



Programa de Pós-Graduação em

BIOTECNOLOGIA

mestrado & doutorado

Edson Adriano de Souza

**GERAÇÃO E ANÁLISE DE GSS (GENOME
SURVEY SEQUENCES) DE *Trypanosoma vivax***

Florianópolis

Fevereiro de 2005



Programa de Pós-Graduação em

BIOTECNOLOGIA

mestrado & doutorado

GERAÇÃO E ANÁLISE DE GSS (GENOME SURVEY SEQUENCES) DE *Trypanosoma vivax*

O presente trabalho foi desenvolvido nos Laboratórios de Protozoologia e de Bioinformática da Universidade Federal de Santa Catarina e no Laboratório de Biologia Molecular de Tripanosomatídeos do Instituto Oswaldo Cruz, Fiocruz, sob a orientação do Dr. Edmundo C. Grisard (UFSC) e do Dr. Alberto Martín Rivera Dávila (IOC/Fiocruz), como requisito parcial para obtenção do grau de Mestre em Biotecnologia junto ao Programa de Pós-Graduação em Biotecnologia da Universidade Federal de Santa Catarina.

Florianópolis

Fevereiro de 2005

Aos meus pais, Cláudio e Maria,
pela coragem e exemplo diante
dos desafios da vida.

AGRADECIMENTOS

Ao Dr. Edmundo Carlos Grisard, por ter aceitado me orientar e pelo convívio maduro e agradável.

Ao Dr. Alberto Martín Rivera Dávila, cuja grande “co-orientação” foi imprescindível para realização do trabalho. Muito obrigado Alberto!

À Cristiane Quimelli Snoeijer e Carlos Rodrigo Bueno, do Laboratório de Protozoologia da Universidade Federal de Santa Catarina, pelas realizações das mini-preps e todo auxílio prestado em laboratório.

À Silvana S. Souza e Luana Tatiana Guerreiro, do Laboratório de Biologia Molecular de Tripanosomatídeos do Instituto Oswaldo Cruz, Fiocruz/ RJ, pela etapa de construção da biblioteca genômica.

Ao colega Glauber Wagner, pela grande colaboração na análise dos dados.

E, também, ao pessoal do Laboratório de Bioinformática do IOC/Fiocruz, Pablo N. Mendes, Linair M. Campos e Paulo F. Pires.

"O homem é apenas seu projeto, só existe na medida em que se realiza, ele é tão somente o conjunto de seus atos"

Jean-Paul Sartre

RESUMO

O *Trypanosoma vivax* é um hemoparasita causador de tripanosomose em bovinos e está difundido na África e na maioria dos países da América do Sul, principalmente no Pantanal do Brasil e Bolívia com surtos recorrentes. As infecções causadas por *T. vivax* provocam doenças e até morte em rebanhos bovinos prejudicando a bovinocultura. No continente africano, a distribuição da parasitose está relacionada com a presença do vetor de desenvolvimento cíclico, a mosca tse-tse (*Glossina* sp.). No continente sul-americano, onde o vetor tse-tse é ausente, a transmissão ocorre de maneira mecânica por outras moscas hematófagas. A doença equivale a 1/3 da distribuição africana com as mesmas características, sintomas e recorrência de surtos. O conhecimento da biologia de *T. vivax* é escasso e atrasado por décadas pelo difícil cultivo *in vitro* e pela efêmera manutenção das cepas em animais de laboratório. Isto acarreta no desconhecimento de marcadores moleculares úteis para testes diagnósticos, bem como de alvos para o desenvolvimento de drogas. No presente trabalho apresentamos os resultados obtidos à partir da construção e seqüenciamento de uma biblioteca de DNA genômico da cepa ILDat 2160 de *T. vivax*, validada por apresentar GSSs correspondentes à seqüências de espécies de kinetoplastídeos relacionados. Foram obtidas 455 GSSs de alta qualidade, dentre as quais 21 seqüências não apresentaram similaridade com seqüências depositadas nos 12 bancos de dados consultados, representando possíveis genes órfãos. Estes resultados demonstram a importância da estratégia de seqüenciamento por GSS para a obtenção de informações sobre a biologia e biologia molecular de *T. vivax*.

ABSTRACT

Trypanosoma vivax is a hemoparasite which causes trypanosomiasis in bovines, being spread out in Africa and in the majority of the South-American countries, mainly in the *Pantanal* of Brazil and Bolivia, with recurrent outbreaks. The infection caused by *T. vivax* provokes illnesses and even death in bovines, thus harming the livestock industry. In the African continent, the distribution of the parasitism is related to the presence of the cyclical development vector, the tse-tse fly (*Glossina* sp.). In the South-American continent, where the tse-tse fly does not exist, the transmission occurs in a mechanical way by other blood-sucking flies. The illness is equivalent to one third of the African distribution with the same characteristics, symptoms and outbreak recurrence. The knowledge of the biology of *T. vivax* is scarce and out-of-date by decades, because of its difficult culture *in vitro* and due to the ephemeral maintenance of stocks in laboratory animals. This results in the lack of knowledge of useful molecular markers for diagnostic tests, as well as of targets for the development of drugs. In the present work, we present the results achieved by the construction and sequencing of a library of genomic DNA of stock ILDat 2160 of *T. vivax*, validated for presenting GSSs corresponding to the sequences of related kinetoplastide species. 455 GSSs of high quality have been obtained, amongst which 21 sequences have not presented similarity with the sequences deposited in the 12 data bases searched, representing possible orphan genes. These results have showed the importance of the strategy of GSS sequencing for the gathering of information on biology and on the molecular biology of *T. vivax*.

LISTA DE FIGURAS

- Figura 1** Representação gráfica do número de entradas (*hits*) obtidas para cada um nos bancos de dados confrontados com as GSS de *Trypanosoma vivax* geradas apresentadas na Tabela 2..... **25**
- Figura 2** Classificação das seqüências de acordo com as três categorias funcionais do Gene Ontology (ASHBURNER *et al.*, 2000) e respectivo percentual da quantidade encontrada..... **44**

LISTA DE TABELAS

Tabela 1	Resultados do seqüenciamento de clones de DNA genômico (GSS) da cepa ILDat 2160 de <i>Trypanosoma vivax</i>	23
Tabela 2	Número de entradas (<i>hits</i>) obtidas para cada um nos bancos de dados confrontados com as GSS de <i>Trypanosoma vivax</i> geradas no presente estudo.....	24
Tabela 3	Relação das GSS de <i>Trypanosoma vivax</i> identificadas com base em seqüências conhecidas de kinetoplastídeos e outros organismos em bancos de dados públicos (<i>e-value</i> limite de $<10^{-5}$)..	26
Tabela 4	Distribuição das identidades das GSS geradas nos banco de dados de domínios de proteínas CDD e InterPro, confirmadas através do banco de dados Blocks / Prints e análises feitas com os programas Glimmer e TestCode.....	42

LISTA DE ABREVIATURAS

% – Porcentagem

°C – Graus centígrados

µg – Micrograma

µL – Microlitro

µm – Micrômetro

Ag-ELISA – Antígeno-imunoensaio enzimático

cDNA – DNA complementar à RNAm

DNA – Ácido desoxirribonucléico

EST – Etiquetas de seqüências expressas (*expressed sequence tags*)

GSS – *Genome survey sequences*

INF-γ – Interferon-gama

IPTG – Isopropil-β-D-tiogalactopiranosídeo

Kb – Kilobase

KV – Kilovolt

LB - Meio Luria-Bertani

LB-ágar – Meio LB preparado em ágar

M – Molaridade

min. – Minuto

mL – Mililitro

mM – Milimolar

mRNA – RNA mensageiro

Oligo dT – Oligonucleotídeo formado pela base timina (T)

ORF – Janela aberta de leitura (*open reading frame*)

pb – pares de base

PCR – Reação em Cadeia da Polimerase (*Polymerase chain reaction*)

PCR-RFLP – Reação em Cadeia da Polimerase associada à digestão com enzimas de restrição

Pmol – picomol

Poli A – Oligonucleotídeo formado pela base adenina (A)

RFLP – Análise de polimorfismos de comprimento de fragmentos de restrição (*Restriction fragment lenght polymorphism*)

RNA – Ácido ribonucléico

rRNA – RNA ribossômico

SDS – dodecil sulfato de sódio

TRIS – Tris-hidroximetil aminometano

tRNA – RNA transportador

VSG – Glicoproteína variante de superfície (*Variant Surface Glycoprotein*)

X-Gal – 5-Bromo-4-Cloro-3-Indolil- β -D-Galactosídeo

Sumário

RESUMO	iv
ABSTRACT	v
LISTA DE FIGURAS	vi
LISTA DE TABELAS	vii
LISTA DE ABREVIATURAS	viii
1. INTRODUÇÃO	1
1.1. <i>Trypanosoma vivax</i> : Morfologia.....	2
1.2. <i>Trypanosoma vivax</i> : Infecção do hospedeiro.....	3
1.3. Diagnóstico da infecção pelo <i>Trypanosoma vivax</i>	7
1.4. Estudos moleculares e o genoma do <i>Trypanosoma vivax</i>	9
2. OBJETIVOS	14
2.1. Objetivo geral.....	14
2.2. Objetivos Específicos.....	14
3. MATERIAL E MÉTODOS	15
3.1. Extração do DNA genômico e construção da biblioteca genômica.....	15
3.2. Semi-normalização da biblioteca genômica.....	16
3.3. Crescimento bacteriano e extração de DNA plasmidial.....	17
3.4. Seqüenciamento de DNA.....	18
3.5. Análise das seqüências obtidas.....	19
4. RESULTADOS	21
4.1. Biblioteca genômica.....	21
4.2. Semi-normalização.....	21
4.3. Seqüenciamento de GSS.....	21
4.4. Classificação das Seqüências em Categorias Funcionais.....	25
4.4.1. Processo Biológico:	29
4.4.1.1. Proteólise.....	30
4.4.1.2. Processo de desenvolvimento.....	30
4.4.1.3. Processamento de DNA.....	31
4.4.1.4. Elementos genéticos móveis.....	32
4.4.1.5. Transporte intracelular.....	32
4.4.1.6. Metabolismo de nucleotídeos.....	33
4.4.1.7. Processo fisiológico.....	33

4.4.2. Componente Celular:	34
4.4.2.1. Citoesqueleto.....	34
4.4.2.2. Estrutura flagelar.....	34
4.4.2.3. Cinetoplasto.....	35
4.4.2.4. Membrana.....	35
4.4.2.5. Núcleo.....	35
4.4.3. Função Molecular:	36
4.4.3.1. Defesa.....	36
4.4.3.2. Replicação de DNA.....	37
4.4.3.3. Enzima.....	37
4.4.3.4. Metabolismo.....	39
4.4.3.5. Reparo de DNA.....	39
4.4.3.6. Transcrição.....	39
4.4.3.7. Transporte de íons.....	40
5. DISCUSSÃO	45
6. CONCLUSÕES	48
7. REFERÊNCIAS BIBLIOGRÁFICAS	50

1. INTRODUÇÃO

O *Trypanosoma vivax* é um hemoparasita de ruminantes que foi introduzido na América Latina com o gado importado da África, provavelmente no final do século 19. Atualmente, o protozoário está difundido na África do Sul e em dez dos 13 países do continente sul-americano, provocando doença e morte em rebanhos bovinos (JONES, DÁVILA, 2001). A presença do *T. vivax* na América do Sul é relatada desde 1919, com a ocorrência de infecções em bovinos na Guiana Francesa. Posteriormente, as infecções foram relatadas na Venezuela (1920), Guadalupe (1926), Martinica (1929), Colômbia (1931), Suriname (1938), Panamá (1941), Guiana (1953) e Brasil (1972). O parasita também foi detectado através de sorologia em 1977 em El-Salvador, Costa Rica, Equador, Peru e Paraguai, e, mais recentemente, em surtos recorrentes no pantanal brasileiro e boliviano (JONES, DÁVILA; 2001). Além da América do Sul, a distribuição do parasita alcança também as Ilhas Maurício, Ilhas do Caribe, Java e Indonésia (GARDINER, 1989).

Na África, sua transmissão ocorre de maneira cíclica pela mosca *Glossina* spp. (tsé-tsé), e secundariamente de forma mecânica por outras moscas hematófagas. O ciclo é confinado à região da probóscide da mosca, podendo haver uma invasão no intestino médio, onde o parasita não sobrevive por longo tempo. A maturação do tripanosoma está relacionada com a modulação da superfície de revestimento por glicoproteínas e a perda da adesão flagelar das formas epimastigotas (HOARE, 1972; GARDINER, 1989).

No continente africano, onde a tripanosomose bovina ocorre de forma distinta em vastas áreas, existe uma associação da invasão do homem e seu gado de corte dentro de áreas da selva infestadas por *Glossina* sp. (BOSSCHE, 2001). Na América

do Sul, onde a *Glossina* sp. é ausente, somente a transmissão mecânica ocorre através de tabanídeos e *Stomoxys* sp. (HOARE, 1972; DÁVILA, 2002).

1.1. *Trypanosoma vivax*: Morfologia

Este tripanosoma apresenta-se como um kinetoplastídeo típico com um flagelo anterior simples (HOARE, 1972; GARDINER, 1989). Formas extracelulares tripomastigota e epimastigota são encontradas no inseto vetor (*Glossina* sp.) que abriga o ciclo do parasita. Na América, os insetos transmitem o parasita de forma mecânica, onde o parasita encontra-se somente na forma tripomastigota (DÁVILA, 2002). O tamanho do comprimento médio do parasita varia de 21 μ m a 25,4 μ m (DÁVILA, RAMIREZ e SILVA, 1997), e se assemelha ao de *Trypanosoma brucei*, sendo distinguido deste pelos movimentos rápidos. O cinetoplasto é outra característica peculiar e tem valor diagnóstico, sendo grande, volumoso, redondo (circular a elíptico) e de posição terminal (HOARE, 1972; GARDINER, 1989). Além dessas características, o parasita possui alta similaridade morfológica com *Trypanosoma uniforme*, o qual possui comprimento menor, situado na faixa de menos 18 μ m. Outras diferenças são encontrados na susceptibilidade e restrição à hospedeiros, patogenicidade, distribuição geográfica e transmissão (GARDINER, 1989).

As formas sangüíneas tripomastigotas geralmente observadas são as largas e as delgadas, sendo a última, mais freqüente quando infectando roedores (GARDINER, 1989). A forma larga é arredondada em direção a extremidade posterior e sofre afilamento em direção a extremidade anterior. A forma delgada apresenta tais diferenças nas extremidades de modo menos visível em microscopia óptica (HOARE, 1972).

Pela presença da forma delgada do parasita, acreditava-se inicialmente em pleomorfismo da espécie; entretanto, demonstrou-se que ela surge após a forma larga e ocupa maior parte da proporção do pico de parasitemia (GARDINER, 1989). Essa forma, considerada a mais infectiva, é denominada forma tardia e possui características distinguíveis da forma inicial como: a) aparência granular quando corada pelo método de Giemsa; b) conteúdo de grandes vacúolos deformáveis quando observados vivos sob contraste óptico em microscopia de Nomarski; c) densidade diferente após centrifugação diferencial; d) tendência à adesão pelas pontas dos flagelos à partículas ou materiais celulares; e) baixa habilidade em iniciar infecções quando comparada com as formas sangüíneas iniciais; f) interrupção da síntese de DNA; g) capacidade aparente de se desenvolver em culturas de baixa temperatura, levando a estágios que normalmente se desenvolvem no inseto vetor (GARDINER, 1989).

1.2. *Trypanosoma vivax*: Infecção do hospedeiro

As formas tardias de *T. vivax*, aptas a infectar o hospedeiro, são usualmente encontradas na probóscide de *Glossina* sp. (GARDINER, 1989). Esse é um dos fatores que dificulta o cultivo do parasita *in vitro* por métodos convencionais, sendo requerido uma condição especial que simule certos fatores que estão contidos na saliva do inseto (HOARE, 1972). Segundo Graham e colaboradores (1998), a permanência do parasita na glândula salivar é importante, pois um dos promotores para glicoproteínas de superfície do parasita tem ativação quando o parasita encontra-se nesta região.

No meio ambiente, ovinos assim como bovinos, são os animais mais susceptíveis às infecções, atingindo taxas de mortalidade de até 70%. Os eqüinos

podem desenvolver doença crônica, porém os suínos são resistentes as infecções, e em camelos ela é considerada benigna. Em animais de laboratório, o curso da infecção por *T. vivax* é variável, tornando difícil o crescimento do parasita em cobaias e a cultura *in vitro*. Em coelhos, a infecção tem duração de até 25 dias, com parasitemia progressiva de dois a seis dias, seguida por rápido declínio e recaída em dois a nove dias mais tarde. Em ratos, a infecção alcança até 23 dias, caracterizada por uma alternância de crises e recaída (HOARE, 1972; GARDINER, 1989). A infectividade do parasita foi mais bem demonstrada com a inoculação experimental em ovelhas (HOARE, 1972).

