



Universidade do Estado do Rio de Janeiro

Centro Biomédico

Faculdade de Ciências Médicas

Tainah Silva Galdino de Paula

**Análise da região 5' não traduzida (5' utr) em Sequências anotadas
como trans-sialidase no genoma de *Trypanosoma cruzi***

Rio de Janeiro

2009

Tainah Silva Galdino de Paula

Análise da região 5' não traduzida (5' utr) em Sequências anotadas como trans-sialidase no genoma de *Trypanosoma cruzi*

Dissertação apresentada, como requisito parcial para obtenção do título de Mestre ao Programa de Pós-graduação em Ciências Médicas da Universidade do Estado do Rio de Janeiro.

Orientador: Dr. Adelfton Brandão
Coorientadora: Dra. Constança Britto

Rio de Janeiro
2009

Tainah Silva Galdino de Paula

Análise da região 5' não traduzida (5' utr) em Sequências anotadas como trans-sialidase no genoma de *Trypanosoma cruzi*

Dissertação apresentada, como requisito parcial para obtenção do título de Mestre ao Programa de Pós-graduação em Ciências Médicas da Universidade do Estado do Rio de Janeiro.

Aprovada em 04 de maio de 2009.

Orientador: Dr. Adéilton Brandão

Coorientadora: Dra. Constança Britto

Banca Examinadora:

Dr. Alexandre Ribeiro Bello
Instituto de Biologia Roberto Alcântara Gomes – UERJ

Dr. Marcelo Alves Ferreira
Fundação Oswaldo Cruz – Fiocruz

Dra. Cláudia D'ávila Levy
Fundação Oswaldo Cruz – Fiocruz

Rio de Janeiro

2009

DEDICATÓRIA

Por você eu dançaria tango no teto, Eu limparia os trilhos do metrô, Eu iria a pé do Rio a Salvador...

Por Você (Roberto Frejat)

Ao Rodrigo Galdino de Paula, meu Gal, por todo amor, amizade, companheirismo e compreensão.

AGRADECIMENTOS

Nenhuma realização pode ser plena se não tivermos ao nosso lado pessoas queridas nos dando apoio, afeto, conselhos e principalmente, nos oferecendo a mão nos momentos de queda.

Ao Dr. Adeílton Alves Brandão, por toda sua paciência, dedicação, e principalmente pela confiança.

À Dra. Constança Britto, por todo apoio, carinho e amizade.

Ao Dr. Cícero Brasileiro de Mello Netto, por ter me incentivado e ajudado no início dessa longa caminhada.

Ao Dr. Octávio Fernandes pela oportunidade de aprendizado.

À minha mãe, Dra. Efigênia, por ter me ensinado o prazer de aprender e, sobretudo, pelo amor, respeito e confiança.

Ao meu pai, João Batista, sempre otimista mediante os meus projetos de vida, além de todo seu carinho e atenção.

Aos meus avós, por todo carinho e dedicação.

Ao meu irmão, João, meu sempre Joãozinho.

Às irmãs que Deus me emprestou, Vanessa Galdino de Paula e Danielle Galdino de Paula.

Aos pais que conquistei, Elizabeth Galdino de Paula e Pedro de Paula Filho, pelo imenso amor que existe entre nós.

À querida tia Nita, por seu grande exemplo de força, fé e amor.

Ao Dr. Dário Kalume, pelas horas de trabalho convertidas em uma grande e divertida amizade, e por toda ajuda profissional, além da revisão dessa Tese!

À super querida Daniela Pita Pereira por sua ajuda durante o *Real Time* e principalmente pela doce amizade.

À Elizabeth Santos (“chatinhaaa”) por estar sempre disposta a ajudar e distribuir carinho.

À doce amiga Estela Maris (Estelinha) e sua linda família por todo carinho e pensamento positivo.

À Dra. Joseli Nogueira pela amizade, pelas horas de descontração e pela ajuda técnica.

À Simone Santos pelas orientações nas horas críticas de bancada e principalmente pelas nossas risadas diárias.

À querida Dra. Maria Inês Dória pelo carinho diário e seu companheirismo.

Ao Dr. Otacílio da Cruz Moreira, uma amizade recente, porém sólida e sincera.

À Dra. Maria Esther Ricci, pelas horas de descontração, pelas loucuras e pela amizade.

Aos amigos Fábio Alves e seu fiel escudeiro Társis Brust (menino prodígio), por todo carinho e ajuda nos momentos difíceis.

À querida Nédia Saad Nheme pela amizade, carinho e apoio.

À Dra. Aline Caseca Volotão por estar sempre disposta a ajudar.

À querida Dra. Martha Suarez Murtis por todo seu carinho e alegria.

Às queridas Simone Kikuchi e Helena Toma por me cederem a maioria das cepas utilizadas nesse trabalho, e pela gostosa amizade.

À Dra. Cátia Sodré pelas boas idéias e pelo carinho.

Aos amigos do Laboratório: Elizabeth Luz, Mariângela Ziccardi, Carla Sodré, Franklyn (“muchacho”), Nathália Motta (“horrorosinha”), Nathália Souza (Natinha), Ingrid, Bill, Grazi, Ju, Vivian, Maria, Thaina (a outra) e Lauren. Vocês são especiais, tenho um carinho grande por cada um.

À D. Maria, Sr. Francisco, Sr. Severino, Geraldo e Sr. Pauzannias por executarem seus trabalhos de maneira sempre carinhosa e alegre.

Aos amigos de VR, pela força, confiança e amizade ao longo do tempo.

Aos amigos do Seara Fraternal, por estarem ao meu lado nessa imensa e linda jornada.

À amiga Paula Fernanda Gonçalves do Santos, pelos desabafos, sorrisos e apoio.

À FAPERJ pelo apoio financeiro durante o desenvolvimento do projeto.

À Plataforma de Seqüenciamento de DNA-RJ / PDTIS-Fiocruz, meu agradecimento especial à Aline Moreira por todo seu bom-humor e empenho.

À Plataforma de PCR em Tempo Real – RJ / PDTIS-Fiocruz.

E, sobretudo, à Deus, por me permitir estar sempre em busca do aprendizado, autoconhecimento, e com muita gente bacana ao meu redor.

OBRIGADA!!!!

Prece de Todo Dia

Senhor, concede-me a consciência dos meus muitos erros.

Não me consintas viver iludido a meu próprio respeito.

Que eu tenha suficiente lucidez para saber quem sou.

Que eu consiga, mais que os outros, detectar as fragilidades que me são comuns.

Dá-me a tua força para que eu possa superar-me, a Tua luz para não caminhar nas trevas,

a Tua paz na luta que me aflinge.

Que eu seja sempre sincero em meus propósitos e humilde em minhas atitudes, verdadeiro

em minhas palavras e fiel aos meus compromissos.

Senhor, não me deixes entregue à invigilância e ao assédio do mal.

Sê meu abrigo e a minha inspiração!....

Assim Seja.

(Vigiai e Orai - Carlos A. Baccelli / Irmão José)

RESUMO

PAULA, Tainah Silva Galdino de. *Análise da região 5' não traduzida (5' utr) em Sequências anotadas como trans-sialidase no genoma de Trypanosoma cruzi*. 2009.69 f. Dissertação (Mestrado em Ciências Médicas) – Faculdade de Ciências Médicas, Universidade do Estado do Rio de Janeiro, rio de Janeiro. 2009.

A trans-sialidase é uma glicoproteína de membrana pertencente a uma família de genes de cópia múltipla, envolvida no processo de invasão celular do *Trypanosoma cruzi* no hospedeiro vertebrado. Esta dissertação foi concebida com um amplo componente analítico que dependia de dados publicamente disponíveis, ou seja, as sequências oriundas do projeto genoma de *T. cruzi* e cDNAs de trans-sialidase depositadas no Genbank-dbEST. Este componente analítico necessitou ser complementado e ampliado com a obtenção experimental de novas sequências, a partir da metodologia baseada na transcrição reversa acoplada a PCR. Os fragmentos obtidos de cepas de *T. cruzi* Dm28c (*T. cruzi* I), Y (*T. cruzi* II), CL-Brener (*T. cruzi* II, cepa híbrida), INPA4167 (zimodema III), 3663 (zimodema III) e Colombiana (zimodema III) foram clonados, sequenciados e analisados composicionalmente. Essas sequências foram editadas e alinhadas usando-se o software CLUSTAL X. Em uma seção específica do Genbank, denominada dbEST, buscamos os cDNAs homólogos a trans-sialidase. Esta busca por similaridade foi realizada individualmente com os números de acesso referentes às sequências supracitadas contra o dbEST utilizando o BLAST a fim de obtermos informações funcionais e evolutivas. Em seguida, desenvolvemos metodologias experimentais que nos permitiu avaliar segmentos da 5' UTR, tais como os sítios de *trans-splicing* adicionais ou múltiplos em TS e seus respectivos sinais (região rica em polipirimidina), variação composicional e tamanho da região das sequências entre diferentes linhagens de *T. cruzi*. O resultado dessa averiguação também nos mostrou a quantidade de cDNAs relacionados com a trans-sialidase no dbEST bem como a relação desses cDNAs com o mini-exon. As cepas do zimodema III apresentaram tamanho médio dos fragmentos de 312 bases, enquanto *T. cruzi* I e *T. cruzi* II apresentaram, respectivamente 209 e 218. *Trans splicing* adicional ou duplicações gênicas com mutações no sítio primário de *trans splicing* não parece ser um fenômeno exclusivo de algum grupo populacional, embora seja mais evidente em *T. cruzi* zimodema III.

Palavras-chave: *Trypanosoma cruzi*. Trans-Sialidase. 5' UTR

ABSTRACT

Trans-sialidase (TS) is a membrane glycoprotein belonging to a family of genes from multiple copies, involved in the process of cell invasion of the vertebrate host by the *Trypanosoma cruzi* (*T. cruzi*) parasite. This dissertation was developed with an ample analytical component that depended on the publicly available genome database or on the sequences from the *T. cruzi* genome project and trans-sialidase cDNAs sequences deposited in the Genbank-dbEST. This analytical component needed to be supplemented and expanded with the acquisition of new test sequences, from the methodology based on reverse transcription coupled to PCR. The fragments obtained from *T. cruzi* strains corresponding to the different groups were cloned, sequenced and compositionally analyzed. These sequences were edited and aligned using the CLUSTAL X software. Homologous sequences to trans-sialidase were searched against the dbEST databank. This search for similarity was performed individually with the access numbers in the sequence mentioned above against the dbEST using the BLAST to have functional and evolutionary information. Experimental methods were developed that allowed us to evaluate segments of 5' UTR, such as the trans splicing additional or multiples sites found in the TS and their respective signals (region rich in pyrimidine), compositionally variation and size of the sequences between different strains of *T. cruzi*. The result of this analysis showed us the amount of cDNAs related to trans-sialidase in the dbEST as well as the relationship of this cDNAs with the mini-exon.

Keywords: *Trypanosoma cruzi*. Trans-Sialidase. 5' UTR

LISTA DE FIGURAS

Figura 01	Biológico do <i>Trypanoma cruzi</i>	13
Figura 02	Formas evolutivas do <i>Trypanosoma cruzi</i>	14
Figura 03	Áreas Endêmicas da Doença de Chagas	15
Figura 04	Habitação característica onde o triatomíneo se instala	16
Figura 05	Esquema de processamento de mRNA	21
Figura 06	Estrutura gênica do mRNA	24
Figura 07	Classes estruturais de trans-sialidase	28
Figura 08	Esquema de invasão celular pelo <i>Trypanosoma cruzi</i>	32
Figura 09	Gel de agarose contendo RNA total	38
Figura 10	Gel de agarose contendo produtos de RT-PCR qualitativo	39
Figura 11	Gel de poliacrilamida de RT-PCT qualitativo com representantes do grupo ZIII	33
Figura 12	Gel de poliacrilamida com produtos de RT-PCR qualitativo de cada cepa contida nesse estudo	33
Figura 13	Gel de agarose com plasmídeos contendo insertos previamente clonados	34
Figura 14	Mapeamento da região 5'UTR	40
Figura 15	Esquema de uma estrutura geral da 5'UTR das sequências genericamente anotadas como trans-sialidase pelo projeto genoma	41
Figura 16	Logo da provável 5'UTR das trans-sialidasas anotadas pelo projeto genoma de <i>T. cruzi</i>	41
Figura 17	Logo representante do primeiro bloco de 53 bases do sítio de trans-splicing	43
Figura 18	Logo representante do segundo bloco logo após o sítio de transsplicing	44
Figura 19	Representação de parte do alinhamento do segundo bloco	45
Figura 20	Representante do terceiro bloco após o sítio de trans-splicing	45

LISTA DE TABELAS

Tabela 01	Tamanhos dos fragmentos (pb) contendo sequências completas de 5' UTR após clonagem e sequenciamento	35
Tabela 02	Média dos valores de Ct e valores de normalização ΔCt das cepas do grupo Zimodema III analisadas por PCR quantitativo	36
Tabela 03	Valores comparativos da expressão gênica relativa da transsialidase de <i>T. cruzi</i> frente ao uso das diferentes cepas do grupo Zimodema III como calibrador para os experimentos de RT-PCR em tempo real	36

LISTA DE ABREVIATURAS

<i>T. cruzi</i> -	<i>Trypanosoma cruzi</i>
<i>E. coli</i> -	<i>Escherichia coli</i>
UTR -	Untranslated Regions
TS -	Trans-Sialidase
dbEST -	Data Base of Expressed Tags
EST -	Expressed Sequence Tag
mRNA -	Ribonucleic Acid Messenger
BLAST -	Basic Local Alignment Search Tool
cDNA -	Complementary Deoxyribonucleic Acid
ME -	Mini-Exon
RT-	PCR - Reverse Transcription – Polymerase Chain Reaction
LB -	Luria-Bertani
CaCl ₂ -	Cloreto de Cálcio
X-Gal -	5-Bromo-4-Cloro-indolyl-β-D-Galactopyranoside
TCNA -	<i>Trypanosoma cruzi</i> Neuraminidase
Ct -	Threshold Cycle
M-MLV -	Moloney Murine Leukemia Vírus Transcriptase
X-Gal -	(5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside)
μg -	Micrograma
μL -	Microlitro
M -	Molar
mM -	Milimolar

SUMÁRIO

	INTRODUÇÃO	13
	O parasito <i>Trypanosoma cruzi</i>	13
	Grupos populacionais de <i>Trypanosoma cruzi</i>	17
	Transcrição gênica	19
	As regiões não traduzidas	22
	Pseudogenes	25
	Trans-sialidase do <i>Trypanosoma cruzi</i>	25
	Mecanismo molecular da invasão celular por <i>T. cruzi</i>	30
1	OBJETIVOS	34
1.1	Objetivo geral	34
1.2	Objetivos específicos	34
2	MATERIAL E MÉTODOS	35
2.1	Análise composicional das sequências de UTR	35
2.2	Cultivo de <i>T. cruzi</i>	36
2.3	Extração de RNA total de <i>T. cruzi</i>	36
2.4	RT-PCR qualitativo	36
2.5	Preparo de células competentes	37
2.6	Clonagem	38
2.7	Sequenciamento	40
2.8	Eletroforese em gel de agarose	40
2.9	Eletroforese em gel de poliacrilamida	41
2.10	RT-PCR quantitativo (RT-PCR em Tempo Real) – Sistema SYBR Green	41
3	RESULTADOS	43
3.1	Extração de RNA total e RT-PCR qualitativo	43
3.2	Clonagem e sequenciamento	45
3.3	RT-PCR quantitativo	47
3.4	Análises das sequências de 5'UTR em trans-sialidase	49
3.5	O tamanho da 5'UTR de trans-sialidase de classe I	57
4	DISCUSSÃO	59
5	CONCLUSÃO	66
6	PERSPECTIVAS	67

REFERÊNCIAS

68

ANEXOS

81

INTRODUÇÃO

O parasito *Trypanosoma cruzi* e a Doença de Chagas

A **Tripanossomíase Americana** ou Doença de Chagas é uma zoonose do continente americano com forte incidência no Brasil e foi descrita por Carlos Chagas em 1909. O agente etiológico desta doença é o protozoário flagelado *Trypanosoma cruzi* da ordem Kinetoplastida, pertencente à família Trypanosomatidae. O parasito passa por três mudanças morfológicas durante seu ciclo de vida oscilando entre os hospedeiros vertebrados, tais como gambá, tatu e o próprio homem, e hospedeiros invertebrados, insetos vetores com numerosas espécies de hemípteros hematófagos da família Reduviidae (denominados triatomíneos e popularmente conhecidos como barbeiros ou chupança) (revisto por De Souza, 2002) (Figura 1).

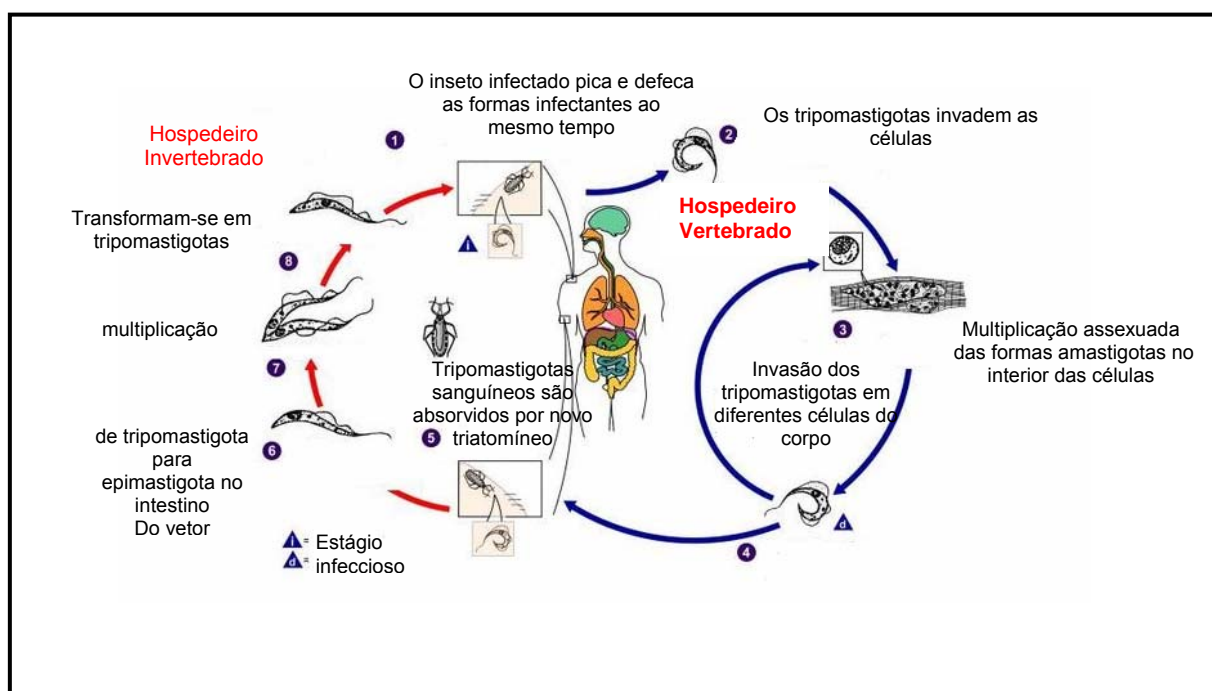


Figura 1 - Ciclo de vida do *Trypanosoma cruzi* (Adaptado de www.dpd.cdc.gov/dpdx)

O *T. cruzi* alterna-se entre estágios que sofrem divisão binária (replicativas) e as formas não replicativas (Figura 2). Como formas replicativas estão incluídas os **epimastigotas**, presentes no tubo digestivo do inseto vetor, e **amastigotas** observados no interior das células de mamíferos (principalmente os macrófagos). As

formas não replicativas infectantes incluem os **tripomastigotas metacíclicos**, encontrados nas fezes e urina do inseto vetor e os **tripomastigotas circulantes**, encontrados no sangue de mamíferos (revisto por De Souza, 2002).

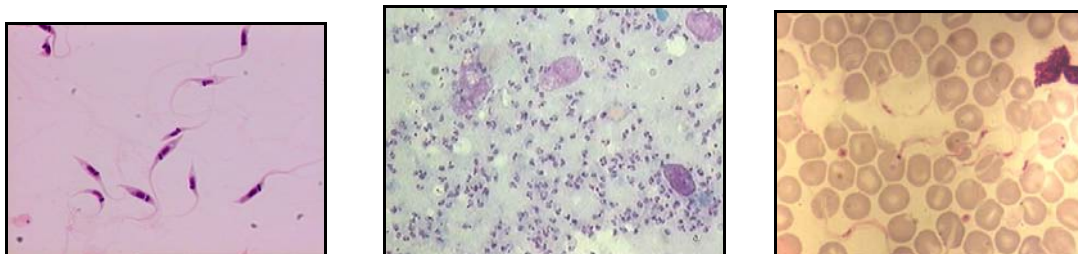


Figura 2 – Formas evolutivas do *Trypanosoma cruzi*. A forma epimastigota é a forma encontrada no tubo digestivo do vetor, é fusiforme e apresenta o cinetoplasto junto ao núcleo; amastigotas, fase intracelular, sem organelas de locomoção aparentes, com pouco citoplasma e núcleo grande. O cinetoplasto fica ao lado do núcleo e é um pouco menor que ele. Está presente na fase crônica da doença; e tripomastigotas fase extracelular, que circula no sangue. Apresenta flagelo e membrana ondulante em toda a extensão lateral do parasito, está presente na fase aguda da doença, constituindo a forma infectante para os vertebrados (<http://www.cienciaviva.org.br>, <http://www.ufrgs.br/para-site/Imagensatlas/Protozoa/Imagens>).

Segundo a Organização Mundial de Saúde (2004), cerca de 18 milhões de pessoas estão acometidas pela moléstia e 100 milhões em situação de risco, principalmente nas regiões endêmicas que compreende a América Latina (Figura 3). Há, portanto, a necessidade de maiores estudos e investimentos acerca da biologia molecular do *Trypanosoma cruzi* e da patogenicidade relacionadas à doença de Chagas.

De acordo com Coura e colaboradores (2002), mais da metade dos casos agudos da doença de Chagas registrados entre 1968 e 2000 na Amazônia brasileira estão relacionadas a transmissão oral através de alimentos contaminados.



Figura 3 - Áreas Endêmicas da Doença de Chagas (WHO, 2002)

O ciclo epidemiológico inicia-se quando o homem invade ecótopos onde eram originalmente ocupados pelo *T. cruzi*, tornando-o suscetível aos hemípteros vetores locais infectados (Dias e Coura, 1997). Além disso, o homem leva para a zona rural habitações precárias isentas de saneamento e com presença de animais domésticos, aumentando a incidência e a prevalência da doença (Figura 4). Portanto, os triatomíneos hematófagos estritos encontram em tais habitações condições ideais de abrigo e oferta alimentar abundante, tornando a transmissão vetorial no mecanismo primário de difusão da doença (Vinhaes e Dias, 2000).



Figura 4 – Exemplo de habitação feita de barro (pau-a-pique) (<http://www.brasiloste.com.br>)

Dessa forma, é notória a observação de que os desequilíbrios econômicos influenciam a distribuição social da parasitose estando diretamente relacionada à qualidade de vida dos habitantes das regiões interioranas. No entanto, com o advento das populações migratórias, houve uma expansão da doença para áreas que antes não eram acometidas pela moléstia principalmente pela transfusão sanguínea, transplante de órgãos, congênita (via transplacentária) e via digestiva, a qual ocorre esporadicamente na espécie humana (Pinto e cols., 2008).

Em humanos, após um curto período de incubação, a infecção apresenta uma fase aguda que pode persistir por cerca de 15 a 30 dias, caracterizada por intensa parasitemia. Nesta fase, morbidade e apresentação clínica estão diretamente associadas com o nível de parasitemia. Alternativamente, a infecção aguda pode também ocorrer como uma forma não aparente, com poucos sintomas. Após a fase aguda da doença, os pacientes entram no estágio crônico assintomático (fase indeterminada), que pode durar por toda a vida na maioria dos indivíduos acometidos pela doença, os quais apresentam altos títulos de anticorpos anti-*T. cruzi*. Os demais indivíduos infectados cronicamente desenvolvem complicações cardíacas ou digestivas 10 a 30 anos após a infecção inicial, em um estágio da doença no qual os parasitos do sangue e tecido são escassos (Tanowitz e cols., 1992; Consenso Brasileiro em Doença de Chagas, 2005).

O diagnóstico da fase aguda de *T. cruzi* geralmente é feito pela detecção de parasitos por métodos parasitológicos convencionais. O exame pode ser feito por análise microscópica de amostra de sangue fresco, não coagulado, ou por marcação dos parasitos com coloração por Giemsa. O diagnóstico da fase crônica da doença de Chagas é geralmente baseado na detecção de anticorpos específicos que se ligam aos antígenos de *T. cruzi*. Os testes sorológicos empregados na fase crônica são: hemaglutinação indireta, imunofluorescência e o ELISA (*enzyme-linked immunosorbent assay*) (Tanowitz e cols., 1992; Consenso Brasileiro em Doença de Chagas, 2005).

Embora nas últimas décadas a articulação de políticas instituídas por países afetados do cone Sul tenha alcançado sucesso quanto ao controle da transmissão da doença pela via vetorial, a elaboração de uma vacina eficaz ainda não é exeqüível, e o tratamento específico apresenta limitações, principalmente nos pacientes portadores da doença crônica tardia (Dias e Schofield, 1999). Portanto, o investimento em estudos mais profundos relacionados a essa patogenia se faz necessário a fim do desenvolvimento de estratégias mais eficientes de controle por quimioterapia ou vacinação. Atualmente a terapêutica utilizada contra o *T. cruzi* ainda são os mesmos de 1968, o Benznidazol (Rochagan[®], Radamil[®], Roche 7-1051; <http://www.roche.com/>) e Nifurtimox (Lampit[®], Bayer 2502; <http://www.bayer.com/>) que apresentam uma série de efeitos colaterais e combate apenas as formas circulantes sanguíneas do parasito (revisto por Coura e De Castro, 2002; Kevin, 2008). As crianças têm uma chance maior de se beneficiarem com o tratamento, pela pouca probabilidade de estarem na fase crônica da moléstia (http://www.dndi.org.br/Portugues/doenca_chagas.aspx).

Grupos populacionais de *T. cruzi*:

Embora o *T. cruzi* seja descrito em um *taxon* único, o parasito apresenta grande heterogeneidade fenotípica e genotípica (Dvorak, 1984; Tibayrenc e cols., 1993). Inicialmente, baseado no método bioquímico por eletroforese de isoenzimas Miles e colaboradores (1977; 1978) observaram perfis eletroforéticos de seis enzimas caracterizando-as, posteriormente, em três padrões de enzimas concernentes a grupos de parasitos (Zimodemas), no qual o zimodema I (ZI) e o zimodema III (ZIII) estão associados ao ciclo silvestre, e o Zimodema II (ZII) ao ciclo doméstico do parasito (Miles e cols., 1978). Aumentando o número de isoenzimas

estudadas e das sub-populações circulantes entre vertebrados e domésticos, silvestres e invertebrados, uma alta heterogeneidade foi evidenciada (Miles e cols., 1980; Bogliolo e cols., 1986; Tibareyenc e cols., 1986, Tibareyenc e Ayala, 1988; Barnabé e cols., 2000).

O progresso tecnológico da biologia molecular culminou no desenvolvimento de novas técnicas possibilitando o estudo da diversidade de *T. cruzi*. Dessa forma, um grau de polimorfismo maior pode ser observado quando comparado com os padrões de isoenzimas.

De acordo com o uso de marcadores moleculares como DNA ribossomal, o gene de mini-exon e análise de microssatélites, foi mostrado um dimorfismo entre os grupos (Souto e Zingales, 1993). O gene de mini-exon está presente no genoma nuclear dos kinetoplastida em torno de 200 cópias organizadas em sequências repetitivas (*tandem*). A amplificação da região intergênica do mini-exon pela Reação em Cadeia da Polimerase (PCR) em diferentes cepas de *T. cruzi* demonstrou um dimorfismo com produtos amplificados em torno de 300pb definindo também dois grupos principais (Murthy e cols., 1992, Souto cols., 1996). Posteriormente, estes dois grupos foram denominados como *T. cruzi* I que se correlaciona ao ZI e *T. cruzi* II que se correlaciona ao ZII (Souto e cols., 1996; Fernandes e cols., 1998a, 1999). Inicialmente, acreditava-se que o grupo *T. cruzi* I era encontrado somente no ciclo silvestre da doença e o *T. cruzi* II no ciclo doméstico (Fernandes e cols., 1998a; Zingales e cols., 1998). Todavia, Fernandes e colaboradores (1999) evidenciaram a presença dos dois grupos no ciclo silvestre da doença e preferência da associação de linhagens filogenéticas principais de *T. cruzi* com diferentes hospedeiros, corroborando a complexidade do ciclo do parasito na natureza (Devera e cols., 2003).

As populações de *T. cruzi* foram agrupadas em duas linhagens filogenéticas principais (*T. cruzi* I e *T. cruzi* II) e um terceiro grupo denominado *T. cruzi* ZIII, cuja posição filogenética não foi definida (Fernandes e cols., 1998; Anônimo, 1999). As subpopulações de *T. cruzi* demonstraram diferenças significantes na composição de sua superfície, tais como a distribuição de carboidratos, lipídeos e proteínas (De Souza, 1995). Essas diferenças podem explicar, pelo menos em parte, a associação entre populações distintas de *T. cruzi* com seus vetores e hospedeiros específicos (Araújo e cols., 2002).

Estudos prévios sobre a complexidade do DNA e da localização genômica de genes de cópia única que codificam enzimas constantemente expressas, sugeriram

que o cromossomo do *T. cruzi* é diplóide (Castro e Castaneda, 1981; Gibson e Miles, 1986; Lanar e cols., 1981). A caracterização citogenética do *T. cruzi* é difícil, pois seus cromossomos não se condensam durante a divisão celular. Outros estudos indicaram que o padrão de migração eletroforética dos cromossomos de *T. cruzi* é único, aparentando estabilidade nas diversas formas evolutivas do parasito (Wagner e So, 1990).

Os microssatélites de *T. cruzi* (minirepetições polimórficas *em tandem*) forneceram uma nova ferramenta para a análise da estrutura do parasito (Oliveira e cols., 1998). Esses marcadores de DNA são extremamente polimórficos e estão dispersos por todo o genoma nuclear do parasito (Macedo e cols., 2004). A análise de microssatélites demonstrou que a percentagem da população multiclonal diminui progressivamente ao comparar isolados de cepas do ciclo silvestre com isolados provenientes do homem, confirmando a idéia prévia de um filtro seletivo humano para a adaptação de determinadas cepas (Macedo e cols., 2001).

Recentemente, Brandão e Fernandes (2006) procurando uma nova evidência da divergência do *taxon T. cruzi* estudaram a região 3' não traduzida (3'UTR) do gene da calmodulina. Esse gene modula o metabolismo de cálcio em várias atividades celulares (Hoefliche e Ikura, 2002). Foram observadas mutações localizadas no segmento de 330pb da 3'UTR com três eventos de inserção /deleção importantes: a presença do dinucleotídeo AG em isolados de ZIII, ausência de 13 bases purinas em ZIII e cinco grupamentos de nucleotídeos GT em *T. cruzi* II. Essas mutações sugerem que *T. cruzi* I e o ZIII compartilham vias funcionais comuns se comparados com os isolados de *T. cruzi* II (Brandão e Fernandes, 2006). Porém a filogenia do ZIII ainda continua controversa.

Transcrição gênica

Usando a hibridização por *Southern*, foram localizados vários genes de cópia múltipla que codificam para rRNA (Farrel e Kirkpatrick, 1987), mini-exon (De Lange e cols., 1984), proteínas específicas de tripomastigotas (Peterson e cols., 1986), entre outros elementos (Wagner e So, 1990). O termo "família multigênica" é amplamente usado para incluir grupos de genes do mesmo organismo que codificam proteínas com sequências similares; e que, portanto podem ter funções relacionadas. A estrutura gênica do *T. cruzi* é bem peculiar; embora seja classificado como um eucarioto, o mesmo possui uma maquinaria transcricional ainda rudimentar quando

comparado com os demais eucariotos superiores; no entanto, quando comparado aos procariotos apresenta um genoma mais complexo. Um exemplo desse paradoxo é o processamento do seu RNA.

O *trans-splicing* é um mecanismo para o processamento do mRNA que requer dois eventos de clivagem que devem ocorrer antes e depois de cada região codificante em um precursor policistrônico para gerar um transcrito maduro, ou seja, mRNAs monocistrônicos (Teixeira e cols., 2008). O *cis-splicing* une exons em uma única molécula de pré-mRNA. O *trans-splicing* geralmente ocorre em organismos unicelulares e nos tripanossomatídeos. Nesses últimos, a maioria dos genes de proteína são agrupados e transcritos policistronicamente. Esses pré-mRNAs são processados a mRNAs monocistrônicos através de um *trans-splicing* (LeBowitz e cols., 1993; Matthews e cols., 1994; Liang e cols., 2003);

O gene de mini-exon está presente no genoma nuclear de kinetoplastida transcritos em cerca de 200 sequências repetidas de um RNA doador. São constituídos de três regiões distintas: exon, intron e região intergênica. O exon é uma sequência com 39 pares de base (pb) altamente conservados e é a região transportada para o RNA (mRNA) (Gonçalves e cols., 2008). O intron, que é a região moderadamente conservada entre espécies do mesmo gênero e a região intergênica ou espaçador não traduzido, pode ser amplificada pela técnica da Reação em cadeia da Polimerase (PCR), que é empregada para classificar taxonomicamente diferentes isolados de *T. cruzi* em dois grupos: *T. cruzi* I e *T. cruzi* II (Fernandes, 1996; Souto e cols., 1996; Fernandes e cols., 1998).

