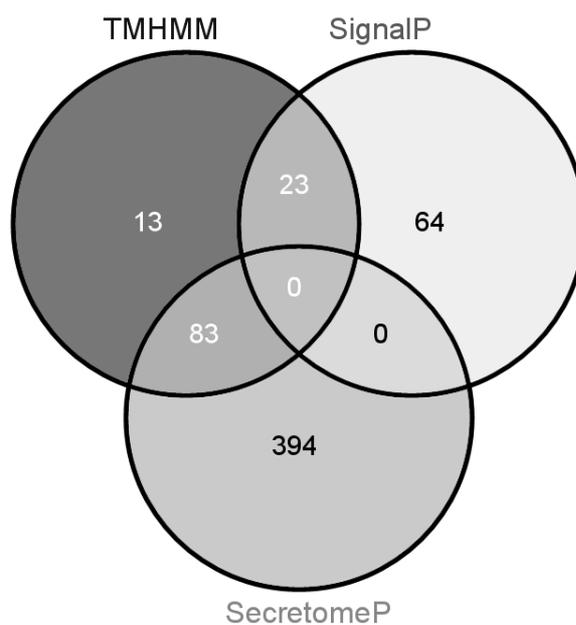


Figura 24. Diagrama de Venn correlacionado as proteínas com preditas hélices transmembrânicas (TMHMM) e as preditas a serem secretadas pelas vias clássica (SignalP) e não-clássica (SecretomeP).

**III. Busca 3: banco de dados das cepas CL
Brener e Sylvio X10/1 de *T. cruzi***

Nesta terceira busca, decidiu-se utilizar o banco de dados contendo 30.158 sequências relativas às cepas CL Brener (19.242 sequências) e Sylvio X10/1 (10.805 sequências) que já tiveram os genomas sequenciados. Ao contrário das buscas 1 e 2, foi utilizado o programa PEAKS para a identificação de proteínas e, no banco de dados, foram incluídas sequências de proteínas contaminantes como queratinas, albuminas e tripsinas. Estipulando FDR = 0,1%, a busca resultou em 2.947 proteínas identificadas com 2 ou mais peptídeos, das quais 1.091 proteínas anotadas como não caracterizadas no banco de dados foram detectadas. Essa foi a maior quantidade de proteínas identificadas entre as três buscas e, principalmente, entre todos os proteomas de *T. cruzi* publicados até o momento independentemente da forma de vida. Portanto, considerando a busca 3, este proteoma de formas epimastigotas da cepa G é o mais abrangente de *T. cruzi* já realizado. Este cenário demonstra também que a completude e a qualidade do banco de busca são fundamentais para identificarem-se proteínas no contexto dessa metodologia.

Apenas 21 proteínas identificadas na busca 1 foram comuns às outras buscas, lembrando que somente 22,67% das identificações pela busca 1 foram provenientes de *T. cruzi*, único organismo considerado nas buscas 2 e 3. Os bancos de dados das buscas 2 e 3 que apresentavam em comum sequências proteicas da cepa CL Brener no banco de dados compartilharam 474 identificações, metade do total de proteínas identificadas pela busca 2. Uma quantidade de 2.460 proteínas foram exclusivas da busca 3 (Figura 25).

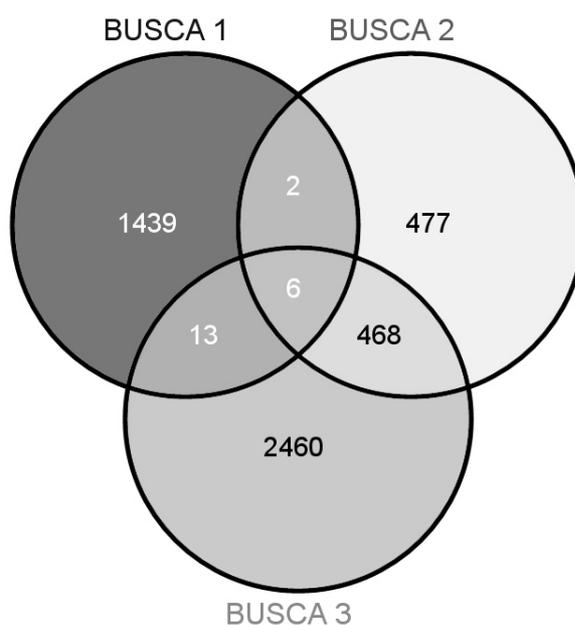


Figura 25. Diagrama de Venn correlacionado as proteínas identificadas nas três buscas realizadas. Busca 1: banco de dados de *Trypanosoma* spp. e outros tripanossomatídeos, busca 2: banco de dados de CL Brener e busca 3: banco de dados de CL Brener e Sylvio X10/1.

Quinze por cento das identificações decorreram de 2 peptídeos, 13,51% de 3 peptídeos e 11,16% de 4 peptídeos. O maior percentual de peptídeos por identificação 30,20% foi obtido com 10 ou mais peptídeos, percentual muito superior aos daqueles das outras duas buscas descritas anteriormente (Figura 26). Isso é muito satisfatório, pois significa que a cobertura das proteínas identificadas pelos peptídeos foi maior nesta busca, fornecendo identificações com maior acuidade. A proteína piruvato fosfato diquinase (Q4E0Q0, também presente na busca 2) foi identificada a partir de 111 peptídeos, quantidade bastante superior às identificações que advieram do maior número de peptídeos das buscas 1 e 2 (31 e 43 peptídeos, respectivamente). Os nomes de entrada dessa busca terminam em `_TRYCC` (CL Brener) ou em `_TRYCR` (Sylvio X10/1), mas a seguir serão apresentadas como números de acesso (forma simplificada).

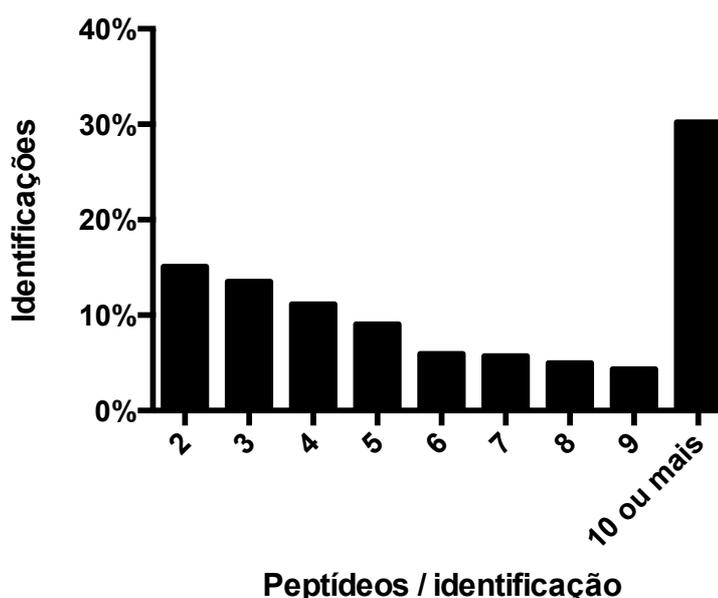


Figura 26. Relação entre o número de peptídeos por identificação e a porcentagem de identificações obtidas por meio de busca no banco de dados de *T. cruzi* CL Brener e Sylvio X10/1.

1. Classificação de *Gene Ontology* (GO)

O Blast2GO originou 6.317 anotações relativas a 1.785 proteínas distintas identificadas (60,57%), porcentagem similar àquela da busca 2 (69,78%).

Em relação à categoria componente celular, 28 termos GO foram assinados. O termo GO citoplasma englobou a maior porcentagem de anotações (16,03%), seguido por complexo proteico (15,71%), celular (12,18%), intracelular (9,90%), etc. (Figura 27A). Esses quatro termos GO de componente celular também foram os mais representativos das outras buscas.

Para processo biológico, 42 termos GO foram usados para anotar proteínas identificadas, dos quais os mais abundantes foram: processo metabólico (12,03%), processo catabólico (9,77%), tradução (8,66%), etc. (Figura 27B). Para todas as três buscas, os termos processo metabólico e processo metabólico contendo nucleobase foram enquadrados entre dos cinco termos mais bem representados.

A quantidade de 33 termos GO foram gerados na categoria função molecular. Com a maior porcentagem de anotações constou ligação a nucleotídeos (17,41%), depois figuraram atividade catalítica (13,45%), ligação a proteínas (12,23%), ligação (11,30%), etc., como visto na Figura 27C. Os cinco termos GO de função molecular mais representativos nesta busca também o foram nas buscas 1 e 2.

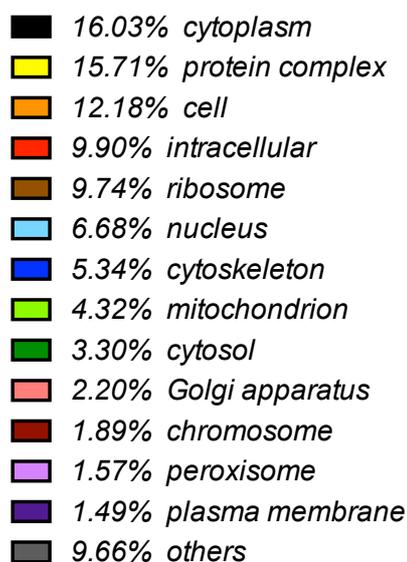
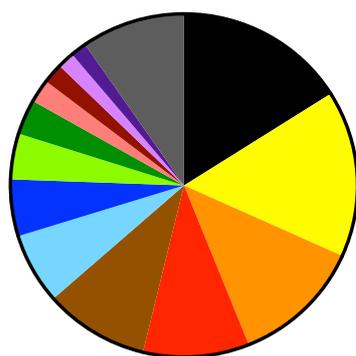
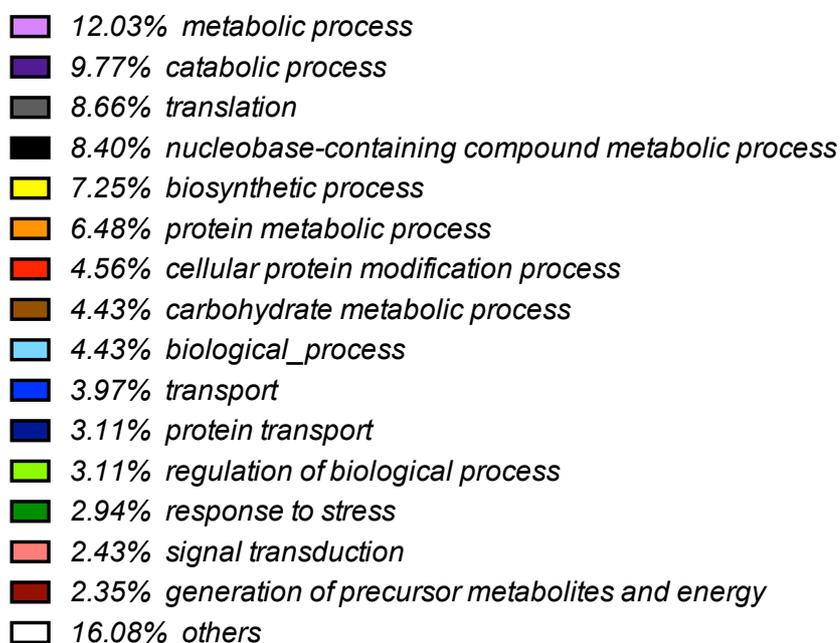
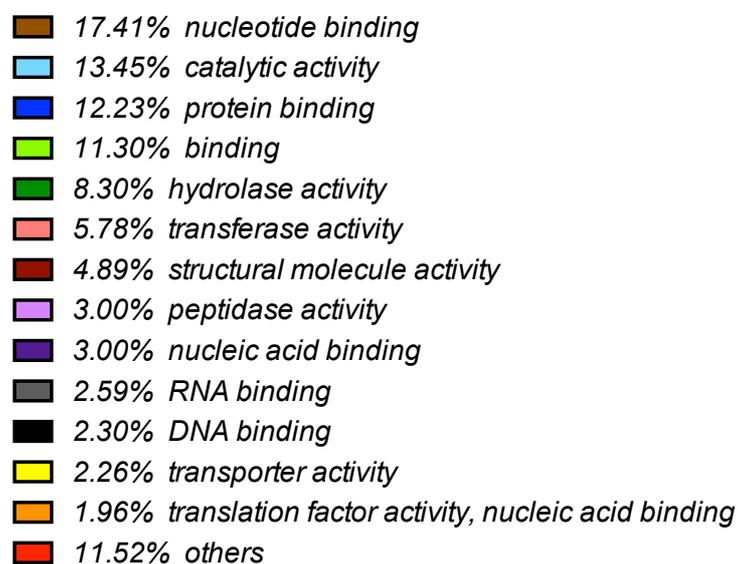
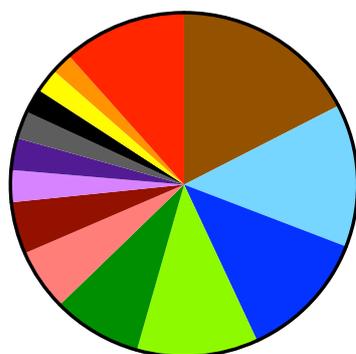
A**B****C**

Figura 27. Gráficos de setores para cada categoria de GO: (A) componente celular, (B) processo biológico e (C) função molecular, de acordo com o programa Blast2GO.

2. Predição de atividades enzimáticas

Um total de 687 anotações de EC foram atribuídas a 603 proteínas. A classe das transferases (EC 2) apresentou a maior quantidade de membros (n = 209). Nas três buscas, as classes mais abundantes foram as das hidrolases e das transferases.

Em relação ao número de atividades enzimáticas, EC 2 foi a mais representativa nesta análise (Tabela 10) assim como nas outras análises. Entre todas as classes de EC desta busca, a atividade enzimática proteína Ser/Thr-quinase (EC 2.7.11.0) teve a maior representatividade, com 26 anotações.

Tabela 10. Abundância de anotações e de atividades enzimáticas para cada uma das seis classes de EC.

| Classe de EC | Anotações ^a | Atividades enzimáticas ^b |
|---------------------|------------------------|-------------------------------------|
| 1 - Óxido-redutases | 140 | 56 |
| 2 - Transferases | 209 | 73 |
| 3 - Hidrolases | 198 | 45 |
| 4 - Liases | 27 | 14 |
| 5 - Isomerases | 42 | 17 |
| 6 - Ligases | 71 | 37 |
| Total | 687 | 242 |

^a Número de anotações de EC pelo Blast2GO

^b Número de *Enzyme* ID distintos do sistema EC

2.1 Peptidases (EC 3.4) putativas

Na subclasse das peptidases, foram classificadas 59 proteínas (Tabela 11)., Uma metionina-aminopeptidase putativa, uma dipeptidil-peptidase putativa e várias aminopeptidases, inclusive a LAP^{Tc} caracterizada pelo nosso grupo de pesquisa (Q4DZJ3) (Cadavid-Restrepo *et al.*, 2011), estão na sub-subclasse das aminopeptidases. Uma dessas aminopeptidases putativas (Q4E686) também foi identificada na busca 2. Consta também a anotação nessa categoria de uma proteína candidata a fator de transcrição do gene *SL* que leva a síntese de SL RNA e esse depois é adicionado a mRNAs maduros por *trans-splicing*. Um mesmo tipo de fator de transcrição, PPB1, anotado sob o número de acesso Q9GU26 foi identificado por Wen *et al.* (2000).

Como serino-carboxipeptidase, foi anotada uma serino-carboxipeptidase putativa (CBP1), assim como na busca 1. Entre as metalo-carboxipeptidases, estão duas proteínas não caracterizadas e uma carboxipeptidase termostável 1 putativa.

Duas ubiquitina C-terminal hidrolases putativas surgem anotadas como ômega-peptidases, das quais uma também foi encontrada na busca 2, visto que trata-se de uma proteína do banco de dados da cepa CL Brener.

Uma OPB e duas POP foram identificadas e cada uma dessas POP foi proveniente de uma cepa considerada no banco de dados. Além disso, estão presentes duas proteínas romboides-símiles putativas na sub-clubclasse das serino-endopeptidases. Essas proteínas romboides estão envolvidas no processo de proteólise intramembrana regulada, com o qual algumas proteínas integrais de membrana inativas na forma associada à membrana sofrem uma clivagem no domínio transmembrânico delas. Isso libera o domínio funcional citoplasmático ou luminal/extracelular (revisadas em Erez *et al.* (2009)). Em *T. cruzi*, essas serino-proteases ainda não foram investigadas.

A catepsina B-símile, mais uma protease bem estudada pelo nosso grupo (Garcia *et al.*, 1998; Nobrega *et al.*, 1998; Cadavid-Restrepo, 2005), foi identificada neste proteoma. Inicialmente, uma cisteíno-protease foi purificada de epimastigotas da cepa Tulahuen e possuía aproximadamente 30 kDa, era glicosilada, mostrava atividade ótima em pH ácido para a hidrólise de BSA e colágeno tipo I (pH 4,0) e de substrato fluorogênico (pH 6.5) e não se tratava de uma isoforma ou produto da cruzipaina. Essa atividade enzimática também foi detectada em extratos de tripomastigotas e amastigotas. A sequência da extremidade N-terminal dessa enzima apresentava similaridade com catepsinas B de outros organismos (Garcia *et al.*, 1998). Concomitantemente, a presença de um gene cópia simples que codifica a catepsina B-símile foi confirmada em *T. cruzi* (Nobrega *et al.*, 1998). Depois foi mostrado que a catepsina B-símile localizou-se em grânulos em epimastigotas e amastigotas, provavelmente lisossomos, e mediou a principal atividade hidrolítica em gelatina sob condições ácidas que foi maior em epimastigotas do que em amastigotas e tripomastigotas (Cadavid-Restrepo, 2005).

Cruzipaina ou cruzaina (revisada em Jose Cazzulo *et al.* (2001) e Alvarez *et al.* (2012)), uma catepsina L-símile e a mais bem caracterizada cisteíno-peptidase de *T. cruzi*, é codificada por um grande número de genes arranjados *in tandem*, atua mais eficientemente em pH ácido e é expressa nas quatro formas do parasito. Em

epimastigotas é mais especificamente localizada nos reservossomos e, em tripomastigotas, algumas isoformas da cruzipaina são secretadas no meio extracelular, relevantemente associada a fator de virulência na doença de Chagas. A cruzipaina e outras catepsinas L-símiles putativas foram identificadas na cepa G.

Entre as metalo-endopeptidases, uma thimet oligopeptidase, uma pitrilisina-símile, três metalopeptidases dependentes de zinco e ATP e uma peptidase putativas e duas proteínas não caracterizadas foram anotadas.

O total de 21 proteínas foram classificadas como treonino-endopeptidases, incluindo diversas subunidades α e β do proteassoma 20S e do complexo HsIVU, tipos de proteínas também anotados na busca 2.

Tabela 11. Proteínas anotadas para a atividade de peptidase (EC 3.4) distribuídas em oito sub-subclasses de EC. (continua)

| Sub-subclasse de EC | Descrição^a | Código identificador^b |
|--|---|---|
| | <i>Spliced leader RNA PSE-promoter transcription factor</i> | K4DW76 |
| | <i>Methionine aminopeptidase</i> | K4DRB9 |
| | <i>Aminopeptidase</i> | K4DX11 |
| | <i>Aminopeptidase P</i> | K4EAQ0 |
| | <i>Aminopeptidase</i> | K4DTU9 |
| Aminopeptidases (EC 3.4.11) | <i>Aminopeptidase</i> | K4EE31 |
| | <i>Aminopeptidase</i> | K4E6R1 |
| | <i>Aminopeptidase</i> | Q4DZJ3 |
| | <i>Aminopeptidase</i> | Q4CXX8 |
| | <i>Aminopeptidase</i> | Q4E686 |
| | <i>Peptidase T</i> | K4DNS4 |
| | <i>Dipeptidyl-peptidase</i> | K4E7R5 |
| Serino-carboxipeptidases (EC 3.4.16) | <i>Serine carboxypeptidase (CBP1)</i> | K4DS29 |
| Metallo-carboxipeptidases (EC 3.4.17) | <i>Uncharacterized protein</i> | K4DZL8 |
| | <i>Uncharacterized protein</i> | K4DXE9 |
| | <i>Thermostable carboxypeptidase 1</i> | Q4D9T0 |
| Ômega-peptidases (EC 3.4.19) | <i>Ubiquitin C-terminal hydrolase</i> | K4DTY6 |
| | <i>Ubiquitin C-terminal hydrolase</i> | Q4E680 |
| Serino-endopeptidases (EC 3.4.21) | <i>Oligopeptidase B</i> | K4DWV0 |
| | <i>Prolyl oligopeptidase</i> | K4E2C3 |
| | <i>Rhomboid-like protein</i> | K4DTW8 |
| | <i>Prolyl oligopeptidase</i> | Q4E132 |
| | <i>Rhomboid-like protein</i> | Q4CV15 |
| Cisteíno-endopeptidases (EC 3.4.22) | <i>Cathepsin B-like protease</i> | O61066 |
| | <i>Cathepsin L-like</i> | K4DV55 |
| | <i>Cathepsin L-like</i> | K4EEA1 |
| | <i>Cysteine peptidase</i> | Q4CMU6 |
| | <i>Cysteine peptidase</i> | Q4E0J7 |

Tabela 11. Anotações para a atividade de peptidase (EC 3.4).

(conclusão)

| Sub-subclasse de EC | Descrição ^a | Código identificador ^b |
|--|--|-----------------------------------|
| Cisteíno-endopeptidases (EC 3.4.22) (continuação) | <i>Cruzipain</i> | Q4DW02 |
| | <i>ATP-dependent zinc metallopeptidase</i> | K4E2P5 |
| | <i>Peptidase</i> | K4E718 |
| | <i>Pitriysin-like metalloprotease</i> | K4DLV5 |
| Metallo-endopeptidases (EC 3.4.24) | <i>Uncharacterized protein</i> | K4DYL4 |
| | <i>Thimet oligopeptidase</i> | K4EBV9 |
| | <i>ATP-dependent zinc metallopeptidase</i> | Q4DJU4 |
| | <i>ATP-dependent zinc metallopeptidase</i> | Q4DEY4 |
| | <i>Uncharacterized protein</i> | Q4DM81 |
| | <i>Proteasome subunit alpha type</i> | K4DKH0 |
| | <i>Proteasome subunit alpha type</i> | K4DY50 |
| | <i>Proteasome subunit alpha type</i> | K4DN23 |
| | <i>Proteasome subunit alpha type</i> | K4EAE2 |
| | <i>Proteasome subunit alpha type</i> | K4DNN2 |
| | <i>Proteasome beta 3 subunit</i> | K4E984 |
| | <i>Proteasome subunit beta type</i> | K4DXV6 |
| | <i>Proteasome subunit beta type</i> | K4E4C0 |
| | <i>Proteasome subunit beta type</i> | K4E0S4 |
| | Treonino-endopeptidases (EC 3.4.25) | <i>Proteasome beta 2 subunit</i> |
| <i>Proteasome alpha 7 subunit</i> | | K4E413 |
| <i>Proteasome subunit beta type</i> | | K4DUM4 |
| <i>Hslvu proteolytic protein</i> | | K4DRE8 |
| <i>Proteasome beta 2 subunit</i> | | Q4CU77 |
| <i>Hslvu complex proteolytic subunit-like</i> | | Q4D729 |
| <i>Proteasome subunit alpha type</i> | | Q4DIK4 |
| <i>Proteasome subunit beta type</i> | | Q4D579 |
| <i>Proteasome subunit alpha type</i> | | Q4DAW6 |
| <i>Proteasome subunit beta type</i> | | Q4E0L9 |
| <i>Proteasome subunit beta type</i> | | Q4D8U7 |
| <i>Proteasome subunit beta type</i> | Q4D6T6 | |

^a Descrição da proteína anotada proveniente do Uniprot^b Número de acesso do Uniprot

2.2. Predição de enzimas envolvidas em vias metabólicas

Noventa e uma vias metabólicas foram preditas por KEGG *pathways* - Blast2GO para as identificações desta busca. Considerando o número de proteínas anotadas em cada via, a via mais representativa foi metabolismo de purinas (n = 97). Assim, para todas as buscas realizadas, essa via constou em primeiro lugar de acordo com essa classificação, condizente com o estágio de vida replicativo de *T. cruzi* analisado. Ambas as vias do ciclo do ácido cítrico e do metabolismo de cisteína e metionina apareceram em segundo lugar em número de anotações (n = 42). Com a maior quantidade de atividades enzimáticas distintas, apresentou-se a via de metabolismo de purinas com 21 atividades e, em seguida, biossíntese de aminoacil-tRNA com 18 (todas de ligase com ECs distintos) (Tabela 12). Assim como na busca 2, entre as dez vias mais representativas, estão presentes o ciclo do ácido cítrico, a glicólise/gliconeogênese, a fosforilação oxidativa, relacionadas com metabolismo energético. Esse resultado sugere que altos níveis de metabolismo energético são indispensáveis para a replicação de formas epimastigotas. As atividades enzimáticas e os números de acesso das cinco vias metabólicas mais representativas em número de anotações estão explicitados na Tabela 13.

Tabela 12. As dez vias metabólicas mais abundantes em número de anotações e os respectivos números de atividades enzimáticas.

| Via metabólica ^a | Anotações ^b | Atividades enzimáticas ^c |
|-------------------------------------|------------------------|-------------------------------------|
| Metabolismo de purinas | 97 | 21 |
| Ciclo do ácido cítrico | 42 | 13 |
| Metabolismo de cisteína e metionina | 42 | 11 |
| Biossíntese de aminoacil-tRNA | 36 | 18 |
| Metabolismo de arginina e prolina | 35 | 16 |
| Metabolismo de fenilalanina | 31 | 6 |
| Glicólise/Gliconeogênese | 30 | 17 |
| Metabolismo de glutatona | 28 | 13 |
| Fosforilação oxidativa | 28 | 6 |
| Metabolismo de pirimidinas | 28 | 12 |

^a Sugerida por KEGG *pathways* - Blast2GO

^b Número de anotações (pode haver uma proteína anotada mais de uma vez na mesma via metabólica)

^c Número de *Enzyme* ID distintos do sistema EC

2.2.1. Metabolismo de purinas

Anotaram-se 96 proteínas diferentes da via de metabolismo de purinas. Foram identificadas 3 PRPP sintetases putativas (K4DYU6, K4E0Q6 e Q4DXW9) e todas as enzimas que estão marcadas no esquema da via de recuperação de purinas (Figura 28). Por esse esquema, percebe-se que uma quantidade muito maior de proteínas pertencentes a essa via foram identificadas em relação às outras buscas.

Onze DNA polimerases mitocondriais do tipo β (família X) e I (família A) (K4E8I0, K4DJT0, K4DK04, K4E5I5, K4DK54, Q4CQE0, Q4E024, Q4E023, Q4D4N5, Q4DSW2 (identificada na busca 2) e Q4D7R3) e duas RNA polimerases dirigidas a DNA putativas (K4E797 e K4E6Y0) foram identificadas.

Adenilato quinases (K4DNU5, K4E1Y3, Q4DLV1 e Q4DXL7), guanilato quinase (Q4DPB1), fosfodiesterases cAMP-específicas (K4DUC7, K4DUK1, Q4CQR0 e Q4E0Y6 (identificada na busca 2)), nucleosídeo difosfato quinase (Q4E256), ribose fosfato pirofosfoquinase (K4DK81), ribonucleosídeo-disfostato redutases (K4E6U8, K4DPY7 e Q4D6V2) putativas também constam nessa classificação.

Enquadradas como fosfatases (EC 3.6.1.3 (ATPase: adenosina trifosfatase) e 3.6.1.15 (nucleosídeo-trifosfato fosfatase) - Tabela 13), foram apresentadas 46 proteínas, das quais alguns tipos são: proteína E de resistência a multidrogas, dineínas, proteínas do proteassoma, HSP, transportadores ABC, proteases Clp, diversas ATPases, metalo-peptidases e katanina putativas.

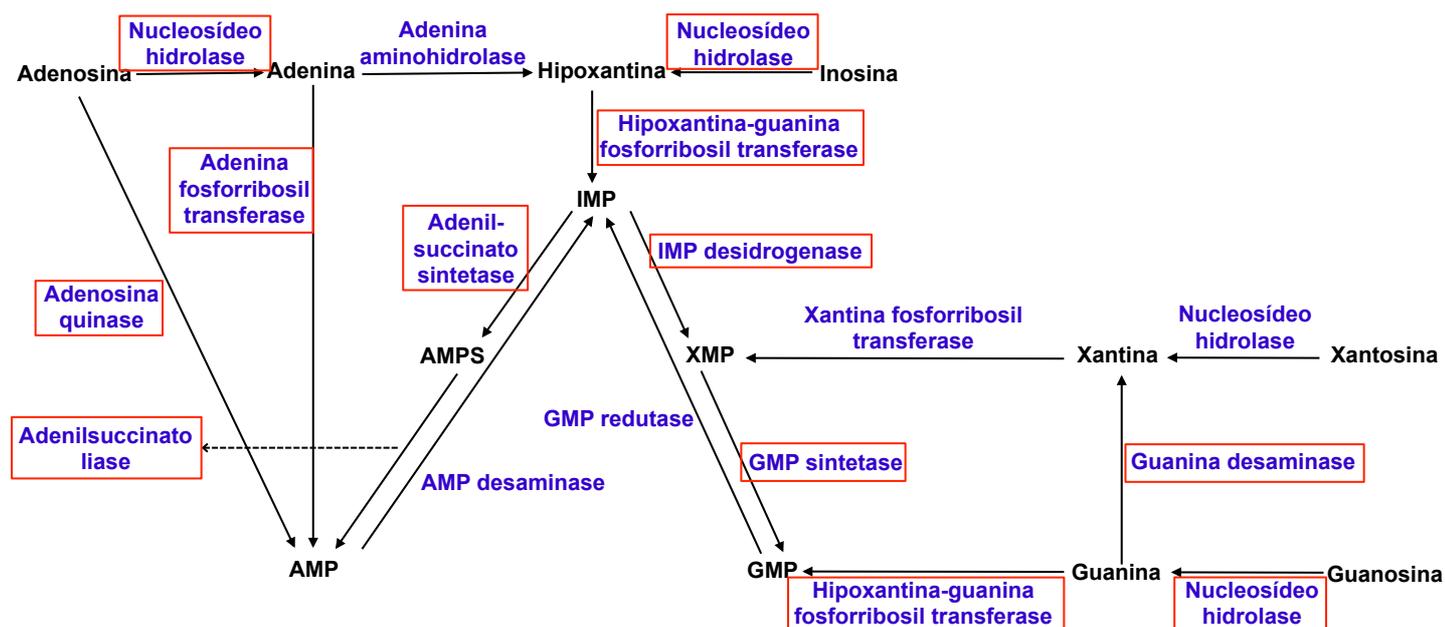


Figura 28. Esquema da via de recuperação/salvação das purinas predita em *Trypanosoma spp.*, com base nas enzimas descritas para *Leishmania spp.*

As enzimas marcadas em vermelho correspondem às proteínas identificadas na cepa G de *T. cruzi* pela busca 3. Modificado de Boitz *et al.* (2013).

AMP: adenilato, AMPS: adenilsuccinato, GMP: guanilato, IMP: inosinato, XMP: xantinilato.

2.2.2. Ciclo do ácido cítrico

Assim como na busca 2, todas as enzimas do complexo da piruvato desidrogenase foram identificadas (subunidade E1 da piruvato desidrogenase (K4DWP2), diidrolipoil desidrogenases (K4DZ35, Q4DWL8 e Q4DD33 (identificada na busca 2)) e diidrolipoil transacetilases (Q4DYI5 e Q4DZT8) putativas). As outras proteínas envolvidas no ciclo que foram identificadas estão destacadas na Figura 28, incluindo três isocitrato desidrogenases, das quais uma foi a isocitrato desidrogenase citosólica Tc_IDH2 (Q4E4L7) também detectada na busca 2. Por não se saber sobre a localização das outras IDH putativas identificadas, na Figura 29, a enzima número 11 correspondente à IDH mitocondrial também foi destacada. Uma fosfoenolpiruvato carboxiquinase glicossomal putativa (K4DSA9) envolvida na reposição de intermediários do ciclo do ácido cítrico também foi constatada. Interessantemente, uma oxoisovalerato desidrogenase putativa (K4DW66) que participa da via de degradação de valina foi listada no ciclo do ácido cítrico.