O fato da ausência da *Glossina* sp., na qual ocorrem mudanças antigênicas relevantes para a infectividade do parasita, na América do Sul é curioso. No entanto, a incidência da doença aqui representa 1/3 da ocorrência na África com as mesmas características, sintomas e recorrência de surtos (JONES, DÁVILA, 2001). A susceptibilidade em relação a diferentes hospedeiros e cepas isoladas também não é explicada. Dwinger e colaboradores (1986) compararam a susceptibilidade de búfalo, gado bovino e caprinos à infecção de diferentes isolados de *T. vivax* (do leste e oeste africano). Os isolados foram comparados nos sintomas de reações de pele, parasitemia e anemia; e os resultados revelaram que búfalos apresentaram resistência as três cepas estudadas, que o gado bovino e caprinos apresentaram todos os sintomas para as cepas isoladas do oeste africano, e que as cepas do leste causaram reações de pele em caprinos e parasitemia em caprinos e gado bovino.

A doença causada por *T. vivax* em ruminantes pode ser resumida pelos seguintes sintomas clínicos: reações na pele, anemia, síndromes hemorrágicas e efeitos sobre a fertilidade (GARDINER, 1989). Nas regiões pantaneiras do Brasil e da Bolívia, os principais sinais da parasitose observados foram a anemia e

leucopenia, que são as causas principais da mortalidade. Outros sintomas observados incluem a lacrimação, fraqueza progressiva, perda de peso marcante, inapetência, diarreia e abortos durante o terceiro mês de gestação (SILVA *et al.*, 1999). O efeito da doença em fêmeas grávidas pode levar a nascimentos prematuros, morte perinatal e morte rápida após o nascimento (OKECH *et al.*, 1996).

O *T. vivax* não possui estágio intracelular, estando restrito ao sangue do hospedeiro, onde responde ao sistema imune de forma interessante. O *T. vivax* dispõe de uma glicoproteína variante de superfície ou VSG (*Variant Surface Glycoprotein*) ancorada na membrana externa, a qual muda continuamente seus epítomos em resposta aos diferentes processos imunológicos em um fenômeno chamado de variação antigênica (BARRY, McCULLOCH, 2001; DONELSON, 2003). O número de genes que permite a variação compreende aproximadamente 1.000 genes VSG e 20 a 40 sítios de expressão em tripanosomos (BORST, ULBERT, 2001; VANHAMME *et al.*, 2001; DONELSON, 2003).

A mudança periódica para expressão das várias VSGs pertence a uma estratégia de sobrevivência geral que aumenta a variação fenotípica. Essa estratégia engloba uma contingência de genes de vírus, bactérias e parasitas super mutantes, permitindo a tais organismos uma adaptação rápida à ambientes hostis ou variáveis (BARRY, McCULLOCH, 2001). Uma característica fundamental da variação antigênica é sua ligação com a dinâmica populacional de tripanosomos dentro de um único hospedeiro. Diferentes variantes surgem hierarquicamente, expressas por diferentes genes, dentro do hospedeiro, de modo aleatório em mecanismos subjacentes ainda não tão bem compreendidos (BARRY, McCULLOCH, 2001).

As glicoproteínas de superfície são expressas inicialmente nas formas tripomastigotas encontradas em *Glossina* sp., e a expressão desses genes torna-se

diferente quando o parasita encontra-se no hospedeiro (GRAHAM, WYMER, BARRY, 1998). Na América do Sul o número de serodemas e os processos imunes envolvidos são desconhecidos, sendo que a transmissão cíclica que conduz à expressão de VSG é ausente (GARDINER, 1989).

As respostas imunológicas das infecções têm sido estudadas principalmente em bovinos. O curso da doença no hospedeiro varia de acordo com a virulência das diferentes cepas, podendo ocasionar morte, doença crônica, ou ainda a autocura nos poucos casos de tolerância (HOARE, 1972). A sobrevivência do parasita depende, principalmente, da evasão do sistema imune do hospedeiro e da modulação periódica de seus antígenos de superfície (ZAMBRANO-VILLA *et al.*, 2002).

A variação antigênica é, entretanto, somente um dos prováveis mecanismos de defesa que permitem o estabelecimento da infecção e o crescimento do parasita. Outros processos de defesa também desempenham um papel na resistência do parasita, como uma proteína que induz as células T do hospedeiro a produzir INF- γ , e há evidência de uma enzima de resistência à ruptura celular semelhante à de *Leishmania* sp. (DONELSON, HILL, EL-SAYED, 1998; EL-SAYED, DONELSON, 1997).

Alguns casos de resistência natural ao *T. vivax* foram observados em bovinos, como em *Bos taurus* do oeste da África. Comparações entre gado imunotolerante e imunossusceptível mostram diferenças nas respostas por anticorpos, nos níveis de complemento e na expressão de citocinas; sendo desconhecido se essas diferenças determinem a resistência. Experimentos recentes sugerem que a resistência natural seja composta por dois mecanismos: um inato, controlando o crescimento do

parasita, e outro limitando a anemia pelo sistema hematopoiético (NAESSENS, TEALE, SILEGHEM, 2002).

No que se refere ao tratamento de animais infectados, um dos maiores problemas atualmente enfrentados é a resistência às drogas tripanocidas correntemente utilizadas, o isometamidium, o homidium e o dimenazene. Na África, essas drogas têm sido usadas sem controle epidemiológico por mais de 40 anos, culminando num processo de resistência emergente. Agravando ainda mais a situação, os altos custos de desenvolvimento e licenciamento para novas drogas despertam pouco interesse pelas indústrias farmacêuticas e dos governos do continente africano na pesquisa e desenvolvimento de novos medicamentos (GEERTS *et al.*, 2001). Na América do Sul o problema da resistência à droga também é observado, tendo sido relatada na Guiana Francesa a resistência de parasitas infectando *Bos indicus*, mesmo à dose máxima de dimenazene (DESQUESNES, ROCQUE, PEREGRINE, 1995).

Algumas abordagens para remediar o problema da resistência a drogas têm sido estudadas. Uma delas foi o esclarecimento do papel dos transportadores de nucleosídeos que mediam a captação de análogos de purinas, podendo também mediar a captação de drogas citotóxicas aos parasitas (BARRETT, FAIRLAMB, 1999). Segundo Landfear (2001), os avanços de genômica podem permitir a clonagem e expressão de genes transportadores de nucleosídeos, sustentando arranjos moleculares para examinar a função dessas permeases.

1.3. Diagnóstico da infecção pelo *Trypanosoma vivax*:

Um importante aspecto das infecções causadas pelo parasita é a necessidade de diagnósticos precisos no estágio inicial da doença a fim de permitir

um tratamento eficaz. Os testes sorológicos existentes como aglutinação em cartão, imunofluorescência indireta e imunoabsorção enzimática apresentam reação cruzada com *T. evansi*, e o diagnóstico direto por microhematócrito tem baixa sensibilidade (MADRUGA, MORZARIA e MAJIWA, 1999). A alternativa tem sido o investimento em técnicas moleculares de diagnóstico que apresentam índices de sensibilidade e especificidade elevados.

Masake e colaboradores (1997) utilizando testes baseados na PCR, detectaram uma maior taxa de resultados positivos (76%) em sangue contaminado nas comparações feitas com o ensaio Ag-ELISA (55%). Na comparação de Desquesnes (1997), a diferença nos resultados foi maior: 64% de resultado positivo em teste de PCR e 4% no ensaio Ag-ELISA. Essa grande diferença também foi constatada nas comparações com outras técnicas parasitológicas como o microhematócrito e o esfregaço sangüíneo corado pelo método de Giemsa, que não detectavam parasitemia em animais sabidamente infectados (ALMEIDA *et al.*, 1997; VENTURA *et al.*, 2001).

A técnica de PCR-RFLP também é considerada como um outro método de diagnóstico potencial. Geysen, Delespaux e Geerts (2003), usaram as enzimas de restrição *MspI* e *Eco571* para caracterizar e distinguir as espécies de tripanosomos de bovinos, especificamente o *T. congolense*, *T. vivax* e *T. theileri*, a partir da amplificação de genes do rRNA. Em outro estudo, Morlais e colaboradores (2001) utilizou seqüências marcadoras do DNA de hospedeiros e de vetores como alvos para a identificação de infecções, indicando que estes marcadores também podem ter uso no estudo de genética de populações.

1.4. Estudos moleculares e o genoma do *Trypanosoma vivax*:

Atualmente, o genoma do *T. vivax* está sendo seqüenciado pelo The Wellcome Trust Sanger Institute - Pathogen Sequencing Unit (PSU) através de um financiamento da Wellcome Trust. O projeto prevê o seqüenciamento do genoma nuclear do parasita, cujo estágio atual já atingiu uma cobertura de cinco vezes o tamanho estimado do genoma (http://www.sanger.ac.uk/Projects/T_vivax/overview.shtml, acesso em 09/02/05), sendo este considerado como o passo inicial para o estudo do genoma do *T. vivax*.

Com o avanço das técnicas de biologia molecular, e principalmente, no que diz respeito ao seqüenciamento de genes, a possibilidade de maior compreensão das bases moleculares dos organismos com estudos mais específicos torna-se uma realidade. Esse avanço tecnológico tem levado a formação de consórcios para o seqüenciamento genômico de organismos de interesse, os quais geram informações que podem ser empregadas em estudos decorrentes como proteomas, análises funcionais e manipulação genética. Os resultados dessas descobertas poderão contribuir significativamente para o desenvolvimento de novos testes diagnósticos e a descoberta ou desenvolvimento de drogas efetivas (DIAS-NETO *et al.*, 2000; DEGRAVE *et al.*, 2001; TARLETON, KISSINGER, 2001).

Para Donelson (2003), o estudo de genomas tem levado à compreensão da organização dos genes dentro dos grupos de sítios de expressão, o que pode finalmente, sugerir melhores modos de controle ou mesmo a eliminação da infecção. Tarleton e Kissinger (2001), argumentam que a informação decorrente dos projetos genoma proverá fontes para o desenvolvimento de vacinas, citando: “o uso do SAGE (*Análise Seriada da Expressão de Genes*) e a análise de microarranjos, bem como das ferramentas de bioinformática e ensaios funcionais que predizem a

localização de produtos gênicos, poderão suprir meios capazes de predizer quais produtos, em local e momento certos, serão bons alvos imunológicos”.

Dentre as técnicas utilizadas na geração de seqüências, a obtenção de seqüências curtas a partir de DNA genômico fragmentado aleatoriamente, as chamadas GSS (*genome survey sequences*), tem permitido levantar informações de genes e revelado um panorama geral dos principais genes de um determinado organismo. Essa abordagem tem se mostrado satisfatória em genomas com baixa concentração de região de íntrons, como no caso dos genomas densos dos tripanosomatídeos e de outros microorganismos (CARLTON *et al.*, 2001; SÁNCHEZ *et al.*, 2001).

Uma outra técnica de seqüenciamento também se baseia na geração de seqüências curtas denominadas EST, *expressed sequence tags* ou etiquetas de seqüências expressas. EST são seqüências derivadas de cDNA sintetizadas a partir de mRNA (SOUZA *et al.*, 2000) e permitem a demonstração do perfil de expressão de genes específicos num tecido ou estágio de um organismo. Huang e colaboradores (1999) utilizaram EST geradas de tecidos de próstata normal e cancerosa para demonstrar o comportamento do câncer específico em termos de expressão de genes.

A utilização das duas técnicas (seqüenciamento por GSS e por EST) no estudo de um dado organismo se mostra extremamente eficaz na identificação de exons, íntrons e dos processamentos nos genes (*splicing*) (DIAS-NETO *et al.*, 2000; KAN *et al.*, 2001).

As seqüências produzidas pelo seqüenciamento de EST representam a expressão de genes estágio específico, enquanto que, as seqüências obtidas por GSS representam o DNA genômico, conseguindo capturar trechos de regiões

codificantes únicas e/ou de baixa expressão. Essas ferramentas são poderosas na geração de sondas gene-específicos para estudos de mapeamento e para confirmação de anotação *in silico* de genomas (CARLTON *et al.*, 2001).

O seqüenciamento de GSS foi utilizado no projeto do patógeno de milho *Spiroplasma kunkelii* e demonstrou ser um procedimento eficiente para a descoberta de genes e caracterização do genoma. Um total de 396 etiquetas de seqüências únicas de alta qualidade representando 9% do genoma do patógeno foram obtidas e após análise, classificadas nas categorias funcionais de processos biológicos, funções moleculares e componentes celulares (BAI, SASKIA, HOGENHOUT, 2002).

No seqüenciamento de genes de *Brucella abortus*, 1.899 clones foram parcialmente seqüenciados produzindo 925 seqüências únicas de alta qualidade, das quais 470 foram classificadas em 15 categorias de funções celulares (SÁNCHEZ *et al.*, 2001). Em *Leishmania braziliensis*, em torno de 25% do genoma haplóide do organismo foi obtido, e os genes seqüenciados foram distribuídos em oito categorias funcionais: ligação (44,6%), atividade catalítica (41,75%), atividade transportadora, atividade de regulação enzimática, atividade de regulação de transcrição, atividade de tradução de sinal, atividade molecular estrutural e atividade reguladora de apoptose (LAURENTINO *et al.*, 2004).

O clone RF122 de *Staphylococcus aureus*, causador da mastite em bovinos e outros ruminantes, teve 10% de seu genoma seqüenciado pela geração de GSS e comparado com os clones infectantes de humanos recentemente seqüenciados. Os resultados mostraram numerosas seqüências únicas para o clone RF122 e diferenças notáveis entre os genomas dos clones de diferentes hospedeiros, identificando também fatores específicos ligados com hospedeiros específicos (HERRON *et al.*, 2002).

Com *Cryptosporidium parvum*, agente infeccioso difundido entre humanos e animais, foram geradas 654 GSS correspondentes a mais de 320 Kb do genoma de 10,4 megabases. Um total de 408 seqüências foram montadas em contigs, correspondendo a 2,5% do genoma total do microorganismo, e comparadas em bancos de dados públicos. Um total de 107 contigs genômicos apresentaram similaridade com genes de proteínas, rRNA e tRNA previamente identificados distribuídos nas funções de rotas glicolíticas, metabolismo de DNA, RNA e proteínas, rotas sinalizadoras de tradução, além de seqüências de elementos repetidos como microsátélites e telômeros (LIU *et al.*, 1999).

O resultado do mesmo projeto de *C. parvum* foi comparado com a análise do seqüenciamento por EST da forma esporozoíta do protozoário. O comprimento médio das seqüências de GSS foi de 496pb contra 476pb das EST. A taxa de redundância de seqüências foi maior em EST, 32,3% contra apenas 6% em GSS (408 contigs únicos de 432 clones). Interessantemente, dentre as 408 seqüências únicas de GSS, 26% (107) apresentaram entradas em bancos de dados e, das seqüências únicas de EST, 37% (142 de 384) encontraram similaridade nos bancos de dados (LIU *et al.*, 1999).

A geração de seqüências consiste numa fase inicial de obtenção de informações sobre um genoma. A fase de análise dos dados obtidos constitui a ferramenta primária de pesquisa para identificação e rotulação das seqüências de genes dentro de categorias de um amplo número de espécies e de funções biológicas diversas, e também um importante recurso para a designação da seqüência genômica (QUACKENBUSH *et al.*, 2001). Bancos de dados como o TIGR (<http://www.tigr.org/tdb/tgi.shtml>) constituem uma das coleções de banco de dados espécie-específicos que usa protocolos refinados para analisar seqüências na

tentativa de identificar genes representados por aqueles dados, e de suprir informações adicionais àqueles genes (QUACKENBUSH *et al.*, 2001).

O seqüenciamento do genoma de *T. vivax* é considerado como de alta relevância pelo fato do parasita causar danos econômicos na produção de gado de corte e leiteiro, e no custo dos tratamentos profilático e terapêutico mediante o uso de drogas. Kristjanson *et al.* (1999), estimaram um valor mínimo de U\$ 288 milhões de dólares anuais para pesquisa de uma vacina contra U\$ 1.340 milhões de dólares anuais gastos em custos da doença, os quais são invariavelmente repassados à produção. Além disso, a geração de conhecimento para o esclarecimento da biologia do hemoparasita, inclusive dos pontos de vista sanitário e epidemiológico, apresenta-se como uma importante justificativa ao estudo.

