

PRISCILA CARNEIRO CAMPOS

Diferenças intra-específicas no sistema de reparo de erros de pareamento e na resposta ao estresse oxidativo em *Trypanosoma cruzi*: o papel da proteína TcMSH2

Orientadora: Santuza Ribeiro Teixeira

Co-orientador: Carlos Renato Machado

Departamento de Bioquímica e Imunologia
Instituto de Ciências Biológicas
Universidade Federal de Minas Gerais
2009



PRISCILA CARNEIRO CAMPOS

Diferenças intra-específicas no sistema de reparo de erros de pareamento e na resposta ao estresse oxidativo em *Trypanosoma cruzi*: o papel da proteína TcMSH2

Tese apresentada ao programa de pós-graduação em Bioquímica e Imunologia do Instituto de Ciências Biológicas da Universidade Federal de Minas Gerais como requisito parcial para obtenção do grau de Doutor em Bioquímica.


Orientadora: Santuza Maria Ribeiro Teixeira

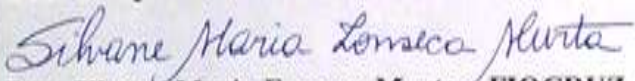
Co-orientador: Carlos Renato Machado

Departamento de Bioquímica e Imunologia
Instituto de Ciências Biológicas
Universidade Federal de Minas Gerais
Novembro de 2009



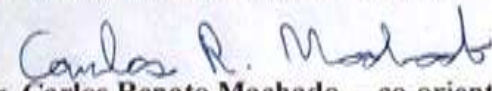
ATA DA DEFESA DA TESE DE DOUTORADO DE PRISCILA CARNEIRO CAMPOS. Aos seis dias do mês de novembro de 2009 às 13:00 horas, reuniu-se no Instituto de Ciências Biológicas da Universidade Federal de Minas Gerais, a Comissão Examinadora da tese de Doutorado, indicada *ad referendum* do Colegiado do Curso, para julgar, em exame final, o trabalho intitulado "Diferenças intra-específicas no sistema de reparo de erros de pareamento e na resposta ao estresse oxidativo em *Trypanosoma cruzi*: o papel da proteína TcMSH2", requisito final para a obtenção do grau de Doutor em Ciências: Biologia Molecular. Abrindo a sessão a Presidente da Comissão, Profa. Santuza Maria Ribeiro Teixeira da Universidade Federal de Minas Gerais, após dar a conhecer aos presentes o teor das Normas Regulamentares do Trabalho Final, passou a palavra a candidata para apresentação de seu trabalho. Seguiu-se a arguição pelos examinadores, com a respectiva defesa da candidata. Logo após a Comissão se reuniu, sem a presença da candidata e do público, para julgamento e expedição do resultado final. Foram atribuídas as seguintes indicações: Dra. Nadja Cristhina de Souza Pinto da Universidade de São Paulo, aprovada; Dra. Silvane Maria Fonseca Murta da Fundação Oswaldo Cruz, aprovada; Dra. Andrea Mara Macedo da Universidade Federal de Minas Gerais, aprovada; Dr. Dawidson Assis Gomes da Universidade Federal de Minas Gerais, aprovada; Dr. Carlos Renato Machado, co-orientador, da Universidade Federal de Minas Gerais, aprovada; Dra. Santuza Maria Ribeiro Teixeira, orientadora, da Universidade Federal de Minas Gerais, aprovada. Pelas indicações a candidata foi considerada APROVADA. O resultado final foi comunicado publicamente a candidata pela Presidente da Comissão. Nada mais havendo a tratar, a Presidente da Comissão encerrou a reunião e lavrou a presente Ata que será assinada por todos os membros participantes da Comissão Examinadora. Belo Horizonte, 6 de novembro de 2009.


Dra. Nadja Cristhina de Souza Pinto - USP


Dra. Silvane Maria Fonseca Murta - FIOCRUZ - MG


Dra. Andrea Mara Macedo - UFMG


Dr. Dawidson Assis Gomes - UFMG


Dr. Carlos Renato Machado - co-orientador - UFMG


Dra. Santuza Maria Ribeiro Teixeira - orientadora - UFMG


Profa. Andrea Mara Macedo
Coordenadora do Curso de Pós-Graduação em
Bioquímica e Imunologia
Instituto de Ciências Biológicas - UFMG

“Aprender é a única coisa de que a mente nunca se cansa, nunca tem medo e nunca se arrepende”.

(Leonardo Da Vinci)

Agradecimentos

Aos meus orientadores, pelo “pareamento perfeito”...

À Santuza Teixeira, por ter me recebido em seu laboratório, me orientado por tantos anos aguentando as minhas crises existenciais, por ter me dado a oportunidade de aprender tanto sobre o *Trypanosoma cruzi*, a biologia molecular e por ser um exemplo de paixão pela pesquisa e pelo conhecimento;

Ao Nem, pela sua co-orientação, pelo seu bom humor constante, seu modo tão empolgado de ver os resultados dos experimentos e sempre achar tudo “muuuuito interessante”;

À Patiu, por ter me permitido a utilização do seu laboratório, pelo seu jeito bem humorado e filosófico de ver a vida dentro da pesquisa e fora dela (e pelo sofá anti-stress);

Ao Álvaro Cantini, que pelos conselhos nos experimentos e por me dar a oportunidade de contribuir em seu novo projeto de pesquisa;

À Dani Bartholomeu, pela parceria nos *Processing Sites*, pelos conselhos e por todas as vezes que eu entrei em seu laboratório pedindo desesperadamente por um reagente;

Ao Richard McCulloch, Gustavo Cerqueira, Alice Machado-Silva, Sérgio Schenkman e Sheila Nardelli, pela colaboração nos artigos;

À Rosiane Silva, pela amizade mesmo na correria, pela ajuda nos experimentos, pelo exemplo de profissionalismo e por ter nos dado a alegria de levar o Eugênio ao Congresso;

À Silvane Murta, por me ensinar a clonar parasitas em meio semi-sólido e por me perdoar a falha de não tê-la avisado da minha defesa de mestrado;

Ao Fabiano, pela grande amizade que cultivamos um pelo outro mesmo sem contato diário, por compartilhar comigo, opiniões, filosofias, questionamentos e “planos A, B, e C”;

Ao Wanderson... por ser o Wanderson! Você me ajudou muito, tanto no laboratório quanto fora dele. Torço muito pelo seu sucesso porque você merece! Fico muito feliz de conseguir manter a nossa amizade mesmo a “milhares” de quilômetros de distância;

À Pati, pela nossa grande amizade, que foi mantida a todo custo, mesmo à distância, com Skype, MSN, e-mail, sinal de fumaça... Espero que ela seja duradoura e intercontinental;

Ao Ramon, que se tornou um grande amigo e que com certeza terá um futuro brilhante... Ninguém passa em primeiro lugar por acaso;

Aos meus grandes amigos do Laboratório, os quais eu gosto muito de conversar, almoçar, chorar as pitangas, rir, jogar truco, além de trabalhar junto, é claro: Gabriela (Gabirola), *Little John*, Natália (cabeçuda), Thiago (superlativo), Jarina (JahJah), Fernanda... Vocês estarão sempre no meu coração;

Aos meus amigos de laboratório Thiago Th2, Dani, Bárbara, Mariana, João, Luciana (a que tem uma isoforma)... A correria não nos permitiu ter tempo pra nos conhecermos melhor, mas podem contar comigo para o que precisarem;

À Vivis (*Mini-me*) e a Camila, por terem me dado a oportunidade de dizer que eu ensinei alguma coisa a alguém... É tão bonitinho ver que vocês já batem asas sozinhas! Tenho muito orgulho de vocês, Ca e Pequetita!

À Dani Marra, por manter a amizade comigo mesmo com os meus deslizes de memória, pela força interior e pelo exemplo de profissionalismo, inteligência e por me mostrar que é possível fazer doutorado, ter filho(sss) e marido ao mesmo tempo;

Ao Laboratório de Genética Bioquímica, pelos reagentes, equipamentos, e pelos grandes amigos que eu fiz por lá, em especial a Carol (minha parceira de MMR e lamúrias!!!), o Bruno (meu parceiro de FITC-avidina e lamúrias!!), a Michelle (minha parceira de lamúrias e lamúrias!!!) e o Matheus (“o cara” do confocal);

À Liza (Psychita!!!!), pela amizade tão bem humorada que cultivamos... Você faz muita falta por aqui;

Aos amigos do Laboratório de Bioenergética e Defesas Antioxidantes em Tripanossomatídeos da Unicamp: Fernanda Gadelha, que com sua simpatia deixou de ser uma simples colaboradora e virou uma grande amiga e aos meus amigos Thiago e Eduardo (homem de Neanderthal), que são figuraças;

Aos amigos que passaram pelo Laboratório, mas a gente não se esquece deles: Junia, Ricardo Trigueiro, Carol Leão...vocês têm de nos visitar;

Aos professores Ricardo Gazzinelli, Alfredo Góes, Carlos Chaves, Wellington Gomes, Sérgio Costa e a todos do departamento de Bioquímica e Imunologia do ICB, que de forma direta ou indireta contribuíram para a elaboração deste trabalho;

Ao Elimar, que esteve sempre presente no meu local de trabalho, exercendo os papéis de técnico, amigo, conselheiro... Você é muito importante pra mim;

À Celise e Grazielle, da secretaria, pela disposição em ajudar sempre;

Ao CNPq, FAPEMIG, Capes e Howard Hughes Medical Institute pelo apoio financeiro;

A todos que de alguma forma contribuíram para este trabalho e que eu não mencionei o nome por lapso de memória;

À minha família: minha irmã Camila, por sermos tão iguais e tão diferentes! Eu amo você mais que tudo neste mundo... À minha prima Zayda, que eu adoro de paixão e a quem devo muito... E aos meus pais, que sempre estiveram presentes na minha mente e no meu coração;

E finalmente, ao meu marido Cacá, a quem devo agradecer por me amar incondicionalmente, mesmo nos momentos mais turbulentos da minha vida... Você suportou os meus choros, angústias, assistiu às minhas vitórias, me deu apoio, esteve sempre ao meu lado. Muito obrigada pelo amor e pela paciência!

Lista de Figuras	7
Lista de Tabelas.....	9
Resumo.....	10
Abstract	12
Introdução	14
1. O parasito <i>Trypanosoma cruzi</i> e a doença de Chagas.....	14
2. Genoma e mecanismos de expressão gênica em <i>Trypanosoma cruzi</i>	17
3. Diversidade genética em <i>Trypanosoma cruzi</i>	20
4. O papel das vias de reparo de DNA na manutenção da integridade genômica e na geração de variabilidade genética em <i>Trypanosoma cruzi</i>	23
Objetivo geral.....	31
Objetivos específicos	32
Material e Métodos	33
1. Parasitos	33
2. Ensaios de viabilidade celular de parasitos após tratamento com agentes genotóxicos	33
3. Análise do acúmulo de 8-oxoguanina	34
4. Estudos do metabolismo oxidativo em cepas de <i>Trypanosoma cruzi</i>	34
a. Produção de NADPH	34
b. Medição do consumo de oxigênio e do controle respiratório	35
c. Produção de peróxido de hidrogênio.....	35
5. Amplificação gênica por PCR (reação em cadeia da polimerase) e construção de vetores	36
6. Análise genotípica através da técnica de RFLP (<i>Restriction Fragment Length Polymorphism</i>)	37
7. Cultivo de bactérias.....	39
8. Transformação de bactérias.....	39
a. Por choque térmico	39
b. Por eletroporação	39

9.	Identificação por PCR de clones bacterianos transformados	40
10.	Obtenção de DNA plasmidiano e digestão do DNA por endonucleases de restrição	40
11.	Produção de proteínas recombinantes e anticorpos policlonais	41
12.	Produção de extratos protéicos totais de epimastigotas de <i>Trypanosoma cruzi</i> e ensaios de <i>Western blot</i>	42
13.	Transfecção de formas epimastigotas de <i>Trypanosoma cruzi</i> por eletroporação	43
14.	Clonagem de parasitos transfectados.....	43
a.	Diluição limitante.....	43
b.	Meio semi-sólido.....	44
15.	Análise por PCR do DNA dos clones de <i>Trypanosoma cruzi</i> transfectados....	44
16.	<i>Southern blot</i>	45
17.	<i>Northern blot</i>	46
18.	Localização subcelular da proteína em fusão com GFP.....	47
19.	Imunofluorescência.....	48
20.	Análises <i>in silico</i> de seqüências de aminoácidos de TcMSH2.....	49
21.	Análises <i>in silico</i> de padrões de processamento do mRNA em <i>T. cruzi</i>	50
	Resultados	52
1.	Inferências filogenéticas a partir de seqüências parciais de aminoácidos para a proteína TcMSH2 nas diferentes cepas de <i>Trypanosoma cruzi</i>	52
3.	Sobrevivência de parasitos pertencentes a diferentes cepas de <i>T. cruzi</i> ao estresse induzido por peróxido de hidrogênio	57
4.	Expressão da proteína recombinante TcMSH2 e produção de anticorpos anti-TcMSH2.....	63
5.	Avaliação da expressão de TcMSH2 nas diferentes cepas de <i>T. cruzi</i>	64
6.	Geração de linhagens heminocautas para o gene <i>Tcmsh2</i> no clone CL Brener	68
7.	Sobrevivência de parasitos CL Brener heminocautas de <i>Tcmsh2</i> ao tratamento com agentes genotóxicos	72
8.	Sobrevivência de parasitos do tipo selvagem e heminocautas para o gene <i>Tcmsh2</i> à exposição prolongada com doses subletais de MNNG.....	75
9.	Localização subcelular da proteína TcMSH2 em <i>T. cruzi</i>	75
10.	Geração de parasitos superexpressores do gene <i>Tcmsh2</i>	80

11. Análises dos padrões de processamento do mRNA em <i>T. cruzi</i>	84
Discussão	90
1. O MMR e o papel da proteína TcMSH2.....	90
2. Padrões de processamento do mRNA em <i>Trypanosoma cruzi</i>	102
Perspectivas.....	107
Bibliografia	108
Anexos	119

Lista de Figuras

Figura 1: <i>Trypanosoma cruzi</i> e suas principais formas ao longo do ciclo de vida do parasito.	15
Figura 2: Representação esquemática do sistema de reparo de erros de pareamento em procariotos e eucariotos.....	27
Figura 3: Representação gráfica da metodologia para análise dos sítios de processamento do mRNA em <i>T. cruzi</i>	51
Figura 4: Construção de um cladograma a partir de sequências parciais de aminoácidos para a proteína TcMSH2 e análise das cepas por RFLP (<i>Restriction Fragment Length Polymorphism</i>).	53
Figura 5: Efeito da cisplatina e do MNNG no crescimento das cepas de <i>Trypanosoma cruzi</i>	56
Figura 6: Efeito do estresse oxidativo induzido por H ₂ O ₂ no crescimento das cepas de <i>Trypanosoma cruzi</i>	58
Figura 7: Análise comparativa do metabolismo oxidativo entre as cepas de <i>T. cruzi</i>	59
Figura 8: Ligação de avidina conjugada a FITC na presença de dGTP e d8-oxoGTP.	61
Figura 9: Análise do acúmulo de 8-oxoguanina em diferentes cepas de <i>T. cruzi</i>	62
Figura 10: Expressão da proteína recombinante TcMSH2::His, avaliação de sua solubilidade e geração de anticorpos policlonais anti-TcMSH2.....	65
Figura 11: Análise da expressão de TcMSH2 nas diferentes cepas de <i>T. cruzi</i>	67
Figura 12: Caracterização dos clones <i>Tcmsh2</i> ^{+/-} de <i>T. cruzi</i>	69
Figura 13: Caracterização dos clones duplo-resistentes e de clones heminocautes para o alelo <i>Tcmsh2</i> -C (NeoR) de <i>Trypanosoma cruzi</i>	71
Figura 14: Efeito do tratamento com cisplatina e MNNG no crescimento de clones de CL Brener <i>Tcmsh2</i> ^{+/-}	73
Figura 15: Análise da resposta de parasitos CL Brener tipo selvagem e clones <i>Tcmsh2</i> ^{+/-} ao estresse oxidativo induzido por peróxido de hidrogênio.	74
Figura 16: Análise do metabolismo oxidativo dos parasitos CL Brener do tipo selvagem e clones <i>Tcmsh2</i> ^{+/-} após exposição prolongada a uma dose sub-letal de MNNG.	76

Figura 17: Localização subcelular da proteína TcMSH2::GFP em CL Brener.	78
Figura 18: Localização subcelular da proteína TcMSH2 em epimastigotas de <i>T. cruzi</i> por imunofluorescência indireta.....	79
Figura 19: Caracterização dos clones superexpressores do gene <i>Tcmsh2</i> de <i>T. cruzi</i>	81
Figura 20: Análise da expressão de TcMSH2 em epimastigotas de CL Brener transfectados com o vetor controle (pROCK-Neo) e com o vetor pROCK-TcMSH2-Neo, por ensaio de Western blot.	83
Figura 21: Tamanhos das regiões 5' e 3'UTR (não traduzidas) e suas distâncias em relação aos sinais de processamento dos mRNAs de <i>Trypanosoma cruzi</i>	86
Figura 22: Frequências dos motivos de polipirimidinas (poliYs) e de cada um dos nucleotídeos, nas regiões adjacentes aos sítios de processamento do mRNA de <i>Trypanosoma cruzi</i>	87
Figura 23: Padrões de processamento do mRNA em <i>Trypanosoma cruzi</i>	88
Figura 24: Correlação entre as isoformas da proteína TcMSH2, a eficiência do MMR e a variabilidade genética nas diferentes linhagens de <i>T. cruzi</i>	101

Lista de Tabelas

Lista de Figuras	7
Lista de Tabelas.....	9
Resumo.....	10
Abstract	12
Introdução	14
Material e Métodos	33
Tabela 1: Descrição dos iniciadores utilizados neste trabalho.....	38
Resultados	52
Tabela 2: Comparação entre os padrões de processamento do mRNA entre <i>Trypanosoma cruzi</i> e <i>Trypanosoma brucei</i>	89
Discussão	90
Perspectivas.....	107
Bibliografia	108
Anexos	119

Trypanosoma cruzi, protozoário causador da doença de Chagas, possui uma população altamente heterogênea, composta de um *pool* de cepas e clones com características distintas, tais como morfologia, taxas de crescimento, virulência e sensibilidade a drogas. Esta diversidade pode ser um reflexo da variabilidade genética intra-específica que, como em outros organismos, é originada principalmente por mutações e rearranjos gênicos. Entretanto, os mecanismos que levam a esta diversidade em *T. cruzi* não estão completamente esclarecidos. Análises filogenéticas serviram de base para uma proposição de classificação do táxon *T. cruzi* em três linhagens principais, denominadas *T. cruzi* I, II e III. Como estudos relacionados a funções envolvendo metabolismo de DNA – principalmente reparo e recombinação – ainda são escassos, o presente trabalho teve como objetivo estudar o papel de uma destas vias de reparo de DNA, o sistema de reparo de erros de pareamento ou MMR, na geração de variabilidade genética em *T. cruzi*. Polimorfismos no gene *Tcmsh2* resultam na expressão de três isoformas da proteína TcMSH2 (um componente chave do MMR) características de cada linhagem. A comparação das taxas de crescimento de cepas pertencentes aos haplogrupos A (*T. cruzi* I) e C (*T. cruzi* II) frente à exposição a agentes genotóxicos, bem como a determinação dos níveis de dano no DNA induzido por estresse oxidativo, forneceram evidências de que cepas pertencentes à linhagem *T. cruzi* I apresentam uma capacidade de reparo mais eficiente. O cultivo de parasitos na presença de cádmio, um inibidor da via de MMR, juntamente com os resultados de ensaios *in vitro* com proteínas recombinantes indicam que diferenças observadas entre as cepas podem ser, em parte, devido às diferenças na atividade de TcMSH2.