Duas aconitases (Q4E5G5 e K4DWM5) e cinco succinato desidrogenases (K4E5H9, Q4D4X1, K4EA42, K4E2R8 e Q4CVH2) putativas (enzimas 2 e 6 da Figura 29, respectivamente) foram identificadas, mas não foram anotadas nesta via pelo Blast2GO. Nos esquemas de vias metabólicas, as proteínas destacadas correspondem a identificações que foram anotadas na via por KEGG *pathways* - Blast2GO e, como no caso desses dois tipos de enzimas do ciclo do ácido cítrico, pode ocorrer de uma proteína putativa ter sido identificada, porém não anotada na respectiva via.

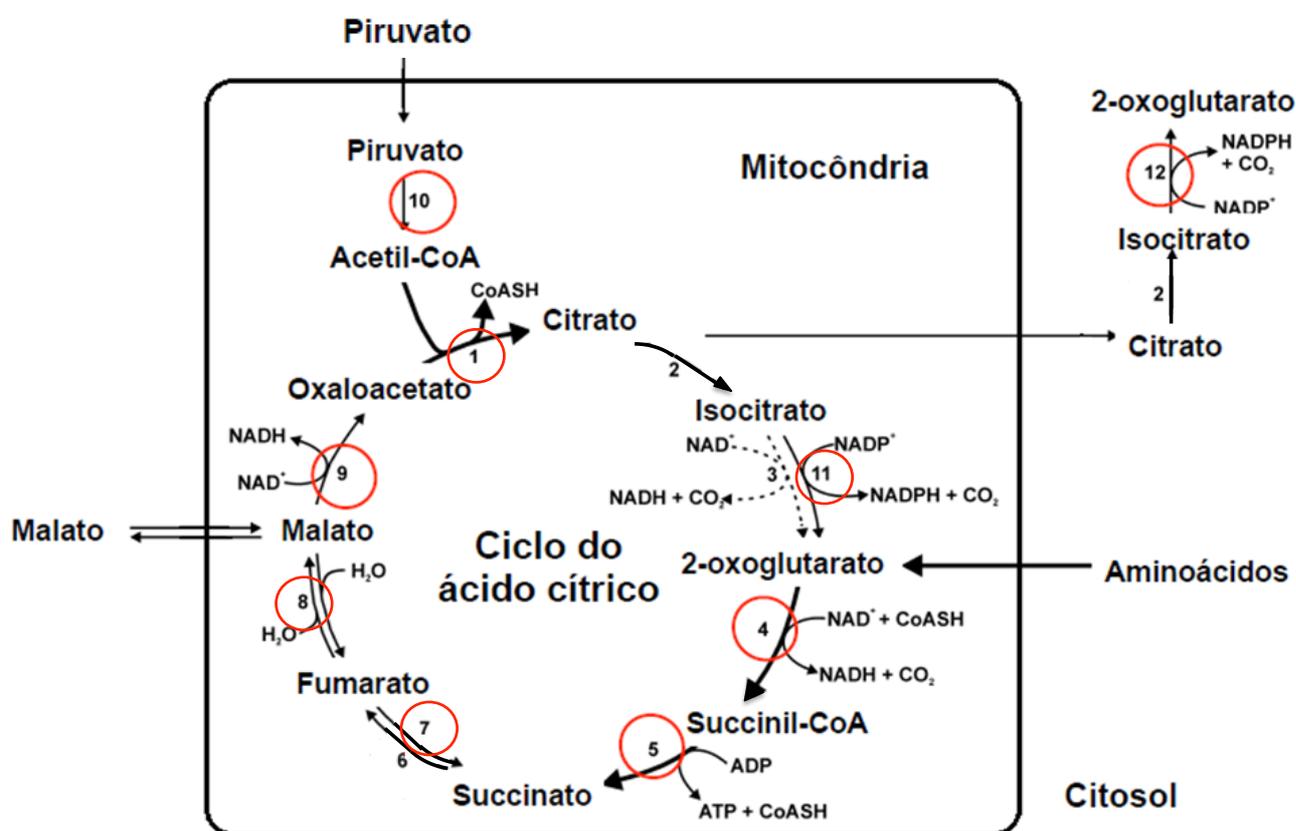


Figura 29. Representação esquemática da possível organização do ciclo do ácido cítrico em *T. cruzi*. As principais reações estão representadas por setas grossas. Os números das enzimas circulados em vermelho correspondem às proteínas identificadas na cepa G de *T. cruzi*. Modificado de Leroux *et al.* (2011).

1: citrato sintase, 2: aconitase, 3: isocitrato desidrogenase dependente de NAD⁺ (ausente em tripanossomas), 4: 2-oxoglutarato desidrogenase, 5: succinil-CoA sintetase, 6: succinato desidrogenase, 7: fumarato redutase dependente de NADH mitocondrial, 8: fumarato hidratase mitocondrial, 9: malato desidrogenase mitocondrial, 10: complexo da piruvato desidrogenase, 11: isocitrato desidrogenase dependente de NADP⁺ mitocondrial, 12: isocitrato desidrogenase dependente de NADP⁺ citosólica.

2.2.3. Metabolismo de cisteína e metionina

Relacionadas ao metabolismo de cisteína e metionina, estão 27 proteínas putativas: 1,2-diidroxi-3-ceto-5-metiltiopenteno dioxigenases (K4E3Q7, K4DZ96 e Q4CVG3), metiltioadenosina fosforilases (K4E900, Q4DZW1 e Q4DQ97), espermidina sintases (K4DV45 e Q4DA73), cisteína sintases (K4E8E0, K4DRJ9, Q4CST7 e Q4CXR9), cistationa beta-sintase (Q4DF36), S-adenosilmetionina sintases (K4E4J2 e Q4CSC4), aspartato aminotransferases (K4EA87 e Q4D1Q4), tirosina aminotransferases (K4DZN8, Q4E4E7, Q4D3Y8 e Q4CVI0), adenosilhomocisteinases (K4DNN4, Q4D455 e Q4DNQ7), S-adenosilmetionina descarboxilases (K4E603, Q4DQQ5 e Q4CWQ0). Colocaremos em evidência a via metabólica do nucleosídeo metiltioadenosina (MTA). A enzima metiltioadenosina fosforilase de *T. cruzi* (TcMTAF) estudada pelo nosso laboratório (Neves, 2006) é responsável pela clivagem de nucleosídeos purínicos, principalmente o nucleosídeo modificado MTA (Figura 30). O MTA está relacionado à síntese de poliaminas, salvação de purinas e de metionina, ele é clivado pela MTAF em adenina e metiltioribose-1-fosfato (MTR1P). A adenina é regenerada a ATP e o MTR1P é regenerado à metionina, através de cascatas enzimáticas. TcMTAF é diferencialmente expressa em epimastigotas, tripomastigotas e amastigotas.

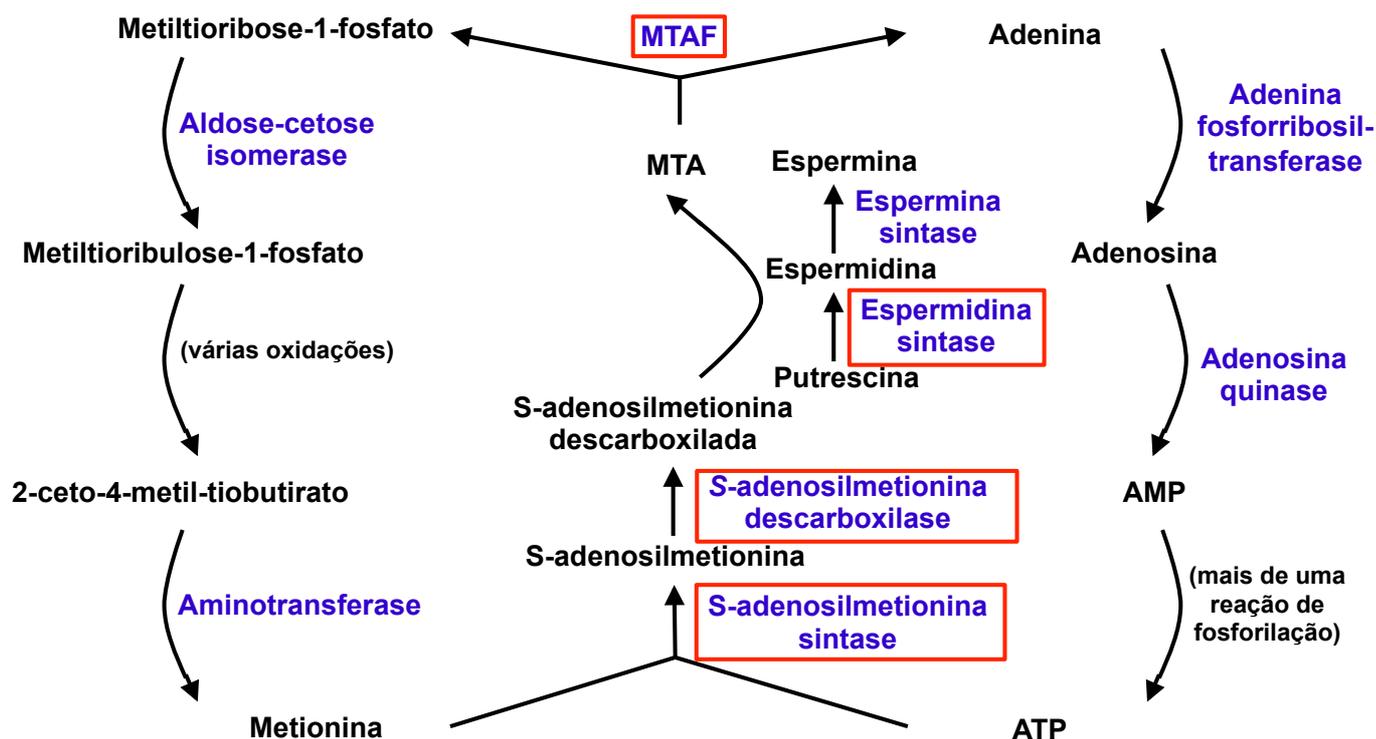


Figura 30. Representação esquemática da via de síntese e de metabolismo de metiltioadenosina (MTA). As proteínas identificadas no proteoma da cepa G (destacadas em vermelho) estão relacionadas à via de síntese de MTA. Em duas reações sucessivas, espermidina sintase e espermina sintase transferem o grupo propilamina da S-adenosilmetionina descarboxilada para putrescina e espermidina, respectivamente. Modificado de Neves (2006).

2.2.4. Biossíntese de aminoacil-tRNA

Uma diversidade de aminoacil-tRNA sintetases ainda maior do que a da busca 2 foi obtida com esta terceira busca. Um total de 18 tipos de aminoacil-tRNA sintetases putativas foi identificado: triptofanil- (K4DJD4 e Q4CTS7), treonil- (K4DU79), leucil- (K4E646, Q4CTR0, Q4CL52 e K4DRR4), isoleucil- (K4E228), prolil- (K4DWJ7, Q4DLK5 e Q4DZZ4), isoleucil- (Q4D2F2), alanil- (K4E8Q2, Q4DQ33 e Q4D747), valil- (K4E5D5), seril- (K4EAD1), aspartil- (K4E578), glicil- (K4DNM9 e Q4D449), cisteinil- (K4DUH4 e Q4CUZ9), glutamil- (K4E4Q3), glutaminil- (K4DUA1), arginil- (K4DZT5), fenilalanil- (K4E2G9 e Q4DY45), histidil- (K4E5W2 e Q4DA54) e asparaginil- (K4E3K3 e Q4E595) tRNA sintetases (Figura 31).

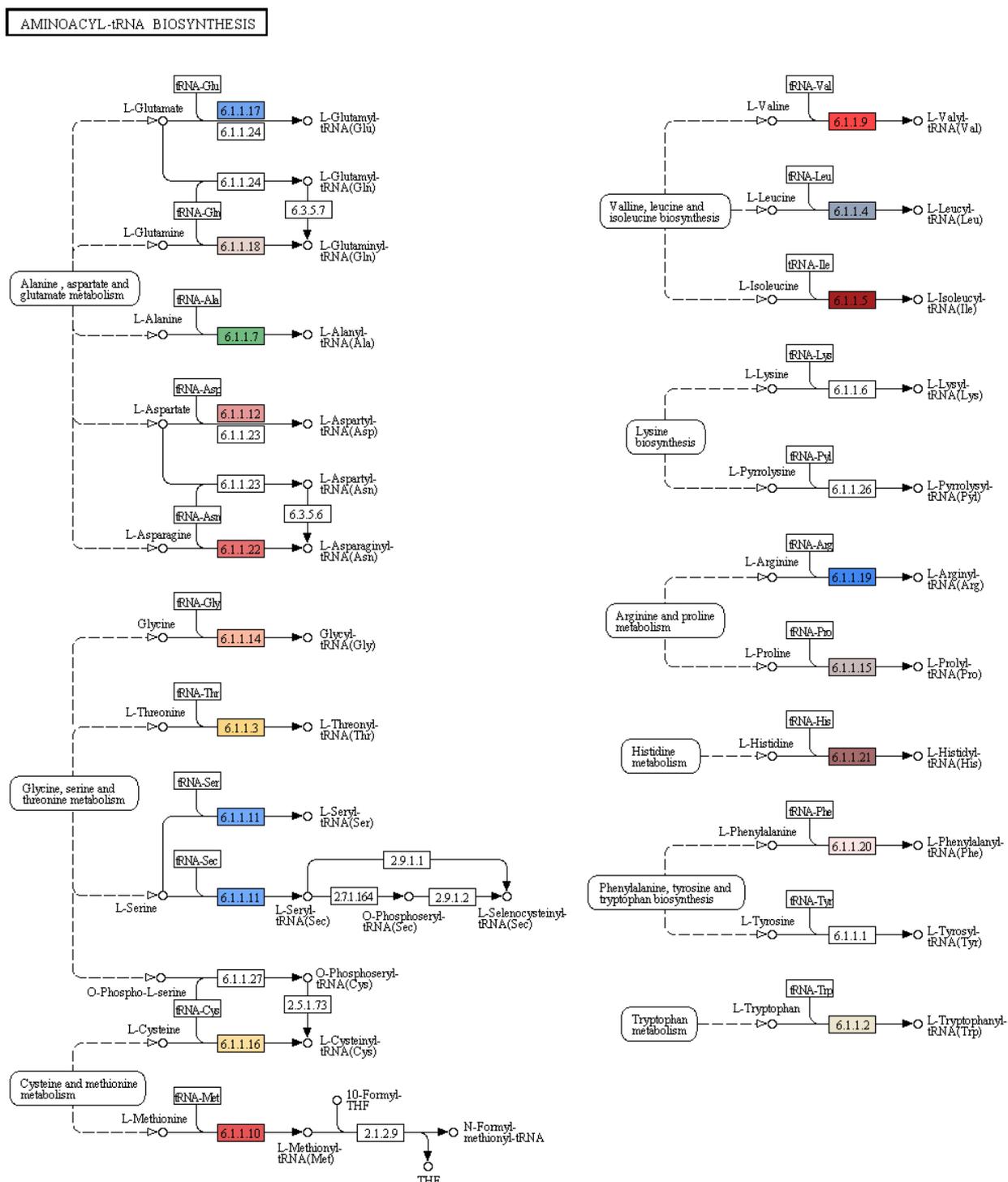


Figura 31. Representação esquemática da via de biossíntese de aminoacil-tRNA gerada pelo programa Blast2GO. Os retângulos em destaque indicam as proteínas identificadas na cepa G de *T. cruzi*.

2.2.5. Metabolismo de arginina e prolina

Por meio da busca no banco de dados de *T. cruzi* CL Brener, uma pirrolina-5-carboxilato redutase putativa (Q4DH60) foi anotada. Nesta busca, a identificamos e

também outra pirrolina-5-carboxilato redutase (K4EAX2) proveniente de Sylvio X10/1. Quatro tirosina aminotransferases (K4DZN8, Q4E4E7, Q4D3Y8 e Q4CVI0), três S-adenosilmetionina descarboxilases (K4E603, Q4DQQ5 e Q4CWQ0) e duas espermidina sintases (K4DV45 e Q4DA73) também anotadas na via de metabolismo de cisteína e metionina; duas pirrolina-5-carboxilato sintetases (K4E0F1 e Q4D112), uma aldeído desidrogenase (Q4DCP7), uma glutamato desidrogenase (K4DUG5), uma calmodulina (K4EBN9), uma prolina oxidase (Q4CVA1), duas aspartato aminotransferases (K4EA87 e Q4D1Q4), uma arginino-quinase (K4DUF3), duas aminoacilases (K4DQT8 e K4DQX8), duas acetilornitina desacetilases (K4DTG9 e Q4D7V2), cinco glutamil carboxipeptidases (K4DKW8, K4ED56, Q4CPW5, Q4CR09 e Q4CZY5), duas glutamina sintetases (K4E9B6 e Q4CN04) e uma carbamil fosfato sintetase (K4DK07) foram as outras enzimas putativas relacionadas a esta via.

Tabela 13. Relação entre as vias metabólicas mais representativas em número de anotações e as respectivas atividades enzimáticas. (continua)

| Via metabólica ^a | Atividade enzimática ^b | Código identificador ^c |
|-----------------------------|---------------------------------------|--|
| | Desidrogenase (EC 1.1.1.205) | K4DPJ9, K4DXK1 |
| | Redutase (EC 1.17.4.1) | K4E6U8, K4DPY7, Q4D6V2 |
| | Fosforribosiltransferase (EC 2.4.2.7) | K4E774, K4E8H9, Q4DVM8, Q4DNZ3 |
| | Fosforibosiltransferase (EC 2.4.2.8) | K4DRA8, Q4DGA2 ^d |
| | Fosforibosiltransferase (EC 2.4.2.22) | Q4DGA2 ^d |
| | Quinase (EC 2.7.1.20) | K4E5U6 |
| | Quinase (EC 2.7.1.40) | K4DWR8 |
| | Quinase (EC 2.7.4.3) | K4DST1, K4DNU5, K4E1Y3, Q4DLV1, Q4DXL7 |
| | Quinase (EC 2.7.4.6) | K4DZS6, Q4E256, Q4DBS8 |
| | Quinase (EC 2.7.4.8) | Q4DPB1 |
| | Difosfoquinase (EC 2.7.6.1) | K4DYU6, K4E0Q6, K4DK81, Q4DXW9 |
| | RNA polimerase (EC 2.7.7.6) | K4E797, K4E6Y0 |
| | DNA polimerase (EC 2.7.7.7) | K4E8I0, K4DJT0, K4DK04, K4E5I5, K4DK54, Q4CQE0, Q4E024, Q4E023, Q4D4N5, Q4DSW2, Q4D7R3 |
| 1) Metabolismo de purinas | Fosfodiesterase (EC 3.1.4.17) | K4DUC7, K4DUK1, Q4CQR0, Q4E0Y6 |
| | Nucleosidase (EC 3.2.2.1) | K4E1L0 |
| | Desaminase (EC 3.5.4.3) | K4DTI7 |
| | Adenil pirofosfatase (EC 3.6.1.3) | K4E642, K4DYP0, K4EEB0, K4DZM2, K4DJQ8, K4E7N2, K4E3W3, K4E4I4, K4E6S4, K4E0Z0, K4E6U6, K4EDU2, K4DJH5, K4EAL1, K4E7U0, Q4D0B9, Q4DEP1, Q4D9J1, Q4DX41, Q4CPR5, Q4CQ22, Q4DZM5, Q4DFN9, Q4DHQ7, Q4DWA3 |
| | Fosfatase (EC 3.6.1.15) | K4ECZ0, K4E2P5, K4EBZ4, K4E670, K4E072, K4DUS1, K4DVN7, K4DYN3, K4DYL4, K4E4J7, K4DUP0, Q4DJU4, Q4CLP1, Q4DEY4, Q4DWB5, Q4DTR4, Q4DW08, Q4DSH8, Q4DM81, Q4E0K2, Q4D0R7 |
| | Liase (EC 4.3.2.2) | K4DU00, Q4CXJ7, Q4DBH5 |
| | Sintetase (EC 6.3.4.4) | K4E5I9 |
| | Sintetase (EC 6.3.5.2) | K4EBX4 |

Tabela 13. Relação entre as vias metabólicas mais representativas em número de anotações e as respectivas atividades enzimáticas. (continuação)

| Via metabólica ^a | Atividade enzimática ^b | Código identificador ^c |
|--|---|---|
| | Desidrogenase (EC 1.1.1.37) | O61084, K4E1P3, K4E6H1, Q4DDK8, Q4D4A0, Q4DRD8, Q4D123 |
| | Desidrogenase (EC 1.1.1.42) | K4DIS2, K4DUK8, Q4E4L7 |
| | Desidrogenase (EC 1.2.4.1) | K4DW66, K4DWP2 |
| | Desidrogenase (EC 1.2.4.2) | K4DRJ2, K4DZA9 ^e , Q4CTE8, Q4DKY2 |
| | Desidrogenase (EC 1.3.99.1) | K4DZS8, K4DZP3, K4DQV8, K4DZK3, Q4CMU8, Q4D778, Q4CYX1, Q4D779, Q4CQU6 |
| 2) Ciclo do ácido cítrico | Desidrogenase (EC 1.8.1.4) | K4DZ35, Q4DWL8, Q4DD33 |
| | Acetiltransferase (EC 2.3.1.12) | K4DZZ0, Q4DZT8 |
| | Sintetase (EC 2.3.3.1) | K4DXA2 |
| | Citrato sintetase (EC 2.3.3.8) | K4DXC1, Q4E2I6 ^f |
| | Carboxiquinase (EC 4.1.1.49) | K4DSA9 ^e |
| | Hidratase (EC 4.2.1.2) | K4EA83, Q4DRX1, Q4CQW5 |
| | Ligase (EC 6.2.1.4) | K4DXC1 ^g , Q4E2I6 ^f |
| | Ligase (EC 6.2.1.5) | K4DXC1 ^g , K4E0F8, Q4E2I6 ^f |
| | Dioxigenase (EC 1.13.11.54) | K4E3Q7, K4DZ96, Q4CVG3, Q4DQC4 |
| | Fosforilase (EC 2.4.2.28) | K4E900, Q4DZW1, Q4DQ97 |
| | Sintetase (EC 2.5.1.16) | K4DV45, Q4DA73 |
| | Sintetase (EC 2.5.1.47) | K4E8E0 ^h , K4DRJ9, Q4DF36, Q4CST7, Q4CXR9 ⁱ |
| | Adenosiltransferase (EC 2.5.1.6) | K4E4J2, Q4CSC4 |
| 3) Metabolismo de cisteína e metionina | Transaminase (EC 2.6.1.1) | K4EA87 ^j , K4DZN8 ^k , Q4E4E7 ^l , Q4D3Y8 ^m , Q4CVI0 ⁿ , Q4D1Q4 ^o |
| | Transaminase (EC 2.6.1.5) | K4EA87 ^j , K4DZN8 ^k , Q4E4E7 ^l , Q4D3Y8 ^k , Q4CVI0 ^l , Q4D1Q4 ^m |
| | Transaminase (EC 2.6.1.57) | K4EA87 ^j , K4DZN8 ^k , Q4E4E7 ^l , Q4D3Y8 ^m , Q4CVI0 ⁿ , Q4D1Q4 ^o |
| | S-adenosilhomocisteína sintetase (EC 3.3.1.1) | K4DNN4, Q4D455, Q4DNQ7 |
| | Descarboxilase (EC 4.1.1.50) | K4E603, Q4DQQ5, Q4CWQ0 |
| | Sintetase (EC 4.2.1.22) | K4E8E0 ^h , Q4CXR9 ⁱ |

Tabela 13. Relação entre as vias metabólicas mais representativas em número de anotações e as respectivas atividades enzimáticas. (conclusão)

| Via metabólica ^a | Atividade enzimática ^b | Código identificador ^c |
|--------------------------------------|-----------------------------------|--|
| | Ligase (EC 6.1.1.2) | K4DJD4, Q4CTS7 |
| | Ligase (EC 6.1.1.3) | K4DU79 |
| | Ligase (EC 6.1.1.4) | K4E646, Q4CTR0, Q4CL52 |
| | Ligase (EC 6.1.1.5) | K4E228, K4DWJ7 ^p , Q4D2F2, Q4DLK5 ^q , Q4DZZ4 |
| | Ligase (EC 6.1.1.7) | K4E8Q2, Q4DQ33, Q4D747, Q4DZZ4 ^f |
| | Ligase (EC 6.1.1.9) | K4E5D5 |
| | Ligase (EC 6.1.1.10) | K4DRR4 |
| 4) Biossíntese de aminoacil-tRNA | Ligase (EC 6.1.1.11) | K4EAD1 |
| | Ligase (EC 6.1.1.12) | K4E578 |
| | Ligase (EC 6.1.1.14) | K4DNM9, Q4D449 |
| | Ligase (EC 6.1.1.15) | K4DWJ7 ^p , Q4DLK5 ^q , Q4DZZ4 ^f |
| | Ligase (EC 6.1.1.16) | K4DUH4, Q4CUZ9 |
| | Ligase (EC 6.1.1.17) | K4E4Q3 |
| | Ligase (EC 6.1.1.18) | K4DUA1 |
| | Ligase (EC 6.1.1.19) | K4DZT5 |
| | Ligase (EC 6.1.1.20) | K4E2G9, K4DYE4, Q4DY45 |
| | Ligase (EC 6.1.1.21) | K4E5W2, Q4DA54 |
| | Ligase (EC 6.1.1.22) | K4E3K3, Q4E595 |
| | Desidrogenase (EC 1.2.1.3) | Q4DCP7 |
| | Desidrogenase (EC 1.2.1.41) | K4E0F1 ^s , Q4D112 |
| | Desidrogenase (EC 1.4.1.2) | K4DUG5 |
| | Redutase (EC 1.5.1.2) | K4EAX2, Q4DH60 |
| | Desidrogenase (EC 1.5.99.8) | K4EBN9, Q4CVA1 |
| | Sintetase (EC 2.5.1.16) | K4DV45, Q4DA73 |
| | Transaminase (EC 2.6.1.1) | K4EA87, K4DZN8, Q4E4E7, Q4D3Y8, Q4CVI0, Q4D1Q4 |
| 5) Metabolismo de arginina e prolina | Quinase (EC 2.7.2.11) | K4E0F1 ^s |
| | Quinase (EC 2.7.3.3) | K4DUF3 |
| | Desidropeptidase (EC 3.5.1.14) | K4DQT8, K4DQX8 |
| | Desacetilase (EC 3.5.1.16) | K4DTG9, K4DKW8, K4ED56, Q4D7V2, Q4CPW5, Q4CR09, Q4CYZ5 |
| | Descarboxilase (EC 4.1.1.50) | K4E603, Q4DQQ5, Q4CWQ0 |
| | Racemase (EC 5.1.1.4) | K4DJR1 |
| | Ligase (EC 6.3.1.2) | K4E9B6, Q4CN04 |
| | Sintetase (EC 6.3.4.5) | K4DK07 ^t |
| | Sintetase (EC 6.3.4.16) | K4DK07 ^t |

^a Via metabólica sugerida por KEGG - Blast2GO ^b Enzyme ID do sistema EC

^c Número de acesso do Uniprot

d, e, f, g, h, i, j, k, l, m, n, o, p, q, r, s, t Identificações com dois ou mais Enzyme ID na mesma via.

3. Predição de proteínas secretadas

Considerando a exclusividade das proteínas para cada uma das vias de secreção, de um total de 2.947 identificações, 280 (9,50%) e 1.556 (52,80%) proteínas foram preditas serem secretadas pelo SignalP e pelo SecretomeP, respectivamente (Figura 32). Portanto, em todas as três buscas realizadas, prevaleceram proteínas secretadas por via não clássica. No entanto, é necessário investigar experimentalmente se essas proteínas são realmente secretadas.

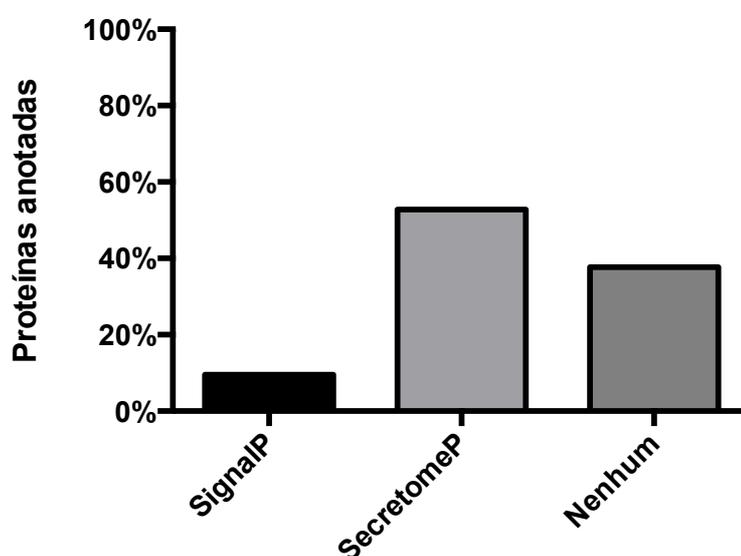


Figura 32. Predição de proteínas secretadas pelas vias clássica (SignalP) e não clássica (SecretomeP) e não anotadas como secretadas (nenhum).

Entre proteínas putativas, baseado na descrição do Uniprot, 82 grupos foram preditos pelo SignalP para secreção por via clássica, dos quais os mais numerosos foram: proteína RHS (20 membros) e DGF-1 (18 membros) (Tabela suplementar 5). Em relação à secreção por via não clássica, 361 grupos de proteínas putativas foram preditos pelo SecretomeP e os mais abundantes foram: proteínas RHS (99 membros) e DGF-1 (51 membros) (Tabela suplementar 6). 39 tipos de proteínas foram comuns a ambas as vias de secreção.

A POP é um exemplo de proteína identificada pelas buscas 2 e 3 com predição para secreção por via não clássica. Previamente, a secreção da POP Tc80 por tripomastigotas e amastigotas foi detectada, enquanto que por epimastigotas não (Santana *et al.*, 1997). Essa predição sugere que epimastigotas também

secretam POP e por mecanismos não convencionais.

Nas buscas 2 e 3, foram identificadas proteínas possivelmente envolvidas na defesa celular e na resposta ao estresse oxidativo como triparedoxina (Q4D1B8) e triparedoxina peroxidases (Q4CM56 e Q4CX87) putativas. Essas proteínas foram incluídas na predição de secreção por via não clássica. Na busca 3, uma anidrase carbônica-símile putativa (Q4CVY4) identificada foi classificada como possível proteína secretada por via clássica.

3.1. Predição de proteínas integrais de membrana

O programa TMHMM fez a predição de domínios transmembrânicos para 508 proteínas (17,24%), com variação de 1 a 20 hélices transmembrânicas para cada identificação. O maior percentual de proteínas com domínios transmembrânicos preditos foi o da busca no banco de dados de *Trypanosoma* spp. e outros tripanossomatídeos (20,68%).

Na intersecção entre proteínas preditas com domínios transmembrânicos e secretadas por via clássica e na intersecção entre proteínas preditas com domínios transmembrânicos e secretadas por via não clássica, estão 108 e 349 proteínas, respectivamente. Assim, somente 21,26% das proteínas com domínios transmembrânicos preditos seriam secretadas por via clássica e, 68,70%, por via não clássica (Figura 33).