Ocorre uma inserção do mini-exon a cada extremidade codificante do mRNA feita pelo reconhecimento de um sítio acceptor rico em AG (sítio de *trans-splicing*) precedido por uma região rica em pirimidina (sinal de *trans-splicing*) (Jager e cols., 2006) que supre a falta de sítios canônicos para a ocorrência da poliadenilação. Após a inserção da molécula do mini-exon às extremidades do mRNA, ocorre uma clivagem no mRNA policistrônico e uma inserção simultânea de uma cauda poli-A formando um transcrito maduro de mRNA monocistrônico. A função exata do SL ainda não está completamente compreendida, mas há evidências de que esse confere estabilidade ao mRNA impedindo a sua degradação além de auxiliar a interação do mRNA com os ribossomos (Ullu e cols., 1996; Silveira, 2000). As moléculas de mRNA maduras nos tripanossomatídeos contêm o mini-exon na

extremidade 5' e uma cauda composta por aproximadamente 30 resíduos de adenina (cauda poliA) na extremidade 3' (Figura 5).

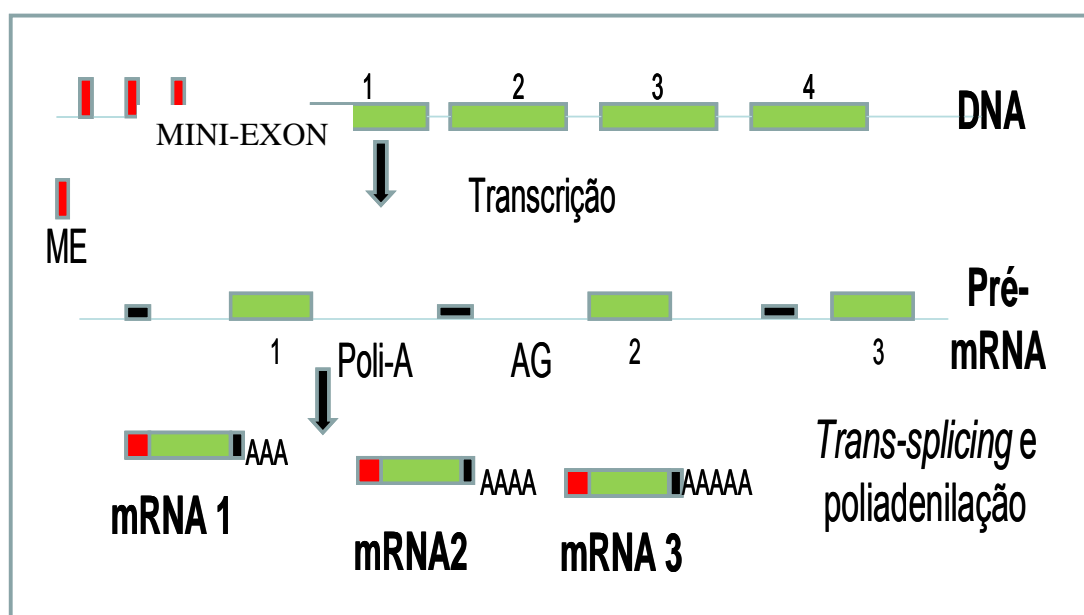


Figura 5 – Esquema adaptado de Bartolomeu e Teixeira (site Fiocruz / Doença de Chagas) ilustrando os processos de formação dos transcritos em tripanossomatídeos. Genes organizados *em tandem* (retângulos verdes) são transcritos policisticamente e as unidades de transcrição são individualizadas através do processamento do pré-mRNA. Sequências de polipirimidinas presentes nas regiões intergênicas, (retângulos pretos) coordenam as reações de clivagem entre os genes (indicadas pelas setas). Cada reação de clivagem está associada à (1) adição da sequência líder (SL / mini-exon) no sítio acceptor AG localizado abaixo de um segmento de polipirimidina (região 5'), e (2) adição da cauda poli-A no sítio de poliadenilação do gene anteriormente localizado (região 3'). Esses processos independentes resultam na produção de mRNAs monocistrônicos maduros.

O mecanismo de controle gênico desses organismos depende quase que exclusivamente de vias regulatórias atuando pós-transcricionalmente. A ausência de controle inicial implica em uma busca pelo conhecimento do processo pós-transcricional, tal como o processamento e estabilização, sendo crucial para o entendimento da expressão gênica nesses organismos (Haile e Papadopoulou, 2007).

Análises comparativas entre os genomas de *Trypanosoma cruzi*, *Trypanosoma brucei* e *Leishmania major* mostraram diferenças não apenas em relação ao ciclo de vida desses organismos, como também em relação aos mecanismos moleculares de controle de expressão gênica, tais como os genes de mudança de classe de glicoproteínas de superfície variáveis (VSG) e transcrição de mRNA por RNA polimerase I (El-Sayed e cols., 2005).

A maquinaria de *splicing* pode reconhecer dois tratos distintos de polipirimidina: um para as reações de *trans-splicing* e outro para poliadenilação (Campos e cols., 2008).

As regiões não traduzidas (UTRs)

De modo geral, os genes dos eucariotos apresentam segmentos em suas extremidades 5' e 3' que são transcritos, porém não são traduzidos sendo denominados regiões não traduzidas (UTRs) ou, de forma simplificada, 5' e 3'UTR (Brandão, 2006). Nesses segmentos, existem elementos do DNA que estão envolvidos na regulação da expressão gênica e na interação com os novos tipos de RNA desempenhando, dessa forma, importante papel regulatório pós-transcricional como a determinação da localização do transcrito, do tempo da transcrição, da proporção e da eficiência da transcrição, da estabilidade do transcrito,- molécula de mRNA, incluindo modulação do transporte de mRNA para o exterior do núcleo, além de sequências adicionais envolvidas no *splicing* (Van der Velden e Thomas, 1999; Jansen, 2001; Bashirullah e cols., 2001; Spence, 2009).

A característica estrutural da 5'UTR tem uma função maior no controle transducional do mRNA. Os RNAs mensageiros codificam proteínas envolvidas no processo de desenvolvimento, tais como os fatores de crescimento e de transcrição ou proto-oncogenes (Kozak, 1987). A seqüência que contém o primeiro códon AUG, em particular o códon localizado na região não traduzida, pode modular a eficiência com a qual é reconhecido como um códon iniciador transducional (Mignone e cols., 2002).

A ausência de um sinal canônico indicando o local exato da transcrição nos tripanossomatídeos é suprido pela existência de sequências de nucleotídeos como o trato de pirimidina (sinal de *trans-splicing*), pela presença do dinucleotídeo AG (sinal de *trans-splicing*), e pela poliadenilação na extremidade 3' fornecendo pontos de reconhecimento para o início da transcrição (Brandão, 2006).

A oportunidade criada pela grande quantidade de sequências nucleotídicas das regiões não traduzidas, oferece uma nova rota nos modelos e estratégias experimentais que traduzem o código do DNA dentro de fenômenos biológicos. Com exceção dos genes constitutivos (*house keeping*) e a importância funcional entre as espécies intimamente relacionadas, a composição e a característica estrutural das UTRs de eucariotos podem ser muito distintas (Mignone e cols., 2005). A inspeção

das sequências do *T. cruzi* mostrou que as UTRs em alguns genes podem ser bons marcadores moleculares e funcionais para a população de *T. cruzi* (Brandão, 2006). Pequenas variações nas UTRs (pequenas inserções ou deleções) podem ser o ponto inicial para o estabelecimento de relacionamento entre a expressão do gene e a estrutura da população. Pequenas mudanças nas UTRs podem afetar a expressão do gene (Weston e cols., 1999; Teixeira e cols., 1999) e isto pode ser uma estratégia evolucionária para ganhar vantagens funcionais sem grandes mutações nos próprios genes. O resultado dessas alterações nesses protozoários flagelados é evidenciado pela maior capacidade de sobrevivência desses parasitos em diferentes ambientes (Brandão, 2006).

Os sinais de controle transducional são freqüentemente encontrados ao longo da 5'UTR (Kozak, 1987). Uma importante região sinalizadora encontrada tanto em procariotos quanto em eucariotos é uma pequena região acima da fase de leitura aberta (*upstream open reading frame- uORF*), ou seja, acima da principal região codificadora. As *uORFS* podem estar envolvidas no controle da tradução agindo como reguladores da estabilidade do mRNA (Figura 6) (Lovett e Rogers, 1996; Morris e Geballe, 2000; Tran e cols., 2008).

Menos de 10% dos mRNA eucariotos contêm códons AUG dentro de regiões porções anteriores ao transcrito, sendo referidas freqüentemente, e erroneamente, como região 5' UTR. No entanto, esses códons são comuns em certas classes de genes, incluindo dois terços dos oncogenes e muitos outros genes envolvidos no controle de diferenciação e crescimento celular (Geballe e Morris, 2000).

As *uORFs* são divididas, de acordo com o papel que desempenham, em seqüência independente e seqüência dependente. A seqüência dependente codifica pequenos peptídeos e alguns desses peptídeos têm participação direta na via de tradução ribossomal ou na terminação transducional pela inibição da atividade da peptidil transferase presente nos ribossomos (Gaba e cols., 2001; Luo e Sachs, 1996).

Para a seqüência independente de *uORF*, os peptídeos codificados não são importantes para o controle da tradução, mas para outros fatores que podem afetar a eficiência da ORF como reconhecimento das *uORFs*, tamanho, localização do códon de parada e seqüências abaixo intercistrônicas (Vilela e McCarthy, 2003; Meijer e Thomas, 2002).

Várias uORFs afetam a tradução da próxima ORF por mecanismos que dependem de sequências de aminoácidos de peptídeos codificados. Experimentalmente, uORFs sequência-dependente são caracterizadas por uma modificação na tradução da ORF seguinte resultando numa mutação pontual mas não de uma mutação silenciosa dos códons dentro das uORFs (Geballe e Morris, 2000). É necessário cautela para a interpretação dos efeitos das mutações pontuais; as alterações nos nucleotídeos podem afetar a transcrição ou o processamento do mRNA e portanto alteram a expressão por uma via independente da transcrição (Geballe e Morris, 2000).

Os produtos de uORFs dos peptídeos e vários procaríotos e eucariotos parecem desempenhar importante papel nos mecanismos de controle transcricional. Alguns genes de resistência a antibióticos, os peptídeos produzidos de uma pequena uORF, se ligam e interferem na estrutura e na atividade do centro ribossomal da peptidiltransferase (Rogers e Lovett, 1996).

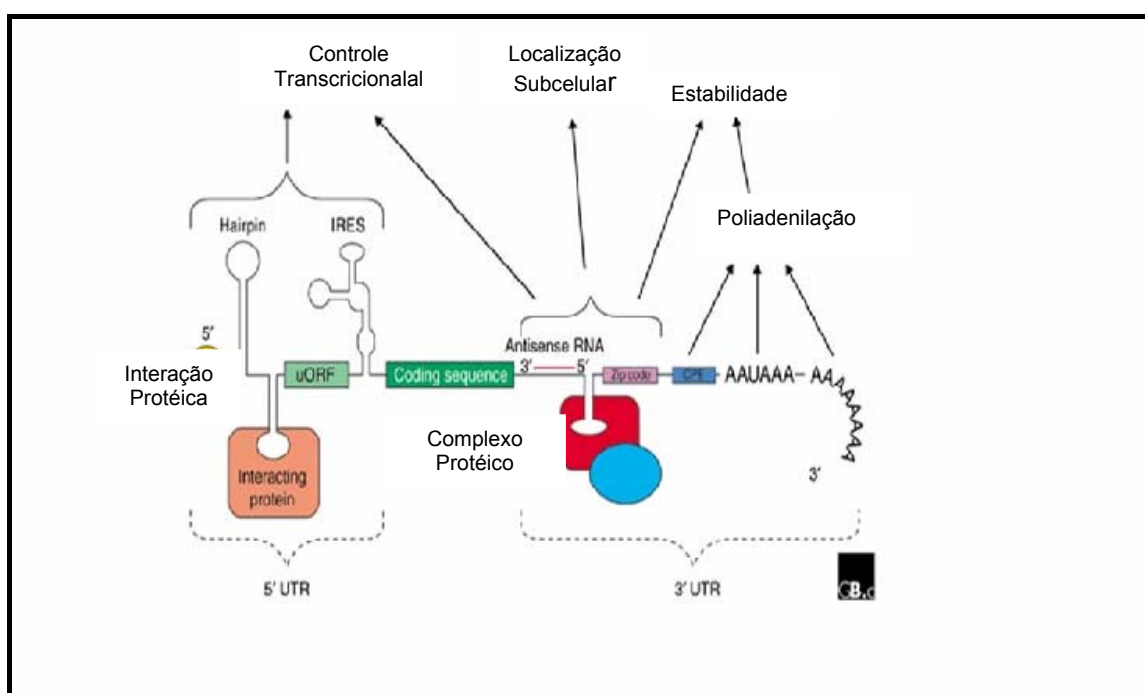


Figura 6 – A estrutura genética de um mRNA. A figura ilustra alguns reguladores pós-transcricionais que afetam a expressão gênica. Abreviaturas obedecendo ao sentido: 5' para 3': M7G (mini-exon); Hairpin (estrutura secundária); uORFs; IRES (sítio interno de entrada para o ribossomo); CPE (elemento de poliadenilação citoplasmático e sinal de poliadenilação (AAAAA)) (adaptado de Mignone e cols., 2002).

Pseudogenes

Os pseudogenes são definidos como sequências genômicas de DNA não funcionais que são originalmente derivadas de genes funcionais, contudo, eles apresentam características degenerativas tais como códons de parada (*stop codons*) prematuros (Balakirev e Ayala, 2003).

Há uma hipótese de que a função dos pseudogenes esteja envolvida durante o desenvolvimento do mRNA servindo como uma fonte de inibidores celulares (McCarrey Riggs, 1986). Os pseudogenes seriam, então, uma fonte de RNA anti-senso que se hibridizariam com o RNA senso de um determinado gene bloqueando a sua expressão. Essa hipótese tem sido confirmada ao se analisar a sequência de genes duplicados que são transcritos em direção oposta permanecendo conservados sempre que a interação senso - anti-senso é vantajosa (McCarrey e Riggs, 1986).

Portanto, há muitas especulações e algumas evidências de que os pseudogenes devam estar envolvidos em funções regulatórias para os genes que eles são derivados (Fotaki e Latrou, 1993).

Trans-sialidases do *T. cruzi*:

O ácido siálico e as sialidases desempenham importante papel na interação celular e na modulação do reconhecimento de microrganismos patogênicos pelas células hospedeiras de mamíferos (Schenkman e cols., 1994). A trans-sialidase (TS) compreende uma família de enzimas expressas e reguladas em tripanossomatídeos que catalizam a transferência do ácido siálico de glicoconjugados das células do hospedeiro para moléculas aceptoras presentes na membrana plasmática do parasito (Schenkman e cols., 1994). Várias evidências sugerem que a TS e os aceptores de ácido siálico na superfície do *T. cruzi* participam da interação parasito-hospedeiro e mediam o estágio inicial da invasão dos tripanossomas nas células do hospedeiro (Schenkman e cols., 1994).

A trans-sialidase de *Trypanosoma cruzi* (TcTS) é codificada por genes integrantes de uma ampla família gênica composta por 1430 membros geralmente chamados de superfamília de trans-sialidase; essa super família codifica centenas

de proteínas com tamanhos variados de 60 a mais de 200 kDa (Douglas e cols., 2008).

Os membros da família TcTS possuem diferentes funções na patogênese da doença de Chagas devido às suas diferenças estruturais e ao seu nível de expressão nos diferentes estágios do ciclo de vida do *T. cruzi* (Frasch, 2000). Em culturas axênicas de *T. cruzi* a forma epimastigota apresenta baixa atividade da TS (Chaves e cols., 1993). A atividade enzimática aumenta quando os parasitas entram na fase de crescimento estacionária (Chaves e cols., 1993).

A superfície do *Trypanosoma cruzi* no estágio invasivo de tripomastigota é coberto por numerosas trans-sialidases e por diversas “mucinas-like”, todas ancoradas à membrana por uma molécula de glicosilfosfatidilinositol (GPI) (Fasch, 2000; Buscaglia e cols., 2006). O *T. cruzi* não pode sintetizar ácido siálico e a principal função das formas TcTS ativas é adquirir unidades de ácido siálico de glicoconjugados presentes nos hospedeiros vertebrados e transferi-los para os resíduos de β -Gal nas mucinas que recobrem a membrana do parasito (Douglas e cols., 2008).

Durante o processo de invasão da célula hospedeira, a forma metacíclica, com grande atividade de TS, entra na célula formando um fagolisossomo conferindo proteção ao parasito. Logo após a invasão, os parasitos rompem a membrana do fagolisossomo, entram no citoplasma e transformam-se nas formas proliferativas, amastigotas, que não possuem TS (Frevet e cols., 1992; Rosenberg e cols., 1991). A atividade da TS reaparece quando amastigotas param de se multiplicar e transformam-se em tripomastigotas, antes de romperem as células infectadas (Frevet e cols., 1992; Rosenberg e cols., 1991). Uma quantidade significativa de TS é liberada no citosol da célula infectada nesse momento. A TS continua sendo expressa pelos tripomastigotas livres, e uma grande quantidade da enzima também é liberada em meio de cultura, secretada pelas formas epimastigotas (Pereira, 1983; Frevet e cols., 1992; Rosenberg e cols., 1991; Schenkman e cols., 1992; Kenneth e cols., 2008).

Quando o ácido siálico é transferido pela TcTS para a superfície do tripomastigota ele fornece uma proteção direta aos resíduos β -Gal nas mucinas de tripomastigotas do reconhecimento pelo sistema imune do hospedeiro, e uma proteção indireta dos resíduos α -Gal por suas mudanças de carga negativa (Schenkman e cols., 1999) além de uma resistência ao complemento e a

sobrevivência imediata dos tripomastigotas liberados na corrente sanguínea (Tomlinson e Rapper, 1998). A TcTS se difunde na corrente sanguínea, levando a uma clivagem do GPI ancorado, e a uma remoção do ácido siálico de superfície de plaquetas, causando plaquetopenia na fase aguda da doença de Chagas (Tribulatti e cols., 2005; Frevert e cols., 1992).

Relatos anteriores mostraram que tripomastigotas derivados de meio de cultura celular expressam e liberam maior quantidade de TS do que as formas metacíclicas diferenciadas em meio TAU (Rubin-de-Celis e cols., 2006). A diferença na expressão pode estar associada ao fato de que as células provenientes do meio de cultura celular escapam do vacúolo parasitóforo convertendo-se mais rapidamente em amastigotas do que as formas metacíclicas diferenciadas em meio TAU (Rubin-de-Celis e cols., 2006).

Os genes de trans-sialidase são clonados de várias cepas de *T. cruzi* e se mostraram pertencer a, pelo menos, duas classes diferentes de genes (Pereira e cols., 1991; Briones e cols., 1995), mostrando um alto grau de similaridade em relação as outras glicoproteínas de superfície do parasito (Nussenzweig e cols., 1994; Campetella e cols., 1992). Uma das classes gênicas, que representa a maioria dos genes das trans-sialidases, proteínas codificadas com um domínio catalítico na extremidade amino-terminal, e um número variável de repetições de 12 aminoácidos na extremidade carboxi-terminal (denominadas de CTR). Esta classe de genes é altamente expressa em tripomastigotas derivadas de células mamíferas (Schenkman e cols., 1996). Quando a trans-sialidase é purificada desses tripomastigotas, eles se comportam como uma proteína oligomérica formada por dois ou três polipeptídeos heterogêneos com tamanho molecular de 120 a 240 kDa (Eichinger e cols., 1994), com todos mostrando a mesma extremidade amino-terminal (Frasch e cols., 1993).

A segunda classe de genes codificadores de proteínas contém apenas o domínio catalítico da trans-sialidase. Esse tipo de enzima é encontrado em formas epimastigotas de cultura. A trans-sialidase purificada de epimastigota é um polipeptídeo monomérico de 90 kDa e não contém o domínio CTR (Chaves e Schenkman, 1993). As duas classes gênicas possuem o mesmo domínio catalítico (Nussenzweig e cols., 1994).

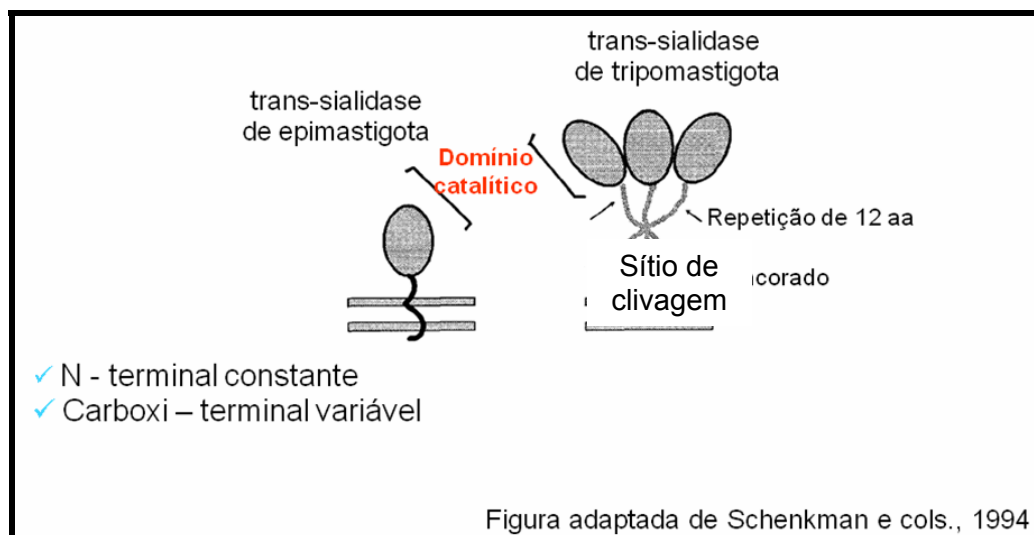


Figura 7 – Classes estruturais das trans-sialidases de *T. cruzi* das formas epimastigotas e tripomastigotas

Os genes de trans-sialidase são organizados em *tandem* e estão distribuídos em vários cromossomos do parasito (Frasch e cols., 1988, 1991). A trans-sialidase expressa em epimastigota é codificada por um gene diferente quando comparada a trans-sialidase expressa no estágio tripomastigota (Briones e cols., 1995).

Há algumas evidências de que os genes sem CTR são regulados por mecanismos pós-transcricional nos tripanossomatídeos, e que a estabilização do mRNA depende da região 3' não traduzida. Como essa porção é altamente divergente entre trans-sialidases de tripomastigotas, isto pode explicar diferentes mecanismos de regulação. Além disso, como o CTR não é requerido para a atividade catalítica (Shenkman e cols., 1993; Eichinger e cols., 1994; Frascch e cols., 1994), poderia se questionar a razão de sua presença nas trans-sialidases das formas tripomastigotas (Briones e cols., 1995).

A interiorização do *T. cruzi* na célula do hospedeiro vertebrado é um processo com múltiplos passos envolvendo o parasito e as moléculas que conduzem a mobilização de cálcio (Docampo e cols., 1996; Burleigh e Andrews, 1998; Yoshida, 2003). Para invadir células de mamíferos, os tripomastigotas metacíclicos e os tripomastigotas oriundos de cultura tecidual captam diferentes grupos de moléculas de superfície que interagem diferencialmente com os componentes do hospedeiro. A glicoproteína gp82 específica de tripomastigotas metacíclicos se liga nas mucinas gástricas (Neira e cols., 2003), enquanto membros da família gp85 dos tripomastigotas de cultura tecidual se ligam aos componentes da matriz extracelular (Quaissi e cols., 1986; Giordano e cols., 1994).

Embora a gp82 dos tripomastigotas metacíclicos e a gp85 tripomastigotas de cultura tecidual apresentem propriedades adesivas diferenciais e interação com receptores diferentes nas suas células alvo, elas são moléculas relacionadas, que possuem considerável identidade entre as sequências e estão incluídas na superfamília das sialidases/gp85 de *T. cruzi* (Araya e cols., 1994; Colli e Alves, 1999).

Outras moléculas, além das TS e gp85/82, também estão envolvidas com o mecanismo de interferência na sinalização da célula hospedeira que culmina com o aumento do cálcio intracelular e migração dos lisossomas para o ponto de contato entre o *T. cruzi* e a célula hospedeira.

A cruzipaína purificada aumentou a invasão do parasito na célula do hospedeiro e disparou a mobilização de cálcio indicando importante papel na geração de agonista para as cininas. A transdução do sinal mediado pelas cininas depende do tipo de célula e do isolado do parasito usado (Scharfstein e cols., 2000).

A oligopeptidase B de *T. cruzi* é um fator solúvel de estrutura desconhecida secretado por células de cultura de tripomastigotas. Este fator é gerado pela ação de uma peptidase alcalina de 120 kDa presente nos precursores infectivos do tripomastigota metacíclicos (Burleigh e Andrews, 1995). Utilizando as metodologias da clonagem e do seqüenciamento com os cDNAs correspondentes, peptidases de células de cultura de tripomastigotas foram identificadas como uma enzima citosólica relacionadas as prolinas e membros da família das serina endopeptidases denominadas oligopeptidase B de *T. cruzi* (Burleigh e cols., 1997).

Tripomastigotas mutantes com oligopeptidase B nula são defeituosos na mobilização de cálcio de estoques presentes nas células dos mamíferos, e no estabelecimento da infecção tanto *in vitro* quanto *in vivo* (Caler e cols., 1998). Baseado em evidências experimentais, tem sido proposto que a geração de agonistas de cálcio pela oligopeptidase B é exportado do parasito ao se ligar a receptores na superfície das células alvo, ativando fosfolipase C e gerando IP3 o qual se liga aos receptores de membrana do retículo endoplasmático e promove a liberação de cálcio (Yoshida, 2006).

Mecanismo molecular da invasão celular por *T.cruzi*:

As características relacionadas a cada cepa do *T. cruzi* são determinantes durante o processo de invasão à célula do hospedeiro vertebrado, devido ao tropismo celular individualizado e da via de sinalização que é ativada (Mortara e cols., 2005; Neira e cols., 2002).

Membros da família da TS junto com glicoproteínas de superfície, tal como a gp85, formam uma população heterogênea de glicosilfosfatidilinositol (GPI) ancorados a glicoproteínas de superfície de tripomastigotas que desempenham importante papel no início do reconhecimento da célula hospedeira (Burleigh e Woolsey, 2002). A atividade da TS também tem sido relacionada com o rompimento do vacúolo parasitóforo pelo *T. cruzi*. Uma variedade de receptores de superfície celular, segundo mensageiros e fatores de transcrição de diferentes vias parecem ser usados para mediar e afetar a invasão e sobrevivência do parasito (Burleigh e Woolsey, 2002).

Os tripomastigotas invasivos ao entrarem em contato com os receptores de superfície na membrana da célula do hospedeiro vertebrado disparam, através de uma cascata de sinais, a liberação de cálcio no meio intracelular (Moreno e cols., 1994; Tardieux e cols., 1994; Caler e cols., 2000; Scharfstein e cols., 2000). A alteração no nível de cálcio regula a entrada gradual do parasito na célula por promover um rearranjo de actina (Rodríguez e cols., 1995) bem como pela fusão dos lisossomos com a membrana plasmática formando um vacúolo derivado da membrana celular (Rodríguez e cols., 1997).

Na interação parasito/hospedeiro, entre os vários eventos, existe aquele que culmina no aumento da concentração intracelular de cálcio tanto no parasito, quanto na célula hospedeira.

A ligação da molécula gp82 nas células do hospedeiro vertebrado dispara um aumento da concentração de cálcio intracelular, (Yoshida e cols., 2003). Epimastigotas não infectivos não induzem a sinalização por cálcio (Tardieux e cols., 1994; Dorta e cols., 1995) a menos que eles sejam transfectados com vetor de expressão carreando cDNA de gp82 de *T. cruzi* (Manque e cols., 2003). A ligação à célula do hospedeiro vertebrado libera um sinal para o parasito que resulta na ativação da proteína tirosina quinase (PTK) e na fosforilação da p175, uma proteína que não é detectável e não infectiva de epimastigota (Favoreto e cols., 1998).

A sinalização da cascata induzida por gp82 inclui a participação da fosfolipase C (PLC), que gera inositol 1,4,5-trifosfato (IP3). Agonistas de cálcio se ligam aos receptores presentes na membrana da célula hospedeira formando um complexo agonista/receptor que deverá ativar a proteína fosfolipase C. A PLC ativa hidrolisa fosfolípídeos de membrana, gerando a molécula de diacilglicerol (DAG) e IP3, que se liga aos receptores presentes na membrana do retículo endoplasmático abrindo os canais de cálcio e liberando, dessa forma, o cálcio para o meio citosólico. O cálcio no citosol se ligará a molécula DAG e o produto dessa ligação ativará a proteína quinase C (PKC) que desencadeará reações de fosforilação e diversas respostas celulares (Yoshida, 2006) (Figura 7).

Diferente de muitos patógenos intracelulares, que evitam a fusão com os lisossomos das células do hospedeiro, a residência no interior do vacúolo parasitóforo é um pré-requisito para a sobrevivência do *T. cruzi* (Ley e cols., 1990). A exposição do *T. cruzi* ao ambiente ácido do vacúolo lisossomal é requerido para a atividade das moléculas secretadas de *T. cruzi* formadoras de poros, denominadas TcTox, que ajudam no rompimento do vacúolo (Andrews e Whitlow, 1989; Andrew, 1990). Além disso, o pH ácido parece ser importante para a diferenciação da forma tripomastigota em amastigota (Tomlinson e cols., 1995). O AMP cíclico (nucleotídeo gerado a partir do ATP pela adenilil ciclase em resposta a estimulação da proteína G por muitos tipos de receptores de superfície celular) é um agonista que promove o movimento dos lisossomos até a periferia da célula, além de facilitar a exocitose de lisossomos regulados por cálcio, aumentando dessa forma, a invasão pelo parasito (Rodríguez e cols., 1999).

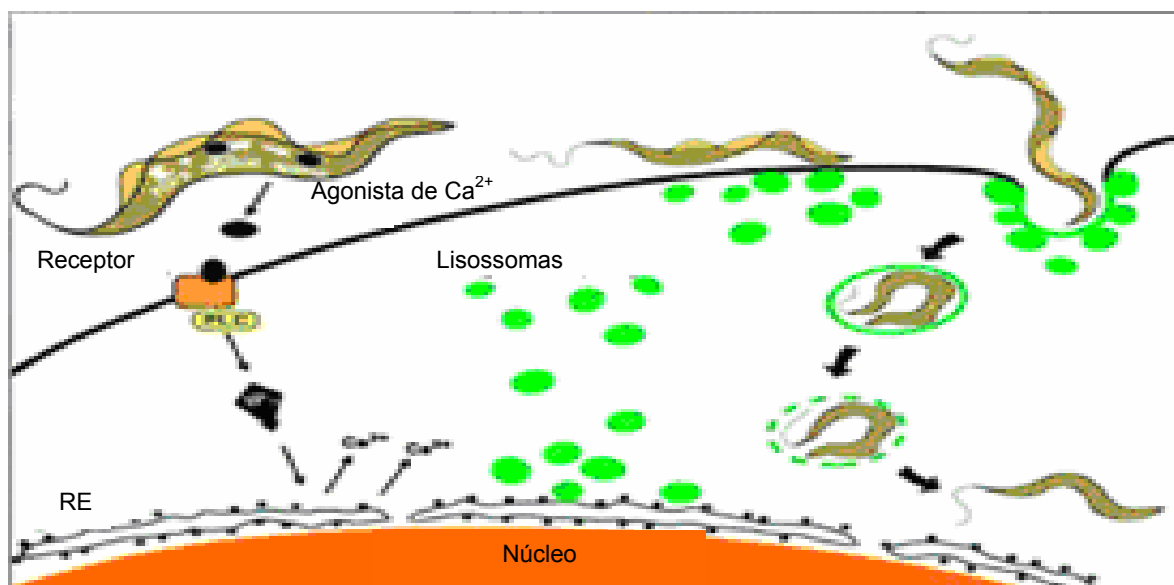


Figura 8 - A figura representa um esquema de invasão celular pelo *T. cruzi*. Os tripomastigotas infectivos se associam com as proteínas de superfície das células do hospedeiro vertebrado, disparando vias de sinalização através da fosfolipase C (PLC) que levarão ao recrutamento e a fusão com lisossomos presentes no citoplasma de células do hospedeiro vertebrado, através da mobilização de cálcio e gerando um vacúolo parasitóforo. Os parasitos são liberados no citoplasma após o rompimento desse vacúolo, e se diferenciam na forma amastigota 24h após a infecção (Andrews e cols., 2000).

Como as TS são proteínas fundamentais para o mecanismo de reconhecimento e invasão do parasito, além de possibilitar a capacidade de adaptação do mesmo ao utilizar o ácido siálico presente na superfície das células do hospedeiro vertebrado, uma caracterização mais profunda da estrutura do mRNA e de sua 5' UTR podem assistir ao desenvolvimento de novas estratégias de profilaxia da doença de Chagas ou mesmo no desenvolvimento de vacinas.