A obtenção de um perfil do genoma de *T. vivax* através da geração de GSS tem seu ineditismo baseado nos poucos estudos de biologia molecular deste protozoário, além de possibilitar a abertura de caminhos para a descoberta de alvos marcadores moleculares espécie-específicos e o desenvolvimento de novos testes diagnósticos moleculares precisos e sensíveis.

O presente trabalho é parte integrante do consórcio Biowebdb (<http://www.biowebdb.org>), o qual visa o desenvolvimento de um sistema de bancos de dados integrado com uma plataforma de análises de genes e genomas. Este consórcio congrega, além do Instituto Oswaldo Cruz (Fiocruz), a Universidade Federal do Rio de Janeiro e a Universidade Federal de Santa Catarina.

2. OBJETIVOS

2.1. Objetivo geral

- Gerar seqüências genômicas da cepa ILDat 2160 de *Trypanosoma vivax* através da técnica de GSS (*Genome Survey Sequences*), e analisar comparativamente estas seqüências com os bancos de dados públicos.

2.2. Objetivos específicos:

- Gerar 250 GSS de formas tripomastigotas da cepa ILDat 2160 de *Trypanosoma vivax*.
- Analisar comparativamente (intraespecífica e interespecificamente) as GSS geradas com seqüências depositadas em bancos de dados públicos.
- Identificar e classificar as GSS geradas por função biológica.

3. MATERIAL E MÉTODOS

3.1. Extração do DNA genômico e construção da biblioteca genômica

O DNA genômico da cepa ILDat 2160 de *T. vivax*, foi gentilmente cedido por Dr. Noel B. Murphy do Trinity College Dublin (Irlanda), tendo sido isolada originalmente no Oeste da África. Aproximadamente 2µg de DNA genômico foram digeridos com enzima de restrição *Sau3AI* (New England BioLabs®) que corta o DNA em fragmentos de tamanhos variados nos sítios de reconhecimento 5'...[^]GATC...3' e 3'...CTAG[^]...5'. O vetor utilizado foi o plasmídeo pUC18 (Sigma®) com 2.686 pares de bases (pb) e digerido com a enzima de restrição *BamH1* (New England BioLabs®) que corta nos sítios 5'...G[^]GATCC...3' e 3'...CCTAG[^]G...5'.

As reações de digestão foram realizadas durante 2 horas a 37°C, sendo os fragmentos de DNA obtidos ligados no vetor plasmidial pUC18 (Sigma®) através da enzima T4 DNA ligase a 16°C por 12 horas. O produto da ligação foi então utilizado na transformação de células competentes (*Escherichia coli*) Dh5α. As células transformadas foram crescidas em placas com meio LB-ágar preparadas com 100µg/mL de ampicilina, 2% de X-Gal e 100mM de IPTG, para seleção de clones com inserto. As placas foram mantidas à 37°C durante 12 horas de acordo com Sambrook e Russel (2001).

As placas com maior número de colônias brancas, positivas para a ligação do inserto, foram selecionadas e congeladas em solução de glicerol (10%) e meio de cultura LB no volume 1:1 a -70°C até o processamento.

A presença de insertos nas colônias selecionadas foi avaliada através das digestões do DNA plasmidial com as enzimas *EcoRI* e *PstI* (New England BioLabs®), por 2 horas a 37°C, sendo os produtos de digestão resolvidos em gel de agarose a

1% em tampão TBE 1x. Após a corrida, o gel foi corado com brometo de etídio (0,1-0,3 $\mu\text{g}/\text{mL}$) por 10 minutos e os produtos visualizados em transiluminador de luz UV e fotografados.

3.2. Semi-normalização da biblioteca genômica

A biblioteca genômica foi diluída nas concentrações 10^{-2} , 10^{-4} , 10^{-6} , 10^{-7} e 10^{-8} visando a obtenção de, no máximo, 500 colônias por placa. Nas quatro primeiras concentrações foram diluídos 10 μL da biblioteca genômica em 990 μL de meio LB e na última concentração 100 μL da mesma. Em seguida as placas foram deixadas por período de 12 horas a 37°C em estufa bacteriológica.

As placas de cada colônia foram transferidas para uma membrana de nitrocelulose de acordo com Sambrook e Russel (2001). Para a desnaturação do DNA, as membranas foram incubadas com uma solução de 0,5M NaOH + 1,5M NaCl durante 5 minutos. Em seguida, as membranas foram incubadas com uma solução de 0,5M Tris HCl pH 7,4 + 1,5M NaCl durante 5 minutos por duas vezes a fim de neutralizar a reação.

As membranas de nitrocelulose foram envolvidas em náilon e colocadas em garrafas para hibridação com 15 mL de solução de pré-hibridação QuickHyb (Stratagene®). Foram utilizados os seguintes oligonucleotídeos como sondas para a normalização: (CA)₁₅, (CT)₁₅, (AAC)₁₀ e TVW (MASIGA *et al.*, 1992); ILO (MASAKE *et al.*, 1997); e, ITS1 e ITS2 (DÁVILA, 2002). A marcação das sondas foi realizada com um volume final de 50 μL por reação com: 5pmol/ μL dos oligonucleotídeos, 10U/ μL da enzima T4 *polinucleotide kinase*, solução tampão 10X e [γ -³²P] ATP radioativo (Pharmacia®) na concentração 0.001mCi/ μL durante 30 minutos a 37°C. As sondas foram adicionadas às garrafas e a reação de hibridação foi realizada por

2 horas e 30 minutos a 65°C. Após a hibridação, as membranas foram lavadas três vezes durante 30 minutos com SSC1X-SDS 0,5% a 55°C e expostas à filme radiográfico (Kodak®) a -70°C por 2-6 horas.

Estes passos foram realizados em sua maioria no Laboratório de Biologia Molecular de Tripanosomatídeos do Instituto Oswaldo Cruz, Fiocruz, no Rio de Janeiro.

3.3. Crescimento bacteriano e extração de DNA plasmidial

Todos os passos seguintes foram realizados nos Laboratório de Protozoologia da UFSC, em Florianópolis.

As colônias transformantes foram trazidas para o Laboratório de Protozoologia da UFSC e cultivadas em placas de 96 poços contendo 3mL de meio LB acrescido de ampicilina na concentração de 100µg/mL, sendo as mesmas identificadas anotando-se os números da biblioteca, placa e a data. As placas foram vedadas com adesivo plástico resistente à álcool, sendo este perfurado com uma agulha em cada poço para facilitar a oxigenação. Foram então incubadas durante 20 horas em estufa bacteriológica à 37°C sob agitação de 100 rpm. Cumprido o tempo o crescimento foi observado através da turbidez do meio e, sendo o crescimento satisfatório, iniciava-se o processo de extração de DNA plasmidial ou, alternativamente, eram armazenadas em geladeira à 4°C por no máximo 12 horas.

A extração de DNA plasmidial foi realizada pelo método de lise alcalina (*mini-prep*) segundo o descrito por Sambrook e Russel (2001).

3.4. Seqüenciamento de DNA

Inicialmente a estratégia utilizada para a obtenção das GSS foi o seqüenciamento das extremidades 5' de cada produto de amplificação.

O seqüenciamento destes produtos foi realizado em um equipamento MegaBACE 1000[®] (Amersham Biosciences, Piscataway), realizando-se a reação de seqüenciamento com o kit DYEnamic[®] ET dye terminator (Amersham Biosciences) conforme especificações do fabricante.

Baseado na experiência prévia do grupo, nós utilizamos a concentração de 3,2 a 5,0 pmol por reação do iniciador M-13forward (5'-CGCCAGGGTTTTCCCAGTCACGAC -3'). Resumidamente, de 2 a 4 μ L do produto PCR foram utilizados diretamente na reação de seqüenciamento. O esquema de amplificação foi de 94°C por 25 segundos seguidos de 30 ciclos com desnaturação de 94°C por 15 segundos, ligação dos iniciadores a 50°C graus por 20 segundos e extensão a 60°C por 1 minuto. Após a reação de seqüenciamento seguiu-se a purificação dos produtos utilizando-se acetato de amônio 750mM e etanol absoluto com o objetivo de se retirar nucleotídeos e iniciadores não incorporados. Após esta purificação os produtos foram eletroinjetados com 3KV por 80 segundos e eletroeluídos por 4 horas a 6KV.

As extremidades 3' dos clones de GSS obtidos foram objeto da continuidade do presente estudo, entretanto os protocolos acima descritos já foram utilizados revelando-se eficazes para geração de GSS a partir desta extremidade dos fragmentos clonados.

3.5. Análise das seqüências obtidas

Foram consideradas como sendo válidas para fins de análise, seqüências de tamanho iguais ou maiores que 250 nucleotídeos e com qualidade *Phred*>20 determinada pelo programa Phred/Phrap/Consed (<http://www.phrap.org>). As GSS de *T. vivax* obtidas no presente estudo foram comparadas à seqüências depositadas em bancos de dados públicos, como por exemplo o GenBank, os quais estão reunidos na plataforma de anotação genômica GARS (Genomic Analysis Resources for Sequence Annotation). Esta plataforma de acesso via internet está em fase de desenvolvimento através de uma parceria entre a UFSC, a UFRJ e a FIOCRUZ/ RJ com o apoio do CNPq.

As análises através da interface Blast do GenBank (<http://www.ncbi.nlm.nih.gov/BLAST/>), utilizaram as rotinas Blastn, Blastx e tBlastx que procuram similaridades com seqüências nucleotídicas e protéicas depositadas no referido banco.

As seqüências geradas que apresentaram baixa ou nenhuma similaridade e estão fora do limite (*cut-off*) estipulado (*e-value* <10⁻⁵) com os dados disponíveis no GenBank foram separadas para análise posterior.

As seqüências que atenderam aos padrões supracitados foram então submetidas à agrupamento e à montagem de *contigs* (agrupamentos de seqüências similares) genômicos, utilizando-se para tanto o programa CAP3 (HUANG e MADAN, 1999). Após esta fase, as seqüências foram então classificadas em categorias funcionais de acordo com as normas do consórcio Gene Ontology (<http://www.geneontology.org>). A estimativa do conteúdo de bases C e G e a presença de janelas abertas de leitura (*open reading frames* – ORF) foram

realizadas com o programa Glimmer (*Gene Locator and Interpolated Markov Modeler*) (DELCHER *et al.*, 1999).

As GSSs foram confrontadas com seqüências genômicas depositadas nos seguintes bancos de dados: Nr do GenBank (NCBI); TrEMBL, Kineto, Minicircles e Uniprot-Sprot do EMBL (*European Molecular Biology Laboratory*); banco de domínios conservados CDD e InterPro; banco de elementos repetitivos RepBASE; banco de dados do genoma de *Gloeobacter violaceus* (CyanoBase); Cyanobacteria (custom subset do GenBank) e Gene Ontology.

Com a finalização da plataforma GARSA e de sua completa disponibilização via internet, as seqüências geradas no presente estudo estarão disponíveis, assim como suas anotações. As seqüências brutas serão depositadas no GenBank.

4. RESULTADOS

4.1. Biblioteca genômica

As digestões do DNA genômico da cepa ILDat 2160 de *T. vivax* realizadas conforme o descrito em 2.1 revelaram que os fragmentos de DNA genômico obtidos apresentavam tamanhos que variaram de 200pb a mais de 3Kb (dados não mostrados).

Após a ligação destes fragmentos em pUC18 e a transformação de células DH5 α , foi obtido um total de 1.086 clones, os quais encontram-se estocados no Laboratório de Protozoologia da UFSC conforme o descrito em 2.1. A digestão dos plasmídeos extraídos das células transformantes revelou a presença de insertos cujo tamanho variou de 600pb a cerca de 3Kb (dados não mostrados).

4.2. Semi-normalização

Após revelação dos filmes radiográficos, as colônias que tiveram resultados negativos para as sondas de hibridação utilizadas foram selecionadas e encaminhadas para seqüenciamento.

As sondas utilizadas hibridaram com regiões conhecidas e altamente repetitivas do DNA de *T. vivax* como o DNA ribossômico, as regiões do gene do mini-exon e de DNA satélite, sendo que sua exclusão permitiram a semi-normalização das bibliotecas e a redução da redundância das seqüências obtidas.

4.2. Seqüenciamento de GSS

O seqüenciamento dos clones obtidos resultou na geração de 456 seqüências de alta qualidade consideradas válidas (Tabela 1) após a avaliação de tamanho e

qualidade e remoção de contaminação vetorial através do pacote Phred/Phrap/Consed. Um total de 630 seqüências de todos os clones gerados foram descartadas por baixa qualidade da seqüência gerada, insertos ou seqüências de tamanho reduzido e/ou a presença de seqüências contaminantes (Tabela 1).

O tamanho das seqüências geradas variou de 100pb para a menor GSS e 1.234pb para a maior, sendo o tamanho médio das GSS válidas obtidas de 300pb, corroborando os resultados de digestão de DNA plasmidial dos clones obtidos.

Tabela 1: Resultados do seqüenciamento de clones de DNA genômico (GSS) da cepa ILDat 2160 de *Trypanosoma vivax*.

Característica	Total	% *
Número de clones seqüenciados	1.086	100,0
Seqüências de alta qualidade	456	41,9
Seqüências únicas (a)	264	
Seqüências agrupadas em clusters	191	
Clusters formados (b)	67	
Número total de seqüências (a + b)	331	
Seqüências descartadas	630	58,1

* O percentual é em relação ao número total de clones seqüenciados

Dentre as 456 GSS válidas, 264 formaram seqüências únicas (*singlets*) e 191 seqüências foram agrupadas em *contigs*, formando um total de 67 *contigs*, totalizando 332 seqüências não redundantes (Tabela 1). As seqüências não redundantes foram submetidas ao alinhamento local através de programas BLAST do GenBank (TblastX, BlastN e BlastX) e comparadas em nível de proteína e DNA quando confrontadas com os 12 bancos de dados utilizados na análise.

O número de entradas obtidas (*hits*) variou dentre os bancos de dados consultados, sendo que o banco nr-NCBI revelou o maior número de entradas (3.302), seguido de modo decrescente pelos bancos TrEMBL (2.934), Kineto (1.111), Swiss-Prot (864), CDD (187) e os demais bancos conforme a Tabela 2.

Tabela 2: Número de entradas (*hits*) obtidas para cada um nos bancos de dados confrontados com as GSS de *Trypanosoma vivax* geradas no presente estudo.

Bancos de Dados		Entradas (hits)	%
1	N (GenBank)	3.302	34,42
2	TrEMBL	2.934	30,58
3	Kineto	1.111	11,58
4	SwissProt	864	9,01
5	InterPro	646	6,73
6	G_200407 (Gene)	469	4,98
7	CDD	187	1,95
8	Gloeobacter	24	0,25
9	Minicircles	22	0,23
10	Synecho	19	0,20
11	Repbase905	11	0,11
12	Cyanobacteria	4	0,04
Total		9.593	

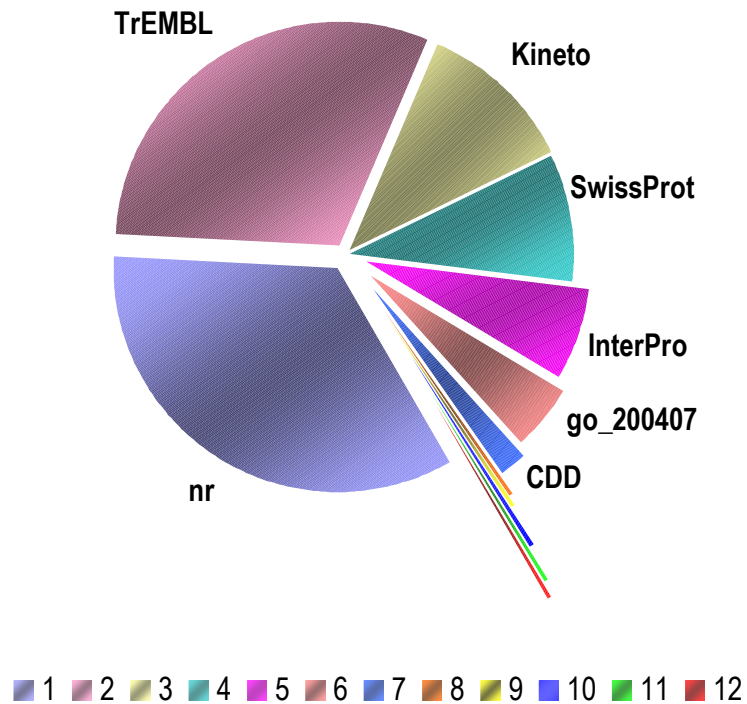


Figura 1: Representação gráfica do número de entradas (*hits*) obtidas para cada um dos bancos de dados confrontados com as GSS de *Trypanosoma vivax* geradas apresentadas na Tabela 2.