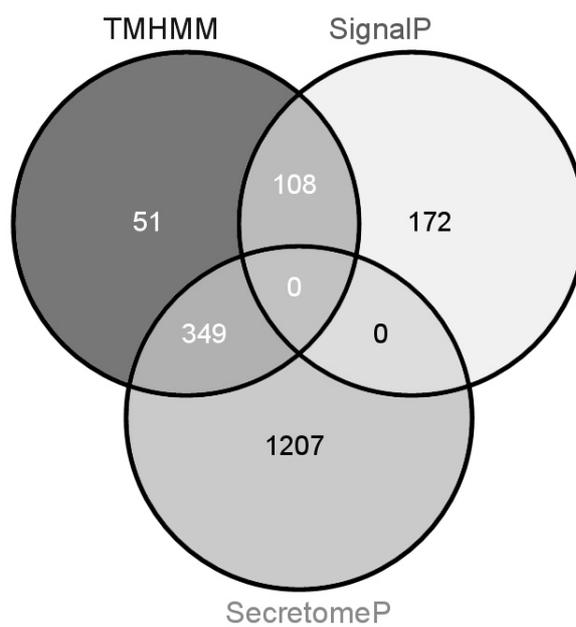


Figura 33. Diagrama de Venn correlacionado as proteínas com preditas hélices transmembrânicas (TMHMM) e as preditas a serem secretadas pelas vias clássica (SignalP) e não-clássica (SecretomeP).

IV. Famílias multigênicas de *T. cruzi* nas buscas realizadas

A seguir, avaliamos a presença de proteínas putativas pertencentes às seis famílias multigênicas mais abundantes em número de membros (El-Sayed *et al.*, 2005) nas três buscas realizadas devido à grande representatividade e à importância biológica dessas famílias em *T. cruzi*. A Tabela 14 resume o número de identificações relativas a essas famílias multicópias obtidas em cada uma das buscas da cepa G.

Tabela 14. Número de proteínas das famílias multigênicas de TS, MASP, mucinas, proteínas RHS, proteínas DGF-1 e gp63 obtidas na cepa G de *T. cruzi* identificadas pelas três buscas realizadas.

| Família multigênica | Busca 1 ^a | Busca 2 ^b | Busca 3 ^c |
|---------------------|----------------------|----------------------|----------------------|
| 1. TS | 35 | ND | 17 |
| 2. MASP | 15 | ND | 1 |
| 3. Mucinas | 9 | ND | ND |
| 4. Proteínas RHS | 4 | 24 | 289 |
| 5. Proteínas DGF-1 | 7 | 9 | 69 |
| 6. gp63 | 2 | ND | ND |
| Total | 72 | 33 | 376 |

^a Busca no banco de dados de *Trypanosoma* spp. e outros tripanossomatídeos

^b Busca no banco de dados de *T. cruzi* CL Brener

^c Busca no banco de dados de *T. cruzi* CL Brener e Sylvio X10/1

ND: não detectada

1. TS

A família gênica mais abundante da cepa CL Brener de *T. cruzi* é a das TS com 1.430 membros, dos quais 693 são pseudogenes (El-Sayed *et al.*, 2005). Em relação à localização cromossomal dessas sequências, as regiões subteloméricas são enriquecidas em genes e pseudogenes de TS, assim como de proteínas RHS e DGF-1, e a frequência das sequências de TS nessas regiões é de 9% (Moraes Barros *et al.*, 2012).

Pela busca no banco de dados de *Trypanosoma* spp. e outros tripanossomatídeos, 35 TS putativas foram constatadas, enquanto que, contraditoriamente, nenhuma TS está presente na lista de identificações da busca mais restritiva, no banco de dados de CL Brener. Essa discrepância em relação à identificação de TS foi um dos aspectos que nos levou a realizar a terceira busca, no

banco de dados de CL Brener e Sylvio X10/1, pela qual identificamos 17 TS putativas que foram exclusivas para essa análise. Assim, duas buscas sugerem a expressão de TS em epimastigotas da cepa G. Na cepa Brazil, 223 TS foram identificadas em tripomastigotas, metacíclicas e amastigotas, mas não em epimastigotas (Atwood *et al.*, 2005).

A secreção de gp82 e de outras TS foi detectada em epimastigotas e tripomastigotas metacíclicas da cepa Dm28c, mas não de gp90 que foi apenas detectada a partir de tripomastigotas metacíclicas (Bayer-Santos *et al.*, 2013). No subproteoma da cepa G, também foram identificadas TS gp82, gp90, TSA-1, SA85 e ASP-2, pertencentes ao grupo II da superfamília das TS, porém somente em tripomastigotas metacíclicas (Cordero *et al.*, 2009).

Existem algumas diferenças descritas para TS de epimastigotas e tripomastigotas. A expressão de TS foi nula ou baixa em cultura de epimastigotas no início da fase log de crescimento e aumentou progressivamente até atingir a fase estacionária (Chaves, Briones e Schenkman, 1993). Tripomastigotas expressam TS enzimaticamente ativas e também TS inativas por causa da mutação Tyr342-His no domínio catalítico N-terminal. TS de tripomastigotas apresentam o domínio repetitivo SAPA (*shed acute phase antigen*) na extremidade C-terminal (Cazzulo e Frasch, 1992; Schenkman *et al.*, 1994; Cremona *et al.*, 1995; Cremona *et al.*, 1999; Frasch, 2000), enquanto TS de epimastigotas, que são ativas, não (Chaves, Briones e Schenkman, 1993; Briones, Egima e Schenkman, 1995). Assim, o domínio SAPA não é necessário para a atividade de trans-sialidase da molécula. Em contrapartida, repetições SAPA são imunodominantes, sendo produzidos, precocemente durante a fase aguda de infecção, anticorpos contra esse domínio. Após a geração desses anticorpos específicos, imunoglobulinas contra a região catalítica são produzidas para neutralizar a atividade enzimática de moléculas de TS circulantes no sangue (Buscaglia *et al.*, 1998). Isso seria um mecanismo do parasito para protelar a inibição da atividade enzimática pelo sistema imune do hospedeiro e, conseqüentemente, prevenir a massiva invasão celular e permitir a sobrevivência do parasito no hospedeiro infectado.

Foi demonstrado que TS de tripomastigotas podem ser ancoradas a GPI (Agusti *et al.*, 1997; Agusti *et al.*, 1998) ou liberadas no meio extracelular e são oligoméricas (Schenkman, Pontes De Carvalho e Nussenzweig, 1992). TS previamente ligadas a GPI e liberadas no meio extracelular por ação de uma

fosfatidilinositol fosfolipase C agiriam como um fator de virulência (Chuenkova e Pereira, 1995). Diferentemente, TS de epimastigotas são monoméricas (Chaves, Briones e Schenkman, 1993) e não apresentam sítio para âncora GPI predito, mas têm um potencial domínio transmembrânico (Briones, Egima e Schenkman, 1995). Sugerindo uma função para a expressão de TS em epimastigotas não infectivos, TS ativas exerceriam influência nas interações inseto-parasito, mais especificamente na diferenciação e na migração através do trato digestivo do triatomíneo.

2. MASP

MASP são proteínas altamente conservadas nos domínios N- e C-terminal e apresentam uma região central variável e repetitiva. Genes de codificam MASP localizam-se dentro de grandes arranjos de genes não dispostos *in tandem*, incluindo genes que codificam mucinas, trans-sialidases-símiles gp85, gp63, proteínas ricas em serina-alanina-prolina (SAP), DGF-1, proteínas RHS. A denominação *mucin-associated surface protein* (MASP) foi dada em razão de uma grande quantidade de genes que codificam MASP estarem localizados próximos a genes que codificam mucinas (Bartholomeu *et al.*, 2009). MASP são codificadas pela segunda família gênica mais numerosa no genoma de CL Brener, com 1.377 membros (433 prováveis pseudogenes) (El-Sayed *et al.*, 2005). Sabe-se pouco ainda sobre a importância biológica das MASP, mas já foi demonstrado que anticorpos IgG específicos para a região catalítica de uma MASP inibiram significativamente a invasão por tripomastigotas metacíclicas (De Pablos *et al.*, 2011).

Na primeira busca, houve a identificação de 15 MASP putativas; na segunda busca, nenhuma; na terceira busca, apenas uma. Portanto, houve uma significativa discordância sobre a abundância de MASP em epimastigotas da cepa G, sendo que duas buscas apontaram para a ausência ou a baixa expressão de MASP. No proteoma da cepa Brazil, nove MASP putativas foram identificadas em epimastigotas e, predominantemente, em tripomastigotas, porém essas detecções fazem parte das identificações derivadas de correspondência de apenas um peptídeo, que perfizeram quase 25% do total (Atwood *et al.*, 2005). Por isso é hipotetizado que MASP apresentam extensivas modificações pós-traducionais, e há evidência de que MASP de tripomastigotas são modificadas por glicanas N-ligadas (Atwood *et al.*, 2006), o que dificultaria a detecção por técnicas proteômicas *shotgun* e forneceria uma

análise subestimada da expressão dessa grande família de proteínas. Outra hipótese seria puramente relacionada à baixa expressão de MASP em epimastigotas.

3. Mucinas

As mucinas são glicoproteínas codificadas por uma família composta por 863 membros (201 prováveis pseudogenes), e, juntamente com genes das famílias de TS e MASP com os quais são fisicamente relacionados, compõem 15% de todos os genes preditos de *T. cruzi* (El-Sayed et al., 2005). São os componentes mais comumente expressos na superfície do *T. cruzi*. São divididas em duas famílias: TcMUC (*T. cruzi mucin*) e TcSMUG (*T. cruzi small mucin-like gene*). A família TcMUC é atualmente subdividida em três grupos: TcMUC I, TcMUC II e *trypomastigote small surface antigen* (TSSA). Os membros dos três grupos codificam para peptídeo-sinal N-terminal e sítio de ligação à âncora GPI C-terminal. Para proteínas TcMUC I e II, é predita uma região central com repetições de resíduos de Thr, Ser e Pro (muitas repetições em TcMUC I e poucas em TcMUC II) que são sítios para O-glicosilação. As mucinas TcMUC são expressas em estágios relacionados ao hospedeiro mamífero, amastigotas e tripomastigotas sanguíneos (revisado por Buscaglia et al. (2006)). Já a família TcSMUG foi diferencialmente expressa em nível de mRNA entre as formas de vida do *T. cruzi*, sendo mais expressa em epimastigotas, devido a mecanismos de regulação pós-transcricional que afetam a estabilidade dos mRNAs (Di Noia et al., 2000).

Como já citado, as mucinas são os principais substratos para sialilação pelas TS. Além disso, as mucinas exercem influência na: proteção do parasito contra degradação proteolítica (no trato intestinal do triatomíneo e no hospedeiro mamífero), proteção contra a lise mediada pela opsonização por anticorpos do hospedeiro mamífero quando estão sialiladas, adesão e invasão das células hospedeiras, estimulação de citocinas pró-inflamatórias e óxido nítrico por macrófagos (revisado por Acosta-Serrano et al. (2001) e Buscaglia et al. (2006)).

Somente na busca 1 foram identificadas mucinas putativas, 6 TcMUC e 3 mucinas-símbles. As sequências das mucinas identificadas derivaram da cepa CL Brener de *T. cruzi* e de *T. cruzi marinkellei*. Interessantemente, os outros dois bancos de dados também continham sequências provenientes de CL Brener, mas nenhuma mucina putativa foi detectada. O fato de nenhuma identificação referente a

mucinas ter sido identificada em nenhuma das quatro formas de vida da cepa Brazil (Atwood *et al.*, 2005) nem nas duas buscas específicas realizadas neste trabalho sugere que houve limitações na obtenção de proteínas altamente glicosiladas pelas metodologias empregadas em ambos os estudos e que seria necessário um enriquecimento das amostras para esse tipo modificação pós-traducional.

4. Proteínas RHS

A grande família multigênica que codifica proteínas RHS foi primeiramente identificada em *T. brucei* e pode ser subdividida em seis subfamílias denominadas RHS1 a RHS6, perfazendo um total de 128 cópias de genes/pseudogenes. Proteínas RHS apresentaram localização nuclear e perinuclear (Bringaud *et al.*, 2002). Em *T. cruzi*, abrange 752 genes na cepa CL Brener (557 pseudogenes) (El-Sayed *et al.*, 2005), os quais são mais frequentemente associados às regiões subteloméricas (19%) do que os que codificam TS (Moraes Barros *et al.*, 2012).

Em relação à identificação de proteínas RHS putativas nos proteomas totais de *T. cruzi* previamente publicados, apenas uma proteína RHS foi detectada em epimastigotas da cepa Dm28c. O método para solubilização de proteínas por ureia/tioureia empregado e a análise por MS de baixo rendimento favoreceram apenas a identificação de proteínas solúveis e altamente concentradas (Parodi-Talice *et al.*, 2004). Em contrapartida, 399 proteínas RHS foram identificadas na cepa Brazil quando consideradas todas as formas de vida do parasito, até porque foi realizada uma análise proteômica de alto rendimento (Atwood *et al.*, 2005). Em epimastigotas da cepa G, a quantidade de proteínas RHS putativas identificadas variou bastante entre as buscas: 4 na busca 1, 24 na busca 2 e 289 na busca 3. Nessa última, o número de identificações de proteínas RHS foi bastante dissonante às outras buscas. Além disso, proteínas RHS foram preditas serem secretadas por ambas as vias de secreção em todas as três buscas em bancos de dados realizadas (Tabelas suplementares 1 a 6), o que diverge da localização subcelular dessas proteínas descrita em *T. brucei*.

5. Proteínas DGF-1

A quinta família multigênica DGF-1 mais representativa engloba 565 genes (136 pseudogenes), dos quais 12% foram encontrados nas regiões subteloméricas (Moraes Barros *et al.*, 2012). Um trabalho mostrou que membros desta família têm

motivos conservados de adesina com similaridade à beta-7 integrina humana e são expressos na superfície de tripomastigotas das cepas G e Y (Kawashita *et al.*, 2009). Anticorpos anti-DGF-1.2 resultaram em uma marcação citoplasmática em tripomastigotas, epimastigotas e amastigotas, porém o sinal de fluorescência foi mais forte nesses últimos. Durante a diferenciação de tripomastigotas em amastigotas dentro da célula mamífera, o pico da intensidade de fluorescência apresentada na forma de pequenos corpos na periferia do parasito ocorre a 18 h pós-infecção e diminui até desaparecer completamente a 40 h pós-infecção. Durante a diferenciação *in vitro*, também se observou uma marcação próxima à membrana do parasito e DGF-1.2 foi secretada no meio de cultura (Lander *et al.*, 2010).

Apesar da abundância dos genes das famílias RHS e DGF-1 no genoma de *T. cruzi*, especialmente próximos aos telômeros, proteínas codificadas por esses genes ainda não tiveram as funções biológicas elucidadas. Assim como para a família de proteínas RHS, os resultados da última busca para proteínas DGF-1 foram discrepantes em relação às outras: 7 na busca 1, 9 na busca 2 e 69 na busca 3.

6. gp63

A glicoproteína de superfície gp63 ou leishmanolisina foi primeiramente descrita em *Leishmania* spp., mas seus homólogos também estão presentes em outros tripanossomatídeos patogênicos, *T. cruzi* e *T. brucei*. gp63 são codificadas por genes altamente conservados arranjados *in tandem*. Metaloproteases dependentes de zinco gp63 parecem ter influência na interação com receptores de macrófagos (Russell, Talamas-Rohana e Zelechowski, 1989; Brittingham *et al.*, 1995), na sobrevivência intralisossomal (Chaudhuri *et al.*, 1989), na resistência à lise mediada pelo sistema complemento (Chaudhuri e Chang, 1988; Brittingham *et al.*, 1995) e a peptídeos antimicrobianos (Kulkarni *et al.*, 2006). Além disso, gp63 degradaram componentes da matriz extracelular, o que favoreceu uma rápida migração do parasito (McGwire, Chang e Engman, 2003).

Em *T. cruzi*, estão presentes 425 cópias (251 pseudogenes) (El-Sayed *et al.*, 2005) e, assim, é observada uma expansão massiva dessa família de genes neste tripanossomatídeo. Há indicativos de que gp63-símbios de *T. cruzi* são importantes na infecção da célula hospedeira, pois anticorpos contra Tcgp63 inibiram a infecção de células mamíferas por tripomastigotas (Cuevas, Cazzulo e Sanchez, 2003;

Kulkarni *et al.*, 2009). 29 gp63-símiles foram expressas e detectadas em todos os estágios da cepa Brazil, exceto amastigotas (Atwood *et al.*, 2005) e, no presente proteoma da cepa G, apenas na busca no banco de dados de *Trypanosoma* spp. e outros tripanossomatídeos (n = 2).

CONCLUSÃO

O proteoma total de formas epimastigotas da cepa G de *T. cruzi* é o mais abrangente até hoje disponível, com a identificação de 2.947 proteínas pela busca no banco de dados das cepas CL Brener e Sylvio X10/1 de *T. cruzi* (busca 3) (Tabela 15). A obtenção de uma quantidade maior de identificações por essa busca em relação às outras buscas poderia estar relacionada ao emprego do algoritmo do programa PEAKS em vez do Sequest e/ou à redundância das sequências entre as diversas espécies dos gêneros *Trypanosoma* e *Leishmania* na busca 1, o que atrapalhou o algoritmo. Houve uma enorme diferença entre a quantidade de identificações das buscas 2 (banco de dados de CL Brener) e 3, o que poderia estar relacionado à utilização de algoritmos diferentes e/ou à inclusão das sequências da cepa Sylvio X10/1 no banco de dados da busca 3. Na tentativa de esclarecer isso, seria interessante refazer a busca 2 com a utilização do PEAKS e também a busca 3 com o Sequest.

Tabela 15. Comparação entre o número de proteínas identificadas neste e em outros proteomas totais previamente publicados de formas epimastigotas de *T. cruzi*.

| Cepa de <i>T. cruzi</i> | Proteínas identificadas ^a | Referência bibliográfica |
|-------------------------|--------------------------------------|-------------------------------------|
| Berenice | 18 | Paba, Santana, <i>et al.</i> (2004) |
| Dm28c | 25 | Parodi-Talice <i>et al.</i> (2004) |
| Brazil | 1.861 | Atwood <i>et al.</i> (2005) |
| G ^b | 1.460 | Este trabalho |
| G ^c | 953 | Este trabalho |
| G ^d | 2.947 | Este trabalho |

^a Número de proteínas identificadas em cada análise proteômica

^b Busca no banco de dados de *Trypanosoma* spp. e outros tripanossomatídeos

^c Busca no banco de dados de *T. cruzi* CL Brener

^d Busca no banco de dados de *T. cruzi* CL Brener e Sylvio X10/1

A busca 3 resultou na identificação de uma maior quantidade de membros das principais famílias multigênicas de *T. cruzi*, como TS, MASP, proteínas RHS e DGF-1 e a busca 1 identificou representantes de todas as seis famílias mais representativas no genoma da cepa CL Brener. No caso específico das TS, as

identificamos nas buscas 1 e 3, enquanto que, em epimastigotas da cepa Brazil, nenhuma TS foi detectada. Uma quantidade contundente de proteínas RHS e DGF-1 foi relacionada na busca 3, mas também estiveram presentes na lista de identificações das outras buscas. MASP, mucinas e gp63 foram pouco representadas em todas as buscas deste proteoma possivelmente em virtude das glicosilações preditas para essas moléculas, fator limitante para a detecção dessas em um extrato de proteínas solubilizadas por SDS. Como objetivamos realizar uma análise proteômica global, não foi realizado nenhum tipo de enriquecimento para modificações pós-traducionais específicas.

A proeminência dos resultados da busca 3 deve-se, também, à identificação de diversas proteínas bem estudadas e consideradas como possíveis alvos de drogas, como LAP, POP, OPB, catepsina B, cruzipaína e MTAF. Adicionalmente, a utilização do banco de dados de *Trypanosoma* spp. e outros tripanossomatídeos levou à identificação de diversas VSG, proteínas presentes em outros tripanossomatídeos como *T. brucei*, *T. vivax* e *T. evansi*, porém inexistentes em *T. cruzi*.

A identificação de proteínas candidatas a diversas vias metabólicas é outro aspecto positivo do nosso proteoma. De modo unânime, a via de metabolismo de purinas apresentou o maior número de proteínas anotadas, o que correlaciona com o estágio de vida do *T. cruzi* analisado, formas replicativas epimastigotas. Nas buscas 2 e 3, a presença das vias ciclo do ácido cítrico, metabolismo de prolina e arginina, fosforilação oxidativa e glicólise entre as dez vias metabólicas mais abundantes em quantidade de anotações também está em concordância com o alto requerimento energético de epimastigotas. Dessas quatro vias, a glicólise figura nessas listas de todas as três buscas, resultado que provavelmente está relacionado à presença de glicose no meio de cultura, a fonte preferencial de carbono para a produção de ATP em epimastigotas.

Em todas as análises, as proteínas preditas a serem secretadas pela via não clássica, na qual as proteínas são secretadas independentemente do RE e/ou Golgi, predominaram sobre as secretadas pela via clássica preditas. Deve-se considerar que o algoritmo do SignalP não está apto para identificar todos os peptídeos-sinais de *T. cruzi*, mas isso está de acordo com outros dados da literatura. Assim, a predição de muitas proteínas identificadas como secretadas demonstra que nosso proteoma compreendeu uma grande diversidade de proteínas.

Outro ponto positivo deste proteoma foi a identificação de potenciais proteínas integrais de membrana, o que demonstra que a extração de proteínas realizada foi bastante eficiente e resultou na identificação de uma ampla gama de proteínas.

Portanto, consideramos a busca no banco de dados das cepas CL Brener e Sylvio X10/1 de *T. cruzi* a mais satisfatória.

Futuramente pretendemos realizar a montagem do genoma da cepa G cujo o DNA genômico já foi sequenciado. Assim, esta análise proteômica auxiliará na anotação dos genes devido à identificação das fases abertas de leitura.

Este trabalho representa um avanço considerável no conhecimento do conjunto de proteínas expressas em *T. cruzi* cepa G. No entanto, ainda há muito o que ser obtido nesta aventura de desvendar os verbos, proteínas, das diferentes formas deste fantástico protozoário. E ainda restarão as perguntas acerca das funções de cada uma destas milhares de proteínas tripanossômicas. Uma tarefa para um mutirão!

REFERÊNCIAS BIBLIOGRÁFICAS

ACOSTA, H. et al. Pyruvate phosphate dikinase and pyrophosphate metabolism in the glycosome of *Trypanosoma cruzi* epimastigotes. **Comp Biochem Physiol B Biochem Mol Biol**, v. 138, n. 4, p. 347-56, Aug 2004. ISSN 1096-4959 (Print)1096-4959. Disponível em: < <http://dx.doi.org/10.1016/j.cbpc.2004.04.017> >.

ACOSTA-SERRANO, A. et al. The mucin-like glycoprotein super-family of *Trypanosoma cruzi*: structure and biological roles. **Mol Biochem Parasitol**, v. 114, n. 2, p. 143-50, May 2001. ISSN 0166-6851 (Print)0166-6851. Disponível em: < <http://dx.doi.org/> >.

AFONSO, A. M.; EBELL, M. H.; TARLETON, R. L. A systematic review of high quality diagnostic tests for Chagas disease. **PLoS Negl Trop Dis**, v. 6, n. 11, p. e1881, 2012. ISSN 1935-2727. Disponível em: < <http://dx.doi.org/10.1371/journal.pntd.0001881> >.

AGUSTI, R. et al. Structure of the glycosylphosphatidylinositol-anchor of the trans-sialidase from *Trypanosoma cruzi* metacyclic trypomastigote forms. **Mol Biochem Parasitol**, v. 97, n. 1-2, p. 123-31, Nov 30 1998. ISSN 0166-6851 (Print)0166-6851. Disponível em: < <http://dx.doi.org/> >.

_____. The trans-sialidase of *Trypanosoma cruzi* is anchored by two different lipids. **Glycobiology**, v. 7, n. 6, p. 731-5, Sep 1997. ISSN 0959-6658 (Print)0959-6658. Disponível em: < <http://dx.doi.org/> >.

ALARCON DE NOYA, B. et al. Large urban outbreak of orally acquired acute Chagas disease at a school in Caracas, Venezuela. **J Infect Dis**, v. 201, n. 9, p. 1308-15, May 1 2010. ISSN 1537-6613 (Electronic)0022-1899 (Linking). Disponível em: < <http://dx.doi.org/10.1086/651608> >.

ALONSO, G. D. et al. TcPDE4, a novel membrane-associated cAMP-specific phosphodiesterase from *Trypanosoma cruzi*. **Mol Biochem Parasitol**, v. 145, n. 1, p. 40-9, Jan 2006. ISSN 0166-6851 (Print)0166-6851. Disponível em: < <http://dx.doi.org/10.1016/j.molbiopara.2005.09.005> >.

_____. TcrPDEA1, a cAMP-specific phosphodiesterase with atypical pharmacological properties from *Trypanosoma cruzi*. **Mol Biochem Parasitol**, v. 152, n. 1, p. 72-9, Mar 2007. ISSN 0166-6851 (Print)0166-6851. Disponível em: < <http://dx.doi.org/10.1016/j.molbiopara.2006.12.002> >.

ALTELAAR, A. F.; MUNOZ, J.; HECK, A. J. Next-generation proteomics: towards an integrative view of proteome dynamics. **Nat Rev Genet**, v. 14, n. 1, p. 35-48, Jan 2013. ISSN 1471-0056. Disponível em: < <http://dx.doi.org/10.1038/nrg3356> >.

ALVAREZ, V. E.; NIEMIROWICZ, G. T.; CAZZULO, J. J. The peptidases of *Trypanosoma cruzi*: digestive enzymes, virulence factors, and mediators of autophagy and programmed cell death. **Biochim Biophys Acta**, v. 1824, n. 1, p. 195-206, Jan 2012. ISSN 0006-3002 (Print)0006-3002. Disponível em: < <http://dx.doi.org/10.1016/j.bbapap.2011.05.011> >.

ANDRADE, S. G. et al. Biological, biochemical and molecular features of *Trypanosoma cruzi* strains isolated from patients infected through oral transmission during a 2005 outbreak in the state of Santa Catarina, Brazil: its correspondence with the new *T. cruzi* Taxonomy Consensus (2009). **Mem Inst Oswaldo Cruz**, v. 106, n. 8, p. 948-56, Dec 2011. ISSN 0074-0276. Disponível em: < <http://dx.doi.org/> >.

ARGOLO, A. M. et al. **Doença de Chagas e seus principais vetores no Brasil**. Rio de Janeiro: Imperial Novo Milênio Gráfica e Editora Ltda., 2008.

ATWOOD, J. A., 3RD et al. Glycoproteomics of *Trypanosoma cruzi* trypomastigotes using subcellular fractionation, lectin affinity, and stable isotope labeling. **J Proteome Res**, v. 5, n. 12, p. 3376-84, Dec 2006. ISSN 1535-3893 (Print)1535-3893. Disponível em: < <http://dx.doi.org/10.1021/pr060364b> >.

_____. The *Trypanosoma cruzi* proteome. **Science**, United States, v. 309, n. 5733, p. 473-6, 2005. ISSN 1095-9203 (Electronic)0036-8075 (Linking). Disponível em: < <http://dx.doi.org/10.1126/science.1110289> >.

AUFDERHEIDE, A. C. et al. A 9,000-year record of Chagas' disease. **Proc Natl Acad Sci U S A**, United States, v. 101, n. 7, p. 2034-9, 2004. ISSN 0027-8424 (Print)0027-8424 (Linking). Disponível em: < <http://dx.doi.org/10.1073/pnas.0307312101> >.

BARBOZA, N. R. et al. Expression profile and subcellular localization of HslV, the proteasome related protease from *Trypanosoma cruzi*. **Exp Parasitol**, v. 130, n. 2, p. 171-7, Feb 2012. ISSN 0014-4894. Disponível em: < <http://dx.doi.org/10.1016/j.exppara.2011.10.011> >.

BARTHOLOMEU, D. C. et al. Genomic organization and expression profile of the mucin-associated surface protein (masp) family of the human pathogen *Trypanosoma cruzi*. **Nucleic Acids Res**, v. 37, n. 10, p. 3407-17, Jun 2009. ISSN 0305-1048. Disponível em: < <http://dx.doi.org/10.1093/nar/gkp172> >.

BASTOS, C. J. et al. Clinical outcomes of thirteen patients with acute chagas disease acquired through oral transmission from two urban outbreaks in northeastern Brazil. **PLoS Negl Trop Dis**, v. 4, n. 6, p. e711, 2010. ISSN 1935-2727. Disponível em: < <http://dx.doi.org/10.1371/journal.pntd.0000711> >.

BASTOS, I. M. et al. Molecular, functional and structural properties of the prolyl oligopeptidase of *Trypanosoma cruzi* (POP Tc80), which is required for parasite entry into mammalian cells. **Biochem J**, v. 388, n. Pt 1, p. 29-38, May 15 2005. ISSN 0264-6021. Disponível em: < <http://dx.doi.org/10.1042/bj20041049> >.

_____. Parasite prolyl oligopeptidases and the challenge of designing chemotherapeutics for Chagas disease, leishmaniasis and African trypanosomiasis. **Curr Med Chem**, v. 20, n. 25, p. 3103-15, 2013. ISSN 0929-8673. Disponível em: < <http://dx.doi.org/> >.

BAYER-SANTOS, E. et al. Proteomic analysis of Trypanosoma cruzi secretome: characterization of two populations of extracellular vesicles and soluble proteins. **J Proteome Res**, v. 12, n. 2, p. 883-97, Feb 1 2013. ISSN 1535-3893. Disponível em: < <http://dx.doi.org/10.1021/pr300947g> >.

BENDTSEN, J. D. et al. Feature-based prediction of non-classical and leaderless protein secretion. **Protein Eng Des Sel**, v. 17, n. 4, p. 349-56, Apr 2004. ISSN 1741-0126 (Print)1741-0126. Disponível em: < <http://dx.doi.org/10.1093/protein/gzh037> >.

BERRIMAN, M. et al. The genome of the African trypanosome Trypanosoma brucei. **Science**, United States, v. 309, n. 5733, p. 416-22, 2005. ISSN 1095-9203 (Electronic)0036-8075 (Linking). Disponível em: < <http://dx.doi.org/10.1126/science.1112642> >.

BOITZ, J. M. et al. Adenylosuccinate synthetase and adenylosuccinate lyase deficiencies trigger growth and infectivity deficits in Leishmania donovani. **J Biol Chem**, v. 288, n. 13, p. 8977-90, Mar 29 2013. ISSN 0021-9258. Disponível em: < <http://dx.doi.org/10.1074/jbc.M112.431486> >.

BORST, P.; ULBERT, S. Control of VSG gene expression sites. **Mol Biochem Parasitol**, v. 114, n. 1, p. 17-27, Apr 25 2001. ISSN 0166-6851 (Print)0166-6851. Disponível em: < <http://dx.doi.org/> >.

BOUVIER, L. A. et al. An expanded adenylate kinase gene family in the protozoan parasite Trypanosoma cruzi. **Biochim Biophys Acta**, v. 1760, n. 6, p. 913-21, Jun 2006. ISSN 0006-3002 (Print)0006-3002. Disponível em: < <http://dx.doi.org/10.1016/j.bbagen.2006.02.013> >.

BRAZ, L. M.; AMATO NETO, V.; OKAY, T. S. Reactivation of Trypanosoma cruzi infection in immunosuppressed patients. **Rev Inst Med Trop Sao Paulo**, Brazil, v. 50, n. 1, p. 65-6, 2008. ISSN 0036-4665 (Print)0036-4665 (Linking). Disponível em: < <http://dx.doi.org/> >.

BRINGAUD, F.; BARRETT, M. P.; ZILBERSTEIN, D. Multiple roles of proline transport and metabolism in trypanosomatids. **Front Biosci (Landmark Ed)**, v. 17, p. 349-74, 2012. ISSN 1093-4715. Disponível em: < <http://dx.doi.org/> >.