Considerando a necessidade específica do *T. cruzi* para obter ácido siálico através da trans-sialidase como mediadora da transferência do hospedeiro para o parasito, entre outras diversas funções dessa enzima na patogênese da doença de Chagas, é que podemos destacar essa enzima como um potencial alvo para drogas (Douglas e cols., 2008).

Dessa forma, visa-se avaliar os polimorfismos na região 5'UTR entre diferentes linhagens, tais como os sítios de *trans-splicing* adicionais ou múltiplos em TS e seus respectivos sinais (região rica em polipirimidina), variação composicional e tamanho da região das sequências. Portanto, uma investigação mais profunda na estrutura do mRNA da família de TS se faz necessária com o intuito de contribuir para o entendimento da evolução deste gene nas populações de *T. cruzi* a fim de

serem utilizados em estudos taxonômicos e auxiliar no desenvolvimento de novas abordagens para análise da expressão de genes de cópias múltiplas em protozoários parasitas.

1 OBJETIVO

1.1 Objetivo geral

Identificar e caracterizar as regiões não-traduzidas (UTRs) dos genes de trans-sialidases de *Trypanosoma cruzi* a partir de seqüências depositadas no GenBank (NCBI), e investigar comparativamente a transcrição desses genes em cepas do parasito classificadas como *T. cruzi I*, *T. cruzi II* e Zimodema III.

1.2 Objetivos específicos

- Realizar análise computacional das 5' UTRs de todas as cópias de trans-sialidases descritas pelo projeto genoma de *T. cruzi* para avaliação estrutural e composicional dessas regiões não traduzidas;
- Analisar a presença de mutações e sítios indicativos de *trans-splicing* na região não traduzida (5'UTR) das trans-sialidases nos três grupos populacionais do parasito (*T. cruzi I*, *T. cruzi II* e Zimodema III);
- Analisar qualitativamente por RT-PCR convencional e quantitativamente por RT-PCR em Tempo Real os níveis de expressão dos genes das trans-sialidases nas cepas estudadas: Dm28c (*T. cruzi I*), Y (*T. cruzi II*), CL-Brener (*T. cruzi II*), INPA 4167 (ZIII), JJ(ZIII) , 3663 (ZIII) e Colombiana (ZIII) .

2 MATERIAL E MÉTODOS

2.1 Análise computacional das seqüências de UTR:

A fim de se construir um arquivo com seqüências de 5' UTR, fez-se uma busca inicial no banco de dados do Genbank-NCBI (versão 158) com as palavras-chaves *Trypanosoma cruzi* e *trans-sialidase*. Essa pesquisa resultou em várias seqüências de UTR putativas redundantes (seqüências repetidas com 100% de identidade) que, após eliminação das mesmas, foi criado um arquivo de 735 seqüências selecionadas. Posteriormente, estabeleceu-se, para cada uma das seqüências, um limite para a abrangência das extremidades 5' UTR que consistiu em um fragmento de 600 nucleotídeos acima do códon iniciador - com o objetivo de se identificar a localização de mini-exons - (provável 5' UTR). Todas essas 735 seqüências foram editadas em formato FASTA.

Em seguida, foi realizado o alinhamento das 735 seqüências utilizando o programa Clustal X (versão 1.83), cujo resultado foi editado no programa MEGA 4.0 a fim de se agrupar as seqüências idênticas.

A análise dos números de acesso obtidos na busca contra o dbEST e a seleção das seqüências com e sem mini-exon geraram dois grupos denominados de *mexon* (referente ao grupo com mini-exon) e *nomexon* (referente ao grupo que não contém mini-exon). Essa classificação foi feita através da observação da presença de parte do fragmento do mini-exon (apenas os 8 últimos nucleotídeos do fragmento de mini-exon: CTATATTG).

Em seguida, foi realizada a junção dos dois grupos de arquivos (*mexon* + *nomexon*) para que se pudesse realizar o alinhamento de todas as seqüências relacionadas à *trans-sialidase* de *T. cruzi*, objetivando o agrupamento de seqüências idênticas. Tais alinhamentos foram feitos de forma sucessiva, mediante a eliminação de redundâncias, e, posteriormente, edições dos fragmentos, igualmente sucessivas no programa MEGA 4.0 até, finalmente a obtenção de um grupo único não-redundante.

2.2 Cultivo de *Trypanosoma cruzi*

Foram cultivadas células no estágio epimastigota das cepas Dm28c (*T. cruzi* I), Y (*T. cruzi* II), CL-Brener (*T. cruzi* II, cepa híbrida), INPA4167 (zimodema III), 3663 (zimodema III) e Colombiana (zimodema III, de acordo com o perfil do mini-exon), a 28° C em meio BHI (*Brain Heart Infusion*) acrescido com 10% de soro fetal bovino estéril.

2.3 Extração de RNA total de *T. cruzi*

RNA total foi extraído de epimastigotas de todas as cepas cultivadas em fase exponencial de crescimento para posteriores análises da expressão diferencial dos genes de trans-sialidase, por RT-PCR (transcriptase-reversa PCR) e por PCR em tempo real. Após o cultivo, as formas epimastigotas foram centrifugadas a 2.500 rpm por 10 minutos a 4° C. O RNA total foi extraído por TRIZOL (Invitrogen®), segundo recomendações do fabricante. Pureza, quantidade e integridade do RNA foram definidas por espectrofotometria e eletroforese em gel de agarose (3 µL de RNA total recém – extraído de cada cepa acrescidos de 10 µL de formamida, foram incubados por 10 minutos antes de serem aplicados no gel a fim de eliminar as estruturas secundárias inerentes ao RNA e melhorar a visualização do mesmo). Cinco microgramas de RNA purificado foram utilizadas nos subseqüentes ensaios de RT-PCR.

2.4 RT- PCR qualitativo

A técnica de RT-PCR (Transcriptase Reversa - Reação em Cadeia da Polimerase) foi utilizada para se amplificar especificamente as regiões não traduzidas 5' (5' UTR) dos cDNAs de trans-sialidase. Inicialmente, realizou-se a síntese da primeira fita do cDNA com 2µL de óligo dT ou com 2µL de Random Primer® (Invitrogen) diluídos na proporção de 1:10. Para tal, 5µg de RNA na presença de 1µL de dNTPs (5mM) foram incubados a 65 °C por 5 mins. Em seguida foram adicionados 4µL de First Strand Buffer, 2µL de 0,1M DTT, 1µL da

transcriptase reversa M-MLV (Invitrogen®) e 2 µL de óligo dT (10 µM) em uma nova incubação a 37 °C por 1 h.

Posteriormente, foi realizada a amplificação da 5' UTR, através da utilização de 1 µL do iniciador específico para 5' UTRTCNA (5' TTCTCCGCGGCAATCCCCCT 3') em combinação com o iniciador para mini-exon (5' CGCTATTATTAGAACAGTTTCT 3'); . A mistura de reação consistiu de 2 µL de cada iniciador a 10 µM , 2 µL de dNTPs (5 mM), 2 µL de MgCl₂ (2 mM), 5 µL tampão de PCR 10X , 0,2 µL da enzima Taq DNA polimerase recombinante (5u/µL) (Invitrogen®), 2 µL de cDNA referente a etapa anterior a PCR, e água para completar o volume final de 48 µL. O regime térmico foi : 35 ciclos de amplificação a 96 °C (10 segundos), 55 °C (30 segundos), 72 °C (30 segundos), seguido de extensão final a 72 °C por 10 minutos.

Os produtos desta amplificação referentes a cada cepa analisada foram clonados (TOPO-TA cloning Kit, Invitrogen® ou pelo pGEM T-Easy Vector, kit Promega®), seqüenciados e analisados quanto a presença de mutações pontuais, inserções e deleções nas regiões não-traduzidas.

2.5 Preparo de células competentes

A metodologia foi iniciada com uma semeadura da cultura de *Escherichia coli* linhagem TOP10 em placa contendo LB sólido sem antibiótico, com crescimento *overnight* a 37 °C. Uma colônia foi retirada e transferida para um erlemeyer de 200 mL contendo LB líquido (Luria-Bertani) sem antibiótico, com forte agitação a 37 °C. Depois de 1h foi iniciada a leitura da D.O. a 550nm e depois, em intervalos de 20 - 30 minutos até chegar ao valor de D.O. = 0.5 (fase log). A cultura foi incubada em banho de gelo por 10 minutos e centrifugada a 4000 rpm por 10 minutos em tubos cônicos de 50 mL estéril. O sobrenadante foi descartado e o pellet foi ressuspensão cuidadosamente com 25 mL de CaCl₂ estéril a 100 mM. Uma nova centrifugação foi realizada por 10 minutos a 4000 rpm. O sobrenadante foi descartado e o pellet foi ressuspensão, cuidadosamente, em 25 mL de CaCl₂ estéril a 100 mM. Uma nova centrifugação foi realizada por 10 minutos a 4000 rpm. O sobrenadante foi descartado e novamente o pellet foi ressuspensão, cuidadosamente, em 2 mL de CaCl₂ estéril a 100 mM acrescido de 10% de glicerol . Foram feitas alíquotas de 100 µL em tubos eppendorf® de 1,5 mL mantendo-as no gelo. Em seguida, as células

foram submetidas ao choque térmico com nitrogênio líquido ou gelo seco em álcool por 1 minuto e rapidamente estocadas em freezer $-70\text{ }^{\circ}\text{C}$.

Com a finalidade de testar a competência dessas células, inoculou-se uma alíquota das colônia bacteriana em 5 mL de meio de cultura LB acrescido do antibiótico seletivo para o plasmídeo utilizado (ampicilina $100\text{ }\mu\text{g/mL}$), e foi realizada uma reação de transformação com o plasmídeo controle (PUC19).

2.6 Clonagem

A obtenção do inserto foi através da técnica de RT-PCR, uma vez que o material de partida foi o RNA total. Para os experimentos de clonagem foram utilizados dois vetores: TOPO TA kit Invitrogen® e o vetor pGEM T-Easy Vector kit Promega®.

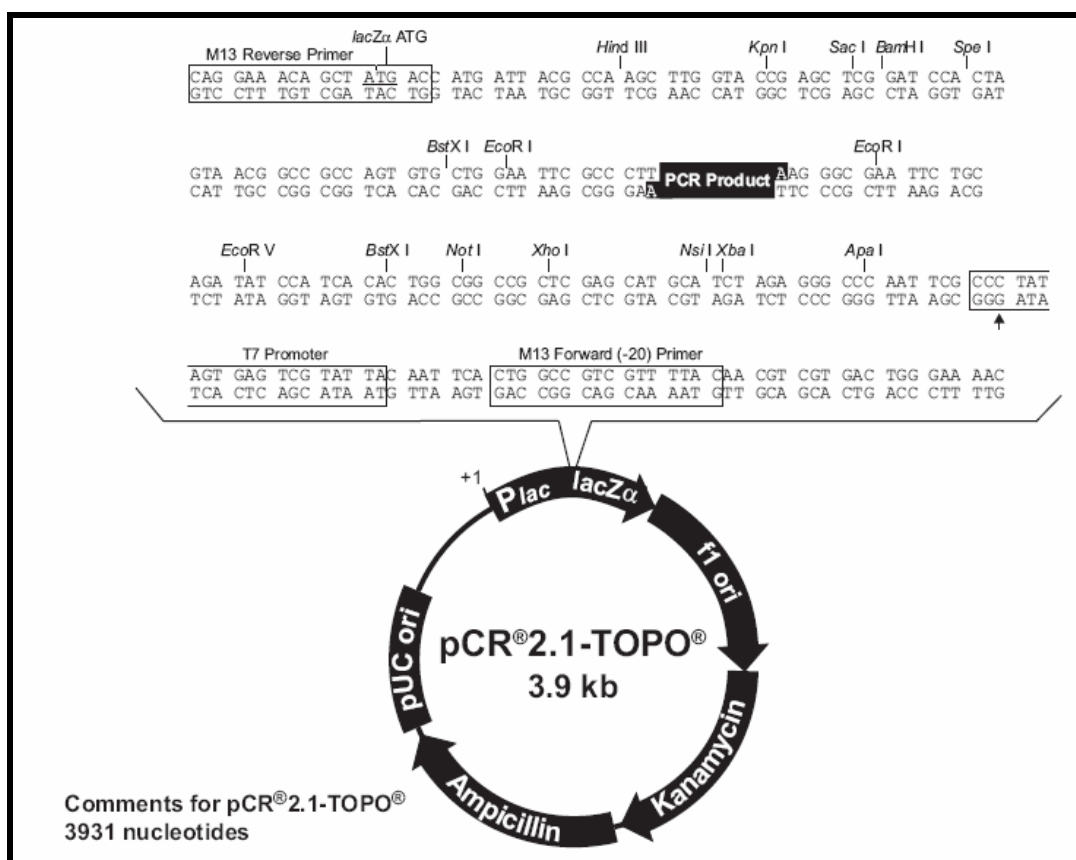


Figura 9 – Topo TA utilizado na primeira etapa da experimentação da clonagem

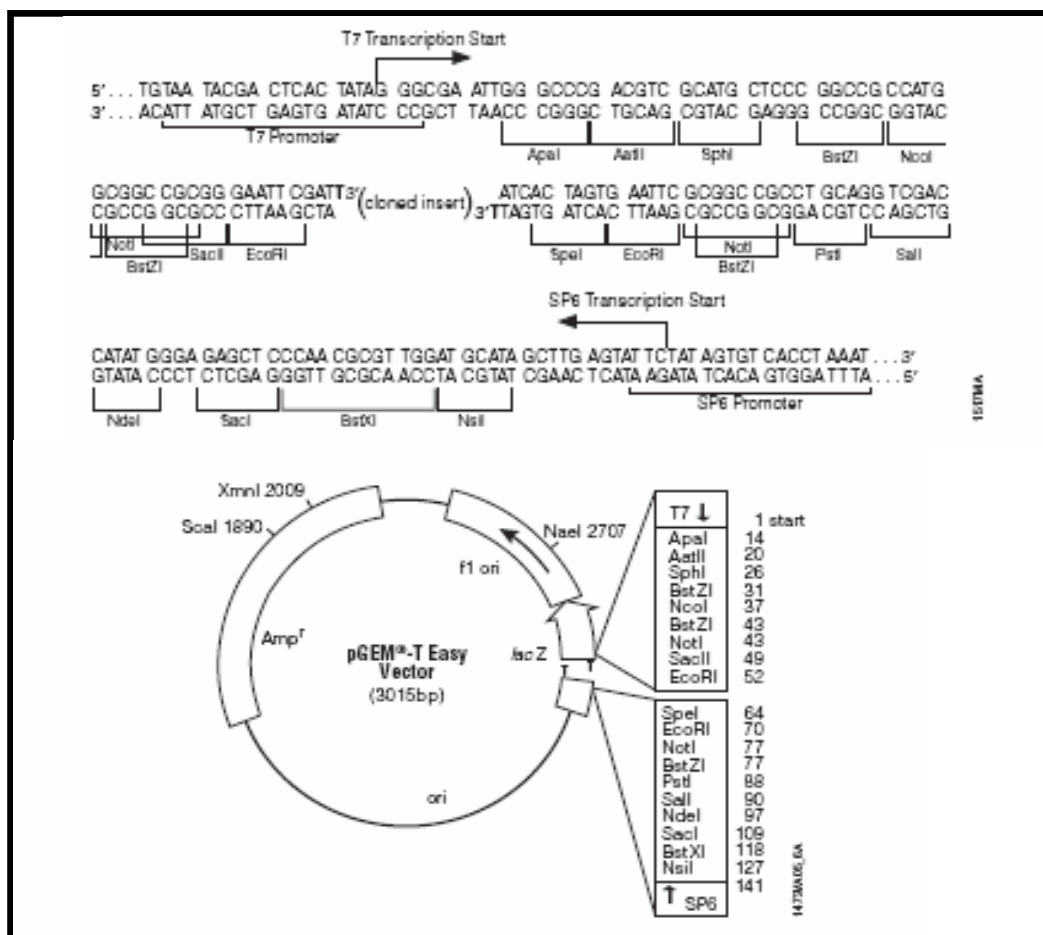


Figura 10 – Vetor da Promega pGem-T easy utilizado na segunda etapa da experimentação da colnagem

Após a etapa de ligação, adicionamos 2 μL da reação de clonagem em um tubo eppendorf® de 1,5 mL contendo previamente as células competentes de *E. coli* da linhagem TOP10 e homogeneizamos levemente. As células foram incubadas em gelo por 15 minutos e, em seguida, foi realizado o choque térmico a 42 °C por 1 minuto. Posteriormente, foi acrescentado 1 mL de LB sem antibiótico ao mesmo tubo eppendorf®, e as células foram incubadas a 37 °C por 50 minutos concluindo, dessa forma, a etapa de transformação.

Placas de LB contendo ampicilina na concentração final 100 $\mu\text{g/mL}$ foram subsequentemente utilizadas, tendo sido adicionados 40 μL de X-Gal em cada placa a fim de selecionar as bactérias que continham o plasmídeo com a região de resistência a ampicilina. O X-Gal indica se a bactéria expressa a enzima β -galactosidase, codificada pelo gene Lac Z (colônia azul); bactérias contendo o plasmídeo recombinante formam colônias brancas (gene Lac Z inativo).

Após selecionarmos as colônias brancas com auxílio de um palito previamente autoclavado, cada palito carreando uma colônia bacteriana foi

introduzido em um tubo Falcon® de 15 mL contendo 3 mL de meio LB com ampicilina na concentração final de 100 µg/mL, e as células foram cultivadas a 37°C *overnight*.

Após o cultivo, as células foram transferidas para tubos *ependorf*®, e centrifugadas a 14000 rpm por 20 minutos. O sobrenadante foi desprezado e o pellet mantido a -20 °C para a etapa subsequente de extração do DNA plasmidial (Minipreparação/ Miniprep).

2.7 Sequenciamento

Após a clonagem foi realizada a extração do DNA plasmidial utilizando-se o kit Wizard® Plus SV Minipreps DNA Purification System (Promega®), segundo as especificações do fabricante. As reações de sequenciamento foram processadas com o *ABI Prim™ Dye Terminator Cycle Sequencing Kit* (Applied Biosystems), e os dados armazenados no *software* de análise das sequências de DNA (Plataforma de Sequenciamento PDTIS/IOC/Fiocruz).

2.8 Eletroforese em gel de agarose

Utilizamos eletroforese em gel de agarose a 1% (g/mL) (Sigma) para as análises dos produtos obtidos pela técnica de RT-PCR convencional e para analisar a qualidade do RNA após a sua extração. O tampão de corrida usado foi TBE 0,5 X. Foram adicionados 2 µL de tampão de amostra a 10 µL de cada amostra, totalizando um volume de 12 µL para ser aplicado no gel. Foi utilizado como marcador de peso molecular o 100 bp ladder (Invitrogen□). Após a corrida, o gel foi corado com brometo de etídeo a 250 mg/mL (invitrogen®) e os resultados foram registrados com o auxílio do sistema de fotodocumentação de imagem de gel (Bio-Rad□).

2.9 Eletroforese em Gel de Poliacrilamida

A concentração do gel de poliacrilamida foi de 10,5%. Utilizamos 3,5 mL de acrilamida 30%, 1 mL de TBE 10X, 1,5 mL de glicerol, 3,6 mL de água, 0,4 mL de persulfato de amônio 1,6% e 3,5 mL de TEMED. Foi adicionado 1 µL de tampão de amostra a 7 µL da amostra, totalizando um volume de 8 µL para ser aplicado no gel. Foi utilizado como marcador de peso molecular o 100 bp ladder (Invitrogen®). Após a corrida, o gel foi revelado com prata utilizando-se o DNA Silver Staining Kit (Amersham Biosciences®), segundo as recomendações do fabricante.

2.10 RT-PCR quantitativo (PCR em tempo real) - Sistema SYBR Green

A técnica foi adotada com o objetivo de detectarmos e quantificarmos simultaneamente os transcritos para o gene da trans-sialidase, contendo a região 5' UTR. Utilizamos o sistema que emprega o corante fluorescente SYBR Green (Applied Biosystems) o qual se intercala inespecificamente na dupla fita de DNA sintetizada, gerada em cada ciclo de amplificação. Foram utilizados os mesmos iniciadores (5' UTRCNA e mini-exon), empregados previamente nos ensaios de RT-PCR convencional. Paralelamente, as amostras foram também testadas com outro conjunto de iniciadores para o gene da calmodulina, como controle endógeno (gene constitutivo para normalização da reação), empregando-se os primers CAL 3 (5'CCCGACGGAGGCGGAGCTGC 3') e CAL 5 (5'GTCCACGTCCGGCCTCGCGGA 3'). Ambos iniciadores foram previamente desenhados e baseados na seqüência do gene da calmodulina.

Os ensaios quantitativos foram realizados no equipamento ABI Prism 7000 (Sequence Detector System – Applied Biosystems), da Plataforma de PCR em Tempo Real PDTIS/FIOCRUZ. Em cada corrida, a discriminação do produto específico para cada cepa de *T. cruzi* utilizada, foi realizada através da análise da curva de dissociação em função do *melting point* geradas (T_m). Os ensaios de PCR em tempo real foram otimizados em volume final de 15 µL nas seguintes condições: 0,8 µmol de cada iniciador, 5 µg de cDNA e 5,9 µL água (ultra pura e estéril, grau Biologia Molecular).e 12 µL do kit *Power SYBR green Universal PCR Master Mix* (Applied Biosystems) e as seguintes condições de ciclagem: passo inicial a 96 °C

por 10 minutos (ativação da enzima Taq DNA polimerase), seguida de 40 ciclos a 96 °C por 10 seg, 55 °C por 30 seg e 72 °C por 30 seg, com uma extensão final de 10 minutos a 72 °C.

O *threshold cycle* (Ct) de uma determinada amostra corresponde ao número de ciclos onde foi registrado um aumento da intensidade de fluorescência acima da faixa basal. O valor de Ct da amostra é calculado durante a fase exponencial da reação de amplificação.

A quantidade relativa dos produtos amplificados foi determinada a partir da diferença entre o Ct referente ao gene da trans-sialidase e o respectivo Ct do gene endógeno (normalizador), através da fórmula de ΔCt : [Ct (trans-sialidase) – Ct (calmodulina)]. Após a normalização (ΔCt) calibramos os resultados obtidos de cada cepa analisada através do cálculo do $\Delta\Delta Ct$ (ΔCt : de cada cepa - ΔCt : da cepa referência). Os dados de quantificação relativa foram expressos em $2^{-\Delta\Delta Ct}$ (Livak e Schmittgen, 2001) utilizando como referência a cepa menos expressa entre as analisadas.

3 RESULTADOS

3.1 Extração de RNA total e RT-PCR qualitativo

Como os genes de trans-sialidase são transcritos relativamente em abundância, mesmo no estágio epimastigota, optamos por não utilizar a seleção de mRNA com poli-A, normalmente indicada quando se trata de genes de baixa expressão. O RNA total foi extraído com o método trizol, o qual demonstrou ser adequado para as etapas de transcrição reversa e PCR (Figura 8). Mesmo no caso das cepas com elevada replicação (como o isolado 4167, que após 4 dias de cultivo em condição padrão atingiram a saturação, com D.O. 600 nm em torno de 0.3 unidades de absorbância), a presença de abundante RNA ribossomal e outros RNA menores não prejudicaram a etapa subsequente.

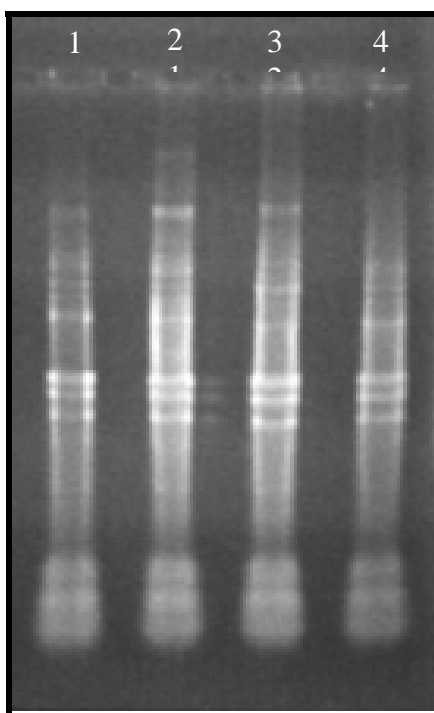


Figura 8 - RNA total das cepas: CL-Brener (canal 1), Dm28c (canal 2), Y (canal 3) e 4167 (canal 4)

Após a extração do RNA e sua conversão a cDNA, a próxima etapa foi estabelecer uma padronização para a RT-PCR qualitativo. Foram usados os *primers* 5'utrtna e ME2 e testadas várias condições de reação empregando-se os cDNAs de CL-Brener, Y, Dm28c e INPA4167. O resultado desta padronização monitorada em eletroforese em gel de agarose está mostrado na figura 9. Relevante neste experimento é a observação de fragmentos adicionais em cepa zimodema III (4167)

na região em torno de 300 e 400 pb, e um fragmento predominante em torno de 400 pb para a cepa Y, indicando a possibilidade de *trans-splicing* adicional ou de cópias com mutações no sítio primário de *trans-splicing*. Essas observações foram posteriormente confirmadas com o sequenciamento desses fragmentos. As cepas Dm28c e o clone CL-Brener apresentaram perfil eletroforético similar com uma banda principal em torno de 300 pb.

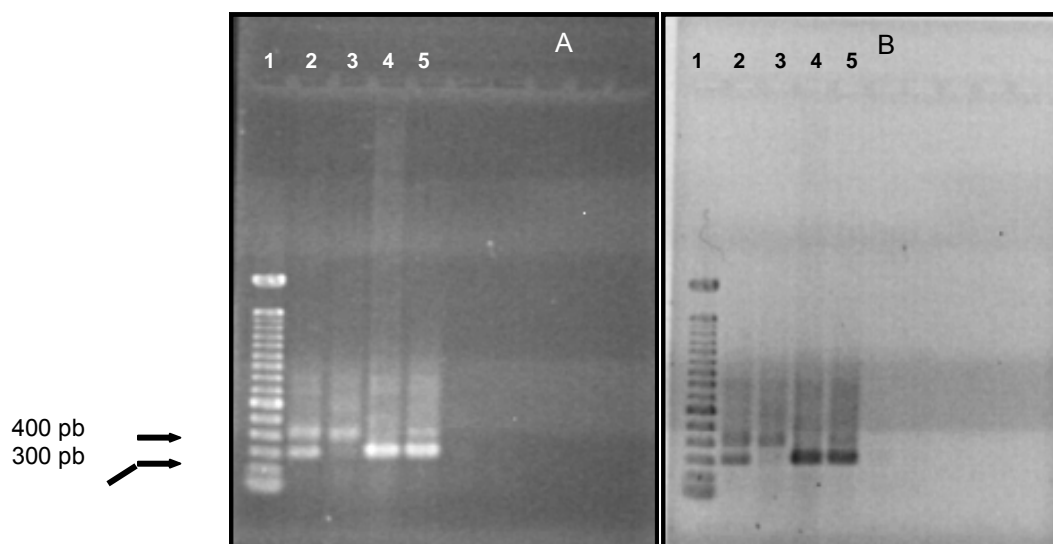


Figura. 9 – Gel de agarose mostrando os produtos de um RT-PCR qualitativo, Canal 1 - marcador de tamanho molecular de 100pb, Canal 2 – cepa INPA4167, Canal 3 - cepa Y, Canal 4 – cepa Dm28c, e Canal 5 -clone CL-Brener. São apresentadas uma versão positiva (gel A) e outra negativa (gel B) da mesma imagem do gel para facilitar a identificação dos fragmentos. As setas representam o marcador em 300 e 400 pb.

A fim de se obter uma melhor resolução do tamanho desses fragmentos, as mesmas amostras aplicadas no gel de agarose foram submetidas a eletroforese em gel de poliacrilamida (10.5%). Como era esperado, o que parecia ser uma única banda no gel de agarose revelaram-se múltiplos fragmentos no gel de poliacrilamida (figura 10 e figura 11). Corroborando mais uma vez a presença de potenciais sítios múltiplos de *trans-splicing* nas trans-sialidases, como poderia ser deduzido pela inspeção inicial das sequências já disponíveis. O inesperado neste caso foi o comportamento da transcrição em algumas cepas, concentrando em fragmentos maiores (cepa Dm28c) ou com sítios múltiplos (cepa Y e zimodema III). Isto revela uma complexidade ainda maior desta classe de trans-sialidases sob a perspectiva das linhagens filogenéticas de *T. cruzi*.

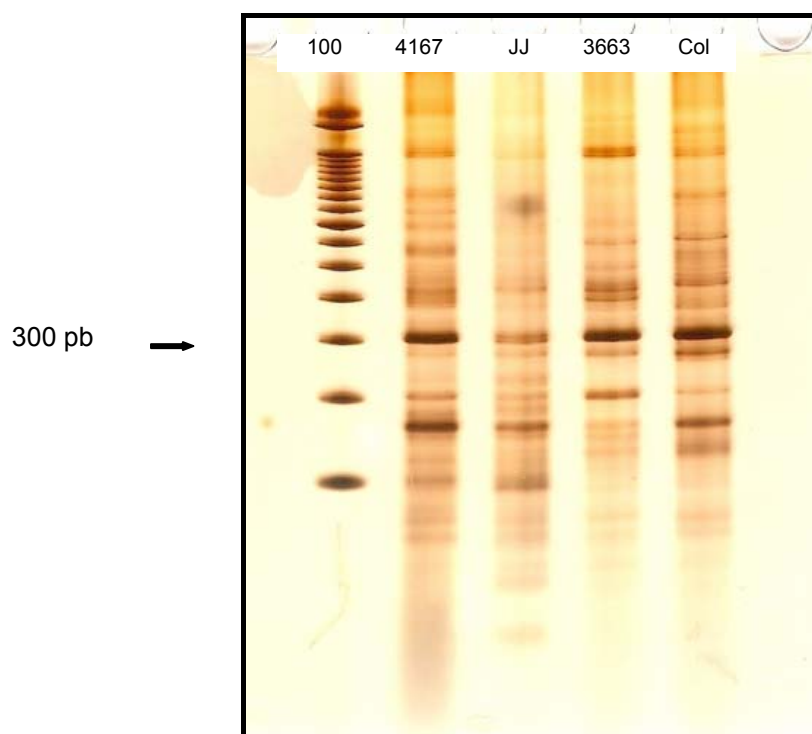


Figura 10 – Gel de poliacrilamida 10,5% com representantes do grupo ZIII. Comparação entre cepas relacionadas ao grupo ZIII..

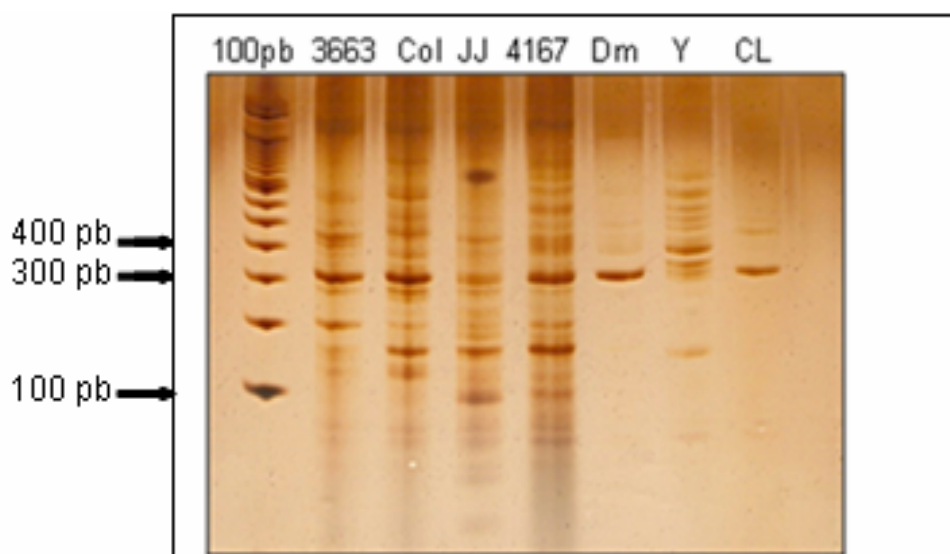


Figura 11 - Gel de poliacrilamida 10,5% contendo todas as cepas estudadas nesse trabalho. Os canais denominados 3663, Col, JJ e 4167 são amostras do grupo Z III e os demais (Dm ou Dm28c, Y e CL) correspondem aos representantes do grupo TcI e TcII respectivamente.

3.2 Clonagem e sequenciamento

Os produtos obtidos através de RT-PCR qualitativo de cada cepa foram utilizados na etapa de clonagem.

Devido à limitação de recursos para sequenciamento e do tempo disponível para a conclusão da dissertação, restringiu-se para até 10 clones sequenciados de cada cepa. Previamente ao seqüenciamento, os clones foram analisados através de gel de agarose (figura 12). Comparado ao controle (Figura 12, canal 2) observa-se que ocorreu inserção completa do fragmento em todos os plasmídios (Figura 12, canais 3 a 21).