A fim de investigar melhor o papel do MMR e, mais precisamente, da proteína TcMSH2 na geração de variabilidade genética em *T. cruzi*, foram geradas linhagens heminocautes para o gene *Tcmsh2*. Os resultados obtidos sugerem que, além do seu papel no MMR, a proteína TcMSH2 exerça uma função MMR-independente importante na resposta ao estresse oxidativo, o que impossibilitaria a deleção de ambos os alelos. Esta função parece ser conservada em um tripanossomatídeo relacionado, *T. brucei*. Como estratégia alternativa, vetores para transfecção de *T. cruzi* foram construídos de maneira a gerar parasitos

superexpressando o gene *Tcmsh2*. Entretanto, devido possivelmente a mecanismos que controlam a expressão do gene *Tcmsh2*, não foi possível obter linhagens transfectadas expressando altos níveis dessa proteína.

Com o objetivo adicional de estudar quais são os sinais importantes para a expressão de um gene exógeno em *T. cruzi*, de forma a aperfeiçoar vetores já utilizados, foram realizadas análises *in silico* onde procuramos investigar melhor os requerimentos necessários para o processamento correto, através de *trans*-splicing e poliadenilação, dos mRNAs no parasito. Essas análises mostraram que a distância mediana entre os motivos de polipirimidinas (sinais já conhecidos como importantes para o processamento do mRNA) e os sítios de *trans*-splicing e os de poliadenilação é de 18 e 40 nt, respectivamente. Os motivos de polipirimidinas apresentaram tamanhos medianos de 11 nt (acima do sítio de *trans*-splicing) e 12 nt (abaixo do sítio de poliadenilação). Os tamanhos médios de 5'UTR (35 nt) e 3'UTR (137 nt) obtidos corroboram resultados anteriores, que indicam que o genoma de *T. cruzi* é mais compacto quando comparado com outros tripanossomatídeos, tais como *T. brucei* e *Leishmania major*.

Trypanosoma cruzi, the protozoan parasite that causes Chagas' disease has a highly heterogeneous population, composed of a pool of strains with distinct characteristics such as morphology, growth rate, virulence, and sensitivity to drugs. This diversity could be a mirror of the intra-specific variability that, as in other organisms, is mainly maintained by mutations and genetic rearrangements. However, the mechanisms underlying this diversity in *T. cruzi* are not completely understood. Phylogenetic analyses support a division of the taxon into three *major* lineages, named *T. cruzi* I, II and III. Since studies related to functions involved in DNA metabolism – particularly DNA repair and recombination – are scarce, the purpose of this work was to investigate the role of one of these DNA repair pathways, the mismatch repair or MMR, in generating genetic variability in *T. cruzi*. Polymorphisms in the *Tcmsh2* gene result into three expressed isoforms of the TcMSH2 protein (the *major* component of MMR), which are correlated to the three lineages of the parasite. By comparing cell survival rates of strains belonging to haplogroups A (*T. cruzi* I) and C (*T. cruzi* II) after exposure to genotoxic agents and measuring the levels of oxidative stress-induced DNA damage, we found evidence that strains belonging to haplogroup A (*T. cruzi* I lineage) present a more efficient DNA repair capacity. Culturing parasites in the presence of cadmium, a MMR inhibitor, as well as *in vitro* assays with recombinant proteins indicate that the differences observed between strains may be, at least in part, due to differences in the activity of TcMSH2.

To further investigate the role of MMR and, more precisely, the TcMSH2 protein, in generating genetic variability in *T. cruzi*, we designed experiments to obtain *Tcmsh2* null mutants. The results indicate that, in addition to its role in MMR, TcMSH2 acts in the parasite response to oxidative stress through a MMR-independent function, and that null mutants are not viable. This function may be conserved in the related trypanosomatid parasite, *T. brucei*. As an alternative approach, we constructed vectors with the aim of generating *T. cruzi* parasites overexpressing the *Tcmsh2* gene. However, probably due to mechanisms controlling the expression of the *Tcmsh2* gene, it was not possible to obtain transfected cells expressing high levels of TcMSH2 protein.

Therefore, with the additional goal of studying gene expression in *T. cruzi* and to optimize expression vectors currently being used, we conducted *in silico* analyses to investigate the sequence requirements involved in *T. cruzi* mRNA processing. We determined that the median distances between polypyrimidine motifs and Spliced Leader addition sites as well as polyadenylation sites is 18 and 40 nt, respectively. The polypyrimidine motifs present median lengths of 11 nt (upstream the Spliced Leader addition site) and 12 nt (downstream the polyadenylation site). The median lengths of 5' UTR (35 nt) and 3' UTR (137 nt) obtained in our analyses corroborate previous results, which showed that *T. cruzi* genome is more compacted among the trypanosomatid parasites *T. brucei* and *Leishmania major*.

1. O parasito *Trypanosoma cruzi* e a doença de Chagas

Trypanosoma cruzi é um protozoário de grande importância médica, por ser o agente etiológico da doença de Chagas (ou tripanossomíase americana), uma moléstia endêmica crônica, debilitante e incapacitante, que se encontra distribuída por toda a América Latina e com uma estimativa de 50 mil mortes/ano e de 100 milhões de pessoas sob risco de contrair a doença (<http://who.int/tdr/diseases/Chagas/>). O parasito pertence à ordem Kinetoplastida, família Trypanosomatidae, caracterizada pela presença de um flagelo e de uma mitocôndria única na qual se encontra o cinetoplasto, uma organela especializada que contém o DNA mitocondrial (Brisse et al., 2000) (Figura 1A). Os membros desta família estão entre os parasitos mais bem sucedidos evolutivamente e com maior distribuição geográfica. Apesar de separados por milhões de anos de evolução, outros parasitos relacionados pertencem à mesma família, como por exemplo, o *T. brucei*, causador da tripanossomíase africana, que aflige a população e afeta a economia da África sub-saariana, e o gênero *Leishmania*, responsável por diversas formas de leishmaniose [revisado por (Barrett et al., 2003; Stevens et al., 2001)].

O ciclo evolutivo do *T. cruzi* é bastante complexo, envolvendo um hospedeiro vertebrado que abrange uma variedade de organismos (incluindo o homem), e um hospedeiro invertebrado constituído por insetos triatomíneos da família Reduviidae. Em condições naturais, o mecanismo de transmissão do *T. cruzi* ao hospedeiro vertebrado ocorre através da deposição de fezes do triatomíneo infectado sobre a pele e as mucosas do hospedeiro durante o repasto sanguíneo, onde freqüentemente pode ser observada uma reação inflamatória que varia conforme o local de infecção (Chagoma ou sinal de Romana). Entretanto, outros mecanismos alternativos de transmissão já foram observados, dentre eles: transfusão sanguínea, transmissão congênita, acidentes de laboratório, transplante de órgão ou por via oral [revisado por (Brener et al., 2000)].

O *T. cruzi* apresenta três formas morfológicas distintas classificadas de acordo com suas dimensões, morfologia e posição do cinetoplasto em relação ao núcleo (Figura 1B). Nos triatomíneos, os parasitos se multiplicam ao longo do trato

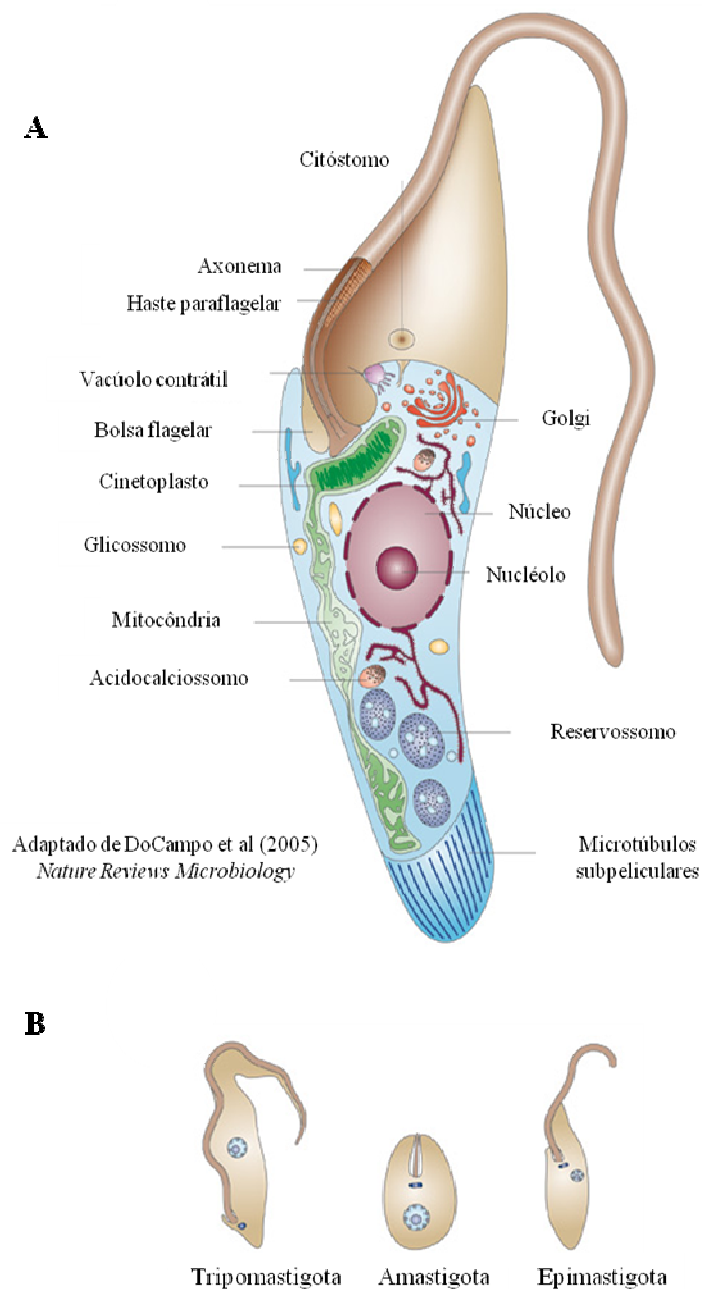


Figura 1: *Trypanosoma cruzi* e suas principais formas ao longo do ciclo de vida do parasito. Em **A**: As principais formas do parasito são definidas pelo formato do corpo, pela presença e adesão do flagelo e pela posição do corpo basal, do cinetoplasto e do núcleo. A forma epimastigota é encontrada no vetor e as formas amastigota e tripomastigota são encontradas no hospedeiro vertebrado, intra e extracelularmente, respectivamente. Em **B**: Representação esquemática de uma seção longitudinal da forma epimastigota de *T. cruzi*.

digestivo como epimastigotas, que são células longas, flageladas, com o cinetoplasto situado próximo ou anterior ao núcleo. Os epimastigotas se diferenciam para a forma tripomastigota metacíclica infectante quando chegam à ampola retal e são eliminados pelo vetor junto com as fezes durante o repasto sangüíneo. Esta forma é mais alongada que a epimastigota e apresenta o cinetoplasto localizado posteriormente ao núcleo, além de um flagelo que percorre ao longo de toda a extensão da célula. No interior das células do hospedeiro vertebrado, o parasito atinge a camada subjacente da pele ou da mucosa e invadem células locais, dando início à multiplicação intracelular sob a forma amastigota, caracterizada pela forma arredondada ou oval, com um resquício de flagelo. Estas formas se multiplicam por divisão binária até se diferenciarem na forma tripomastigota antes do rompimento da célula parasitada. Estes parasitas, através da circulação sanguínea vão para outros órgãos e tecidos do organismo onde reiniciam um novo ciclo intracelular. Ao picar o animal ou o homem infectado pelo *T. cruzi*, o triatomíneo ingere as formas tripomastigotas sangüíneas, reiniciando assim um novo ciclo evolutivo (Brener et al., 2000).

Uma das características da doença de Chagas é que, muito freqüentemente, as manifestações clínicas não existem ou passam despercebidas. Ainda sim, pode-se classificar a doença em duas fases distintas: a aguda e a crônica. Na fase aguda, a Doença de Chagas caracteriza-se pela alta parasitemia e pela inflamação mononuclear local derivada da ruptura de células parasitadas, que podem ser macrófagos, células musculares lisas e estriadas, fibroblastos e eventualmente neurônios (Rodrigues Coura and de Castro, 2002). Em um período de poucos dias ou até duas semanas, pode ser detectada no sangue a presença de complexos imunes, além de necrose inflamatória local. A inflamação severa é freqüentemente acompanhada por necrose de células parasitadas ou não parasitadas, especialmente no coração. Também têm sido observadas nesta fase a agregação de plaquetas, degranulação de eosinófilos, patologia microvascular, edema, trombose e isquemia (Teixeira et al., 2006). À medida que o sistema imune começa a produzir anticorpos contra o *T. cruzi*, o número de parasitos diminui e a fase crônica começa. A partir deste ponto, a doença pode progredir de uma forma assintomática ou evoluir para alterações cardíacas e do trato digestivo, onde o coração, cólon e esôfago se tornam notadamente aumentados e suas funções severamente comprometidas, podendo levar à morte ou incapacitação do indivíduo (Brener et al., 2000). A variabilidade genética

e biológica presente tanto táxon *T. cruzi* quanto nos seus hospedeiros, é um fator que pode influenciar no curso da infecção (Andrade et al., 1999). Assim sendo, as diferentes manifestações clínicas da doença de Chagas e suas taxas regionais de morbidade podem ser um reflexo da heterogeneidade das populações naturais de *T. cruzi* (Buscaglia and Di Noia, 2003; Devera et al., 2002).

Infelizmente, ainda não existe tratamento quimioterápico eficiente para os indivíduos chagásicos. E embora seja discutível a sua utilização para o tratamento da fase crônica da doença, duas drogas, o Nifurtimox e o Benzonidazol estão disponíveis para os casos diagnosticados durante a fase aguda. Ambas as drogas são altamente tóxicas para os pacientes, causando uma variedade de efeitos colaterais indesejáveis (Urbina and Docampo, 2003). Portanto, a procura de novas abordagens quimioterapêuticas para o tratamento da doença de Chagas permanece como assunto de extrema importância, considerando a prevalência e a mortalidade da doença. Além disso, uma vacina efetiva contra o *T. cruzi*, que seria outro modo de controle da endemia, parece ser uma possibilidade ainda remota, embora existam crescentes avanços na área (Rodrigues et al., 2009).

2. Genoma e mecanismos de expressão gênica em *Trypanosoma cruzi*

Em 2005 foi completado o sequenciamento do genoma de *T. cruzi*, tendo como referência o clone CL Brener (El-Sayed et al., 2005a). O tamanho do genoma diplóide do parasito foi estimado entre 106 e 110 Mb (El-Sayed et al., 2005a). Neste trabalho foi também observado que mais de 50% do genoma do parasito consistia de seqüências repetitivas, principalmente de moléculas de superfície, retrotransposons e repetições sub-teloméricas. As características do genoma de *T. cruzi*, por ser altamente repetitivo, tornaram difícil a montagem dos 4008 *contigs* e, mesmo com as ferramentas mais avançadas de bioinformática, não foi possível completar a montagem final do genoma, inviabilizando a determinação do número de cromossomos na espécie.

A caracterização do *T. cruzi* no nível citogenético é complexa, principalmente pelo fato de seus cromossomos não se condensarem durante a divisão celular. Entretanto, pela técnica de eletroforese de campo alternado, estimou-se o número de cromossomos em algumas cepas. Eles apresentam variação em relação ao tamanho e

às bandas detectadas nas diferentes cepas analisadas. No clone CL Brener, por exemplo, foram identificadas 55 bandas cromossômicas, variando de 550 Kb a 3,1 Mb (Branche et al., 2006). Uma estimativa recente, utilizando os dados gerados no projeto de sequenciamento do genoma do clone CL Brener e baseada em mapas de sintenia entre os tripanossomatídeos *T. cruzi*, *Leishmania major* e *T. brucei*, sugere que o genoma do parasito seria composto de 41 pares de cromossomos, com tamanhos que variam de 78 Kb a 2,4 Mb (Weatherly et al., 2009).

O sequenciamento completo do genoma de *T. cruzi* revelou também características interessantes do ponto de vista genético: o clone CL Brener de *T. cruzi* apresenta uma variabilidade alélica, estimada em 5,4 %, e as regiões codificadoras apresentam um maior grau de conservação (2,2 %). Por estas razões, tornou-se necessário o sequenciamento de ambos os haplótipos, cujas diferenças principais se devem a inserções/deleções em regiões intergênicas e subteloméricas e/ou ampliações de sequências repetitivas. Com base nestas análises, estima-se de que o genoma haplóide de CL Brener contenha 12000 genes (El-Sayed et al., 2005a).

O parasito *T. cruzi*, assim como outros membros da família Trypanosomatidae, divergiram muito cedo da linhagem principal que deu origem aos eucariotos ao longo do processo evolutivo (Briones et al., 1999; Junqueira et al., 2005). Provavelmente por esta razão, estes organismos apresentam muitas características bioquímicas e processos moleculares exclusivos que diferem bastante daqueles presentes em outros eucariotos, como por exemplo, a transcrição gênica em unidades policistrônicas e a ausência de introns no genoma [revisado por (Campbell et al., 2003)]. O mecanismo de transcrição gênica em *T. cruzi*, como em outros tripanossomatídeos, exibe características distintas dos eucariotos. Nos eucariotos, a transcrição de genes estruturais é feita, a partir de seus respectivos promotores, pela enzima RNA polimerase II como RNAs monocistrônicos que são posteriormente modificados por vários processos, no núcleo da célula, antes de serem transportados e traduzidos no citoplasma. Estas modificações incluem a remoção dos íntrons via *cis*-splicing, a adição de um resíduo de 7-metil guanósina (em inglês, *cap*) na extremidade 5' da molécula, a metilação dessa extremidade e de nucleotídeos internos e a poliadenilação na extremidade 3'. Entretanto, nos tripanossomatídeos, os mRNAs são transcritos em unidades policistrônicas que são submetidas aos processos acoplados de *trans*-splicing e poliadenilação, gerando mRNAs individuais

maduros. Para isto, uma seqüência nucleotídica líder de 39 nucleotídeos, conhecida como *spliced leader* (SL) ou miniéxon, é transferida para a extremidade 5' de cada RNA individual. Pelo fato de envolver dois "éxons" presentes em duas moléculas de RNA diferentes, esta reação foi denominada de *trans*-splicing. Esta seqüência líder, contendo a estrutura cap na extremidade 5', é doada para os pré-mRNAs, em um sítio acceptor AG localizado abaixo de um segmento de polipirimidinas. Após o processo, os mRNAs maduros adquirem um caráter descontínuo devido à ligação de dois pré-RNAs [revisito por (Teixeira and daRocha, 2003)]. Uma vez que o sítio acceptor de SL é reconhecido, é sugerido que o complexo responsável pela adição da cauda de poliA seja recrutado para a proximidade do sítio acceptor do SL e este começa a examinar aquela região intergênica na direção 3'-5' até identificar uma suposta seqüência de adição de cauda de poliA (Liang et al., 2003). Uma vez que a transcrição é policistrônica, a regulação da expressão gênica em tripanossomatídeos ocorre principalmente no nível pós-transcricional, incluindo o controle da etapa de maturação e estabilidade dos transcritos, vários estudos mostraram que esta regulação é dependente de elementos em *cis* localizados nas regiões 3' UTR (da Silva et al., 2006; Nozaki and Cross, 1995; Vanhamme and Pays, 1995). Além da maquinaria de *trans*-splicing, estudos indicaram a existência de quatro genes interrompidos por íntrons, o que sugere que o mecanismo de *cis*-splicing seja funcional em tripanossomatídeos (Ivens et al., 2005; Mair et al., 2000; Palfi et al., 2005).