BRINGAUD, F. et al. A new, expressed multigene family containing a hot spot for insertion of retroelements is associated with polymorphic subtelomeric regions of Trypanosoma brucei. **Eukaryot Cell**, v. 1, n. 1, p. 137-51, Feb 2002. ISSN 1535-9778 (Print)1535-9786. Disponível em: < <http://dx.doi.org/> >.

BRINGAUD, F.; RIVIERE, L.; COUSTOU, V. Energy metabolism of trypanosomatids: adaptation to available carbon sources. **Mol Biochem Parasitol**, v. 149, n. 1, p. 1-9, Sep 2006. ISSN 0166-6851 (Print)0166-6851. Disponível em: < <http://dx.doi.org/10.1016/j.molbiopara.2006.03.017> >.

BRIONES, M. R.; EGIMA, C. M.; SCHENKMAN, S. Trypanosoma cruzi trans-sialidase gene lacking C-terminal repeats and expressed in epimastigote forms. **Mol Biochem**

Parasitol, v. 70, n. 1-2, p. 9-17, Mar 1995. ISSN 0166-6851 (Print)0166-6851. Disponível em: < <http://dx.doi.org/> >.

BRITTINGHAM, A. et al. Role of the Leishmania surface protease gp63 in complement fixation, cell adhesion, and resistance to complement-mediated lysis. **J Immunol**, v. 155, n. 6, p. 3102-11, Sep 15 1995. ISSN 0022-1767 (Print)0022-1767. Disponível em: < <http://dx.doi.org/> >.

BRITTO, C. C. Usefulness of PCR-based assays to assess drug efficacy in Chagas disease. **Mem Inst Oswaldo Cruz**, v. 104 Suppl 1, p. 122-35, Jul 2009. ISSN 0074-0276. Disponível em: < <http://dx.doi.org/> >.

BURLEIGH, B. A. et al. A cytosolic serine endopeptidase from Trypanosoma cruzi is required for the generation of Ca²⁺ signaling in mammalian cells. **J Cell Biol**, v. 136, n. 3, p. 609-20, Feb 10 1997. ISSN 0021-9525 (Print)0021-9525. Disponível em: < <http://dx.doi.org/> >.

BUSCAGLIA, C. A. et al. The repetitive domain of Trypanosoma cruzi trans-sialidase enhances the immune response against the catalytic domain. **J Infect Dis**, v. 177, n. 2, p. 431-6, Feb 1998. ISSN 0022-1899 (Print)0022-1899. Disponível em: < <http://dx.doi.org/> >.

_____. Trypanosoma cruzi surface mucins: host-dependent coat diversity. **Nat Rev Microbiol**, v. 4, n. 3, p. 229-36, Mar 2006. ISSN 1740-1526 (Print)1740-1526. Disponível em: < <http://dx.doi.org/10.1038/nrmicro1351> >.

BUTTERWORTH, A. C. et al. Saturation mutagenesis, complement selection, and steady-state kinetic studies illuminate the roles of invariant residues in active site loop I of the hypoxanthine phosphoribosyltransferase from Trypanosoma cruzi. **Biochim Biophys Acta**, v. 1699, n. 1-2, p. 87-94, Jun 1 2004. ISSN 0006-3002 (Print)0006-3002. Disponível em: < <http://dx.doi.org/10.1016/j.bbapap.2004.01.006> >.

CADAVID-RESTREPO, G. et al. The major leucyl aminopeptidase of Trypanosoma cruzi (LAPTc) assembles into a homohexamer and belongs to the M17 family of metallopeptidases. **BMC Biochem**, v. 12, p. 46, 2011. ISSN 1471-2091. Disponível em: < <http://dx.doi.org/10.1186/1471-2091-12-46> >.

CADAVID-RESTREPO, G. E. **Cinética de atividade e expressão da catepsina B e caracterização bioquímica de uma leucil aminopeptidase de Trypanosoma cruzi**. 2005. (Tese de doutorado). Faculdade de Medicina, Universidade de Brasília (UnB)

CAMARA MDE, L. et al. Molecular and functional characterization of a Trypanosoma cruzi nuclear adenylate kinase isoform. **PLoS Negl Trop Dis**, v. 7, n. 2, p. e2044, 2013. ISSN 1935-2727. Disponível em: < <http://dx.doi.org/10.1371/journal.pntd.0002044> >.

CANNATA, J. J.; CAZZULO, J. J. The aerobic fermentation of glucose by Trypanosoma cruzi. **Comp Biochem Physiol B**, v. 79, n. 3, p. 297-308, 1984. ISSN 0305-0491 (Print)0305-0491. Disponível em: < <http://dx.doi.org/> >.

CARDINALE, G. J.; ABELES, R. H. Purification and mechanism of action of proline racemase. **Biochemistry**, v. 7, n. 11, p. 3970-8, Nov 1968. ISSN 0006-2960 (Print)0006-2960. Disponível em: < <http://dx.doi.org/> >.

CARREIRA, J. C. et al. Trypanosoma cruzi in the scent glands of Didelphis marsupialis: the kinetics of. **Exp Parasitol**, v. 97, n. 3, p. 129-40, Mar 2001. ISSN 0014-4894 (Print)0014-4894. Disponível em: < <http://dx.doi.org/10.1006/expr.2001.4603> >.

CAZZULO, J. J. Aerobic fermentation of glucose by trypanosomatids. **Faseb j**, v. 6, n. 13, p. 3153-61, Oct 1992. ISSN 0892-6638 (Print)0892-6638. Disponível em: < <http://dx.doi.org/> >.

CAZZULO, J. J.; FRASCH, A. C. SAPA/trans-sialidase and cruzipain: two antigens from Trypanosoma cruzi contain immunodominant but enzymatically inactive domains. **Faseb j**, v. 6, n. 14, p. 3259-64, Nov 1992. ISSN 0892-6638 (Print)0892-6638. Disponível em: < <http://dx.doi.org/> >.

CHAGAS, C. Nova tripanozomíaze humana: estudos sobre a morfologia e o ciclo evolutivo do Schizotrypanum cruzi n. gen., n. sp., agente etiológico de nova entidade morbida do homem. **Mem. Inst. Oswaldo Cruz**, v. 1, n. 2, p. 159-218, 08/1909 1909. ISSN 0074-0276. Disponível em: < http://www.scielo.br/scielo.php?script=sci_abstract&pid=S0074-02761909000200008&lng=en&nrm=iso&tlng=pt >. Disponível em: < http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0074-02761909000200008&lng=en&nrm=iso&tlng=pt >. Disponível em: < http://www.scielo.br/scielo.php?script=sci_pdf&pid=S0074-02761909000200008&lng=en&nrm=iso&tlng=pt >.

CHALKLEY, R. J. et al. Comprehensive analysis of a multidimensional liquid chromatography mass spectrometry dataset acquired on a quadrupole selecting, quadrupole collision cell, time-of-flight mass spectrometer: II. New developments in Protein Prospector allow for reliable and comprehensive automatic analysis of large datasets. **Mol Cell Proteomics**, v. 4, n. 8, p. 1194-204, Aug 2005. ISSN 1535-9476 (Print)1535-9476. Disponível em: < <http://dx.doi.org/10.1074/mcp.D500002-MCP200> >.

CHAMOND, N. et al. Trypanosoma cruzi proline racemases are involved in parasite differentiation and infectivity. **Mol Microbiol**, v. 58, n. 1, p. 46-60, Oct 2005. ISSN 0950-382X (Print)0950-382X. Disponível em: < <http://dx.doi.org/10.1111/j.1365-2958.2005.04808.x> >.

CHAUDHURI, G.; CHANG, K. P. Acid protease activity of a major surface membrane glycoprotein (gp63) from Leishmania mexicana promastigotes. **Mol Biochem Parasitol**, v. 27, n. 1, p. 43-52, Jan 1 1988. ISSN 0166-6851 (Print)0166-6851. Disponível em: < <http://dx.doi.org/> >.

CHAUDHURI, G. et al. Surface acid proteinase (gp63) of Leishmania mexicana. A metalloenzyme capable of protecting liposome-encapsulated proteins from

phagolysosomal degradation by macrophages. **J Biol Chem**, v. 264, n. 13, p. 7483-9, May 5 1989. ISSN 0021-9258 (Print)0021-9258. Disponível em: < <http://dx.doi.org/> >.

CHAVES, L. B.; BRIONES, M. R.; SCHENKMAN, S. Trans-sialidase from *Trypanosoma cruzi* epimastigotes is expressed at the stationary phase and is different from the enzyme expressed in trypomastigotes. **Mol Biochem Parasitol**, v. 61, n. 1, p. 97-106, Sep 1993. ISSN 0166-6851 (Print)0166-6851. Disponível em: < <http://dx.doi.org/> >.

CHO, C.; VALE, R. D. The mechanism of dynein motility: insight from crystal structures of the motor domain. **Biochim Biophys Acta**, v. 1823, n. 1, p. 182-91, Jan 2012. ISSN 0006-3002 (Print)0006-3002. Disponível em: < <http://dx.doi.org/10.1016/j.bbamcr.2011.10.009> >.

CHUENKOVA, M.; PEREIRA, M. E. *Trypanosoma cruzi* trans-sialidase: enhancement of virulence in a murine model of Chagas' disease. **J Exp Med**, v. 181, n. 5, p. 1693-703, May 1 1995. ISSN 0022-1007 (Print)0022-1007. Disponível em: < <http://dx.doi.org/> >.

CONESA, A. et al. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. **Bioinformatics**, v. 21, n. 18, p. 3674-6, Sep 15 2005. ISSN 1367-4803 (Print)1367-4803. Disponível em: < <http://dx.doi.org/10.1093/bioinformatics/bti610> >.

CORDERO, E. M. et al. Proteomic analysis of detergent-solubilized membrane proteins from insect-developmental forms of *Trypanosoma cruzi*. **J Proteome Res**, v. 8, n. 7, p. 3642-52, Jul 2009. ISSN 1535-3893 (Print)1535-3893. Disponível em: < <http://dx.doi.org/10.1021/pr800887u> >.

CORTEZ, C. et al. Differential infectivity by the oral route of *Trypanosoma cruzi* lineages derived from Y strain. **PLoS Negl Trop Dis**, v. 6, n. 10, p. e1804, 2012. ISSN 1935-2727. Disponível em: < <http://dx.doi.org/10.1371/journal.pntd.0001804> >.

CORTEZ, M. et al. Infection by *Trypanosoma cruzi* metacyclic forms deficient in gp82 but expressing a related surface molecule, gp30. **Infect Immun**, v. 71, n. 11, p. 6184-91, Nov 2003. ISSN 0019-9567 (Print)0019-9567. Disponível em: < <http://dx.doi.org/> >.

_____. *Trypanosoma cruzi* surface molecule gp90 downregulates invasion of gastric mucosal epithelium in orally infected mice. **Microbes Infect**, v. 8, n. 1, p. 36-44, Jan 2006. ISSN 1286-4579 (Print)1286-4579. Disponível em: < <http://dx.doi.org/10.1016/j.micinf.2005.05.016> >.

COURA, J. R. Transmission of chagasic infection by oral route in the natural history of Chagas disease. **Rev Soc Bras Med Trop**, v. 39 Suppl 3, p. 113-7, 2006. ISSN 0037-8682 (Print)0037-8682. Disponível em: < <http://dx.doi.org/> >.

_____. Present situation and new strategies for Chagas disease chemotherapy: a proposal. **Mem Inst Oswaldo Cruz**, v. 104, n. 4, p. 549-54, Jul 2009. ISSN 0074-0276. Disponível em: < <http://dx.doi.org/> >.

COURA, J. R.; BORGES-PEREIRA, J. Chronic phase of Chagas disease: why should it be treated? A comprehensive. **Mem Inst Oswaldo Cruz**, v. 106, n. 6, p. 641-5, Sep 2011. ISSN 0074-0276. Disponível em: < <http://dx.doi.org/> >.

COURA, J. R.; DE CASTRO, S. L. A critical review on Chagas disease chemotherapy. **Mem Inst Oswaldo Cruz**, v. 97, n. 1, p. 3-24, Jan 2002. ISSN 0074-0276 (Print)0074-0276. Disponível em: < <http://dx.doi.org/> >.

COURA, J. R. et al. Emerging Chagas disease in Amazonian Brazil. **Trends Parasitol**, v. 18, n. 4, p. 171-6, Apr 2002. ISSN 1471-4922 (Print)1471-4922. Disponível em: < <http://dx.doi.org/> >.

COVARRUBIAS, C. et al. Interaction with host factors exacerbates Trypanosoma cruzi cell invasion capacity upon oral infection. **Int J Parasitol**, v. 37, n. 14, p. 1609-16, Dec 2007. ISSN 0020-7519 (Print)0020-7519. Disponível em: < <http://dx.doi.org/10.1016/j.ijpara.2007.05.013> >.

COX, J.; MANN, M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. **Nat Biotechnol**, v. 26, n. 12, p. 1367-72, Dec 2008. ISSN 1087-0156. Disponível em: < <http://dx.doi.org/10.1038/nbt.1511> >.

COX, J. et al. Andromeda: a peptide search engine integrated into the MaxQuant environment. **J Proteome Res**, v. 10, n. 4, p. 1794-805, Apr 1 2011. ISSN 1535-3893. Disponível em: < <http://dx.doi.org/10.1021/pr101065j> >.

CRAIG, R.; BEAVIS, R. C. TANDEM: matching proteins with tandem mass spectra. **Bioinformatics**, v. 20, n. 9, p. 1466-7, Jun 12 2004. ISSN 1367-4803 (Print)1367-4803. Disponível em: < <http://dx.doi.org/10.1093/bioinformatics/bth092> >.

CREMONA, M. L. et al. Enzymically inactive members of the trans-sialidase family from Trypanosoma cruzi display beta-galactose binding activity. **Glycobiology**, v. 9, n. 6, p. 581-7, Jun 1999. ISSN 0959-6658 (Print)0959-6658. Disponível em: < <http://dx.doi.org/> >.

_____. A single tyrosine differentiates active and inactive Trypanosoma cruzi trans-sialidases. **Gene**, v. 160, n. 1, p. 123-8, Jul 4 1995. ISSN 0378-1119 (Print)0378-1119. Disponível em: < <http://dx.doi.org/> >.

CUERVO, P. et al. Proteomic characterization of the released/secreted proteins of Leishmania (Viannia) braziliensis promastigotes. **J Proteomics**, v. 73, n. 1, p. 79-92, Nov 2 2009. Disponível em: < <http://dx.doi.org/10.1016/j.jprot.2009.08.006> >.

CUEVAS, I. C.; CAZZULO, J. J.; SANCHEZ, D. O. gp63 homologues in Trypanosoma cruzi: surface antigens with metalloprotease activity and a possible role in host cell infection. **Infect Immun**, v. 71, n. 10, p. 5739-49, Oct 2003. ISSN 0019-9567 (Print)0019-9567. Disponível em: < <http://dx.doi.org/> >.

D'ANGELO, M. A. et al. A novel calcium-stimulated adenylyl cyclase from *Trypanosoma cruzi*, which interacts with the structural flagellar protein paraflagellar rod. **J Biol Chem**, v. 277, n. 38, p. 35025-34, Sep 20 2002. ISSN 0021-9258 (Print)0021-9258. Disponível em: < <http://dx.doi.org/10.1074/jbc.M204696200> >.

_____. Identification, characterization and subcellular localization of TcPDE1, a novel cAMP-specific phosphodiesterase from *Trypanosoma cruzi*. **Biochem J**, v. 378, n. Pt 1, p. 63-72, Feb 15 2004. ISSN 0264-6021. Disponível em: < <http://dx.doi.org/10.1042/bj20031147> >.

DC-RUBIN, S. S.; SCHENKMAN, S. *Trypanosoma cruzi* trans-sialidase as a multifunctional enzyme in Chagas' disease. **Cell Microbiol**, v. 14, n. 10, p. 1522-30, Oct 2012. ISSN 1462-5814. Disponível em: < <http://dx.doi.org/10.1111/j.1462-5822.2012.01831.x> >.

DE DIEGO, J. L. et al. The ubiquitin-proteasome pathway plays an essential role in proteolysis during *Trypanosoma cruzi* remodeling. **Biochemistry**, v. 40, n. 4, p. 1053-62, Jan 30 2001. ISSN 0006-2960 (Print)0006-2960. Disponível em: < <http://dx.doi.org/> >.

DE GAUDENZI, J. G. et al. Gene expression regulation in trypanosomatids. **Essays Biochem**, England, v. 51, p. 31-46, 2011. ISSN 1744-1358 (Electronic)0071-1365 (Linking). Disponível em: < <http://dx.doi.org/10.1042/bse0510031> >.

DE MOURA, M. B. et al. Cloning and characterization of DNA polymerase eta from *Trypanosoma cruzi*: roles for translesion bypass of oxidative damage. **Environ Mol Mutagen**, v. 50, n. 5, p. 375-86, Jun 2009. ISSN 0893-6692. Disponível em: < <http://dx.doi.org/10.1002/em.20450> >.

DE PABLOS, L. M. et al. Differential expression and characterization of a member of the mucin-associated surface protein family secreted by *Trypanosoma cruzi*. **Infect Immun**, v. 79, n. 10, p. 3993-4001, Oct 2011. ISSN 0019-9567. Disponível em: < <http://dx.doi.org/10.1128/iai.05329-11> >.

DEANE, M. P.; LENZI, H. L.; JANSEN, A. *Trypanosoma cruzi*: vertebrate and invertebrate cycles in the same mammal host. **Mem Inst Oswaldo Cruz**, v. 79, n. 4, p. 513-5, Oct-Dec 1984. ISSN 0074-0276 (Print)0074-0276. Disponível em: < <http://dx.doi.org/> >.

DEANE, M. P.; LENZI, H. L.; JANSEN, A. M. Double development cycle of *Trypanosoma cruzi* in the opossum. **Parasitol Today**, v. 2, n. 5, p. 146-7, May 1986. ISSN 0169-4758 (Print)0169-4758. Disponível em: < <http://dx.doi.org/> >.

DI NOIA, J. M. et al. AU-rich elements in the 3'-untranslated region of a new mucin-type gene family of *Trypanosoma cruzi* confers mRNA instability and modulates translation efficiency. **J Biol Chem**, v. 275, n. 14, p. 10218-27, Apr 7 2000. ISSN 0021-9258 (Print)0021-9258. Disponível em: < <http://dx.doi.org/> >.

DIAS, J. C.; AMATO NETO, V. Prevention concerning the different alternative routes for transmission of *Trypanosoma cruzi* in Brazil. **Rev Soc Bras Med Trop**, v. 44 Suppl 2, p. 68-72, 2011. ISSN 0037-8682. Disponível em: < <http://dx.doi.org/> >.

DIAS, J. P. et al. Acute Chagas disease outbreak associated with oral transmission. **Rev Soc Bras Med Trop**, v. 41, n. 3, p. 296-300, May-Jun 2008. ISSN 0037-8682 (Print)0037-8682. Disponível em: < <http://dx.doi.org/> >.

DIAZ-BENJUMEA, R. et al. Characterization of a novel cAMP-binding, cAMP-specific cyclic nucleotide phosphodiesterase (TcrPDEB1) from *Trypanosoma cruzi*. **Biochem J**, v. 399, n. 2, p. 305-14, Oct 15 2006. ISSN 0264-6021. Disponível em: < <http://dx.doi.org/10.1042/bj20060757> >.

DNDI. Plataforma de Pesquisa Clínica em Doença de Chagas. 2009. Disponível em: < <http://www.dndi.org.br/pt/doencas-negligenciadas/doenca-de-chagas/plataformachagas.html> >. Acesso em: 09 de maio de 2013.

EDELMAN, A. M.; BLUMENTHAL, D. K.; KREBS, E. G. Protein serine/threonine kinases. **Annu Rev Biochem**, v. 56, p. 567-613, 1987. ISSN 0066-4154 (Print)0066-4154. Disponível em: < <http://dx.doi.org/10.1146/annurev.bi.56.070187.003031> >.

EDMAN, P. A method for the determination of amino acid sequence in peptides. **Arch Biochem**, v. 22, n. 3, p. 475, Jul 1949. Disponível em: < <http://dx.doi.org/> >.

EL-SAYED, N. M. et al. The genome sequence of *Trypanosoma cruzi*, etiologic agent of Chagas disease. **Science**, United States, v. 309, n. 5733, p. 409-15, 2005. ISSN 1095-9203 (Electronic)0036-8075 (Linking). Disponível em: < <http://dx.doi.org/10.1126/science.1112631> >.

ELIAS, J. E.; GYGI, S. P. Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry. **Nat Methods**, v. 4, n. 3, p. 207-14, Mar 2007. ISSN 1548-7091 (Print)1548-7091. Disponível em: < <http://dx.doi.org/10.1038/nmeth1019> >.

ENG, J. K.; MCCORMACK, A. L.; YATES, J. R. An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. **Journal of the American Society for Mass Spectrometry**, v. 5, n. 11, p. 976-989, November 1994 1994. Disponível em: < [http://dx.doi.org/10.1016/1044-0305\(94\)80016-2](http://dx.doi.org/10.1016/1044-0305(94)80016-2) >.

EREZ, E.; FASS, D.; BIBI, E. How intramembrane proteases bury hydrolytic reactions in the membrane. **Nature**, v. 459, n. 7245, p. 371-8, May 21 2009. ISSN 0028-0836. Disponível em: < <http://dx.doi.org/10.1038/nature08146> >.

FAVORETO, S., JR.; DORTA, M. L.; YOSHIDA, N. *Trypanosoma cruzi* 175-kDa protein tyrosine phosphorylation is associated with host cell invasion. **Exp Parasitol**, v. 89, n. 2, p. 188-94, Jun 1998. ISSN 0014-4894 (Print)0014-4894. Disponível em: < <http://dx.doi.org/10.1006/expr.1998.4285> >.

FERNANDES, A. J. et al. [Natural infection of the anal glands of the opossum (*Didelphis albiventris*) by *Trypanosoma cruzi* in the municipality of Bambuí--MG, Brazil]. **Mem Inst Oswaldo Cruz**, v. 84, n. 1, p. 87-93, Jan-Mar 1989. ISSN 0074-0276 (Print)0074-0276. Disponível em: < <http://dx.doi.org/> >.

FERNANDES, M. C. et al. Extracellular amastigotes of *Trypanosoma cruzi* are potent inducers of phagocytosis in mammalian cells. **Cell Microbiol**, v. 15, n. 6, p. 977-91, Jun 2013. ISSN 1462-5814. Disponível em: < <http://dx.doi.org/10.1111/cmi.12090> >.

FERREIRA, R. C.; FERREIRA, L. C. Mutagenicity of nifurtimox and benznidazole in the Salmonella/microsome assay. **Braz J Med Biol Res**, v. 19, n. 1, p. 19-25, 1986. ISSN 0100-879X (Print)0100-879x. Disponível em: < <http://dx.doi.org/> >.

FISCHER, B. et al. NovoHMM: a hidden Markov model for de novo peptide sequencing. **Anal Chem**, v. 77, n. 22, p. 7265-73, Nov 15 2005. ISSN 0003-2700 (Print)0003-2700. Disponível em: < <http://dx.doi.org/10.1021/ac0508853> >.

FLAWIA, M. M.; TELLEZ-INON, M. T.; TORRES, H. N. Signal transduction mechanisms in *Trypanosoma cruzi*. **Parasitol Today**, v. 13, n. 1, p. 30-3, Jan 1997. ISSN 0169-4758 (Print)0169-4758. Disponível em: < <http://dx.doi.org/> >.

FRAIDENRAICH, D. et al. Stimulation of *Trypanosoma cruzi* adenylyl cyclase by an alpha D-globin fragment from *Triatoma hindgut*: effect on differentiation of epimastigote to trypomastigote forms. **Proc Natl Acad Sci U S A**, v. 90, n. 21, p. 10140-4, Nov 1 1993. ISSN 0027-8424 (Print)0027-8424. Disponível em: < <http://dx.doi.org/> >.

FRANK, A.; PEVZNER, P. PepNovo: de novo peptide sequencing via probabilistic network modeling. **Anal Chem**, v. 77, n. 4, p. 964-73, Feb 15 2005. ISSN 0003-2700 (Print)0003-2700. Disponível em: < <http://dx.doi.org/> >.

FRANZEN, O. et al. Shotgun sequencing analysis of *Trypanosoma cruzi* I Sylvio X10/1 and comparison with T. cruzi VI CL Brener. **PLoS Negl Trop Dis**, v. 5, n. 3, p. e984, 2011. ISSN 1935-2735 (Electronic)1935-2727 (Linking). Disponível em: < <http://dx.doi.org/10.1371/journal.pntd.0000984> >.

FRASCH, A. C. Functional diversity in the trans-sialidase and mucin families in *Trypanosoma cruzi*. **Parasitol Today**, v. 16, n. 7, p. 282-6, Jul 2000. ISSN 0169-4758 (Print)0169-4758. Disponível em: < <http://dx.doi.org/> >.

FREIRE-DE-LIMA, L. et al. Sialic acid: a sweet swing between mammalian host and *Trypanosoma cruzi*. **Front Immunol**, v. 3, p. 356, 2012. ISSN 1664-3224. Disponível em: < <http://dx.doi.org/10.3389/fimmu.2012.00356> >.

FREITAS, L. M. et al. Genomic analyses, gene expression and antigenic profile of the trans-sialidase superfamily of *Trypanosoma cruzi* reveal an undetected level of complexity. **PLoS One**, v. 6, n. 10, p. e25914, 2011. ISSN 1932-6203. Disponível em: < <http://dx.doi.org/10.1371/journal.pone.0025914> >.

GARCIA, M. P. et al. Characterisation of a *Trypanosoma cruzi* acidic 30 kDa cysteine protease. **Mol Biochem Parasitol**, v. 91, n. 2, p. 263-72, Mar 15 1998. ISSN 0166-6851 (Print)0166-6851. Disponível em: < <http://dx.doi.org/> >.

GILBER, S. R. et al. Comparison of conventional serology and PCR methods for the routine diagnosis of *Trypanosoma cruzi* infection. **Rev Soc Bras Med Trop**, v. 46, n. 3, p.

310-5, May-Jun 2013. ISSN 0037-8682. Disponível em: < <http://dx.doi.org/10.1590/0037-8682-0046-2013> >.

GONZALES-PERDOMO, M.; ROMERO, P.; GOLDENBERG, S. Cyclic AMP and adenylate cyclase activators stimulate *Trypanosoma cruzi* differentiation. **Exp Parasitol**, v. 66, n. 2, p. 205-12, Aug 1988. ISSN 0014-4894 (Print)0014-4894. Disponível em: < <http://dx.doi.org/> >.

GONZALEZ, J. et al. Proteasome activity is required for the stage-specific transformation of a protozoan parasite. **J Exp Med**, v. 184, n. 5, p. 1909-18, Nov 1 1996. ISSN 0022-1007 (Print)0022-1007. Disponível em: < <http://dx.doi.org/> >.

GORLA, N. B. Sister-chromatid exchange in splenic lymphocytes of mice after exposure to. **Mutat Res**, v. 188, n. 2, p. 129-33, Jun 1987. ISSN 0027-5107 (Print)0027-5107. Disponível em: < <http://dx.doi.org/> >.

GORLA, N. B. et al. Assessment of cytogenetic damage in chagasic children treated with benznidazole. **Mutat Res**, v. 206, n. 2, p. 217-20, Oct 1988. ISSN 0027-5107 (Print)0027-5107. Disponível em: < <http://dx.doi.org/> >.

_____. Thirteenfold increase of chromosomal aberrations non-randomly distributed in. **Mutat Res**, v. 224, n. 2, p. 263-7, Oct 1989. ISSN 0027-5107 (Print)0027-5107. Disponível em: < <http://dx.doi.org/> >.

GREBAUT, P. et al. Identification of total and differentially expressed excreted-secreted proteins from *Trypanosoma congolense* strains exhibiting different virulence and pathogenicity. **Int J Parasitol**, v. 39, n. 10, p. 1137-50, Aug 2009. ISSN 0020-7519. Disponível em: < <http://dx.doi.org/10.1016/j.ijpara.2009.02.018> >.

GRELLIER, P. et al. *Trypanosoma cruzi* prolyl oligopeptidase Tc80 is involved in nonphagocytic mammalian cell invasion by trypomastigotes. **J Biol Chem**, v. 276, n. 50, p. 47078-86, Dec 14 2001. ISSN 0021-9258 (Print)0021-9258. Disponível em: < <http://dx.doi.org/10.1074/jbc.M106017200> >.

GURGEL-GONÇALVES, R. et al. Geographic distribution of chagas disease vectors in Brazil based on ecological niche modeling. **J Trop Med**, v. 2012, p. 705326, 2012. ISSN 1687-9694 (Electronic)1687-9686 (Linking). Disponível em: < <http://dx.doi.org/10.1155/2012/705326> >.

GUTTERIDGE, W. E.; GABORAK, M. A re-examination of purine and pyrimidine synthesis in the three main forms of *Trypanosoma cruzi*. **Int J Biochem**, v. 10, n. 5, p. 415-22, 1979. ISSN 0020-711X (Print)0020-711x. Disponível em: < <http://dx.doi.org/> >.

HAN, X. et al. PeaksPTM: Mass spectrometry-based identification of peptides with unspecified modifications. **J Proteome Res**, v. 10, n. 7, p. 2930-6, Jul 1 2011. ISSN 1535-3893. Disponível em: < <http://dx.doi.org/10.1021/pr200153k> >.

HAN, Y.; MA, B.; ZHANG, K. SPIDER: software for protein identification from sequence tags with de novo sequencing error. **Proc IEEE Comput Syst Bioinform Conf**, p. 206-15, 2004. ISSN 1551-7497 (Print)1551-7497. Disponível em: < <http://dx.doi.org/> >.

HANNAERT, V. et al. Evolution of energy metabolism and its compartmentation in Kinetoplastida. **Kinetoplastid Biol Dis**, v. 2, n. 1, p. 11, Oct 28 2003. ISSN 1475-9292 (Print)1475-9292. Disponível em: < <http://dx.doi.org/10.1186/1475-9292-2-11> >.

HENDRICKSON, N.; ALLEN, T.; ULLMAN, B. Molecular characterization of phosphoribosylpyrophosphate synthetase from *Leishmania donovani*. **Mol Biochem Parasitol**, v. 59, n. 1, p. 15-27, May 1993. ISSN 0166-6851 (Print)0166-6851. Disponível em: < <http://dx.doi.org/> >.

HOFT, D. F. Differential mucosal infectivity of different life stages of *Trypanosoma cruzi*. **Am J Trop Med Hyg**, v. 55, n. 4, p. 360-4, Oct 1996. ISSN 0002-9637 (Print)0002-9637. Disponível em: < <http://dx.doi.org/> >.

HOFT, D. F. et al. Gastric invasion by *Trypanosoma cruzi* and induction of protective mucosal immune responses. **Infect Immun**, v. 64, n. 9, p. 3800-10, Sep 1996. ISSN 0019-9567 (Print)0019-9567 (Linking). Disponível em: < <http://dx.doi.org/> >.

HOHMANN, S.; MEACOCK, P. A. Thiamin metabolism and thiamin diphosphate-dependent enzymes in the yeast *Saccharomyces cerevisiae*: genetic regulation. **Biochim Biophys Acta**, v. 1385, n. 2, p. 201-19, Jun 29 1998. ISSN 0006-3002 (Print)0006-3002. Disponível em: < <http://dx.doi.org/> >.

ISSA, V. S.; BOCCHI, E. A. Antitrypanosomal agents: treatment or threat? **Lancet**, England, v. 376, n. 9743, p. 768; author reply 768-9, 2010. ISSN 1474-547X (Electronic)0140-6736 (Linking). Disponível em: < [http://dx.doi.org/10.1016/s0140-6736\(10\)61372-4](http://dx.doi.org/10.1016/s0140-6736(10)61372-4) >.