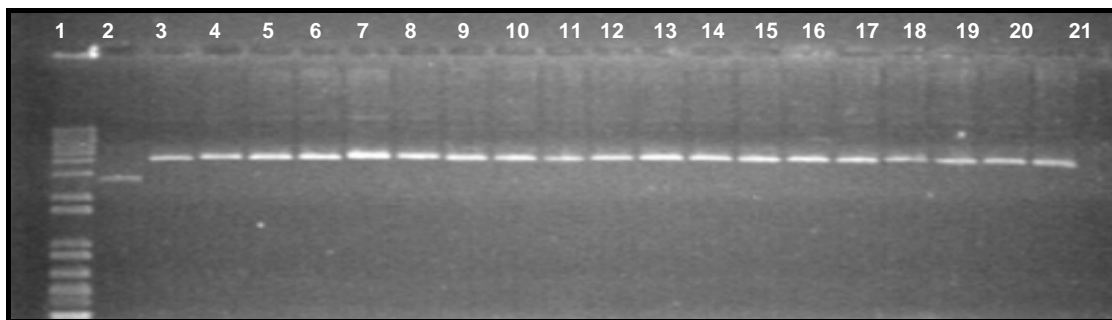


Figura12 – Amostra de um gel de agarose de 1% visando analisar o DNA plasmidial. Canal 1 - Marcador de tamanho molecular de 1kb (invitrogen®), Canal 2 - Plasmídeo extraído de colônia azul (sem inserto). Canais 3 a 21 – Plasmídios contendo o inserto.

Nem todas as seqüências obtidas ofereceram informações totalmente confiáveis, principalmente sobre o códon ATG iniciador (essencial para a precisa delimitação da 5' UTR). Assim, de 170 clones sequenciados, 104 seqüências foram consideradas com um bom perfil de qualidade, havendo 7 seqüências da cepa JJ (das quais 5 continham mini-exon), 19 seqüências da cepa 4167 (das quais 15 apresentavam mini-exon), 29 seqüências da cepa 3663 (das quais 22 apresentavam mini-exon), 15 seqüências da cepa Colombiana (das quais 11 apresentavam mini-exon), 16 seqüências da cepa Dm28c e todas com a presença de mini-exon, 9 seqüências da cepa Y com todas contendo mini-exon, e 11 seqüências da cepa CL-Brener com presença de mini-exon em todas as seqüências. Após edição e eliminação de seqüências redundantes, resultaram 36 com 5' UTR completas, apresentando fragmento do mini-exon e códon iniciador bem definidos (tabela 2). Estes números, a despeito da diversidade das trans-sialidases, permitiram confirmar a observação feita anteriormente de que nas cepas do grupo *T. cruzi* zimodema III, as 5' UTR desta classe de trans-sialidase consistentemente apresentam mais de um fragmento em tamanhos bem distintos, em franco contraste com as cepas das outras populações (TcI e Tc II) (vide tabela 2 item 4.5). Acreditamos que esta peculiaridade das cepas zimodema III são provenientes, sobretudo, do *trans-splicing* adicional. Um maior número de cepas analisadas poderá consolidar esta observação

como mais uma importante característica desse grupo populacional de *T. cruzi*. A tabela 2 resume os tamanhos das 5' UTR das cepas utilizadas.

A média dos tamanhos dos fragmentos das cepas escolhidas considera nucleotídeos a partir do primeiro nucleotídeo ao lado do ATG mais próximo ao óligo TCNA. A cepa Dm28c teve uma média de 214 nucleotídeos, a cepa Y obteve uma média de 178 nucleotídeos, a média da JJ foi de 71 nucleotídeos, a Colombiana revelou média de 80 nucleotídeos, a 3663 obteve média de 238 nucleotídeos, a 4167 indicou 313 nucleotídeos e a CL-Brener indicou uma média de 215 nucleotídeos. De modo geral, as cepas do zimodema III apresentaram tamanho médio de 312 bases, enquanto *T. cruzi* I e *T. cruzi* II apresentaram, respectivamente 209 e 218.

Tabela 1. Tamanhos dos fragmentos (pb) contendo sequências completas de 5' UTR após clonagem e sequenciamento

Cepa	Dm28c	JJ	Y	CL-Brener	3663	4167	Colombiana
Máximo	222	315	299	324	323	324	176
Mínimo	161	-	89	177	244	315	-
Número de clones	7	1	7	7	6	7	1

3.3 RT- PCR quantitativo

Todos os ensaios foram realizados utilizando triplicatas intra e entre experimentos. Foram calculadas as médias dos Cts obtidos de cada cepa analisada entre as triplicatas e com esses valores foram calculados o Δ Ct (tabela 2).

Tabela 2 – Média dos valores de Ct e valores de normalização Δ Ct das cepas do grupo Zimodema III analisadas por PCR quantitativo

Cepas de <i>T. cruzi</i>	Δ Ct*	Média Ct alvo**	Média Ct endógeno**
4167	5,85	31,12	25,26
3663	8,93	34,87	25,94
Colombiana	9,43	35,14	25,70
JJ	3,82	29,08	25,26

*média Ct alvo - média Ct endógeno ** média dos Cts obtidos em três experimentos diferentes

Depois do cálculo do ΔCt , calculamos o $\Delta\Delta Ct$, sempre tendo como calibrador uma cepa referência para cada cepa analisada (calibração) (tabela 2).

Tabela 3. Valores comparativos da expressão gênica relativa da trans-sialidase de *T. cruzi* frente ao uso das diferentes cepas do grupo Zimodema III como calibrador para os experimentos de RT-PCR em tempo real

Calibração	Calibrador 4167	Calibrador 3663	Calibrador Col	Calibrador JJ
Ct Comparativo	3663-4167	0,12	4167-3663	5,08
	Col-4167	0,08	Col-3663	0,70
	JJ-4167	4,08	JJ-3663	34,5
			4167-Col	11,9
			3663-Col	1,41
			JJ-Col	48,8
			4167-JJ	0,24
			3663-JJ	0,02
			Col-JJ	0,02

*Calibração ($\Delta\Delta Ct$)= ΔCt amostra – ΔCt calibrador, **Ct comparativo (quantificação relativa) = $2^{-\Delta\Delta Ct}$

Os valores do cálculo para a quantificação relativa $2^{-\Delta\Delta Ct}$, foram calculados, considerando a cepa Colombiana como referência, pois é a cepa menos expressa.

Dentre os quatro isolados estudados do grupo Zimodema III, a análise pela técnica da RT-PCR quantitativo mostrou que a cepa JJ apresentou quatro vezes maior expressão de transcrição de trans-sialidasas comparada à cepa 4167 que, por sua vez, foi a segunda mais expressa do grupo.

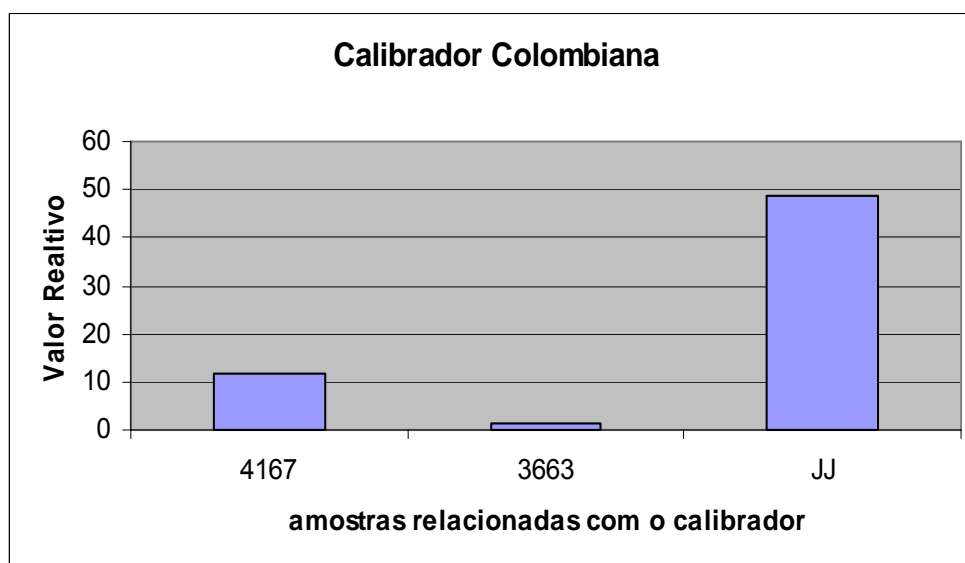


Gráfico 1 – Utilizando a cepa Colombiana como cepa referência para quantificação relativa onde identificamos a cepa JJ com o maior nível de expressão em relação as demais cepas do grupo zimodema III; o mesmo evento ocorre ao utilizar as outras cepas como calibrador.

3.4 Análise das sequências de 5' UTR em trans-sialidases

Após a retirada dos números de acesso redundantes, foi criado um arquivo contendo as informações referentes a 735 sequências genômicas. Durante o processo de edição e comparação dessas sequências, percebemos que poderíamos desenvolver uma abordagem para avaliar os segmentos das regiões não traduzidas pela inspeção de sequências em genes duplicados (Vide artigo anexo). Neste trabalho, embora não seja o objetivo principal da dissertação, aplicamos o uso do alinhamento de sequências idênticas para delimitar os limites das UTRs: o início da 5' UTR e o final da 3' UTR. Até o presente, em ambos os casos é necessário o uso de abordagens inteiramente experimentais para se determinar estes limites. Por exemplo, para a determinação do início da 5' UTR, o sequenciamento de fragmentos obtidos por PCR e transcrição reversa com o iniciador mini-exon permitem determinar precisamente onde ocorreu o trans-splicing. Para a 3' UTR, o sequenciamento de cDNAs iniciados por oligo dT é a opção mais utilizada. O método que apresentamos permite essa localização apenas com alinhamento de sequências, porém é limitada por não ser aplicada a genes de cópia única.

Após identificação no Genbank de 735 sequências anotadas como trans-sialidase no genoma de *T. cruzi*, criou-se um único arquivo alinhado e as sequências idênticas foram separadas. Em seguida, essas sequências foram comparadas com outras oriundas de RNA mensageiro (cDNAs caracterizados como pertencentes a família das trans-sialidases). O número inicial de 735 foi reduzido, então, para 540 sequências anotadas como similar a trans-sialidase. Com cada uma das sequências, foi realizada uma busca por similaridade a fim de se identificar cDNAs originados das trans-sialidases. Este procedimento permitiria definir os prováveis sítios de *trans-splicing* quando se fizesse a comparação do cDNA contendo fragmento de mini-exon com a respectiva sequência genômica. O resultado dessa busca com o BLAST gerou um arquivo de resultados com 28.525 seqüências de alinhamentos de similaridade (seqüências de cDNAs de *T. cruzi* com similaridade as 540 seqüências genômicas). Este elevado número de seqüências similares obtido com essa busca foi causado por: i) as próprias sequências genômicas que são similares entre si; e ii) o elevado número de ESTs de *T. cruzi* que são originados do mesmo gene

transcrito, isto é, o dbEST é banco altamente redundante. Estas sequências similares redundantes foram em seguida eliminadas.

As trans-sialidases são codificadas por genes que estão presentes em grande número de cópias e distribuídos em vários cromossomos no genoma de *T. cruzi* e, além disso, há diferenças na composição da proteína, sobretudo na extremidade carboxila, o que permite a organização das trans-sialidases em quatro classes (Cross, 1994). No entanto, a anotação atual do genoma de *T. cruzi* apenas relaciona um determinado gene como sendo trans-sialidase, sem especificar exatamente a sua classe. Assim, ao se fazer uma busca com as palavras-chaves trans-sialidase cruzada com *T. cruzi* contra o banco de dados do genoma, obteve-se um número expressivo de sequências, porém a sua respectiva classe não foi prontamente identificada.

Como um dos objetivos nesse estudo foi determinar o sítio de *trans-splicing*, a comparação inicial com os cDNAs de trans-sialidase não permitiu se identificar a classe a que pertencia, pois o fragmento selecionado para análise não incluía a sequência codificadora, apenas aquela que provavelmente incluía a 5'UTR, isto é, os nucleotídeos que estavam antes do códon iniciador, anotados pelo Projeto Genoma de *T. cruzi* como sendo o início da trans-sialidase. Supondo que não tenham ocorrido erros de anotação do genoma (por exemplo, códon iniciador identificado como verdadeiro, mas não ser de fato o início da proteína), então a 5' UTR estaria contido na região entre 300-400 pares de bases localizada antes do provável início da CDS. A comparação da sequência desse segmento com um cDNA de trans-sialidase que traz o fragmento do RNA *spliced leader*, permite localizar exatamente onde começa a 5' UTR. Na comparação direta com a sequência genômica é possível determinar duas informações ainda não validadas experimentalmente: o sítio de *trans-splicing* (dinucleotídeo AG que sinaliza a posição no RNA mensageiro na qual o mini-exon será acoplado) e o sinal de *trans-splicing* (região rica em pirimidina que antecede o sítio de *trans-splicing*). Portanto, para as mais de 500 sequências genericamente anotadas como trans-sialidase, identificou-se o sítio de *trans-splicing* e o provável sinal de *trans-splicing* através de alinhamento com respectivo cDNA (depositados no Genbank). Estas informações serão oportunamente comunicadas aos curadores do genoma de *T. cruzi* para atualização das informações sobre *splicing* das trans-sialidases. Este mapeamento tanto dos sítios como dos respectivos prováveis sinais de *trans-splicing* permitiu traçar a

estrutura da 5' UTR em trans-sialidase mostrada graficamente na figura 13 e no logo da figura 14. Blocos de composição relativamente conservada entre as cópias genômicas puderam ser identificados. Uma melhor visualização destes blocos composicionais na sequência da 5' UTR de trans-sialidasas é proporcionado pelo Weblogo (Crooks e cols., 2004). Este é um tipo de representação gráfica de sequências mais frequentes baseado no conceito de "teoria da informação molecular ou bits de informação" (Schneider e cols., 1990). Esta representação difere da tradicional sequência consenso, a qual é obtida pelo alinhamento das sequências e representação somente das bases que aparecem com mais frequência. Ao contrário da sequência consenso, o Weblogo não elimina as bases menos frequentes, ele as representa no gráfico de sequências em uma escala proporcional ao nível de informação que ela representa (bits).

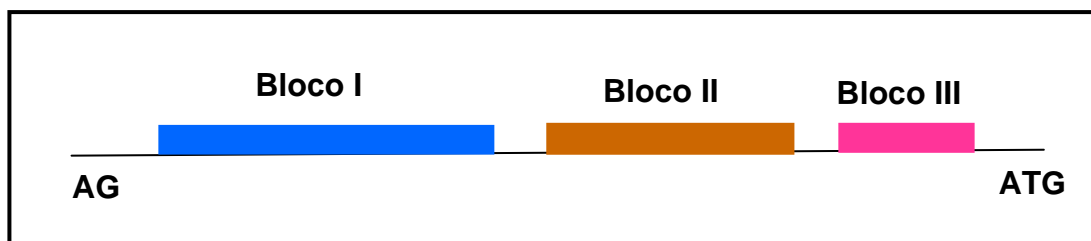


Figura 14 – Esquema da possível estrutura geral da 5' UTR das sequências genericamente anotadas como trans-sialidases pelo Projeto Genoma de *T. cruzi*. AG e ATG indicam, respectivamente, o sítio de *trans-splicing* e o códon iniciador da trans-sialidase. A linha nas extremidades e entre os blocos representa os segmentos de composição muito variável, os quais não apresentam elementos conservados entre as cópias de trans-sialidases. A representação gráfica não está em escala proporcional.

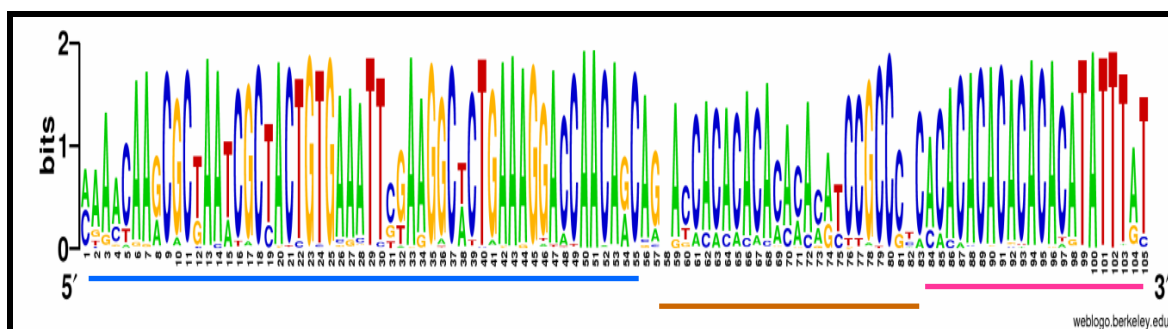


Figura 15 – Logo da provável 5' UTR das trans-sialidases anotadas no genoma de *T. cruzi*. Esta representação não inclui as seqüências da classe I (que inclui o gene TCNA). O logo inclui os três blocos composicionais identificados na 5' UTR das seqüências genômicas analisadas.

Esta representação mostra que a 5' UTR de trans-sialidase, como era esperado para um gene de cópias múltiplas, varia amplamente com o número de cópias, embora os genes que aparentemente estão em *tandem* no genoma não apresentem tanta divergência na seqüência. De modo geral, a variação na 5' UTR reflete o polimorfismo observado até agora nas trans-sialidases. Se levarmos em consideração somente polimorfismos em determinada parte desta região não traduzida (por exemplo, no segmento que compreende as 10 bases antes do códon iniciador, ATG) podemos reconhecer mais de 10 classes de seqüências. Supondo que todas estas cópias são transcritas em algum momento do ciclo de vida do parasita, então será possível detectar diversas variantes ao nível do RNA mensageiro para uma mesma classe de trans-sialidase (neste caso, levando em conta que fase aberta de leitura (ORF) não tenha sido polimórfica). Para estas seqüências genômicas analisadas neste trabalho, o tamanho da 5' UTR de trans-sialidase varia de 90 a 120 bases. Não foi considerado aqui os casos de cópias que apresentam sítios adicionais de *trans splicing*, resultando em 5' UTR com tamanhos maiores, geralmente entre 200 e 300 bases. Estas seqüências com *trans-splicing*

adicional podem ser o resultado de duplicações gênicas (o mais provável para esta família) ou, então, mecanismo regulatório utilizado por *T. cruzi* para enfrentar os diversos choques físico-químicos durante o seu ciclo de vida. Um trabalho experimental mais sofisticado será necessário para determinar qual a função que este tipo de transcrição desempenha no parasita.

A estrutura da 5' UTR esquematizada na figura 15 mostra que existem três blocos principais ao longo da UTR que apresenta composição nucleotídica bem característica: a) um bloco com pouca variação (talvez seja o bloco funcionalmente ativo da 5'UTR para todas as classes); b) um bloco bastante variável, composto por repetições simples de CA. Este bloco pode ter apenas um monômero (CA) ou repetições desse monômero, podendo chegar até 10; c) um bloco antecedendo o códon iniciador e composto mais ou menos 10 bases. Este bloco apresenta, nas 7 bases imediatamente antes do códon iniciador, mutações que poderiam ser utilizadas para classificar as 5' UTRs destas cópias das trans-sialidases em 5 grupos, de acordo com a presença de A, G, T, ou C em uma ou duas posições finais:

- i) **TATTTAT**
- ii) **TATTTGT**
- iii) **TATTTAC**
- iv) **TATTTTT**
- v) **TATTTCT**

Existem outras substituições dentro deste bloco, o que aumentaria estes grupos. No entanto, estas substituições descritas aqui são as que aparecem em maior frequência. Obviamente, para delimitar melhor estes grupos, deve ser feita uma comparação destes grupos baseados na 5' UTR com aqueles derivados das sequências codificantes (4 grupos descritos, Cross e cols., 1994).

A partir do alinhamento das 540 sequências anotadas genericamente como trans-sialidases, os blocos composicionalmente conservados de sequência foram submetidos ao Weblogo. O logo da figura 15 representa o primeiro bloco de 53 bases logo depois do sítio de *trans splicing*. Neste bloco se sobressaem pequenos tratos de purinas (AAGG, GAA, AAAGAA), os quais são interrompidos por algumas

pirimidinas (CT, CTCT, TTC). Em algumas cópias de trans-sialidases, este arranjo de purinas contendo o dinucleotídeo AG precedido por C ou T pode dar origem a um sítio adicional de trans splicing, gerando 5'UTR menores.

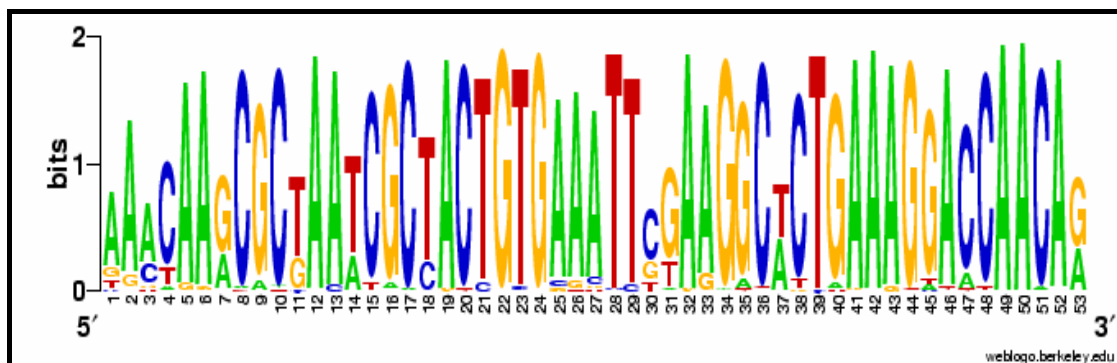


Figura 16 - Representação do primeiro bloco de 53 bases logo depois do sítio de *trans splicing*

O logo da figura 17 representa o segmento mais variável da 5'UTR das trans-sialidases. Este bloco, com conteúdo informativo em aproximadamente 40 bases, é composto por repetições simples do dinucleotídeo CA, variando de 4 a 16 repetições.

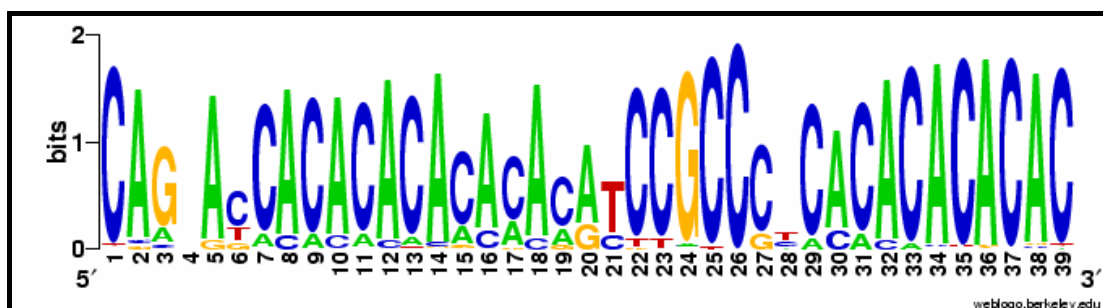


Figura 17 – Representação do segundo bloco (segmento mais variável da 5' UTR) logo após o sítio de *trans splicing*

Uma parte do alinhamento desse bloco nas trans-sialidase é mostrado na figura 18, para se ter uma idéia de quão variável é este bloco. Aproximadamente no meio deste bloco existe um pentanucleotídeo formado por CCGCC. Em algumas cópias, este penta nucleotídeo aparece como TTGCC, o que está mostrado no logo pelo seu baixo valor em bits. Este segundo bloco é muito importante na composição da 5' UTR de trans-sialidase, pois quase toda a variação observada, seja em tamanho ou composição, decorre das mutações nele ocorridas. Em geral, as

Ocorrem também mutações no pentanucleotídeo em algumas classes de trans-sialidases. Uma comparação interessante do ponto de vista evolutivo para as trans-sialidases poderia ser feita com as sequências oriundas de *Trypanosoma rangeli* e tripanosomas africanos (*T. brucei*, *T. congolense*), os quais também possuem genes similares. Se esses blocos apresentarem função relevante no metabolismo do RNA de trans-sialidases, é possível que eles sejam conservados, seja no viés composicional das sequências ou na estrutura secundária da 5'UTR.

3.5 O tamanho da 5' UTR de trans-sialidase classe I

A primeira trans-sialidase descrita em *T. cruzi* foi originalmente denominada *T. cruzi* Neuraminidase, por sua similaridade com as neuraminidases até então descritas, sobretudo de bactérias (Pereira e cols., 1991). A sequência nucleotídica depositada no Genbank (número de acesso M61732) e designação gene TCNA não apresenta a anotação das regiões não traduzidas (5' e 3' UTR). A sequência foi apresentada como sendo diretamente obtida de clone genômico, e não cDNA ou mRNA. As descrições de trans-sialidases que se seguiram a este depósito, apresentavam em alguns casos a descrição da 5' UTR, principalmente quando a sequência era obtida diretamente de cDNA. Esta classe de trans-sialidase, que difere das sequências genômicas apresentadas acima, apresenta uma sequência codificante (ORF) de aproximadamente 3.500 bases, sendo uma das maiores entre as classes de trans-sialidases. Até o presente, a descrição das 5' UTRs deste gene permanece inédita, e é por este motivo que escolhemos desenvolver este trabalho de análise experimental das 5'UTRs de trans-sialidase em *T. cruzi* com foco neste gene específico, sendo desenhado para este fim um oligonucleotídeo para amplificar a UTR desta classe em cepas representando os três grupos populacionais de *T. cruzi* definidos pelo marcador molecular derivado do gene de mini-exon.

Realizamos inicialmente uma determinação do tamanho da 5'UTR e em seguida comparamos estas sequências com aquela descrita por Pereira e cols., apenas o segmento correspondente a 5' UTR. O iniciador utilizado nos experimentos de RT-PCR amplificou um segmento compreendido entre as 30 primeiras bases da ORF (30 bases depois do códon ATG, amplificando, portanto a extremidade amina) e o início da 5' UTR, delimitado pelo oligonucleotídeo derivado do *spliced-leader* RNA (mini-exon).

Considerando apenas os tamanhos das 5'UTRs obtidas aleatoriamente após o RT-PCR, observa-se que o *trans-splicing* adicional ou duplicações gênicas com mutações no sítio primário de *trans-splicing* não parece ser um fenômeno exclusivo de algum grupo populacional. Praticamente, todos os grupos populacionais de *T. cruzi* apresentaram variações no tamanho da 5' UTR para esta classe de trans-sialidase. No entanto, como fora observado na eletroforese após o RT-PCR, tanto em gel de agarose como em gel de poliacrilamida, as cepas do zimodema III (JJ, 4167, 3663) apresentam uma intensidade maior de fragmentos com tamanho em torno de 300 bases. Isto foi confirmado pelo seqüenciamento aleatório dos fragmentos de RT-PCR clonados. Um caso interessante é o clone *T. cruzi* CL-Brener, que isoladamente apresenta tamanho médio da 5'UTR de 252 bases, mas que ao contrário da outra cepa do grupo II utilizada (a cepa Y), apresentou fragmentos em torno de 300 bases. O clone CL-Brener tem sido apontado como um híbrido das populações de *T. cruzi*, e dependendo dos marcadores utilizados, pode apresentar traços do grupo I ou traços do zimodema III. Originalmente, *T. cruzi* CL-Brener foi descrito como pertencente exclusivamente a linhagem filogenética *T. cruzi* II, sendo por isto selecionado para o projeto genoma de *T. cruzi*. O marcador utilizado para caracterizar as cepas neste trabalho, o PCR multiplex baseado no gene de mini-exon, não permite, no entanto, detectar o hibridismo das cepas. Em todo caso, é interessante que também pela transcrição das trans-sialidasas, este comportamento "híbrido" do clone Brener da cepa CL de *T. cruzi* seja evidenciado. Estamos assumindo neste caso que a 5' UTR possa refletir este fenômeno.

Uma varredura no dbEST- GenBank para trans-sialidasas em *Trypanosoma cruzi* mostrou que há 203 cDNAs relacionados com a trans-sialidase. Dessas 203 cópias no dbEST 65 possuem mini-exon, o que nos possibilitou realizar uma comparação com as seqüências genômicas para determinarmos os sítios de *trans splicing*.

4 DISCUSSÃO

Este projeto teve início com uma varredura no GenBank-NCBI para a seleção de sequências genômicas associadas aos segmentos codificantes identificados como trans-sialidase no genoma de *T. cruzi*. O alinhamento das 735 sequências anotadas como trans-sialidase confirmou a composição muito variável do segmento não codificante desses genes, fato que nos trouxe dificuldades de análise. Cerca de 400 sequências puderam ser alinhadas e agrupadas de acordo com a anotação, porém, várias se apresentaram de forma completamente distinta, não havendo nenhuma possibilidade de alinhamento. Talvez isto seja uma consequência de montagem inadequada dos "contigs" durante o sequenciamento do genoma de *T. cruzi*, pois além da dificuldade normal de se reconstruir genomas com alto conteúdo de repetições, as trans-sialidasas são sequências com blocos de repetição. Há, então, a possibilidade de se gerar sequências quiméricas, com a reunião de cópias originalmente distintas ou localizadas em cromossomas diferentes como sendo partes da mesma sequência. Outro fato que pode contribuir para a ambiguidade inicial no alinhamento geral das trans-sialidasas é anotação de genes similares ou correlatos como sendo trans-sialidasas autênticas. É sabido que certos motivos presentes nas trans-sialidasas também aparecem em outras proteínas de superfície, como as mucinas e outras glicoproteínas. Assim, o programa de anotação poderia tomar este motivo como identificador de uma trans-sialidase, e denominado todo um conjunto de sequências apenas similares como sendo idênticas a trans-sialidasas. Por fim, a anotação torna-se mais complexa devido a evidência de que, depois do início do projeto genoma, a cepa escolhida, o clone Brener da cepa CL de *T. cruzi*, aparenta ser um híbrido. Assim, muitas sequências gênicas trarão polimorfismos de base única que ora são típicas de *T. cruzi* grupo I, ora típicas de *T. cruzi* II ou mesmo *T. cruzi* zimodema 3. Isto sem dúvida foi um complicador para a anotação do genoma, mesmo durante a reconstrução da sequência cromossômica original a partir dos "contigs" de sequenciamento.

Na varredura do dbEST utilizando o BLAST, além de se identificar quais cepas já possuem sequências relacionadas a trans-sialidase depositadas neste banco de dados, também obteve-se informação do estágio evolutivo do qual foi extraído o RNA mensageiro. Por exemplo, para o clone CL-Brener foram depositados cDNAs originados dos estágios epimastigota, tripomastigota e

amastigota, enquanto que para a cepa Tulahuen, encontra-se somente o estágio de epimastigota. Este tipo de informação, apesar de publicamente disponível, vem sendo negligenciada no seu potencial para avançar novas estratégias experimentais e analíticas no estudo da transcrição dos tripanossomatídeos, como foi discutido por Brandão (2008). Implícito nos objetivos dessa dissertação, está uma avaliação do aproveitamento dessas informações no dbEST no que tange as trans-sialidases. Pode-se facilmente avaliar quais sequências presentes nas UTRs são mais frequentes nos estágios amastigota, epimastigota ou tripomastigota. Para isto, no entanto, é preciso definir a estrutura geral de uma 5' UTR em trans-sialidase. Esta informação, todavia, não tem recebido a mesma atenção que a respectiva sequência codificante. As ORFs das trans-sialidases foram e continuam sendo muito estudadas em *T. cruzi*, seja pela importância que estas proteínas têm no processo de invasão celular, seja pelo potencial de diagnóstico, imunológico e vacinal que apresentam. Porém, com o crescente aumento das evidências experimentais apontando para um papel relevante de segmentos não traduzidos (5' UTR e 3' UTR) nos processos pós-transcricionais nos tripanossomatídeos, é mister que mais informações sobre estes segmentos sejam gerados. Neste sentido, os bancos de sequências transcritas como o dbEST desempenham uma função importante por facilitar o trabalho de análise preliminar deste tipo de sequência.

As regiões não traduzidas já eram estudadas nos anos 1990, no entanto, a análise de fenômenos relacionados aos sítios (dinucleotídeo AG) e sinais (região rica em polipirimidina) de *trans-splicing* adicional, ainda não havia sido muito explorada. A possibilidade de mapear as regiões não traduzidas, especificamente a 5' UTR, identificando o início e o tamanho dessa região utilizando apenas ferramentas de bioinformática sem haver a necessidade de uma metodologia experimental, nos levou a realizar um estudo paralelo ao foco principal da Tese (ver artigo em anexo). Neste trabalho puramente analítico, testamos uma abordagem para se detectar sequências de ambas as UTRs apenas pela comparação de um determinado gene que se apresente no genoma em uma ou mais cópias. Inicialmente, supomos que independente da organização das cópias, seria possível identificar as UTRs pelo alinhamento. No entanto, quando os genes estão organizados em *tandem* (isto é, a uma cópia do gene segue-se outra imediatamente) e se este gene apresenta indícios de transcrição policistrônica (o caso dos tripanossomatídeos) não será possível avaliar com precisão os limites da UTRs. No

entanto, para outras formas de organização gênica, este método parece ser adequado, como o caso dos genes duplicados e localizados em partes distintas do genoma. Uma vez que não há possibilidade de transcrição policistrônica, os limites das UTRs podem ser exatamente definidos. No caso de genes em *tandem* que não apresentem espaçador muito pequeno, talvez este método possa funcionar.

Ao iniciarmos a parte experimental do nosso trabalho, realizamos a extração do RNA total de três cepas: Dm28c (representante do grupo TcI), Y (representante do grupo TcII) e CL-Brener.