4.3. Classificação das Seqüências em Categorias Funcionais

Um total de 108 seqüências com *e-value* superior à $1e < 10^{-5}$ foram analisadas quanto à sua identidade e função. A classificação foi realizada de acordo com o Gene Ontology (Tabela 3), um consórcio que produz um vocabulário controlado para a anotação genômica (ASHBURNER *et al.*, 2000). Algumas das identidades reveladas foram confirmadas com novas análises realizadas utilizando-se um banco de dados de domínios conservados de proteínas (Tabela 4).

Tabela 3. Relação das GSS de *Trypanosoma vivax* identificadas com base em seqüências conhecidas de kinetoplastídeos e outros organismos em bancos de dados públicos (*e-value* limite de $<10^{-5}$).

Função e Clone	Código ou número de acesso	Descrição	Organismo	Banco de dados	P	Escore
Processos Biológicos						
Proteólise						
TVAD007001E12	UBIQ_TRYBB	ubiquitina	<i>T. brucei brucei</i>	sprot	4.0e-10	64
TVAD007011B05	UBIQ_TRYBB	ubiquitina	<i>T. brucei brucei</i>	sprot	4.0e-12	69
TVAD007011H10	UBIQ_TRYCR	ubiquitina	<i>T. cruzi</i>	sprot	3.0e-08	56
Processo de desenvolvimento						
TVAD007005M96	26269	extensina 2	<i>Core eudicots</i>	CDD	1.0e-37	148
TVAD007100A06	24679	atrofina-1	<i>Eutheria (placentals)</i>	CDD	2.0e-06	46
TVAD007120B05	26269	extensina 2	<i>Core eudicots</i>	CDD	9.0e-47	180
Processamento de DNA						
TVAD007003M72	27539	histona H1	<i>Chlamydiaceae</i>	CDD	1.0e-06	46
TVAD007011C07	H3_LEIIN	histona H3	<i>L. infantum</i>	sprot	3.0e-06	49
TVAD007017D02	L27660	histona H3	<i>T. cruzi</i>	kineto	2.0e-06	54
TVAD007018C07	L27659	histona H3	<i>T. cruzi</i>	Kineto	6.0e-11	68
TVAD007100H03	H3_LEIIN	histona H3	<i>L. infantum</i>	sprot	3.0e-07	53
TVAD007001A08	P04915	histona H4	<i>Physarum polycephalum</i>	sprot	2.0e-12	71
TVAD007001E07	Q8WXA9	proteína fator de processamento	<i>Homo sapiens</i>	sprot	4.0e-08	52
Transporte intracelular						
TVAD007001D12	9659	complexo Vps53	<i>Tetraodon nigroviridis</i>	CDD	7.0e-11	63
TVAD007100D12	9659	complexo Vps53	<i>Tetraodon nigroviridis</i>	CDD	2.0e-11	65
Elementos genéticos móveis						
TVAD007002F08	Q8T9M5	Retrotransposon RHS4a	<i>T. brucei</i>	TrEMBL	2.0e-12	72
Metabolismo de nucleotídeos						
TVAD007003M65	Q9XZY8	AMP deaminase	<i>L. major</i>	TrEMBL	3.0e-07	55
Processos fisiológicos						
TVAD007021A03	Q6KEQ0	precursor da alatotropina	<i>Spodoptera frugiperda</i>	TrEMBL	1.0e-06	54
TVAD007021A04	Q6KEQ0	precursor da alatotropina	<i>Spodoptera frugiperda</i>	TrEMBL	9.0e-06	54
TVAD007021A05	Q6KEQ0	precursor da alatotropina	<i>Spodoptera frugiperda</i>	TrEMBL	6.0e-06	51
TVAD007021A06	Q6KEQ0	precursor da alatotropina	<i>Spodoptera frugiperda</i>	TrEMBL	6.0e-06	51
TVAD007021A07	Q6KEQ0	precursor da alatotropina	<i>Spodoptera frugiperda</i>	TrEMBL	8.0e-06	51
TVAD007021A08	Q6KEQ0	precursor da alatotropina	<i>Spodoptera frugiperda</i>	TrEMBL	6.0e-06	51
TVAD007021A11	Q6KEQ0	precursor da alatotropina	<i>Spodoptera frugiperda</i>	TrEMBL	7.0e-06	51
TVAD007021B01	Q6KEQ0	precursor da alatotropina	<i>Spodoptera frugiperda</i>	TrEMBL	6.0e-06	51
TVAD007021B04	Q6KEQ0	precursor da alatotropina	<i>Spodoptera frugiperda</i>	TrEMBL	6.0e-06	51
TVAD007021B05	Q6KEQ0	precursor da alatotropina	<i>Spodoptera frugiperda</i>	TrEMBL	5.0e-05	48

 continuação da tabela 3
Componente celular**Citoesqueleto**

TVAD007002C10	P42528	actina	<i>Dictyostelium discoideum</i>	sprot	4.0e-11	65
TVAD007002C08	P34036	dineína	<i>Dictyostelium discoideum</i>	sprot	1.0e-09	60
TVAD007100B07	Q39565	dineína	<i>Chlamydomonas reinhardtii</i>	sprot	6.0e-41	165
TVAD007200F02	P34036	dineína	<i>Dictyostelium discoideum</i>	sprot	1.0e-09	60

Estrutura flagelar

TVAD007001B03	Q26789	proteína flagelar	<i>T. brucei brucei</i>	sprot	3.0e-13	72
TVAD007001B10	AF005193	proteína flagelar	<i>T. cruzi</i>	kineto	3.0e-28	127
TVAD007010B01	P22225	proteína flagelar	<i>T. brucei brucei</i>	sprot	1.0e-10	64
TVAD007010F02	P22225	proteína flagelar	<i>T. brucei brucei</i>	sprot	1.0e-10	64
TVAD007010F06	AF005193	proteína flagelar	<i>T. cruzi</i>	kineto	2.0e-07	56
TVAD007015C06	P22225	proteína flagelar	<i>T. brucei brucei</i>	sprot	1.0e-10	64
TVAD007015C10	P22225	proteína flagelar	<i>T. brucei brucei</i>	sprot	3.0e-07	52
TVAD007016E11	EMBL:M97548	proteína flagelar	<i>T. cruzi</i>	kineto	7.0e-25	113
TVAD007018B02	M97548	proteína flagelar	<i>T. cruzi</i>	kineto	2.0e-23	109
TVAD007018B10	M97548	proteína flagelar	<i>T. cruzi</i>	kineto	3.0e-16	86
TVAD007018E01	P22225	proteína flagelar	<i>T. brucei brucei</i>	sprot	1.0e-10	64
TVAD007018F12	P22225	proteína flagelar	<i>T. brucei brucei</i>	sprot	3.0e-07	52

Cinetoplasto

TVAD007001B02	M19750	kDNA minicírculo	<i>T. congolense</i>	minicírculos	1.0e-06	43
TVAD007001B09	EMBL:X66923	DNA cinetoplástico	<i>C. fasciculata</i>	kineto	4.0e-06	54
TVAD007002A10	M19178	kDNA minicírculo	<i>T. cruzi</i>	minicírculos	8.0e-06	35
TVAD007002G01	M74195	kDNA minicírculo	<i>Phytomonas</i> sp.	minicírculos	1.0e-06	41
TVAD007002G07	X56188	kDNA minicírculo	<i>T. cruzi</i>	kineto	2.0e-07	58
TVAD007005166	X56188	kDNA minicírculo	<i>T. cruzi</i>	kineto	1.0e-07	58
TVAD007100B10	X56188	kDNA minicírculo	<i>T. cruzi</i>	kineto	2.0e-07	58
TVAD007100C10	M19750	kDNA minicírculo	<i>T. congolense</i>	minicírculos	5.0e-06	40
TVAD007120A01	X04690	kDNA minicírculo	<i>T. cruzi</i>	kineto	1.0e-06	41
TVAD007120B02	AA052853	kDNA minicírculo	<i>T. brucei</i>	kineto	2.0e-14	82
TVAD007200C02	AQ904936	kDNA minicírculo	<i>T. cruzi</i>	kineto	1.0e-07	58
TVAD007200D01	M15511	kDNA minicírculo	<i>T. congolense</i>	minicírculos	6.8e-08	37
TVAD007200D07	M15511	kDNA minicírculo	<i>T. congolense</i>	minicírculos	2.0e-07	46
TVAD007200G01	M19190	kDNA minicírculo	<i>T. cruzi</i>	minicírculos	1.0e-06	40
TVAD007700N94	M15511	kDNA minicírculo	<i>T. congolense</i>	minicírculos	5.0e-08	48
TVAD007800117	M74195	kDNA minicírculo	<i>Phytomonas</i> sp.	minicírculos	5.0e-06	37
TVAD007800165	AF401099	kDNA minicírculo	<i>T. congolense</i>	kineto	4.0e-07	56

Membrana

TVAD007003K04	5010	proteína 45/70kd	<i>Tymovirus</i>	CDD	7.0e-09	59
TVAD007003M83	Q26721	adenilato ciclase	<i>T. brucei brucei</i>	sprot	2.0e-29	126

Núcleo

TVAD007001H10	Q9NU22	midasina	<i>H. sapiens</i>	sprot	1.0e-20	97
---------------	--------	----------	-------------------	-------	---------	----

continuação da tabela 3

Função molecular**Defesa**

TVAD007002H05	P21441	proteína resistência multidroga/ p-glicoproteína	<i>L. tarentolae</i>	sprot	3.0e-44	160
TVAD007003M01	Q8WPT4	glicoproteína variante superfície/ VSG(H25N7.03)	<i>T. brucei</i>	trembl	3.0e-07	55
TVAD007003M12	Q8WPT4	glicoproteína variante superfície/ VSG (H25N7.03)	<i>T. brucei</i>	trembl	2.0e-07	125

Replicação de DNA

TVAD007002D08	AQ657065	RNA helicase II	<i>T. brucei</i>	kineto	4.0e-06	52
---------------	----------	-----------------	------------------	--------	---------	----

Enzima

TVAD007001C12	AF359249	treonina quinase	<i>T. brucei</i>	sprot	1.0e-18	91
TVAD007002C04	IPR006662	tioredoxina		interpro		
TVAD007002G11	Q9GRP3	possível fosfatase	<i>L. major</i>	trembl	1.0e-11	68
TVAD007003K30	AA720380	ATP sintetase	<i>T. brucei</i>	CDD	1.0e-07	52
TVAD007100B12	P37798	acetil CoA carboxilase (subunidade biotina)	<i>P. aeruginosa</i>	sprot	2.0e-07	53
TVAD007100E07	IPR001211	fosfolipase A2		interpro		
TVAD007200H05	25296	proteína quinase	<i>L. infantum</i>	CDD	4.0e-06	49
TVAD007700117	HL3B_ARATH	fosfopantotenoilcisteína sintetase	<i>A. thaliana</i>	sprot	6.0e-17	84

Metabolismo

TVAD007001A02	10280	fosfoglicerato mutase	<i>T. brucei</i>	sprot	2.0e-11	67
---------------	-------	-----------------------	------------------	-------	---------	----

Reparo de DNA

TVAD007200E08	15765	proteína PSP1 C-terminal	<i>Bacillus subtilis</i>	CDD	9.0e-19	89
---------------	-------	--------------------------	--------------------------	-----	---------	----

Transcrição e regulação

TVAD007001A12	Q8RWB4	proteína domínio ring finger	<i>Arabidopsis thaliana</i>	TrEMBL	1.0e-09	64
TVAD007001G04	5398	proteína domínio WD	<i>L. major</i>	CDD	3.0e-08	51

Transporte de íons

TVAD007001F09	Q86AJ7	proteína transportadora de zinco	<i>Oryza sativa</i>	TrEMBL	4.0e-06	51
TVAD007001F10	IPR000345	citocromo C		interpro		
TVAD007100E01	NP-061865	proteína canal de potássio	<i>H. sapiens</i>	CDD	2.0e-11	70
TVAD007200F07	IPR000345	citocromo C		interpro		
TVAD007021A01	Q7Z8K5	magnésio superóxido desmutase	<i>Antrodia camphorata</i>	TrEMBL	6.0e-06	51
TVAD007021A02	Q7Z8K5	magnésio superóxido desmutase	<i>Antrodia camphorata</i>	TrEMBL	6.0e-06	51

Proteínas hipotéticas

TVAD007001A09	BA000045	proteína hipotética	<i>G. violaceus</i>	cyanobacteria.gig	3.0e-04	37
TVAD007001E04	BX530564	proteína hipotética	<i>L. brasiliensis</i>	kineto	6.0e-06	52
TVAD007002A05	BH196243	proteína hipotética	<i>T. cruzi</i>	nt	6.0e-06	28
TVAD007002C05	AQ944111	proteína hipotética	<i>T. brucei</i>	kineto	1.0e-11	72
TVAD007002D01	AL499624	proteína hipotética	<i>L. brasiliensis</i>	nt	4.0e-17	86
TVAD007002D02	Q9NJ10	proteína hipotética	<i>L. donovani</i>	nt	7.0e-22	104
TVAD007002F09	AI075532	proteína hipotética	<i>T. cruzi</i>	nt	2.0e-07	50
TVAD007002G05	AI667941	proteína hipotética	<i>T. cruzi</i>	nt	4.0e-10	61
TVAD007003K31	AC096672	proteína hipotética	<i>T. brucei</i>	nt	1.0e-16	88

continuação da tabela 3

TVAD007003K34	AQ952263	proteína hipotética	<i>T. brucei</i>	nt	3.0e-09	58
TVAD007014A07	Q7YVG7	proteína hipotética	<i>T. brucei</i>	TrEMBL	7.0e-06	50
TVAD007015A10	Q7YVG7	proteína hipotética	<i>T. brucei</i>	TrEMBL	2.0e-09	62
TVAD007015E06	AZ218285	proteína hipotética	<i>T. brucei</i>	nt	1.0e-06	50
TVAD007017G01	Q7YVG7	proteína hipotética	<i>T. brucei</i>	TrEMBL	3.0e-06	52
TVAD007017G02	Q7YVG7	proteína hipotética	<i>T. brucei</i>	TrEMBL	1.0e-06	53
TVAD007021A12	EM_NEW:CF888951	proteína hipotética	<i>T. cruzi</i>	kineto	3.0e-22	107
TVAD007100D04	EMBL:AL456009	proteína hipotética	<i>T. brucei</i>	kineto	3.0e-12	72
TVAD007120A03	EMBL:AA052853	proteína hipotética	<i>T. brucei brucei</i>	kineto	9.0e-19	96
TVAD007120A05	EMBL:AA052853	proteína hipotética	<i>T. brucei brucei</i>	kineto	3.0e-18	96
TVAD007120A09	EMBL:AA052853	proteína hipotética	<i>T. brucei brucei</i>	kineto	5.0e-13	76
TVAD007120A10	EMBL:BF299436	proteína hipotética	<i>T. cruzi</i>	kineto	2.0e-06	54
TVAD007120B04	EMBL:AA052853	proteína hipotética	<i>T. brucei brucei</i>	kineto	7.0e-10	66
TVAD007200E05	AI707443	proteína hipotética	<i>T. brucei</i>	kineto	7.0e-07	46

Número de seqüências com similaridade significativa usando o *e-value* <10⁻⁵ como limite: 108 seqüências

Número de seqüências sem entradas (*hits*) significantes: 205 seqüências

Número de seqüências sem *hits*: 21 seqüências

Total: 331 seqüências

Nos tópicos seguintes, passamos a descrever as seqüências com similaridade significativa sob o *e-value* <10⁻⁵ nas categorias funcionais padronizadas pelo Gene Ontology (Figura 2) e suas descrições em espécies relacionadas de kinetoplastídeos assim como em outros organismos com os quais as GSS geradas no presente estudo apresentaram identidade.

4.4.1. Processo Biológico:

Nesta categoria estão incluídas as seqüências cuja identidade denota o(s) processo(s) biológico(s) no qual o gene ou o seu produto participa no metabolismo do organismo. Do total de seqüências categorizadas, 25% foram classificadas nesta categoria (Tabela 5).