IVENS, A. C. et al. The genome of the kinetoplastid parasite, *Leishmania major*. **Science**, United States, v. 309, n. 5733, p. 436-42, 2005. ISSN 1095-9203 (Electronic)0036-8075 (Linking). Disponível em: < <http://dx.doi.org/10.1126/science.1112680> >.

JENSEN, R. E.; ENGLUND, P. T. Network news: the replication of kinetoplast DNA. **Annu Rev Microbiol**, v. 66, p. 473-91, 2012. ISSN 1545-3251 (Electronic)0066-4227 (Linking). Disponível em: < <http://dx.doi.org/10.1146/annurev-micro-092611-150057> >.

JOSE CAZZULO, J.; STOKA, V.; TURK, V. The major cysteine proteinase of *Trypanosoma cruzi*: a valid target for chemotherapy of Chagas disease. **Curr Pharm Des**, v. 7, n. 12, p. 1143-56, Aug 2001. ISSN 1381-6128 (Print)1381-6128. Disponível em: < <http://dx.doi.org/> >.

JUNQUEIRA, A. C.; DEGRAVE, W.; BRANDAO, A. Minicircle organization and diversity in *Trypanosoma cruzi* populations. **Trends Parasitol**, England, v. 21, n. 6, p. 270-2, 2005. ISSN 1471-4922 (Print)1471-4922 (Linking). Disponível em: < <http://dx.doi.org/10.1016/j.pt.2005.04.001> >.

JUNQUEIRA, M. R. **Systematic Characterization of Mammalian Protein Complexes by Shotgun Liquid Chromatography Tandem Mass Spectrometry**. 2009. (Tese de doutorado). Max Planck Institute of Molecular Cell Biology and Genetics.

KALLI, A. et al. Evaluation and Optimization of Mass Spectrometric Settings during Data-dependent Acquisition Mode: Focus on LTQ-Orbitrap Mass Analyzers. **J Proteome Res**, v. 12, n. 7, p. 3071-86, Jul 5 2013. ISSN 1535-3893. Disponível em: < <http://dx.doi.org/10.1021/pr3011588> >.

KAWASHITA, S. Y. et al. Homology, paralogy and function of DGF-1, a highly dispersed Trypanosoma cruzi specific gene family and its implications for information entropy of its encoded proteins. **Mol Biochem Parasitol**, v. 165, n. 1, p. 19-31, May 2009. ISSN 0166-6851 (Print)0166-6851. Disponível em: < <http://dx.doi.org/10.1016/j.molbiopara.2008.12.010> >.

KIDDER, G. W. Adenosine kinase from Trypanosoma cruzi. **Biochem Biophys Res Commun**, v. 107, n. 1, p. 381-8, Jul 16 1982. ISSN 0006-291X (Print)0006-291x. Disponível em: < <http://dx.doi.org/> >.

KIERSZENBAUM, F.; GOTTLIEB, C. A.; BUDZKO, D. B. Antibody-independent, natural resistance of birds to Trypanosoma cruzi infection. **J Parasitol**, v. 67, n. 5, p. 656-60, Oct 1981. ISSN 0022-3395 (Print)0022-3395. Disponível em: < <http://dx.doi.org/> >.

KIM, H. M.; YU, Y.; CHENG, Y. Structure characterization of the 26S proteasome. **Biochim Biophys Acta**, v. 1809, n. 2, p. 67-79, Feb 2011. ISSN 0006-3002 (Print)0006-3002. Disponível em: < <http://dx.doi.org/10.1016/j.bbagr.2010.08.008> >.

KOLLIEN, A. H.; SCHAUB, G. A. The development of Trypanosoma cruzi (Trypanosomatidae) in the reduviid bug. **J Eukaryot Microbiol**, v. 45, n. 1, p. 59-63, Jan-Feb 1998. ISSN 1066-5234 (Print)1066-5234. Disponível em: < <http://dx.doi.org/> >.

KROPF, S. P.; SA, M. R. The discovery of Trypanosoma cruzi and Chagas disease (1908-1909): tropical medicine in Brazil. **Hist Cienc Saude Manguinhos**, v. 16 Suppl 1, p. 13-34, Jul 2009. ISSN 0104-5970 (Print)0104-5970 (Linking). Disponível em: < <http://dx.doi.org/> >.

KULKARNI, M. M. et al. The major surface-metalloprotease of the parasitic protozoan, Leishmania, protects against antimicrobial peptide-induced apoptotic killing. **Mol Microbiol**, v. 62, n. 5, p. 1484-97, Dec 2006. ISSN 0950-382X (Print)0950-382x. Disponível em: < <http://dx.doi.org/10.1111/j.1365-2958.2006.05459.x> >.

_____. Trypanosoma cruzi GP63 proteins undergo stage-specific differential posttranslational modification and are important for host cell infection. **Infect Immun**, v. 77, n. 5, p. 2193-200, May 2009. ISSN 0019-9567. Disponível em: < <http://dx.doi.org/10.1128/iai.01542-08> >.

KUNZ, S. et al. Cyclic nucleotide specific phosphodiesterases of the kinetoplastida: a unified nomenclature. **Mol Biochem Parasitol**, v. 145, n. 1, p. 133-5, Jan 2006. ISSN

0166-6851 (Print)0166-6851. Disponível em: < <http://dx.doi.org/10.1016/j.molbiopara.2005.09.018> >.

LANDER, N. et al. Localization and developmental regulation of a dispersed gene family 1 protein in *Trypanosoma cruzi*. **Infect Immun**, v. 78, n. 1, p. 231-40, Jan 2010. ISSN 0019-9567. Disponível em: < <http://dx.doi.org/10.1128/iai.00780-09> >.

LEE, M. C. et al. Bi-directional protein transport between the ER and Golgi. **Annu Rev Cell Dev Biol**, v. 20, p. 87-123, 2004. ISSN 1081-0706 (Print)1081-0706. Disponível em: < <http://dx.doi.org/10.1146/annurev.cellbio.20.010403.105307> >.

LENT, H.; WYGODZINSKY, P. Revision of the Triatominae (Hemiptera, Reduviidae), and their significance as vector of Chagas' disease. **Bull Am Mus Nat History**, v. 163, p. 123-520, 1979.

LEROUX, A. E. et al. Functional characterization of NADP-dependent isocitrate dehydrogenase isozymes from *Trypanosoma cruzi*. **Mol Biochem Parasitol**, v. 177, n. 1, p. 61-4, May 2011. ISSN 0166-6851. Disponível em: < <http://dx.doi.org/10.1016/j.molbiopara.2011.01.010> >.

_____. Comparative studies on the biochemical properties of the malic enzymes from *Trypanosoma cruzi* and *Trypanosoma brucei*. **FEMS Microbiol Lett**, v. 314, n. 1, p. 25-33, Jan 2011. ISSN 0378-1097. Disponível em: < <http://dx.doi.org/10.1111/j.1574-6968.2010.02142.x> >.

LEWINSOHN, R. Carlos Chagas and the discovery of Chagas' disease (American trypanosomiasis). **J R Soc Med**, v. 74, n. 6, p. 451-5, Jun 1981. ISSN 0141-0768 (Print)0141-0768 (Linking). Disponível em: < <http://dx.doi.org/> >.

LI, Z. H. et al. A 43-nucleotide U-rich element in 3'-untranslated region of large number of *Trypanosoma cruzi* transcripts is important for mRNA abundance in intracellular amastigotes. **J Biol Chem**, v. 287, n. 23, p. 19058-69, Jun 1 2012. ISSN 0021-9258. Disponível em: < <http://dx.doi.org/10.1074/jbc.M111.338699> >.

MA, B. et al. PEAKS: powerful software for peptide de novo sequencing by tandem mass spectrometry. **Rapid Commun Mass Spectrom**, v. 17, n. 20, p. 2337-42, 2003. ISSN 0951-4198 (Print)0951-4198. Disponível em: < <http://dx.doi.org/10.1002/rcm.1196> >.

MAGALHAES, A. D. et al. *Trypanosoma cruzi* alkaline 2-DE: Optimization and application to comparative proteome analysis of flagellate life stages. **Proteome Sci**, v. 6, p. 24, 2008. ISSN 1477-5956. Disponível em: < <http://dx.doi.org/10.1186/1477-5956-6-24> >.

MAIR, G. et al. A new twist in trypanosome RNA metabolism: cis-splicing of pre-mRNA. **Rna**, v. 6, n. 2, p. 163-9, Feb 2000. ISSN 1355-8382 (Print)1355-8382. Disponível em: < <http://dx.doi.org/> >.

MAKAROV, A. Electrostatic axially harmonic orbital trapping: a high-performance technique of mass analysis. **Anal Chem**, v. 72, n. 6, p. 1156-62, Mar 15 2000. ISSN 0003-2700. Disponível em: < <http://dx.doi.org/> >.

MAKAROV, A.; SCIGELOVA, M. Coupling liquid chromatography to Orbitrap mass spectrometry. **J Chromatogr A**, v. 1217, n. 25, p. 3938-45, Jun 18 2010. ISSN 0021-9673. Disponível em: < <http://dx.doi.org/10.1016/j.chroma.2010.02.022> >.

MALAGA, S.; YOSHIDA, N. Targeted reduction in expression of Trypanosoma cruzi surface glycoprotein gp90 increases parasite infectivity. **Infect Immun**, v. 69, n. 1, p. 353-9, Jan 2001. ISSN 0019-9567 (Print)0019-9567. Disponível em: < <http://dx.doi.org/10.1128/iai.69.1.353-359.2001> >.

MARCILI, A. et al. Trypanosoma cruzi in Brazilian Amazonia: Lineages TCI and TCIIa in wild primates, Rhodnius spp. and in humans with Chagas disease associated with oral transmission. **Int J Parasitol**, v. 39, n. 5, p. 615-23, Apr 2009. ISSN 0020-7519. Disponível em: < <http://dx.doi.org/10.1016/j.ijpara.2008.09.015> >.

MARIN-NETO, J. A. et al. Rationale and design of a randomized placebo-controlled trial assessing the. **Am Heart J**, v. 156, n. 1, p. 37-43, Jul 2008. ISSN 0002-8703. Disponível em: < <http://dx.doi.org/10.1016/j.ahj.2008.04.001> >.

MARTINEZ-CALVILLO, S. et al. Gene expression in trypanosomatid parasites. **J Biomed Biotechnol**, v. 2010, p. 525241, 2010. ISSN 1110-7243. Disponível em: < <http://dx.doi.org/10.1155/2010/525241> >.

MARTINS-MELO, F. R. et al. Mortality due to Chagas disease in Brazil from 1979 to 2009: trends and regional differences. **J Infect Dev Ctries**, v. 6, n. 11, p. 817-24, 2012. ISSN 1972-2680. Disponível em: < <http://dx.doi.org/10.3855/jidc.2459> >.

MATOBA, K. et al. Crystallization and preliminary X-ray analysis of aspartate transcarbamoylase from the parasitic protist Trypanosoma cruzi. **Acta Crystallogr Sect F Struct Biol Cryst Commun**, v. 65, n. Pt 9, p. 933-6, Sep 1 2009. ISSN 1744-3091. Disponível em: < <http://dx.doi.org/10.1107/s1744309109031959> >.

MATTHEWS, K. R. The developmental cell biology of Trypanosoma brucei. **J Cell Sci**, v. 118, n. Pt 2, p. 283-90, Jan 15 2005. ISSN 0021-9533 (Print)0021-9533. Disponível em: < <http://dx.doi.org/10.1242/jcs.01649> >.

MCGWIRE, B. S.; CHANG, K. P.; ENGMAN, D. M. Migration through the extracellular matrix by the parasitic protozoan Leishmania is enhanced by surface metalloprotease gp63. **Infect Immun**, v. 71, n. 2, p. 1008-10, Feb 2003. ISSN 0019-9567 (Print)0019-9567. Disponível em: < <http://dx.doi.org/> >.

MEDINA-ACOSTA, E.; CROSS, G. A. Rapid isolation of DNA from trypanosomatid protozoa using a simple 'mini-prep' procedure. **Mol Biochem Parasitol**, v. 59, n. 2, p. 327-9, Jun 1993. ISSN 0166-6851 (Print)0166-6851 (Linking). Disponível em: < <http://dx.doi.org/> >.

MERRITT, E. A. et al. Crystal structures of trypanosomal histidyl-tRNA synthetase illuminate differences between eukaryotic and prokaryotic homologs. **J Mol Biol**, v. 397,

n. 2, p. 481-94, Mar 26 2010. ISSN 0022-2836. Disponível em: < <http://dx.doi.org/10.1016/j.jmb.2010.01.051> >.

MILAGROS CAMARA MDE, L. et al. Identification and validation of Trypanosoma cruzi's glycosomal adenylate kinase containing a peroxisomal targeting signal. **Exp Parasitol**, v. 130, n. 4, p. 408-11, Apr 2012. ISSN 0014-4894. Disponível em: < <http://dx.doi.org/10.1016/j.exppara.2012.01.020> >.

MONTEIRO, W. M. et al. Trypanosoma cruzi IV causing outbreaks of acute Chagas disease and infections by different haplotypes in the Western Brazilian Amazonia. **PLoS One**, v. 7, n. 7, p. e41284, 2012. ISSN 1932-6203. Disponível em: < <http://dx.doi.org/10.1371/journal.pone.0041284> >.

_____. Trypanosoma cruzi TcIII/Z3 genotype as agent of an outbreak of Chagas disease in the Brazilian Western Amazonia. **Trop Med Int Health**, v. 15, n. 9, p. 1049-51, Sep 2010. ISSN 1360-2276. Disponível em: < <http://dx.doi.org/10.1111/j.1365-3156.2010.02577.x> >.

MORAES BARROS, R. R. et al. Anatomy and evolution of telomeric and subtelomeric regions in the human protozoan parasite Trypanosoma cruzi. **BMC Genomics**, v. 13, p. 229, 2012. ISSN 1471-2164. Disponível em: < <http://dx.doi.org/10.1186/1471-2164-13-229> >.

MORTARA, R. A. et al. Polymorphism of the 35- and 50-kilodalton surface glycoconjugates of Trypanosoma cruzi metacyclic trypomastigotes. **Infect Immun**, v. 60, n. 11, p. 4673-8, Nov 1992. ISSN 0019-9567 (Print)0019-9567. Disponível em: < <http://dx.doi.org/> >.

MOTTA, F. N. et al. The Trypanosoma cruzi virulence factor oligopeptidase B (OPBTc) assembles into an active and stable dimer. **PLoS One**, v. 7, n. 1, p. e30431, 2012. ISSN 1932-6203. Disponível em: < <http://dx.doi.org/10.1371/journal.pone.0030431> >.

MURCIA, L. et al. Diagnóstico y tratamiento de la enfermedad de Chagas. **Enferm Infecc Microbiol Clin**, v. 31 Suppl 1, p. 26-34, Feb 2013. ISSN 0213-005x. Disponível em: < [http://dx.doi.org/10.1016/s0213-005x\(13\)70111-3](http://dx.doi.org/10.1016/s0213-005x(13)70111-3) >.

NARA, T. et al. Molecular interaction of the first 3 enzymes of the de novo pyrimidine biosynthetic pathway of Trypanosoma cruzi. **Biochem Biophys Res Commun**, v. 418, n. 1, p. 140-3, Feb 3 2012. ISSN 0006-291x. Disponível em: < <http://dx.doi.org/10.1016/j.bbrc.2011.12.148> >.

NAULA, C.; PARSONS, M.; MOTTRAM, J. C. Protein kinases as drug targets in trypanosomes and Leishmania. **Biochim Biophys Acta**, v. 1754, n. 1-2, p. 151-9, Dec 30 2005. ISSN 0006-3002 (Print)0006-3002. Disponível em: < <http://dx.doi.org/10.1016/j.bbapap.2005.08.018> >.

NEIRA, I. et al. Involvement of Trypanosoma cruzi metacyclic trypomastigote surface molecule gp82 in adhesion to gastric mucin and invasion of epithelial cells. **Infect**

Immun, v. 71, n. 1, p. 557-61, Jan 2003. ISSN 0019-9567 (Print)0019-9567. Disponível em: < <http://dx.doi.org/> >.

NEVES, D. **Metiltioadenosina fosforilase de Trypanosoma cruzi, um alvo potencial para quimioterapia da doença de Chagas, apresenta ampla especificidade a substratos e elevada estabilidade estrutural**. 2006. (Tese de doutorado). Faculdade de Medicina, Universidade de Brasília (UnB)

NICKEL, W.; RABOUILLE, C. Mechanisms of regulated unconventional protein secretion. **Nat Rev Mol Cell Biol**, v. 10, n. 2, p. 148-55, Feb 2009. ISSN 1471-0072. Disponível em: < <http://dx.doi.org/10.1038/nrm2617> >.

NICKEL, W.; SEEDORF, M. Unconventional mechanisms of protein transport to the cell surface of eukaryotic cells. **Annu Rev Cell Dev Biol**, v. 24, p. 287-308, 2008. ISSN 1081-0706 (Print)1081-0706. Disponível em: < <http://dx.doi.org/10.1146/annurev.cellbio.24.110707.175320> >.

NOBREGA, A. A. et al. Oral transmission of Chagas disease by consumption of acai palm fruit, Brazil. **Emerg Infect Dis**, v. 15, n. 4, p. 653-5, Apr 2009. ISSN 1080-6040. Disponível em: < <http://dx.doi.org/10.3201/eid1504.081450> >.

NOBREGA, O. T. et al. Cloning and sequencing of tccb, a gene encoding a Trypanosoma cruzi cathepsin B-like protease. **Mol Biochem Parasitol**, v. 97, n. 1-2, p. 235-40, Nov 30 1998. ISSN 0166-6851 (Print)0166-6851. Disponível em: < <http://dx.doi.org/> >.

NOIREAU, F.; DIOSQUE, P.; JANSEN, A. M. Trypanosoma cruzi: adaptation to its vectors and its hosts. **Vet Res**, v. 40, n. 2, p. 26, Mar-Apr 2009. ISSN 0928-4249 (Print)0928-4249. Disponível em: < <http://dx.doi.org/10.1051/vetres/2009009> >.

PABA, J. et al. Proteomic analysis of Trypanosoma cruzi developmental stages using isotope-coded affinity tag reagents. **J Proteome Res**, v. 3, n. 3, p. 517-24, May-Jun 2004. ISSN 1535-3893 (Print)1535-3893. Disponível em: < <http://dx.doi.org/> >.

_____. Proteomic analysis of the human pathogen Trypanosoma cruzi. **Proteomics**, v. 4, n. 4, p. 1052-9, Apr 2004. ISSN 1615-9853 (Print)1615-9853 (Linking). Disponível em: < <http://dx.doi.org/10.1002/pmic.200300637> >.

PANDEY, A.; MANN, M. Proteomics to study genes and genomes. **Nature**, v. 405, n. 6788, p. 837-46, Jun 15 2000. ISSN 0028-0836 (Print)0028-0836. Disponível em: < <http://dx.doi.org/10.1038/35015709> >.

PARODI-TALICE, A. et al. Proteome analysis of the causative agent of Chagas disease: Trypanosoma cruzi. **Int J Parasitol**, England, v. 34, n. 8, p. 881-6, 2004. ISSN 0020-7519 (Print)0020-7519 (Linking). Disponível em: < <http://dx.doi.org/10.1016/j.ijpara.2004.05.002> >.

PEREIRA-CHIOCCOLA, V. L. et al. Mucin-like molecules form a negatively charged coat that protects Trypanosoma cruzi trypomastigotes from killing by human anti-alpha-

galactosyl antibodies. **J Cell Sci**, v. 113 (Pt 7), p. 1299-307, Apr 2000. ISSN 0021-9533 (Print)0021-9533. Disponível em: < <http://dx.doi.org/> >.

PERKINS, D. N. et al. Probability-based protein identification by searching sequence databases using mass spectrometry data. **Electrophoresis**, v. 20, n. 18, p. 3551-67, Dec 1999. ISSN 0173-0835 (Print)0173-0835. Disponível em: < [http://dx.doi.org/10.1002/\(sici\)1522-2683\(19991201\)20:18%3C3551::aid-elps3551%3E3.0.co](http://dx.doi.org/10.1002/(sici)1522-2683(19991201)20:18%3C3551::aid-elps3551%3E3.0.co) >.

PETERSEN, T. N. et al. SignalP 4.0: discriminating signal peptides from transmembrane regions. **Nat Methods**, United States, v. 8, n. 10, p. 785-6, 2011. ISSN 1548-7105 (Electronic)1548-7091 (Linking). Disponível em: < <http://dx.doi.org/10.1038/nmeth.1701> >.

QUEIROZ, R. M. L. **Proteômica de *Trypanosoma cruzi*: variações em subproteomas durante a amastigogênese**. 2013. (Tese de doutorado). Faculdade de Medicina, Universidade de Brasília (UnB)

RAJAO, M. A. et al. DNA polymerase kappa from *Trypanosoma cruzi* localizes to the mitochondria, bypasses 8-oxoguanine lesions and performs DNA synthesis in a recombination intermediate. **Mol Microbiol**, v. 71, n. 1, p. 185-97, Jan 2009. ISSN 0950-382x. Disponível em: < <http://dx.doi.org/10.1111/j.1365-2958.2008.06521.x> >.

RASSI, A., JR.; RASSI, A.; MARCONDES DE REZENDE, J. American trypanosomiasis (Chagas disease). **Infect Dis Clin North Am**, United States, v. 26, n. 2, p. 275-91, 2012. ISSN 1557-9824 (Electronic)0891-5520 (Linking). Disponível em: < <http://dx.doi.org/10.1016/j.idc.2012.03.002> >.

RASSI, A., JR.; RASSI, A.; MARIN-NETO, J. A. Chagas disease. **Lancet**, England, v. 375, n. 9723, p. 1388-402, 2010. ISSN 1474-547X (Electronic)0140-6736 (Linking). Disponível em: < [http://dx.doi.org/10.1016/s0140-6736\(10\)60061-x](http://dx.doi.org/10.1016/s0140-6736(10)60061-x) >.

RAWLINGS, N. D.; BARRET, A. J. Introduction: Metallopeptidases e their clans. In: (Ed.). **Handbook of Proteolytic Enzymes**. 2 ed. San Diego, USA.: Elsevier/Academic Press, 2004. p.231-263.

RAWLINGS, N. D.; BARRETT, A. J.; BATEMAN, A. MEROPS: the database of proteolytic enzymes, their substrates and inhibitors. **Nucleic Acids Res**, v. 40, n. Database issue, p. D343-50, Jan 2012. ISSN 0305-1048. Disponível em: < <http://dx.doi.org/10.1093/nar/gkr987> >.

REINA-SAN-MARTIN, B. et al. A B-cell mitogen from a pathogenic trypanosome is a eukaryotic proline racemase. **Nat Med**, v. 6, n. 8, p. 890-7, Aug 2000. ISSN 1078-8956 (Print)1078-8956. Disponível em: < <http://dx.doi.org/10.1038/78651> >.

RENUSE, S.; CHAERKADY, R.; PANDEY, A. Proteogenomics. **Proteomics**, v. 11, n. 4, p. 620-30, Feb 2011. ISSN 1615-9853. Disponível em: < <http://dx.doi.org/10.1002/pmic.201000615> >.

RODRIGUES, A. A. et al. IFN-gamma plays a unique role in protection against low virulent *Trypanosoma cruzi* strain. **PLoS Negl Trop Dis**, United States, v. 6, n. 4, p. e1598, 2012. ISSN 1935-2735 (Electronic)1935-2727 (Linking). Disponível em: < <http://dx.doi.org/10.1371/journal.pntd.0001598> >.

RODRIGUEZ, A. et al. A trypanosome-soluble factor induces IP3 formation, intracellular Ca²⁺ mobilization and microfilament rearrangement in host cells. **J Cell Biol**, v. 129, n. 5, p. 1263-73, Jun 1995. ISSN 0021-9525 (Print)0021-9525. Disponível em: < <http://dx.doi.org/> >.

RUDNICK, G.; ABELES, R. H. Reaction mechanism and structure of the active site of proline racemase. **Biochemistry**, v. 14, n. 20, p. 4515-22, Oct 7 1975. ISSN 0006-2960 (Print)0006-2960. Disponível em: < <http://dx.doi.org/> >.

RUIZ, R. C. et al. Infectivity of *Trypanosoma cruzi* strains is associated with differential expression of surface glycoproteins with differential Ca²⁺ signalling activity. **Biochem J**, v. 330 (Pt 1), p. 505-11, Feb 15 1998. ISSN 0264-6021 (Print)0264-6021 (Linking). Disponível em: < <http://dx.doi.org/> >.

RUSSELL, D. G.; TALAMAS-ROHANA, P.; ZELECHOWSKI, J. Antibodies raised against synthetic peptides from the Arg-Gly-Asp-containing region of the *Leishmania* surface protein gp63 cross-react with human C3 and interfere with gp63-mediated binding to macrophages. **Infect Immun**, v. 57, n. 2, p. 630-2, Feb 1989. ISSN 0019-9567 (Print)0019-9567. Disponível em: < <http://dx.doi.org/> >.

SANTANA, J. M. et al. Purification and characterization of a new 120 kDa alkaline proteinase of *Trypanosoma cruzi*. **Biochem Biophys Res Commun**, v. 187, n. 3, p. 1466-73, Sep 30 1992. ISSN 0006-291X (Print)0006-291x. Disponível em: < <http://dx.doi.org/> >.

_____. A *Trypanosoma cruzi*-secreted 80 kDa proteinase with specificity for human collagen types I and IV. **Biochem J**, v. 325 (Pt 1), p. 129-37, Jul 1 1997. ISSN 0264-6021 (Print)0264-6021. Disponível em: < <http://dx.doi.org/> >.

SCHENKMAN, S. et al. Structural and functional properties of *Trypanosoma* trans-sialidase. **Annu Rev Microbiol**, v. 48, p. 499-523, 1994. ISSN 0066-4227 (Print)0066-4227. Disponível em: < <http://dx.doi.org/10.1146/annurev.mi.48.100194.002435> >.

SCHENKMAN, S.; PONTES DE CARVALHO, L.; NUSSENZWEIG, V. *Trypanosoma cruzi* trans-sialidase and neuraminidase activities can be mediated by the same enzymes. **J Exp Med**, v. 175, n. 2, p. 567-75, Feb 1 1992. ISSN 0022-1007 (Print)0022-1007. Disponível em: < <http://dx.doi.org/> >.

SCHIJMAN, A. G. et al. International study to evaluate PCR methods for detection of *Trypanosoma cruzi*. **PLoS Negl Trop Dis**, v. 5, n. 1, p. e931, 2011. ISSN 1935-2727. Disponível em: < <http://dx.doi.org/10.1371/journal.pntd.0000931> >.

SCHOIJET, A. C. et al. Defining the role of a FYVE domain in the localization and activity of a cAMP phosphodiesterase implicated in osmoregulation in *Trypanosoma cruzi*. **Mol**

Microbiol, v. 79, n. 1, p. 50-62, Jan 2011. ISSN 0950-382x. Disponível em: < <http://dx.doi.org/10.1111/j.1365-2958.2010.07429.x> >.

SCIGELOVA, M.; MAKAROV, A. Orbitrap mass analyzer--overview and applications in proteomics. **Proteomics**, v. 6 Suppl 2, p. 16-21, Sep 2006. ISSN 1615-9861 (Electronic)1615-9853 (Linking). Disponível em: < <http://dx.doi.org/10.1002/pmic.200600528> >.

SHEVCHENKO, A. et al. In-gel digestion for mass spectrometric characterization of proteins and proteomes. **Nat Protoc**, England, v. 1, n. 6, p. 2856-60, 2006. ISSN 1750-2799 (Electronic)1750-2799 (Linking). Disponível em: < <http://dx.doi.org/10.1038/nprot.2006.468> >.

SILVA, A. R. **Transmissão sexual do Trypanosoma cruzi em Mus musculus**. 2013. (Dissertação de mestrado). Faculdade de Medicina, Universidade de Brasília (UnB)

SILVEIRA, A. C.; DIAS, J. C. [The control of vectorial transmission]. **Rev Soc Bras Med Trop**, v. 44 Suppl 2, p. 52-63, 2011. ISSN 0037-8682. Disponível em: < <http://dx.doi.org/> >.

SOUTO, R. P. et al. DNA markers define two major phylogenetic lineages of Trypanosoma cruzi. **Mol Biochem Parasitol**, v. 83, n. 2, p. 141-52, Dec 20 1996. ISSN 0166-6851 (Print)0166-6851. Disponível em: < <http://dx.doi.org/> >.

SOUTO, R. P.; ZINGALES, B. Sensitive detection and strain classification of Trypanosoma cruzi by amplification of a ribosomal RNA sequence. **Mol Biochem Parasitol**, v. 62, n. 1, p. 45-52, Nov 1993. ISSN 0166-6851 (Print)0166-6851. Disponível em: < <http://dx.doi.org/> >.

SOUZA, R. T. et al. Genome size, karyotype polymorphism and chromosomal evolution in Trypanosoma cruzi. **PLoS One**, United States, v. 6, n. 8, p. e23042, 2011. ISSN 1932-6203 (Electronic)1932-6203 (Linking). Disponível em: < <http://dx.doi.org/10.1371/journal.pone.0023042> >.

SOUZA-LIMA RDE, C. et al. Outbreak of acute Chagas disease associated with oral transmission in the Rio Negro region, Brazilian Amazon. **Rev Soc Bras Med Trop**, v. 46, n. 4, p. 510-4, Jul-Aug 2013. ISSN 0037-8682. Disponível em: < <http://dx.doi.org/10.1590/0037-8682-1367-2013> >.

SPECTOR, T.; BERENS, R. L.; MARR, J. J. Adenylosuccinate synthetase and adenylosuccinate lyase from Trypanosoma cruzi, Specificity studies with potential chemotherapeutic agents. **Biochem Pharmacol**, v. 31, n. 2, p. 225-9, Jan 15 1982. ISSN 0006-2952 (Print)0006-2952. Disponível em: < <http://dx.doi.org/> >.

STEEN, H.; MANN, M. The ABC's (and XYZ's) of peptide sequencing. **Nat Rev Mol Cell Biol**, v. 5, n. 9, p. 699-711, Sep 2004. ISSN 1471-0072 (Print)1471-0072. Disponível em: < <http://dx.doi.org/10.1038/nrm1468> >.

STEINDEL, M. et al. Characterization of *Trypanosoma cruzi* isolated from humans, vectors, and animal reservoirs following an outbreak of acute human Chagas disease in Santa Catarina State, Brazil. **Diagn Microbiol Infect Dis**, United States, v. 60, n. 1, p. 25-32, 2008. ISSN 0732-8893 (Print)0732-8893 (Linking). Disponível em: < <http://dx.doi.org/10.1016/j.diagmicrobio.2007.07.016> >.

_____. Presence of *Trypanosoma cruzi* in the anal glands of naturally infected opossum. **Mem Inst Oswaldo Cruz**, v. 83, n. 1, p. 135-7, Jan-Mar 1988. ISSN 0074-0276 (Print)0074-0276. Disponível em: < <http://dx.doi.org/> >.

TAYLOR, M. C. et al. *Trypanosoma cruzi* adenylyl cyclase is encoded by a complex multigene family. **Mol Biochem Parasitol**, v. 104, n. 2, p. 205-17, Nov 30 1999. ISSN 0166-6851 (Print)0166-6851. Disponível em: < <http://dx.doi.org/> >.

TEIXEIRA, A. R.; CALIXTO, M. A.; TEIXEIRA, M. L. Chagas' disease: carcinogenic activity of the antitrypanosomal nitroarenes in. **Mutat Res**, v. 305, n. 2, p. 189-96, Mar 1 1994. ISSN 0027-5107 (Print)0027-5107. Disponível em: < <http://dx.doi.org/> >.