Em 1978 ao analisar o perfil isoenzimático de amostras de pacientes chagásicos da região Norte do país, Milles e colaboradores encontraram um terceiro grupo de parasitos denominados ZIII. O termo zimodema tem por definição uma população de parasito com uma única combinação de perfis de enzimas (definição da Organização Mundial de Saúde em 1978).

Algumas peculiaridades a respeito do grupo Zimodema III foram observadas durante o presente trabalho. Ao se analisar o perfil eletroforético em géis de agarose das amostras das cepas inicialmente estudadas dos representantes dos três grupos populacionais de *T. cruzi*, notamos a presença constante de duas bandas, uma com o tamanho em torno de 300 pares de base, e outra banda em torno de 150 pares de base. Essa observação levou a estender a análise sobre o grupo zimodema III a fim de averiguar se as duas bandas permaneciam nas outras cepas.

A princípio, a única representante ZIII era a cepa INPA 4167; posteriormente, as cepas JJ, 3663 e Colombiana foram integradas. As primeiras cepas são consideradas zimodema III através do marcador molecular derivado do gene de mini-exon (Fernandes e cols., 1998). A cepa Colombiana apresenta um genótipo que depende do marcador utilizado. Pelo gene de mini-exon, é tipada como *T. cruzi* I, porém com outro marcador (3' UTR calmodulina) apresenta mutações típicas de isolados Zimodema III. É possível que com essa cepa ocorra um fenômeno típico dos organismos que são cultivados em vários laboratórios simultaneamente: a identificação equivocada ou mistura com outras cepas. No nosso laboratório, amostras desta cepa foram recebidas em duas ocasiões distintas. Na primeira vez, a tipagem foi realizada com um DNA previamente preparado para eletroforese em gel de campo pulsado, não sendo possível analisar outros parâmetros, como por exemplo, a transcrição gênica. Na segunda vez, a amostra veio para cultivo axênico para a obtenção de formas epimastigotas, e neste caso ela tem sido utilizada para

outros experimentos envolvendo transcrição e análise proteômica. Ao ser tipada com os marcadores citados (mini-exon e 3' UTR calmodulina) apresentou perfil divergente: TcI pelo primeiro e zimodema III pelo segundo. Não há relatos na literatura de que esta cepa/isolado seja um híbrido, no entanto as variações encontradas no seu genótipo de acordo com o marcador utilizado revelam uma grande possibilidade de que seja considerado como tal. Será necessário uma combinação de marcadores moleculares diversos e outros parâmetros biológicos para determinar a real natureza desta cepa/isolado.

De modo geral, no que se refere ao ZIII, ainda não existe um consenso sobre a sua origem filogenética. Trabalhos recentes utilizando técnicas moleculares sugeriram que o ZIII está mais próximo filogeneticamente de *T. cruzi* I do que de *T. cruzi* II. Desta forma, *T. cruzi* I e ZIII parecem estar compartilhando rotas funcionais comuns (Brandão e Fernandes, 2006).

Recentemente, Brandão e Fernandes (2006) procurando uma nova evidência da divergência entre *T. cruzi* estudaram a região 3' não traduzida (3'UTR) do gene da calmodulina. Esse gene modula o metabolismo de cálcio em várias atividades celulares (Hoelflich e Ikura, 2002). Foram observadas mutações localizadas no segmento de 330pb da 3'UTR com três eventos de inserção / deleção importantes: a presença do dinucleotídeo AG em isolados de ZIII, ausência de 13 bases purinas em ZIII e cinco grupamentos de nucleotídeos GT em *T. cruzi* II. Essas mutações sugerem que *T. cruzi* I e o ZIII compartilham vias funcionais comuns se comparados com os isolados de *T. cruzi* II (Brandão e Fernandes, 2006). Porém a filogenia do ZIII ainda continua controversa.

Detectamos eventos mutacionais ao analisarmos o alinhamento das sequências do grupo Zimodema III após sequenciarmos os clones pertencentes a esse grupo, como por exemplo, a existência de um bloco conservado formado pelos nucleotídeos AAC, se repetindo cinco vezes e a ocorrência de uma substituição gerando um sítio de *splicing* (AG) num trecho e uma deleção em outro.

Para algumas cópias de trans-sialidase, houve uma diferença de expressão entre as moléculas de RNA, ocorrendo mais *trans-splicing* habitual do que *trans-splicing* adicional em *T. cruzi*. É possível que esse fato tenha ocorrido devido ao critério randômico adotado. Um exemplo desse fato é uma classe de Dm28c sequenciada na qual observamos sete (7) *trans-splicing* alternativo (habitual) e apenas um (1) *trans-splicing* adicional.

Briones (1995) clonou várias cepas de *T. cruzi* e mostrou que existe pelo menos duas classes gênicas de trans-sialidase com alto grau de similaridade relacionadas com outras proteínas de superfície do parasito (Schenkman e cols., 1994). A classe que representa a maioria dos genes de TS codifica proteínas com um domínio catalítico na extremidade amino-terminal, e um número variável de repetições contendo 12 aminoácidos chamados de CTR (repetições na extremidade da carboxila). Essa classe é altamente expressa em tripomastigotas. A TS purificada dos tripomastigotas se comportam como proteínas oligoméricas formadas por dois ou três polipeptídeos heterogêneos com tamanho molecular entre 120-240 kDa (Einchinger e cols., 1994), todas mostrando o mesmo amino terminal (Pollevick e cols., 1993). A segunda classe de genes codifica proteínas contendo apenas o domínio catalítico da TS. Essa classe é encontrada apenas nos estágios epimastigotas. A TS purificada desse estágio é um polipeptídeo monomérico de 90 kDa e não contém CTR (Chaves e cols., 1993).

O iniciador utilizado no presente trabalho (derivado das bases no início da região codificante do gene TCNA) não abrange as regiões CTRs devido a sua localização, uma vez que este iniciador se encontra (se hibridiza) mais próximo a extremidade amino e a CTR localiza-se na extremidade da carboxila. Possivelmente, a 5' UTR das sequências de TS que contenham CTR sejam diferentes das sequências das TS que não apresentam CTR. Não avaliamos as diferenças das 5' UTRs relativamente a região codificantes das trans-sialidases. Todavia, considerando que as UTRs em geral tendem a refletir os eventos mutacionais que ocorrem no segmento codificante (mas não sob as mesmas taxas de mutação), é de se esperar que estes segmentos sejam igualmente agrupados de acordo com as substituições que ocorrem na ORF e que sobretudo tenham alterado a atividade da proteína codificada. Como já foi demonstrado para o gene de calmodulina (Brandão e Fernandes, 2006), mutações nas UTRs podem refletir a estrutura populacional apresentada pelo organismo em questão, mesmo que o gene seja muito conservado evolutivamente. Dai resulta a importância que se reveste o estudo desses segmentos, seja sob o viés composicional das sequências, ou pelo padrão de transcrição do respectivo gene. Procuramos demonstrar aqui, por meio de uma metodologia de fácil replicação, embora relativamente custosa, que estes genes de cópia múltipla e com grande variação composicional podem servir de parâmetro para analisar ao nível transcricional as populações de *T. cruzi*. Para melhor avaliar o

impacto desta observação experimental na estrutura populacional de *T. cruzi* e mesmo no comportamento de algumas cepas/isolados é importante que tenhamos um mapa o mais acurado possível sobre que partes desses genes estão sendo transcritos nas cepas e isolados.

O nosso objetivo primordial nesta análise das sequências anotadas como trans-sialidase pelo genoma de *T. cruzi* era determinar com precisão o sítio de trans-splicing para cada cópia identificada. No entanto, dada as dificuldades de se lidar com sequências repetitivas e com alto grau de variabilidade, o alinhamento inicial não produziu o resultado esperado. Os grupos obtidos pelo alinhamento no Clustal W não puderam ser correlacionados com a sequência inicialmente escolhida para a etapa experimental: o gene denominado TCNA, descrito por Pereira e colaboradores em 1991 como *T. cruzi* neuraminidase. Este termo foi abandonado posteriormente em prol da designação funcional mais geral, trans-sialidase. Logo depois, outras trans-sialidasas começaram a ser identificadas, sobretudo por varreduras em bibliotecas de expressão com soro de pacientes chagásicos. Isto determinou que várias proteínas de superfície (GP82, GP85, GP90, etc.) fossem identificadas e nomeadas como trans-sialidase similares. Isto tudo resultou na consolidação desses genes em uma família gênica, com ampla distribuição no genoma de *T. cruzi*. Curiosamente, a ênfase era somente na porção codificante, restando os fragmentos laterais (antes do códon iniciador (ATG) e depois do códon terminador (TAG, TAA ou TGA) apenas como componentes menores do gene (e do respectivo RNA mensageiro). Este fato configurou uma escassez de informações sobre estes segmentos não codificantes, porém funcionais, situação que perdura mesmo após a conclusão do projeto genoma de *T. cruzi*. Em outras palavras, podemos dizer que há uma janela de oportunidade para o estudo destes segmentos, pois todo o esforço para caracterizar a atividade funcional/enzimática das trans-sialidasas não levou em consideração que a transcrição, tradução e destino final de muitas proteínas dependem em alguma medida de sinais/motivos presentes na 5' UTR ou 3' UTR. Com as trans-sialidasas não poderiam ser diferente, e recentemente observações experimentais exclusivas das 3' UTR destes genes foram publicadas enfatizando a importância funcional de certos motivos nestes segmentos (Jager, 2008). Nosso esforço neste projeto foi inicialmente concentrado na 5' UTR sob a perspectiva populacional de *T. cruzi*, e dadas as nossas limitações estruturais e materiais não foi possível avançar além do simples mapeamento e identificação de eventos

alternativos na transcrição das cepas/isolados do zimodema III. Acreditamos, porém, que as informações que conseguimos gerar sobre a estrutura provável de uma 5' UTR das sequências genômicas anotadas como trans-sialidase e a posição de seus blocos composicionais conservados poderão auxiliar o desenvolvimento de novas estratégias para o estudo dos mecanismos de transcrição de genes de cópia múltipla utilizadas por este protozoário.

5 CONCLUSÃO

- Os fragmentos amplificados correspondentes à 5'UTR de trans-sialidases de cepas do grupo zimodema III, revelaram a presença de dois fragmentos, pela visualização em gel de agarose, de aproximadamente 300 e 400 pb, sugerindo a ocorrência de sítios adicionais de *trans-splicing* e/ou mutações no sítio primário de *trans-splicing*; o que foi posteriormente confirmado por seqüenciamento;
- Entre as 4 cepas estudadas de zimodema III por RT-PCR em Tempo Real, a cepa JJ foi o que apresentou maior nível de transcrição de trans-sialidase;
- Os fragmentos amplificados correspondentes à 5'UTR de trans-sialidases de cepas do grupo zimodema III, revelaram a presença de dois fragmentos, pela visualização em gel de agarose, de aproximadamente 300 e 400 pb, sugerindo a ocorrência de sítios adicionais de *trans-splicing* e/ou mutações no sítio primário de *trans-splicing*; o que foi posteriormente confirmado por seqüenciamento;
- Entre as 4 cepas estudadas de zimodema III por RT-PCR em Tempo Real, a cepa JJ foi o que apresentou maior nível de transcrição de trans-sialidase;
- Os fragmentos amplificados correspondentes à 5'UTR de trans-sialidases de cepas do grupo zimodema III, revelaram a presença de dois fragmentos, pela visualização em gel de agarose, de aproximadamente 300 e 400 pb, sugerindo a ocorrência de sítios adicionais de *trans-splicing* e/ou mutações no sítio primário de *trans-splicing*; o que foi posteriormente confirmado por seqüenciamento;
- Entre as 4 cepas estudadas de zimodema III por RT-PCR em Tempo Real, a cepa JJ foi o que apresentou maior nível de transcrição de trans-sialidase.

6 PERSPECTIVAS

- Avaliar o potencial da 5' UTR da TS em aplicações de tipagem e diagnóstico;
- Finalizar os experimentos de transcrição por PCR em Tempo Real com as cepas Tc I e Tc II;

- Avaliar a variação da estrutura da 5' UTR proposta (3 blocos composicionais) nos estágios amastigotas e tripomastigotas nos grupos populacionais de *T. cruzi*.

REFERÊNCIAS

- Andrews**, N. W; Einstein, M; Nussenzweig, V. 1989. Presence of antibodies to the major surface glycoprotein of *Trypanosoma cruzi* amastigotes in sera from Chagasic patients. *Am J Trop Med Hyg.* 40: 46 – 9
- Andrews**, N. W; Abrams, C. K; Slatin, S. L; Griffiths, G. 1990. A *T. cruzi* secreted protein immunologically related to the complement component C9: evidence for membrane pore-forming activity at low pH. *Cell.* 61: 1277-87
- Anomimo**. 1999. Recommendations from a Sattelite Meeting. *Mem. Inst. Oswaldo Cruz Andrade*, S.G. 1979. Biological Characterization of Strains of *Trypanosoma cruzi*. Congresso Internacional de Doença de Chagas, Rio de Janeiro, Abstract n 9 – 13
- Araújo**, C.A.C; Mello, C.B; Jansen, A.M. 2002. *Trypanosoma cruzi* I and *Trypanosoma cruzi* II: Recognition of sugar structures by *Arachis hypogea* (Peanut Agglutinin) Lectin. *J. Parasitol.* 88: 582–586.
- Araya**, J.E., Cano, M.I., Yoshida, N.; da Silveira, J.F. 1994. Cloning and characterization of a gene for the stage-specific 82-kDa surface antigen of metacyclic trypomastigotes of *Trypanosoma cruzi*. *Mol Biochem Parasitol* 65: 161–169.
- Balakirev**, E. S; Ayala, F. J .2003. Molecular population genetics of the beta-esterase gene cluster of *Drosophila melanogaster*. *J Genet* vol 82: 115-31
- Barnabé**, C; Brisse, S; Tibayrenc, M. 2000. Population structure and genetic typing of *Trypanosoma cruzi*, the agent of Chagas disease: a multilocus enzyme electrophoresis approach. *Parasitology.* Vol.120. p. 513-52
- Barret**, M.P; Burchmore, R.J; Stich, A; Lazzari, J.O; Frasch, A.C.; Cazzulo, J.J; Krishna, S. The trypanosomiases. *Lancet*, v.362, n.9394, p.1469-80, 2003
- Bashirullah**, A; Cooperstock, R. L; Lipshitz, H. D. 2001. Spatial and temporal control of RNA stability. *Proc Natl Acad Sci U S A.* 98: 7025-8
- Bogliolo**, A. R; Chiari, E; Silva-Pereira, R. O; Silva-Pereira, A. A. 1986. A comparative study of *Trypanosoma cruzi* enzyme polymorphism in South America. *Braz J Med Biol Res.*Vol. 19. P. 673-83
- Borst**, P. 1984. Comparison of the genes coding for the common 5' terminal sequence of messenger RNAs in three Trypanosome species. *Nucleic Acids Res.* 12: 4431 4443

- Brandão**, A. 2006. The Untranslated regions of genes from *Trypanosoma cruzi*: perspectives for functional characterization of strains and isolates. Mem Inst Oswaldo Cruz. 101 (7): 775 – 777
- Brandão**, A; Fernandes, O. 2006. *Trypanosoma cruzi*: Mutations in the 3' untranslated region of calmodulina gene are specific for lineages *T. cruzi I*, *T. cruzi II*, and the zymodeme III isolates. Experimental Parasitology, Vol. 112, p. 247 - 252
- Brandão**, A. 2008. Trypanosomatid EST: a neglected information resource regarding flagellated protozoa? Mem Inst Oswaldo Cruz. 103: 622-6
- Breniere**, S. F; Bosseno, M. F; Barnabe, C; Urdaneta-Morales, S; Tibayrenc, M. 1993. Copy number differences in the 195 bp repeated satellite DNA from *Trypanosoma cruzi* and *Trypanosoma rangeli*: potential use for epidemiologic surveys. Mem Inst Oswaldo Cruz. 88: 163-5
- Briones**, M. R; Egima, C. M; Eichinger, D; Schenkman, S. 1995. Trans-sialidase genes expressed in mammalian forms of *Trypanosoma cruzi* evolved from ancestor genes expressed in insect forms of the parasite. J Mol Evol. 41: 120-31
- Burleigh**, B.A; Andrews, N.W. 1995. The mechanisms of *Trypanosoma cruzi* invasion of mammalian cells. Annu Rev Microbiol 49: 175–200
- Burleigh**, B.A.; Caler, E.V.; Webster, P; Andrews, N.W. 1997. A cytosolic serine endopeptidase from *Trypanosoma cruzi* is required for the generation of Ca²⁺ signaling in mammalian cells. J Cell Biol 136: 609–620
- Burleigh**, B. A; Andrews, N. W. 1998. Signaling and host cell invasion by *Trypanosoma cruzi*. Curr Opin Microbiol. Vol. 1. P. 461-5
- Burleigh**, B. A; Woolsey, A. M. 2002. Cell signalling and *Trypanosoma cruzi* invasion. Cellular Microbiology. Vol.11, 701–711
- Buscaglia**, C.A. *et al.* 2006. *Trypanosoma cruzi* surface mucins: host-dependent coat diversity. Nat. Rev. Microbiol. Vol. 4. P. 229-236
- Caler**, E. V; Vaena de Avalos, S; Haynes, P. A; Andrews, N. W; Burleigh, B. A. 1998. Oligopeptidase B-dependent signaling mediates host cell invasion by *Trypanosoma cruzi*. EMBO J. Vol. 17. P. 4975-86
- Caler**, E. V; Morty, R. E; Burleigh, B. A; Andrews, N. W. 2000. Dual role of signaling pathways leading to Ca(2+) and cyclic AMP elevation in host cell invasion by *Trypanosoma cruzi*. Infect Immun. 68: 6602-10
- Campetella**, O; Sanchez, D; Cazzulo, J. J; Frasch, A. C. 1992. A superfamily of *Trypanosoma cruzi* surface antigens. Parasitol Today. Vol. 8. P. 378-81

- Campos**, P.C; Bartholomeu, D. C; DaRocha, W. D; Cerqueira, G.C; Teixeira, M.R. 2008. Sequences involved in mRNA processing in *Trypanosoma cruzi*. International Journal for Parasitology. Vol. 38. p.1383 – 1389
- Castro**, C; Craig, S.P; Castaneda, M. 1981. Genome Organization and Ploidy Number in *Trypanosoma cruzi*. Mol. Biochem. Parasitol. Vol. 4. p. 273-282
- Chagas**, C. 1909. Nova tripanozomiaze humana. Estudos sobre a morfologia e o ciclo evolutivo do *Schizotrypanum cruzi* n. gen, n. sp., agente etiológico de nova entidade morbida do homem. Mem Inst Oswaldo Cruz. Vol.1. p. 11-80
- Colli**, W; Alves, M. J. 1999. Relevant glycoconjugates on the surface of *Trypanosoma cruzi*. Mem Inst Oswaldo Cruz. Vol. 94. P. 37-49
- Consenso** Brasileiro em Doença de Chagas 2005. Rev Soc Bras Med Trop. Vol. 3. p. 7-29: www.parasitologia.org.br/atualidades/consensochagas.pdf, acesso em 21/03/2009
- Cortez**, M; Silva, M. R; Neira, I; Ferreira, D; Sasso, G. R; Luquetti, A. O; Rassi, A; Yoshida, N. 2006. *Trypanosoma cruzi* surface molecule gp90 downregulates invasion of gastric mucosal epithelium in orally infected mice. Microbes Infect. Vol. 8. p. 36-44
- Coura**, J.R; Junqueira, A.C.V; Fernandes, O, Valente, S.A.S; Miles, M.A. 2002. Emerging Chagas disease in Amazonian Brazil. Trends Parasitol. Vol. 1. p. 171-176.
- Coura**, J.R, de Castro, S.L. 2002. A critical review on Chagas disease chemotherapy. Mem Inst Oswaldo Cruz. Vol. 97. p. 3-24
- Crooks**, G.E; Hon, G; Chandonia, J.M; Brenner, S.E. 2004. WebLogo: A sequence logo generator, Genome Research, Vol. 14. p. 1188-1190 - <http://weblogo.berkeley.edu>
- Cross**, G.A.M; Nozaki T. 1994. Functional complementation of glycoprotein 72 in a *Trypanosoma cruzi* glycoprotein 72 null mutant. Mol Biochem Parasitol. Vol. 67. p. 91-102.
- Cuevas**, I. C; Cazzulo, J. J; Sanchez, D. O. 2003. gp63 homologues in *Trypanosoma cruzi*: surface antigens with metalloprotease activity and a possible role in host cell infection. Infect Immun. Vol. 71.p. 5739-49
- De Souza**, W. 1995. Structural organization of the cell surface of pathogenic protozoa. Micron, Vol. 26. p. 405 – 430
- De Souza**, W. 2002. From the cell biology to the development of new chemotherapeutic. Kinetoplastid Biology and Disease. Vol.1. p. 3

- De Lange**, T; Michels, P. A; Veerman, H. J; Cornelissen, A. W; Borst, P. 1984. Many trypanosome messenger RNAs share a common 5' terminal sequence. *Nucleic Acids Res.* Vol. 12. P. 3777-90
- Devera**, R; Fernandes, O; Coura, J. R. 2003. Should *Trypanosoma cruzi* be called "cruzi" complex? a review of the parasite diversity and the potential of selecting population after in vitro culturing and mice infection. *Mem Inst Oswaldo Cruz.* Vol. 98. P. 1-12
- Dias**, J. C. P; Coura, J. R., 1997. Epidemiologia. In: *Clínica e Terapêutica da Doença de Chagas. Uma Abordagem Prática para o Clínico Geral* (J. C. P. Dias & J. R. Coura, org.)p. 33-66, Rio de Janeiro: Editora Fiocruz.
- Dias**, J.C.P; Silveira, A.C; Schofield, C.J. 1999. The Evolution of Chagas Disease (American Trypanosomiasis). Control after 90 Years since Carlos Chagas Discovery. *Mem Inst Oswaldo Cruz, Rio de Janeiro.* Vol. 94. p. 103-121
- Docampo**, R; Scott, D.A, Vercesi, A.E; Moreno, S.N.J. 1996. The role of Ca^{2+} in processo of cell invasion by intracellular parasites. *Parasitol Today.* Vol. 12. P. 61-65
- Dorta**, M.L; Ferreira, A.T; Oshiro, M.E.M; Yoshida, N. 1995. Ca^{2+} signal induced by *Trypanosoma cruzi* metacyclic trypomastigote surface molecules implicated in mammalian cell invasion. *Mol Biochem Parasitol* Vol. 73. p. 285–289
- Douglas**, K. T; Bryce, R. A; Neres, J. 2008. Rational drug design in parasitology: trans-sialidase as a case study for Chagas disease. *Drug discovery today*, Vol. 13
- EL-SAYED**, N.M. *et al.* 2005. The genome sequence of *Trypanosoma cruzi*, etiologic agent of Chagas disease. *Science.* v 15, p. 409-15
- Eichinger**, D; Chaves, L. B; Pontes de Carvalho, L. C; Schenkman, S. 1994. A proteolytic fragment of *Trypanosoma cruzi* trans-sialidase lacking the carboxyl-terminal domain is active, monomeric, and generates antibodies that inhibit enzymatic activity. *J Biol Chem.* Vol. 269. P. 7970-5
- Farrell**, J. P; Kirkpatrick, C. E. 1987. Experimental cutaneous leishmaniasis. II. A possible role for prostaglandins in exacerbation of disease in *Leishmania major*-infected BALB/c mice. *J Immunol.* Vol. 138. P. 902-7
- Favoreto**, S., Jr; Dorta, M. L; Yoshida, N. 1998. *Trypanosoma cruzi* 175-kDa protein tyrosine phosphorylation is associated with host cell invasion. *Exp Parasitol.* Vol. 89. P. 188-94

- Fernandes**, A. J; Diotaiuti, L; Dias, J.C.P; Romanha, A.J; Chiari, E. 1994. Relationships Between *Trypanosoma cruzi* Transmission Cycles in the County of Bambuí, Minas Gerais, Brazil. Vol.10 (4). P. 473-480
- Fernandes**, O; Sturm, N. R; Derre, R; Campbell, D. A. 1998. The mini-exon gene: a genetic marker for zymodeme III of *Trypanosoma cruzi*. Mol Biochem Parasitol. Vol. 95. p. 129-33
- Fernandes**, O; Mangia, R.H; Lisboa, C.V; Pinho, A.P; Morel, B; Zingales, B; Campbell, D.A; Jansen, A.M. 1999. The complexity of the sylvatic cycle of *Trypanosoma cruzi* in Rio de Janeiro state (Brazil) revealed by the non-transcribed spacer of the mini-exon gene. Parasitology. Vol. 118. P. 161-166
- Formas** evolutivas do *Trypanosoma cruzi*: <http://www.cienciaviva.org.br>, <http://www.ufrgs.br/para-site/Imagensatlas/Protozoa/Imagens>, acesso em 29 de janeiro de 2009
- Fotaki**, M. E.Iatrou, K. 1993. Silk moth chorion pseudogenes: hallmarks of genomic evolution by sequence duplication and gene conversion. J Mol Evol. Vol 37, issue 2: 211-20
- Frasch**, A. C; Affranchino, J. L; Ibanez, C. F; Reyes, M. B; Macina, R. A; Luquetti, A.O; Rassi, A; Aslund, L; Pettersson, U. 1988. Analysis of cloned *Trypanosoma cruzi* proteins that are antigenic during the acute and chronic phase of Chagas' disease. Mem Inst Oswaldo Cruz. Vol. 83. P. 345-6
- Frasch**, A. C; Cazzulo, J. J; Aslund, L; Pettersson, U. 1991. Comparison of genes encoding *Trypanosoma cruzi* antigens. Parasitol Today. Vol. 7. P. 148-51
- Frasch**, A.C; Sanchez, D.O; Campetella, O; Trombetta, S; Sousa, M; Henriksson, J; Hellman, U; Pettersson, U; Cazzulo, J.J; Pollevick, G.D. 1993. Members of the SAPA/trans-sialidase protein family have identical N-terminal sequences and a putative signal peptide. [Mol Biochem Parasitol. Vol.59. p.171-4](#)
- Frasch**, A. C; Uttaro, A. D; Parodi, A. J; Campetella, O. E. 1994. A recombinant *Trypanosoma cruzi* trans-sialidase lacking the amino acid repeats retains the enzymatic activity. Mol Biochem Parasitol. Vol. 64. P. 337-40
- Frasch**, A.C.C. 2000. Functional diversity in the trans-sialidase and mucin families in *Trypanosoma cruzi*. Parasitol. Today. Vol. 16. p. 282-286
- Frevert**, U; Schenkman, S; Nussenzweig, V. 1992. Stage-specific expression and intracellular shedding of the cell surface trans-sialidase of *Trypanosoma cruzi*. Infect Immun. Vol. 60 . p. 2349-60

- Fotaki**, M. E; Iatrou, K. 1993. Silk moth chorion pseudogenes: hallmarks of genomic evolution by sequence duplication and gene conversion. *J Mol Evol.* Vol. 37. P. 211-20
- Frevert**, U; Schenkman, S; Nussenzweig, V. 1992. Stage-specific expression and intracellular shedding of the cell surface trans-sialidase of *Trypanosoma cruzi*. *Infect Immun.* Vol. 60. p. 2349-60
- Gaba**, A; Wang, Z; Krishnamoorthy, T; Hinnebusch, A.G; Sachs, M.S. 2001. Physical evidence for distinct mechanisms of translational control by upstream open reading frames. *EMBO J.*Vol.20. p. 6453-6463
- Gibson**, W. C; Miles, M.A. 1986. The karyotype and ploidy of *Trypanosoma cruzi*. *EMBO J.* Vol. 5. p.1299-1305
- Giordano**, R.; Chammas, R.; Veiga, S.S.; Colli, W; Alves, M.J.M. 1994. An acidic component of the heterogeneous Tc-85 protein family from the surface of *Trypanosoma cruzi* is a laminin binding glycoprotein. *Mol Biochem Parasitol.* Vol. 65. P. 85-94
- Habitação** de pau-a-pique. Extraído de <http://www.brasiloste.com.br>, em 15 de dezembro de 2008
- Haile**, S; Papadopoulou, B. 2007. Developmental regulation of gene expression in trypanosomatid parasitic protozoa. *Curr Opin Microbiol.* Vol. 10. 569-77
- Hoeflich**, K. P; Ikura, M. 2002. Calmodulin in action: diversity in target recognition and activation mechanisms. *Cell.* Vol. 108. P. 739-42
- Horiuchi**, T; Aigaki, T. 2006. Alternative *trans*-splicing: a novel mode of pre mRNA processing. *Biol. Cell.* 98: 135–140
- Jäger**, A. V; De Gaudenzi, J.G; Cassola, A; D'Orso, I; Frasch, A.C. 2006 (aceito em 2007). mRNA maturation by two-step trans-splicing/polyadenylation processing in trypanosomes
- Jäger**, A.V; Muiá, R.P; Campetella, O. 2008. Stage-specific expression of *Trypanosoma cruzi* trans-sialidase involves highly conserved 3' untranslated regions. [FEMS Microbiol Lett.](#)Vol. 283. p. 182-8
- Jansen**, R.P. 2001. mRNA localization: message on the move. *Nat Rev Mol Cell Biol.* Vol. 2. p. 247-256
- Kirchhoff**, L. V; Engel, J. C; Dvorak, J. A; Sher, A. 1984. Strains and clones of *Trypanosoma cruzi* differ in their expression of a surface antigen identified by a monoclonal antibody. *Mol Biochem Parasitol.* Vol.11. p. 81-9

- Kirkpatrick**, C. E; Nolan, T. J; Farrell, J. P. 1987. Rate of Leishmania-induced skin-lesion development in rodents depends on the site of inoculation. *Parasitology*. Vol. 94. p. 451-65
- Kozak**, M. 1987. An analysis of 5'-noncoding sequences from 699 vertebrate messenger RNAs. *Nucleic Acid Res*. Vol. 15. P. 8125-8148
- Lanar**, D. F; Levy, L. S; J. E. Manning. 1981. Complexity and content of the DNA in *Trypanosoma cruzi*. *Mol. Biochem Parasitol*. Vol. 3. P. 327-341
- Lenzi**, H. L; Castelo-Branco, M. T; Pelajo-Machado, M; Oliveira, D; N; Gattass, C. R. 1998. *Trypanosoma cruzi*: compromise of reproductive system in acute murine infection. *Acta Trop*. Vol. 71. p. 117-29
- LeBowitz**, J. H; Smith, H. Q; Rusche, L; Beverley, S. M. 1993. Coupling of poly(A) site selection and trans-splicing in *Leishmania*. *Genes Dev*. Vol. 7. P. 996-1007
- Ley**, V; Robbins, E.S; Nessenzweig, V; Andrews, N.W. 1990. The exit of *Trypanosoma cruzi* from the phagosome is inhibited by raising the pH of acidic compartments. *J Exp Med*. Vol. 171. p. 401-413
- Liang**, X.H; Haritan, A; Uliel, S; Michaeli, S. 2003. Trans and cis-splicing in trypanosomatids: mechanism, factors, and regulation. *Eucaryot Cell*. Vol. 2. P. 830-840
- Livak**, K. J; Schmittgen†, T. D. 2001. Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the 2^{-ΔΔCT} Method. *Methods*, Vol. 25. p. 402-408
- Lovett**, P. S; Rogers, E. J. 1996. Ribosome regulation by the nascent peptide. *Microbiol Ver*. 60: 366-85
- Luo**, Z; Sachs, M. S. 1996. Role of an upstream open reading frame in mediating arginine-specific translational control in *Neurospora crassa*. *J Bacteriol*. Vol. 178. p. 2172-7
- Macedo**, A.M. e colaboradores. 2001. Usefulness of micro satellites typing in population genetic studies of *Trypanosoma cruzi*. *Mem. Inst. Oswaldo Cruz*. Vol. 96. p. 407-413
- Macedo**, A. M; Machado, C. R; Oliveira, R. P; Pena, S. D. 2004. *Trypanosoma cruzi*: genetic structure of populations and relevance of genetic variability to the pathogenesis of chagas disease. *Mem Inst Oswaldo Cruz*. Vol. 99. p. 1-12
- Manque**, P. M; Neira, I; Atayde, V. D; Cordero, E; Ferreira, A. T; da Silveira, J. F; Ramirez, M; Yoshida, N. 2003. Cell adhesion and Ca²⁺ signaling activity in stably

transfected *Trypanosoma cruzi* epimastigotes expressing the metacyclic stage-specific surface molecule gp82. *Infect Immun*. Vol. 71. p. 1561-5

Matthews, K. R; Tschudi, C; Ullu, E. 1994. A common pyrimidine-rich motif governs trans-splicing and polyadenylation of tubulin polycistronic pre-mRNA in trypanosomes. *Genes Dev*. Vol. 8. P. 491-501

McCarrey, J. R; Riggs, A. D. 1986. Determinator-inhibitor pairs as a mechanism for threshold setting in development: a possible function for pseudogenes. *Proc Natl Acad Sci U S A*. vol 83, issue 3: 679-83.