4.4.1.1. Proteólise:

Três seqüências (TVAD007001E12, 11B05, 11H10) são similares à ubiquitina, uma proteína conservada em eucariotos que regula processos celulares como apoptose, transcrição e progressão do ciclo celular (ARCHIBALD, TEH, KEELIN, 2003). Os genes ligados a proteína variam em número e tamanho e são repetitivos em tripanossomatídeos, *T. cruzi* apresenta número maior de genes que *Leishmania donovani* reunidos em segmento de 27 Kb (KIRCHHOFF *et al.*, 1987; AJIOKA *et al.*, 1988).

4.4.1.2. Processo de Desenvolvimento:

As seqüências TVAD007005M96 e TVAD0007120B05 apresentaram alta similaridade com a extensina, uma proteína altamente glicosilada, rica em repetições do pentapeptídeo serina (01 unidade) e prolina (04 unidades), e constitui mais de 20% do peso seco da parede celular de plantas atuando na resposta ao estresse mecânico por pragas e no desenvolvimento do pistilo (GOLDMAN *et al.*, 1992). A proteína também é encontrada na parede vegetativa e gamética da alga verde *Chlamydomonas reinhardtii* como principal componente (WOESSNER, GOODENOUGH, 1989; WOESSNER *et al.*, 1994) e no embrião de *Volvox carterii* controlando a biogênese da matriz extracelular (ERTL *et al.*, 1992).

A seqüência TVAD007100A06 é similar à atrofina-1, proteína com excesso de glutamina distribuída no núcleo e citoplasma e causadora de desordens degenerativas neuronais (MARGOLIS *et al.*, 1996; WOOD *et al.*, 2000). Em *Drosophila*, uma proteína tipo atrofina está associada à segmentação correta durante a embriogênese e ao desenvolvimento de membros (ERKNER *et al.*, 2002).

4.4.1.3. Processamento de DNA:

Cinco seqüências (TVAD007001A08, 003M72, 011C07, 017D02, 018C07, 100H03) são similares às histonas H1, H3 e H4. As histonas de tripanossomatídeos apresentam propriedades bioquímicas diferentes das histonas de eucariotos superiores e por este motivo promovem uma interação fraca entre as histonas e o DNA tornando a cromatina menos compacta na regulação da expressão de genes (BENDER, BETSCHART, BECKER, 1992; HECKER *et al.*, 1994).

As histonas H1 de tripanossomos africanos são encontradas em número de três subtipos e duas unidades transcricionais nestes organismos devido a substituições e deleções na seqüência do gene (GRUTER, BETSCHART, 2001). Em *T. cruzi*, foi observado que a fosforilação das proteínas H1 é paralela ao seu ciclo de vida, histonas de formas tripomastigotas são estruturas fosforiladas enquanto as de formas epimastigotas não são fosforiladas (MARQUES PORTO *et al.*, 2002).

Além da função de montagem do nucleossomo, as histonas exercem outras funções na gênese e organização dos cromossomos. A histona H3 em humanos está envolvida no comportamento do centrômero, que compartilha dois domínios similares (62% de identidade) ao núcleo da proteína H3 (SULLIVAN, HECHENBERGER, MASRI, 1994). Em levedura, o gene CSE4 essencial para segregação das cromátides na mitose codifica uma proteína com domínio também similar (64% de identidade) a histona H3, sugerindo a atuação desta no centrômero (STOLER *et al.*, 1995).

O conjunto de proteínas histonas se mostra divergente em eucariotos inferiores (*Entamoeba* sp., *Trichomonas* sp., *Trypanosoma* sp., *Leishmania* sp., *Crithidia* sp., *Giardia* sp., dinoflagelados e ciliados), enquanto que, nos eucariotos superiores é conservado. A fixação evolutiva inicial ocorreu em eucariotos

intermediários como os platelmintos *Echinococcus granulosus* e *Fasciola hepatica* que apresentam histonas H1 similares as de protistas e histonas H2A, H2B, H3 e H4 similares as de eucariotos superiores (GALINDO *et al.*, 2004).

A seqüência TVAD007001E07 é similar a proteína de processamento SR rica em repetições terminais de serina (S) e arginina (R) descrita em humano. A proteína SR foi identificada em *T. cruzi* e apresentou função de processamento em ensaio com células HeLa. A estrutura da proteína também compartilha componentes comuns à proteína de humano e *Schizosaccharomyces pombe* sugerindo o surgimento do mecanismo de processamento na linhagem eucariótica através de eucariotos inferiores (PORTAL *et al.*, 2003).

4.4.1.4. Elementos genéticos móveis:

A seqüência TVAD007002F08 é similar ao RHS (*retrotransposon hot spot*), um gene situado à jusante dos sítios de expressão de glicoproteínas de superfície localizados na região telomérica. O gene é comum em tripanossomatídeos, no genoma de *T. brucei* são encontrados cerca de 280 genes dessa família distribuídos no cromossomo I (homólogos aos de *T. cruzi*), no cromossomo Ia e II (BRINGAUD *et al.*, 2002).

4.4.1.5. Transporte intracelular:

Os clones da seqüência TVAD007001D12 e sua seqüência reversa TVAD007100D12 apresentam similaridade à proteína vacuolar VPS53. Esta proteína é uma subunidade de um complexo protéico encontrado em *Saccharomyces cerevisiae* responsável pela reciclagem de vesículas endossômicas (LEMMON, TRAUB, 2000; CONIBEAR, STEVENS; 2000). O complexo não é descrito em tripanossomatídeos, mas outra proteína vacuolar VPS descrita em levedura

recentemente como mediadora da interação do complexo é altamente similar a proteína dinamina remodeladora de membranas encontrada em *Leishmania major* e *T. brucei* (CONIBEAR, CLECK, STEVENS, 2003; MORGAN, GOULDING, FIELD, 2004).

4.4.1.6. Metabolismo de nucleotídeos:

A seqüência TVAD007003M65 é similar a AMP deaminase. A proteína em humanos é caracterizada pelo domínio C-terminal conservado e domínio N-terminal diversificado que influencia no comportamento catalítico, interações proteína-proteína e distribuição intracelular (SABINA, MAHNKE-ZIZELMAN; 2000). A proteína é descrita em tripanossomatídeos ao lado de outras enzimas responsáveis pelo metabolismo de purinas e pirimidinas, porém em quantidades insignificantes (HASSAN, COOMBS, 1986).

4.4.1.7. Processo Fisiológico:

Dez seqüências (TVAD007021A03, 021A04, 021A05, 021A06, 021A07, 021A08, 021A11, 021B01, 021B04, 021B05) apresentam similaridade com alatotropina. O neuropeptídeo é encontrado em *Spodoptera frugiperda* regulando a produção de hormônios juvenis e tem expressão no cérebro, nervos dos músculos viscerais, células endócrinas do intestino e hemolinfa do inseto (ABDEL-LATIEF *et al.*, 2000; STAY, 2000). O neuropeptídeo também ocorre em *Drosophila* atuando no desenvolvimento e é estruturalmente relacionado a peptídeos encontrados em moluscos e anelídeos (ELEKONICH, HORODYSKI; 2003). A proteína não é descrita em protozoários.

4.4.2. Componente Celular:

Nesta categoria estão incluídas as seqüências cuja identidade se refere ao local na célula onde os genes produzidos são ativados. 33,33% das seqüências foram classificadas nesta categoria (Tabela 5).

4.4.2.1. Citoesqueleto:

A seqüência TVAD007002C10 é similar à actina, uma proteína distribuída no citoesqueleto de tripanossomatídeos responsável pela arquitetura e mobilidade da célula (MORTARA, 1989). A distribuição da proteína em *T. brucei* varia com as formas de vida. Nas formas procíclicas a distribuição se mantém no citoesqueleto, enquanto que, nas formas sanguíneas a distribuição concentra-se no flagelo. Essa distribuição maior no flagelo é atribuída ao tráfego vesicular na bolsa flagelar e atividade endocítica, funções para as quais a actina também contribui (GARCIA-SALCEDO *et al.*, 2004).

Três seqüências (TVAD007002C08, 100B07, 200F02) são similares à dineína, proteína citoplasmática com atividade ATPase que atua como propulsora na mobilidade de vesículas e microtúbulos (KOONCE, GRISSOM, McINTOSH, 1992). As dineínas são encontradas em organismos com cílios e flagelos num número estimado de quatorze genes que expressam diferentes formas de proteínas especializadas de acordo com sua estrutura (ASAI, WILKES, 2004).

4.4.2.2. Estrutura flagelar:

Doze seqüências são similares à proteína flagelar PFR (*paraflagellar rod*) (TVAD007001B03, 001B10, 010B01, 010F02, 010F06, 015C06, 015C10, 016E11, 018B02, 018B10, 018E01, 018F12). As proteínas são intraflagelares presentes ao

longo da estrutura do axonema encontradas nos tipos PFR-A e PFR-C e limitadas aos grupos dos euglenóides, dinoflagelados e kinetoplastídeos. Em *T. brucei*, foi verificado que os genes da proteína são cluster repetidos, o primeiro gene difere dos consecutivos que são iguais entre si (FLORIN, RUDOLF, SEEBECK, 1994).

4.4.2.3. Cinetoplasto:

Dezessete seqüências são similares ao DNA cinetoplástico, sendo dezesseis delas similares a kDNA de minicírculo (TVAD007001B02, 001B09, 002A10, 002G01, 002G07, 005166, 100B10, 100C10, 120A01, 120B02, 200C02, 200D01, 200D07, 200G01, 700N94, 800117, 800165). O kDNA de minicírculo compreende moléculas que variam de 0,5 a 1,5 Kb de tamanho e atuam em funções como a codificação de gRNAs (RNAs guias), moléculas envolvidas na edição de RNA, um processo restrito a tripanossomatídeos (THERTULEIN, SIMPSON-HAIDARIS, HAIDARIS, 1994).

4.4.2.4. Membrana:

A seqüência TVAD007003M83 é similar a adenilato ciclase. A proteína em tripanossomatídeos possui um domínio N-terminal extracelular e um domínio intracelular catalítico altamente conservado que difere de mamíferos pela presença de dois subdomínios homólogos. O gene é localizado no sítio de expressão de glicoproteína de superfície (VSG) e similar aos de *T. brucei*, *T. congolense*, *T. vivax* e *T. mega* (ALEXANDRE *et al.*, 1990; NAULA, SEEBECK, 2000).

4.4.2.5. Núcleo:

A seqüência TVAD007001H10 é similar à midasina. A proteína possui domínios característicos AAA conservados idênticos aos de *Drosophila*

melanogaster, *Arabidopsis thaliana*, *Schizosaccharomyces pombe*. Esses domínios AAA (ATPase Associada a várias Atividades celulares) participam nas funções de proteólise, modelamento de proteínas, tráfego em membranas, replicação de DNA, metabolismo de íons e mobilidade intracelular (GARBARINO, GIBBONS, 2002). No núcleo, as midasinas agem como chaperonas na montagem e desmontagem de complexos moleculares (VALE, 2000).

4.4.3. Função Molecular:

Nesta categoria estão incluídas as seqüências cuja identidade se refere às atividades bioquímicas do produto dos genes. 20,37% das seqüências foram classificadas nesta categoria (Tabela 5).

4.4.3.1. Defesa:

A seqüência TVAD007002H05 é similar a glicoproteína do sistema ABC de resistência multidroga. Membros dessa família de glicoproteínas têm sido identificados em *Plasmodium* spp., *Leishmania* spp. e *Entamoeba* spp. cujo número de genes e sítios de expressão implicam na resistência a drogas em cada um desses gêneros (BORST, OUELLETTE, 1995; ULLMAN, 1995).

As seqüências TVAD007003M01 e TVAD007003M12 são similares a glicoproteínas variante de superfície (VSG). Os genes dessas glicoproteínas são expressos em sítios teloméricos cuja ativação de um novo gene depende do desligamento ou deleção de um gene VSG antigo (RUDENKO *et al.*, 1998).

4.4.3.2. Replicação de DNA:

A seqüência TVAD007002D08 é similar a RNA helicase II. O gene da proteína em *T. cruzi* é descrito como uma seqüência longa de 2.703 pares de bases que codifica uma proteína de 901 aminoácidos e sete a oito motivos conservados que caracterizam a família DEAD-box (LORENZI, VÁZQUEZ, LEVIN, 2000; DIAZ ANEL *et al.*, 2000). A função descrita dessas proteínas em *T. brucei* é controle do metabolismo de ácidos nucléicos nuclear, e em tripanossomatídeos em geral, controle da edição de RNA mitocondrial e nuclear (MISSEL. GORINGER, 1994; MISSEL *et al.*, 1999; WORTHEY *et al.*, 2003).

4.4.3.3. Enzima:

As seqüências TVAD007001C12 e TVAD007200H05 são similares a quinases. As proteínas quinases são caracterizadas pela presença de domínios catalíticos e são divididas em treoninas quinases e tirosinas quinases (HANKS, HUNTER; 1995). Em *T. brucei*, foi demonstrado que as quinases são requeridas durante o ciclo de desenvolvimento sugerindo a função de controle da proliferação e diferenciação de protozoários (PARSONS, VALENTINE, CARTER; 1993). Uma dessas quinases, a proteína TBPK50, corrobora a função pois apresenta similaridade à proteína Orb6 de *Schizosaccharomyces pombe* descrita como controladora do ciclo celular (VERDE, WILEY, NURSE; 1998 e GARCIA-SALCEDO *et al.*, 2002). A função de proliferação também foi verificada em *T. cruzi* através da atuação de fosfatidilinositol quinase e quinase B na invasão por tripomastigotas em células Vero não fagocitárias e macrófagos humanos (WILKOWSKY *et al.*, 2001),

A seqüência TVAD007002C04 é similar a tioredoxina. A enzima atua em proteínas disulfídicas como oxiredutora cuja função em tripanossomatídeos era atribuída a tripanotiona redutase (LÜDEMANN *et al.*, 1998; FRIEMANN *et al.*, 2003).

A seqüência TVAD007002G11 apresenta similaridade com uma possível fosfatase. Em tripanossomatídeos, a função das fosfatase não é bem definida. Mas, um ensaio com ácido ocadáico em *T. brucei*, uma molécula inibidora de atividade fosfatase, sugere que a enzima esteja envolvida em atividades de coordenação na mitose, divisão de DNA mitocondrial e citocinese (DAS *et al.*, 1994). Outro estudo recente revela a importância da proteína fosfatase 2As na transformação das formas celulares de *T. cruzi* (GONZALEZ *et al.*, 2003).

A seqüência TVAD007003K30 apresenta similaridade com ATP sintetase. A enzima descrita em *T. brucei* possui um motivo ligação de membrana e outro catalítico regulado por peptídeo inibidor, cuja remoção pode diminuir até três vezes sua atividade (RILO, CATALDI DE FLOMBAUM, STOPPANI, 1989; CHI, CHOI, WILLIAMS, 1996).

A seqüência TVAD007001A07 é similar a fosfolipase A-1. A enzima nas formas sangüíneas de tripanossomos africanos é utilizada para o metabolismo de lisofosfolípídeos exógenos capturados de fluídos dos tecidos hospedeiros (MELLORS, SAMAD, 1989; BOWES *et al.*, 1993). A alta atividade de fosfolipases está associada patogenicidade destes organismos, em *T. cruzi*, parasita intracelular, atribui-se a fosfolipase A-1 o papel de mecanismo de ruptura de membrana das células hospedeiras (WAINSZELBAUM *et al.*, 2001).

A seqüência TVAD007700117 é similar ao gene da fosfopantotenoilcisteína sintase que em *Arabidopsis thaliana* está envolvido no crescimento da planta e no controle ao estresse salino e osmótico (ESPINOSA-RUIZ *et al.*, 1999). A enzima

também é conhecida pela participação na biossíntese de pantotenato e encontrada em organismos como levedura, planta, humano e bactéria; neste último apresenta diferenças na estrutura quaternária da proteína que são úteis a ação de componentes antibacterianos (O'TOOLE, CYGLER, 2003).

4.4.3.4. Metabolismo:

A seqüência TVAD007100B12 é similar a fosfoglicerato mutase. A enzima em tripanossomatídeos atua sem cofator contrastando com as enzimas encontradas em todos os vertebrados. A comparação da seqüência do gene da enzima de *T. brucei* com outros organismos revela que há similaridade maior com plantas (59-62%), seguido de eubactérias (29-35%) e invertebrados (em torno de 28%) (CHEVALIER *et al.*, 2000).