TEIXEIRA, A. R. et al. Malignant, non-Hodgkin's lymphomas in *Trypanosoma cruzi*-infected rabbits treated. **J Comp Pathol**, v. 103, n. 1, p. 37-48, Jul 1990. ISSN 0021-9975 (Print)0021-9975. Disponível em: < <http://dx.doi.org/> >.

TEIXEIRA, D. E. et al. Interactive multimedia to teach the life cycle of *Trypanosoma cruzi*, the causative agent of Chagas disease. **PLoS Negl Trop Dis**, United States, v. 6, n. 8, p. e1749, 2012. ISSN 1935-2735 (Electronic)1935-2727 (Linking). Disponível em: < <http://dx.doi.org/10.1371/journal.pntd.0001749> >.

TIBAYRENC, M. Genetic epidemiology of parasitic protozoa and other infectious agents: the need for an integrated approach. **Int J Parasitol**, England, v. 28, n. 1, p. 85-104, 1998. ISSN 0020-7519 (Print)0020-7519 (Linking). Disponível em: < <http://dx.doi.org/> >.

TOTH, E. A.; YEATES, T. O. The structure of adenylosuccinate lyase, an enzyme with dual activity in the de novo purine biosynthetic pathway. **Structure**, v. 8, n. 2, p. 163-74, Feb 15 2000. ISSN 0969-2126 (Print)0969-2126. Disponível em: < <http://dx.doi.org/> >.

ULLMAN, B.; CARTER, D. Molecular and biochemical studies on the hypoxanthine-guanine phosphoribosyltransferases of the pathogenic haemoflagellates. **Int J Parasitol**, v. 27, n. 2, p. 203-13, Feb 1997. ISSN 0020-7519 (Print)0020-7519. Disponível em: < <http://dx.doi.org/> >.

VALADI, H. et al. Exosome-mediated transfer of mRNAs and microRNAs is a novel mechanism of genetic exchange between cells. **Nat Cell Biol**, v. 9, n. 6, p. 654-9, Jun 2007. ISSN 1465-7392 (Print)1465-7392. Disponível em: < <http://dx.doi.org/10.1038/ncb1596> >.

VALENTE, S. A. et al. Analysis of an acute Chagas disease outbreak in the Brazilian Amazon: human cases, triatomines, reservoir mammals and parasites. **Trans R Soc**

Trop Med Hyg, v. 103, n. 3, p. 291-7, Mar 2009. ISSN 0035-9203. Disponível em: < <http://dx.doi.org/10.1016/j.trstmh.2008.10.047> >.

VAN WEELDEN, S. W. et al. New functions for parts of the Krebs cycle in procyclic *Trypanosoma brucei*, a cycle not operating as a cycle. **J Biol Chem**, v. 280, n. 13, p. 12451-60, Apr 1 2005. ISSN 0021-9258 (Print)0021-9258. Disponível em: < <http://dx.doi.org/10.1074/jbc.M412447200> >.

VERMELHO, A. B. et al. Trypanosomatidae Peptidases: A Target for Drugs Development. **Current Enzyme Inhibition**, v. 3, n. 1, p. 19-48, February 2007 2007. Disponível em: < <http://www.ingentaconnect.com/content/ben/cei/2007/00000003/00000001/art00003> >.

WASINGER, V. C. et al. Progress with gene-product mapping of the Mollicutes: *Mycoplasma genitalium*. **Electrophoresis**, v. 16, n. 7, p. 1090-4, Jul 1995. ISSN 0173-0835 (Print)0173-0835. Disponível em: < <http://dx.doi.org/> >.

WEN, L. M. et al. PPB1, a putative spliced leader RNA gene transcription factor in *Trypanosoma cruzi*. **Mol Biochem Parasitol**, v. 110, n. 2, p. 207-21, Oct 2000. ISSN 0166-6851 (Print)0166-6851. Disponível em: < <http://dx.doi.org/> >.

WENDLER, P. et al. Structure and function of the AAA+ nucleotide binding pocket. **Biochim Biophys Acta**, v. 1823, n. 1, p. 2-14, Jan 2012. ISSN 0006-3002 (Print)0006-3002. Disponível em: < <http://dx.doi.org/10.1016/j.bbamcr.2011.06.014> >.

WHO. **Control of Chagas disease**. Geneva, p.1-109. 2002

_____. **Reporte sobre la enfermedad de Chagas**. p.1-96. 2007

_____. **Research Priorities for Chagas Disease, Human African Trypanosomiasis and Leishmaniasis**. 2012

_____. WHO | Chagas disease (American trypanosomiasis). WHO, 2013-04-12 16:32:06 2013. Disponível em: < <http://www.who.int/mediacentre/factsheets/fs340/en/> >.

WILKINS, M. R. et al. From proteins to proteomes: large scale protein identification by two-dimensional electrophoresis and amino acid analysis. **Biotechnology (N Y)**, v. 14, n. 1, p. 61-5, Jan 1996. ISSN 0733-222X (Print)0733-222x. Disponível em: < <http://dx.doi.org/> >.

WU, C. C.; MACCOSS, M. J. Shotgun proteomics: tools for the analysis of complex biological systems. **Curr Opin Mol Ther**, v. 4, n. 3, p. 242-50, Jun 2002. ISSN 1464-8431 (Print)1464-8431. Disponível em: < <http://dx.doi.org/> >.

YOSHIDA, N. Surface antigens of metacyclic trypomastigotes of *Trypanosoma cruzi*. **Infect Immun**, v. 40, n. 2, p. 836-9, May 1983. ISSN 0019-9567 (Print)0019-9567 (Linking). Disponível em: < <http://dx.doi.org/> >.

_____. Molecular basis of mammalian cell invasion by *Trypanosoma cruzi*. **An Acad Bras Cienc**, v. 78, n. 1, p. 87-111, Mar 2006. ISSN 0001-3765 (Print)0001-3765. Disponível em: < <http://dx.doi.org/S0001-37652006000100010> >.

_____. *Trypanosoma cruzi* infection by oral route: how the interplay between parasite and. **Parasitol Int**, v. 57, n. 2, p. 105-9, Jun 2008. ISSN 1383-5769 (Print)1383-5769. Disponível em: < <http://dx.doi.org/10.1016/j.parint.2007.12.008> >.

_____. Molecular mechanisms of *Trypanosoma cruzi* infection by oral route. **Mem Inst Oswaldo Cruz**, v. 104 Suppl 1, p. 101-7, Jul 2009. ISSN 1678-8060 (Electronic)0074-0276 (Linking). Disponível em: < <http://dx.doi.org/> >.

YOSHIDA, N. et al. Signal transduction induced in *Trypanosoma cruzi* metacyclic trypomastigotes during the invasion of mammalian cells. **Braz J Med Biol Res**, v. 33, n. 3, p. 269-78, Mar 2000. ISSN 0100-879X (Print)0100-879x. Disponível em: < <http://dx.doi.org/> >.

_____. Metacyclic neutralizing effect of monoclonal antibody 10D8 directed to the 35- and 50-kilodalton surface glycoconjugates of *Trypanosoma cruzi*. **Infect Immun**, v. 57, n. 6, p. 1663-7, Jun 1989. ISSN 0019-9567 (Print)0019-9567. Disponível em: < <http://dx.doi.org/> >.

ZHANG, J. et al. PEAKS DB: de novo sequencing assisted database search for sensitive and accurate peptide identification. **Mol Cell Proteomics**, v. 11, n. 4, p. M111.010587, Apr 2012. ISSN 1535-9476. Disponível em: < <http://dx.doi.org/10.1074/mcp.M111.010587> >.

ZINGALES, B. et al. A new consensus for *Trypanosoma cruzi* intraspecific nomenclature: second revision meeting recommends TcI to TcVI. **Mem Inst Oswaldo Cruz**, v. 104, n. 7, p. 1051-4, Nov 2009. ISSN 1678-8060 (Electronic)0074-0276 (Linking). Disponível em: < <http://dx.doi.org/> >.

_____. The revised *Trypanosoma cruzi* subspecific nomenclature: rationale, epidemiological relevance and research applications. **Infect Genet Evol**, v. 12, n. 2, p. 240-53, Mar 2012. ISSN 1567-1348. Disponível em: < <http://dx.doi.org/10.1016/j.meegid.2011.12.009> >.

APÊNDICE A. Tabelas suplementares da busca 1

Tabela suplementar 1. Proteínas putativas preditas para secreção por via clássica pelo programa SignalP.

| Descrição ^a | Código identificador ^b |
|--|---|
| <i>75 kDa invariant surface glycoprotein (isg75)</i> | Q57VX5_TRYB2 |
| <i>Acid phosphatase</i> | Q38AR1_TRYB2 |
| <i>Amino acid permease-like protein</i> | K2NAC7_TRYCR |
| <i>ATPase subunit 9</i> | Q4DDG2_TRYCC, C9ZS37_TRYB9 |
| <i>DNA repair and recombination protein</i> | G0UBP0_TRYVY |
| <i>Mucin TcMUCII</i> | Q4CMK8_TRYCC, Q4DNP7_TRYCC, Q4DYE7_TRYCC, Q4E1Z0_TRYCC, Q4D1D0_TRYCC |
| <i>Mucin-associated surface protein (MASP)</i> | K2MJD6_TRYCR, K2MQI2_TRYCR, Q4DT31_TRYCC, Q4D4H9_TRYCC, Q4DAB0_TRYCC, Q4E2A6_TRYCC, Q4E201_TRYCC, Q4DLR4_TRYCC, Q4D521_TRYCC, Q4CNE9_TRYCC, Q4CS52_TRYCC |
| <i>Mucin-like glycoprotein or protein</i> | K2M2H7_TRYCR, O15779_TRYCR |
| <i>p63-1 surface protease homolog</i> | G0UC84_TRYVY |
| <i>Piwi-like protein 1</i> | G0U5M0_TRYVY |
| <i>Proteophosphoglycan ppg4</i> | F9WQI2_TRYVY |
| <i>Receptor-type adenylate cyclase</i> | Q57WZ8_TRYB2, F9WRK6_TRYVY, G0TUP8_TRYVY, F9WUQ7_TRYVY |
| <i>Retrotransposon hot spot (RHS) protein</i> | Q4CPY3_TRYCC |
| <i>Serine carboxypeptidase (CBP1)</i> | G0U597_TRYVY |
| <i>Surface glycoprotein</i> | Q4W4T2_TRYRA |
| <i>T. brucei spp.-specific protein</i> | C9ZY57_TRYB9 |
| <i>Trans-sialidase (TS)</i> | K2MQE5_TRYCR, K2M785_TRYCR, K2M735_TRYCR, K4DVQ9_TRYCR |
| <i>tRNA pseudouridine synthase A-like protein</i> | K2NR90_TRYCR |
| <i>Variant surface glycoprotein (VSG)</i> | Q4FKU4_TRYB2, Q57XG6_TRYB2, Q4FKF4_TRYB2, Q57X36_TRYB2, Q387P1_TRYB2, Q380W7_TRYB2, Q4FKR0_TRYB2, Q380X9_TRYB2, Q38G44_TRYB2, Q38G45_TRYB2, Q380X5_TRYB2, Q571X1_TRYBR, Q571V7_TRYBB, A5JS52_9TRYP, Q57WM4_9TRYP, Q9GYV4_9TRYP, A5JS53_9TRYP, C9ZQ05_TRYB9, C9ZJV0_TRYB9, F9WAF3_TRYCI, F9WHB2_TRYCI, F9WJD3_TRYCI, F9WFI6_TRYCI, F9WJ0_TRYCI, F9W755_TRYCI, F9WAT0_TRYCI, F9W7A5_TRYCI, F9WUD5_TRYVY, Q26979_TRYCO |

^a Descrição da proteína anotada proveniente do Uniprot

^b Nome de entrada ou ID do Uniprot

Tabela suplementar 2. Proteínas putativas preditas para secreção por via não clássica pelo programa SecretomeP. (continua)

| Descrição ^a | Código identificador ^b |
|---|--|
| <i>3,2-trans-enoyl-CoA isomerase</i> | G0TTQ2_TRYVY |
| <i>33 kDa inner dynein arm light chain</i> | K2LYG6_TRYCR, Q4D1F5_TRYCC, G0UXV3_TRYCI |
| <i>50S ribosomal protein L7Ae</i> | Q4DTK4_TRYCC |
| <i>AAA ATPase</i> | G0V0C1_TRYCI |
| <i>ABC transporter</i> | G0UZS8_TRYCI, G0U9M4_TRYVY, G0UAK0_TRYVY |
| <i>Acyl-CoA dehydrogenase, mitochondrial</i> | G0UR95_TRYCI |
| <i>Adenylosuccinate lyase</i> | G0U2G5_TRYVY |
| <i>ADP/ATP mitochondrial carrier protein</i> | G0TYL3_TRYVY |
| <i>Amastin</i> | K4E706_TRYCR |
| <i>Amino acid transporter</i> | Q57V81_TRYB2, K2MMV7_TRYCR, G0USE1_TRYCI, G0TZY3_TRYVY, G0U9A8_TRYVY |
| <i>Ankyrin repeat protein</i> | G0U8C3_TRYVY |
| <i>Arginase</i> | Q581Y0_TRYB2 |
| <i>ATP-dependent RNA helicase</i> | G0UB40_TRYVY |
| <i>ATP-NAD kinase-like protein</i> | G0TYY4_TRYVY |
| <i>Beta transducin-like protein</i> | G0UB29_TRYVY |
| <i>Calmodulin or calmodulin-like protein</i> | G0U2A2_TRYVY, K4E482_TRYCR |
| <i>Calpain-like cysteine peptidase</i> | Q4CV42_TRYCC |
| <i>Carrier protein</i> | K2MVR5_TRYCR |
| <i>Cation transporter</i> | G0UCB5_TRYVY |
| <i>Centrin</i> | Q4DBB8_TRYCC |
| <i>Chaperone protein DNAj</i> | G0UQA9_TRYCI, G0U416_TRYVY, G0TYT7_TRYVY |
| <i>Cholinephosphate cytidyltransferase A</i> | G0UY44_TRYCI |
| <i>Coatomer subunit gamma</i> | G0UD96_TRYVY |
| <i>Complement regulatory protein</i> | K4DM44_TRYCR |
| <i>Conserved SH3 domain protein</i> | G0TX63_TRYVY |
| <i>Conserved zinc-finger protein</i> | G0TX23_TRYVY |
| <i>Cop-coated vesicle membrane protein</i> | G0U938_TRYVY |
| <i>Cop9 signalosome complex subunit</i> | G0U761_TRYVY |
| <i>CYC2-like cyclin</i> | G0U860_TRYVY |
| <i>Cytosolic leucyl aminopeptidase</i> | G0URS4_TRYCI |
| <i>Dihydrolipoamide dehydrogenase, point mutation</i> | G0TZQ1_TRYVY |
| <i>Dipeptidyl-peptidase 8-like serine peptidase</i> | G0U6Y5_TRYVY |
| <i>Diphthine synthase</i> | G0TUV9_TRYVY |
| <i>Dispersed gene family protein 1 (DGF-1)</i> | Q4CKE7_TRYCC, K4DV97_TRYCR, K4DJ35_TRYCR, K4EC51_TRYCR, K4DTT1_TRYCR, K4DMB4_TRYCR |

Tabela suplementar 2. Proteínas putativas preditas para secreção por via não clássica pelo programa SecretomeP. (continuação)

| Descrição ^a | Código identificador ^b |
|---|-----------------------------------|
| <i>DNA repair and recombination helicase protein PIF1</i> | G0UR10_TRYCI |
| <i>DNA repair helicase and transcription factor protein</i> | G0U9E8_TRYVY |
| <i>DNA-directed RNA polymerase II</i> | G0UY95_TRYCI |
| <i>Dual specificity protein phosphatase</i> | G0TZG6_TRYVY |
| <i>Dynein intermediate chain</i> | G0TRW8_TRYVY |
| <i>Dynein light chain</i> | G0U890_TRYVY |
| <i>Electron transfer protein</i> | G0TXN5_TRYVY |
| <i>Elongation of very long chain fatty acids protein</i> | G0TV10_TRYVY |
| <i>Endoplasmic reticulum oxidoreductin</i> | G0U171_TRYVY |
| <i>Endoribonuclease</i> | G0U9Q7_TRYVY |
| <i>Endosomal integral membrane protein</i> | K2MRS2_TRYCR |
| <i>Endosomal trafficking protein RME-8</i> | G0TX19_TRYVY |
| <i>Ethanolamine phosphotransferase</i> | K2NNJ7_TRYCR |
| <i>Eukaryotic translation initiation factor 3 subunit G</i> | G0TU72_TRYVY |
| <i>Exportin 1</i> | G0U8U7_TRYVY |
| <i>Farnesyltransferase</i> | G0TZN0_TRYVY |
| <i>FG-GAP repeat protein</i> | G0TZB7_TRYVY |
| <i>Flagellum-adhesion glycoprotein</i> | G0U0Y0_TRYVY |
| <i>Formin A</i> | Q4QE97_LEIMA |
| <i>Glyceraldehyde-3-phosphate dehydrogenase</i> | G0U2V4_TRYVY |
| <i>Glycerol-3-phosphate dehydrogenase</i> | G0UIV0_TRYCI, G0TR30_TRYVY |
| <i>Glycosyl transferase-like protein</i> | G0UCN3_TRYVY |
| <i>Glycosylphosphatidylinositol (GPI) anchor</i> | G0TU70_TRYVY |
| <i>Glycosyltransferase family 28 protein</i> | G0TWL3_TRYVY |
| <i>GTP binding protein</i> | G0UJE3_TRYCI, G0UL50_TRYCI |
| <i>Guanine nucleotide-binding beta subunit-like protein</i> | Q385S8_TRYB2 |
| <i>Helicase</i> | C9ZQA9_TRYB9 |
| <i>Histone deacetylase 2</i> | G0U975_TRYVY |
| <i>Hypoxanthine-guanine phosphoribosyltransferase</i> | G0U5D5_TRYVY |
| <i>Inositol polyphosphate 1-phosphatase</i> | G0TZN1_TRYVY |
| <i>Iron-containing hydrogenase</i> | D0A4K1_TRYB9 |

Tabela suplementar 2. Proteínas putativas preditas para secreção por via não clássica pelo programa SecretomeP. (continuação)

| Descrição ^a | Código identificador ^b |
|--|--|
| <i>Kinesin</i> | Q4DI76_TRYCC, K4DWY3_TRYCR |
| <i>Kinetoplast DNA-associated protein</i> | Q381E6_TRYB2 |
| <i>Kynureninase</i> | G0UTE1_TRYCI |
| <i>Lactoylglutathione lyase-like protein</i> | Q4D7B4_TRYCC |
| <i>Leucine-rich repeat protein (LRRP)</i> | G0UYI7_TRYCI, G0UVT3_TRYCI, G0TZH4_TRYVY, F9WTW8_TRYVY |
| <i>Lipin</i> | G0UQH2_TRYCI |
| <i>Major facilitator superfamily protein (MFS)</i> | G0UBV1_TRYVY, G0UYF0_TRYCI |
| <i>Major surface protease</i> | M1FV55_TRYRA |
| <i>Mannosyltransferase</i> | G0TX00_TRYVY |
| <i>Mercaptopyruvate sulfurtransferase</i> | G0UQL9_TRYCI |
| <i>Mitochondrial carrier protein</i> | G0TVC9_TRYVY, G0U0R9_TRYVY |
| <i>Mitochondrial elongation factor G</i> | G0U6F2_TRYVY |
| <i>MP81</i> | Q6T429_TRYCR |
| <i>Mucin TcMUCII</i> | K4DSJ3_TRYCR |
| <i>Mucin-associated surface protein (MASP)</i> | K2M445_TRYCR, K4DWD2_TRYCR, K4DIP3_TRYCR, K4E839_TRYCR |
| <i>Mucin-like glycoprotein</i> | K2MVL0_TRYCR |
| <i>Multidrug resistance protein A</i> | G0U0F8_TRYVY |
| <i>NAD-dependent deacetylase</i> | Q4CNV0_TRYCC |
| <i>Nucleic acid binding protein</i> | K2N9Z3_TRYCR |
| <i>Nucleolar RNA-binding protein</i> | Q4DHV9_TRYCC |
| <i>Nucleoside phosphorylase</i> | G0US35_TRYCI, G0U124_TRYVY |
| <i>Peptidyl-prolyl cis-trans isomerase</i> | G0TY50_TRYVY |
| <i>Phosphatidic acid phosphatase protein</i> | C9ZZW0_TRYB9 |
| <i>Phosphatidylinositol 3-related kinase</i> | Q4DJT9_TRYCC, K4DNM3_TRYCR |
| <i>Phosphatidylserine decarboxylase</i> | Q38DZ5_TRYB2 |
| <i>Phosphoinositide-specific phospholipase C</i> | G0UBF3_TRYVY |
| <i>Phospholipid transporting ATPase-like protein</i> | G0U8H6_TRYVY |
| <i>Phosphoribosylpyrophosphate synthetase</i> | Q57ZX2_TRYB2 |
| <i>Phosphotransferase</i> | Q385M0_TRYB2 |
| <i>Polynucleotide kinase 3'-phosphatase</i> | G0TWG9_TRYVY |
| <i>Predicted zinc finger protein</i> | G0U2T1_TRYVY |
| <i>Procylic form surface glycoprotein</i> | G0U835_TRYVY |
| <i>Protein kinase</i> | K2MZG2_TRYCR, F9WN97_TRYVY |
| <i>Protein phosphatase methylesterase 1</i> | G0U8T5_TRYVY |

Tabela suplementar 2. Proteínas putativas preditas para secreção por via não clássica pelo programa SecretomeP. (continuação)

| Descrição ^a | Código identificador ^b |
|--|--|
| <i>Pseudouridylate synthase</i> | G0UT16_TRYCI |
| <i>Ras-related GTP-binding protein</i> | G0U8H2_TRYVY |
| <i>Receptor-type adenylate cyclase GRESAG 4</i> | F9WMJ7_TRYVY |
| <i>Regulator of chromosome condensation</i> | G0TXR8_TRYVY |
| <i>Retrotransposon hot spot (RHS) protein</i> | K2NLF6_TRYCR, F9WVV7_TRYVY |
| <i>Rhomboid-like protein</i> | G0U0B5_TRYVY |
| <i>RNA polymerase B subunit RPB8</i> | G0UC34_TRYVY |
| <i>RNA-binding protein</i> | G0U842_TRYVY, G0U0Y9_TRYVY |
| <i>RNasePH-like protein</i> | G0UJ16_TRYCI |
| <i>Serine/threonine protein kinase</i> | G0UVI7_TRYCI, G0URC1_TRYCI, K4E2G5_TRYCR |
| <i>Small nuclear ribonucleoprotein SmD3</i> | Q4DX43_TRYCC |
| <i>SNF2 DNA repair protein</i> | G0UBS6_TRYVY |
| <i>Squalene monooxygenase</i> | G0UAL6_TRYVY |
| <i>Surface protease gp63</i> | K2M384_TRYCR, K4DJ96_TRYCR |
| <i>Syntaxin</i> | Q388D5_TRYB2, Q8S4W2_PHYIN |
| <i>T. brucei spp.-specific protein</i> | D0A7Y2_TRYB9, C9ZZQ6_TRYB9, D0AA20_TRYB9 |
| <i>T. congolense-specific, putative cell surface-expressed gene family</i> | F9W4N7_TRYCI |
| <i>Target of rapamycin</i> | K4EBX0_TRYCR |
| <i>Thymidine kinase</i> | F9WAY4_TRYCI |
| <i>TPR-repeat-containing chaperone protein</i> | G0U6G7_TRYVY |
| <i>DNAJ</i> | |
| <i>Trans-sialidase (TS)</i> | K2NWD0_TRYCR, K2MYQ4_TRYCR, K2NJQ9_TRYCR, K2LU45_TRYCR, K2NBE9_TRYCR, K2M324_TRYCR, K2NN88_TRYCR, K2NM96_TRYCR, K2N3M2_TRYCR, K2PEX6_TRYCR, K2MW42_TRYCR, K2MWI6_TRYCR, K2MHH2_TRYCR, K2N4I7_TRYCR, Q4DR51_TRYCC, Q4DSY7_TRYCC, Q4E222_TRYCC, Q4D349_TRYCC, Q4DAK1_TRYCC, Q4DGS3_TRYCC, Q4CKL4_TRYCC, Q4DYT7_TRYCC, K4DQI5_TRYCR, K4DRU0_TRYCR, K4EAY0_TRYCR, K4DSK4_TRYCR, K4DS27_TRYCR, K4DJC1_TRYCR, K4E9H7_TRYCR, K4ECF8_TRYCR, K4DWC7_TRYCR |
| <i>tRNA-methyl transferase</i> | G0U0B6_TRYVY |
| <i>Tubulin-tyrosine ligase-like protein</i> | G0UBN5_TRYVY |
| <i>Ubiquitin</i> | Q4CNV2_TRYCC |
| <i>Ubiquitin activating enzyme</i> | G0TRW6_TRYVY, G0U237_TRYVY |

Tabela suplementar 2. Proteínas putativas preditas para secreção por via não clássica pelo programa SecretomeP. (conclusão)

| Descrição ^a | Código identificador ^b |
|--|--|
| <i>Ubiquitin carboxyl-terminal hydrolase</i> | K2NP27_TRYCR |
| <i>Ubiquitin-conjugating enzyme-like</i> | Q4D439_TRYCC |
| <i>Ubiquitin-protein ligase</i> | G0U7Q4_TRYVY |
| <i>UDP-galactose transporter</i> | K2N909_TRYCR |
| <i>UDP-glucuronic acid/UDP-N-acetylgalactosamine transporter</i> | K4E8J5_TRYCR |
| <i>UDP-sugar pyrophosphorylase</i> | Q4QE75_LEIMA |
| <i>Vacuolar protein sorting complex subunit</i> | G0UY66_TRYCI |
| <i>Vacuolar proton-ATPase-like protein</i> | Q4DY50_TRYCC |
| <i>Variable surface glycoprotein</i> | Q968L1_TRYEV |
| <i>Variant surface glycoprotein (VSG)</i> | G1CRM2_9TRYP, Q380Y0_TRYB2, Q38G27_TRYB2, Q380U6_TRYB2, Q4FKR1_TRYB2, Q4FKK6_TRYB2, Q4FKS4_TRYB2, Q4FKL1_TRYB2, Q4FKC8_TRYB2, Q4FKL6_TRYB2, B3GVR3_TRYBB, C9ZKT0_TRYB9, F9WFY6_TRYCI, F9WA69_TRYCI, F9WCQ2_TRYCI, F9WDK7_TRYCI, F9W865_TRYCI, F9WD08_TRYCI, F9WFL2_TRYCI, F9WD34_TRYCI, F9WJB0_TRYCI, F9WR09_TRYVY, F9WQ58_TRYVY |
| <i>Zinc carboxypeptidase</i> | G0UCS1_TRYVY, G0TU56_TRYVY |

^a Descrição da proteína anotada proveniente do Uniprot

^b Nome de entrada ou ID do Uniprot

APÊNDICE B. Tabelas suplementares da busca 2

Tabela suplementar 3. Proteínas putativas preditas para secreção por via clássica.

| Descrição ^a | Código identificador ^b |
|---|---|
| <i>2-oxoglutarate dehydrogenase E1 component</i> | Q4CTE8, Q4DKY2 |
| <i>3,2-trans-enoyl-CoA isomerase</i> | Q4CRI1 |
| <i>60S ribosomal protein</i> | Q4E2Y1, Q4D531 |
| <i>Acyl-CoA dehydrogenase</i> | Q4DLR8, Q4DF20 |
| <i>Aspartate aminotransferase</i> | Q4D1Q4 |
| <i>ATP synthase</i> | Q4CNR6 |
| <i>ATP-dependent Clp protease subunit</i> | Q4DTH5 |
| <i>Calreticulin</i> | Q4CPZ0, Q4DDX3 |
| <i>COP-coated vesicle membrane protein erv25</i> | Q4E4C4 |
| <i>Cytochrome c oxidase assembly factor</i> | Q4CUN8 |
| <i>Cytochrome-B5 reductase</i> | Q4CM25 |
| <i>Dihydrolipoamide acetyltransferase</i> | Q4DYI5, Q4DZT8 |
| <i>Dispersed gene family protein 1 (DGF-1)</i> | Q4E3G7, Q4CWM9, Q4DXG7 |
| <i>DnaJ chaperone protein</i> | Q4DMC3 |
| <i>Dynein light chain</i> | Q4CZW3 |
| <i>Electron-transfer-flavoprotein</i> | Q4DCB8 |
| <i>Glucose-regulated protein 78</i> | Q4D620 |
| <i>Glutaredoxin-like protein</i> | Q4D0D1 |
| <i>Heat shock protein</i> | Q4D0L7, Q4E2Q2 |
| <i>Kinesin</i> | Q4DIG0 |
| <i>Lectin</i> | Q4D030 |
| <i>Lipophosphoglycan biosynthetic protein</i> | Q4DW89 |
| <i>Mitochondrial DNA polymerase I protein B</i> | Q4DSW2 |
| <i>NADH-dependent fumarate reductase</i> | Q4CYX1 |
| <i>p22 protein</i> | Q4E0T1 |
| <i>Peptidyl-prolyl cis-trans isomerase</i> | Q4CXV1, Q4D932 |
| <i>Pitriysin-like metalloprotease</i> | Q4DU56 |
| <i>Protein kinase-like protein</i> | Q4DQP8 |
| <i>Protein transport protein sec31</i> | Q4E0P0 |
| <i>Pyruvate dehydrogenase E1 beta subunit</i> | Q4DB91 |
| <i>Retrotransposon hot spot (RHS) protein</i> | Q4DKM0, Q4CYY4, Q4CXF4, Q4D1V4, Q4DZC9 |
| <i>RNA-binding protein</i> | Q4DY30 |
| <i>Serine/threonine protein kinase</i> | Q4E2L0 |
| <i>Signal recognition particle protein</i> | Q4D940 |
| <i>Translation initiation factor IF-2</i> | Q4DX02 |
| <i>Ubiquitin-conjugating enzyme E2</i> | Q4CLZ1 |
| <i>Urocanate hydratase</i> | Q4D9S6 |
| <i>Vacuolar-type proton translocating pyrophosphatase 1</i> | Q4DGG5 |

^a Descrição da proteína anotada proveniente do Uniprot

^b Número de acesso do Uniprot

Tabela suplementar 4. Proteínas putativas preditas para secreção por via não clássica. (continua)

| Descrição ^a | Código identificador ^b |
|--|--|
| <i>1-acyl-sn-glycerol-3-phosphate acyltransferase</i> | Q4DTE0 |
| <i>10 kDa heat shock protein</i> | Q4DFA8 |
| <i>2-hydroxy-3-oxopropionate reductase</i> | Q4DFE2 |
| <i>2-oxoglutarate dehydrogenase</i> | Q4D1R9 |
| <i>25 kDa translation elongation factor 1-beta</i> | Q4CQA2 |
| <i>26S proteasome regulatory non-ATPase subunit</i> | Q4CND5 |
| <i>3-ketoacyl-CoA thiolase</i> | Q4D7M6 |
| <i>3-oxo-5-alpha-steroid 4-dehydrogenase</i> | Q4CM68 |
| <i>3,2-trans-enoyl-CoA isomerase</i> | Q4CTQ3 |
| <i>40S ribosomal protein</i> | Q4CTR5, Q4CRH9, Q4D6I5, Q4DK39, Q4E093, Q4DFU6, Q4DTQ1, Q4DW38, Q4D5P4, Q4E0Q3, Q4DSU0, Q4CQ63 |
| <i>60S ribosomal protein</i> | Q4DD50, Q4D3B0, Q4E661, Q4D7A0, Q4CRX5, Q4DZP2, Q4DKZ4, Q4E2W6, Q4E0P2, Q4DST7, Q4DN72, Q4CQW2, Q4CPG4, Q4CU61, Q4DIV9, Q4DNZ7, Q4CRI7 |
| <i>69 kDa paraflagellar rod protein</i> | Q4DIF6 |
| <i>ABC transporter</i> | Q4CX07, Q4DAY6, Q4DWA3 |
| <i>Acetyl-CoA synthetase</i> | Q4DYN2 |
| <i>Acetylornithine deacetylase-like</i> | Q4D7V2 |
| <i>Acetyltransferase</i> | Q4DEI2 |
| <i>Activated protein kinase C receptor</i> | Q4DTN2 |
| <i>Acyl-CoA dehydrogenase</i> | Q4DK44 |
| <i>Acyltransferase</i> | Q4E308 |
| <i>Adenine phosphoribosyltransferase</i> | Q4DNZ3, Q4DNZ4 |
| <i>Adenosine kinase</i> | Q4CVK9 |
| <i>Adenylate kinase</i> | Q4D818 |
| <i>Adenylosuccinate synthetase 2</i> | Q4CWX1 |
| <i>ADP-ribosylation factor GTPase activating protein 1</i> | Q4CTU1 |
| <i>ADP, ATP carrier protein 1</i> | Q4DD13 |
| <i>Alcohol dehydrogenase</i> | Q4DY23 |
| <i>Alpha tubulin</i> | Q4CLA1 |
| <i>Aminopeptidase</i> | Q4D5P1 |
| <i>Asparaginyl-tRNA synthetase</i> | Q4E595 |
| <i>ATP-dependent Clp protease subunit</i> | Q4CLP1, Q4CRW0 |
| <i>ATP-dependent RNA helicase</i> | Q4DRK9, Q4DH96, Q4D640 |
| <i>ATP-dependent zinc metallopeptidase</i> | Q4DJU4 |