O estudo do genoma de tripanossomatídeos teve um avanço considerável a partir da década de 80, quando os primeiros genes nestes organismos começaram a ser clonados e ter sua função inferida, tendo por base a sua homologia com genes de outros organismos. Estudos em larga escala de regiões codificadoras e da estrutura dos genes foram possíveis através do sequenciamento de ESTs (do inglês *expressed sequence tags*) (el-Sayed et al., 1995). Esta metodologia permitiu ao nosso grupo a caracterização de genes expressos nas formas amastigotas de *T. cruzi* (Cerqueira et al., 2005). Além disso, foram realizados estudos de localização de diversas proteínas e a sua expressão em sistemas heterólogos (Clayton, 1999). Grandes avanços nesta área foram obtidos com o desenvolvimento da transfecção gênica, que se iniciou com a expressão transiente de marcadores repórteres após a eletroporação de parasitos com vetores circulares (Bellofatto and Cross, 1989). Posteriormente, diversos

trabalhos foram publicados, utilizando metodologias distintas para a manipulação de genes em Tripanossomatídeos (Allaoui et al., 1999; Caler et al., 1998; Chung and Swindle, 1997; Cooper et al., 1993; DaRocha et al., 2004b; dos Santos and Buck, 2000; Hariharan et al., 1993; MacRae et al., 2006; Manning-Cela et al., 2001; Otsu et al., 1993; Vazquez and Levin, 1999). Em *T. cruzi*, os vetores podem ser mantidos no parasito na forma epissomal (extracromossômica) ou integrados no genoma através da recombinação homóloga. Qualquer que seja a metodologia de transfecção, para vários tripanossomatídeos é possível o estudo de importantes aspectos da sua biologia através da deleção, interrupção ou da superexpressão de genes de interesse ou a interferência de RNA (RNAi) (Fire et al., 1998; Ngo et al., 1998). Com as notáveis exceções de *T. cruzi*, *Leishmania donovani* e *L. major*, nos quais não foi obtida nenhuma evidência de um mecanismo funcional de RNAi (DaRocha et al., 2004a; Robinson and Beverley, 2003), em *T. brucei*, *T. congolense* e *L. braziliensis*, a interferência de RNA pode ser uma importante ferramenta nos estudos de genômica funcional destes parasitos (Peacock et al., 2007; Ullu et al., 2004).

3. Diversidade genética em *Trypanosoma cruzi*

É provável que o parasito *T. cruzi* seja um dos microorganismos patogênicos nos quais a diversidade genética intra-específica seja a mais bem conhecida (Tibayrenc, 2003). Os estudos bioquímicos e moleculares de *T. cruzi*, realizados principalmente a partir da década de 70, permitiram um crescente entendimento da grande variabilidade intra-específica no parasito. Os estudos pioneiros feitos por Miles e colaboradores (1978) forneceram os primeiros indícios da elevada diversidade genética de *T. cruzi*, relatando a existência de três tipos isoenzimáticos principais, denominados zimodemas. Estes zimodemas foram inicialmente tomados como unidades de análise para pesquisas epidemiológicas, mostrando que os zimodemas Z1 e Z3 estavam associados com o ciclo silvestre e Z2 com o ciclo doméstico (Miles et al., 1980; Miles et al., 1978). Dois anos mais tarde, foi demonstrado por RFLP (*Restriction Fragment Length Polymorphism*) de minicírculos de kDNA que havia uma surpreendente diversidade genética em *T. cruzi*. Este trabalho demonstrou pela primeira vez a existência de um número variado de clones dentro das cepas analisadas (Morel et al., 1980). Estudos subsequentes

usando 15 *loci* isoenzimáticos evidenciaram uma diversidade genética ainda maior, propondo a divisão do táxon em 43 genótipos (Tibayrenc et al., 1986). Posteriormente, o mesmo autor propôs a organização do táxon em 2 linhagens filogenéticas principais altamente heterogêneas, com diversas propriedades biológicas distintas (Tibayrenc, 1995). Esta dicotomia foi também evidenciada em trabalhos posteriores, que propuseram uma divisão em linhagens 1, 2 e 1/2, de acordo com marcadores de rDNA (Souto et al., 1996; Souto and Zingales, 1993). Em 1999, foi adotado o modelo de classificação que dividia o gênero *Trypanosoma cruzi* em linhagens I e II (Momen, 1999), sendo que algumas linhagens não puderam ser classificadas em nenhum dos dois grupos. Alguns trabalhos têm sugerido que *T. cruzi* I seja um grupo relativamente homogêneo, enquanto que *T. cruzi* II apresenta uma alta variabilidade, sendo posteriormente dividido em 5 clusters filogenéticos, IIa, IIb, IIc, IId e IIe (Brisse et al., 2000). Em seguida, duas novas propostas foram apresentadas para classificar o táxon. A primeira, baseada em dimorfismos apresentados pelo domínio D7 do gene de rDNA 24S α pelos diferentes isolados e dividindo o táxon em 4 haplogrupos (Kawashita et al., 2001); e a segunda, baseada em polimorfismos em seqüências codificadoras de dois genes mitocondriais (subunidade II da citocromo oxidase e subunidade 1 da NADH desidrogenase) e nucleares (dihidrofolato redutase-timidilato sintase e tripanotiona redutase) e dividindo em 3 (A, B, C) ou 4 (A, B, C, D) grupos (Machado and Ayala, 2002).

Em 2003, nosso grupo propôs uma classificação para o táxon *T. cruzi*, baseada em polimorfismos presentes em um gene da via de reparo de erros de pareamento (*Tcmsh2*), que divide o táxon *T. cruzi* em três haplogrupos, denominados A, B e C (Augusto-Pinto et al., 2003). Em 2006, estudos realizados com um número maior de marcadores moleculares, incluindo *loci* de microssatélites, polimorfismos observados em genes mitocondriais (NADH desidrogenase subunidade 1 e citocromo B) e ribossômicos (24 α S), corroboraram esta divisão (de Freitas et al., 2006). Mesmo após diversas propostas de nomenclatura para o táxon *T. cruzi*, desde 1999 até julho de 2009, a classificação em duas linhagens [*T. cruzi* I e II, (Momen, 1999)] vinha sendo adotada pela maioria da comunidade científica. Em agosto desse ano, durante a reunião anual da Sociedade Brasileira de Protozoologia, uma nova proposta de nomenclatura foi apresentada com o objetivo de unificar as nomenclaturas utilizadas nas pesquisas com o *T. cruzi* (Zingales et al., manuscrito em preparação). Nesta

reunião, chegou-se ao consenso de que o táxon *T. cruzi* é dividido em seis Unidades Discretas de Tipagem ou DTUs (*T. cruzi* I a VI), permitindo uma correlação direta à divisão proposta por Brisse e colaboradores em 2000. Assim, a linhagem *T. cruzi* I permanece inalterada, *T. cruzi* IIb equivale a *T. cruzi* II, *T. cruzi* IIc equivale a *T. cruzi* III, *T. cruzi* IIa equivale a *T. cruzi* IV, *T. cruzi* IId equivale a *T. cruzi* V e a linhagem *T. cruzi* IIe equivale à linhagem *T. cruzi* VI. É importante ressaltar que as diversas classificações do táxon *T. cruzi* têm sido propostas com o objetivo de tentar compreender melhor as relações filogenéticas entre as cepas e clones que compõem esta complexa população de *T. cruzi*. O tema ainda é objeto de discussão. O clone CL Brener, tomado como referência para o sequenciamento completo do *T. cruzi* (El-Sayed et al., 2005a) e atualmente classificado como *T. cruzi* VI, seria decorrente de um evento de hibridização entre as linhagens *T. cruzi* II e *T. cruzi* III (de Freitas et al., 2006). Outra hipótese se baseia na origem de CL Brener a partir de eventos de hibridização entre linhagens II e III, sendo que a linhagem parental III poderia ser originada de evento de hibridização entre linhagens *T. cruzi* I e *T. cruzi* II [revisto por (Cribb et al., 2004)].

Vários estudos avaliaram diferenças na variabilidade genética entre as linhagens de *T. cruzi*. Estudos filogenéticos envolvendo sequências codificadoras dos genes diidrofolato redutase-timidilato sintase e tripanotiona redutase em 31 cepas diferentes mostraram que todas as sequências obtidas a partir da linhagem *T. cruzi* I são monofiléticas, enquanto que as obtidas a partir da antiga linhagem *T. cruzi* II são parafiléticas e pertencem a 3 clados (Machado and Ayala, 2002). Análises de famílias multigênicas de *T. cruzi* mostraram que sequências parálogas codificadoras do gene amastina, uma proteína de superfície expressa predominantemente nas formas amastigotas do parasito, e o gene TcAG48, uma proteína antigênica de ligação ao RNA, apresentam uma maior variabilidade intragenômica nas cepas JG e Esmeraldo cl 3 (*T. cruzi* II) e no clone CL Brener (*T. cruzi* VI), quando comparadas às obtidas em Col1.7G2, Silvio X-10 cl1 e Dm28c (*T. cruzi* I) (Cerqueira et al., 2008). Da mesma forma, sequências codificadoras da proteína imunorregulatória Tc52 são homogêneas em cepas pertencentes à linhagem *T. cruzi* I, mas elas apresentam alto grau de divergência entre as cepas da linhagem *T. cruzi* II, IV e V (Mathieu-Daude et al., 2007).

Diversos trabalhos indicaram que as cepas pertencentes às linhagens *T. cruzi* II são predominantemente encontradas no ciclo doméstico e, conseqüentemente, mais associadas com maior frequência a infecções em humanos nos países endêmicos, como o Brasil e a Argentina (Freitas et al., 2005; Zingales et al., 1999), enquanto que as cepas pertencentes à linhagem *T. cruzi* I têm sido identificadas como agentes principais da doença de Chagas no norte da Bacia amazônica (Llewellyn et al., 2009). Os estudos filogenéticos podem fornecer informações de grande importância epidemiológica, pois nos permitem entender melhor a estrutura populacional de *T. cruzi*, bem como explorar a origem dos polimorfismos genéticos observados nas linhagens distribuídas entre o ciclo silvestre e o ciclo doméstico do parasito (Brisse et al., 1998).

4. O papel das vias de reparo de DNA na manutenção da integridade genômica e na geração de variabilidade genética em *Trypanosoma cruzi*

A habilidade de um organismo de se reproduzir está fortemente relacionada ao desafio de copiar fielmente o seu genoma, além de manter sua estabilidade após sucessivos ciclos de divisão celular. A manutenção da integridade genômica é crucial para o metabolismo do DNA e depende de um correto funcionamento das maquinarias de replicação e reparo existentes, além dos processos de recombinação (Buermeyer et al., 1999; de Boer, 2002). O genoma de uma célula é um dos principais alvos para danos induzidos por agentes endógenos e exógenos que, por sua vez, geram uma série de lesões na molécula, desde pequenos adutos e lesões de *cross-linking*, até as altamente nocivas quebras de fita dupla (ou DSBs). Diversos agentes genotóxicos são conhecidos, mas alguns têm recebido maior atenção, dentre eles, os geradores de espécies reativas de oxigênio (peróxido de hidrogênio e o azul de metileno, por exemplo), que são responsáveis por uma das lesões mais tóxicas sofridas pelo DNA, a 8-oxo-7,8-dihidro-2'-desoxiguanosina ou 8-oxoG (Aust and Eveleigh, 1999; DeFedericis et al., 2006; Shibutani et al., 1991; Slupphaug et al., 2003); o *N*-metil-*N'*-nitro-*N*-nitrosoguanidina ou MNNG, agente metilante altamente mutagênico e carcinogênico, capaz de desencadear uma interrupção no ciclo celular e/ou apoptose (Beardsley et al., 2005; Cejka and Jiricny, 2008); e o agente quimioterápico *cis*-diaminodicloroplatina (também conhecida como CDDP ou

cisplatina), que induz à formação de adutos de platina na posição N⁷ de purinas do DNA e, se não reparados, podem levar à ativação de vias apoptóticas (Stojic et al., 2004).

Considerando a complexidade da molécula de DNA, não é surpreendente que existam diferentes tipos de lesões com os mais variados graus de severidade. Em todos os casos, os danos gerados no DNA devem ser reparados. O reparo do DNA é um processo bioquimicamente complexo, o qual emprega inúmeras proteínas com funções distintas. Vários mecanismos de reparo celular envolvendo etapas multi-enzimáticas são capazes de remover rapidamente diferentes tipos de lesões, de acordo com a natureza do dano ou a sua extensão. Os complexos protéicos envolvidos no reparo são as chamadas “vias de reparo de DNA”, responsáveis pela manutenção do genoma em quaisquer condições fisiológicas. Em eucariotos, por exemplo, estão presentes 6 vias principais de reparo de DNA, dentre elas: o reparo por fusão de extremidades não-homólogas (NHEJ) e o reparo por recombinação homóloga (HRR), que lidam com quebras duplas no DNA (DSB); o sistema de reparo de erros de pareamento (MMR), o reparo por excisão de nucleotídeos (NER) e o reparo por excisão de bases (BER), cujos substratos são bases ou nucleotídeos aberrantes em uma das fitas do DNA ou alças de inserção/deleção (InDels); e o reparo por reversão da lesão (RL), que não requer molde para a síntese de DNA e tem como função principal a remoção de grupos metil anômalos ou dímeros de timina, sem causar quebra na dupla hélice. Estas vias operam de maneira coordenada de forma a minimizar o dano na informação genética; como consequência, a perda de um destes mecanismos pode levar em um acúmulo de mutações e/ou morte celular (de Boer, 2002; Gill and Fast, 2007).

Entretanto, a existência de numerosas vias de reparo de DNA não é suficiente para lidar com todos os danos sofridos pela molécula. Como consequência, ocorre um aumento nas taxas de mutação, deletérias em sua maioria, mas que, de modo mais sutil, podem favorecer a diversidade genética (de Boer, 2002; Elena and Lenski, 2003; Machado et al., 2006). Nesse sentido, as maquinarias de reparo de DNA e, em especial o MMR, são um componente crucial na manutenção deste equilíbrio (Machado et al., 2006), considerando que a inativação específica deste sistema de reparo pode elevar espontaneamente as taxas de mutação de 50 a 1000 vezes, além de aumentar as taxas de substituição de bases e de modificações de janelas de leitura

dos genes (Iyer et al., 2006). Mutações em linhagens germinativas dos genes da maquinaria de MMR (do inglês *methyl-directed mismatch repair*) estão associadas a uma susceptibilidade a um tipo de tumor denominado câncer colorretal humano não-polipóico hereditário (ou HNPCC, do inglês *hereditary nonpolyposis colorectal human cancer*) ou síndrome de Lynch (Akiyama et al., 1997; Kruger et al., 2002; Muller-Koch et al., 2001; Trojan et al., 2002). O HNPCC é uma síndrome autossômica dominante caracterizada pela sua precocidade (os pacientes possuem em média < 45 anos) e pela ocorrência de lesões neoplásicas em diversos tecidos, como a pele, o tecido endometrial, ovariano, gástrico e renal (Hsieh and Yamane, 2008). Estudos sobre reparo de DNA em bactérias, fungos e células de mamíferos mostraram que a perda do MMR resulta em um fenótipo mutante, no qual se encontram elevadas as taxa de substituições de base e mutações de mudança de janela de leitura, bem como uma elevada frequência de isolados resistentes a drogas (Branch et al., 1993; Clark et al., 1999; Hsieh and Yamane, 2008; Mena et al., 2008).

Para visualizar melhor o papel do MMR na manutenção do genoma, é necessário entender o funcionamento da via. O MMR em procariotos compreende 10 fatores (Lahue et al., 1989), que podem ser subdivididos em duas classes: os envolvidos exclusivamente no reparo pós-replicativo de erros de pareamento (proteínas MutS, MutL e MutH) e os que participam também de outras vias de metabolismo de DNA (MutU ou UvrD, DNA helicase II, Ssb, RecJ, exonucleases, DNA polimerases e ligases). O homodímero MutS se liga ao substrato [sendo essa ligação de alta (pareamentos G:T) ou baixa (pareamentos C:C) afinidade], recruta o homodímero MutL e a proteína MutH para o sítio de reparo e, de modo ATP-dependente, forma um grampo (também chamado de α -loop) usando a seqüência que contém o erro (Allen et al., 1997). A proteína MutH cliva a fita de DNA não metilada, a DNA helicase desfaz o dúplice e a exonuclease degrada o fragmento contendo o erro. Esta ação gera uma região de fita simples no DNA (estabilizada pelas proteínas Ssb), que é preenchida pela DNA polimerase, selada pela DNA ligase e metilada pela *Dam* metilase, tornando o sítio refratário a futuras ações do sistema MutSLH (Figura 2). Em eucariotos, as funções da via de MMR são cumpridas por heterodímeros de proteínas MSH, denominadas MSH1 a MSH7, além de outros componentes, tais como MLH1, MLH3, PMS1 e PMS2. Em *Saccharomyces cerevisiae*, o complexo MSH2/MSH6 (ou MutS α) reconhece malpareamentos e

pequenas alças de inserção ou deleção (InDels) de uma ou duas bases, enquanto que o complexo MSH2/MSH3 (ou MutS β), reconhece uma gama mais diversa de InDels. Além de erros de pareamento e InDels são conhecidos outros substratos do MMR, como O⁶-metilguanina, 8-oxoG, adutos carcinogênicos, fotoprodutos derivados da radiação UV e adutos de cisplatina (Cejka and Jiricny, 2008; Feng et al., 1991; Li et al., 1996; Stojic et al., 2004; Yamada et al., 1997).

O papel do MMR também foi descrito na manutenção da estabilidade do DNA mitocondrial (mtDNA) em leveduras, desempenhado pela proteína MSH1 de *S. cerevisiae* (Pogorzala et al., 2009; Reenan and Kolodner, 1992a, b). Dado o potencial para o dano oxidativo no mtDNA, outras vias de reparo de DNA também estão presentes na mitocôndria. Em leveduras, por exemplo, foi descrita uma via de BER (reparo por excisão de bases) que participa do reparo ao dano oxidativo. A uracil DNA glicosilase (Ung1p), oxoguanina DNA glicosilase (Ogg1p) e a pirimidina glicosilase (Ntg1p) têm sido identificadas em leveduras [revisado por (Stuart and Brown, 2006)].

Outra característica interessante do MMR diz respeito à diversidade de seus componentes protéicos entre os organismos. Em *Bacillus anthracis* e *Helicobacter pylori*, por exemplo, foi descrito um gene codificador da proteína MutS, embora em *H. pylori*, a mutS pertença a uma subfamília de proteínas com função não relacionada ao MMR (Eisen, 1998; Read et al., 2003; Wang et al., 2005). Em *Drosophila melanogaster* e *Caenorhabditis elegans* não foi encontrado gene homólogo a MSH3, sugerindo que a remoção de InDels nestes organismos seja MSH3-independente (Marti et al., 2002). Por outro lado, *Schizosaccharomyces pombe* apresenta MSH3, porém esta proteína parece desempenhar outra função dentro da célula, como em processos de recombinação (Tornier et al., 2001). Em *Plasmodium falciparum*, somente cinco genes codificadores de proteínas do MMR foram identificados: *msh6*, *mlh1*, *pms1* e, surpreendentemente, duas variantes da proteína MSH2, com genes denominados *Pfmsh2-1* e *Pfmsh2-2* (Bethke et al., 2007).