Medicamentos:

<http://www. Roche.com>, acessado em 15 de janeiro de 2009

<http://www. Bayer.com>, acessado em 15 de janeiro de 2009

Meijer, H. A; Thomas, A. A. 2002. Control of eukaryotic protein synthesis by upstream open reading frames in the 5'-untranslated region of an mRNA. *Biochem J*. Vol. 367. p. 1 – 11

Meirelles, M. N; Juliano, L; Carmona, E; Silva, S. G; Costa, E. M; Murta, A. C; Scharfstein, J. 1992. Inhibitors of the major cysteinyl proteinase (GP57/51) impair host cell invasion and arrest the intracellular development of *Trypanosoma cruzi* in vitro. *Mol Biochem Parasitol*. Vol. 52. P. 175-84

Mignone, F; Liuni, S; Grillo, G; Licciulli, F;; Gissi, C; Saccone, C. Pesole, G 2002. UTRdb and UTRsite: specialized databases of sequences and functional elements of 5' and 3' untranslated regions of eukaryotic mRNAs. Update. *Nucleic Acids Res*. Vol 30. p. 335-40

Mignone, F; Grillo, G; Licciulli, F; Iacono, M; Liuni, S; Kersey, P. J; Duarte, J; Saccone, C; Pesole, G. 2005. UTRdb and UTRsite: a collection of sequences and regulatory motifs of the untranslated regions of eukaryotic mRNAs. *Nucleic Acids Res*. Vol. 33. P.141-6

Milles, M.A; Toye, P.J; Oswald, S.C; Godfrey, D.G.1977. The Identification by Isoenzyme Patterns of Two Distinct Strain – Groups of *Trypanosoma cruzi*, Circulating Independently in a Rural area of the Royal Society of Tropical Medicine and Hygiene. Vol. 71. p. 217–225

Miles, M.A; Souza, A; Pova, M; Shaw, J.J.; Lainson, R; Toye, P.J. 1978. Isozymic heterogeneity of *Trypanosoma cruzi* in the first autochthonous patients with Chagas' disease in Amazonian Brazil. *Nature*. Vol. 272. P. 819-21

- Milles**, M.A; Lanham, S.M, de Souza, A.A; Povoá, M. 1980. Further enzymic characters of *Trypanosoma cruzi* and their evaluation for strain identification. Transactions of the Royal Society of Tropical Medicine and Hygiene. Vol. 74, 221-237
- Moreno**, S. N. J; Silva, J; Vercesi, A. E; Docampo, R. 1994. J. Exp. Med. Vol.180. p. 1535–1540
- Morris**, D. R; Geballe, A. P. 2000. Upstream open reading frames as regulators of mRNA translation. Mol Cell Biol. Vol. 20. p. 8635-42
- Mortara**, R. A; Andreoli, W. K; Taniwaki, N. N; Fernandes, A. B; Silva, C. V; Fernandes, M. C; L'Abbate, C; Silva, S. 2005. Mammalian cell invasion and intracellular trafficking by *Trypanosoma cruzi* infective forms. An Acad Bras Cienc. Vol. 77. p. 77-94
- Murta**, A.C; Persechini, P.M; Souto, Padron, T; De Souza, W; Guimarães, J.A; Scharfstein, J. 1990. Structural and functional identification of GP57/51 antigen of *Trypanosoma cruzi* as a cysteine proteinase. Mol Biochem Parasitol. Vol. 43. p. 27–38
- Murthy**, V. K; Dibbern, K. M; Campbell, D. A. 1992. PCR amplification of mini-exon genes differentiates *Trypanosoma cruzi* from *Trypanosoma rangeli*. Mol Cell Probes. Vol. 6. P. 237-43
- Neira**, I; Ferreira, A. T; Yoshida, N. 2002. Activation of distinct signal transduction pathways in *Trypanosoma cruzi* isolates with differential capacity to invade host cells. Int J Parasitol. Vol. 32. p. 405-14
- Neira**, I; Silva, F.A; Cortez, M; Yoshida, N. 2003. Involvement of *Trypanosoma cruzi* metacyclic tripomastigote surface molecule gp82 in adhesion to gastric mucin and invasion of epithelial cells. Infect Immun. Vol. 71. P. 557-561
- Neres**, J; Bryce, R. A; Douglas, K. T. 2008. Rational drug design in parasitology: trans-sialidase as a case study for Chagas disease. Drug Discov Today. 13:110-7
- Nussenzweig**, V; Eichinger, D; Pereira, M. E; Schenkman, S. 1994. Structural and functional properties of *Trypanosoma* trans-sialidase. Annu Rev Microbiol. Vol. 48. p. 499-523
- Oliveira**, R.P; Broude, N.E; Macedo, A. M; Cantor, C.R; Smith, C. L; Pena, S. D. J. 1998. Probing the genetic population structure of *Trypanosoma cruzi* with polymorphic microsatellites. Proc Natl Acad Sci USA. Vol. 95. p. 3776-3780

Organização Mundial de Saúde (WHO), 2004. www.who.int, acessado em 20 de Dezembro de 2008

Paiva, C. N; Souto-Padron, T; Costa, D. A; Gattass, C. R. 1998. High expression of a functional cruzipain by a non-infective and non-pathogenic *Trypanosoma cruzi* clone. *Parasitology*. Vol. 117. p. 483-90

Pereira, M. E. A. 1983. Rapid and sensitive assay for neuraminidase using peanut lectin hemagglutination: application to *Vibrio cholera* and *Trypanosoma cruzi*. *J Immunol Methods*. Vol 63: 25-34

Pereira, M. E Prioli, R. P; Mejia, J. S. 1991. On the interaction of *Trypanosoma cruzi* neuraminidase and human lipoproteins. *Eur J Epidemiol*. Vol. 7. p. 344-8

Pinto, A. Y. N ; Valente, S.A ; Valente, V. C; Ferreira Junior, A; Coura, J.R. 2008. Acute phase of Chagas disease in the Brazilian Amazon region. Study of 233 cases from Pará, Amapá and Maranhão observed between 1988 and 2005. *Revista da Sociedade Brasileira de Medicina Tropical*. Vol 41. P. 602-614

Pollevick, G. D; Sanchez, D. O; Campetella, O; Trombetta, S; Sousa, M; Henriksson, J; Hellman, U; Pettersson, U; Cazzulo, J. J; Frasc, A. C. 1993. Members of the SAPA/trans-sialidase protein family have identical N-terminal sequences and a putative signal peptide. *Mol Biochem Parasitol*. Vol. 59. 171-

Ouaisi, M.A; Cornette, J; Afchain, D; Capron, A; Gras-Masse, H; Tartar , A. 1986. *Trypanosoma cruzi* infection inhibited by peptides modeled from a fibronectin cell attachment domain. [Science](http://www.sciencedirect.com). Vol. 234. P. 603-7

Rodriguez, A; Rioult, M.G; Ora, A; Andrews,N. 1995. A trypanosome-soluble factor induces IP3 formation, intracellular Ca²⁺ mobilization and microfilament rearrangement in host cells. *Journal of Cell Biology*. Vol.129. p. 1263-1273

Rodríguez, A., Webster, P., Ortego, J., and Andrews, N.W. 1997. Lysosomes behave as Ca²⁺-regulated exocytic vesicles in fibroblasts and epithelial cells. *J Cell Biol* in press

Rodríguez, A; Martinez, I; Chung, A; Berlot, C.H; Andrews, N.W. 1999. cAMP regulates Ca²⁺-dependent exocytosis of lysosomes and lysosome-mediated cell invasion by trypanosomes. *J Biol Chem*. Vol. 274. p. 16754–16759.

Rosenberg, I; Prioli, R. P; Ortega-Barria, E; Pereira, M. E. 1991. Stage-specific phospholipase C-mediated release of *Trypanosoma cruzi* neuraminidase. *Mol Biochem Parasitol*. Vol. 46. p. 303 – 5

- Rubin-de-Celis**, S. S; Uemura, H; Yoshida, N; Schenkman, S. 2006. Expression of trypomastigote trans-sialidase in metacyclic forms of *Trypanosoma cruzi* increases parasite escape from its parasitophorous vacuole. *Cell Microbiol.* Vol 8. p. 1888-98.
- Silveira**, A.C. 2000. Current situation with Chagas disease vector control in the Americas. *Cad. Saúde Publica.* Vol. 2. p. 35-42
- Scharfstein**, J; Schmitz, V; Morandi; Capella, M. M; Lima, A. P; Morrot, A; Juliano, L; Muller-Esterl, W. 2000. Host cell invasion by *Trypanosoma cruzi* is potentiated by activation of bradykinin B (2) receptors. *J Exp Med.* Vol.192. p 1289-300
- Schenkman**, S; Mortara, R.A. 1992. HeLa cells extend and internalize pseudopodia during active invasion by *Trypanosoma cruzi* trypomastigotes. *J Cell Sci.* Vol. 101. p. 895-905
- Schenkman**, S; Briones, B. M. R; Chaves, L. 1993. Trans-sialidase from *Trypanosoma cruzi* epimastigotes is expressed at the stationary phase and is different from the enzyme expressed in trypomastigotes *Mol Biochem Parasitol.* Vol. 61. p. 97-106
- Schenkman**, S; Uemura, H; Schenkman, R.P.F; Freitas Junior, L.H.G; Briones, M.R.S; Egima, C.M. 1996. Organization of trans-sialidase genes in *Trypanosoma cruzi*. *Molecular and Biochemical Parasitology.* Vol. 77. P. 115-125
- Schenkman**, S; Alves, M.J.M; Colli, W; Freitas-Junior, L.H.G; Abuin, G. 1999. Expression of trans-sialidase and 85-kDa Glycoprotein genes in *Trypanosoma cruzi* is differentially regulated at the post-transcriptional level by labile protein factors. *The Journal of Biological Chemistry.* Vol. 274. P. 13041-13047
- Schneider**, T.D; Stephens, R.M. 1990. Sequence Logos: A New Way to Display Consensus Sequences. *Nucleic Acids Res.* Vol. 18. p. 6097-6100
- Spence**, J. 2009. Pathway prediction by bioinformatic analysis of the untranslated regions of the CFTR mRNA. *Genomics*
- Souto**, R. P; Zingales, B. 1993. Sensitive detection and strain classification of *Trypanosoma cruzi* by amplification of a ribosomal RNA sequence. *Mol Biochem Parasitol.* Vol. 62. P. 45-52
- Souto**, R. P; Fernandes, O; Macedo, A. M; Campbell, D. A; Zingales, B. 1996. DNA markers define two major phylogenetic lineages of *Trypanosoma cruzi*. *Mol Biochem Parasitol.* Vol. 83. P. 141-52
- Tanowitz**, H. B; Kirchhoff, L. V; Simon, D; Morris, S. A; Weiss, L.M; Wittfner, M. 1992. Chagas' Disease. *Clinical Microbiology Reviews.* p. 400-419

- Tardieux**, I; Nathanson, M.H. 1994. Role in host cell invasion of *Trypanosoma cruzi* – induced cytosolic – free Ca²⁺ transients. J. Exp med. Vol. 179. p 1017 – 22
- Teixeira**, S. M; Kirchhoff, L. V; Donelson, J. E. 1999. *Trypanosoma cruzi*: suppression of tuzin gene expression by its 5'-UTR and spliced leader addition site. Exp Parasitol. Vol. 93. P. 143-51
- Tibayrenc**, M; Ward, P; Moya, A; Ayala, F.J. 1986. Natural populations of *Trypanosoma cruzi*, the agent of Chagas disease, have a complex multiclonal structure. Proc Natl Acad Sci USA. Vol. 83. p. 115-119.
- Tomlinson**, S; Vandekerckhove, F; Frevert, U; Nussenzweig, V. 1995. The Induction of *Trypanosoma cruzi* trypomastigote to amastigote transformation by low pH. Parasitology. Vol. 110. p. 547–554
- Tomlinson**, S; Raper, J. 1998. Natural human immunity to trypanosomes. Parasitol. Today. Vol. 14. P. 354-359
- Tran**, M. K; Schultz, C. J; Baumann, Ute. 2008. Conserved upstream open reading frames in higher plants. BMC Genomics. Vol. 9. p. 361
- Tribulatti**, M. V ; Mucci, J. Van Rooijen, N; Leguizamon, M. S; Campetella, O. 2005. The trans-sialidase from *Trypanosoma cruzi* induces thrombocytopenia during acute Chagas' disease by reducing the platelet sialic acid contents. Infect Immun. Vol. 73. p. 201-7
- Ullu**, E; Tschudi, C; Gunzi, A. 1996. *Trans-splicing in trypanosomatid protozoa*. In: Smith, D. F; Parsons, M (eds). Molecular Biology of Parasitic Protozoa. IRL Press Oxford, p. 115-129
- Van der Velden**, A.W; Thomas, A.A. 1999. The role of the 5' untranslated region of an mRNA in translation regulation during development. Int J Biochem Cell Biol. Vol. 31. p. 87-106
- Vilela**, C; McCarthy, J.E. 2003. Regulation of fungal gene expression via short open reading frames in the mRNA 5'untranslated region. Mol Microbiol. 49 (4). p 859-67
- Vinhaes**, M.C; Dias, J.C.P. 2000. Doença de Chagas no Brasil. Caderno de Saúde Pública, Rio de Janeiro, Vol. 16. p. 7-12
- Yoshida**, N; Favoreto, S. Jr; Ferreira, A.T; Manque, P.M. 2000. Signal transduction induced in *Trypanosoma cruzi* metacyclic trypomastigotes during the invasion of mammalian cells. Braz J Med Biol Res. Vol. 33. p 269-78
- Yoshida**, N. 2003. *Trypanosoma cruzi* cell invasion mechanisms. World Class Parasites. Tyler KM and Miles MA (eds), Kluwer Academic Publishers. P. 69-79

- Yoshida**, N. 2006. Molecular basis of mammalian cell invasion by *Trypanosoma cruzi*. *An Acad Bras Cienc*. Vol 7
- Yoshida**, N; Cortez, M. 2008. *Trypanosoma cruzi*: parasite and host cell signaling during the invasion process. *Subcell Biochem*. Vol. 47. p 82-91
- Wagner**, W; So, M. 1990. Genomic Variation of *Trypanosoma cruzi*: Involvement of Multicopy Genes. *Infection and Immunity*. Vol. 58. p. 3217-3224
- Weston**, D; La Flamme, A. C; Van Voorhis, W. C. 1999. Expression of *Trypanosoma cruzi* surface antigen FL-160 is controlled by elements in the 3' untranslated, the 3' intergenic, and the coding regions. *Mol Biochem Parasitol*. Vol. 102. 53-66
- Zingales**, B; Souto, R. P; Mangia, R. H; Lisboa, C. V; Campbell, D. A; Coura, J. R; Jansen, A; Fernandes, O. 1998. Molecular epidemiology of American trypanosomiasis in Brazil based on dimorphisms of rRNA and mini-exon gene sequences. *Int J Parasitol*. Vol. 28. P. 105-12

ANEXOS 1

- Meio LB (Luria-Bertani) líquido

- Triptona (10g/L)
- Extrato de levedura (5g/L)
- Cloreto de sódio (10g/L)

- Meio LB (Luria-Bertani) sólido

- Triptona (10g/L)
- Extrato de levedura (5g/L)
- Cloreto de sódio (10g/L)
- Agar (20g/L)

4- Tampão Tris-EDTA (TE)

- 10mM de Tris (pH 8,0)
- 1mM de EDTA

5- Tampão TBE 0,5X (Tris-borato/EDTA eletroforesis buffer)

(Tris base 89 mM , ácido bórico 89 mM e EDTA 2mM)

- 5,4g de Tris base
- 2,25g de ácido bórico
- 0,465g de EDTA
- Água destilada até completar o volume de 1L

6- Brometo de etídio (0,5µg / mL)

- 25µL de brometo de etídeo
- 500mL de água destilada

7- X-gal 20mg/mL (5-bromo-4-chloro-3-indolyl- β -D-galactopyranoside)

- 1g de X-gal
- 50 mL de dimetilformamida

ANEXO 2

Representação dos dados obtidos através do site do NCBI, informando o número de acesso, o Locus gênico, a seqüência nucleotídica e as regiões 5' (Downstream) e 3'(Upstream) de cada uma das 735 seqüências supracitadas.

Gene ID	Locus tag	Genomic	89205	Possível	Possível
3538544	Tc00.1047053503533.30	AAHK0100	7512...4758	7512...8113	4757...2
3538584	Tc00.1047053507145.60	AAHK0100	9250...6793	9250...9851	6792..49
3539873	Tc00.1047053508745.30	AAHK0100	9383...6848	9383...9983	6848...5
3541150	Tc00.1047053503447.50	AAHK0100	15630...1297	15631...1623	12974...1
3539879	Tc00.1047053508745.11	AAHK0100	1270...1	1271...1871	1..1
3534682	Tc00.1047053509629.10	AAHK0100	120...2884	121..1	2884..44
3552801	Tc00.1047053510307.230	AAHK0100	48957..46788	48957..49557	46788..4
3547633	Tc00.1047053506741.100	AAHK0100	22810...2549	22810...22210	25496...1
3548980	Tc00.1047053506279.60	AAHK0100	16676...1387	16675..17275	13880...1
3538570	Tc00.1047053508963.30	AAHK0100	6170..9304	6170..5570	9304..11
3552802	Tc00.1047053510307.240	AAHK0100	54264..51112	54264..54864	51112..4
3545052	Tc00.1047053511923.30	AAHK0100	2795..5416	2794..2194	5416..72
3547604	Tc00.1047053511233.140	AAHK0100	33445..36099	33445..32845	36099..3
3538509	Tc00.1047053508627.40	AAHK0100	11360..8184	11360..11960	8183..63
3547636	Tc00.1047053506741.130	AAHK0100	32438..35248	32438..31838	35248..3
3538521	Tc00.1047053504533.40	AAHK0100	8081..10837	8081..7481	10837..1
3538538	Tc00.1047053504261.20	AAHK0100	1008..3836	1008..408	3836..56
3545008	Tc00.1047053507091.60	AAHK0100	13699..10907	13699..14299	10907..9
3550260	Tc00.1047053510713.160	AAHK0100	55311..52726	55311..55911	52726..5
3550253	Tc00.1047053510713.30	AAHK0100	22962..20476	22962..23562	20476..1
3545054	Tc00.1047053511923.60	AAHK0100	13920..16538	13920..13320	16538..1
3550256	Tc00.1047053510713.90	AAHK0100	35812..33158	35812..36412	33158..3
3545022	Tc00.1047053506975.90	AAHK0100	26538..29654	26538..25938	29654..3
3534654	Tc00.1047053504341.10	AAHK0100	4481..1698	4481..5081	1698..1
3552807	Tc00.1047053510307.284	AAHK0100	64123..60722	64123..64723	60722..5
3533328	Tc00.1047053510907.10	AAHK0100	4229..1371	4229..4829	1371..1
3539803	Tc00.1047053506279.140	AAHK0100	29588..26856	29588..30188	26856..2
3539803	Tc00.1047053508451.60	AAHK0100	17152..14774	17152..17752	14774..1
3538591	Tc00.1047053506187.20	AAHK0100	4709..2067	4709..5309	2067..26
3547623	Tc00.1047053511077.7	AAHK0100	720..1	720..1320	1..1
3541142	Tc00.1047053508109.60	AAHK0100	14209..16995	14209..13609	16995..1
3545021	Tc00.1047053506975.80	AAHK0100	19832..21952	19832..16232	21952..2
3554169	Tc00.1047053510377.10	AAHK0100	3125..765	3125..3725	765..1
3546349	Tc00.1047053506909.100	AAHK0100	18858..21506	18858..18258	21506..2
3538516	Tc00.1047053510491.60	AAHK0100	11140..8483	11140..11740	8483..66
3534689	Tc00.1047053509513.10	AAHK0100	3318..5582	3318..2718	5582..73
3533390	Tc00.1047053508225.4	AAHK0100	1280..1	1280..1880	1..1
3537232	Tc00.1047053511323.10	AAHK0100	1280..1	1280..1880	1..1
3541148	Tc00.1047053503447.20	AAHK0100	2550..286	2550..3150	286..1
3537272	Tc00.1047053511551.10	AAHK0100	565..2949	565..1	2949..47
3554171	Tc00.1047053510377.50	AAHK0100	11625..9295	11625..12225	9295..74
3554153	Tc00.1047053510377.330	AAHK0100	68132..70558	68132..67532	70558..7
3538535	Tc00.1047053510709.70	AAHK0100	10234..1290	10234..9634	12903..1
3532029	Tc00.1047053507095.30	AAHK0100	1604..2221	1604..1004	2221..40
3532053	Tc00.1047053511777.10	AAHK0100	1163..2263	1163..563	2263..40
3548978	Tc00.1047053506279.10	AAHK0100	5537..2751	5537..6137	2751..95
3546340	Tc00.1047053506909.160	AAHK0100	31414..34218	31414..30814	34218..3
3552744	Tc00.1047053509897.200	AAHK0100	63047..60432	63047..63647	60432..5
3547519	Tc00.1047053511861.50	AAHK0100	16198..13037	16198..16798	13037..1
3543692	Tc00.1047053511401.30	AAHK0100	7279..4502	7279..7879	4502..27
3541014	Tc00.1047053511237.130	AAHK0100	20685..21253	20685..20085	21253..2
3537143	Tc00.1047053506373.39	AAHK0100	11478..12029	11478..10878	12029..1
3548807	Tc00.1047053509905.200	AAHK0100	47037..44773	47037..47637	44773..4
3541029	Tc00.1047053507523.70	AAHK0100	12438..15062	12438..11838	15062..1
3533297	Tc00.1047053503403.20	AAHK0100	3747..1363	3747..4347	1363..1
3533228	Tc00.1047053510803.19	AAHK0100	3681..4918	3681..3081	4918..67
3546202	Tc00.1047053510715.90	AAHK0100	24542..21855	24542..25142	21855..2
3534555	Tc00.1047053510921.40	AAHK0100	5843..6985	5843..5243	69.858.7
3538401	Tc00.1047053507121.11	AAHK0100	2154..1	2154..2754	1..1
3550118	Tc00.1047053509419.40	AAHK0100	8030..5538	8030..8630	5538..37
3543695	Tc00.1047053511401.90	AAHK0100	18559..15764	18559..19159	15764..1
3552734	Tc00.1047053509897.20	AAHK0100	7650..5023	7650..8250	5023..32
3548804	Tc00.1047053509905.170	AAHK0100	40333..37199	40333..40933	37199..3
3552739	Tc00.1047053509897.140	AAHK0100	41122..38513	41122..41722	38513..3
3548803	Tc00.1047053509905.150	AAHK0100	35005..32786	35005..35605	32786..3
3531976	Tc00.1047053460111.10	AAHK0100	2100..1	2100..2700	1..1
3552736	Tc00.1047053509897.70	AAHK0100	18937..16160	18937..19537	16160..1
3539773	Tc00.1047053510207.30	AAHK0100	8939..11437	8939..8339	11437..1
3533258	Tc00.1047053511831.10	AAHK0100	1901..4981	1901..1301	4981..67
3543697	Tc00.1047053511401.120	AAHK0100	25750..22976	25750..26350	22976..2
3535841	Tc00.1047053510331.10	AAHK0100	5086..2135	5086..5686	2135..33

3537108	Tc00.1047053510163.60Tc	AAHK0100	8826..11159	8826..8226	11159..1
3554025	Tc00.1047053507611.170	AAHK0100	52013..49650	52013..52613	49650..4
3550149	Tc00.1047053508837.70	AAHK0100	12223..10037	12223..12823	10037..8
3539757	Tc00.1047053508161.50	AAHK0100	18278..16280	18278..18878	16280..1
3544971	Tc00.1047053506737.90	AAHK0100	20248..17567	20248..20848	17567..1
3544989	Tc00.1047053509265.110	AAHK0100	28349..23103	28349..28949	23103..2
3542321	Tc00.1047053508125.149	AAHK0100	23267..24838	23267..22667	24838..2
3547518	Tc00.1047053511861.20	AAHK0100	10725..9133	10725..11325	9133..73
3552725	Tc00.1047053508325.230	AAHK0100	54175..58734	54175..53575	58734..6
3537133	Tc00.1047053511779.60	AAHK0100	8627..11395	8627..8027	11395..1
3546206	Tc00.1047053510715.150	AAHK0100	33133..3595	33133..32533	35958..3
3548864	Tc00.1047053506577.80	AAHK0100	24568..21851	24568..25168	21851..2
3542368	Tc00.1047053511449.70	AAHK0100	24568..21851	24568..25168	21851..2
3550119	Tc00.1047053509419.70	AAHK0100	16377..13513	16377..16977	13513..1
3544988	Tc00.1047053509265.90	AAHK0100	21377..18729	21131..21731	18729..1
3550126	Tc00.1047053509419.90	AAHK0100	20457..22835	20457..19857	22835..2
3538419	Tc00.1047053508759.20	AAHK0100	2051..4732	2051..1451	4732..65
3552749	Tc00.1047053509897.270	AAHK0100	79434..76798	79434..80034	76798..7
3544982	Tc00.1047053509265.120	AAHK0100	32734..30479	32734..33334	30479..2
3533256	Tc00.1047053421173.4	AAHK0100	2455..1	2455..3055	1..1
3552738	Tc00.1047053509897.110	AAHK0100	30003..27418	30003..30603	27418..2
3535802	Tc00.1047053508511.49		7723..9287	7723..7123	9287..11
3544958	Tc00.1047053510095.20	AAHK0100	6868..4559	6868..7468	4559..27
3539754	Tc00.1047053508161.20	AAHK0100	2761..5322	2761..2161	5322..71
3547592	Tc00.1047053510503.50	AAHK0100	10935..13634	10935..10335	13634..1
3552796	Tc00.1047053510307.210	AAHK0100	38615..36669	38615..39215	36669..3
3535780	Tc00.1047053509187.10	AAHK0100	509..2776	509..1	2776..45
3535763	Tc00.1047053503759.10	AAHK0100	944..3217	944..344	3217..50
3552664	Tc00.1047053509875.80	AAHK0100	33088..30821	33088..33688	30821..2
3546199	Tc00.1047053510715.20	AAHK0100	6292..3497	6292..6892	3497..16
3533163	Tc00.1047053508913.25	AAHK0100	3667..3056	3667..3067	3056..12
3535766	Tc00.1047053509387.19	AAHK0100	7937..9251	7937..7337	9251..11
3534464	Tc00.1047053505363.19	AAHK0100	5527..6865	5527..4927	6865..86
3553993	Tc00.1047053507071.390	AAHK0100	97318..10017	97318..96718	100173..
3553926	Tc00.1047053510279.320	AAHK0100	96682..93950	96682..97282	93950..9
3531822	Tc00.1047053425435.10	AAHK0100	1830..1	1830..2430	1..1
3534478	Tc00.1047053503705.19	AAHK0100	5863..6877	5863..5263	6877..86
3553921	Tc00.1047053510279.230	AAHK0100	55899..53002	55899..56499	53002..5
3548702	Tc00.1047053510497.40	AAHK0100	8691..6085	8691..9291	6085..42
3552633	Tc00.1047053509287.80	AAHK0100	18869..21565	18869..18269	21565..2
3535765	Tc00.1047053509387.10	AAHK0100	2616..5732	2616..2016	5732..75
3542260	Tc00.1047053508903.110	AAHK0100	24722..21619	24722..25322	21619..1
3535779	Tc00.1047053510233.39	AAHK0100	7177..9252	7177..6577	9252..11
3553952	Tc00.1047053508355.80	AAHK0100	24041..21252	24041..24641	21252..1
3553970	Tc00.1047053507071.80	AAHK0100	16433..19303	16433..15833	19303..2
3537067	Tc00.1047053507839.40	AAHK0100	11304..9043	11304..11904	9043..72
3553985	Tc00.1047053507071.230	AAHK0100	63689..66463	63689..63089	66463..6
3534481	Tc00.1047053509185.30	AAHK0100	3683..6046	3683..3083	6046..78
3551316	Tc00.1047053506717.80	AAHK0100	26653..29820	26653..26053	29820..3
3539618	Tc00.1047053508375.20	AAHK0100	4644..7436	4644..4044	7436..92
3553902	Tc00.1047053510279.110	AAHK0100	48173..50908	48173..47573	50908..5
3553982	Tc00.1047053507071.190	AAHK0100	48173..50908	48173..47573	50908..5
3552651	Tc00.1047053509875.120	AAHK0100	41712..39541	41712..42312	39541..3
3553999	Tc00.1047053507071.10	AAHK0100	4784..2094	4784..5384	2094..29
3534463	Tc00.1047053505363.10	AAHK0100	937..3135	937..337	3135..49
3548771	Tc00.1047053507953.100	AAHK0100	17493..15145	17493..18093	15145..1
3533104	Tc00.1047053510107.15	AAHK0100	715..1	715..1315	1..1
3534405	Tc00.1047053503907.10	AAHK0100	5334..2575	5334..5934	2575..77
3547483	Tc00.1047053508593.20	AAHK0100	9562..6839	9562..10162	6839..50
3533181	Tc00.1047053504145.10	AAHK0100	667..2907	667..67	2907..47
3534475	Tc00.1047053511833.10	AAHK0100	2408..5548	2408..1808	5548..73
3547482	Tc00.1047053508593.10	AAHK0100	4815..2092	4815..5415	2092..29
3547423	Tc00.1047053506939.120	AAHK0100	22032..24923	22032..21432	24923..2
3538313	Tc00.1047053510125.20	AAHK0100	9639..6475	9639..10239	6475..46
3551373	Tc00.1047053506067.300	AAHK0100	62846..65602	62846..62246	65602..6
3538399	Tc00.1047053507121.20	AAHK0100	6583..4106	6583..7183	4106..23
3546111	Tc00.1047053506339.10	AAHK0100	343..3066	343..1	3066..48
3552641	Tc00.1047053509287.230	AAHK0100	57573..59741	57573..56973	59741..6
3548774	Tc00.1047053507953.140	AAHK0100	27688..25370	27688..28288	25370..2
3549958	Tc00.1047053506499.170	AAHK0100	37999..40359	37999..37399	40359..4
3552514	Tc00.1047053507063.70	AAHK0100	14301..13701	16451..18251	16451..1
3543457	Tc00.1047053508977.120	AAHK0100	24485..21714	24485..25085	21714..1
3536966	Tc00.1047053509533.10	AAHK0100	6576..4354	6576..7176	4354..25
3534357	Tc00.1047053510419.20	AAHK0100	1839..1239	1839..1239	4661..64
3551222	Tc00.1047053510729.280	AAHK0100	55426..58368	55426..54826	58368..6
3534375	Tc00.1047053506035.4	AAHK0100	1518..1	1518..2118	1..1
3552563	Tc00.1047053504239.260	AAHK0100	44064..46403	44064..43464	46403..4
3551235	Tc00.1047053510197.180	AAHK0100	54828..51958	54828..55428	51958..5
3548662	Tc00.1047053508871.110	AAHK0100	29708..32035	29708..29108	32035..3
3543452	Tc00.1047053508977.10	AAHK0100	2585..102	2585..3185	102..1
3546035	Tc00.1047053508789.120	AAHK0100	23129..23729	23129..23729	20358..1
3534382	Tc00.1047053505231.20	AAHK0100	2638..4947	2638..2038	4947..67
3551229	Tc00.1047053510197.30	AAHK0100	11040..13292	11040..10440	13292..1
3547393	Tc00.1047053510055.20	AAHK0100	1949..1341	1949..2549	1341..1