4.4.3.5. Reparo de DNA:

A seqüência TVAD007200E08 apresenta similaridade com a PSP1 (*paraspeckle protein*), umas das proteínas repressoras de mutações na subunidade alfa da DNA polimerase de *Saccharomyces cerevisiae* (FORMOSA, NITTIS, 1998). A proteína também é descrita em humanos e apresenta alta similaridade com a PSP1 descrita em rato (95% de identidade); entretanto, a mesma similaridade não ocorre quando são comparadas entre si as proteínas PSP1 de *Drosophila melanogaster* e *Caenorhabditis elegans* (FOX *et al.*, 2000).

4.4.3.6. Transcrição:

A seqüência TVAD007001A02 é similar à proteína que apresenta domínio *ring finger* ligação de zinco que atua em eventos de ubiquitinação. Um exemplo de

proteína com este tipo de domínio é a ubiquitina ligase que reconhece substratos específicos e promove sua ligação à ubiquitina (JOAZEIRO *et al.*, 1999; FREEMONT, 2000). Estes domínios ligação de zinco também mediam a interação de proteínas em diversos organismos em outras funções como as de: proteínas oncogênicas humanas, proteínas patogênicas virais, proteínas de vias de tradução e da biogênese de peroxissomos (SAURIN *et al.*, 1996).

A seqüência TVAD007001G04 é similar a proteína domínio WD, uma unidade conservada com cerca de 40 aminoácidos em até quatro repetições seguida de um final triptofano (W) e ácido aspártico (D). O domínio foi identificado em organismos como levedura, *Caenorhabditis elegans* e *Arabidopsis thaliana* e está envolvido em funções como modulação na tradução de sinais, processamento de DNA e montagem do citoesqueleto (SMITH *et al.*, 1999).

4.4.3.7. Transporte de íons:

Dois seqüências (TVAD007001F10 e 200F07) são similares ao citocromo C redutase. O complexo citocromo C em tripanossomatídeos está relacionado ao processo de respiração, fato verificado pela presença de subunidades da enzima nas formas procíclicas de *T. brucei* e promastigotas de *Leishmania mexicana amazonensis*, e pela ausência em *Phytomonas serpens*, organismo cuja respiração é deficiente (MASLOV *et al.*, 2002). A enzima é presente em menor quantidade nas formas sanguíneas de *T. brucei* sugerindo que o suprimento de energia nesse estágio ocorra pela glicólise (PRIEST, HAJDUK, 1994).

A seqüência TVAD007001F09 apresenta similaridade com uma proteína transportadora de zinco encontrada na membrana de *Arabidopsis thaliana*. Em tripanossomatídeos, não há descrição de proteínas de membrana com o íon zinco, a

presença de íons é mais comum em proteínas internas como a enolase de *T. brucei* envolvida na glicólise (DA SILVA GIOTTO *et al.*, 2003).

A seqüência TVAD007001F10 apresenta similaridade com proteína canal de potássio. Em *T. cruzi*, o canal de potássio é presente na membrana plasmática ajudando a atividade de ATPase na regulação do pH (VAN DER HEYDEN, DOCAMPO, 2000).

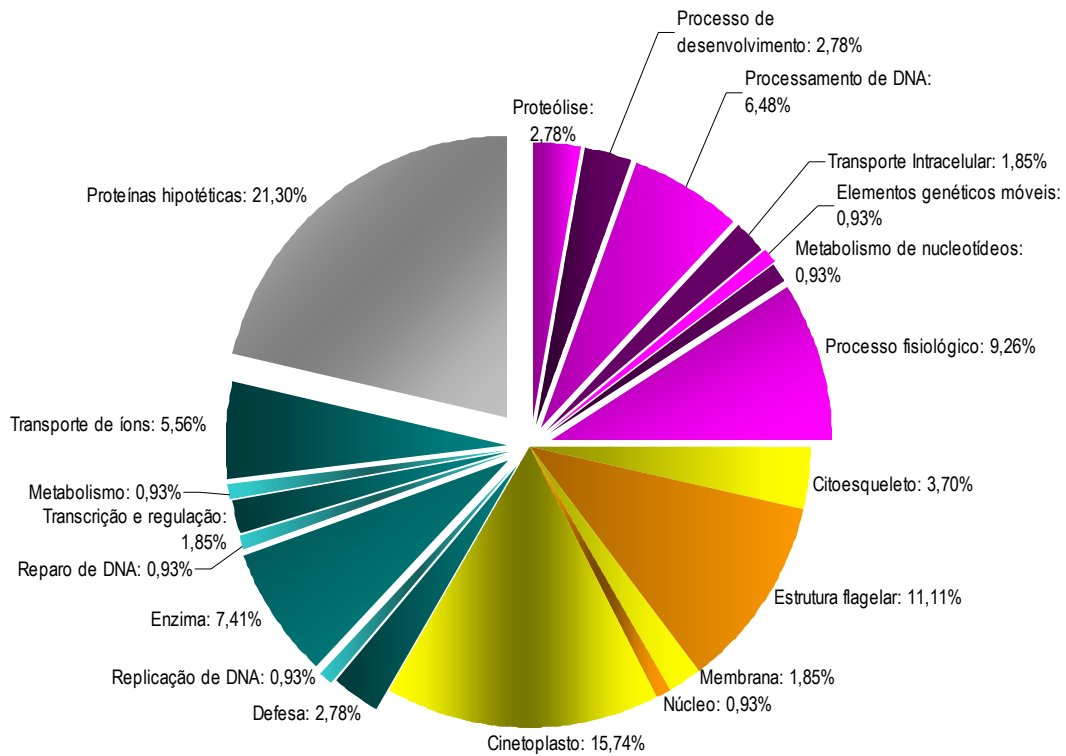
Cabe salientar que as GSS geradas que apresentaram alta qualidade, mas que seus alinhamentos com seqüências depositadas em bancos de dados públicos utilizados neste estudo ficaram com o *e-value* $>10^{-5}$, assim como as GSS que não apresentaram similaridade com nenhuma seqüência depositada nestes bancos através da rotina de alinhamento local BLAST, serão novamente analisadas através do método probabilístico construído com base no modelo de Markov, o que será objeto de um novo estudo e será realizado em colaboração com o IOC/Fiocruz e com a UFRJ.

Tabela 4. Distribuição das identidades das GSS geradas nos banco de dados de domínios de proteínas CDD e InterPro, confirmadas através do banco de dados Blocks / Prints e análises feitas com os programas Glimmer e TestCode.

Clone	Descrição	CDD			InterPro		Blocks/Prints	Glimmer		TestCode	
		acesso	e-value	escore	acesso	e-value	identidade	e-value	ORF		gene
TVAD007001A02	fosfoglicerato mutase	10280	2.0e-24	105	IPR001345	1.4e-08	bi/ fosfoglicerato mutase	8.6e-15	sim	não	não codificante
TVAD007001A08	histona H4	14811	9.0e-17	81	IPR001951	2.1e-12	-		sim	sim	codificante
TVAD007001A12	proteína domínio <i>ring finger</i>	14831	3.0e-11	62	IPR001841	1.3e-07	proteína tipo <i>Zn-finger</i> TRAF	2e-05	sim	não	não codificante
TVAD007001B03	proteína flagelar	16444	4.0e-07	47	-		assinatura orexina (hipocretina)	0.12	sim	não	não codificante
TVAD007001B09	DNA cinetoplástico				IPRO01211		receptor ácido lisofosfatidico	0.3	sim	não	não codificante
TVAD007001C12	treoninas quinase	25296	3.0e-22	98	IPR000719	2e-32	-		sim	não	não codificante
TVAD007001D12	complexo VPS53	9659	9.0-12	63	-		proteína tipo VPS53	4.9e-07	sim	não	codificante
TVAD007001E07	proteína de processamento	25873	2.0e-06	47	-		receptor A ácido gama aminobutírico	0.0097	sim	não	não codificante
TVAD007001E12	ubiquitina	25454	1.0e-07	50	-		-		sim	não	não codificante
TVAD007001H10	midasina	14393	7.0e-20	90	-		-		sim	não	não codificante
TVAD007002C04	tio redoxina				IPR006662		proteína tipo 2 policistica renal	0.027	sim	não	não codificante
TVAD007002C08	dineína	26007	7.0e-14	70			-		sim	não	não codificante
TVAD007002C10	actina	22736	5.0e-10	57	IPR004000	8.3e-06	-		sim	não	não codificante
TVAD007002D08	RNA helicase II	10384	1.0e-07	49			-		sim	não	não codificante
TVAD007002H05	proteína de resistência multidroga	10852	1.0e-22	100	IPR003439	1.0e-15	-		sim	não	não codificante
TVAD007003K04	proteína 45/70Kd Tymovirus	5010	2.0e-09	58	IPR000276		-		sim	não	não codificante
TVAD007003K30	ATP sintetase	17083	9.0e-09	53	-		ATPase H ⁺ -transportadora	0.0045	sim	não	não codificante
TVAD007003M72	histona H1	27539	1.0e-06	46			-		sim	sim	codificante

continuação da tabela 4

TVAD007005M96	extensina 2	26497	4.0e-42	163	IPR003832	5.4e-07	assinatura extensina	2.1e-13	sim	não	não codificante
TVAD007010B01	proteína flagelar	16444	5.0e-06	44	-	-	-	-	sim	não	-
TVAD007010F02	proteína flagelar	16444	5.0e-06	44	-	-	-	-	sim	não	-
TVAD007011B05	ubiquitina	5394	3.0e-08	51	-	-	-	-	sim	não	-
TVAD007011C07	histona H3	383	6.0e-07	47	-	-	-	-	sim	não	não codificante
TVAD007015C06	proteína flagelar	16444	5.0e-06	44	-	-	-	-	sim	não	-
TVAD007018E01	proteína flagelar	16444	5.0e-06	44	-	-	-	-	sim	não	-
TVAD007100A06	atrofina 1	24679	2.0e-06	46	-	-	HTLV Tax	0.064	sim	sim	codificante
TVAD007100B07	dineína	26007	7.0e-47	180	-	-	-	-	sim	não	não codificante
TVAD007100B12	acetil CoA carboxilase	13908	2.0e-12	65	-	-	-	-	sim	não	não codificante
TVAD007100D12	complexo VPS53	9659	2.0e-12	65	-	-	-	-	sim	não	codificante
TVAD007100E07	fosfolipase A2				IPR001221		-	-	sim	não	não codificante
TVAD007100H03	histona H3	383	7.0e-08	50	IPR009072	7.8e-06	-	-	sim	sim	não codificante
TVAD007120B02	kDNA minicírculo				IPR006209		-	-	sim	não	não codificante
TVAD007120B05	extensina 2	26269	9.0e-47	180	IPR003832	7.0e-07	-	-	sim	não	não codificante
TVAD007200E08	proteína PSP1 C-terminal	15765	9.0e20	89	IPR007557	2.1e-17	-	-	sim	não	não codificante
TVAD007200F02	dineína	26007	2.0e-13	68	-	-	-	-	sim	não	não codificante
TVAD007200H05	proteína quinase	25296	4.0e-07	49	IPR000719		quinase MAP	0.0045	sim	sim	codificante
TVAD007700117	fosfopantotenoilcisteína sintetase	10326	1.0e-11	62	IPR003382	2.5e-17	-	-	sim	não	não codificante
TVAD007001G04	proteína domínio WD	5398	3.0e-08	51	IPR001680	2.9e-10	-	-	sim	não	não codificante
TVAD007001F09	proteína transportadora de zinco	25919	2.0e-06	45	-	-	-	-	sim	não	não codificante
TVAD007001F10	citocromo C				IPR000345		-	-	sim	não	não codificante
TVAD007100E01	proteína canal de potássio	25886	3.0e-10	59	IPR003131	5.4e-08	-	-	sim	não	não codificante
TVAD007200F07	citocromo C				IPR000345		-	-	sim	não	não codificante



Legenda:

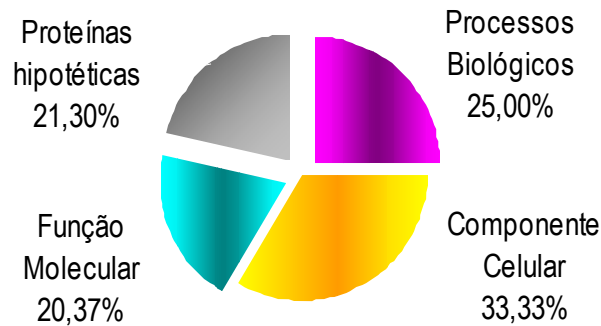


Figura 2: Classificação das seqüências de acordo com as três categorias funcionais do Gene Ontology (ASHBURNER *et al.*, 2000) e respectivo percentual da quantidade encontrada usando o *e-value* $<10^{-5}$. Das 331 seqüências não-redundantes obtidas, desconsiderando-se o *e-value* estipulado, 56,5% obtiveram similaridade em bancos de dados com proteínas descritas, 43,5% delas foram similares a proteínas hipotéticas.

5. DISCUSSÃO

Das 1.086 GSS geradas foram obtidas 331 seqüências não-redundantes perfazendo um total de 135.250 bases do genoma de *T. vivax*, que representam uma amostra de 0,54% do genoma do organismo cujo tamanho está estimado em torno de 25 Mb (<http://www.sanger.ac.uk/Projects/Protozoa>). Considerando-se o desconhecimento do genoma do *T. vivax* e que a maior parcela das seqüências geradas foram identificadas como proteínas similares ao cinetoplasto e flagelo, estruturas características de organismos da Ordem Kinetoplastida, o número de seqüências geradas corresponde a uma amostra representativa do genoma do organismo.

Das 108 GSS geradas e que apresentaram o *e-value* $<10^{-5}$, um total de 71 (68,27%) apresentam similaridade com seqüências de proteínas descritas em kinetoplastídeos. As identidades das demais seqüências apresentam similaridade com proteínas descritas em insetos (9,62%), fungos (5,77%), mamíferos (4,81%), plantas (3,85%), bactérias (2,88%), peixes (1,92%), algas (1,92%) e vírus (0,96%).

Das GSS similares à seqüências de kinetoplastídeos, 29 referem-se ao minicírculo de kDNA ou à proteínas do flagelo, cujas funções estão relacionadas, respectivamente, ao envolvimento na edição de RNA e composição do axonema flagelar. Proteínas do flagelo são restritas à estruturas de organismos flagelados como os grupos dos kinetoplastídeos, euglenóides e dinoflagelados; e proteínas do minicírculo de kDNA são exclusivas do grupo Kinetoplastida (THERTULEIN, SIMPSON-HAIDARIS, HAIDARIS, 1994; MAGA, LeBOWITZ, 1999).

A similaridade com a actina e a dineína de fungo também revela proteínas comumente encontradas em tripanossomatídeos, cuja função é indispensável para a montagem do citoesqueleto e funcionamento do flagelo. Em organismos mutantes

de *Leishmania mexicana*, isentos de proteínas flagelares nativas, foi verificado que a mobilidade celular teve redução em até cinco vezes. Além da mobilidade, estas proteínas desempenham papel na aderência flagelar em tecidos hospedeiros e do vetor (MAGA, LeBOWITZ, 1999).

Algumas seqüências são similares a proteínas não descritas no gênero *Trypanosoma* como a alatotropina, presente em insetos, a atrofina em neurônio humano, a extensina e fosfopantotenoilcisteína sintetase em plantas.

A alatotropina é um peptídeo atuante no controle hormonal e no desenvolvimento de insetos, cuja descrição e função são desconhecidas em protozoários. A proteína extensivamente estudada em insetos foi descrita recentemente como semelhante à estrutura de proteína encontrada em diversos filos de invertebrados (ELEKONICH, HORODYSKI; 2003).

A atrofina-1 causa distúrbios neuronais em humanos e não apresenta função normal conhecida nestes; diferente da função das proteínas tipo atrofina que atuam na embriogênese de *Drosophila spp.* Também, não há relatos deste gene em tripanossomatídeos (WOOD *et al.*, 2000; ERKNER *et al.*, 2002).

A proteína extensina, comumente encontrada em plantas, regula o desenvolvimento do pistilo em angiospermas e compõe a estrutura das paredes celulares de algas. A proteína é altamente conservada em espécies distantes de algas, mas não tem ocorrência e função em tripanossomatídeos (GOLDMAN *et al.*, 1992; WOESSNER *et al.*, 1994).

Outra proteína não descrita em *Trypanosoma spp.* é a fosfopantotenoilcisteína sintetase. Em *A. thaliana*, a proteína está envolvida no crescimento e tolerância salina da planta e em bactérias e leveduras atuam na

biossíntese de pantotenato (ESPINOSA-RUIZ *et al.*, 1999; O'TOOLE, CYGLER, 2003).