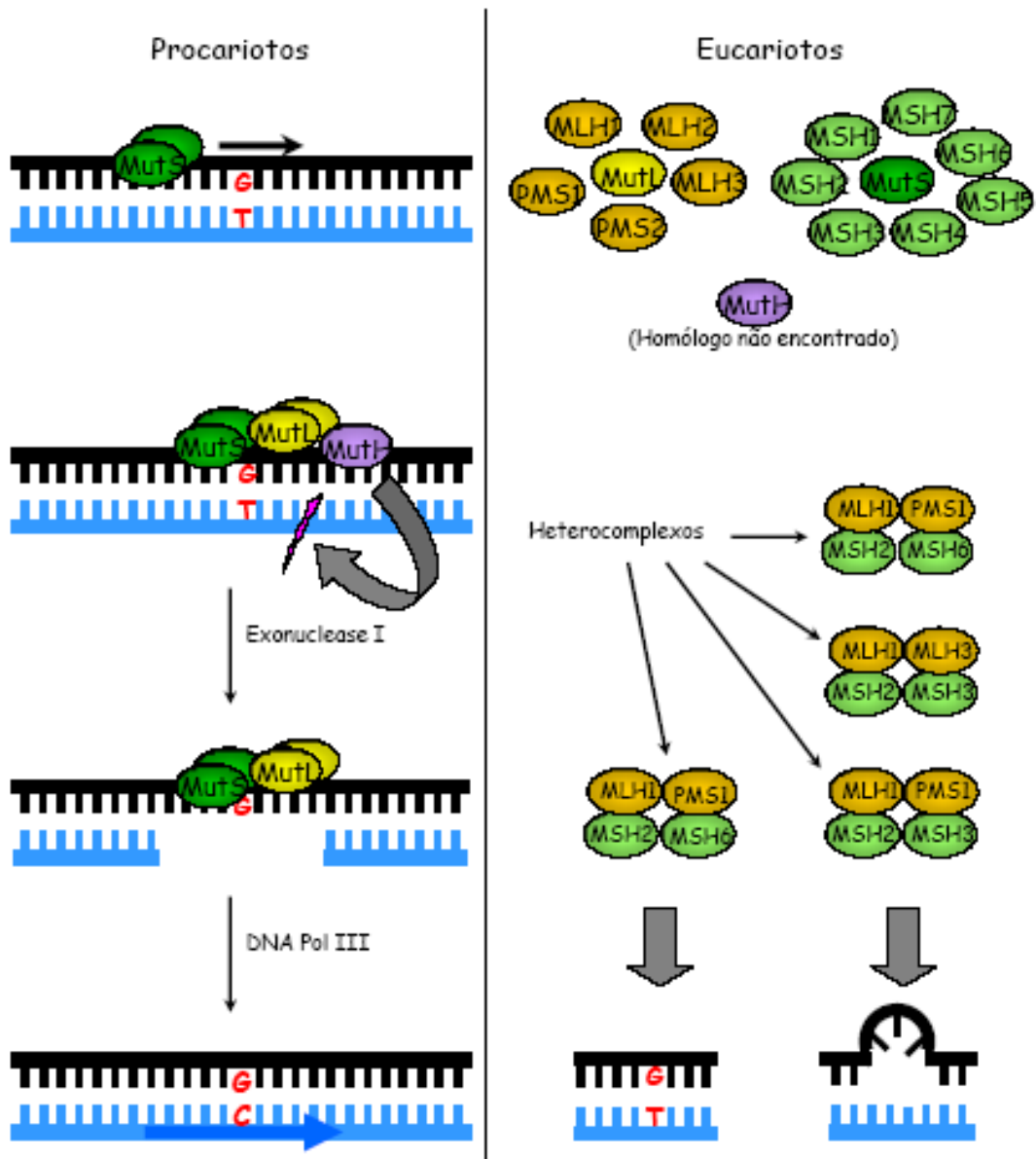


Figura 2: Representação esquemática do sistema de reparo de erros de pareamento em procariotos e eucariotos. Em Procariotos, o homodímero mutS se liga ao erro de pareamento e recruta o homodímero mutL que, por sua vez, ativa a endonuclease mutH. A proteína mutH corta a fita não-metilada que contém a base incorreta. Este fragmento é clivado e degradado. O DNA é então re-sintetizado e ligado. Em Eucariotos, a proteína mutH não é encontrada e as proteínas MSH e MLH interagem formando heterocomplexos que reparam diversos tipos de substratos como, por exemplo, erros de pareamento e alças de fitas simples geradas durante a replicação de microssatélites.

A grande diversidade genética observada em *T. cruzi* (e mais precisamente, nas cepas da linhagem *T. cruzi* II) pode exercer um papel importante na patogênese e na sobrevivência do parasito dentro dos diferentes hospedeiros. É então razoável postular que componentes de vias de reparo de DNA participem de processos que resultam na grande variabilidade genética observada na população do parasito.

A proteína TcMSH2 foi o primeiro componente da maquinaria de MMR descrita em *T. cruzi* (Augusto-Pinto et al., 2001). Por análises do RNA, foi demonstrado que o gene *Tcmsh2* era expresso em todos os estágios do ciclo de vida do parasito, mas em quantidades maiores nas formas replicativas epimastigota e amastigota. Em 2004, cinco genes envolvidos no MMR em *Trypanosoma brucei* foram caracterizados. Neste trabalho, foi mostrado que linhagens nocaute para os genes *MSH2* e *MLH1* apresentavam maior tolerância à alquilação gerada por MNNG e maior instabilidade de microssatélites após a cultura ser mantida por aproximadamente 25 gerações *in vitro* (Bell et al., 2004). Estes fenótipos são comumente associados a linhagens MMR-deficientes (de Wind and Hays, 2001). Somente com o sequenciamento completo dos genomas de *Leishmania major* (Ivens et al., 2005), *T. brucei* (Berriman et al., 2005) e *T. cruzi* (El-Sayed et al., 2005a), foi possível obter uma visão mais ampla da maquinaria de MMR nestes organismos. Foram identificados mais 4 genes homólogos de mutS em *T. cruzi*: *msh3*, *msh4*, *msh5* e *msh6* e sugeriu-se que a maquinaria de MMR em *T. cruzi*, *T. brucei* e *Leishmania major* esteja completa. Entretanto, não foi observada a presença de seqüências homólogas ao gene *msh4* (envolvida em processos de recombinação homóloga) em *L. major*, tampouco do gene *msh1* (envolvido no reparo mitocondrial) nos três tripanossomatídeos. É importante salientar que, todos os trabalhos na literatura ressaltam o papel crucial da proteína MutS e, principalmente, da proteína MSH2 homóloga em eucariotos, em organismos nos quais elas tenham sido identificadas.

A hipótese de um papel do MMR na geração de variabilidade genética entre as cepas de *T. cruzi* surgiu inicialmente após estudos que mostraram que cepas representativas das linhagens *T. cruzi* II e *T. cruzi* III apresentavam uma variação alélica em *loci* de microssatélites mitocondriais após tratamento com peróxido de hidrogênio (H₂O₂), o mesmo não ocorrendo em uma cepa representativa da linhagem *T. cruzi* I (Augusto-Pinto et al., 2003). Análises de polimorfismos para o gene

Tcmsh2 (de cópia única em *T. cruzi*) a partir de diferentes cepas do parasito, sugerem a existência de três isoformas da proteína: TcMSH2-A, -B e -C. Estes polimorfismos foram a base para a proposição da divisão do táxon *T. cruzi* em três haplogrupos, denominados A, B e C (Augusto-Pinto et al., 2003). Além disso, parasitos dos então denominados haplogrupos B e C (baseados nas isoformas da proteína TcMSH2) foram mais resistentes ao tratamento com cisplatina, de acordo com o que foi previamente reportado em outras linhagens celulares deficientes do MMR (Stojic et al., 2004). Estes estudos sugerem que, pelo menos sob condições de estresse genotóxico, as cepas pertencentes à linhagem *T. cruzi* I (que apresenta a isoforma A da proteína TcMSH2) apresentem uma atividade de MMR mais eficiente do que as cepas pertencentes à linhagem *T. cruzi* II (Augusto-Pinto et al., 2003). Esses dados corroboram também a hipótese de que a diferença na diversidade genética encontrada entre as linhagens *T. cruzi* I e II podem ser uma consequência de diferentes níveis de atividade da via de MMR. É importante observar que a proposta de classificação em haplogrupos A, B e C e a de divisão do táxon *T. cruzi* em três linhagens (de Freitas et al., 2006) se correlacionam: o haplogrupo A corresponderia à linhagem I, haplogrupo C corresponderia à linhagem II e o haplogrupo B agruparia as cepas posteriormente classificadas como linhagem III. Além disso, o caráter híbrido de *T. cruzi* III (haplogrupo B), principalmente da cepa CL Brener (*T. cruzi* II/III ou haplogrupo B/C), é ressaltado em todos os casos.

A variabilidade intra-específica observada no táxon *T. cruzi*, associada aos resultados obtidos nos estudos do MMR no parasito, nos levou a propor uma hipótese: é possível que a existência de diferentes isoformas da proteína TcMSH2 possa estar correlacionada a níveis de eficiência diferentes do MMR. Isso pode ter conseqüências diretas na geração de variabilidade, considerando que cepas pertencentes ao haplogrupo TcMSH2-A ou linhagem *T. cruzi* I apresentam menor variabilidade genética do que cepas dos haplogrupos B e C (*T. cruzi* III e II, respectivamente). Se considerarmos a variabilidade um elemento crucial na adaptação do parasito dentro do hospedeiro vertebrado, (Azuaje et al., 2007; Buscaglia et al., 2006), isto poderia explicar a baixa incidência de infecções causadas por membros da linhagem *T. cruzi* I (Di Noia et al., 2002; Llewellyn et al., 2009; Zingales et al., 1998).

Assim sendo, para uma melhor compreensão dos mecanismos envolvidos na geração de variabilidade genética em *T. cruzi* e, mais precisamente, o papel da maquinaria de MMR neste processo, investigamos as respostas aos danos no DNA do parasito, induzidos por diversos agentes genotóxicos, em cepas representativas das duas principais linhagens de *T. cruzi* (*T. cruzi* I e *T. cruzi* II), bem como as atividades de diferentes isoformas da proteína TcMSH2. Além disso, nós procuramos também testar diretamente o papel da proteína TcMSH2 na resposta diferencial apresentada pelas diferentes cepas de *T. cruzi* através de estratégias envolvendo manipulação genética do parasito. Para este fim, foram gerados parasitos mutantes para o gene *Tcmsh2* e também parasitos onde uma cópia adicional do gene foi inserida no genoma, com o objetivo de obter linhagens superexpressando a proteína TcMSH2. Entretanto, a obtenção de linhagens nocaute ou superexpressoras de um gene exógeno em qualquer organismo depende do conhecimento do seu genoma, de estratégias de manipulação gênica, bem como sobre os requerimentos necessários para um correto processamento dos mRNAs. Existem ainda algumas questões referentes ao processo de *trans*-splicing que ainda não estão completamente esclarecidas, como por exemplo, os sinais determinantes na escolha do sítio acceptor de A_(n), bem como na geração de transcritos derivados de *splicing* alternativo. Somente um estudo sobre esta questão foi relatado até o momento, analisando seqüências de cDNAs disponíveis de *Trypanosoma brucei* na busca por sítios potenciais de adição de SL e de poliA (Benz et al., 2005). No nosso trabalho, foi realizado um mapeamento genômico *in silico*, analisando todos os cDNAs e ESTs disponíveis contendo a seqüência SL e/ou a cauda de A_(n). Foram identificados os tamanhos medianos de 5'UTR (região 5' não traduzida) e de 3'UTR (região 3' não traduzida), dos motivos de polipirimidinas bem como as distâncias medianas entre o motivo de polipirimidinas e o sítio acceptor de SL ou de A_(n). Uma vez obtidos, estes dados serão extremamente relevantes para uma maior compreensão dos mecanismos de expressão gênica em tripanossomatídeos, bem como o melhoramento das técnicas de expressão de genes exógenos no parasito.

Objetivo geral

Investigar o papel do sistema de reparo de erros de pareamento e da proteína TcMSH2 na geração de variabilidade genética e na resposta ao estresse oxidativo em *Trypanosoma cruzi*.

Objetivos específicos

1. Comparar a viabilidade celular de cepas pertencentes a linhagens distintas de *Trypanosoma cruzi* frente ao tratamento com agentes genotóxicos;
2. Analisar o dano oxidativo induzido por peróxido de hidrogênio sobre o DNA de cepas pertencentes a linhagens distintas do parasito;
3. Comparar o metabolismo oxidativo em diferentes cepas do parasito;
4. Obter linhagens de *T. cruzi* nas quais o gene *Tcmsh2* foi deletado por recombinação homóloga;
5. Obter linhagens transfectadas de *T. cruzi* superexpressando o gene *Tcmsh2*
6. Realizar ensaios de viabilidade celular frente ao tratamento com agentes genotóxicos e análise do dano oxidativo induzido por peróxido de hidrogênio em parasitos nocautes e linhagens transfectadas superexpressoras;
7. Determinar a localização subcelular de TcMSH2;
8. Analisar seqüências do genoma do *T. cruzi* envolvidas no processamento do mRNA (sinais putativos de adição de *Spliced Leader* e poliadenilação) necessárias para a expressão de genes exógenos em *T. cruzi*.

1. Parasitos

Foram utilizadas neste estudo a cepa JG e os clones Col1.7G2, Silvio X-10 c11, Esmeraldo c13 e CL Brener, cedidos pelo Dr. Egler Chiari, do departamento de Parasitologia/ICB/UFMG e pela Dra. Bianca Zingales, do departamento de Bioquímica/IQ/USP. Parasitos foram mantidos a 28°C em sua forma epimastigota e em fase logarítmica em meio LIT (10g/L de infuso de fígado; 4g/L NaCl; 400mg/L KCl; 8g/L Na₂HPO₄; 1g/L glicose; 5g/L triptose; 10 mL de hemina a 0,2% em NaOH (p/v); pH 7,2) suplementado com soro fetal bovino a 10%, penicilina a 100 U/mL e estreptomicina a 100 µg/mL, então denominado meio LIT completo. Estas condições foram utilizadas em todos os experimentos, com modificações apenas para os experimentos de transfecção, nos quais após a eletroporação, os parasitos eram transferidos para o meio LIT completo, suplementado com 200 µg/mL do antibiótico adequado para a seleção das culturas transfectadas. Após o processo de seleção, as culturas foram mantidas na presença do antibiótico e/ou congeladas em N₂ para posterior utilização (em soro fetal bovino suplementado com 10% DMSO).

2. Ensaios de viabilidade celular de parasitos após tratamento com agentes genotóxicos

Os parasitos foram incubados em placas de 24 poços com diferentes concentrações de MNNG (*N*-methyl-*N'*-nitro-*N*-nitrosoguanidine) (fornecido pelo Dr. Álvaro Augusto da Costa Leitão, Instituto de Biofísica Carlos Chagas Filho, RJ), cisplatina (fornecida pela Quiral Química do Brasil S/A), ou H₂O₂ (Merck). Nos tratamentos com MNNG, as placas foram adicionalmente cobertas como papel alumínio. Após incubação por 3 ou 5 dias, 10 µL de cada poço foram retirados, diluídos em PBS 1x e corados, usando o corante vital Eritrosina B na concentração final de 0.04% (em PBS). As culturas foram contadas com o auxílio de uma câmara de Neubauer. Cada experimento foi realizado em triplicata e repetido três vezes e os resultados foram expressos como porcentagens de crescimento (média ± desvio padrão) em relação às culturas não tratadas.

3. Análise do acúmulo de 8-oxoguanina

Este experimento foi baseado em um protocolo descrito por Struthers e colaboradores (Struthers et al., 1998), que descreve a capacidade da avidina de reconhecer e ligar-se à base oxidada 8-oxoguanina, por possuírem similaridades estruturais. Epimastigotas em fase exponencial foram coletados por centrifugação (960 x g por 10 minutos), lavados com PBS gelado, ressuspensos em 500 µL PBS e incubados no gelo e no escuro na presença de 0,2 ou 0,3 mM de H₂O₂. Os parasitos foram então lavados 2x e ressuspensos em 50 µL PBS. Todas as etapas seguintes foram feitas no escuro, com PBS gelado e à temperatura de 4°C, exceto quando indicado. Os parasitos foram fixados com paraformaldeído (4% em PBS) por 20 minutos, lavados e ressuspensos em PBS para uma concentração final de 6 x 10⁸ células/mL. Alíquotas de 50 µL das suspensões de células foram distribuídas em lâminas de vidro de 8 poços com câmaras. Após 1 hora de incubação na geladeira, as células foram permeabilizadas com Triton X-100 (0,2% em PBS), tratadas com RNase A (100 µg/mL em PBS) por 15 minutos a 37°C e lavadas 3x com PBS. As lâminas são incubadas primeiramente com PBS + soro fetal bovino a 5% por 1 hora e depois com uma solução de avidina conjugada a FITC (em PBS, concentração final de 5 µg/mL) por mais 1 hora à temperatura ambiente. Após 4 lavagens com PBS, as lâminas foram montadas com uma solução de montagem (Glicerol/Tris-HCl pH 9,0 na proporção 9:1 + 2,3% DABCO) e visualizadas em microscopia de epifluorescência.

4. Estudos do metabolismo oxidativo em cepas de *Trypanosoma cruzi*

a. Produção de NADPH

A produção de NADPH pela via das pentoses também foi medida, através da análise da atividade enzimática das enzimas glicose-6-fosfato (G6PD) e 6-fosfogliconato (6PGD) desidrogenases. Os parasitos foram ressuspensos em PBS na presença de um coquetel de inibidores de proteases (para uma concentração final da leupeptina de 13 µM). A atividade das enzimas G6PD e 6PGD foi medida em

uma solução contendo KCl a 50 mM, MgCl₂ a 2 mM, Tris-HCl pH 7.6 a 50 mM, Triton X-100 a 0,1%, NADP⁺ a 250 μM, glicose-6-fosfato e 6-fosfogliconato a 1 mM cada. A redução de NADP⁺ foi monitorada espectrofotometricamente a 340 nm (Mielniczki-Pereira et al., 2007). A atividade da enzima G6PD foi obtida através da diferença da quantidade de NADPH, obtida pela enzima 6PGD, da quantidade total produzida pelas duas enzimas.

b. Medição do consumo de oxigênio e do controle respiratório

As medições do consumo de oxigênio foram determinadas à temperatura ambiente, usando um oxígrafo (Hansatech Instruments) acoplado a um computador. Os parasitos (10⁸ células) foram lavados com PBS e ressuspendidos em 1 mL de meio de respiração contendo sacarose a 125 mM, KCl a 65 mM, tampão HEPES pH7.2 a 10 mM, MgCl₂ a 1 mM, fosfato de sódio a 2 mM, EGTA a 500 μM, succinato a 5 mM, BSA a 1 mg/mL e digitonina a 20 μM (Col1.7G2, Silvio X-10 c11 e Esmeraldo c13) ou 30 μM (JG). O controle respiratório foi medido no mesmo experimento após a adição de ADP a 375 μM e carboxiatractilosídeo a 2 mM. Os experimentos foram repetidos três vezes, em duplicata.

c. Produção de peróxido de hidrogênio

Os parasitos (10⁸ células) foram lavados com PBS e ressuspendidos em um meio de reação contendo sacarose a 125 mM, KCl a 65 mM, tampão HEPES pH7.2 a 10 mM, MgCl₂ a 1 mM, fosfato de sódio a 2 mM, EGTA a 500 μM, succinato a 5 mM, BSA a 1 mg/mL digitonina a 20 μM (Col1.7G2, Silvio X-10 c11 e Esmeraldo c13), 30 μM (JG) ou 40 μM (CL Brener), peroxidase a 1U/mL e da sonda Amplex Red™ (Invitrogen, Molecular Probes) a 50 μM. As medições de produção de peróxido de hidrogênio foram determinadas em um fluorímetro (Hitachi F-4500), com excitação a 563 nm e emissão a 587 nm. Foram realizados três experimentos independentes, em duplicata. Para determinar quantitativamente a produção de H₂O₂, o aparelho foi calibrado pela adição de quantidades conhecidas de uma solução recém preparada de H₂O₂, quantificada espectrofotometricamente a 230 nm.

5. Amplificação gênica por PCR (reação em cadeia da polimerase) e construção de vetores

Os vetores contendo a região codificadora do gene *Tcmsh2* (denominados TOPOpCR4*Tcmsh2*) foram gentilmente cedidos pela Dra. Alice Machado Silva. Eles foram gerados através da inserção da sequência completa do gene *Tcmsh2* no vetor TOPO pCR4 (Invitrogen®), amplificado por PCR a partir do DNA genômico de culturas de CL Brener e utilizando iniciadores adequados para a inserção de um sítio de restrição para as endonucleases *XbaI* ou *EcoRV* na extremidade 5' do fragmento amplificado e *NotI* ou *SacI* na extremidade 3'.

Os fragmentos gênicos contendo os genes de resistência higromicina fosfotransferase (HYG) ou neomicina fosfotransferase (NEO) foram obtidos através de digestão dos plasmídios pROCK-GFP-HYG ou pROCK-GFP-NEO (DaRocha et al., 2004b) com as endonucleases *XhoI* e *NheI* e ligados ao TOPOpCR4*Tcmsh2*, previamente digeridos com as mesmas enzimas. Os plasmídios resultantes (TOPO N-Neo-C e TOPO N-Hyg-C) apresentam cada um dos genes de resistência flanqueados por regiões 5' e 3' da região codificadora do gene *Tcmsh2*, de forma que a deleção por recombinação homóloga dos dois alelos fosse facilitada. Antes da transfecção, ambos os plasmídios TOPO N-Neo-C e TOPO N-Hyg-C foram digeridos com *XbaI/NotI* ou *EcoRV/SacI*, conforme o caso. Os fragmentos resultantes de 1.7 e 1.9 Kb, respectivamente foram purificados em gel de agarose a 1%. Cerca de 50 µg destes fragmentos purificados foram usados na transfecção de epimastigotas de *T. cruzi* por eletroporação (DaRocha et al., 2004b).