Tabela suplementar 4. Proteínas putativas preditas para secreção por via não clássica. (continuação)

| Descrição ^a | Código identificador ^b |
|--|---|
| <i>Basic transcription factor 3a</i> | Q4CYN3 |
| <i>Beta prime C</i> | Q4DAM1 |
| <i>Beta tubulin</i> | Q4DQP2 |
| <i>Biotin/lipoate protein ligase</i> | Q4CPE2 |
| <i>Calcium channel protein</i> | Q4DZ99 |
| <i>Calcium motive p-type ATPase</i> | Q4DU38 |
| <i>Calcium-translocating P-type ATPase</i> | Q4DIM7 |
| <i>Calpain cysteine peptidase</i> | Q4DQN6 |
| <i>Calpain-like cysteine peptidase</i> | Q4CYU3, Q4CXJ8, Q4DQN5, Q4E0D8 |
| <i>Carnitine/choline acetyltransferase</i> | Q4DKK5 |
| <i>Casein kinase</i> | Q4CSP0 |
| <i>Cation-transporting ATPase</i> | Q4DK98 |
| <i>Centrin</i> | Q4CTH2 |
| <i>Chaperonin</i> | Q4E151 |
| <i>Citrate synthase</i> | Q4CVJ5 |
| <i>Co-chaperone GrpE</i> | Q4CPP0 |
| <i>Coatomer beta subunit</i> | Q4CW60 |
| <i>Coatomer epsilon subunit</i> | Q4CP08 |
| <i>Cofilin/actin depolymerizing factor</i> | Q4CVE9 |
| <i>Copper homeostasis protein</i> | Q4CRX3 |
| <i>Cystathione gamma lyase</i> | Q4DUG8 |
| <i>Cytochrome b-domain protein</i> | Q4DJQ7 |
| <i>Cytochrome c</i> | Q4CV48 |
| <i>Cytochrome c oxidase subunit IV</i> | Q4E2P2 |
| <i>Cytochrome c oxidase subunit V</i> | Q4D214 |
| <i>Cytoskeleton-associated protein CAP5.5</i> | Q4CP00, Q4DJS7 |
| <i>Cytosolic leucyl aminopeptidase</i> | Q4D4Y0 |
| <i>D-isomer specific 2-hydroxyacid dehydrogenase-protein</i> | Q4DAE1 |
| <i>Dipeptidyl-peptidase 8-like serine peptidase</i> | Q4CX66 |
| <i>Dispersed gene family protein 1 (DGF-1)</i> | Q4CZM4, Q4CQN2, Q4CQT3, Q4CSQ4, Q4DSY5, Q4DH83 |
| <i>DNAK protein</i> | Q4DGL0 |
| <i>Dynein arm light chain</i> | Q4DLN1 |
| <i>Dynein heavy chain</i> | Q4CRR4 |
| <i>Dynein intermediate chain</i> | Q4D4E6 |
| <i>Electron transfer flavoprotein</i> | Q4DR32 |
| <i>Elongation factor 1-gamma (EF-1-gamma)</i> | Q4CXW2 |

Tabela suplementar 4. Proteínas putativas preditas para secreção por via não clássica.
(continuação)

| Descrição ^a | Código identificador ^b |
|---|-----------------------------------|
| <i>Endo-beta-N-acetylglucosaminidase</i> | Q4DJ21 |
| <i>Endosomal trafficking protein RME-8</i> | Q4DT81 |
| <i>Enoyl-CoA hydratase</i> | Q4CKX1, Q4DMX9 |
| <i>Eukaryotic initiation factor 5a</i> | Q4E4N4 |
| <i>Eukaryotic release factor 3</i> | Q4CXW9 |
| <i>Eukaryotic translation initiation factor eIF-4E</i> | Q4D511 |
| <i>Farnesyl pyrophosphate synthase</i> | Q4DFM3 |
| <i>Fatty acid desaturase</i> | Q4CKQ6 |
| <i>Fatty acid elongase</i> | Q4DUK6 |
| <i>Fibrillarin</i> | Q4DZU7 |
| <i>Formin C</i> | Q4D3C2 |
| <i>Gim5A protein</i> | Q4CR48 |
| <i>Glucosamine-6-phosphate isomerase</i> | Q4DD39 |
| <i>Glucosamine-fructose-6-phosphate aminotransferase</i> | Q4DZ02 |
| <i>Glucose transporter</i> | Q4CKT5 |
| <i>Glutamate dehydrogenase</i> | Q4D5C2 |
| <i>Glutamine-dependent carbamoyl-phosphate synthetase</i> | Q4D6X8 |
| <i>Glutathione peroxidase</i> | Q4DEJ5 |
| <i>Glutathione-S-transferase/glutaredoxin</i> | Q4D776 |
| <i>Glyceraldehyde 3-phosphate dehydrogenase</i> | Q4DCN9, Q4DZT1 |
| <i>Glycosomal malate dehydrogenase</i> | Q4DDK8, Q4DRD8 |
| <i>Glycosomal membrane protein</i> | Q4CYY9, Q4D418 |
| <i>GMP synthase</i> | Q4E1D4 |
| <i>GTP-binding nuclear protein rtb2</i> | Q4DIB9 |
| <i>Guanine deaminase</i> | Q4D711 |
| <i>Guanylate kinase</i> | Q4E4D6 |
| <i>Heat shock 70 kDa protein</i> | Q4DXG9, Q4CVR9, Q4DTM8, Q4DTM9 |
| <i>Heat shock protein DnaJ</i> | Q4E4K5, Q4D832 |
| <i>Helicase</i> | Q4CSN6 |
| <i>High mobility group protein</i> | Q4D714 |
| <i>Histone H2B</i> | Q4CTD7 |
| <i>Histone H3</i> | Q4CWE9 |
| <i>Homoserine kinase</i> | Q4CME2 |
| <i>Hslvu complex proteolytic subunit-like</i> | Q4D729 |
| <i>I/6 autoantigen</i> | Q4DFL2 |
| <i>IgE-dependent histamine-releasing factor</i> | Q4CW52 |
| <i>Immunodominant antigen</i> | Q4D528 |
| <i>Importin beta-1 subunit</i> | Q4DZ97 |

Tabela suplementar 4. Proteínas putativas preditas para secreção por via não clássica.
(continuação)

| Descrição ^a | Código identificador ^b |
|---|---|
| <i>Isoform 2 of Proline racemase A</i> | Q4DA80-2 |
| <i>Katanin</i> | Q4DLN7 |
| <i>Kinesin</i> | Q4DRD6, Q4DDL0, Q4E3N3, Q4D0K8, Q4E0R3 |
| <i>Kinetoplast DNA-associated protein</i> | Q4CXF1, Q4CN27 |
| <i>Kinetoplastid membrane protein KMP-11 (Fragment)</i> | Q4D3A5 |
| <i>LA RNA binding protein</i> | Q4DYI6 |
| <i>Lanosterol synthase</i> | Q4E2Y7 |
| <i>Leucyl-tRNA synthetase</i> | Q4CL52 |
| <i>Lipophosphoglycan biosynthetic protein</i> | Q4D253 |
| <i>Lysophospholipase</i> | Q4DA08 |
| <i>Malate dehydrogenase</i> | Q4DXL5, Q4CTR7 |
| <i>MCAK-like kinesin</i> | Q4DFR8 |
| <i>Methionyl-tRNA synthetase</i> | Q4D6H2 |
| <i>Mevalonate kinase</i> | Q4DJT8 |
| <i>Microtubule-associated protein</i> | Q4CMT2 |
| <i>Mitochondrial carrier protein</i> | Q4D8C1 |
| <i>Mitochondrial oligo_U binding protein TBRGG1</i> | Q4D590 |
| <i>Mitochondrial phosphate transporter</i> | Q4DJJ9 |
| <i>Mitochondrial processing peptidase</i> | Q4DPR3 |
| <i>Mitochondrial RNA binding protein 1</i> | Q4D0R4 |
| <i>Mitochondrial RNA-binding protein 2</i> | Q4E544 |
| <i>Myo-inositol-1-phosphate synthase</i> | Q4DFC2 |
| <i>Myosin heavy chain</i> | Q4CNB2 |
| <i>NADH-dependent fumarate reductase</i> | Q4D779 |
| <i>Nascent polypeptide associated complex subunit</i> | Q4D9L7 |
| <i>Nitrilase</i> | Q4D9K8 |
| <i>Nuclear pore complex protein (NUP155)</i> | Q4E149 |
| <i>Nucleobase transporter</i> | Q4DHG0 |
| <i>Nucleoside phosphorylase</i> | Q4DKS7, Q4DAD9 |
| <i>NUP-1 protein</i> | Q4DB82 |
| <i>P-type H⁺-ATPase</i> | Q4CQ52, Q4CRY5, Q4DE37 |
| <i>p21 antigen protein</i> | Q4CSP8 |
| <i>Pantothenate kinase subunit</i> | Q4DHB2 |
| <i>Paraflagellar rod protein 3</i> | Q4D634 |
| <i>Peptidase M20/M25/M40</i> | Q4DGX6 |
| <i>Peptidase T</i> | Q4DPI6 |
| <i>Peptide methionine sulfoxide reductase</i> | Q4DK88 |

Tabela suplementar 4. Proteínas putativas preditas para secreção por via não clássica.
(continuação)

| Descrição ^a | Código identificador ^b |
|--|---|
| <i>Peptidyl-prolyl cis-trans isomerase</i> | Q4E4L9, Q4CU89 |
| <i>Peroxin-2</i> | Q4E2J8 |
| <i>Peroxisomal membrane protein 4</i> | Q4DBE4 |
| <i>Peroxisome targeting signal 1 receptor</i> | Q4DPQ9 |
| <i>Phosphoacetylglucosamine mutase</i> | Q4DA33 |
| <i>Phosphomannomutase</i> | Q4E4A3, Q4DEE6 |
| <i>Poly(A) polymerase</i> | Q4DA83 |
| <i>Poly(A)-binding protein</i> | Q4E4V0 |
| <i>Prohibitin</i> | Q4D8I7 |
| <i>Prolyl oligopeptidase</i> | Q4E132 |
| <i>Prostaglandin F synthase</i> | Q4DJ07 |
| <i>Prostaglandin F2alpha synthase</i> | Q4E4V7, Q4CNI1 |
| <i>Proteasome activator protein PA26</i> | Q4CPR5 |
| <i>Proteasome alpha 2 subunit</i> | Q4CKY8 |
| <i>Proteasome beta 3 subunit</i> | Q4DHA9 |
| <i>Proteasome regulatory ATPase subunit 3</i> | Q4DW08 |
| <i>Proteasome regulatory non-ATPase subunit</i> | Q4CSX5 |
| <i>Proteasome regulatory non-ATPase subunit 7</i> | Q4E268 |
| <i>Proteasome regulatory non-ATPase subunit 8</i> | Q4DH98 |
| <i>Proteasome subunit beta type</i> | Q4E0L9 |
| <i>Protein kinase</i> | Q4DMI7 |
| <i>Protein transport protein Sec24C</i> | Q4D7L4 |
| <i>Protein tyrosine phosphatase</i> | Q4D674, Q4CTD5 |
| <i>Pumilio/PUF RNA binding protein 2</i> | Q4DL20 |
| <i>Pumilio/PUF RNA binding protein 6</i> | Q4D0T3 |
| <i>Pyroline-5-carboxylate synthetase-like protein</i> | Q4CVK6 |
| <i>Pyruvate dehydrogenase E1 component alpha subunit</i> | Q4DL16 |
| <i>Quinone oxidoreductase</i> | Q4DTF4 |
| <i>Rab7 GTP binding protein</i> | Q4E4T4 |
| <i>Ras-related protein rab-5</i> | Q4D504 |
| <i>Retrotransposon hot spot (RHS) protein</i> | Q4DWU4 |
| <i>Ribose 5-phosphate isomerase</i> | Q4DBP9 |
| <i>Ribosomal P protein AGP2beta-1</i> | Q4CVQ5 |
| <i>Ribosomal protein</i> | Q4E5H8, Q4D604, Q4E581, Q4D6N9, Q4CUC8 |
| <i>RNA-binding protein</i> | Q4CXT3, Q4D3A9, Q4DCC7 |
| <i>RNA-editing complex protein</i> | Q4D022 |
| <i>Serine/threonine protein kinase</i> | Q4DDT6 |

Tabela suplementar 4. Proteínas putativas preditas para secreção por via não clássica.
(conclusão)

| Descrição ^a | Código identificador ^b |
|---|-----------------------------------|
| <i>Serine/threonine-protein phosphatase</i> | Q4DDM9 |
| <i>Short chain 3-hydroxyacyl-coa dehydrogenase</i> | Q4DMG1, Q4DZT0 |
| <i>Small GTP-binding protein Rab1</i> | Q4CZR0 |
| <i>Small Rab GTP binding protein</i> | Q4DSV2 |
| <i>Sterol 24-c-methyltransferase</i> | Q4CM63 |
| <i>Sterol C-24 reductase</i> | Q4DIV5 |
| <i>Succinyl-CoA synthetase alpha subunit</i> | Q4E2I6 |
| <i>Succinyl-CoA:3-ketoacid-coenzyme A transferase</i> | Q4D0L3 |
| <i>Superoxide dismutase</i> | Q4DCQ3, Q4DI29 |
| <i>Thioredoxin-like protein</i> | Q4DDZ3 |
| <i>Transketolase</i> | Q4DU82, Q4CW17 |
| <i>Tricarboxylate carrier</i> | Q4CVF7 |
| <i>Trifunctional enzyme alpha subunit</i> | Q4CWB6 |
| <i>Trypanothione reductase</i> | Q4CMQ7 |
| <i>Tryparedoxin</i> | Q4D1B8 |
| <i>Tryparedoxin peroxidase</i> | Q4CX87, Q4CM56 |
| <i>Tubulin-tyrosine ligase-like protein</i> | Q4DRI1 |
| <i>U-box domain protein</i> | Q4DW87, Q4D7U8 |
| <i>U-rich RNA-binding protein UBP-2</i> | Q4E1N4 |
| <i>Ubiquitin</i> | Q4CNV2 |
| <i>Ubiquitin carboxyl-terminal hydrolase</i> | Q4E680 |
| <i>Ubiquitin fusion degradation protein 2</i> | Q4E0T5 |
| <i>Ubiquitin hydrolase</i> | Q4DFD2 |
| <i>Ubiquitin ligase</i> | Q4D6V3, Q4DFS7 |
| <i>Ubiquitin-activating enzyme E1</i> | Q4DRK7, Q4D7C0 |
| <i>Uracil phosphoribosyltransferase</i> | Q4E301 |
| <i>UV excision repair RAD23-like protein</i> | Q4CYX4 |
| <i>Vacuolar-type Ca²⁺-ATPase</i> | Q4E3S6 |
| <i>Vesicle-associated membrane protein</i> | Q4DDN8, Q4DP28 |
| <i>Vesicle-fusing ATPase</i> | Q4D5R2 |
| <i>Vesicle-associated membrane protein</i> | Q4D0X1 |

^a Descrição da proteína anotada proveniente do Uniprot

^b Número de acesso do Uniprot

APÊNDICE C. Tabelas suplementares da busca 3

Tabela suplementar 5. Proteínas putativas preditas para secreção por via clássica.

(continua)

| Descrição ^a | Código identificador ^b |
|--|--|
| <i>10 kDa heat shock protein</i> | Q4D000 |
| <i>2-oxoglutarate dehydrogenase</i> | K4DRJ2, Q4CTE8, Q4DKY2 |
| <i>2-oxoisovalerate dehydrogenase</i> | K4EBU1, K4DW66 |
| <i>3,2-trans-enoyl-CoA isomerase</i> | K4E5W6 |
| <i>40S ribosomal protein</i> | Q4CP93, Q4D7C8 |
| <i>60S ribosomal protein</i> | Q4CSS8, Q4D7Y0, Q4E4R0, Q4E4R1, K4E9F6, Q4E2Y1, Q4DI49, Q4E3R2, Q4D531, Q4DJ32 |
| <i>Acyl-CoA dehydrogenase</i> | Q4DLR8, Q4DYZ3 |
| <i>ADP-ribosylation factor</i> | Q4DL77 |
| <i>Aspartate aminotransferase</i> | K4EA87, Q4D1Q4 |
| <i>ATP synthase subunit alpha</i> | K4E934 |
| <i>ATP-dependent Clp protease subunit</i> | K4EBZ4 |
| <i>Basic transcription factor 3a</i> | Q4D103 |
| <i>Beta galactofuranosyl glycosyltransferase</i> | Q4DIN0 |
| <i>Calreticulin</i> | K4E613 |
| <i>Carbonic anhydrase-like protein</i> | Q4CVY4 |
| <i>Cathepsin B-like protease</i> | O61066 |
| <i>COP-coated vesicle membrane protein</i> | K4EBJ8, Q4CU63, Q4DJZ3, K4E084 |
| <i>Copper homeostasis protein</i> | K4EAR4 |
| <i>Cruzipain</i> | Q4DW02 |
| <i>Cysteine peptidase</i> | Q4CMU6, Q4E0J7 |
| <i>Cysteine synthase</i> | K4E8E0 |
| <i>Cytochrome c oxidase assembly factor</i> | K4DQP9 |
| <i>Cytochrome-b5 reductase</i> | K4E3X5 |
| <i>Dihydrolipoamide acetyltransferase</i> | K4DVJ5, K4DZZ0, Q4DZT8 |
| <i>Dipeptidyl-peptidase</i> | K4E7R5 |
| <i>Dispersed gene family protein 1 (DGF-1)</i> | Q4CSJ3, Q4CWM9, Q4CTS2, Q4CSI0, Q4CNE4, Q4CMY8, Q4CMN7, Q4D259, Q4D5T8, Q4DVR7, Q4DQ05, Q4DJP4, Q4DJA7, Q4D8Y4, Q4DXG7, Q4D1G6, Q4D335, Q4DQA7 |
| <i>dTDP-glucose 4,6-dehydratase</i> | K4DRP0 |
| <i>Dynein light chain</i> | Q4CZW3, Q4DN14 |
| <i>Electron transfer flavoprotein</i> | K4E9M0, Q4DG52 |
| <i>Eukaryotic initiation factor 4a</i> | Q4E0F6 |
| <i>Fibrillarin</i> | Q4DGF1 |
| <i>Flagellar calcium-binding protein</i> | K4EAZ1 |

Tabela suplementar 5. Proteínas putativas preditas para secreção por via clássica.

(continuação)

| Descrição ^a | Código identificador ^b |
|---|--|
| <i>Flagellum-adhesion glycoprotein</i> | K4EDX3 |
| <i>Glucose-regulated protein 78</i> | K4DT97 |
| <i>Glycerol kinase</i> | K4DXG8 |
| <i>Glycosylphosphatidylinositol (GPI) anchor</i> | K4E4M2 |
| <i>Golgi/lysosome glycoprotein</i> | K4DLS7 |
| <i>Heat shock protein HsIVU - ATPase subunit HsIU</i> | K4E7U0, Q4DEP1 |
| <i>Heat shock protein</i> | K4DYB4, Q4E2Q2 |
| <i>Histone H3</i> | Q4CZH3 |
| <i>Inositol polyphosphate 1-phosphatase</i> | K4E7I0 |
| <i>Kinesin</i> | K4DTK2 |
| <i>Lanosterol 14-alpha-demethylase</i> | K4EAL9 |
| <i>Lectin</i> | Q4DHU0, Q4D030 |
| <i>Lipophosphoglycan biosynthetic protein</i> | Q4D7V0, Q4DW89 |
| <i>Lysosomal alpha-mannosidase</i> | K4E2E6 |
| <i>Lysosomal/endosomal membrane protein p67</i> | K4E8C9 |
| <i>Membrane-bound acid phosphatase</i> | K4E6J4 |
| <i>Mitochondrial DNA polymerase beta-PAK</i> | K4DK04, Q4E023 |
| <i>Mitochondrial DNA polymerase I protein B</i> | K4E8I0, Q4DSW2 |
| <i>Mitochondrial DNA polymerase I protein C</i> | Q4D7R3 |
| <i>Mitochondrial processing peptidase</i> | Q4DB63 |
| <i>NADH-dependent fumarate reductase</i> | K4DZP3, Q4CYX1 |
| <i>OSM3-like kinesin</i> | K4EB74 |
| <i>p22 protein</i> | Q4E0T1 |
| <i>Peptidyl-prolyl cis-trans isomerase</i> | Q4DPB9, Q4CXV1, Q4D7C3, Q4D932 |
| <i>Phosphoglycerate mutase</i> | K4DR72 |
| <i>Pitriylsin-like metalloprotease</i> | K4DLV5 |
| <i>Poly(ADP-ribose) polymerase</i> | K4DJL0 |
| <i>Protein disulfide isomerase</i> | K4E7C1, K4E1S3 |
| <i>Protein kinase</i> | K4E7I3 |
| <i>Protein transport protein Sec24C</i> | K4DYN7 |
| <i>Protein transport protein Sec31</i> | K4E9C7, Q4E0P0 |
| <i>Ras-related GTP-binding protein</i> | Q4E5H4 |
| <i>Reiske iron-sulfur protein</i> | K4EA56, Q4DS82 |
| <i>Reticulon-like protein</i> | K4E4U7 |
| <i>Retrotransposon hot spot (RHS) protein</i> | K4E7K3, K4E7J6, K4DSI4, Q4DZC9, Q4CQQ0, K4E9C1, K4DNJ0, K4DJ48, Q4DKM0, Q4D8B1, Q4D1V4, Q4CQB8, Q4E0I0, Q4CXF4, Q4CYY4, Q4D9N3, |

Tabela suplementar 5. Proteínas putativas preditas para secreção por via clássica.
(conclusão)

| Descrição ^a | Código identificador ^b |
|---|--|
| <i>Retrotransposon hot spot (RHS) protein (cont.)</i> | Q4CU66, Q4D1Y4, Q4D8R0, Q4DM59 |
| <i>Ribonucleoprotein p18</i> | Q4CZG5 |
| <i>RNA helicase</i> | K4E8G4 |
| <i>RNA-binding protein</i> | K4DKY5, K4DQW8, Q4DCC6 |
| <i>Serine carboxypeptidase (CBP1)</i> | K4DS29 |
| <i>Solaneyl diphosphate synthase</i> | D9YN59 |
| <i>Squalene monooxygenase</i> | K4E4U2 |
| <i>Succinate dehydrogenase</i> | K4E5H9 |
| <i>Thioredoxin</i> | K4E390, Q4DP22 |
| <i>Tob55</i> | K4DT74 |
| <i>Transcription modulator/accessory protein</i> | K4E9M8 |
| <i>Transport protein particle (TRAPP) subunit</i> | Q4D2W2 |
| <i>Ubiquitin-conjugating enzyme E2</i> | Q4CLZ1 |
| <i>UDP-Gal or UDP-GlcNAc-dependent glycosyltransferase</i> | K4E6V6, Q4D4N0, Q4CTY1, Q4CNJ4, Q4CPP7, Q4DR76, Q4CMM5, Q4CPI4, Q4CSG1 |
| <i>Urocanate hydratase</i> | Q4D9S6 |
| <i>Vacuolar-type proton translocating pyrophosphatase 1</i> | K4E6E7 |

^a Descrição da proteína anotada proveniente do Uniprot

^b Número de acesso do Uniprot

Tabela suplementar 6. Proteínas putativas preditas para secreção por via não clássica.
(continua)

| Descrição ^a | Código identificador ^b |
|---|--|
| <i>1,2-dihydroxy-3-keto-5-methylthiopentene dioxygenase</i> | Q4CVG3, K4DZ96 |
| <i>1-acyl-sn-glycerol-3-phosphate acyltransferase</i> | K4E3R6 |
| <i>10 kDa heat shock protein</i> | Q4DFB0, Q4DFA8 |
| <i>2-aminoethylphosphonate:pyruvate aminotransferase</i> | K4DWU9 |
| <i>2-hydroxy-3-oxopropionate reductase</i> | K4E5F1 |
| <i>2-oxoisovalerate dehydrogenase</i> | K4DNB1 |
| <i>25 kDa translation elongation factor 1-beta</i> | Q4CQA2, K4E311, Q4DG86 |
| <i>26S proteasome regulatory non-ATPase subunit</i> | K4DJN6 |
| <i>3,2-trans-enoyl-CoA isomerase</i> | K4DXD8 |
| <i>3-oxo-5-alpha-steroid 4-dehydrogenase</i> | Q4CM68, K4DWA1 |
| <i>40 kDa cyclophilin</i> | Q6V7K6, Q4E4G0 |
| <i>40S ribosomal protein</i> | Q4CTR5, Q4DQ35, Q4D0U7, Q4CRH9, Q4D4L4, Q4CWD6, Q4D6I5, Q4DGZ5, Q4E0N6, Q4CXX2, K4E7X0, Q4DK40, Q4DK39, Q4E093, Q4DFU6, Q4DFU5, Q4DTQ1, Q4DW38, Q4DN73, K4DU81, Q4D7C9, Q62594, Q4D5P4, Q4E0Q3, Q4E3P1, K4EBD4, Q4DU31, K4EAG9, K4E1A5, Q4CX08, Q4DW55, K4DQZ4, Q4CQU0 |
| <i>50S ribosomal protein</i> | Q4DTK4 |
| <i>6-phospho-1-fructokinase</i> | K4E1M0 |
| <i>6-phosphogluconolactonase</i> | K4DT66, Q4CQA8 |
| <i>60S ribosomal protein</i> | K4E0W5, Q4CVQ4, K4DV40, Q4DD50, Q4CNF4, Q4CQ81, Q4D3B0, Q4DHz9, K4E2A2, Q4E661, K4E2X9, Q4CVX1, Q4D7A0, Q4DRQ7, Q4CRX5, Q4DC23, Q4DZP2, Q4DHK0, Q4DKZ4, Q4DZ31, Q4E2W6, Q4DPQ8, Q4DG45, Q4E0P2, Q4D8K4, Q4DIC9, Q4D8K3, Q4D1H8, Q4DU35, Q4DST7, Q4DB26, Q4DN72, Q4DL18, K4E3A3, Q4CQW2, K4EB05, Q4CU61, Q4DIV9, Q4CP06, K4DNL7, Q4DNZ7, K4DMD5, Q4CVX9 |
| <i>69 kDa paraflagellar rod protein</i> | Q4DIF6 |
| <i>AAA ATPase</i> | K4DVN7 |
| <i>ABC transporter</i> | K4DUS3, K4E123, K4EDU2, K4DJH5, Q4DHQ7, Q4DWA3 |

Tabela suplementar 6. Proteínas putativas preditas para secreção por via não clássica.
(continuação)

| Descrição ^a | Código identificador ^b |
|--|--|
| <i>Acetyl-CoA synthetase</i> | K4EE89 |
| <i>Acetylornithine deacetylase-like</i> | K4DTG9, Q4D7V2 |
| <i>Acetyltransferase</i> | K4E1D4, Q4DEI2 |
| <i>Actin</i> | K4DQH2 |
| <i>Actyltransferase-like</i> | K4E8I2 |
| <i>Acyl-CoA dehydrogenase</i> | K4DYV9, K4E6E2, K4E922, Q4DHU6 |
| <i>Acyltransferase</i> | K4DUE3 |
| <i>Adenine phosphoribosyltransferase</i> | K4E774, K4E8H9, Q4DVM8, Q4DNZ3, Q4DVM9 |
| <i>Adenosine 5'-monophosphoramidase</i> | Q4DJ40 |
| <i>Adenosine kinase</i> | K4DTE0, K4E5U6 |
| <i>Adenylate kinase</i> | K4E1Y3, Q4DLV1, Q4DXL7 |
| <i>Adenylosuccinate lyase</i> | Q4CXJ7, Q4DBH5 |
| <i>Adenylosuccinate synthetase</i> | K4E5I9 |
| <i>ADP-ATP carrier protein 1</i> | Q4DD13, K4DX91 |
| <i>ADP-ribosylation factor GTPase activating protein 1</i> | K4E285 |
| <i>ADP/ATP mitochondrial carrier protein</i> | Q4DDY8 |
| <i>Alanine racemase</i> | K4EDS3 |
| <i>Aldehyde dehydrogenase</i> | Q4DCP7, K4DRH8 |
| <i>Aldose 1-epimerase</i> | K4EE94, Q4DQX3 |
| <i>Alpha tubulin</i> | Q4CLA1 |
| <i>Ama1 protein</i> | Q4D749 |
| <i>Amastin</i> | Q4CKT7, K4E706, Q4CVL1 |
| <i>Amino acid transporter</i> | K4E5V4, K4E5Y6 |
| <i>Aminoacylase</i> | K4DQT8, K4DQX8 |
| <i>Aminomethyltransferase</i> | K4DNF4, Q4D100 |
| <i>Aminopeptidase P</i> | K4EAQ0 |
| <i>Aminopeptidase</i> | Q4CXX8, Q4CPW3, K4EE31, K4DX11 |
| <i>Arginase</i> | K4DXN0 |
| <i>Ascorbate-dependent peroxidase</i> | K4E1Q7 |
| <i>AsparaginyI-tRNA synthetase</i> | Q4E595 |
| <i>ATP synthase</i> | K4DNM1, Q4E1T6, Q4DB72 |
| <i>ATP-dependent Clp protease subunit</i> | Q4CLP1 |
| <i>ATP-dependent DEAD/H RNA helicase</i> | K4E6M7, Q4D8X3 |
| <i>ATP-dependent RNA helicase</i> | K4E6U0, K4DXM8, K4DX70 |
| <i>ATP-dependent zinc metallopeptidase</i> | Q4DJU4, Q4DEY4, K4E2P5 |
| <i>Axoneme central apparatus protein</i> | K4DTD7, Q4CZ83 |
| <i>Basic transcription factor 3a</i> | Q4CYN3 |