3534389	Tc00.1047053507283.10	AAHK0100	57..2252	57..1	2252..40
3533002	Tc00.1047053503661.11	AAHK0100	2066..1	2066..2666	1..1
3551240	Tc00.1047053510197.229	AAHK0100	61757..63930	61757..61157	63930..6
3548666	Tc00.1047053508871.170	AAHK0100	44283..41611	44283..44883	41611..3
3534388	Tc00.1047053508107.30	AAHK0100	5874..3010	5874..6474	3010..12
3549987	Tc00.1047053506241.30	AAHK0100	20228..17967	20228..20822	17967..1
3548699	Tc00.1047053511911.60	AAHK0100	19810..16682	19810..20410	16682..1
3533075	Tc00.1047053510813.20	AAHK0100	3938..4563	3938..3338	4563..63
3542178	Tc00.1047053511875.20	AAHK0100	3881..6610	3881..3281	6610..84
3552589	Tc00.1047053504239.434	AAHK0100	76741..77871	76741..76141	77871..7
3548656	Tc00.1047053508871.10	AAHK0100	1446..3794	1446..846	3794..55
3539562	Tc00.1047053510205.40	AAHK0100	11159..8835	11159..11759	8835..70
3534395	Tc00.1047053503767.10	AAHK0100	1735..3897	1735..1135	3897..56
3531777	Tc00.1047053398265.9	AAHK0100	310..1763	310..1	1763..35
3551233	Tc00.1047053510197.130	AAHK0100	47920..45050	47920..48520	45050..4
3546026	Tc00.1047053510853.40	AAHK0100	12535..9689	12535..13133	9689..78
3543456	Tc00.1047053508977.80	AAHK0100	17170..14408	17170..17770	14408..1
3533086	Tc00.1047053504189.5	AAHK0100	768..1	768..1368	1..1
3549937	Tc00.1047053504155.20	AAHK0100	1916..4609	1916..1316	4609..64
3544789	Tc00.1047053510371.90	AAHK0100	20609..18153	20609..21209	18153..1
3543490	Tc00.1047053506703.90	AAHK0100	11996..14845	11996..11396	14845..1
3549985	Tc00.1047053506955.240	AAHK0100	44890..42083	44890..45490	42083..4
3540855	Tc00.1047053506053.50	AAHK0100	14566..17337	14566..13966	17337..1
3549966	Tc00.1047053506955.270	AAHK0100	53485..50783	53485..54085	50783..4
3538146	Tc00.1047053505993.30	AAHK0100	7310..4179	7310..7910	4179..23
3555087	Tc00.1047053508221.750	AAHK0100	175449..1728	175449..1760	
3551184	Tc00.1047053507875.70	AAHK0100	17906..20509	17906..17306	20509..2
3540703	Tc00.1047053511885.20	AAHK0100	8042..10372	8042..7442	10372..1
3536877	Tc00.1047053508753.20	AAHK0100	4541..7054	4541..3941	7054..88
3555082	Tc00.1047053508221.650	AAHK0100	152697..1497	152697..1532	149725..
3536824	Tc00.1047053509031.3	AAHK0100	1437..1	1437..2037	1..1
3534290	Tc00.1047053509411.10	AAHK0100	1401..3932	1401..801	3932..57
3547299	Tc00.1047053506443.100	AAHK0100	17703..20471	17703..17103	20471..2
3544610	Tc00.1047053510237.80	AAHK0100	11564..14275	11564..10964	14275..1
3548565	Tc00.1047053507723.50	AAHK0100	17767..15053	17767..18367	15053..1
3555079	Tc00.1047053508221.590	AAHK0100	138408..1355	135556..1337	135556..
3547216	Tc00.1047053511301.160	AAHK0100	27254..29374	27254..26654	29374..3
3532968	Tc00.1047053509927.10	AAHK0100	1544..1	1544..2144	1..1
3545967	Tc00.1047053511921.50	AAHK0100	6299..9094	6299..5699	9094..10
3542064	Tc00.1047053508313.14	AAHK0100	1694..1	1694..2294	1..1
3547239	Tc00.1047053506667.90	AAHK0100	20462..23008	20462..19862	23008..2
3555095	Tc00.1047053508221.870	AAHK0100	200760..2034	200760..2001	203426..
3536839	Tc00.1047053511105.60	AAHK0100	11460..9091	11460..12060	9091..72
3539410	Tc00.1047053508297.7	AAHK0100	1339..1	1339..1939	1..1
3545972	Tc00.1047053511921.150	AAHK0100	28752..31526	28752..28152	31526..3
3534215	Tc00.1047053509257.10	AAHK0100	1223..3337	1223..623	3337..51
3555088	Tc00.1047053508221.790	AAHK0100	185594..1832	185594..1861	183219..
3531637	Tc00.1047053423205.10	AAHK0100	1198..1607	1198..598	1607..34
3540774	Tc00.1047053509581.10	AAHK0100	8856..5413	8856..9456	5413..36
3534213	Tc00.1047053511669.10	AAHK0100	4671..1534	4671..5271	1534..1
3539496	Tc00.1047053506159.5	AAHK0100	737..1	737..1337	1..1
3547230	Tc00.1047053506309.60	AAHK0100	17047..14579	17047..17647	14579..1
3538101	Tc00.1047053508229.40	AAHK0100	6263..8680	6263..5663	8680..10
3538154	Tc00.1047053509115.20	AAHK0100	4452..1825	4452..5052	1825..25
3555017	Tc00.1047053508221.180	AAHK0100	35166..37814	35166..34566	37814..3
3547244	Tc00.1047053506667.160	AAHK0100	35408..37843	35408..34808	37843..3
3551178	Tc00.1047053510643.40	AAHK0100	17822..20983	17822..17222	20983..2
3549887	Tc00.1047053506973.120	AAHK0100	45984..48674	45984..45384	48674..5
3549894	Tc00.1047053506973.10	AAHK0100	2148..4886	2148..1548	4886..66
3532981	Tc00.1047053468217.30	AAHK0100	4398..2973	4398..4998	2973..11
3535554	Tc00.1047053507219.9	AAHK0100	2468..1	2468..3068	1..1
3543379	Tc00.1047053507357.100	AAHK0100	19363..21771	19363..18763	21771..2
3555057	Tc00.1047053508221.400	AAHK0100	82151..81551	82151..81551	84688..8
3540779	Tc00.1047053509427.10	AAHK0100	781..2976	781..181	2976..47
3532938	Tc00.1047053508309.5	AAHK0100	777..1	777..1377	1..1
3535532	Tc00.1047053511413.19	AAHK0100	7344..8790	7344..6744	8790..10
3535517	Tc00.1047053506455.30	AAHK0100	4435..6696	4435..3835	6696..84
3547201	Tc00.1047053508515.150	AAHK0100	34484..31341	34484..35084	31341..2
3549768	Tc00.1047053511149.180	AAHK0100	41459..38613	41459..42059	38613..3
3534168	Tc00.1047053509739.10	AAHK0100	3919..1436	3919..4519	1436..1
3535472	Tc00.1047053506347.5	AAHK0100	740..1	740..1340	1..1
3548431	Tc00.1047053507085.30	AAHK0100	5320..7998	5320..4720	7998..97
3551030	Tc00.1047053508365.190	AAHK0100	32024..29235	32024..32624	29235..2
3554909	Tc00.1047053511183.230	AAHK0100	60414..57859	60414..61014	57859..5
3540653	Tc00.1047053506951.80	AAHK0100	11767..14439	11767..11167	14439..1
3538092	Tc00.1047053509685.40	AAHK0100	9967..12594	9967..9367	12594..1
3554915	Tc00.1047053511183.60	AAHK0100	12466..9926	12466..13066	9926..81
3544583	Tc00.1047053506129.30	AAHK0100	8832..6646	8832..9432	6646..48
3548468	Tc00.1047053506757.60	AAHK0100	9284..11707	9284..8684	11707..1
3541902	Tc00.1047053508979.50	AAHK0100	13717..10916	13717..14317	10916..9
3539358	Tc00.1047053509971.90	AAHK0100	17296..17296	16168..15598	17296..1
3554906	Tc00.1047053511183.190	AAHK0100	50095..47531	50095..50695	47531..4
3539370	Tc00.1047053504787.40	AAHK0100	4697..7162	4697..4097	7162..89
3543247	Tc00.1047053508047.60	AAHK0100	12171..9400	12171..12771	9400..76

3544584	Tc00.1047053506129.50	AAHK0100	15299..10809	15299..15809	10809..9
3536736	Tc00.1047053504101.50	AAHK0100	7310..4980	7310..7610	4980..31
3534126	Tc00.1047053511399.24	AAHK0100	2642..1	2642..3242	1..1
3534142	Tc00.1047053508171.10	AAHK0100	475..3603	475..1	3603..54
3548439	Tc00.1047053510543.70	AAHK0100	17227..19404	17227..16627	19404..2
3540691	Tc00.1047053507819.30	AAHK0100	9210..11573	9210..8610	11573..1
3548470	Tc00.1047053506757.90	AAHK0100	19260..21764	19260..18660	21764..2
3538022	Tc00.1047053504031.50	AAHK0100	12158..9363	12158..12758	9363..75
3539310	Tc00.1047053506603.40	AAHK0100	10666..8336	10666..11266	8336..65
3540654	Tc00.1047053506951.90	AAHK0100	15723..18473	15723..15123	18473..2
3554943	Tc00.1047053511183.400	AAHK0100	107374..1102	107374..1067	110238..
3539339	Tc00.1047053506597.40	AAHK0100	11851..14184	11851..11251	14184..1
3554945	Tc00.1047053511183.430	AAHK0100	115640..1182	115640..1150	118201..
3538016	Tc00.1047053508305.40	AAHK0100	10620..7930	10620..11220	7930..61
3538062	Tc00.1047053511873.10	AAHK0100	3973..6531	3973..3373	6531..83
3541934	Tc00.1047053508717.60	AAHK0100	22323..19966	22323..22923	19966..1
3549784	Tc00.1047053504115.10	AAHK0100	8329..5168	8329..8929	5168..33
3544529	Tc00.1047053504425.10	AAHK0100	743..3751	743..143	3751..55
3538097	Tc00.1047053508229.60	AAHK0100	10752..13916	10752..10152	13916..1
3549787	Tc00.1047053504115.60	AAHK0100	27717..24949	27717..28317	24949..2
3548485	Tc00.1047053511827.110	AAHK0100	24707..26863	24707..24107	26863..2
3548472	Tc00.1047053506757.120	AAHK0100	29753..32194	29753..29153	32194..3
3540569	Tc00.1047053509867.70	AAHK0100	12126..14993	12126..11526	14993..1
3536627	Tc00.1047053504193.15	AAHK0100	818..1	818..1418	1..1
3539248	Tc00.1047053505207.10	AAHK0100	4171..1700	4171..4771	1700..1
3550963	Tc00.1047053505997.80	AAHK0100	24075..20959	24075..24675	20959..1
3535382	Tc00.1047053504007.20	AAHK0100	7898..5274	7898..8498	5274..34
3553536	Tc00.1047053511603.90	AAHK0100	18131..15393	18131..18731	15393..1
3543154	Tc00.1047053507479.20	AAHK0100	2800..5238	2800..2200	5238..70
3550999	Tc00.1047053508433.40	AAHK0100	10795..13329	10795..10195	13329..1
3541817	Tc00.1047053510403.30	AAHK0100	9630..7369	9630..10230	7369..55
3543169	Tc00.1047053508859.118	AAHK0100	25366..27028	25366..24766	27028..2
3552245	Tc00.1047053506537.80	AAHK0100	26196..28595	26196..25595	28595..3
3535340	Tc00.1047053506841.20	AAHK0100	7323..4966	7323..7923	4966..31
3540514	Tc00.1047053509735.10	AAHK0100	808..3102	808..208	3102..49
3540509	Tc00.1047053509735.50	AAHK0100	8187..10934	8187..7587	10934..1
3553573	Tc00.1047053509979.320	AAHK0100	79756..77408	79756..80356	77408..7
3545719	Tc00.1047053510201.10	AAHK0100	1642..4518	1642..1042	4518..63
3548357	Tc00.1047053509217.40	AAHK0100	8144..10858	8144..7544	10858..1
3534051	Tc00.1047053504183.20	AAHK0100	2206..4689	2206..1606	4689..64
3553505	Tc00.1047053510275.350	AAHK0100	73073..75757	73073..72473	75757..7
3531446	Tc00.1047053401961.10	AAHK0100	1384..1	1384..1984	1..1
3552252	Tc00.1047053506537.200	AAHK0100	53103..56441	53103..52503	56441..5
3534069	Tc00.1047053507505.10	AAHK0100	1..2071	1..1	2071..38
3545745	Tc00.1047053508045.120	AAHK0100	33986..30978	33986..34586	30978..2
3543155	Tc00.1047053507479.70	AAHK0100	21806..24652	21806..21206	24652..2
3549651	Tc00.1047053508023.100	AAHK0100	24441..21673	24441..25041	21673..1
3541848	Tc00.1047053510697.50	AAHK0100	6707..9583	6707..6107	9583..11
3543119	Tc00.1047053506345.40	AAHK0100	10152..7729	10152..10752	7729..59
3553553	Tc00.1047053511603.450	AAHK0100	78762..81101	78762..78168	81101..8
3545728	Tc00.1047053509777.30	AAHK0100	8646..6493	8646..9246	6493..46
3532714	Tc00.1047053511917.10	AAHK0100	1..2269	1..1	2269..28
3552216	Tc00.1047053506763.110	AAHK0100	17921..21082	17921..17321	21082..2
3549693	Tc00.1047053509157.170	AAHK0100	31566..34526	31566..30966	34526..3
3549658	Tc00.1047053510025.50	AAHK0100	8030..10612	8030..7430	10612..1
3532791	Tc00.1047053503427.9	AAHK0100	2099..1	2099..2699	1..1
3544462	Tc00.1047053511587.90	AAHK0100	22268..19122	22268..22868	19122..1
3535306	Tc00.1047053509931.20	AAHK0100	3570..6704	3570..2970	6704..85
3534024	Tc00.1047053472357.10	AAHK0100	1541..3982	1541..941	3982..57
3535365	Tc00.1047053511123.11	AAHK0100	1477..1	1477..2077	1..1
3543121	Tc00.1047053506345.90	AAHK0100	20674..18107	20674..21274	18107..1
3534065	Tc00.1047053422867.10	AAHK0100	38..2152	38..1	2152..39
3548257	Tc00.1047053507047.40	AAHK0100	16174..13874	16174..16747	13874..1
3533990	Tc00.1047053506283.20	AAHK0100	5206..6138	5206..4606	6138..79
3536526	Tc00.1047053509575.4	AAHK0100	702..1	702..1302	1..1
3544346	Tc00.1047053507687.10	AAHK0100	1..1446	1..1	1446..32
3531369	Tc00.1047053432995.9	AAHK0100	1..1333	1..1	1333..31
3536586	Tc00.1047053511061.40	AAHK0100	8573..6039	8573..9173	6039..42
3553410	Tc00.1047053506173.70	AAHK0100	13502..16153	13502..12902	16153..1
3546930	Tc00.1047053508637.10	AAHK0100	3947..849	3947..4547	849..1
3539103	Tc00.1047053508013.40	AAHK0100	7331..4482	7331..7931	4482..26
3554739	Tc00.1047053510359.610	AAHK0100	127322..1251	127322..1279	125130..
3540486	Tc00.1047053507907.20	AAHK0100	3141..5333	3141..2541	5333..71
3533956	Tc00.1047053463323.10	AAHK0100	3704..1446	3704..4304	1446..1
3536569	Tc00.1047053507473.30	AAHK0100	3849..6110	3849..3249	6110..79
3545681	Tc00.1047053506001.40	AAHK0100	10637..7986	10637..11237	7986..61
3545681	Tc00.1047053508873.390	AAHK0100	83939..81456	83939..84539	81456..7
3535253	Tc00.1047053510033.30	AAHK0100	5552..7885	5552..4952	7885..96
3532684	Tc00.1047053510049.10	AAHK0100	3778..1915	3778..4378	1915..11
3532643	Tc00.1047053511891.10	AAHK0100	958..3219	958..358	3219..50
3537851	Tc00.1047053509817.50	AAHK0100	11792..9645	11792..12392	9645..78
3554702	Tc00.1047053508873.60	AAHK0100	13884..11146	13884..14484	11146..9
3549563	Tc00.1047053506139.120	AAHK0100	29283..32271	29983..29383	32271..3
3533904	Tc00.1047053506397.10	AAHK0100	4115..1869	4115..4715	1869..69

3532692	Tc00.1047053510881.9	AAHK0100	2165..1	2165..2765	1..1
3554758	Tc00.1047053510359.530	AAHK0100	106097..1085	106097..1054	108532..
3545651	Tc00.1047053506965.170	AAHK0100	32368..30089	32368..32968	30089..2
3549507	Tc00.1047053504769.100	AAHK0100	17742..20039	17742..17142	20039..2
3533975	Tc00.1047053470827.20	AAHK0100	6111..4843	6111..6711	4843..30
3533981	Tc00.1047053505699.10	AAHK0100	3770..732	3770..4370	732..1
3532652	Tc00.1047053510831.21	AAHK0100	2119..3713	2119..1519	3713..55
3548265	Tc00.1047053511843.70	AAHK0100	31722..34484	31722..31122	34484..3
3539146	Tc00.1047053506389.10	AAHK0100	1..1637	1..1	1637..34
3539195	Tc00.1047053509097.20	AAHK0100	5005..2666	5005..5605	2666..86
3539137	Tc00.1047053506713.80	AAHK0100	15105..16703	15105..14505	16703..1
3535286	Tc00.1047053506021.20	AAHK0100	4838..7222	4838..4238	7222..90
3542999	Tc00.1047053503587.30	AAHK0100	8450..11062	8450..7850	11062..1
3549403	Tc00.1047053511063.40	AAHK0100	9089..6579	9089..9689	6579..47
3539008	Tc00.1047053505609.40	AAHK0100	12235..14547	12235..11635	14547..1
3532514	Tc00.1047053511615.10	AAHK0100	3419..1671	3419..4019	1671..1
3545514	Tc00.1047053507653.20	AAHK0100	4853..7225	4853..4253	7225..90
3548164	Tc00.1047053506613.40	AAHK0100	10968..8101	10968..11568	8101..63
3546881	Tc00.1047053511311.30	AAHK0100	14197..12089	14197..14797	12089..1
3554665	Tc00.1047053511255.190	AAHK0100	40919..38334	40919..41519	38334..3
3546880	Tc00.1047053511311.20	AAHK0100	10216..8048	10216..10816	8048..62
3544241	Tc00.1047053508247.90	AAHK0100	15210..17657	15210..14610	17657..1
3533842	Tc00.1047053508697.10	AAHK0100	1423..4209	1423..823	4209..60
3533883	Tc00.1047053509771.10	AAHK0100	4226..1827	4226..4826	1827..27
3553397	Tc00.1047053510693.240	AAHK0100	58479..61349	58479..57879	61349..6
3537794	Tc00.1047053507773.10	AAHK0100	3182..420	3182..3782	420..1
3553399	Tc00.1047053510693.270	AAHK0100	73389..76175	73389..72789	76175..7
3552032	Tc00.1047053507167.80	AAHK0100	19486..17324	19486..20086	17324..1
3542950	Tc00.1047053509899.100	AAHK0100	24217..21578	24217..24817	21578..1
3536417	Tc00.1047053511197.10	AAHK0100	1..1862	1..1	1862..36
3539039	Tc00.1047053510411.10	AAHK0100	6765..4204	6765..7365	4204..24
3540382	Tc00.1047053507997.14	AAHK0100	1695..1	1695..2295	1..1
3536471	Tc00.1047053505931.30	AAHK0100	8694..6610	8694..9294	6610..48
3546806	Tc00.1047053510933.10	AAHK0100	3590..942	3590..4190	942..1
3539072	Tc00.1047053510979.40	AAHK0100	8553..6007	8553..9153	6007..42
3536441	Tc00.1047053508853.20	AAHK0100	8553..6007	8553..9153	6007..42
3539029	Tc00.1047053507255.20	AAHK0100	3223..5583	3223..2623	5583..73
3533860	Tc00.1047053508103.30	AAHK0100	5541..2947	5541..6141	2947..11
3552054	Tc00.1047053507237.10	AAHK0100	2441..165	2441..3041	165..1
3552000	Tc00.1047053510625.40	AAHK0100	22773..25487	22773..22173	25487..2
3536437	Tc00.1047053508455.20	AAHK0100	7383..4723	7383..7983	4723..41
3537743	Tc00.1047053507979.30	AAHK0100	7156..4652	7156..7756	4652..28
3550760	Tc00.1047053510361.80	AAHK0100	14052..16832	14052..13452	16832..1
3546871	Tc00.1047053506131.170	AAHK0100	33431..30843	33431..34031	30843..2
3540395	Tc00.1047053507355.50	AAHK0100	6247..9378	6247..5647	9378..11
3531276	Tc00.1047053455171.9	AAHK0100	1..1233	1..1	1233..30
3546839	Tc00.1047053511089.80	AAHK0100	23290..20606	23290..23890	20606..1
3554697	Tc00.1047053508873.330	AAHK0100	68298..65581	68298..68898	65581..6
3546886	Tc00.1047053511311.7	AAHK0100	2039..1	2039..2639	1..1
3544254	Tc00.1047053510847.10	AAHK0100	4917..1900	4917..5517	1900..10
3541674	Tc00.1047053509075.50	AAHK0100	14662..16869	14662..14062	16869..1
3544294	Tc00.1047053508121.150	AAHK0100	23518..26187	23518..22918	26187..2
3539005	Tc00.1047053505609.10	AAHK0100	2336..4603	2336..1736	4603..64
3546809	Tc00.1047053510933.80	AAHK0100	15222..12538	15222..15822	12538..1
3546822	Tc00.1047053508293.60	AAHK0100	13142..16042	13142..12542	16042..1
3535162	Tc00.1047053416041.14	AAHK0100	1776..1	1776..2376	1..1
3550736	Tc00.1047053511797.110	AAHK0100	13706..11100	13706..14306	11100..9
3537715	Tc00.1047053506723.20	AAHK0100	5015..2445	5015..5615	2445..64
3540375	Tc00.1047053507997.30	AAHK0100	6276..3742	6276..6876	3742..19
3541601	Tc00.1047053510553.40	AAHK0100	11306..14173	11306..10706	14173..1
3546851	Tc00.1047053506595.40	AAHK0100	11196..13181	11196..10596	13181..1
3549415	Tc00.1047053508495.100	AAHK0100	23082..25460	23082..22482	25460..2
3540307	Tc00.1047053511219.40	AAHK0100	7252..9642	7252..6652	9642..11
3550756	Tc00.1047053510361.10	AAHK0100	1011..3620	1011..411	3620..54
3554650	Tc00.1047053511255.700	AAHK0100	132493..1349	132493..1318	134997..
3542973	Tc00.1047053510005.20	AAHK0100	9925..7868	9925..10525	7868..60
3554616	Tc00.1047053511255.10	AAHK0100	329..2968	329..1	2968..47
3554644	Tc00.1047053511255.350	AAHK0100	74609..71970	74609..75209	71970..7
3536369	Tc00.1047053507313.20	AAHK0100	4595..7732	4595..3995	7732..95
3535052	Tc00.1047053503429.39	AAHK0100	6671..7878	6671..6071	7878..96
3537672	Tc00.1047053508581.10	AAHK0100	4123..1628	4123..4723	1628..1
3550646	Tc00.1047053506961.25	AAHK0100	6847..3713	6847..7447	3713..19
3536317	Tc00.1047053503667.10	AAHK0100	1..1904	1..1	1904..37
3537653	Tc00.1047053509495.30	AAHK0100	6340..3518	6340..6940	3518..17
3540205	Tc00.1047053506171.60	AAHK0100	12933..14957	12933..12333	14957..1
3537674	Tc00.1047053508581.40	AAHK0100	12909..10807	12909..13509	10807..9
3533732	Tc00.1047053506003.39	AAHK0100	3845..5683	3845..3245	5683..74
3536314	Tc00.1047053509417.9	AAHK0100	1958..1	1958..2558	1..1
3533728	Tc00.1047053508935.10	AAHK0100	1271..4003	1271..671	4003..58
3537655	Tc00.1047053507753.10	AAHK0100	5660..8788	5660..5060	8788..10
3550657	Tc00.1047053507069.160	AAHK0100	36451..39324	36451..35851	39324..4
3537682	Tc00.1047053510111.20	AAHK0100	7735..5258	7735..8335	5258..34
3541528	Tc00.1047053507633.80	AAHK0100	11377..14190	11377..10777	14190..1
3550642	Tc00.1047053506961.150	AAHK0100	54055..50876	54055..54655	50876..4

3537677	Tc00.1047053506609.30	AAHK0100	4828..2396	4828..5428	2396..59
3544138	Tc00.1047053507233.10	AAHK0100	723..3821	723..123	723..56
3537640	Tc00.1047053507213.40	AAHK0100	5100..7283	5100..4500	7283..90
3535028	Tc00.1047053511919.20	AAHK0100	2225..5020	2225..1625	5020..68
3533730	Tc00.1047053507497.11	AAHK0100	639..1	639..1239	1..1
3535059	Tc00.1047053507125.10	AAHK0100	4334..1968	4334..4934	1968..16
3532437	Tc00.1047053509737.9	AAHK0100	1020..1	1020..1620	1..1
3546766	Tc00.1047053508523.30	AAHK0100	13854..11377	13854..14454	11377..9
3549328	Tc00.1047053509699.40	AAHK0100	8961..6247	8961..9561	6247..44
3550650	Tc00.1047053507069.40	AAHK0100	7852..4991	7852..8452	4991..31
3535090	Tc00.1047053506157.40	AAHK0100	7991..6747	7991..8591	6747..49
3548074	Tc00.1047053511185.80	AAHK0100	14331..11776	14331..15361	11776..9
3549352	Tc00.1047053503861.40	AAHK0100	14761..11648	14761..15361	11648..9
3540216	Tc00.1047053510239.80	AAHK0100	14126..16456	14126..13526	16456..1
3541505	Tc00.1047053511129.40	AAHK0100	8410..5288	8410..9010	5288..34
3548012	Tc00.1047053506967.60	AAHK0100	17009..19912	17009..16409	19912..2
3546768	Tc00.1047053508523.50	AAHK0100	34879..33737	34879..35479	33737..3
3551950	Tc00.1047053504427.230	AAHK0100	47221..50214	47221..46621	50214..5
3540241	Tc00.1047053507163.70	AAHK0100	14790..12028	14790..15390	12028..1
3531176	Tc00.1047053423707.10	AAHK0101	203..1132	203..1	1132..29
3533715	Tc00.1047053413293.9	AAHK0100	927..1	927..1527	1..1
3537668	Tc00.1047053511057.40	AAHK0100	5759..8092	5759..5159	8092..98
3548008	Tc00.1047053506967.10	AAHK0100	2359..20	2359..2959	20..1
3550528	Tc00.1047053508761.110	AAHK0100	20260..23046	20260..19660	23046..2
3554466	Tc00.1047053506885.210	AAHK0100	54353..51402	54353..54953	51402..4
3544022	Tc00.1047053508677.40	AAHK0100	10520..7662	10520..11120	7662..58
3533672	Tc00.1047053510093.20	AAHK0100	4790..1644	4790..5390	1644..1
3534929	Tc00.1047053504343.10	AAHK0100	4054..1652	4054..4654	1652..1
3550523	Tc00.1047053508761.40	AAHK0100	5819..8596	5819..5219	8596..10
3551812	Tc00.1047053506683.190	AAHK0100	56075..53814	56075..56675	53814..5
3544071	Tc00.1047053511643.110	AAHK0100	26758..23948	26758..27358	23948..2
3541431	Tc00.1047053505919.20	AAHK0100	5362..7548	5362..4762	7548..93
3534916	Tc00.1047053504479.10	AAHK0100	3147..535	3147..3747	535..1
3553199	Tc00.1047053508165.14	AAHK0100	1964..1	1964..2564	1..1
3553160	Tc00.1047053510483.360	AAHK0100	82895..85435	82895..82295	85435..8
3542717	Tc00.1047053509765.129	AAHK0100	23734..25968	23734..23134	25968..2
3544044	Tc00.1047053511451.80	AAHK0100	22052..19278	22052..22652	19278..1
3542708	Tc00.1047053506615.150	AAHK0100	24388..22058	24388..24988	22058..2
3553139	Tc00.1047053510483.60	AAHK0100	12282..9511	12282..12882	9511..77
3550574	Tc00.1047053506133.190	AAHK0100	54570..51787	54570..55170	51787..4
3533616	Tc00.1047053506353.10	AAHK0100	1171..3468	1171..571	3468..52
3534947	Tc00.1047053508069.5	AAHK0100	1016..1	1016..1616	1..1
3553135	Tc00.1047053508139.240	AAHK0100	81234..83543	81234..80634	83543..8
3538848	Tc00.1047053510635.10	AAHK0100	2041..4536	2041..1441	4536..63
3551814	Tc00.1047053506683.240	AAHK0100	67367..65154	67367..67967	65154..6
3544067	Tc00.1047053511643.40	AAHK0100	8741..6102	8741..9341	6102..43
3537570	Tc00.1047053510061.30	AAHK0100	6195..3577	6195..6795	3577..17
3554481	Tc00.1047053506885.480	AAHK0100	114832..1177	114832..1142	117708..
3532387	Tc00.1047053508603.10	AAHK0100	3141..982	3141..3741	982..1
3550571	Tc00.1047053506133.130	AAHK0100	30969..33735	30969..30369	33735..3
3551809	Tc00.1047053506683.110	AAHK0100	30208..27068	30208..30808	27068..2
3542715	Tc00.1047053509765.50	AAHK0100	6434..9565	6434..5834	9565..11
3533648	Tc00.1047053504491.20	AAHK0100	5540..3963	5540..6140	3963..21
3536228	Tc00.1047053511101.11	AAHK0100	546..1	546..1146	1..1
3544092	Tc00.1047053509785.50	AAHK0100	7980..10424	7980..7380	10424..1
3533637	Tc00.1047053510843.19	AAHK0100	4710..5513	4710..4110	5513..73
3553128	Tc00.1047053508139.130	AAHK0100	31203..33377	31203..30603	33377..3
3550595	Tc00.1047053510465.100	AAHK0100	35322..32611	35322..35922	32611..3
3544069	Tc00.1047053511643.70	AAHK0100	15647..12885	15647..16247	12885..1
3540101	Tc00.1047053510551.50	AAHK0100	19175..16812	19175..19775	16812..1
3553136	Tc00.1047053508139.250	AAHK0100	85495..87804	85495..84895	87804..8
3533596	Tc00.1047053510981.20	AAHK0100	5439..3198	5439..6039	3198..13
3553018	Tc00.1047053511585.230	AAHK0100	63901..66090	63901..63301	66090..6
3534853	Tc00.1047053505779.20	AAHK0100	7454..5259	7454..8054	5259..34
3542676	Tc00.1047053507059.50	AAHK0100	18107..15327	18107..18707	15327..1
3553056	Tc00.1047053506459.230	AAHK0100	59520..57010	59520..60120	57010..5
3549148	Tc00.1047053511487.90	AAHK0100	26800..29139	26800..26200	29139..3
3543983	Tc00.1047053505365.60	AAHK0100	12356..15481	12356..11756	15481..1
3551761	Tc00.1047053506813.80	AAHK0100	18310..17710	20976..22776	20976..2
3554329	Tc00.1047053504081.540	AAHK0100	116235..1185	116235..1156	118598..
3546543	Tc00.1047053508607.50	AAHK0100	27092..29656	27092..26492	29656..3
3546520	Tc00.1047053511599.80	AAHK0100	26495..28978	26495..25895	28978..3
3534852	Tc00.1047053505779.10	AAHK0100	3133..824	3133..3733	824..1
3534828	Tc00.1047053509281.20	AAHK0100	5046..2860	5046..5646	2860..10
3546565	Tc00.1047053506471.120	AAHK0100	35249..38065	35249..34649	38065..3
3541380	Tc00.1047053504313.10	AAHK0100	2100..4964	2100..1500	4964..67
3534832	Tc00.1047053503601.10	AAHK0100	934..3315	934..334	3315..51
3533547	Tc00.1047053510569.10	AAHK0100	1..1377	1..1	1377..31
3536177	Tc00.1047053420293.20	AAHK0100	1392..3578	1392..792	3578..53
3546559	Tc00.1047053506471.30	AAHK0100	5189..7468	5189..4589	7468..92
3542689	Tc00.1047053511757.70	AAHK0100	15930..18098	15930..15330	18098..1
3533532	Tc00.1047053463279.20	AAHK0100	5333..3093	5333..5933	3093..12
3551769	Tc00.1047053506813.220	AAHK0100	61338..63923	61338..60738	63923..6
3541315	Tc00.1047053508285.60	AAHK0100	17377..19752	17377..16777	19752..2