Construções para a deleção do gene *Tcmsh2* foram também obtidas por uma técnica chamada de *gene SOEing* (Horton et al., 1990). Nesta estratégia, fragmentos amplificados correspondentes às regiões alvo para recombinação são ligados via PCR às extremidades 5' e 3' de um gene de resistência (também amplificado por PCR). Neste estudo, regiões 5' e 3' (não codificadoras) do gene *Tcmsh2* foram amplificadas usando os iniciadores diretos MSH2UpsFOR e MSH2DowNeoFOR e os iniciadores reversos MSH2UpsNeoREV e MSH2DowREV, gerando fragmentos de aproximadamente 0.5 Kb. Estes fragmentos foram posteriormente ligados a um cassette contendo a região codificadora do gene HYG, gerado em uma segunda etapa de PCR. Na terceira etapa, dois ou mais produtos de PCR (com extremidades

homólogas entre si) são usados como molde da PCR. Após este processo, os fragmentos foram então clonados no vetor pCR4TOPO ou pCR2.1 (Invitrogen) de acordo com instruções do fabricante. As sequências de todos os iniciadores utilizados neste trabalho estão detalhadas na Tabela 1 e os sítios de anelamento dos mesmos nos alelos de *Tcmsh2* estão detalhados no Anexo 1.

As construções para a superexpressão do gene *Tcmsh2* em epimastigotas de *T. cruzi* foram obtidas através da digestão do plasmídio pROCK-GFP-NEO (DaRocha et al., 2004b) para a retirada da região codificadora de GFP e posterior adição da sequência codificadora completa de *Tcmsh2*, originada do plasmídio TOPOpCR4*Tcmsh2*. Os plasmídios foram digeridos com *XhoI* ou *NotI* (pROCK-GFP-NEO ou TOPOpCR4*Tcmsh2*, respectivamente), tratados com a T4 DNA polimerase, e digeridos com *XbaI*. O inserto foi ligado ao vetor via T4 DNA ligase, gerando o plasmídio pROCK-*Tcmsh2*-NEO. O plasmídio foi construído de forma a sofrer integração no genoma do parasito no locus de β -tubulina, através de sua linearização com a endonuclease *NotI*.

6. Análise genotípica através da técnica de RFLP (*Restriction Fragment Length Polymorphism*)

Um fragmento de 875 pb foi obtido por PCR a partir do DNA de diferentes culturas com os iniciadores TmutS30 e TmutS41 (Tabela 1), utilizando o seguinte procedimento: 10ng DNA, 1 μ L de cada iniciador (a 5 pmoles/ μ L), 200 nM de cada dNTP, 2 μ L de tampão IB 10X e 0,5 mL de *Taq* polimerase (Phoneutria), em um volume final de 20 μ L. Os produtos de PCR foram digeridos com a endonuclease *HhaI* por 6 horas nas seguintes condições: 500 ng DNA, 5 μ L de tampão C, 1 μ L de enzima em um volume final de 50 μ L. Os produtos da digestão foram separados em gel de agarose a 1,5% contendo brometo de etídio 0,1 mg/mL e visualizados em luz ultravioleta.

Tabela 1: Descrição dos iniciadores utilizados neste trabalho. O sentido de anelamento dos iniciadores está indicado por setas (→Direto; ←Reverso). Os sítios de anelamento de cada iniciador é mostrado no Anexo 1.

	Iniciador/Sentido	Sequência (5' → 3')
Análise de clones transfectados e/ou PCR de colônia	HX1→	TTCTTCAAATATGCAGCGG
	TmutS30→	GACGAACTGATGGAAGTGA
	TmutS41←	CAAACCAAACCCATCGTAAG
	7442F→	GTGGAATGGGTTGCTCGTGCT
	7442R←	CCCACCCTGATTCCCTTGCCCCG
	8158F→	TGAAATTTTGCTACGAGTGT
	8158R←	CCCACCCTGATTCCCTTGCCCCG
	NEON11←	CGGCCGGAGAACCTGCGTGCAA
	HYGN11←	GAGCTGCATCAGGTCGGAGACGC
<i>Gene SOEing</i>	MSH2UpsFOR→	ACGTTGTCTTTTTATCGTGGCAGCC
	MSH2UpsNeoREV←	CCATATTGGCTGCAGGGTCGTCCCCG CCGCTCGCCCACCC
	MSH2DowNeoFOR→	TTCTTGACGAGTTCTTCTGAGTGTA ATGAAAAGGGAAG
	MSH2DowREV←	CCGGGTGGACAAAAACAAACGTT
	NEOFOR→	CGACCCTGCAGCCAATATGGGATCG
	NEOREV←	TCAGAAGAAGTCTCAAGAAGGCG

7. Cultivo de bactérias

As linhagens de bactérias *Escherichia coli* utilizadas neste estudo foram a XL1-Blue (*endA1 gyrA96(nalR) thi-1 recA1 relA1 lac glnV44 F'[::Tn10 proAB+ lacIq Δ(lacZ)M15] hsdR17(r_K⁻ m_K⁺)*), DH5α (*F- endA1 glnV44 thi-1 recA1 relA1 gyrA96 deoR nupG Φ80dlacZΔM15 Δ(lacZYA-argF)U169, hsdR17(r_K⁻ m_K⁺), λ⁻*) e BL-21 (*F ompT gal dcm lon hsdS_B(r_B⁻ m_B⁻) λ(DE3) pLysS(cm^R)*). Para o crescimento destas linhagens, foi utilizado o meio de cultura 2xYT (bacto-triptona a 1.6%, extrato de levedura a 1% e NaCl a 0.5%, pH 7.0) líquido ou sólido (quando acrescido de ágar a 1.5%), preparado segundo protocolos-padrão (Sambrook et al., 1989).

8. Transformação de bactérias

a. Por choque térmico

Culturas termocompetentes de *Escherichia coli* foram preparadas de acordo com protocolos-padrão (Sambrook et al., 1989) e utilizadas para transformação bacteriana, através da técnica de choque térmico. A transformação foi feita adicionando 1 a 5 µL do de DNA a uma alíquota de 50 ou 100 µL de bactéria e incubando a mistura em gelo por 30 minutos. A cultura recebeu choque térmico de 42°C por 1 minuto e voltou para o gelo por mais 1 minuto. Adicionou-se 500 µL de meio 2xYT e incubou-se a 37°C por 1h. A cultura transformada foi plaqueada em placas de meio 2xYT-ágar + 100 µg/mL de ampicilina e incubada por 37°C por 16h.

b. Por eletroporação

Culturas eletrocompetentes de *Escherichia coli* foram preparadas de acordo com protocolos-padrão (Sambrook et al., 1989) e utilizadas para transformação bacteriana, através da técnica de eletroporação. A transformação foi feita adicionando 2 µL do de DNA a uma alíquota de 40 µL de bactéria e incubando a mistura em gelo por 5 minutos. A cultura foi transferida para uma cuveta de 0.2 cm e submetida a choque em um eletroporador Micropulser (Bio-Rad). Quinhentos microlitros de meio 2xYT foram adicionados à cuveta e esta foi incubada a 27°C sob

agitação constante por 45 minutos. A cultura transformada foi plaqueada em placas de meio 2xYT-ágar + 100 µg/mL de ampicilina e incubada por 37°C por 16h.

9. Identificação por PCR de clones bacterianos transformados

As colônias transformadas foram repicadas simultaneamente para outras placas de 2xYT-ágar + 100 µg/mL de ampicilina e para tubos contendo 30 µL de tampão de lise para PCR (Tris pH 8,0 a 10 mM; EDTA a 100 µM; Tween 20 a 0,1%) e posteriormente fervidas a 95°C por 10 minutos. Os restos celulares foram separados por centrifugação a 20000 x g por 2 minutos e 4 µL do sobrenadante foram utilizados para a reação de PCR. Cada reação continha 2 µL de tampão IB 10x, 4 µL do lisado de bactérias transformadas, 50 µM de cada dNTP, 0,2µL de Taq (Phoneutria) e 5 µM de cada iniciador (Tabela 1), em um volume final de 20 µL. As reações foram feitas em termociclador PTC-100TM (MJ Research Inc) utilizando o seguinte programa: desnaturação inicial de 95°C por 3 minutos; 35 ciclos de 95°C por 1 minuto, anelamento a 58°C por 1 minuto e extensão a 72°C por 2 minutos; 72°C por 10 minutos. Os produtos da amplificação foram submetidos à eletroforese em gel de agarose a 1% contendo brometo de etídio 0,1 mg/mL e visualizados em luz UV.

10. Obtenção de DNA plasmidiano e digestão do DNA por endonucleases de restrição

Para a obtenção de DNA plasmidiano, colônias bacterianas foram inoculadas em 3 mL de meio 2xYT + 100 µg/mL ampicilina e incubados a 37°C por 16 horas sob agitação constante. Procedeu-se a extração do DNA em pequena (miniprep) ou larga escala (maxiprep). Para a primeira técnica, utilizou-se o kit GFX PLASMID MINI KIT (GE Healthcare), conforme instruções do fabricante, ou uma técnica de miniprep plasmidiano, baseada no método de lise alcalina (Sambrook et al., 1989); e para a segunda, utilizando o kit QIAGEN PLASMID MAXI KIT, conforme descrito no manual da QIAGEN (Chatsworth, CA, EUA). A análise da eficiência da extração do DNA foi feita aplicando-se 1 µL das amostras em géis de agarose a 1% e/ou por espectrofotometria em 260 nm.

As reações de digestão em pequena escala foram realizadas usando 1 a 5 µg de DNA, 3 a 12U de cada endonuclease de restrição, 3 µL de tampão 10x sugerido pelo fabricante, em um volume final de 30 µL e incubadas de 2 a 6 horas a 37°C, conforme a eficiência de digestão da endonuclease. Para reações de digestão em grande escala, 100 a 300 µg de DNA foram digeridos em reações de volume final de 360 µL, utilizando 50U de cada endonuclease, 36µL de tampão 10x e água para completar o volume adequado. A reação era incubada a 37°C por 16 horas. Ao final de todas as digestões, a endonuclease era inativada a 75°C por 15 minutos e a reação era estocada a 4°C até sua utilização. Após a digestão de DNAs (plasmídicos ou genômicos), as reações eram aplicadas em gel de agarose de 1 a 1,5% contendo brometo de etídio 0,1 mg/mL, conforme o tamanho das bandas esperadas na digestão. O gel sofreu eletroforese a 80V até a separação dos fragmentos, os fragmentos de interesse foram excisados do gel e purificados usando o kit GFX GEL e PCR PURIFICATION (GE Healthcare), conforme instruções do fabricante.

11. Produção de proteínas recombinantes e anticorpos policlonais

O plasmídio TOPOpCR4*Tcmsh2* foi digerido com *XbaI/HindIII* e o fragmento resultante foi clonado no vetor pET21a+ (Novagen), previamente digerido com *NheI/HindIII*, para a geração de uma proteína recombinante contendo os primeiros 442 aminoácidos da proteína TcMSH2 fusionada a uma cauda de polihistidina. Colônias isoladas de *Escherichia coli* BL-21 (DE3) pLysS, previamente transformadas com o plasmídio denominado pET21a+*Tcmsh2*(t) foram inoculadas em 3 mL de meio 2xYT + 100 µg/µL de ampicilina e incubadas por 16 horas a 37°C sob agitação constante. Esse pré-inóculo era então diluído 1:50 em 10 mL de meio 2xYT + 100 µg/µL de ampicilina e incubado a 37°C até atingir a DO₆₀₀ entre 0,4 e 0,5. Nesta etapa, adicionou-se IPTG para uma concentração final de 1 mM e a indução teve um período de 4 horas. A eficiência de indução foi verificada aplicando-se alíquotas intermediárias em gel de poliacrilamida a 10%. A cultura restante foi centrifugada a 3800 x g e solubilizada através de 5 ciclos de congelamento-descongelamento. A fração solúvel foi separada da insolúvel por centrifugação a 20000 x g por 10 minutos a 4°C. A fração insolúvel foi aplicada em gel de poliacrilamida a 10% (composta predominantemente de corpos de inclusão) e

a banda de tamanho correspondente a 48 KDa foi excisada do gel. A banda sofreu sucessivos ciclos de congelamento com N₂ e maceração (até converter-se em um pó), emulsificada com adjuvante de Freund e inoculada subcuneamente em 4 camundongos BALB/c fêmeas para a geração de anticorpos policlonais anti TcMSH2(t). A imunização foi reforçada com uma nova dose (2 semanas após a primeira inoculação) e a sangria final foi feita 8 semanas após a inoculação inicial.

12. Produção de extratos protéicos totais de epimastigotas de *Trypanosoma cruzi* e ensaios de *Western blot*

Os parasitos foram coletados do meio LIT por centrifugação a 960 x g por 10 minutos, lavados com tampão PBS e ressuspensos em tampão da amostra para SDS-PAGE [tampão inicial composto de Tris-HCl pH 6.8 a 500 mM e SDS a 0.4%, acrescido de glicerol a 40%, azul de bromofenol a 0.02%, β-mercaptoetanol a 4%, SDS a 20% e um coquetel de inibidores de proteases diluído 1:100 (*Protease Inhibitor Mix*TM, GE Healthcare)] para a concentração de 10⁶ células/mL. Este lisado foi então fervido por 10 minutos e centrifugado a 20000 x g por 2 minutos. Cinco microlitros deste sobrenadante foram separados em gel de poliacrilamida a 12,5% por 1 hora e meia no gelo e o gel foi corado com azul de Coomassie.

Para a análise dos extratos protéicos totais por *Western blot*, as amostras foram separadas em gel de poliacrilamida (ver descrição acima), transferidas para uma membrana de nitrocelulose por 16 horas a 4°C. Para verificar a eficiência da transferência, a membrana foi posteriormente corada com Ponceau (solução a 0.1% diluído em 5% ácido acético glacial) e descorada com água. A membrana foi bloqueada por 2 horas com TBS-T (Tris a 10 mM, NaCl a 150 mM e Tween-20 a 0,1%) + 5% de leite em pó e incubada com o anticorpo primário (diluído 1:500 em TBS-T + 5% leite em pó) por 2 horas.

Após a incubação com o anticorpo primário, a membrana foi lavada três vezes com TBS-T e incubada com o anticorpo secundário conjugado à peroxidase (diluído 1:4000 em TBS-T + 5% leite em pó) por 1 hora e meia e posteriormente lavada 5 vezes com TBS-T. Todas as etapas do *Western blot* foram realizadas à temperatura ambiente. A membrana foi revelada com o kit ECL Plus (GE

Healthcare), conforme instruções do fabricante, usando filme fotográfico ou através do aparelho STORM™ (GE Healthcare).

13. Transfecção de formas epimastigotas de *Trypanosoma cruzi* por eletroporação

Cerca de 4×10^7 parasitos em fase exponencial de crescimento foram utilizados em cada sistema de transfecção. Os parasitos foram coletados do meio LIT por centrifugação a $960 \times g$ por 10 minutos, lavados por duas vezes com tampão PBS e ressuspendidos em 400 μL de tampão de eletroporação [tampão CITOMIX (KCl a 120 mM; CaCl_2 a 150 μM ; K_2HPO_4 a 10 mM; EDTA a 2mM; MgCl_2 a 5 mM; HEPES a 25mM; pH 7.6) e tampão Sacarose (sacarose a 277mM; KH_2PO_4 a 7mM; MgCl_2 a 1mM; pH 7.4), na proporção 3:1], para uma concentração final de parasitos de $10^8/\text{mL}$. Às cuvetas de 0.2 cm utilizadas para cada sistema foram adicionados 50 μg do DNA apropriado em 50 μL final ou 50 μL de tampão TE (para eletroporação das culturas controle) e, após a adição dos parasitos, mantidas em gelo por 15 minutos. Após este processo, a cuveta foi colocada no eletroporador Gene Pulser™ (BioRad Laboratories) ajustado para pulsos de 0.3kV, 500 μF , com 2 pulsos com 10s de intervalo entre eles e constantes de tempo entre 3 e 6 ms (DaRocha et al., 2004b). Após a eletroporação, as células foram incubadas à temperatura ambiente por 10 minutos, transferidas para garrafas de 25 cm^2 estéreis contendo 5 mL de meio LIT completo + 200 $\mu\text{g}/\text{mL}$ de antibiótico (G418 e/ou higromicina B) e incubadas a 28°C até a seleção das culturas controle (aproximadamente 30-35 dias).

14. Clonagem de parasitos transfectados

a. Diluição limitante

Na técnica por diluição limitante, foram preparados 25 mL de uma solução de LIT completo + 10% sangue + 200 $\mu\text{g}/\text{mL}$ de antibiótico (concentração final). Com estas alíquotas fez-se a diluição dos parasitos (30, 3 e 0.3 parasitos/mL). As diluições foram transferidas para placas de 96 poços, de modo que um terço de cada placa fosse preenchido com cada uma das diluições e que cada poço recebesse 200 μL de

cada diluição. As placas foram seladas com Parafilm™, armazenadas em câmara úmida a 28°C e monitoradas a cada 3 dias até o surgimento de poços com alta densidade de parasitos, sendo então considerados clones.

b. Meio semi-sólido

Na técnica de crescimento em meio semi-sólido, o meio estéril [agarose *low melting* a 0.75%; BHI (*Brain Heart Infusion*) a 48.4%; meio LIT a 48.4%] foi misturado a 2.5% de sangue desfibrinado e 250 µg/mL de antibiótico (concentração final). Cada placa de Petri recebeu 20 mL desta mistura e após a gelificação do meio, em cada placa foram espalhados 100 µL de cultura de parasitos transfectados a uma densidade celular de 10^3 , 10^4 ou 10^5 /mL. As placas foram embaladas uma a uma com papel Kraft estéril e mantidas em posição invertida por um período de 30-35 dias. Os clones obtidos em ambos os processos foram transferidos para placas de 24 poços contendo 2 mL de meio LIT completo + 200 µg/mL do antibiótico adequado, seladas com Parafilm™ e mantidas em sistema de câmara úmida a 28°C por um período de 5 dias, sendo então transferidos para garrafas de cultura de 25 cm² com 5 mL de meio LIT completo + 200 µg/mL de antibiótico. A identificação preliminar dos clones foi feita por PCR (Tabela 1) e posteriormente confirmada por *Southern* e *Northern* blots.

15. Análise por PCR do DNA dos clones de *Trypanosoma cruzi* transfectados

O DNA genômico utilizado como molde para a PCR foi obtido a partir de 1 mL de cada poço contendo culturas de clones transfectados. A amostra foi centrifugada a 20000 x g por 2 minutos, o sobrenadante descartado e o restante ressuspenso em 20 µL de água Milli-Q. Após fervura por 10 minutos, a amostra foi novamente centrifugada e 10 µL do sobrenadante foram utilizados em reações de PCR. Cada reação continha 4 µL de tampão 5x, 10 µL do lisado de parasitos transfectados, 100 µM de cada dNTP, 0,5 µL de GoTaq™ (Promega) e 5 µM de cada iniciador (Tabela 1), em um volume final de 20 µL. As reações foram feitas em termociclador PTC-100™ (MJ Research Inc) utilizando o seguinte programa: desnaturação inicial de 95°C por 3 minutos; 35 ciclos de 95°C por 1 minuto, anelamento a 58°C por 1 minuto e extensão a 72°C por 2 minutos; 72°C por 10

minutos. Os produtos da amplificação foram submetidos à eletroforese em gel de agarose a 1% contendo brometo de etídio 0,1 mg/mL e visualizados em luz UV. Os clones que apresentaram resultados positivos neste ensaio foram transferidos para garrafas de cultura contendo 5 mL de meio LIT completo e os demais foram descartados.