Tabela suplementar 6. Proteínas putativas preditas para secreção por via não clássica.
(continuação)

| Descrição ^a | Código identificador ^b |
|---|--|
| <i>Beta galactofuranosyl glycosyltransferase</i> | Q4DFT4, Q4DCT9, Q4D9G4, Q4D109, Q4DYV9, Q4CYW4, Q4DWU8, Q4DTL8, Q4DEF5, Q4E0D2, Q4DKM5, Q4D6R6 |
| <i>Beta prime cop protein</i> | K4E2T6, Q4CV82, Q4DAM1 |
| <i>Beta tubulin</i> | Q4DQP2 |
| <i>Bifunctional dihydrofolate reductase-thymidylate synt.</i> | Q8T5T8 |
| <i>C-5 sterol desaturase</i> | K4E5Q3 |
| <i>Calcium channel protein</i> | K4E1H7, Q4DZ99 |
| <i>Calcium motive p-type ATPase</i> | K4EB30, Q4DU38 |
| <i>Calcium-translocating P-type ATPase</i> | K4DJX8 |
| <i>Calmodulin</i> | K4E482, K4EBN9 |
| <i>Calpain cysteine peptidase</i> | K4DUN8, Q4CMV9, Q4E0D8, Q4D066, Q4CW64, K4DU84, K4DTH5, K4DTF4 |
| <i>cAMP specific phosphodiesterase</i> | K4DUC7, K4DUK1, Q4CQR0 |
| <i>Carnitine/choline acetyltransferase</i> | Q4DKK5 |
| <i>CAS/CSE/importin domain protein</i> | K4E5W4, Q4DX95 |
| <i>Casein kinase</i> | Q4DN97, Q4CSP0 |
| <i>Cation-transporting ATPase</i> | K4DMA8 |
| <i>Cell cycle associated protein MOB1</i> | K4E9B3, Q4DZ03 |
| <i>Centrin</i> | Q4E591, Q4CTH2, Q4E3T4 |
| <i>Chaperone DNAJ protein</i> | K4E449, K4E185, K4DTU5, Q4DLU9 |
| <i>Chaperonin</i> | K4E2E8 |
| <i>Choline/carnitine O-acetyltransferase</i> | K4DXA1, Q4CZN3 |
| <i>Citrate synthase</i> | K4DXA2 |
| <i>Cleavage and polyadenylation specificity factor</i> | K4E2Z4, Q4DNK0 |
| <i>Co-chaperone GrpE</i> | K4E538, Q4CPP0 |
| <i>Coatomer epsilon subunit</i> | K4E2P8 |
| <i>Cofilin/actin depolymerizing factor</i> | Q4D8D3 |
| <i>Copper homeostasis protein</i> | Q4CRX3 |
| <i>CTP synthase</i> | K4DTM2 |
| <i>CYC2-like cyclin 6</i> | Q4DY40 |
| <i>Cyclophilin</i> | K4E9L6, Q4DFL3, Q4DQ45, Q4DJC5, Q4DNS3 |
| <i>Cystathione gamma lyase</i> | K4DYZ3, Q4DUG8 |
| <i>Cysteine desulfurase</i> | K4DYK7 |
| <i>Cysteine peptidase inhibitor</i> | Q4DH32 |
| <i>Cysteine peptidase</i> | Q4CTX9, K4DV55, K4EEA1 |
| <i>Cytochrome b-domain protein</i> | K4E8Y5, Q4DJQ7 |

Tabela suplementar 6. Proteínas putativas preditas para secreção por via não clássica.
(continuação)

| Descrição ^a | Código identificador ^b |
|--|--|
| <i>Cytochrome c oxidase assembly factor</i> | Q4DE34, Q4D9Q9, K4DU78 |
| <i>Cytochrome c oxidase assembly protein</i> | K4E1Y0, Q4DEH9 |
| <i>Cytochrome c oxidase subunit 10</i> | Q4CLQ9, Q4CM35, Q4CZ48 |
| <i>Cytochrome c oxidase subunit IV</i> | K4DVU1, Q4E5P5 |
| <i>Cytochrome c oxidase subunit V</i> | Q4D214 |
| <i>Cytochrome c oxidase VII</i> | Q4CXC3, Q4DTY0 |
| <i>Cytochrome c</i> | Q4CV48 |
| <i>Cytoskeleton-associated protein CAP5.5</i> | K4E5Y1 |
| <i>D-isomer specific 2-hydroxyacid dehydrogenase</i> | K4E3Q2, Q4CU50, Q4DAE1 |
| <i>Deoxyuridine triphosphatase</i> | Q4E2Y0 |
| <i>Dihydrolipoamide branched chain transacylase</i> | K4DM80 |
| <i>Dihydroxyacetone phosphate acyltransferase</i> | K4E377 |
| <i>Diphosphomevalonate decarboxylase</i> | Q4E1U8 |
| <i>Dispersed gene family protein 1 (DGF-1)</i> | K4DMS3, K4DQI7, K4EDG8, Q4CST2, Q4D805, Q4CTC1, Q4CUD3, Q4CVS2, Q4CSF7, Q4CXH5, Q4CT27, Q4DH14, Q4DF03, Q4DL68, Q4D6C1, Q4CQN1, Q4CQI5, Q4D051, Q4CNW4, Q4CQT3, Q4CY27, Q4CNB9, Q4DJH9, Q4CTB2, Q4CQI0, Q4DS00, Q4DHR5, Q4DKH7, Q4E5L0, Q4DT01, Q4DDW5, Q4E489, Q4DSH9, Q4DIL7, Q4D7A5, Q4DTW5, Q4DPJ9, Q4E1B3, Q4D1D7, Q4DVF3, Q4DNC0, Q4DJH8, Q4DE09, Q4DUJ4, Q4E0M6, Q4D276, Q4D523, Q4D378, Q4DTL6, Q4DJI4, Q4CZ77 |
| <i>DNA repair protein</i> | K4DWN3 |
| <i>DNA-directed RNA polymerase</i> | K4E6Y0 |
| <i>DnaJ chaperone protein</i> | K4DUP5, Q4CXG2 |
| <i>dTDP-glucose 4,6-dehydratase</i> | K4E503 |
| <i>Dynammin-like protein</i> | K4DVQ5 |
| <i>Dynein arm light chain</i> | K4DR31, Q4DLN1, Q4DTQ6, K4E9R5 |
| <i>Dynein-associated protein</i> | Q4DG38 |
| <i>Electron transfer flavoprotein</i> | K4DYT4, Q4DR32 |
| <i>Elongation factor 1-alpha</i> | Q4CXI1 |
| <i>Elongation factor 1-gamma (EF-1-gamma)</i> | K4DW77, Q4CXW2 |
| <i>Endo-beta-N-acetylglucosaminidase</i> | K4E4K6 |
| <i>Enoyl-CoA hydratase</i> | K4DTD3, Q4E679, K4DZ33, Q4DMX9 |

Tabela suplementar 6. Proteínas putativas preditas para secreção por via não clássica.
(continuação)

| Descrição ^a | Código identificador ^b |
|---|---|
| <i>Epsilon tubulin</i> | K4DM83, K4DMA6 |
| <i>ER lumen protein retaining receptor</i> | K4E771, Q4DMI8 |
| <i>Eukaryotic initiation factor 5a</i> | Q4E4N4 |
| <i>Eukaryotic release factor 3</i> | K4E9H0, Q4CXW9 |
| <i>Eukaryotic translation initiation factor 1A</i> | K4DSL1 |
| <i>Eukaryotic translation initiation factor 3 subunit I</i> | Q69EZ4, Q4D452, Q4DNQ4 |
| <i>Eukaryotic translation initiation factor 3 subunit L</i> | Q4D5W3 |
| <i>Eukaryotic translation initiation factor 3 subunit</i> | K4DK02 |
| <i>Eukaryotic translation initiation factor 5</i> | K4E173 |
| <i>Eukaryotic translation initiation factor eIF-4E</i> | K4DZ28, Q4D511 |
| <i>Exportin 1</i> | K4E2X7 |
| <i>Fatty acid desaturase</i> | K4EAI2, Q4CKQ6, Q4DTH7 |
| <i>Fatty acid elongase</i> | K4DK32, K4DKF5, K4DK66, Q4DHY1 |
| <i>Ferric reductase</i> | Q4D1K3 |
| <i>Fibrillarin</i> | K4E142, K4DX81, Q4DZU7, Q4DIZ5, Q4DYJ3 |
| <i>Flagellar radial spoke component</i> | K4E802, Q4DE57 |
| <i>Formin C</i> | Q4D3C2 |
| <i>Fumarate hydratase</i> | K4EA83, Q4DRX1, Q4CQW5 |
| <i>Gim5A protein</i> | K4DY43 |
| <i>Glucokinase</i> | K4E0J8 |
| <i>Glucosamine-6-phosphate isomerase</i> | K4DZ69 |
| <i>Glucose-6-phosphate 1-dehydrogenase</i> | K4E1P8 |
| <i>Glucose-regulated protein 78</i> | K4E7W8 |
| <i>Glutamamyl carboxypeptidase</i> | K4E778, K4ED56 |
| <i>Glutamate dehydrogenase</i> | K4E4Z8, Q4DWV8, Q4D5C2 |
| <i>Glutamine synthetase</i> | Q4CN04 |
| <i>Glutaredoxin</i> | Q4DGB1 |
| <i>Glutathione peroxidase</i> | Q4DJ87, Q4DEJ4, Q4DEJ5, Q4DEJ6, Q4DJ86 |
| <i>Glutathione-S-transferase/glutaredoxin</i> | Q4DMF2 |
| <i>Glyceraldehyde 3-phosphate dehydrogenase</i> | Q4DZT1, K4E4T5 |
| <i>Glycosomal malate dehydrogenase</i> | O61084, Q4DDK8, Q4DRD8 |
| <i>Glycosomal membrane protein</i> | K4DU68, K4EA72, Q4CYY9 |
| <i>Glycyl-tRNA synthetase</i> | K4DNM9 |
| <i>Glycylpeptide N-tetradecanoyltransferase</i> | K4DXF7 |
| <i>GMP synthase</i> | K4EBX4 |
| <i>Golgi SNARE protein-like</i> | K4DR83, Q4D726 |

Tabela suplementar 6. Proteínas putativas preditas para secreção por via não clássica.
(continuação)

| Descrição ^a | Código identificador ^b |
|---|---|
| <i>GPI alpha-mannosyltransferase III</i> | K4DLS0 |
| <i>GPI anchor biosynthesis protein</i> | Q4CZK3 |
| <i>GPI transamidase component Tta2</i> | K4DLV0 |
| <i>GTP-binding nuclear protein rtb2</i> | K4DWD6, Q4CP51, Q4DIB9 |
| <i>GTP-binding protein</i> | Q9XZK0 |
| <i>Guanine deaminase</i> | K4DTI7 |
| <i>Guanylate kinase</i> | K4E0J2, Q4DPB1 |
| <i>Haloacid dehalogenase-like hydrolase</i> | K4E7Z4 |
| <i>Heat shock 70 kDa protein</i> | Q4DXG9, K4E4W4, Q4CVR9, Q4CU95, Q4DTM8, Q4DTM9 |
| <i>Heat shock protein DnaJ</i> | Q4E4K5, Q4D832 |
| <i>Heat shock protein-like protein</i> | K4E9T7 |
| <i>Helicase</i> | Q4CSN6 |
| <i>Hexose transporter</i> | Q4CQF8, Q4DKD0 |
| <i>High mobility group protein</i> | K4E3C6 |
| <i>Histone H2A</i> | Q4DYB0, Q4CYH8, Q4DYB1 |
| <i>Histone H2B</i> | Q4CTD7, Q4CTD8 |
| <i>Histone H3</i> | Q4DRD1, Q4CWE9 |
| <i>Homoserine kinase</i> | K4E4R7, Q4CME2 |
| <i>Hslvu proteolytic protein</i> | Q4D729, K4DRE8 |
| <i>Hydrolase alpha/beta fold family</i> | K4DPI3, Q4CQ95 |
| <i>Hypoxanthine-guanine phosphoribosyltransferase</i> | K4DRA8, Q4DGA2 |
| <i>I/6 autoantigen</i> | K4E8Y0 |
| <i>IgE-dependent histamine-releasing factor</i> | Q4CW52, Q4CWM1 |
| <i>Immunodominant antigen</i> | K4DKE0 |
| <i>Importin beta-1 subunit</i> | Q4DZ97 |
| <i>Importin subunit alpha</i> | K4E5S9 |
| <i>Inosine-5'-monophosphate dehydrogenase</i> | K4DPJ9, K4DXK1 |
| <i>Inosine-adenosine-guanosine-nucleoside hydrolase</i> | K4E1L0 |
| <i>Isocitrate dehydrogenase [NADP]</i> | K4DIS2 |
| <i>Katanin</i> | Q4DTR4 |
| <i>Kinesin</i> | K4DNY8, K4E9E7, Q4DDL0, Q4DYM0, Q4E0R3 |
| <i>Kinetoplast DNA-associated protein</i> | Q4CXF1, Q4DNI4, Q4CN27 |
| <i>Kinetoplastid membrane protein KMP-11</i> | Q4D7Y4 |
| <i>Kynureninase</i> | K4E455, Q4DQ63 |
| <i>La RNA binding protein</i> | K4DUV1 |
| <i>Lanosterol 14-alpha-demethylase</i> | K4E5U7 |

Tabela suplementar 6. Proteínas putativas preditas para secreção por via não clássica.
(continuação)

| Descrição ^a | Código identificador ^b |
|---|--|
| <i>Lanosterol synthase</i> | K4EA17 |
| <i>Leucyl-tRNA synthetase</i> | Q4CL52, K4DRR4 |
| <i>Lipophosphoglycan biosynthetic protein</i> | Q4D253 |
| <i>Lysophospholipase</i> | K4EAB5, Q4DA08 |
| <i>Malate dehydrogenase</i> | Q4D4A0, K4E1P3, K4E6H1 |
| <i>Mannose-6-phosphate isomerase</i> | K4E2C0 |
| <i>MCAK-like kinesin</i> | K4E4U9, K4EAX0, Q4DGE0, Q4D1U1, Q4DFR8 |
| <i>Methionine aminopeptidase</i> | K4DRB9 |
| <i>Methylthioadenosine phosphorylase</i> | Q4DQ97 |
| <i>Mevalonate kinase</i> | K4E5Y8 |
| <i>Mevalonate-diphosphate decarboxylase</i> | K4DWR3 |
| <i>Microtubule-associated protein</i> | Q4D5A7, Q4CMT2 |
| <i>Mitochondrial carrier protein</i> | K4EA35, K4DZ79, K4DTW7, Q4DR73, Q4D8C1, Q4DFR9, Q4DMB4, Q4DTH1, Q4DA69 |
| <i>Mitochondrial DNA polymerase beta</i> | K4DJT0, Q4E024 |
| <i>Mitochondrial DNA topoisomerase II</i> | K4ECA7 |
| <i>Mitochondrial oligo U binding protein TBRGG1</i> | Q4D590 |
| <i>Mitochondrial processing peptidase</i> | K4DUT8, K4E7E9, Q4DI18, K4DSZ5 |
| <i>Mitochondrial RNA binding protein</i> | K4DZV7, Q4D0R4, Q4CVB9, K4EDR7, Q4DCA9, Q4DBM0, Q4DWJ0, Q4E544 |
| <i>Monoxygenase</i> | K4E959, K4DW79 |
| <i>Mucin-associated surface protein (MASP)</i> | K4DTC4 |
| <i>Multidrug resistance protein E</i> | K4E642 |
| <i>Multidrug resistance-associated protein</i> | K4DSV9 |
| <i>Myosin heavy chain</i> | K4DW36 |
| <i>N-acetylglucosamine-6-phosphate deacetylase</i> | Q4CYW6 |
| <i>N-acetyltransferase complex ARD1 subunit</i> | Q4DSU6 |
| <i>NAD(P)-dependent oxidoreductase</i> | K4DPL0 |
| <i>NADH dehydrogenase</i> | Q4DCU2, Q4D9I1 |
| <i>NADH-cytochrome b5 reductase</i> | K4EAL6, Q4DYC3 |
| <i>NADH-dependent fumarate reductase</i> | K4DQV8, K4DZK3, K4DZS8, Q4D779 |
| <i>Nascent polypeptide associated complex subunit</i> | Q4D9L7 |
| <i>Nitrate reductase</i> | Q4DA56 |
| <i>Nitrilase</i> | K4E0Q3 |
| <i>Nuclear movement protein</i> | Q4DA46 |
| <i>Nucleobase transporter</i> | K4DU75 |

Tabela suplementar 6. Proteínas putativas preditas para secreção por via não clássica.
(continuação)

| Descrição ^a | Código identificador ^b |
|--|--|
| <i>Nucleolar RNA-binding protein</i> | Q4D244, Q4D245, Q4DHV9, Q4DHW0 |
| <i>Nucleoporin</i> | K4E164 |
| <i>Nucleoside diphosphate kinase</i> | Q4E256 |
| <i>Nucleoside phosphorylase</i> | Q4DKS7 |
| <i>Nucleosome assembly protein</i> | K4DXN7, Q4D5X8, Q4DSA7, K4DU97, Q4CVR4, Q4D2Z8 |
| <i>NUP-1 protein</i> | K4DPP5, Q4DB82 |
| <i>Oligosaccharyl transferase subunit</i> | Q4D2C9 |
| <i>Orotidine-5-phosphate decarboxylase/orotate phosphoribosyltransferase</i> | K4EC08 |
| <i>Oxidoreductase</i> | Q4DAS0 |
| <i>P-type H⁺-ATPase</i> | Q4CQ52, Q4CRY5 |
| <i>p21 antigen protein</i> | Q4CSP8 |
| <i>Pantothenate kinase subunit</i> | K4E945 |
| <i>Paraflagellar rod protein 3</i> | K4EBQ5, K4E370, Q4D634 |
| <i>Peptidase M20/M25/M40</i> | K4EBF2 |
| <i>Peptidase t</i> | K4DNS4 |
| <i>Peptide methionine sulfoxide reductase</i> | Q4DK88 |
| <i>Peptidyl-prolyl cis-trans isomerase</i> | K4DSN8, Q4E4L9, Q4CY23 |
| <i>Peroxin-2</i> | Q4E2J8 |
| <i>Peroxiredoxin</i> | O79469 |
| <i>Peroxisomal membrane protein 4</i> | K4DZE8 |
| <i>Phosphatidylinositol 3-kinase</i> | K4EDT8 |
| <i>Phosphoacetylglucosamine mutase</i> | K4DMG9, Q4DA33 |
| <i>Phosphoglycerate kinase</i> | K4E587, Q4DX96, Q4D192 |
| <i>Phosphoglycerate mutase</i> | Q4DLI0 |
| <i>Phosphomannomutase</i> | K4E0A5, K4E3H7, Q4DX52 |
| <i>Phosphonopyruvate decarboxylase</i> | K4E7W9 |
| <i>Phosphoribosylpyrophosphate synthetase</i> | K4E0L3 |
| <i>Plasma membrane Ca²⁺ ATPase</i> | K4DM23 |
| <i>Poly-zinc finger protein 2</i> | Q4D6T9 |
| <i>Poly(A) binding protein</i> | Q27335, Q4E4V0 |
| <i>Procyclic form surface glycoprotein</i> | Q4DZR9 |
| <i>Profilin</i> | Q4CMM4, Q4CZF5 |
| <i>Prohibitin</i> | K4EBM1, K4ECB4, Q4D8I7 |
| <i>Proline oxidase</i> | Q4CVA1 |
| <i>Prolyl oligopeptidase</i> | Q4E132, K4E2C3 |
| <i>Prostaglandin F synthase</i> | Q4DJ07, K4DRH6, Q4E4V7 |

Tabela suplementar 6. Proteínas putativas preditas para secreção por via não clássica.
(continuação)

| Descrição ^a | Código identificador ^b |
|--|--|
| <i>Proteasome activator protein pa26</i> | K4DZM2, Q4CPR5 |
| <i>Proteasome beta 2 subunit</i> | K4DTA7, Q4CU77 |
| <i>Proteasome regulatory ATPase subunit 3</i> | K4DUP0, Q4DW08, Q4E0K2 |
| <i>Proteasome regulatory non-ATPase subunit</i> | K4E1J2, Q4DKG4, K4DZM6, Q4E268, K4DY52, K4E2K3, K4DMB8 |
| <i>Proteasome subunit alpha type</i> | K4DN23, Q4DAW6 |
| <i>Proteasome subunit beta type</i> | K4E4C0, K4DUM4, Q4D579, Q4E0L9 |
| <i>Protein disulfide isomerase</i> | K4DU85 |
| <i>Protein kinase</i> | Q4E4T3, K4DUQ8, K4DZN9, Q4CXN4, Q4D6L8 |
| <i>Protein phosphatase</i> | K4DME1, K4E388, Q4E575, Q4E576 |
| <i>Protein translation factor SUI1 homolog</i> | Q4DM75 |
| <i>Protein transport protein Sec13</i> | K4DWV6 |
| <i>Protein transport protein Sec23A</i> | K4DIR6 |
| <i>Protein transport protein Sec24A</i> | K4DJF8, Q4DWT3 |
| <i>Protein tyrosine phosphatase</i> | K4E0I3 |
| <i>Protein-L-isoaspartate O-methyltransferase</i> | K4E2B2, Q4CZD3 |
| <i>Proton motive ATPase</i> | Q4CQ53, K4DTT7, K4DWB7 |
| <i>Pumilio protein</i> | K4DLB7 |
| <i>Pumilio/PUF RNA binding protein 2</i> | K4DWC6 |
| <i>Pyridoxal kinase</i> | K4E5S5 |
| <i>Pyroglutamyl-peptidase I (PGP)</i> | Q4DFZ2, K4DJ60 |
| <i>Pyrroline-5-carboxylate synthetase-like protein</i> | K4E0F1 |
| <i>Quinone oxidoreductase</i> | K4E1N3 |
| <i>Quinonoid dihydropteridine reductase</i> | K4E4Y8 |
| <i>R27-2 protein</i> | Q4DV01 |
| <i>Rab1 small GTP-binding protein</i> | K4ED44 |
| <i>Rab7 GTP binding protein</i> | Q4E4T4 |
| <i>Rac serine-threonine kinase</i> | K4DW26 |
| <i>Ran-binding protein</i> | Q4DKU8 |
| <i>Ras-related protein rab-5</i> | K4E8J7 |
| <i>Replication Factor A</i> | K4DPS6, K4E6Z1 |
| <i>Retrotransposon hot spot (RHS) protein</i> | K4DL33, K4DZH4, K4DRH2, K4DKS7, K4DQJ6, K4DRA7, K4DN42, K4DS17, K4DJF0, K4DRI3, K4E338, K4E7N3, K4DJ07, K4DL44, K4E6F3, K4DL11, K4DPN7, K4DPQ6, K4DMW1, K4E546, K4DLE2, K4DPY1, K4DPW1, K4DIY1, |

Tabela suplementar 6. Proteínas putativas preditas para secreção por via não clássica.
(continuação)

| Descrição ^a | Código identificador ^b |
|--|--|
| <i>Retrotransposon hot spot (RHS) protein</i> (continuação) | K4DV10, K4DKT3, K4DII2, K4DZ72, K4DQ94, K4DMN5, K4DMD8, K4EDP2, K4E946, K4EDN2, K4DML2, K4DMS1, K4DKI7, K4DMY0, K4EA98, K4DKZ4, K4DPL1, K4ECC3, K4E463, K4DIP2, K4DRE6, K4E7W2, K4E116, K4DN54, Q4CX54, Q4CXJ0, Q4CQP8, Q4DC18, Q4CZV4, Q4E0D4, Q4CNJ6, Q4D2F5, Q4D6R8, K4DS53, K4DRN4, K4EAZ5, K4DVK8, K4DNX1, K4E5T1, K4DXL2, K4DPL3, K4DIH5, K4EB42, K4DTD0, K4DYM5, K4DSC9, K4E6P5, K4DMS7, K4EBS9, K4DUW6, K4DRU6, K4DPC2, K4DKN0, K4DIU0, K4DPV5, K4DJ08, K4EAY3, K4DQS3, K4E411, K4DS18, K4DLR5, K4DSG8, K4DYW7, K4DRE7, Q4DKD3, Q4DK56, Q4D2G1, Q4DKH6, Q4DM58, Q4CUW8, Q4DWU4, Q4CYR3, Q4E157, Q4DM62, Q4DR27 |
| <i>Rhomboid-like protein</i> | Q4CV15, K4DTW8 |
| <i>Riboflavin kinase</i> | K4DYC9 |
| <i>Ribonuclease</i> | Q4D0A9, Q4D0B0, K4DUI3, Q4CU74, Q4D5U5 |
| <i>Ribonucleoprotein p18</i> | K4E7G6 |
| <i>Ribonucleoside-diphosphate reductase small chain</i> | K4DPY7, Q4D6V2 |
| <i>Ribose 5-phosphate isomerase</i> | Q4DBP9 |
| <i>Ribosomal protein</i> | K4EBB3, Q4DJZ6, Q4CTG9, Q4DED2, Q4E078, Q4D604, Q4DMS7, Q4E581, K4DWL1, Q4E5H8, Q4DJY1, Q4D6N9, Q4D6H7, Q4CYE4, Q4CXR4, Q4CMS5, Q4CPB5, K4E4B4, Q4CUC8, Q4CUC7, Q4DBY4 |
| <i>RNA-binding protein</i> | K4DXA0, Q4CXT3, K4DX96, K4DU33, K4DXD7, K4ECS4, K4DXV7, K4EAR6, K4E555, Q4DYL2, Q4DYL1, Q4E5D6, Q4E1N2, Q4DE30, Q4E1N5 |
| <i>RNA-editing complex protein</i> | K4DMC3, Q4DK95, Q4D022 |
| <i>S-adenosylmethionine decarboxylase proenzyme</i> | K4E603, Q4DQQ5, Q4CWQ0 |

Tabela suplementar 6. Proteínas putativas preditas para secreção por via não clássica.
(continuação)

| Descrição ^a | Código identificador ^b |
|--|---|
| <i>S-adenosylmethionine synthase</i> | K4E4J2, Q4CSC4 |
| <i>Sarcoplasmic reticulum glycoprotein</i> | Q4DCC9 |
| <i>Serine/threonine protein kinase</i> | Q4E650, K4E1U9, K4DM34, Q4CXF6, K4E714 |
| <i>Serine/threonine-protein phosphatase</i> | K4DLP7, K4E6R0, Q4CV19, Q4DMJ3 |
| <i>SH3 domain protein</i> | K4E5I3, Q4DMT4 |
| <i>Short-chain dehydrogenase</i> | K4DZW5, Q4DZT0, K4DSN3, Q4D5D5, Q4D3C6, Q4DKY8, K4DRS1 |
| <i>Small glutamine-rich tetratricopeptide repeat protein</i> | K4E551 |
| <i>Small GTP-binding protein</i> | Q4D7V7, K4DM88, K4EDG3, K4EBK1, Q4DUG2, Q4D1E8, Q4DZ75, K4E9U2, Q4D092 |
| <i>Small nuclear ribonucleoprotein Smd3</i> | Q4DX43 |
| <i>Small Rab GTP binding protein</i> | Q4DSV2 |
| <i>SNARE protein</i> | K4DXT0 |
| <i>snoRNP protein GAR1</i> | K4E3Z3, Q4D9X5 |
| <i>Spermidine synthase</i> | K4DV45, Q4DA73 |
| <i>Spliced leader RNA PSE-promoter transcription factor</i> | K4DW76 |
| <i>Sterol 24-c-methyltransferase</i> | K4DT61 |
| <i>Sterol C-24 reductase</i> | K4E0G3, Q4DR82, Q4DIV5 |
| <i>Stress-inducible protein STI1-like</i> | K4DM28 |
| <i>Structural maintenance of chromosome (SMC) family protein</i> | K4DWP0 |
| <i>Structural maintenance of chromosome protein 4</i> | K4DQR0 |
| <i>Succinate dehydrogenase</i> | K4E2R8, Q4CVH2 |
| <i>Succinyl-CoA synthetase alpha subunit</i> | Q4E2I6 |
| <i>Succinyl-CoA:3-ketoacid-coenzyme A transferase</i> | K4DYX5, Q4D0L3 |
| <i>Superoxide dismutase</i> | K4DJW9, K4DQ73, Q4D5A6, Q4CUQ5 |
| <i>T-complex protein 1 subunit alpha</i> | K4DZT6 |
| <i>Thiolase protein-like protein</i> | K4E746 |
| <i>Thioredoxin</i> | Q4DDZ3, Q4DFA1 |
| <i>Trans-sialidase</i> | K4ECX2, K4DNT9, K4DIM0, K4DXB0, Q4DDX6, Q4DBD2, Q4CQC9, K4DQE4, K4DSP2, K4DVP7, K4DVC9, Q4DDU4, Q4CZA1, Q4CXJ1, Q4DBD3, Q4CYG9 |
| <i>Transketolase</i> | K4EAC1 |
| <i>Translation elongation factor 1-beta</i> | K4E1H5, Q4DWC3 |
| <i>Translation initiation factor EIF-2b alpha subunit</i> | K4DUD4, Q4E0M1, Q4DW27 |

Tabela suplementar 6. Proteínas putativas preditas para secreção por via não clássica.
(conclusão)

| Descrição ^a | Código identificador ^b |
|--|--|
| <i>Transport protein particle (TRAPP) subunit</i> | K4DYF8, Q4D0X7 |
| <i>Tricarboxylate carrier</i> | K4E4C4, Q4CVF7 |
| <i>Trifunctional enzyme alpha subunit</i> | Q4CWB6 |
| <i>Trypanothione reductase</i> | K4E0T9 |
| <i>Tryparedoxin peroxidase</i> | Q4CX87, Q4CM56 |
| <i>Tryparedoxin</i> | Q4D1B8 |
| <i>Tyrosine aminotransferase</i> | K4DU56, Q4E4E9 |
| <i>U-box domain protein</i> | K4DJJ0 |
| <i>U-rich RNA-binding protein UBP-2</i> | Q4E1N4 |
| <i>Ubiquinone biosynthesis protein COQ7</i> | K4DSR1 |
| <i>Ubiquitin carboxyl-terminal hydrolase</i> | K4DTY6, Q4E680 |
| <i>Ubiquitin fusion degradation protein 2</i> | K4E526 |
| <i>Ubiquitin hydrolase</i> | K4ECZ8, K4E5A2 |
| <i>Ubiquitin ligase</i> | Q4D6V3, K4DVQ6 |
| <i>Ubiquitin</i> | Q4CNV2 |
| <i>Ubiquitin-activating enzyme e1</i> | K4DXM3 |
| <i>Ubiquitin-conjugating enzyme</i> | K4EC09, Q4E344, Q4DYH2, K4DV56 |
| <i>Ubiquitin/ribosomal protein S27a</i> | Q4D1T6 |
| <i>UDP-Gal or UDP-GlcNAc-dependent glycosyltransferase</i> | K4DU89, Q4CPN8, Q4D0U0, Q4D4N1, Q4CKW9 |
| <i>UDP-galactose 4-epimerase</i> | Q4DUZ8, K4DYX9, Q4DN47 |
| <i>UDP-sugar pyrophosphorylase</i> | K4DPF2 |
| <i>Universal minicircle sequence binding prot. (UMSBP)</i> | Q4D6T7, Q4D6T8 |
| <i>Uracil phosphoribosyltransferase</i> | K4E368, Q4E301 |
| <i>UV excision repair RAD23-like protein</i> | K4E4N6, Q4D7G6 |
| <i>Vacuolar ATP synthase</i> | K4DWM1, Q4E5H0, K4E806, Q4D3F5, Q4D3E6 |
| <i>Vacuolar proton translocating ATPase subunit A</i> | K4EDC9, Q4DSC7, Q4DK78 |
| <i>Vesicle-associated membrane protein</i> | K4EAS6, K4DKF0, Q4DDN8, Q4DS88, Q4DP28 |
| <i>Vesicle-fusing ATPase</i> | K4DUS1 |
| <i>Vesicular-fusion ATPase-like protein</i> | K4DYN3, Q4D0R7 |
| <i>Vesicle-associated membrane protein</i> | K4DNY5, Q4E1T0 |
| <i>Voltage-dependent anion-selective channel</i> | K4ECN2 |
| <i>Zinc finger protein 2</i> | Q4D7N5, Q4CLW6 |

^a Descrição da proteína anotada proveniente do Uniprot

^b Número de acesso do Uniprot