As 108 seqüências obtidas sob o *e-value* $<10^{-5}$ quando classificadas por categorias funcionais mostram números percentuais não discrepantes em relação à composição total da amostra; 25% estão agrupadas em “processos biológicos”, 33,33% em “componente celular”, 20,37% em “função molecular” e 21,30% restante em proteínas hipotéticas. Este resultado mostra a eficiência da técnica de seqüenciamento através de GSS em cobrir o genoma revelando um panorama do organismo em linhas gerais.

Com relação às seqüências classificadas como proteínas hipotéticas, excetuando a seqüência similar à proteína hipotética de *Gloeobacter violaceus*, as demais são similares a proteínas das espécies de kinetoplastídeos como *T. brucei*, *T. cruzi*, *T. brucei brucei*, *L. brasiliensis* e *L. donovani*.

Por fim, 21 seqüências (6,34%) do total de GSS não-redundantes, não apresentaram similaridade com seqüências depositadas em bancos de dados públicos. Estas seqüências representam até o momento possíveis ‘genes órfãos’, embora sejam passíveis de processo de anotação e análise mais detalhadas. Este conjunto de seqüências também é de extrema importância, pois, pode representar genes exclusivos da espécie úteis como marcadores moleculares para o desenvolvimento de novos testes diagnósticos.

Todas as GSS geradas no presente estudo serão depositadas no GenBank após a publicação dos resultados. As seqüências anotadas estarão disponíveis através da plataforma Biowebdb (<http://www.biowebdb.org>).

6. CONCLUSÕES

- 1 A biblioteca de DNA genômico da cepa ILDat2160 de *T. vivax* foi validada por apresentar GSS correspondentes à seqüências de espécies relacionadas.
- 2 Foram seqüenciados 1.086 clones da biblioteca de DNA genômico da cepa ILDat2160 de *T. vivax* obtendo-se um total de 455 GSS de alta qualidade e 331 seqüências não redundantes.
- 3 A identidade das GSS geradas apresentaram-se distribuídas homoganeamente dentre as categorias funcionais da classificação do Gene Ontology (GO), representando uma varredura inédita do genoma do organismo assim como a primeira descrição de alguns genes para o *T. vivax*.
- 4 Das 331 seqüências não redundantes, 108 apresentaram similaridade significativa com proteínas depositadas em bancos de dados públicos usando-se o *e-value* $<10^{-5}$ como limite e 21 seqüências não apresentaram similaridade com seqüências depositadas nos 12 bancos de dados consultados, representando possíveis genes órfãos.
- 5 As GSS de alta qualidade geradas corresponderam a 135.250 bases representando uma amostra de aproximadamente 0,54% do genoma de *T. vivax*.

- 6 Das 108 seqüências sob o *e-value* $<10^{-5}$, 71 (68,27%) apresentaram similaridade com seqüências descritas em kinetoplastídeos.
- 7 Das seqüências similares a kinetoplastídeos, 29 foram identificadas como kDNA minicírculo e proteínas flagelares PFR-A e PFR-C.
- 8 Das seqüências similares a outros organismos, 14 identificadas em alatotropina, atrofina-1, extensina e fosfopantotenoilcisteína sintetase não são descritas no gênero *Trypanosoma*.
- 9 As GSS geradas que não apresentaram similaridade com os bancos de dados através da rotina de alinhamento local BLAST estão sendo submetidas à análise através do método probabilístico construído com base no modelo de Markov.
- 10 A plataforma interativa de anotação genômica GARSA (*Genomic Analysis Resources for Sequence Annotation*) em desenvolvimento apresenta-se como uma alternativa viável, fácil e de tempo reduzido para a análise de seqüências, tendo facilitado e permitido o êxito do presente estudo.

7. REFERÊNCIAS BIBLIOGRÁFICAS

ABDEL-LATIEF, M.; MEYERING-VOS, M.; HOFFMANN, K. H. Molecular characterization of cDNAs from the fall armyworm *Spodoptera frugiperda* encoding *Manduca sexta* allatotropin and allatostatin preprohormone peptides. **Insect Biochemistry and Molecular Biology**, v. 33, p. 467-476, 2003.

ABDEL-LATIEF, M.; MEYERING-VOS, M.; HOFFMANN, K. H. Molecular characterization of cDNAs from the fall armyworm *Spodoptera frugiperda* encoding *Manduca sexta* allatotropin and allatostatin preprohormone peptides. **Insect Biochemistry and Molecular Biology**, v. 33, p. 467-476, may-2003.

AJIOKA, J. W. *et al.* Gene Discovery by EST Sequencing in *Toxoplasma gondii* Reveals Sequences Restricted to the Apicomplexa. **Genome Research**, v. 8, p. 18-28, 1998.

ALEXANDRE, S. *et al.* Differential of a family of putative adenylate/ guanylate cyclase genes in *Trypanosome brucei*. **Molecular Biochemical Parasitology**, v. 43, n. 02, p. 279-288, 1990.

ALMEIDA, P. J.L.P. *et al.* Diagnostic evaluation of PCR in goats experimentally infected with *Trypanosoma vivax*. **Acta Tropica**, v. 66, p. 45-50, fev-1997.

ARCHIBALD, J. M.; THE, E. M.; KEELING, P. J. Novel ubiquitin fusion proteins: ribosomal protein P1 and Actin. **Journal of Molecular Biology**, v. 328, p. 771-778, may-2003.

ASAY, D. J.; WILKES, D. E. The dynein heavy chain family. **Journal Eukaryotic Microbiology**, v. 51, n. 01, p. 23-29, jan-2004.

ASHBURNER *et al.* Gene Ontology: tool for the unification of biology. **Nature Genetics**, v. 25, p. 25-29, 2000.

BAI, X.; SASKIA, A.; HOGENHOUT, A. A genome sequence survey of the mollicute corn stunt spiroplasma *Spiroplasma kunkelii*. **FEMS Microbiology Letters**, v. 210, p. 7-17, marc-2002.

BARRETT, M. P.; FAIRLAMB, A. H. The biochemical basis of arsenical-diamidine cross-resistance in African trypanosomes. **Parasitology Today**, v. 15, n. 4, p. 136-140, apr-1999.

BARRY, J. D. *et al.* Why are parasite contingency genes often associated with telomeres? **International Journal for Parasitology**, v. 33, p. 29-45, 2003.

BARRY, J. D.; Mc CULLOCH, R. Antigenic variation in trypanosomes: enhanced phenotypic variation in a eukaryotic parasite. **Advanced Parasitology**, Glasgow, v. 49, p. 1-70, 2001.

BENDER, K.; BETSCHART, B.; HECKER, H. Histone-DNA interactions in the chromatin of procyclic *Trypanosoma brucei brucei*. **Parasitology Research**, v. 78, n. 06, p. 495-500, 1992.

BORST, P.; OUELLETTE, M. New mechanisms of drug resistance in parasitic protozoa. **Annual Review Microbiology**, v. 49, p. 427-460, 1995.

BORST, P.; ULBERT, S. Control of VSG gene expression sites. **Molecular Biochemistry Parasitology**, v. 114, n. 1, p. 17-27, apr-2001.

BOSSCHE, P. V. Some general aspects of the distribution and epidemiology of bovine trypanosomosis in southern Africa. **International Journal for Parasitology**, v. 1, n. 31, p. 592-598, 2001.

BOWES, A. E. *et al.* The acquisition of lyso-phosphatidylcholine by African trypanosomes. **Journal Biological Chemical**, v. 268, p. 13885-13992, 1993.

BRINGAUD, F. *et al.* A new expressed multigene family containing a hot spot of insertion for retroelements is associated with subtelomeric regions in *Trypanosoma brucei*. **Eukaryotic Cell**, v. 01, p. 137-151, feb-2002.

CAMARGO, A. A. *et al.* The contribution of 700,000 ORF sequence tags to the definition of the human transcriptome. **Proceedings of the National Academy of Sciences**, v. 98, n. 21, p. 12103-12108, oct-2001.

CARLTON, J. M. *et al.* Profiling the malaria genome: a gene survey of three species of malaria parasite with comparison to other apicomplexan species. **Molecular Biochemistry Parasitology**, v. 118, n. 02, p. 201-210, dec-2001.

CHEVALIER, N. *et al.* *Trypanosoma brucei* contains a 2,3-biphosphoglycerate independent phosphoglycerate mutase. **European Journal of Biochemical**, v. 267, n.05, p. 1464-1472, 2000.

CHI, T. B.; CHOI, S. Y.; WILLIAMS, N. The ATP synthase of *Trypanosoma brucei* is developmentally regulated by an inhibitor peptide. **Arch Biochem Biophys**. V. 333, n. 01, p. 291-297, sep-1996.

CONIBEAR, E.; CLECK, J. N.; STEVENS, T. H. Vps51p mediates the association of the GARP (Vps52/53/54) complex with the late Golgi t-SNARE T1g1p. **Molecular Biology of the Cell**, v. 14, p. 1610-1623, 2003.

CONIBEAR, E.; STEVENS, T. H. Vps52p, Vps53p, and Vps54p form a novel multisubunit complex required for protein sorting at the yeast late Golgi. **Molecular Biology of the Cell**, v. 11, p. 305-323, 2000.

DA SILVA GIOTTO, M. T. *et al.* The crystal structure of *Trypanosoma brucei* enolase: visualisation of the inhibitory metal binding site III and potential as target for selective, irreversible inhibition. **Journal Molecular Biological**, v. 331, n.03, p. 653-665.

DAS, A. *et al.* The protein phosphatase inhibitor okadaic acid induces defects in cytokinesis and organellar genome segregation in *Trypanosoma brucei*. **Journal cell Science**, v. 107, p. 3477-3483, 1994.

DAVILA, A. M. R. **Tripanosomose animal na América do Sul: Epizootiologia, evolução e tecnologias da informação**. Rio de Janeiro, 2002. 133. Tese (Doutorado em Biologia Celular e Molecular) – Instituto Oswaldo Cruz, Fundação Oswaldo Cruz.

DAVILA, A. M. R.; RAMIREZ, L.; SILVA, R. A. M. S. Morphological and Biometrical Differences among *Trypanosoma vivax* Isolates from Brazil and Bolivia. **Memórias do Instituto Oswaldo Cruz**, v.92, n. 3, p. 357-358, Maio/Jun-1997.

DEGRAVE, W. M. *et al.* Parasite genome initiatives. **International Journal for Parasitology**, v. 31, p. 532-536, may-2001.

DELCHER, A. L. *et al.* Improved microbial gene identification with Glimmer. **Nucleic Acids Research**, v. 27, n. 23, p. 4636-4641, dec-1999.

DESQUESNES M. *et al.* Detection and identification of *Trypanosoma* of African livestock through a single PCR based on internal transcribed spacer 1 of rDNA. **International Journal for Parasitology**, v. 31, p. 610-614, mai-2001.

DESQUESNES M. Evaluation of a simple PCR technique for the diagnosis of *Trypanosoma vivax* infection in the serum of cattle in comparison to parasitological techniques and antigen-enzyme-linked immune sorbent assay. **Acta Tropica**, v. 65, p. 139-148, mai-1997.

DESQUESNES, M.; ROCQUE, S. La.; PEREGRINE, A. S. French Guyana stocks of *Trypanosoma vivax* resistant to diminazene aceturate but sensitive to isometamidium chloride. **Acta Tropica**, v. 60, p. 133-136, jul-1995.

DIAS-NETO, E. *et al.* Shotgun sequencing of the human transcriptome with ORF expresses sequence tags. **Proceedings of the National Academy of Sciences**, v. 97, n. 7, p. 3491-3496, mar-2000.

DIAZ ANEL, A. M. *et al.* MRNA encoding a putative RNA helicase of the DEAD-box gene family is up-regulated in trypomastigotes of *Trypanosoma cruzi*. **Journal of Eukaryotic Microbiology**, v. 47, n. 06, p. 555-560, 2000.

DONELSON, J. E. Antigenic variation and the African trypanosome genome. **Acta Tropica**, v. 85, p. 391-404, 2003.

DONELSON, J. E.; HILL, K. L.; EL-SAYED, N. M. A. Multiple mechanisms of immune evasion by African Trypanosomes. **Molecular and Biochemical Parasitology**, v.91, p. 51-66, 1998.

DWINGER, R. H. *et al.* Susceptibility of buffaloes, cattle and goats to infection with different stocks of *Trypanosoma vivax* transmitted by *Glossina morsitans centralis*. **Research Veterinary Science**, v. 41, n. 3, p. 307-315, nov-1986.

EBRALIDSE, K. K.; GRACHEV, S. A.; MIRZABEKOV, A. D. A highly basic histone H4 domain bound to the sharply bent region of nucleosomal DNA. **Nature**, v. 331, n. 6154, p. 365-367, 1988.

ELEKONICH, M. M.; HORODYSKI, F. M. Insect allatotropins belong to a family of structurally-related myoactive peptides present in several invertebrate phyla. **Peptides**, v. 24, n. 10, p. 1623-1632, 2003.

ELEKONICH, M. M.; HORODYSKI, F. M. Insect allatotropins belong to a family of structurally-related myoactive peptides present in several invertebrate phyla. **Peptides**, v. 24, p. 1623-1632, 2003.

EL-SAYED, N. M. A.; DONELSON, J. E. African Trypanosomes Have Differentially Expressed Genes Encoding Homologues of the *Leishmania* GP63 Surface Protease. **The Journal of Biological Chemistry**, v.272, n.42, p. 26742-26748, out-1997.

EL-SAYED, N. M. *et al.* The African trypanosome genome. **International Journal for Parasitology**, v. 30, n. 4, p. 329-345, apr-2000.

ERKNER, A. *et al.* Grunge, related to human Atrophin-like proteins, has multiple functions in *Drosophila* development. **Development**, v. 129, n. 05, p. 1119-1129, 2002.

ERTL, H. *et al.* A novel extensin that may organize extracellular matrix biogenesis in *Volvox carteri*. **The EMBO Journal**, v. 11, n. 06, p. 2055-2062, jun-1992.

ESPINOSA-RUIZ, A. *et al.* *Arabidopsis thaliana* AtHAL3: a flavoprotein related to salt and osmotic tolerance and plant growth. **The Plant Journal**, v. 20, n. 05, p. 529-539, 1999.

FAN, J. *et al.* Generation, Identification and Evaluation of Expressed Sequence Tags from Different Developmental Stages of the Asian Blood Fluke *Schistosoma japonicum*. **Biochemical and Biophysical Research Communications**, v. 252, p. 348-356, 1998.

FLORIN, J.; RUDOLF, M.; SEEBECK, T. The major components of the paraflagellar rod of *Trypanosoma brucei* are two similar, but distinct proteins which are encoded

by two different gene loci. **Journal Biological Chemistry**, v. 269, n. 46, p. 28745-28751, 1994.

FORMOSA, T.; NITTIS, T. Suppressors of the temperature sensitivity of DNA polymerase α mutations in *Saccharomyces cerevisiae*. **Molecular General Genetics**, v. 257, p. 461-468, 1998.

FOX, A. H *et al.* Paraspeckles: a novel nuclear domain. **Current Biology**, v. 12, p. 13-25, 2002.

FRANCO, G. R. *et al.* Identification of new *Schistosoma mansoni* genes by the EST strategy using a directional cDNA library. **Gene**, v. 152, p. 141-147, 1995.

FREEMONT, P. S. Ring for destruction? **Current Biology**, v. 10, n. 02, p. 84-87, 2000.

FRIEMANN, R. *et al.* Structure of thioredoxin from *Trypanosoma brucei brucei*. **FEBS Letters**, v. 554, n. 03, p. 301-305, 2003.

GALINDO, M. *et al.* Chromatin from two classes of platyhelminthes display both protist H1 and higher eukaryote core histones. **FEBS Letters**, v. 567, p. 225-229, 2004.

GARBARINO, J. E.; GIBBONS, I. R. Expression and genomic analysis of midasin, a novel and highly conserved AAA protein distantly related to dynein. **BMC Genomics**, v. 03, n. 18, p. 1-16, 2002.

GARCIA-SALCEDO, J. A. *et al.* A differential role for actin during the life cycle of *Trypanosoma brucei*. **The EMBO Journal**, v. 23, n. 04, p. 780-789, 2004.

GARCIA-SALCEDO, J. A. *et al.* A protein kinase specifically associated with proliferative forms of *Trypanosoma brucei* is functionally related to a yeast kinase involved in the co-ordination of cell shape and division. **Molecular Microbiology**, v. 45, n. 02, p. 307-319, 2002.

GARDINER, P. R. Recent studies of the biology of *Trypanosoma vivax*. **Advanced Parasitology**, v. 28, p. 229-317, 1989.