16. Southern blot

Para a extração do DNA genômico dos parasitos, foram utilizadas culturas em fase exponencial de crescimento, em um total de 10^8 células. O procedimento é dividido em três etapas: lise celular, desproteinização e precipitação do DNA. Na primeira etapa, a cultura foi transferida a tubos de 15 mL e centrifugadas a $960 \times g$ por 10 minutos. Os parasitos foram então lavados por 2 vezes com tampão PBS 1x, ressuspendidos em uma solução de lise (Tris-HCl a 10 mM pH 8.0; EDTA a 100 μ M pH 8.0; RNase A a 20 μ g/mL; SDS a 0.5%) e incubados por 1 hora a 37°C. Em seguida, adicionou-se proteinase K para uma concentração final de 100 μ g/mL e o lisado foi homogeneizado e incubado a 50°C por 3 horas. Na segunda etapa, adicionou-se ao lisado 200 μ L de fenol, a mistura foi homogeneizada e centrifugada a $960 \times g$ por 10 minutos a 4°C. A fase aquosa foi transferida a outro tubo, onde foi adicionado 200 μ L de fenol: clorofórmio:álcool isoamílico na proporção 25:24:1. A mistura foi novamente homogeneizada e centrifugada a $960 \times g$ por 10 minutos a 4°C. A fase aquosa foi transferida para outro tubo, onde foi misturado a 200 μ L de clorofórmio:álcool isoamílico na proporção 24:1. A mistura foi mais uma vez homogeneizada e centrifugada a $960 \times g$ por 10 minutos a 4°C. A fase aquosa foi transferida para um novo tubo. Para a terceira etapa, adicionou-se NaCl em concentração final de 250 mM e 2 volumes de etanol absoluto. A mistura foi homogeneizada e incubada por no mínimo 3 horas a -20°C, sendo então centrifugada a $20000 \times g$ por 15 minutos a 4°C. O pellet foi lavado por 2 vezes em etanol 70% e ressuspendido em 50 mL de água ou tampão TE.

As análises do DNA genômico de *T. cruzi* por Southern blot foram feitas de acordo com protocolos-padrão (Sambrook et al., 1989). Dez microgramas de DNA genômico de cada amostra de cultura foram digeridos com diferentes endonucleases de restrição e submetidos à eletroforese em gel de agarose a 0,8% por 16 h a 24V. O

gel foi depurinado em solução de depurinação (HCL a 0,25N) por 15 minutos sob leve agitação. Após lavagem com água durante 5 minutos, o gel foi transferido para uma solução de desnaturação (NaOH a 0.5N; NaCl a 1.5M), agitando levemente e durante 30 minutos. Este processo foi repetido mais uma vez e o gel foi então novamente lavado com água durante 5 minutos. Posteriormente, o gel foi neutralizado em tampão de neutralização (Tris-HCl a 0.5M pH 8.0; NaCl a 1.5M), sob agitação leve durante 30 minutos. Este processo foi repetido mais uma vez e, por fim, foi realizada a transferência para a membrana de nylon Hybond-N™ (GE Healthcare) por capilaridade em solução SSC 10X (NaCl a 1.5M; citrato de sódio a 150 mM; pH 7.0). A membrana foi lavada em SSC 2X e imobilizada em luz ultravioleta no aparelho UVSTRATALINKER™ (Stratagene). A membrana foi pré-hibridada com solução de hibridização [Formamida a 50%; SSC 6X; Denhardt's 10X; SDS a 0,2%; DNA de esperma de salmão a 0,2% (concentração final de 75µg/mL)] por 3 horas a 42°C, sob rotação. A membrana foi então hibridada por 24 horas com 50 µL de sonda marcada radioativamente com 50 µCi de $\alpha^{32}\text{P}$ -dCTP por “*random priming*” (kit MEGAPRIME Random Primer Labeling™, GE Healthcare, conforme instruções do fabricante) e previamente desnaturada a 100°C durante 5 minutos, no mesmo tampão de pré-hibridização, a 42°C, sob rotação. Em seguida, procederam-se duas lavagens de 30 minutos cada com uma solução de SSC 2X, SDS a 0.2% e NaPPi a 0.1%, uma vez à temperatura ambiente e uma vez a 65°C. A solução foi então lavada mais uma vez sob condição de alta estringência em solução de SSC a 0.2% e SDS a 0,2%. Após esta última lavagem, a membrana foi selada e exposta a um filme de raios-X a -80°C em um cassete com intensificador. O filme foi revelado em diferentes tempos de exposição.

17. Northern blot

Cerca de 5×10^8 parasitos na forma epimastigota (em fase exponencial) foram utilizados para extração e purificação de RNA. As amostras de RNA foram obtidas utilizando-se RNAsy MINI KIT (QIAGEN) conforme instruções do fabricante. O método de extração consiste na lise de células em condições desnaturantes e da purificação do RNA por cromatografia de afinidade, utilizando água como eluente. Após quantificação em espectrofotômetro a 260nm, foi feita a análise dos RNAs por

eletroforese em géis de agarose a 1.2% em tampão MOPS pH 7.0 (MOPS a 20 mM; acetato de sódio a 5 mM; EDTA a 1mM) e formaldeído 5% (Teixeira et al., 1995). A cada amostra foram adicionados 2.5 µL de tampão da amostra (0.75mL de formamida, 0.15 mL de tampão MOPS 10x, 0.24 mL de formaldeído, 0.1 mL de água livre de RNase/DNase, 0.1 mL de glicerol e 0.08 mL de azul de bromofenol a 10%). As amostras foram incubadas por 15 minutos a 65°C, e a elas foi adicionado um µL de uma solução de brometo de etídio a 1,0mg/mL antes da aplicação no gel. O procedimento foi realizado a 100V durante 45 minutos à temperatura ambiente.

Para os ensaios de *Northern blot*, 25 µg do RNA total foram separados em gel de agarose a 1.2% em tampão MOPS pH 7.0 por 3 horas. Após a eletroforese, os géis foram tratados 2 vezes com SSC 10x (NaCl a 3M; citrato de sódio a 0.3 mM; pH 7.0) por 20 minutos. Os RNAs foram transferidos por capilaridade, em SSC 10x, para membranas de nitrocelulose QIABRANE NYLON (QIAGEN) também tratadas previamente com SSC 10x por 5 minutos. Após a transferência, as membranas foram lavadas em SSC 6x e os RNAs foram imobilizados por irradiação com luz UV no aparelho UVSTRATALINKER (Stratagene). As membranas foram pré-hibridadas por 3h, a 42°C, sob rotação, em solução de formamida a 50%, SSC 6x, Denhardt's 10x, SDS a 0,2% e DNA de esperma de salmão a 75 µg/mL. Posteriormente, as membranas foram hibridadas por 14h a 42°C com 50 µl de sonda purificada, marcada radiativamente com 50 µCi de [α -³²P] dCTP por "random priming" (utilizando os kits MEGAPRIME Random Primer Labeling ou Ready-To-Go (GE Healthcare), conforme instruções do fabricante. Após a hibridação, as membranas foram lavadas 3 vezes com solução de lavagem (SSC 2x; SDS a 0.1%) a 65°C por 30 minutos e reveladas à temperatura ambiente por intervalos variados de tempo usando o sistema de documentação Phosphorimager® (GE Healthcare).

18. Localização subcelular da proteína em fusão com GFP

O gene *Tcmsh2* foi amplificado a partir de DNA genômico utilizando os iniciadores direto TcMSH2Xba0 e reverso LocMSH2EcoRI (que inserem sítios de restrição para as endonucleases *XbaI/EcoRI* nas extremidade 5' e 3' do fragmento amplificado e geram uma mutação no códon de parada). O fragmento amplificado foi inserido no vetor TOPO 2.1 (Invitrogen®), digerido com as endonucleases

XbaI/EcoRI e subclonado no vetor de expressão em *T. cruzi* pTREX-Neo-GFP (DaRocha et al., 2004b; Vazquez and Levin, 1999), digerido previamente com as mesmas enzimas. Bactérias competentes foram transformadas com este plasmídeo e colônias isoladas foram usadas na obtenção de DNA em grande escala pela técnica de maxiprep (ver seção 11). Epimastigotas de *T. cruzi* em fase exponencial (clone CL Brener) foram transfectados com 200 µg do então denominado pTREX-Neo-MSH2::GFP (ou do plasmídeo vazio pTREX-Neo-GFP) conforme descrito anteriormente (Teixeira et al., 1995), para a expressão da proteína TcMSH2 fusionada a GFP em sua porção C-terminal. Após 24 horas de transfecção, os parasitos foram colocados em lâminas, fixados em paraformaldeído a 4% (em PBS) por 15 minutos à temperatura ambiente, permeabilizados com solução de permeabilização (PBS + Triton X-100 a 0.1% + RNase a 0.1 mg/mL) por 15 minutos a 37°C, lavados com PBS e contra-corados com uma solução de PBS + DAPI a 0.01 mM. As lâminas foram visualizadas em um microscópio de epifluorescência acoplado a uma câmera CCD, em objetiva de imersão 100x.

19. Imunofluorescência

Neste experimento, 2×10^5 parasitos foram lavados com PBS e transferidos para lâminas, fixados com paraformaldeído a 2% (em PBS) completas por 20 minutos e lavados 3x com PBS. As células fixadas foram permeabilizadas com Triton X-100 a 0.1% (em PBS) por 5 minutos e bloqueadas com PBS contendo 1% BSA por 30 minutos. A lâmina foi incubada com anticorpo primário não-imune ou policlonal gerado contra a proteína TcMSH2(t) recombinante (vide seção 12) por 1 hora, lavada 3x com PBS e incubada com o anticorpo secundário anti-camundongo conjugado a Alexa Fluor 488. A contra-coloração foi feita incubando-se a lâmina com 0.01 mM DAPI e após a lavagem 3x com PBS, a lâmina foi montada com Vectashield™ (Vector Laboratories). Todo o experimento foi realizado à temperatura ambiente e a lâmina foi visualizada em microscópio de epifluorescência acoplado a uma câmera CCD, em objetiva de imersão 100x. Os dados foram obtidos pela aluna de doutorado Sheila Cristina Nardelli, da Universidade Federal de São Paulo.

20. Análises *in silico* de seqüências de aminoácidos de TcMSH2

As seqüências parciais de aminoácidos deduzidas do gene de interesse em diferentes cepas de *Trypanosoma cruzi* foram obtidas através de pesquisa em banco de dados Genbank (<http://www.ncbi.nlm.nih.gov/Genbank/index.html>) ou GeneDB (<http://www.genedb.org>) e utilizadas para a construção de um cladograma no programa MEGA versão 4.0 (Tamura et al., 2007), pelo método de *Neighbor-Joining*, com *Bootstrap* como método de significância das ramificações derivadas de 1000 replicatas (mínimo de 60%). As 28 seqüências utilizadas na análise foram originadas das cepas: P255 (número de acesso no Genbank: AAY51668), MNcl2 (AAY51667 e AAX77022), CBBcl3 (AAX77026), SilvioX10cl1 (AAX77025), Y (AAX77023), SC43cl1 (AAX77021), M5631 (AAX77020), TU18cl2 (AAX77018 e AAX77017), RBI (AAM14741), D7 (AAM14740), 239 (AAM14739 e AAM14738), Col1.7G2 (AAM14737), 115 (AAM14729 e AAM14736), 226 (AAM14734), 1005 (AAM14733), 577 (AAM14732), JG (AAM14730), 167 (AAM14726 e AAM14728), 231 (AAM14727), Tulahuén (AAG00261), CL Brener (EAN89935 e EAN98026) e Esmeraldo cl3 (gentilmente cedida pelo Dr. Najib M. El-Sayed).

As seqüências completas do gene *Tcmsh2* derivadas do clone CL Brener foram obtidas via banco de dados GeneDB com números de acesso Tc00.1047053509643.80 e Tc00.1047053507711.320, enquanto que a seqüência derivada do clone Col1.7G2 foi obtida por amplificação e seqüenciamento, posteriormente depositada no GenBank sob o número de acesso GQ869654. As seqüências foram traduzidas *in silico* e alinhadas usando os programas ClustalW versão 1.8 (Thompson et al., 1994) ou Multalin (<http://bioinfo.genotoul.fr/multalin/>). Todas as seqüências de aminoácidos foram analisadas quanto à busca por domínios e motivos usando bancos de dados dos programas Pfam (Sonnhammer et al., 1997), Prosite (Bairoch, 1991) e Interpro (Apweiler et al., 2000).

21. Análises *in silico* de padrões de processamento do mRNA em *T. cruzi*

Foi construído um sub-banco de dados contendo seqüências intercodificadoras de *T. cruzi*, obtidas do banco de dados derivados do seqüenciamento do genoma do parasito (El-Sayed et al., 2005a). Foram mapeados ESTs (*Expressed Sequence Tags*) e cDNAs disponíveis, utilizando como sondas as seqüências referentes ao *Spliced Leader* (SL) e à cauda de poliadenilação ($A_{(n)}$). Eram consideradas completas apenas as seqüências que apresentavam em uma das extremidades seqüências similares às das sondas utilizadas e na outra extremidade o códon de iniciação ou terminação. Estas extremidades comuns foram posteriormente excisadas de todas as seqüências. Estas análises permitiram a construção de sub-bibliotecas, que continham seqüências referentes às regiões 5'UTR, 3'UTR e intergênicas. Neste último subgrupo, foram realizadas buscas por motivos ricos em polipirimidinas (polyYs), tendo como requisito básico apresentar um tamanho mínimo de 9 nucleotídeos, sendo tolerado uma única interrupção por uma purina interna. De posse destes dados, foram feitas análises de tamanho de regiões 5' e 3'UTR, tamanho de polyYs e sua composição, bem como as distâncias destes motivos em relação aos sítios putativos de SL e de $A_{(n)}$. Todas as análises foram realizadas através da plataforma Linux. Um esquema representativo da metodologia de mapeamento do genoma de *T. cruzi*, através das análises de todos os cDNAs e ESTs disponíveis contendo SL e/ou cauda de $A_{(n)}$, está mostrado na Figura 3.

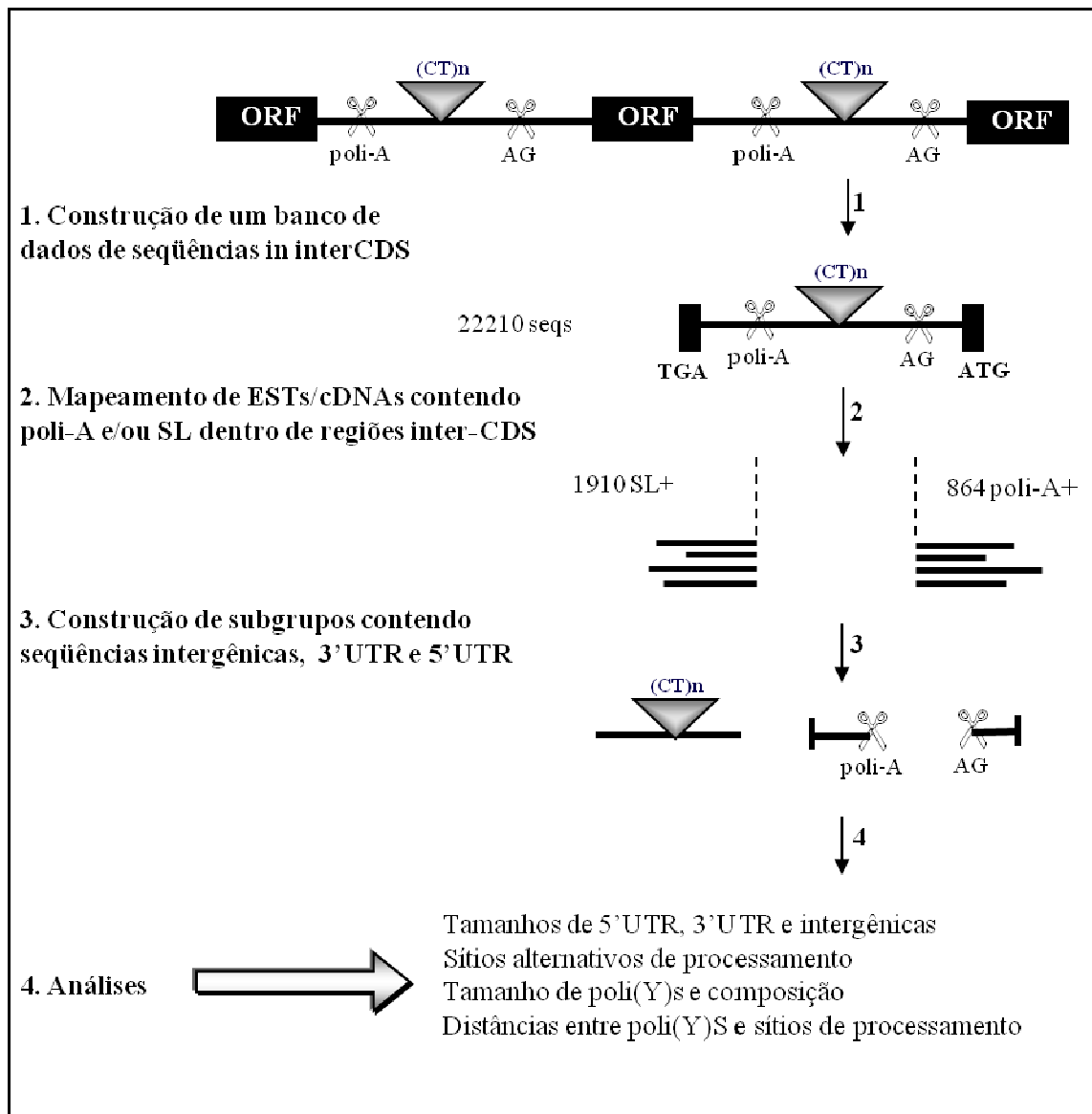


Figura 3: Representação gráfica da metodologia para análise dos sítios de processamento do mRNA em *T. cruzi*. Triângulos cinza representam os motivos de polipirimidinas [(CT)_n]; poli-A: poliadenilação; AG: dinucleotídeo aceptor de *Spliced Leader* (SL); ORF: (janela aberta de leitura ou *open reading frame*); tesouras indicam sítios de adição de SL ou poli-A.

1. Inferências filogenéticas a partir de seqüências parciais de aminoácidos para a proteína TcMSH2 nas diferentes cepas de *Trypanosoma cruzi*

Resultados anteriores obtidos pelo nosso grupo demonstraram a existência de polimorfismos no gene *Tcmsh2* de *T. cruzi*, que poderiam estar relacionados a uma eficiência diferencial do MMR nas diversas cepas analisadas. Neste trabalho, as 17 seqüências do gene *Tcmsh2* estudadas indicaram a existência de três isoformas da proteína TcMSH2, o que resultou na proposição da divisão do táxon *T. cruzi* em três haplogrupos: A, B e C (Augusto-Pinto et al., 2003). Em busca de dados adicionais relacionados ao gene *Tcmsh2*, pesquisamos novamente em bancos de dados, na busca por seqüências em outras cepas de *T. cruzi*. Esta pesquisa resultou na elaboração do cladograma apresentado na Figura 4A, que reúne as seqüências anteriormente depositadas por Augusto-Pinto e colaboradores, como também as depositadas por Westenberger e colaboradores (Augusto-Pinto et al., 2003; Westenberger et al., 2005), totalizando 28 seqüências parciais para o gene *Tcmsh2* de *T. cruzi*. Neste cladograma, pode ser visualizada a divisão das cepas em três ramificações, correspondentes aos haplogrupos A, B e C (Augusto-Pinto et al., 2003), as quais podem ser correlacionadas às linhagens *T. cruzi* I, II e III (de Freitas et al., 2006). A Figura 4B apresenta as respectivas substituições de aminoácidos presentes nas cepas analisadas. Estas substituições foram encontradas em todas as seqüências analisadas, o que sugere uma “assinatura” haplogrupo-específica. Foi também observado neste cladograma que algumas cepas apresentaram polimorfismo alélico, o que pode resultar na classificação em 2 grupos como, por exemplo, o clone CL Brener, que apresenta seqüências de TcMSH2 dos tipos B e C. Em todas as análises descritas neste trabalho, os clones Col1.7G2 e Silvio X-10 c11 foram selecionados como representantes da linhagem *T. cruzi* I (haplogrupo A), e a cepa JG e o clone Esmeraldo c13 foram selecionados como representantes da linhagem *T. cruzi* II (haplogrupo C). Para confirmar a identidade destas cepas, o DNA foi isolado e submetido a ensaios de RFLP (*Restriction Fragment Length Polymorphism*). Os resultados mostraram perfis de digestão característicos dos haplogrupos aos quais as cepas pertencem (Figura 4C).