3546563	Tc00.1047053506471.100	AAHK0100	23193..25694	23193..22593	25694..2
3541364	Tc00.1047053509251.20	AAHK0100	3186..5810	3186..2586	5810..76
3554357	Tc00.1047053508163.250	AAHK0100	79076..76203	79076..79676	76203..7
3534867	Tc00.1047053506427.10	AAHK0100	1532..4468	1532..932	4468..62
3532224	Tc00.1047053510985.10	AAHK0100	484..1	484..1084	1..1
3553052	Tc00.1047053506459.180	AAHK0100	39544..37220	39544..40144	37220..3
3542674	Tc00.1047053507059.20	AAHK0100	6446..3639	6446..7046	3639..18
3536181	Tc00.1047053510787.10	AAHK0100	1060..4674	1060..460	4674..64
3550414	Tc00.1047053506409.170	AAHK0100	28874..26073	28874..29474	26073..2
3551764	Tc00.1047053506813.140	AAHK0100	31541..34018	31541..30941	34018..3
3550401	Tc00.1047053506331.130	AAHK0100	31736..29490	31736..32336	29490..2
3533530	Tc00.1047053398477.10	AAHK0100	4494..2131	4494..5094	2131..33
3550496	Tc00.1047053509437.160	AAHK0100	53910..51142	53910..54510	51142..4
3551767	Tc00.1047053506813.190	AAHK0100	54284..57145	54284..53684	57145..5
3541367	Tc00.1047053509251.80	AAHK0100	14012..16618	14012..13412	16618..1
3554277	Tc00.1047053511173.280	AAHK0100	71231..73954	71231..70631	73954..7
3554286	Tc00.1047053511173.80	AAHK0100	21187..23991	21187..20587	23991..2
3543812	Tc00.1047053509529.40	AAHK0100	11508..8902	11508..12108	8902..71
3534714	Tc00.1047053507427.10	AAHK0100	137..2623	137..1	2623..44
3541272	Tc00.1047053503609.50	AAHK0100	10603..7772	10603..11203	7772..59
3550399	Tc00.1047053506331.90	AAHK0100	20008..17702	20008..20608	17702..1
3550358	Tc00.1047053510021.120	AAHK0100	27052..24629	27052..27652	24629..2
3542504	Tc00.1047053506765.40	AAHK0100	5136..7628	5136..4536	7628..94
3533496	Tc00.1047053507411.10	AAHK0100	3099..313	3099..3699	313..1
3554267	Tc00.1047053511173.440	AAHK0100	100970..1032	100970..1003	103276..
3533477	Tc00.1047053506515.29	AAHK0100	4462..5278	4462..3862	5278..70
3532102	Tc00.1047053507051.11	AAHK0100	954..1	954..1554	1..1
3554280	Tc00.1047053511173.370	AAHK0100	88417..91239	88417..87817	91239..9
3537357	Tc00.1047053510817.50	AAHK0100	9267..7063	9267..9867	7063..52
3552935	Tc00.1047053506245.60	AAHK0100	11536..9554	11536..12136	9554..77
3543867	Tc00.1047053509663.50	AAHK0100	18834..21962	18834..18234	21962..2
3552947	Tc00.1047053506245.250	AAHK0100	66700..69594	66700..66100	69594..7
3534775	Tc00.1047053503441.5	AAHK0100	1057..1	1057..1657	1..1
3545190	Tc00.1047053503955.40	AAHK0100	8841..6688	8841..9441	6688..48
3533423	Tc00.1047053505155.4	AAHK0100	770..1	770..1370	1..1
3545108	Tc00.1047053507821.130	AAHK0100	26043..28514	26043..25443	28514..3
3532108	Tc00.1047053510919.9	AAHK0100	1878..1	1878..2478	1..1
3534781	Tc00.1047053506253.39	AAHK0100	5725..7385	5725..5125	7385..91
3551617	Tc00.1047053510431.300	AAHK0100	60793..62241	60793..60193	62241..6
3550328	Tc00.1047053508501.320	AAHK0100	57345..56830	57345..57945	56830..5
3533439	Tc00.1047053505317.9	AAHK0100	1669..1	1669..2269	1..1
3533424	Tc00.1047053504593.10	AAHK0100	5220..2990	5220..5820	2990..11
3552942	Tc00.1047053506245.190	AAHK0100	54199..51578	54199..54799	51578..4
3550367	Tc00.1047053510021.220	AAHK0100	53351..55876	53351..52751	55876..5
3545158	Tc00.1047053509527.120	AAHK0100	18934..21708	18934..18334	21708..2
3545104	Tc00.1047053507821.10	AAHK0100	175..3012	175..1	3012..48
3554295	Tc00.1047053504081.80	AAHK0100	23949..21394	23949..24549	21394..1
3536050	Tc00.1047053508485.10	AAHK0100	5471..3207	5471..6071	3207..14
3532103	Tc00.1047053509531.10	AAHK0100	798..1	798..1398	1..1
3554269	Tc00.1047053511173.470	AAHK0100	108788..1115	108788..1081	111586..
3547707	Tc00.1047053508219.20	AAHK0100	1919..4585	1919..1319	4585..63
3533453	Tc00.1047053509333.10	AAHK0100	4315..1295	4315..4915	1295..1
3546419	Tc00.1047053508061.20	AAHK0100	2944..5434	2944..2344	5434..72
3543829	Tc00.1047053503827.100	AAHK0100	22858..25731	22858..22258	25731..2
3554257	Tc00.1047053511173.130	AAHK0100	34026..36830	34026..33426	36830..3
3533493	Tc00.1047053507879.10	AAHK0100	4559..2178	4559..5159	2178..37
3537349	Tc00.1047053506911.30	AAHK0100	7462..9867	7462..6862	9867..11
3550363	Tc00.1047053510021.180	AAHK0100	42020..39504	42020..42620	39504..3
3543847	Tc00.1047053508521.20	AAHK0100	7224..4108	7224..7824	4108..23
3546424	Tc00.1047053508061.154	AAHK0100	36201..38032	36201..35601	38032..3
3543853	Tc00.1047053508521.100	AAHK0100	25893..22543	25893..26493	22543..2
3547782	Tc00.1047053507035.11	AAHK0100	1651..1	1651..2251	1..1
3532008	Tc00.1047053510171.10	AAHK0100	953..2171	953..353	2171..39
3551532	Tc00.1047053510363.250	AAHK0100	47801..45324	47801..48401	45324..4
3533360	Tc00.1047053509225.11	AAHK0100	1..276	1..1	276..207
3534620	Tc00.1047053508099.50	AAHK0100	7117..6680	7117..7717	6680..48
3543797	Tc00.1047053507835.80	AAHK0100	29254..28613	29254..29854	28613..2
3532000	Tc00.1047053508021.10	AAHK0100	1..203	1..1	203..200
3541152	Tc00.1047053510731.10	AAHK0100	1..371	1..1	371..217
3542441	Tc00.1047053510971.30	AAHK0100	17921..17019	17921..18521	17019..1
3543747	Tc00.1047053506751.50	AAHK0100	11664..9472	11664..12264	9472..76
3534655	Tc00.1047053504341.20	AAHK0100	7161..6625	7161..7761	6625..48
3533345	Tc00.1047053510157.5	AAHK0100	344..1	344..944	1..1
3532083	Tc00.1047053511223.20	AAHK0100	2330..1905	2330..2930	1905..10
3543754	Tc00.1047053505949.30	AAHK0100	4884..7271	4884..4284	7271..90
3531924	Tc00.1047053507133.20	AAHK0100	1683..1953	1683..1083	1953..37
3537101	Tc00.1047053506413.89	AAHK0100	9861..11965	9861..9261	11965..1
3537131	Tc00.1047053511779.10	AAHK0100	1..707	1..1	707..250
3550124	Tc00.1047053509419.131	AAHK0100	31144..31869	31144..30544	31869..3
3533208	Tc00.1047053511087.19	AAHK0100	4113..4876	4113..3513	4876..66
3541027	Tc00.1047053507523.10	AAHK0100	1..2542	1..1	2542..43
3533257	Tc00.1047053421173.14	AAHK0100	4978..4376	4978..5578	4376..25
3531925	Tc00.1047053507133.10	AAHK0100	1..335	1..1	335..213
3544876	Tc00.1047053511625.130	AAHK0100	28879..26183	28879..29479	26183..2

3539662	Tc00.1047053508589.5	AAHK0100	569..1	569..1169	1..1
3534489	Tc00.1047053507759.10	AAHK0100	1716..4844	1716..1116	4844..66
3548756	Tc00.1047053507953.10	AAHK0100	835..2097	835..235	2097..38
3531888	Tc00.1047053432997.10	AAHK0100	1..1828	1..1	1828..36
3535767	Tc00.1047053509387.5	AAHK0100	1..395	1..1	395..219
3540974	Tc00.1047053511221.60	AAHK0100	13573..16341	13573..12973	16341..1
3537035	Tc00.1047053506899.10	AAHK0100	1..1112	1..1	1112..29
3542251	Tc00.1047053510047.80	AAHK0100	16202..13524	16202..16802	13524..1
3552187	Tc00.1047053459061.10	AAHK0100	1914..1125	1914..2514	1125..1
3543454	Tc00.1047053508977.40	AAHK0100	9883..7109	9883..10483	7109..53
3535627	Tc00.1047053508883.48	AAHK0100	7224..8967	7224..6624	8967..10
3544771	Tc00.1047053506335.130	AAHK0100	28721..26481	28721..29321	26481..2
3533036	Tc00.1047053509763.10	AAHK0100	1..1133	1..1	1133..29
3538220	Tc00.1047053508043.10	AAHK0100	1..847	1..1	847..264
3531742	Tc00.1047053510567.10	AAHK0100	1209..1726	1209..609	1726..35
3535622	Tc00.1047053508157.50	AAHK0100	8959..7732	8959..9559	7732..59
3531728	Tc00.1047053409139.9	AAHK0100	1702..631	1702..2302	631..1
3539583	Tc00.1047053507937.3	AAHK0100	1..356	1..1	356..215
3547392	Tc00.1047053510055.10	AAHK0100	1..967	1..1	967..276
3534384	Tc00.1047053506653.10	AAHK0100	1..803	1..1	803..260
3539559	Tc00.1047053510205.10	AAHK0100	1..485	1..1	485..228
3548697	Tc00.1047053511911.10	AAHK0100	2867..5077	2867..2267	5077..68
3552442	Tc00.1047053506789.20	AAHK0100	1615..4446	1615..1015	4446..62
3539491	Tc00.1047053505975.20	AAHK0100	6842..2502	6842..7442	2502..70
3531699	Tc00.1047053404431.10	AAHK0100	1..1077	1..1	1077..28
3535582	Tc00.1047053509377.20	AAHK0100	4938..2584	4938..5538	2584..78
3534260	Tc00.1047053410199.30	AAHK0100	4290..4745	4290..3690	4745..65
3535545	Tc00.1047053506231.10	AAHK0100	1..985	1..1	985..278
3531614	Tc00.1047053509481.10	AAHK0100	1580..457	1580..2180	457..1
3540724	Tc00.1047053506451.20	AAHK0100	9187..11601	9187..8587	11601..1
3531623	Tc00.1047053483183.10	AAHK0100	1..1593	1..1	1593..33
3531621	Tc00.1047053402919.10	AAHK0100	1..599	1..1	599..239
3536880	Tc00.1047053508753.3	AAHK0100	1..542	1..1	542..234
3555036	Tc00.1047053508221.1050	AAHK0100	234085..2327	234085..2346	232743..
3531680	Tc00.1047053508639.10	AAHK0100	1656..817	1656..2256	817..1
3531630	Tc00.1047053408445.10	AAHK0100	521..1	521..1121	1..1
3532972	Tc00.1047053509841.10	AAHK0100	1..696	1..1	696..249
3532995	Tc00.1047053507995.20	AAHK0100	4417..3509	4417..5017	3509..17
3536822	Tc00.1047053509031.20	AAHK0100	9920..11418	9920..9320	11418..1
3531609	Tc00.1047053435665.10	AAHK0100	1..171	1..1	171..197
3534217	Tc00.1047053508449.31	AAHK0100	6480..5838	6480..7080	5838..40
3531679	Tc00.1047053401569.10	AAHK0100	1..1173	1..1	1173..29
3551194	Tc00.1047053507875.220	AAHK0100	59321..61693	59321..58721	61693..6
3532999	Tc00.1047053508053.10	AAHK0100	4420..3669	4420..5020	3669..18
3532811	Tc00.1047053510209.9	AAHK0100	1..427	1..1	427..222
3538018	Tc00.1047053508305.9	AAHK0100	2045..1	2045..2645	1..1
3540694	Tc00.1047053511855.10	AAHK0100	1..1316	1..1	1316..31
3531586	Tc00.1047053508007.10	AAHK0100	1..985	1..1	985..278
3538043	Tc00.1047053418185.70	AAHK0100	12295..12912	12295..11695	12912..1
3551002	Tc00.1047053508433.70	AAHK0100	21197..21535	21197..20597	21535..2
3532838	Tc00.1047053506377.10	AAHK0100	1..586	1..1	586..238
3543243	Tc00.1047053508047.11	AAHK0100	525..1	525..1125	1..1
3532822	Tc00.1047053506665.30	AAHK0100	3247..4079	3247..2647	4079..58
3531546	Tc00.1047053508855.10	AAHK0100	1..561	1..1	561..236
3538098	Tc00.1047053508229.10	AAHK0100	1..1040	1..1	1040..28
3534124	Tc00.1047053507561.10	AAHK0100	1..342	1..1	342..214
3539340	Tc00.1047053506597.9	AAHK0100	351..1	351..951	1..1
3531544	Tc00.1047053492827.10	AAHK0100	1230..1481	1230..630	1481..32
3531534	Tc00.1047053448329.10	AAHK0100	1..276	1..1	276..207
3539296	Tc00.1047053506887.10	AAHK0100	1..587	1..1	587..238
3543113	Tc00.1047053507867.5	AAHK0100	681..1	681..1281	1..1
3547018	Tc00.1047053509549.140	AAHK0100	40233..38468	40233..40833	38468..3
3537928	Tc00.1047053510489.50	AAHK0100	13825..13016	13825..14425	13016..1
3531450	Tc00.1047053510707.10	AAHK0100	2..214	2..1	214..201
3547020	Tc00.1047053509549.10	AAHK0100	2932..5331	2932..2332	5331..71
3534045	Tc00.1047053503993.10	AAHK0100	408..2483	408..1	2483..42
3544483	Tc00.1047053508633.20	AAHK0100	6553..4166	6553..7153	4166..23
3531453	Tc00.1047053480561.10	AAHK0100	1..417	1..1	417..221
3531443	Tc00.1047053433559.10	AAHK0100	953..1381	953..353	1381..31
3534000	Tc00.1047053509743.40	AAHK0100	6151..5682	6151..6751	5682..38
3550809	Tc00.1047053510013.50	AAHK0100	8058..6889	8058..8658	6889..50
3531328	Tc00.1047053433733.10	AAHK0100	1..1097	1..1	1097..28
3548296	Tc00.1047053509815.5	AAHK0100	255..1	255..855	1..1
3546966	Tc00.1047053508857.15	AAHK0100	703..1	703..1303	1..1
3546967	Tc00.1047053508857.30	AAHK0100	7661..5505	7661..8261	5505..37
3543026	Tc00.1047053506281.7	AAHK0100	276..1	276..876	1..1
3533939	Tc00.1047053503713.50	AAHK0100	5567..6021	5567..4967	6021..78
3536543	Tc00.1047053508481.10	AAHK0100	6374..3636	6374..6974	3636..18
3549508	Tc00.1047053504769.140	AAHK0100	25353..26459	25353..24753	26459..2
3531353	Tc00.1047053452475.10	AAHK0100	704..18	704..1304	18..1
3537849	Tc00.1047053510805.5	AAHK0100	1..377	1..1	377..217
3540485	Tc00.1047053507907.10	AAHK0100	1..1095	1..1	1095..28
3532653	Tc00.1047053418405.30	AAHK0100	3715..2948	3715..4315	2948..11
3542954	Tc00.1047053511851.5	AAHK0100	1..442	1..1	442..224

3536423	Tc00.1047053509329.31	AAHK0100	10473..10108	10473..11073	10108..8
3532566	Tc00.1047053503659.30	AAHK0100	3574..2783	3574..4174	2783..98
3537713	Tc00.1047053511595.59	AAHK0100	12411..13209	12411..11811	13209..1
3539041	Tc00.1047053510411.40	AAHK0100	16434..14709	16434..17034	14709..1
3537799	Tc00.1047053504201.30	AAHK0100	6135..3451	6135..6735	3451..16
3541638	Tc00.1047053511349.100	AAHK0100	21051..18511	21051..21651	18511..1
3531246	Tc00.1047053460061.20	AAHK0100	858..1207	858..258	1207..30
3541676	Tc00.1047053509075.11	AAHK0100	1..300	1..1	300..210
3535159	Tc00.1047053506453.10	AAHK0100	1..1083	1..1	1083..28
3531238	Tc00.1047053483813.10	AAHK0100	682..1198	682..82	1198..29
3539069	Tc00.1047053510979.5	AAHK0100	394..1	394..994	1..1
3536400	Tc00.1047053511839.40	AAHK0100	6499..9090	6499..5899	9090..10
3533816	Tc00.1047053507861.24	AAHK0100	5826..5347	5826..6426	5347..35
3532596	Tc00.1047053508101.9	AAHK0100	1971..1	1971..2571	1..1
3533861	Tc00.1047053508103.14	AAHK0100	360..1	360..960	1..1
3535069	Tc00.1047053506071.19	AAHK0100	7614..7479	7914..8514	7479..56
3532477	Tc00.1047053510929.10	AAHK0100	1..1024	1..1	1024..28
3532404	Tc00.1047053506059.10	AAHK0100	3203..2560	3203..3803	2560..76
3531192	Tc00.1047053509271.10	AAHK0100	1..635	1..1	635..243
3538928	Tc00.1047053511117.30	AAHK0100	9859..7598	9859..10459	7598..57
3550622	Tc00.1047053507747.180	AAHK0100	36334..39006	36334..35734	39006..4
3538919	Tc00.1047053511677.20	AAHK0100	16137..15760	16137..16737	15760..1
3538988	Tc00.1047053510115.70	AAHK0100	16284..15441	16284..16884	15441..1
3531124	Tc00.1047053473709.10	AAHK0101	142..1079	142..1	1079..28
3535058	Tc00.1047053506065.21	AAHK0100	7594..7899	7594..6994	7899..96
3532425	Tc00.1047053510097.10	AAHK0100	304..714	304..1	714..251
3538986	Tc00.1047053510115.30	AAHK0100	8115..5269	8115..8715	5269..34
3538947	Tc00.1047053510235.3	AAHK0100	1..304	1..1	304..210
3531169	Tc00.1047053426675.9	AAHK0101	1..827	1..1	827..262
3536267	Tc00.1047053510501.40	AAHK0100	10074..9250.	10074..10674	9250..74
3538805	Tc00.1047053508055.30	AAHK0100	11646..9286	11646..12246	9286..74
3532393	Tc00.1047053510379.10	AAHK0100	1..451	1..1	451..225
3541422	Tc00.1047053509081.166	AAHK0100	21817..22415	21817..21217	22415..2
3537514	Tc00.1047053508089.10	AAHK0100	2459..306	2459..3059	306..1
3532309	Tc00.1047053511567.11	AAHK0100	232..1	232..832	1..1
3533655	Tc00.1047053507987.10	AAHK0100	1..1111	1..1	1111..29
3553198	Tc00.1047053508165.430	AAHK0100	89205..88056	89205..89805	88056..8
3538801	Tc00.1047053511547.20	AAHK0100	3091..5829	3091..2491	5829..76
3538898	Tc00.1047053506895.80	AAHK0100	15969..14980	15969..16569	14980..1
3538815	Tc00.1047053509665.30	AAHK0100	8638..11178	8638..8038	11178..1
3546605	Tc00.1047053503957.10	AAHK0100	1..777	1..1	777..257
3536266	Tc00.1047053506217.40	AAHK0100	6585..9728	6585..5985	9728..11
3537521	Tc00.1047053506341.50	AAHK0100	7233..10358	7233..6633	10358..1
3532201	Tc00.1047053508343.10	AAHK0100	2659..1634	2659..3259	1634..1
3547820	Tc00.1047053507699.230	AAHK0100	43229..41937	43229..43829	41937..4
3532226	Tc00.1047053508115.20	AAHK0100	2739..2465	2739..3339	2465..66
3546557	Tc00.1047053506843.5	AAHK0100	330..1	330..930	1..1
3540086	Tc00.1047053506191.58	AAHK0100	19099..18878	19099..19699	18878..1
3534878	Tc00.1047053511597.50	AAHK0100	6796..7562	6796..6196	7562..93
3543947	Tc00.1047053510535.109	AAHK0100	29256..29711	29156..28556	29711..3
3532227	Tc00.1047053511059.10	AAHK0100	1..424	1..1	424..222
3538701	Tc00.1047053506449.10	AAHK0100	26..541	26..1	541..234
3554369	Tc00.1047053508163.234	AAHK0100	72119..71268	72119..72719	71268..6
3534876	Tc00.1047053511597.10	AAHK0100	1..1018	1..1	1018..28
3550474	Tc00.1047053508173.120	AAHK0100	27602..26460	27602..28202	26460..2
3549163	Tc00.1047053507949.210	AAHK0100	38919..36667	38919..39519	36667..3
3534864	Tc00.1047053508563.20	AAHK0100	5955..3421	5955..6555	3421..16
3543952	Tc00.1047053503717.10	AAHK0100	5243..2508	5243..5843	2508..70
3538754	Tc00.1047053508401.10	AAHK0100	1..484	1..1	484..108
3542681	Tc00.1047053504099.50	AAHK0100	18609..16432	18609..19209	16432..1
3537484	Tc00.1047053509843.20	AAHK0100	2829..5447	2829..2229	5447..72
3543959	Tc00.1047053503717.120	AAHK0100	21374..23788	21374..20774	23788..2
	Tc00.1047053507731.60	AAHK0100	9948..8521	9948..10548	8521..67
3542690	Tc00.1047053511757.100	AAHK0100	24104..25824	24104..23504	25824..2
3546517	Tc00.1047053511599.10	AAHK0100	1..1163	1..1	1163..29
3534837	Tc00.1047053510833.10	AAHK0100	1..696	1..1	696..249
	Tc00.1047053508887.5	AAHK0100	1..325	1..1	325..212
3533452	Tc00.1047053511835.10	AAHK0100	628..3681	628..28	3681..54
3554283	Tc00.1047053511173.10	AAHK0100	1..1035	1..1	1035..28
3547790	Tc00.1047053511667.30	AAHK0100	11679..11341	11679..12279	11341..9
3546404	Tc00.1047053511767.10	AAHK0100	1..720	1..1	720..252
3532154	Tc00.1047053508849.11	AAHK0100	420..1	420..1020	1..1
3543861	Tc00.1047053509663.69	AAHK0100	29285..29495	29285..28685	29495..3
3549081	Tc00.1047053510943.206	AAHK0100	49833..49562	49833..50433	49562..4
3539994	Tc00.1047053511593.91	AAHK0100	18531..18948	18531..17931	18948..2
3537319	Tc00.1047053504639.60	AAHK0100	12342..11161	12342..12942	11161..9

De: abrandao@fiocruz.br Assunto: enquiry article submission Data: Qua, Abril 22, 2009
11:36 Para: tig@elsevier.com

Dear Dr. Treasa Creavin
Editor Trends in Genetics

This letter is to enquiry on the possibility of submitting an analysis article dealing with the question of Untranslated Regions definition in trypanosomatid genes. Following the initial enthusiasm with the conclusion of genome projects for the species of medical importance (Trypanosoma cruzi, T. brucei and Leishmania major), a daunting task remains to be addressed: to associate the genome codes to the pathogenic capabilities of these parasites. In this article we discuss an approach to uncover the trans-splicing and poly-A sites (the boundaries of UTRs) in duplicated genes in trypanosomatid genomes. Please find in the attached file an overview of the proposed article.

Sincerely yours,

Adeilton Brandão, D. Sc.
Researcher
Instituto Oswaldo Cruz-Fiocruz
abrandao@fiocruz.br

The conservation of upstream and downstream sequences of duplicated genes in trypanosomatid genomes: an appraisal of the usefulness to define UTRs .

Tainah Silva¹ and Adeilton Brandão²

1-Programa de Pós-graduação em Ciências Médicas

Universidade do Estado do Rio de Janeiro-RJ

2- Instituto Oswaldo Cruz

Fiocruz- Rio de Janeiro, Brazil

Introduction

The Trypanosomatidae family include protozoa parasites like *Trypanosoma cruzi*, *Leishmania braziliensis* and *T. brucei rhodesiense*, causing respectively Chagas Disease, leishmaniasis and African trypanosomiasis (sleeping sickness). They present several features that distinguish them from the rest of eukaryotes. Among these are mitochondrial DNA with two different circular molecules; b) modification of mitochondrial transcripts by RNA editing; c) Polycistronic transcription, trans-splicing and post-transcriptional regulation of genes; d) presence of repeated elements (VIPER, SINE) in the Untranslated Regions of several genes; e) absence of introns in most of the genes; f) presence of glycosomes (Simpson et al., 2000; Campbell et al, 2003; Moyersoan et al., 2004; Palenchar & Bellofatto, 2006; Vazquez et al., 2000). Besides these features, it is also possible to categorize genes in trypanosomatids by their genomic organization: single copy genes, tandem arrays, multicopy, oligo copies and two – three copies in different chromosomes.

The genome sequence of *T. brucei*, *T. cruzi* and *L. major* (El-Sayed et al., 2005) has been completed and three other projects are underway: *L. braziliensis*, *T. congolense*, *T. vivax* (www.genedb.org). Apart from the challenges associated with the analysis of this huge amount of information, i. e. translating it to the biological context of these protozoa parasites, there are specific questions from the genome information view that should be addressed. For example, the annotation of the transcribed segments other than the recognized open reading frames (ORF), specifically the Untranslated Regions (UTR) that appear in mRNA (Brandão, 2006). These UTRs flank the ORF and several experiments demonstrated that they contain structured and non-structured sequence elements involved in the control of mRNA stability, localization and translation efficiency (Mignone *et al.*, 2002; McNicoll et al., 2005).

The problem of annotating the UTRs and other non-coding functional segments is particularly important in trypanosomatid genes, since a similar signal sequence like the one used to predict poly-A sites (the end of 3' UTR) in other eukaryotes has not been identified so far . Only experimental methods such as primer extension, 5' and 3' RACE and complete cDNA sequencing allow the accurate definition of the transcribed segments in trypanosomatids. Algorithms using different approaches to define the untranslated regions have been proposed but they suffer from low accuracy (Gopal et al., 2003). While an efficient computational tool for description of trypanosomatid UTRs is waiting to come, alternative approaches may be useful to address this question in genome analysis. We are proposing here one such approach for generation of information on the recognition of the transcription points in non-tandem duplicated trypanosomatids genes. Considering that the coding sequences and

their flanking segments (which includes the UTRs) remains conserved or appear with minor variations when present in more than one copy in the genome, it is possible to determine with accuracy the beginning and termination of transcription in trypanosomatids simply comparing the copies among themselves. The DNA segments that are common to the copies outside the open reading frame should define the 5' and 3' UTR. In addition, other functional segments close to the UTRs tend also to be conserved, and this can be a good starting point to map those segments that do not appear in the mRNA but are functionally active. We have tested this approach in selected repeated genes with the same ORF length from trypanosomatids species whose genomes are either complete or under sequencing (*T. cruzi*, *T. brucei*, *L. major*, *T. congolense*, *T. vivax*, *L. braziliensis*). The UTRs predicted by these approach were confirmed by comparing the upstream with cDNA sequences that have the spliced leader RNA (mini-exon) attached to 5' end.

Basic methodology

The steps for this sequence analysis include:

- a) choosing gene sequences that have an equal ORF length and more than one copy in a given genome
- b) extracting from the database a sequence fragment that contains, the entire ORF, an upstream sequence of length at least 40% the ORF length and a downstream fragment of length equal to 50%. These percent numbers are based on observations that the maximum length the 5' and 3' UTR reach in trypanosomatid transcripts is respectively, 40% and 60% of ORF length (Adeilton Brandao, unpublished).
- c) aligning the sequences (by using an algorithm like ClustalW) and taking the initiator codon as starting point to locate the position where the upstream sequences cease to be identical in each copy analyzed. The end of identity in upstream sequences indicates the start of 5' UTR. Continuing the inspection upstream of this point, a dinucleotide AG would appear as the last common identical bases. Inspecting the bases upstream of this dinucleotide AG should reveal the tract of pyrimidines (C and T) in variable amount, pointing to trans-splicing signal. This motif is not conserved as the UTR in related copies of a gene.
- d) applying the same procedure for the 3' UTR, by starting downstream of the terminator codon.

For this analysis the gene sequence that appear in the genome may be selected according to the five basic classes of duplicated genes in trypanosomatids:

- 1) tandem repeated genes;
- 2) duplicated genes with short intergenic region;

- 3) duplicated genes with intergenic region containing another different gene;
- 4) genes duplicated in the same chromosome but separated by a large intergenic region, with or without intervening coding sequences;
- 5) duplicated genes in different chromosomes.

Sequence fragments upstream of the initiation codon and downstream of the termination codon were cut from the respective gene with a length proportional to the ORF length as defined above. The flanking sequences is aligned under default parameters by Clustal W. This served to determine the DNA segment other than the protein coding sequence that is common to all copies in each gene set. To find the length of the probable UTR the alignments are visually inspected and the positions of sequence identity are marked as reference for the locations of initiation and termination codons in the original annotation. Afterwards, each sequence in the alignment is compared to homologous cDNA sequence that exhibit the spliced leader RNA at the 5' end and poly-A tail at the 3' end.

Appraisal and concluding remarks

Although we have not developed a high throughput computational tool, the strategy devised here may be easily incorporated in the existing algorithms for gene analysis in the almost intronless genomes like the trypanosomes. The target are the genes that appear repeated at least twice in the genome and the rationale we sought to propose such an alternative is that genes that have been duplicated in different locations in the genome maintain the sequence features that are under functional pressures. That is the case for untranslated regions since they are transcribed into messenger RNA. If we align the flanking sequences of two or more copies of the same gene in one specific genome, mark the position where the bases do not change and compare the distance to either the initiation or the termination codon we can calculate the length of UTR and predict the transcription initiation and termination with high accuracy. Thus, the simple alignment of flanking sequences of two duplicated genes in trypanosomatids can tell us where their UTR locate. Nonetheless, application of the approach described here to the first three classes of duplicated genes did not produce the result we were expecting. This is because in most of the genes analyzed, the intergenic region did not vary among copies, even for those with intergenic region with other genes in-between. Probably these genes are polycistronic transcribed and since the UTRs and intergenic spacers are present in pre-messenger RNA, the alignment of these flanking segments will not distinguish the UTRs from the intergenic spacer. In this case, the UTRs can only be determined by the aid of sequences from ESTs or full cDNA.

But for the classes 4 and 5 the approach worked well. In the selected duplicated genes that matched the classes 4 and 5 UTRs were easily determined by aligning the ORF flanking sequences. Irrespective of gene length, a segment common to the copies in the respective class appears after the alignment. These common segments are supposed to be the UTRs and to confirm this we BLAST searched the dbEST-Genbank to find homologous ESTs bearing the mini-exon fragment and or poly-A tail that mark the beginning and the end of transcription. Although experimental findings established that a poly-pyrimidine rich segment precedes the trans-splicing signal, represented by the nucleotide AG, no consensus has been determined about the length and composition of this poly-pyrimidine region. It seems that each gene set (the gene and its copies) posses a unique tract of poly pyrimidine and even after the rearrangements that normally occur in the genome of these protozoa the overall structure of this CT rich tract remains constant. For example, the 2 copies of the putative 20s proteasome beta 7 subunit in *L. braziliensis* (gene id LbrM20.5430 and LbrM20.5400) shows conservation of one CT tract that probably contains the trans-splicing signal. The existing algorithms for analysis of UTRs in trypanosomatids may incorporate the approach discussed here, though the tandem repeated genes and the polycistronic transcription restrict the applicability of a computational tool to wide genome analysis. To cope with these unique features of trypanosomatids the best solution is the combination of distinct algorithms in a suite of computational tools that simultaneously analyze each parameter of the genome organization, structure and transcription. One example of a trypanosomatid feature that cannot be analyzed separately in a functional context is the coupling of trans splicing and poly-A addition. Both phenomena show specific characteristics from the point of nucleotide composition, organization and biochemical timing. However, any attempt to separately analyzing them in an algorithm is not a successful strategy by observation that they are simultaneous in certain circumstances, and the signals that are use in independent manner can be the same for a concerted action. Thus, upon annotation of genomes, they must be thoroughly inspected in order to evaluate the influence of untranslated regions and other non-coding RNAs.

References

- Camargo EP. Phytomonas and other trypanosomatid parasites of plants and fruit. Adv Parasitol. 1999;42:29-112.
- Podlipaev S. The more insect trypanosomatids under study-the more diverse Trypanosomatidae appears. Int J Parasitol. 2001, 31:648-52.

- Campbell DA, Thomas S, Sturm NR. Transcription in kinetoplastid protozoa: why be normal? *Microbes Infect.* 2003, 5:1231-40
- Vickerman K. The evolutionary expansion of the trypanosomatid flagellates. *Int J Parasitol.* 1994, 24:1317-31.
- Podlipaev SA. Insect trypanosomatids: the need to know more. *Mem Inst Oswaldo Cruz.* 2000, 95:517-22
- Momen H. Some current problems in the systematics of Trypanosomatids. *Int J Parasitol.* 2001, 31:640-2.
- Simpson L, Thiemann OH, Savill NJ, Alfonzo JD, Maslov DA. Evolution of RNA editing in trypanosome mitochondria. *Proc Natl Acad Sci U S A.* 2000, 97:6986-93.
- Brandao, A. The untranslated regions of genes from *Trypanosoma cruzi*: perspectives for functional characterization of strains and isolates. *Mem Inst Oswaldo Cruz.* 2006, 101:775-777
- McNicoll F, Muller M, Cloutier S, Boilard N, Rochette A, Dube M, Papadopoulou B Distinct 3'-untranslated region elements regulate stage-specific mRNA accumulation and translation in *Leishmania*. *J Biol Chem.* 2005, 280:35238-46
- Gopal S, Cross GA, Gaasterland T. An organism-specific method to rank predicted coding regions in *Trypanosoma brucei*. *Nucleic Acids Res.* 2003, 31:5877-85
- El-Sayed NM, Myler PJ, Berriman M, Crabtree J, Aggarwal G, Caler E, Renauld H, Wortley EA, Ghedin E, Peacock C, Bartholomeu DC, Haas BJ, Tran AN, Wortman JR, Alsmark UC, Angiuoli S, Anupama A, Badger J, Bringaud F, Cadag E, Carlton JM, Cerqueira GC, Creasy T, Delcher AL, Djikeng A, Embley TM, Hauser C, Ivens AC, Kummerfeld SK, Pereira-Leal JB, Nilsson D, Peterson J, Salzberg SL, Shallom J, Silva JC, Westenberger S, Melville SE, Andersson B, Hall N. Comparative genomics of trypanosomatid parasitic protozoa. *Science.* 2005, 309:404-9.
- Mignone F, Gissi C, Liuni S, Pesole G (2002) Untranslated regions of mRNAs. *Genome Biology* 3:reviews0004.1-0004.10
- Moyersoens J, Choe J, Fan E, Hol WG, Michels PA. Biogenesis of peroxisomes and glycosomes: trypanosomatid glycosome assembly is a promising new drug target. *FEMS Microbiol Rev.* 2004, 28:603-43.
- Palenchar JB, Bellofatto V. Gene transcription in trypanosomes.. *Mol Biochem Parasitol.* 2006, 146:135-41.

Vazquez M, Ben-Dov C, Lorenzi H, Moore T, Schijman A, Levin MJ..The short interspersed repetitive element of *Trypanosoma cruzi*, SIRE, is part of VIPER, an unusual retroelement related to long terminal repeat retrotransposons. Proc Natl Acad Sci U S A. 2000, 97:2128-33.