GEERTS, S. *et al.* African bovine trypanosomiasis: the problem of drug resistance. **Trends in Parasitology**, v. 17, n.1, p. 25-28, jan-2001.

GEYSEN, D.; DELESPAUX, V.; GEERTS, S. PCR-RFLP using Ssu-rDNA amplification as an easy method for species-specific diagnosis of *Trypanosoma* species in cattle. **Veterinary Parasitology**, v. 110, p. 171-80, jan-2003.

GOLDMAN, M. H. M. *et al.* Developmental expression of tobacco pistil-specific genes encoding novel extensin-like proteins. **The Plant Cell**, v. 04, p. 1041-1051, 1992.

GONZALEZ, J. *et al.* A novel protein phosphatase 2A (PP2A) is involved in the transformation of human protozoan parasite *Trypanosoma cruzi*. **Biochemical Journal**, v. 374, p. 647-656, 2003.

GRAHAM, S. V.; WYMER, B.; BARRY, J. D. Activity of a trypanosome metacyclic variant surface glycoprotein gene promoter is dependent upon life cycle stage and chromosomal context. **Molecular Cell Biology**, v.18, p. 1137-1146, mar-1998.

GRUTER, E.; BETSCHAT, B. Isolation, characterization and organization of histone H1 genes in African trypanosomes. **Parasitology Research**, v. 87, n. 12, p. 977-984. dec-2001.

HANKS, S. K.; HUNTER, T. The eukaryotic protein kinase superfamily: kinase (catalytic) domain structure and classification. **The FASEB Journal**, v. 09, n. 08, p. 576-596, 1995.

HASSAN, H. F.; COOMBS, G. H. A comparative study of the purine and pyrimidine metabolizing enzymes of a range of trypanosomatids. **Comparative Biochemical Physiological Biological**, v. 84, .02, p. 219-223, 1986.

HECKER, H. *et al.* The chromatin of trypanosomes. **International Journal of Parasitology**, v. 24, n. 06, p. 809-811, sep-1994.

HERRON, L. L. *et al.* Genome sequence survey identifies unique sequences and key virulence genes with unusual rates of amino acid substitution in bovine *Staphylococcus aureus*. **Infection and Immunity**, v. 70, n. 07, p. 3978-3981, aug-2002.

HOARE, C. A. The Salivaria: subgenus *Duttonella*. In: **The trypanosomes of mammals: A zoological monograph**. Oxford and Edinburgh: Blackwell Scientific Publications, 1972. p. 401-424.

HUANG, M. G. et al. Prostate Cancer Expression Profiling by cDNA Sequencing Analysis. **Genomics**, v. 59, p. 178-186, 1999.

HUANG, X., MADAN, A. CAP3: A DNA Sequence Assembly Program. **Genome Research**, vol. 9, p. 868-877, 1999.

IYER, L. M. et al. Evolutionary history and higher order classification of AAA+ ATPases. **Journal Structure Biology**, v. 146, p. 11-31, 2004.

JOAZEIRO et al. The tyrosine kinase negative regulator c-Cbl as ring-type, E2-dependent ubiquitin-protein ligase. **Science**, v. 826, n. 5438, p. 309-312, 1999.

JONES, T. W.; DAVILA, A. M. *Trypanosoma vivax* - out of Africa. **Trends in Parasitology**, v. 17, n. 2, p. 99-101, feb-2001.

KAN, Z. et al. Gene structure prediction and alternative splicing analysis using genomically aligned EST's. **Genome Research**, v. 11, p. 889-900, 2001.

KIRCHHOFF, L. V. et al. Ubiquitin genes in Trypanosomatidae. **The Journal of Biological Chemistry**, v. 253, n. 25, p. 12698-12704, sept-1988.

KOONCE, M. P.; GRISSOM, P. M.; McINTOSH, J. R. Dynein from *Dictyostelium*: primary structure comparisons between a cytoplasmic motor enzyme and flagellar dynein. **The Journal of Cell Biology**, v. 119, p. 1597-1604, dec-1992.

KRISTJANSON, P. M. et al. Measuring the cost of African animal trypanosomosis, the potential of control and results to research. **Agricultural Systems**, v. 59, p. 79-98, 1999.

LANDFEAR, S. M. Molecular genetics of nucleoside transporters in *Leishmania* and African trypanosomes. **Biochemical Pharmacology**, v. 62, p. 149-155, 2001.

LAURENTINO, E. C. *et al.* A survey of *Leishmania braziliensis* genome by shotgun sequencing. **Molecular and Biochemical Parasitology**, v. 137, n. 01, p. 81-86, sep-2004.

LEMMON, S. K.; TRAUB, L. M. Sorting in the endosomal system in yeast and animal cells. **Current Opinion in Cell Biology**, v. 12, p. 457-466, aug-2000.

LIU, C. *et al.* A random survey of the *Cryptosporidium parvum* genome. **Infection and Immunity**, v. 67, n. 8, p. 3660-3669, ago-1999.

LORENZI, H. A.; VÁZQUEZ, M. P.; LEVIN, M. J. The genes for a DEAH RNA helicase, a NifU like protein and the translation factor eIF6 constitute the SZ5 locus of *Trypanosoma cruzi*. **Molecular and Biochemical Parasitology**, v. 111, p. 207-211, 2000.

LÜDEMANN, H. *et al.* *Trypanosoma brucei* tryparedosin, a thioredoxin-like protein in African trypanosomes. **FEBS Letters**, v. 431, p. 381-385, 1998.

MADRUGA, C. R.; MORZARIA, S.; MAJIWA, P. O. Caracterização genética do *Trypanosoma vivax* isolado no Pantanal do Mato Grosso e o diagnóstico diferencial da infecção por *T. evansi* pela reação em cadeia da polimerase (PCR). Disponível em <<http://www.cnpqc.embrapa.br/publicacoes/pa/pa49.html>> Acesso em: 25 ago 2003.

MAGA, J. A.; LeBOWITZ, J. H. Unravelling the kinetoplastid paraflagellar rod. **Trends in Cell Biology**, v. 09, p. 409-413, out-1999.

MARGOLIS, R. L. *et al.* DRPLA gene (atrophin-1) sequence and mRNA expression in human brain. **Brain Research Molecular**, v. 36, n. 02, p. 219-226, 1996.

MARQUES PORTO, R. *et al.* Histone H1 is phosphorylated in non-replicating and infective forms of *Trypanosoma cruzi*. **Molecular Biochemistry Parasitology**, v. 119, n. 02, p. 265-271, feb-2002.

MASAKE, R. A. *et al.* Sensitive and Specific Detection of *Trypanosoma vivax* Using the Polymerase Chain Reaction. **Experimental Parasitology**, v. 85, p. 193-205, 1997.

MASIGA, D. K. *et al.* Sensitive detection of trypanosomes in tsetse flies by DNA amplification. **International Journal for Parasitology**, v. 22, p. 909-918, 1992.

MASLOV, D. A. *et al.* A putative novel nuclear-encoded subunit of the cytochrome c oxidase complex in trypanosomatids. **Molecular Biochemical Parasitology**, v. 125, n. 1-2, p. 113-125, 2002.

MELLORS, A.; SAMAD, A. the acquisition of lipids African trypanosomes. **Parasitology Today**, v. 05, n. 08, p. 239-244, 1989.

MISSEL, A. *et al.* DEAD-box protein HEL64 from *Trypanosoma brucei*: subcellular localization and gene knockout analysis. **Parasitology Research**, v. 85, n. 04, p. 324-330, 1999.

MISSEL, A.; GORINGER, H. U. *Trypanosoma brucei* mitochondria contain RNA helicase activity. **Nucleic Acids Research**, v. 22, n. 20, p. 4050-4056, 1994.

MORGAN, G. W; GOULDING. D; FIELD, M. C. The single dynamin-like protein of *Trypanosoma brucei* regulates mitochondrial division and is not required for endocytosis. **Journal of Biological Chemical**, v. 279, n. 11, p. 10692-10672, dec-2004.

MORLAIS, I. *et al.* New molecular marker for *Trypanosoma (Duttonella) vivax* identification. **Acta Tropica**, v. 80, p. 207-13, dez-2001.

MORTARA, R. A. Studies on trypanosomatids actin, I. Immunochemical and biochemical identification. **Journal of Protozoology**, v. 36, n.01, p. 08-13, jan-1989.

NAESSENS, J.; TEALE, A. J.; SILEGHEM, M. Identification of mechanisms of natural resistance to African trypanosomiasis in cattle. **Veterinary Immunology. Immunopathology**, v. 10, n 87, p. 187-194, sep-2002.

NAULA, C.; SEEBECK, T. Cyclic AMP signaling in trypanosomatids. **Parasitology Today**, v. 16, n. 01, p. 35-38, 2000.

O'TOOLE, N.; CYGLER, M. The final player in the coenzyme A biosynthetic pathway. **Structure**, v. 11, p. 899-900, 2003.

OKECH, G. *et al.* The effect of experimental infection of Boran cattle in early and mid-pregnancy with *Trypanosoma vivax*. **Theriogenology**, v. 46, p. 859-869, 1996.

PARSONS, M.; VALENTINE, M.; CARTER, V. protein kinases in divergent eukaryotes: identification of protein kinase activities regulated during trypanosome development. **Proceedings of the National Academy of Sciences**, v. 90, n. 07, p. 2656-2660, 1993.

PORTAL D *et al.* An early ancestor in the evolution of splicing: a *Trypanosoma cruzi* serine-arginine-rich protein (TcSR) is functional in cis-splicing. **Molecular Biochemical Parasitology**, v. 127, p. 37-46, 2003.

PRIEST, J. W.; HAJDUK, S. L. Developmental regulation of *Trypanosoma brucei* cytochrome c reductase during bloodstream to procyclic differentiation. **Molecular Biochemical Parasitology**, v. 65, n. 02, p. 291-304, 1994.

Protozoan Genomes <<http://www.sanger.ac.uk/Projects/Protozo>> 15 jan. 2005.

QUACKENBUSH, J. *et al.* The TIGR Gene Indices: analysis of gene transcript sequences in highly sampled eukaryotic species. **Nucleic Acids Research**, v. 29, n. 1, p. 159-164, 2001.

RECKENFELDERBAUMER, N. *et al.* Identification and functional characterization of thioredoxin from *Trypanosoma brucei brucei*. **Journal of Biological Chemistry**, v. 275, n. 11, p. 7547-7552, 2000.

RUDENKO, G. *et al.* Selection for activation of a new variant surface glycoprotein gene expression site in *Trypanosoma brucei* can result in deletion of the old one. **Molecular Biochemical Parasitology**, v. 95, n. 01, p. 97-109, 1998.

SABINA, R. L.; MAHNKE-ZIZELMAN, D. K. Towards and understanding of the functional significance of N-terminal domain divergence in human AMP deaminase isoforms. **Pharmacology & Therapeutics**, v. 87, p. 279-283, 2000.

SAMBROOK, J., RUSSEL, D. W. **Molecular cloning: a laboratory manual**. Third edition. Cold Spring Harbor Laboratory Press, New York. Vol. 2, 2001.

SAURIN, A. J. *et al.* Does this have a familiar RING? **Trends Biochemical Sciences**, v. 24, n. 5, p. 208-214, jun-1996.

SÁNCHEZ, D. O. *et al.* Gene discovery through genomic sequencing of *Brucella abortus*. **Infection and Immunity**, v. 69, n. 02, p. 865-868, feb-2001.

SEIDL, A.; DAVILA, A. M.; SILVA, R. A. Estimated financial impact of *Trypanosoma vivax* on the Brazilian pantanal and Bolivian lowlands. **Memórias do Instituto Oswaldo Cruz**, v. 94, p. 269-272, 1999.

SILVA, R. A. *et al.* Hematology of natural bovine trypanosomosis in the Brazilian Pantanal and Bolivian wetlands. **Veterinary Parasitology**, v. 85, n. 1, p. 87-93, aug-1999.

SMITH, T. F. *et al.* The WD repeat: a common architecture for diverse functions. **Trends Biochemical Sciences**, v. 24, p. 181-185, may-1999.

SOUZA, S. J. *et al.* Identification of human chromosome 22 transcribed sequences with ORF expresses sequence tags. **Proceedings of the National Academy of Sciences**, v. 97, p. 12690-12693, nov-2000.

STAY, B. A review of the role of neurosecretion in the control of juvenile hormone synthesis: a tribute to Berta Scharrer. **Insect Biochemistry and Molecular Biology**, v. 30, p. 653-662, 2000.

STOLER *et al.* A mutation in CSE4, an essential gene coding a novel chromatin-associated protein in yeast, causes chromosome nondisjunction and cell cycle arrest at mitosis. **Genes & Development**, v. 09, n. 05, p. 573-586, 1995.

SULLIVAN, K. F.; HECHENBERGER, M.; MASRI, K. Human CENP-A contains a histone H3 related histone fold domain that is required for targeting to the centromere. **The Journal of Cell Biology**, v. 127, n. 03, p. 581-592, 1994.

SWINDLE, J. *et al.* The genomic organization and transcription of the ubiquitin genes of *Trypanosoma cruzi*. **The EMBO Journal**, v. 07, n. 04, p. 1211-1217, apr-1988.

TARLETON. R. L.; KISSINGER, J. Parasite genomics: current status and future prospects. **Current Opinion in Immunology**, v.13, p.395-402, 2001.

THERTULIEN, R.; SIMPSON-HAIDARIS, P. J; HAIDARIS, C. G. Intracellular localization of a *Trypanosoma cruzi* kDNA minicircle transcript using RNA: RNA in situ hybridization. **Journal Eukaryotic Microbiology**, v. 41, n. 04, p. 402-407, jul-aug-1994.

ULLMAN, B. Multidrug resistance and P-glycoproteins in parasitic protozoa. **Jounal Bioenergetic Biomembranes**, v. 27, n. 01, p. 77-84, 1995.

VALE, R. D. AAA proteins: Lords of the ring. **The Journal of Cell Biology**, v. 150, n. 01, p. 13-19, 2000.

VAN DER HEYDEN, N.; DOCAMPO, R. Intracellular pH in mammalian stages of *Trypanosoma cruzi* is K⁺- dependent and regulated by H⁺- ATPases. **Molecular Biochemical Parasitology**, v. 105, n. 02, p. 237-251, 2000.

VANHAMME, L. *et al.* An update on antigenic variation in African trypanosomes. **Trends in Parasitology**, v. 17, n. 7, p. 338-343, jul-2001.

VENTURA, R. M. *et al.* *Trypanosoma vivax*: Characterization of the Spliced-Leader Gene of a Brazilian Cepa and Species-Specific Detection by PCR Amplification of an Intergenic Spacer Sequence. **Experimental Parasitology**, v. 99, p. 37-48, 2001.

VERDE, F.; WILEY, D. J.; NURSE, P. Fission yeast orb6, a ser/thr protein kinase related to mammalian rho kinase and myotonic dystrophy kinase, is required for maintenance of cell polarity and coordinates cell morphogenesis with the cell cycle. **Proceedings of the National Academy of Sciences**, v. 95, n. 13, p. 7526-7531, 1998.

WAINSZELBAUM, M. *et al.* Lysosomal phospholipase A1 in *Trypanosoma cruzi*: an enzyme with a possible role in the pathogenesis of Chagas' disease. **Biochemical Journal**, v. 355, p. 765-770, 2001.

WILKOWSKY, S. E. *et al.* *Trypanosoma cruzi*: phosphatidylinositol 3-kinase and protein kinase B activation is associated with parasite invasion. **Experimental Cell Research**, v. 264, n. 02, p. 211-218, 2001.

WOESSNER, J. P. *et al.* Domain conservation in several volvocalean cell wall proteins. **Plant Molecular Biology**, v. 26, n. 03, p. 947-960, nov-1994.

WOESSNER, J. P.; GOODENOUGH, U. W. Molecular characterization of a zygote wall protein: an extensin-like molecule in *Chlamydomonas reinhardtii*. **Plant Cell**, v. 01, n. 09, p. 901-911, sep-1989.

WOOD, J. D. *et al.* A. Atrophin-1, the dentate-rubral and pallido-luysian atrophy gene product, interacts with ETO/MTG8 in the nuclear matrix and represses transcription. **Journal of Cell Biology**, v. 150, n. 05, p. 939-948, 2000.

WORTHEY, E. A. *et al.* Comparative analysis of editosome proteins in trypanosomatids. **Nucleic Acids Research**, v. 31, n. 22, p. 6392-6408, 2003.

ZAMBRANO-VILLA, S. *et al.* How protozoan parasites evade the immune response. **Trends in Parasitology**, v. 18, n. 6, p. 272-278, jun-2002.