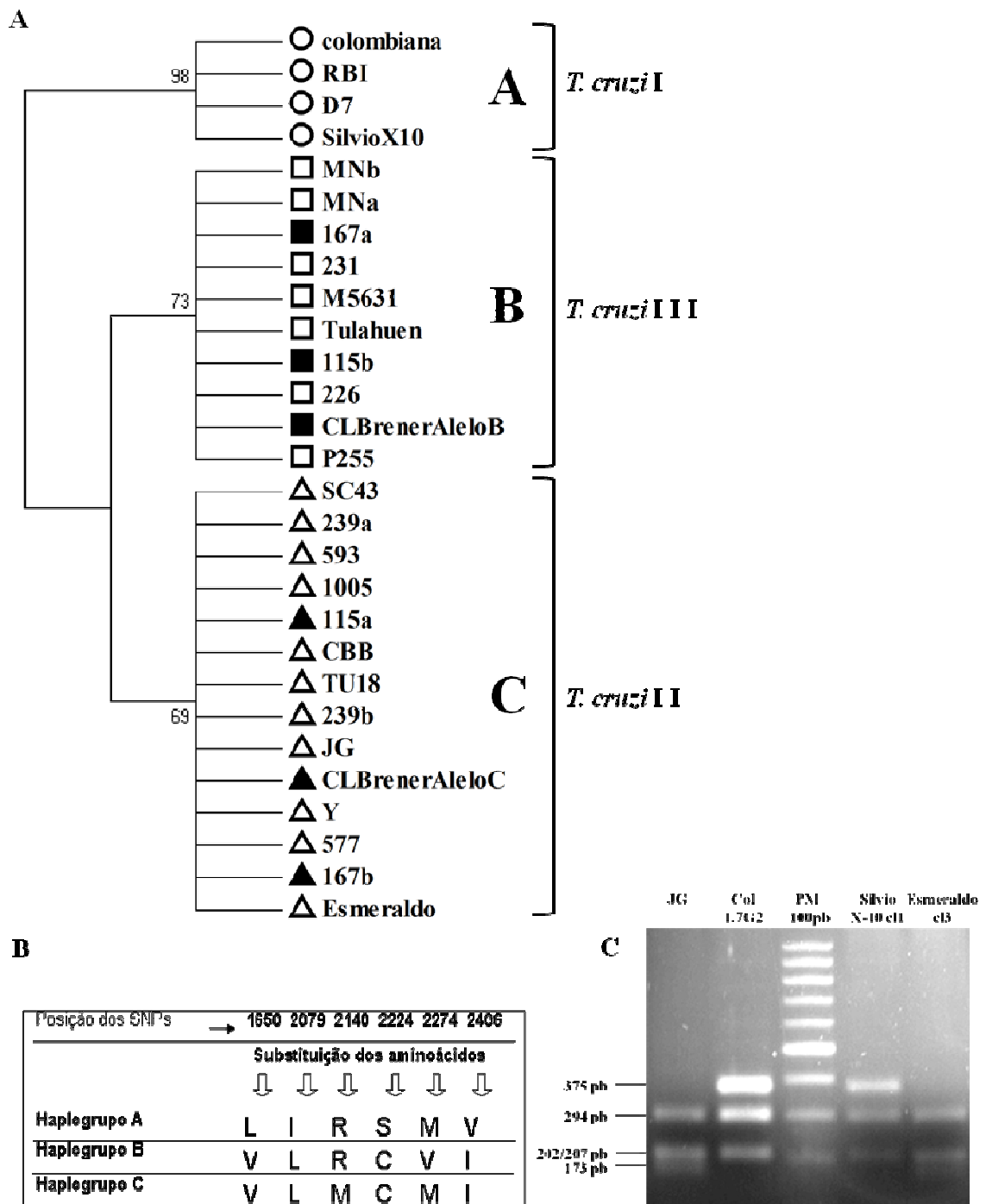


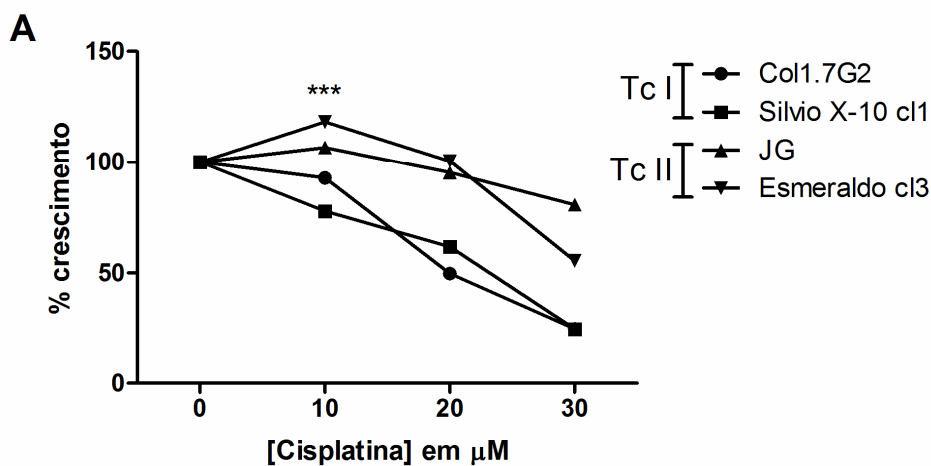
Figura 4: Construção de um cladograma a partir de seqüências parciais de aminoácidos para a proteína TcMSH2 e análise das cepas por RFLP (*Restriction Fragment Length Polymorphism*). Em A: Cladograma construído a partir de seqüências de aminoácidos depositadas no GenBank (NCBI) para o gene de TcMSH2. A distribuição das cepas entre os haplogrupos baseados em TcMSH2 A (○), B (□) e C (△) foi correlacionada com a classificação proposta por Freitas et al., (2006) em linhagens *T. cruzi* I, II e II. Os símbolos cheios (■ ou ▲) indicam as

cepas híbridas (B/C) pertencentes às linhagem *T. cruzi* II e III. Os valores indicados nas ramificações são índices de *bootstrap* observados em 1000 replicatas. Em **B**: Posição dos SNPs (*Single Nucleotide Polymorphisms*) e suas respectivas substituições de aminoácidos correspondentes observadas para cada haplogrupo. Em **C**: Análise por RFLP das cepas Silvio X-10 c11 e Esmeraldo c13, usando as cepas Col1.7G2 e JG como referência para as linhagens *T. cruzi* I e II, respectivamente. Os fragmentos de 875 pb foram obtidos e digeridos conforme descrito em Material e Métodos. Produtos da digestão foram separados em gel de agarose a 1,5%. Os perfis de digestão teóricos do gene *Tcmsh2* das cepas Col1.7G2 foram comparados aos tamanhos dos fragmentos resultantes da digestão (em pb).

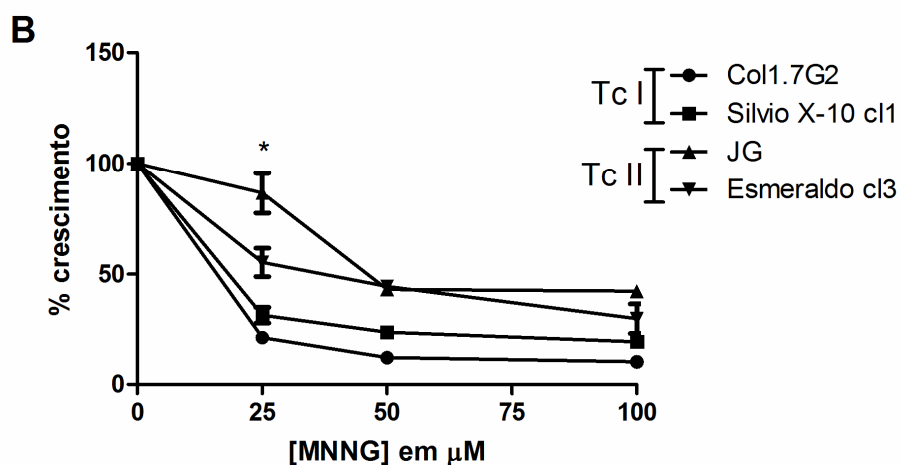
2. Ensaio de sobrevivência de parasitos pertencentes a diferentes cepas de *T. cruzi* ao tratamento com cisplatina e MNNG

Com o objetivo de analisar a eficiência do MMR nos clones Col1.7G2, Silvio X-10 c11 e Esmeraldo c13 e da cepa JG, os parasitos foram tratados com diferentes concentrações do agente genotóxico cisplatina, e após 5 dias, a viabilidade celular medida como uma porcentagem de sobrevivência dos mesmos com relação às culturas não tratadas. A cisplatina é uma droga amplamente utilizada no tratamento do câncer. Ela forma ligações cruzadas intra e entre fitas do DNA, que levam à morte celular se não adequadamente reparadas. As células proficientes em MMR são mais susceptíveis à droga, pois as proteínas da via de MMR podem agir como transdutoras em cascatas de sinalização para a morte celular (Borst et al., 2008). A Figura 5A mostra que as cepas Silvio X-10 c11 e Col1.7G2 (MSH2 haplogrupo A ou linhagem *T. cruzi* I) foram mais sensíveis à cisplatina do que as cepas Esmeraldo c13 e JG (MSH2 haplogrupo C ou linhagem *T. cruzi* II). O experimento foi repetido na presença de cádmio, um inibidor do MMR (Jin et al., 2003), para verificar se o MMR está envolvido nas diferenças de sensibilidade a este agente genotóxico. Os dados obtidos foram apresentados pela Dra. Carolina Furtado em sua Tese de Doutorado (Torres-Silva, 2009), que indicaram a participação de componentes da via de MMR na resposta ao tratamento com cisplatina, uma vez que a diferença previamente observada entre as cepas foi abolida na presença do inibidor.

Outro agente genotóxico utilizado é o *N*-metil-*N'*-nitro-*N*-nitrosoguanidina (MNNG), que gera metilações na posição O⁶ de guaninas (Stojic et al., 2004). Este tipo de dano é também um substrato para o MMR. As quatro cepas de *T. cruzi* foram expostas a diferentes concentrações de MNNG e a viabilidade da cultura foi determinada após 72 horas (Figura 5B). Em todas as concentrações, as culturas de Col1.7G2, Silvio X-10 c11 (MSH2 haplogrupo A ou *T. cruzi* I) apresentaram maior sensibilidade ao tratamento com MNNG do que JG e Esmeraldo c13 (MSH2 haplogrupo C ou *T. cruzi* II). As diferenças observadas entre as cepas de *T. cruzi* I e II não são tão evidentes quanto as que foram observadas após o tratamento com cisplatina; entretanto, elas foram igualmente abolidas quando o mesmo tratamento foi repetido na presença de cádmio (Dra. Carolina Furtado, Tese de Doutorado, 2009).



*** $p < 0,001$ quando comparado entre os grupos *T. cruzi* I e *T. cruzi* II



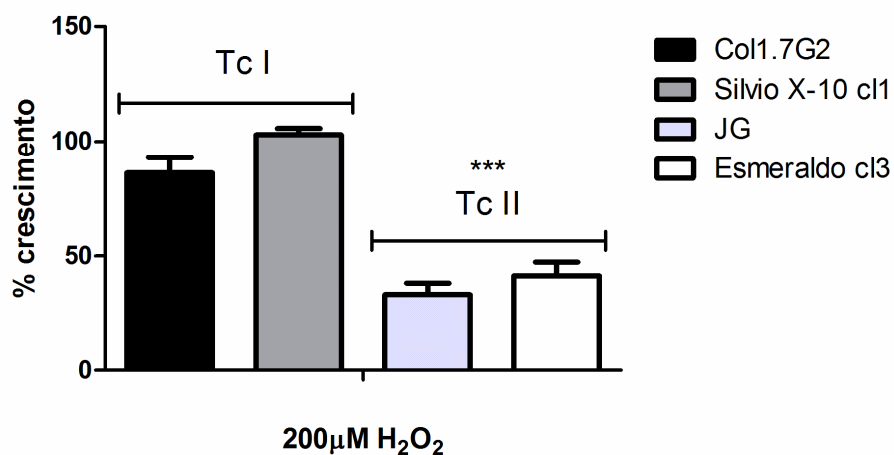
* $p < 0,05$ quando comparado entre os grupos *T. cruzi* I e *T. cruzi* II

Figura 5: Efeito da cisplatina e do MNNG no crescimento das cepas de *Trypanosoma cruzi*. Culturas de epimastigotas em fase exponencial de diferentes cepas *T. cruzi* foram tratadas com 10, 20 e 30 μM cisplatina (A) ou com 25, 50 e 100 μM MNNG (B). As células viáveis foram quantificadas após 3 (para o tratamento com MNNG) ou 5 dias (para o tratamento com cisplatina) com eritrosina B. Os resultados são mostrados como médias \pm desvio padrão de 3 experimentos independentes, realizados em triplicata.

3. Sobrevivência de parasitos pertencentes a diferentes cepas de *T. cruzi* ao estresse induzido por peróxido de hidrogênio

O dano oxidativo é uma ameaça constante à estabilidade genômica, que pode surgir do próprio metabolismo oxidativo endógeno e causar graves danos em uma célula. As lesões nas bases que podem surgir a partir do dano oxidativo ao DNA, tais como a 8-oxoG, se não removidas pelo sistema de reparo por excisão de bases (BER), podem ser reconhecidas pelo MMR (Slupphaug et al., 2003). Para investigar a resposta de parasitos pertencentes às duas linhagens (*T. cruzi* I e II) ao dano oxidativo, a capacidade de diferentes cepas de sobreviver na presença de peróxido de hidrogênio (H₂O₂) foi testada. Foram adicionados 200 µM de H₂O₂ às culturas de epimastigotas e a viabilidade celular foi determinada após 5 dias. Conforme mostrado na Figura 6, as linhagens de *T. cruzi* I e II exibiram diferentes padrões de resposta ao H₂O₂, mas diferentemente dos resultados observados nos tratamentos com cisplatina e MNNG, as cepas da linhagem *T. cruzi* I (Col1.7G2 e Silvio X-10 c11) foram mais resistentes ao tratamento do que as cepas da linhagem *T. cruzi* II (JG e Esmeraldo c13). Além disso, ao contrário do observado nos tratamentos com cisplatina e ao MNNG, o cádmio não aboliu a resposta diferencial das cepas de *T. cruzi* ao tratamento com H₂O₂. Como mostrado nos experimentos realizados pela Dra. Carolina Furtado em sua Tese de Doutorado (Torres-Silva, 2009), quando as cepas de Esmeraldo c13 e Silvio X-10 c13 são incubadas com peróxido de hidrogênio na presença do inibidor do MMR, a taxa de crescimento de Silvio X-10 c13 (*T. cruzi* I) permaneceu mais alta do que a observada em Esmeraldo c13 (*T. cruzi* II).

Os resultados obtidos nos ensaios de produção de NADPH pela via das pentoses (Figura 7A), da produção basal de H₂O₂ (Figura 7B) e do consumo basal de O₂ (Figura 7C) sugerem que os resultados obtidos após a indução de estresse oxidativo não estão correlacionados com diferenças no metabolismo oxidativo das células analisadas.



***p<0,001 quando comparado entre os grupos *T. cruzi* I e *T. cruzi* II

Figura 6: Efeito do estresse oxidativo induzido por H₂O₂ no crescimento das cepas de *Trypanosoma cruzi*. Culturas de epimastigotas em fase exponencial de crescimento foram tratadas com 200 μM H₂O₂. As células viáveis foram determinadas após 5 dias de tratamento com o corante vital eritrosina B. Os resultados são mostrados como médias ± desvios-padrão de 3 experimentos independentes, realizados em triplicata.

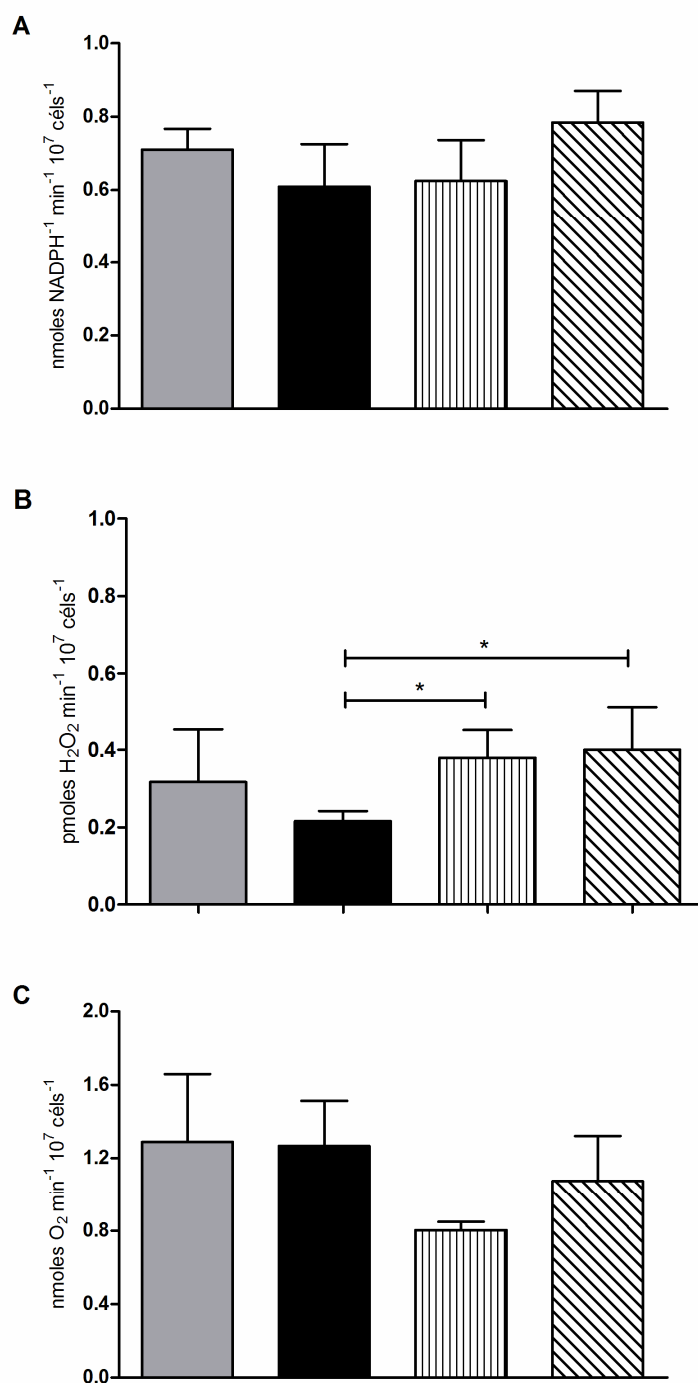


Figura 7: Análise comparativa do metabolismo oxidativo entre as cepas de *T. cruzi*. Em **A** é mostrada uma medição da produção de NADPH nas diferentes cepas de *T. cruzi* pela via das pentoses, através da medida de absorvância a 340 nm. Em **B** é mostrada a produção basal de H₂O₂ pelas diferentes cepas de *T. cruzi*, medida por fluorimetria, com excitação a 563 nm e emissão a 587 nm. Em **C** é mostrado o consumo basal de oxigênio pelas diferentes cepas de *T. cruzi*. As células foram incubadas em meio de respiração (vide Material e Métodos) e os dados registrados em um oxígrafo. Os experimentos foram repetidos três vezes, em duplicata (p<0.05).

4. Análise do acúmulo de 8-oxoguanina no genoma das diferentes cepas de *T. cruzi*

Para analisar diretamente o efeito do estresse oxidativo induzido por H_2O_2 , foi feita uma análise quantitativa do acúmulo de 8-oxoG no genoma de diferentes cepas de *T. cruzi* – antes e depois do tratamento com H_2O_2 – utilizando uma metodologia baseada na ligação da avidina à base oxidada 8-oxoG. Este reconhecimento ocorre em função das similaridades estruturais compartilhadas pela biotina (ligante da avidina) e pela 8-oxoG (Struthers et al., 1998). Neste trabalho, a especificidade da ligação foi mostrada através de ensaios na presença ou ausência de inibidores, tais como guanina e biotina. Inicialmente, testamos também a especificidade desta ligação, incubando o conjugado avidina-FITC ao dGTP ou ao d8-oxoGTP antes da incubação com as células para a detecção de 8-oxoG. Conforme mostrado na Figura 8, a incubação prévia de avidina-FITC com quantidades crescentes de d8-oxoGTP resultou na diminuição da fluorescência, o que não ocorreu quando a avidina-FITC foi previamente incubada com dGTP. Este resultado está de acordo com o observado em estudos anteriores (Struthers et al., 1998). O próximo passo foi detectar o acúmulo de 8-oxoG no genoma das diferentes cepas de *T. cruzi*. A Figura 9 resume os resultados obtidos neste experimento, no qual é mostrado que as cepas pertencentes à linhagem *T. cruzi* I (Col1.7G2 e Silvio X-10 cl1) apresentaram níveis mais baixos da base oxidada do que as cepas da linhagem *T. cruzi* II (JG e Esmeraldo cl3) (Figura 9A). Esta diferença se deve principalmente a um maior acúmulo de 8-oxoG no kDNA de cepas da linhagem *T. cruzi* II, um efeito visivelmente acentuado após 20 minutos de tratamento com 200 μM H_2O_2 (Figura 9B).

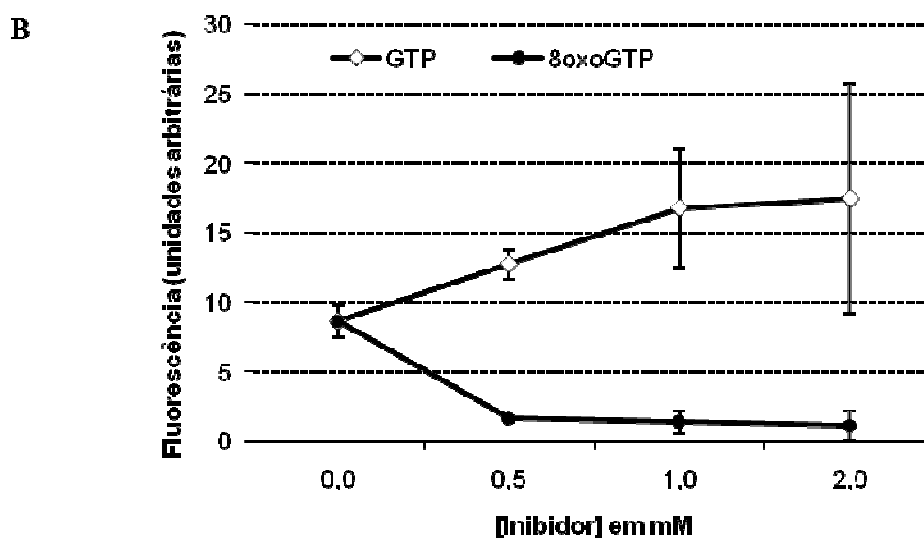
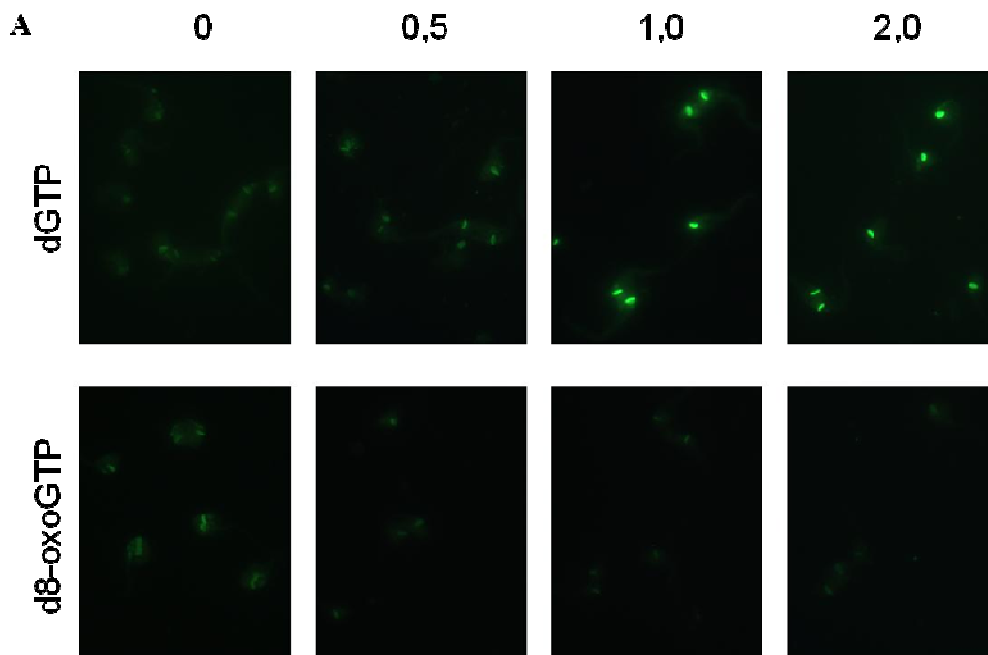


Figura 8: Ligação de avidina conjugada a FITC na presença de dGTP e d8-oxoGTP. Culturas de epimastigotas em fase exponencial foram fixadas com paraformaldeído e permeabilizadas com Triton X-100. Após a transferência para os poços da lâmina de vidro, os parasitos foram incubados com 5 ug/mL de avidina conjugada a FITC (ou avidina conjugada a FITC previamente incubada com diferentes concentrações de dGTP ou d8-oxoGTP) por 1 hora. A lâmina foi visualizada em microscopia de epifluorescência (A) em objetiva com imersão 100X. A fluorescência foi quantificada com o programa ImageJ (<http://rsbweb.nih.gov/ij/>) e os dados analisados, conforme mostrado em B, no qual cada ponto representa a média \pm desvio-padrão de 25 células.

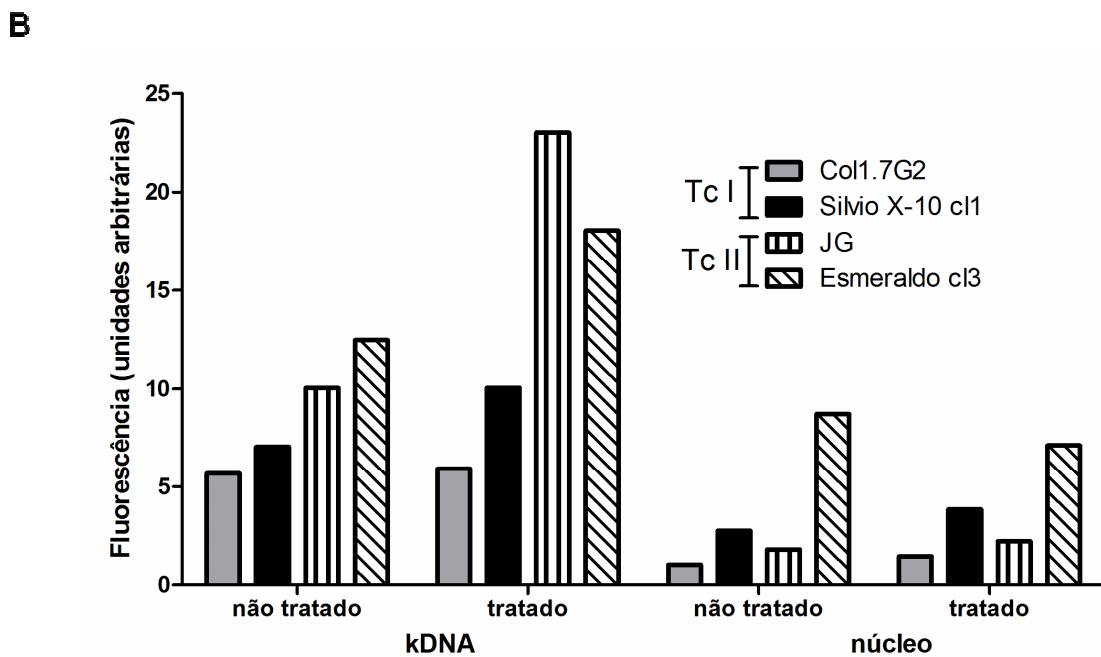
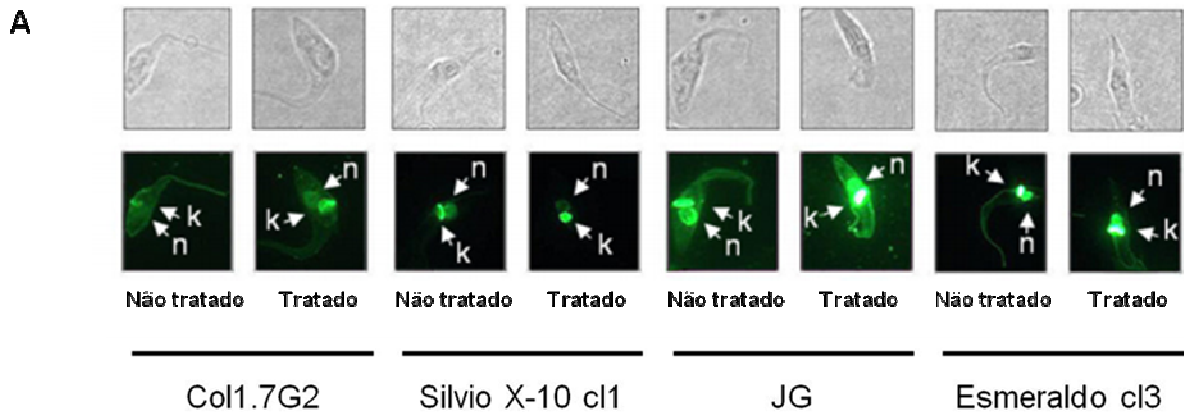


Figura 9: Análise do acúmulo de 8-oxoguanina em diferentes cepas de *T. cruzi*. Culturas de epimastigotas em fase exponencial foram tratadas com 200 μ M H₂O₂ por 20 minutos, fixadas com paraformaldeído e permeabilizadas com Triton X-100. Os parasitos foram incubados com 5 μ g/mL de avidina conjugada a FITC por 1 hora. A lamina foi visualizada em microscopia de campo claro e por fluorescência (A) em objetiva com imersão 100X. A fluorescência foi quantificada com o programa ImageJ (<http://rsbweb.nih.gov/ij/>) e os dados analisados, conforme mostrado em (C). Cada barra representa a média \pm desvio-padrão de 25 células.

4. Expressão da proteína recombinante TcMSH2 e produção de anticorpos anti-TcMSH2

Para avaliar a expressão da proteína TcMSH2 no *T. cruzi*, bem como determinar a sua localização subcelular, camundongos foram imunizados com a proteína TcMSH2 recombinante para a obtenção de anticorpos policlonais anti-TcMSH2. Os experimentos de expressão e imunização foram realizados em colaboração com a aluna de mestrado Viviane Grazielle da Silva. O fragmento gênico que codifica os primeiros 442 aminoácidos da porção N-terminal de TcMSH2 (a principal proteína da via de MMR) foi inserida no vetor pET21a+. Não obtivemos colônias bacteriana expressando a proteína TcMSH2 completa, possivelmente devido a uma toxicidade da proteína, que não permitiu o cultivo de colônias bacterianas viáveis. Bactérias *Escherichia coli* BL-21 (DE3) pLysS foram transformadas com este plasmídeo e, conforme mostrado na Figura 10A, a cultura bacteriana transformada foi capaz de produzir uma proteína recombinante fusionada a uma cauda de poli-histidina na sua porção C-terminal, com tamanho aproximado de 49 KDa. A expressão da proteína truncada TcMSH2::His atingiu seu pico máximo de expressão após 2 horas de indução com IPTG (Figura 10A). Entretanto, a proteína recombinante mostrou-se insolúvel (Figura 10B). Para verificar se a banda de 49 KDa corresponde à proteína recombinante em fusão com histidina foi feito um ensaio de *Western blot*, no qual o extrato protéico foi transferido para uma membrana de nitrocelulose e incubado com anticorpo monoclonal anti poli-histidina C-terminal. O controle positivo foi feito utilizando a proteína recombinante TcRpL7a purificada (Pais et al., 2008), que também possui uma cauda de poli-histidina na porção C-terminal. Os resultados mostrados na Figura 10C mostram que a proteína recombinante de 49 KDa produzida está em fusão com poli-histidinas.

A fração insolúvel do extrato bacteriano induzido foi separada em gel de poliacrilamida a 12,5%, a banda correspondente à proteína TcMSH2::His foi cortada do gel e preparada de acordo com Material e Métodos para a imunização de camundongos. Esta mistura foi inoculada em quatro camundongos BALB/c, para a produção de anticorpos policlonais contra a proteína recombinante. A fim de determinar a especificidade do anticorpo policlonal produzido nos camundongos, extratos de bactérias transformadas induzidas e não induzidas, bem como um extrato

de epimastigotas de *T. cruzi*, foram separados em gel de poliacrilamida a 12,5%, transferidos para membranas de nitrocelulose e incubadas com o soro de camundongos não imunes e imunizados com a proteína recombinante TcMSH2::His. A Figura **10D** mostra que o soro não imune não foi capaz de reconhecer bandas em nenhuma das amostras testadas. Por outro lado, o soro anti-TcMSH2::His reconhece a banda correspondente à proteína recombinante (49 KDa) apenas no extrato bacteriano induzido e uma banda de tamanho aproximado de 75 KDa no extrato de epimastigotas de *T. cruzi*. A Figura **10D** representa os resultados obtidos com o soro de um dos camundongos, mas experimentos realizados com soros dos demais animais tiveram resultados similares.

5. Avaliação da expressão de TcMSH2 nas diferentes cepas de *T. cruzi*

Os resultados obtidos nos ensaios de sobrevivência dos parasitos na presença de agentes genotóxicos e de H₂O₂ sugerem que cepas pertencentes às linhagens *T. cruzi* I e II apresentam respostas distintas aos tratamentos. Para verificar se as diferenças na resposta a estes agentes não são resultantes de variações na expressão de componentes da via de MMR, foram feitos ensaios de *Northern* e *Western* blots, nos quais foram verificados os níveis protéicos e dos transcritos para a principal proteína da via, TcMSH2. No experimento de *Northern* blot, foi utilizado como sonda um fragmento correspondente a uma porção da região que codifica o gene *Tcmsh2*. A Figura **11A** mostra que não há uma diferença significativa no nível do mRNA das cepas analisadas, o que sugere que o nível dos transcritos para o gene *Tcmsh2* não é o fator determinante na resposta diferencial das cepas. Com a produção de anticorpos policlonais contra a proteína recombinante TcMSH2::His, foi possível investigar se a similaridade observada no nível dos transcritos para o gene *Tcmsh2* se refletiria na expressão da proteína. A Figura **11B** mostra um ensaio de *Western* blot, no qual se observa que o nível da proteína TcMSH2 (banda de 75 KDa) permanece o mesmo em todas as cepas analisadas, como observado no experimento de *Northern* blot. Isto indica que as diferenças fenotípicas observadas nas cepas analisadas não estão correlacionadas com diferenças nos níveis de expressão da proteína TcMSH2.

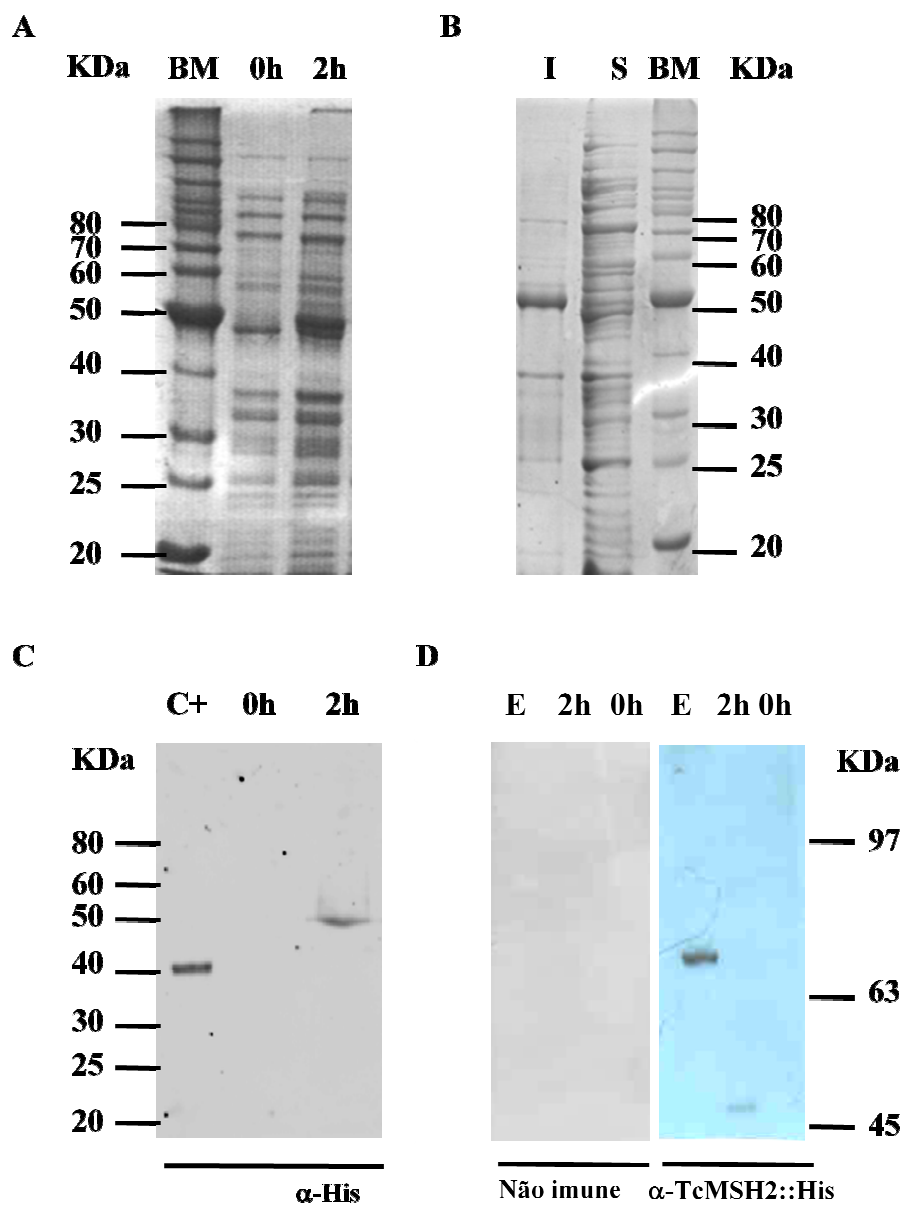


Figura 10: Expressão da proteína recombinante TcMSH2::His, avaliação de sua solubilidade e geração de anticorpos policlonais anti-TcMSH2. Foram retiradas alíquotas da cultura bacteriana transformada com o plasmídeo pET21a+_TcMSH2::His e induzida com 1 mM IPTG, nos intervalos de 0, 2 e 4 horas. Estas amostras foram centrifugadas, ressuspensas em tampão da amostra e separadas em gel de poliacrilamida a 10% corado com Azul de Comassie, conforme mostrado em **A**. A amostra correspondente ao intervalo de 2 horas foi fracionada e as frações solúvel (S) e insolúvel (I) foram separadas em gel de poliacrilamida a 10% e coradas com Azul de Comassie (**B**). Em **C** é mostrado um ensaio de *Western blot* com anticorpo primário anti-His (diluição 1:2000) dos extratos protéicos mostrados

em **A**. O controle positivo foi feito com a proteína recombinante TcRpL7a em fusão com poli-histidinas. A especificidade dos soros obtidos foi testada na diluição 1:500, através de ensaios de *Western blot* (mostrado em **D**), nos quais foram utilizados extratos de epimastigotas de *T. cruzi* (canaleta E) e extratos de bactérias transformadas induzidas (canaleta 2h) e não induzidas (canaleta 0h). As amostras foram também incubadas com soro não imune. O padrão de peso molecular utilizado em todos os experimentos é o Benchmark (Invitrogen).

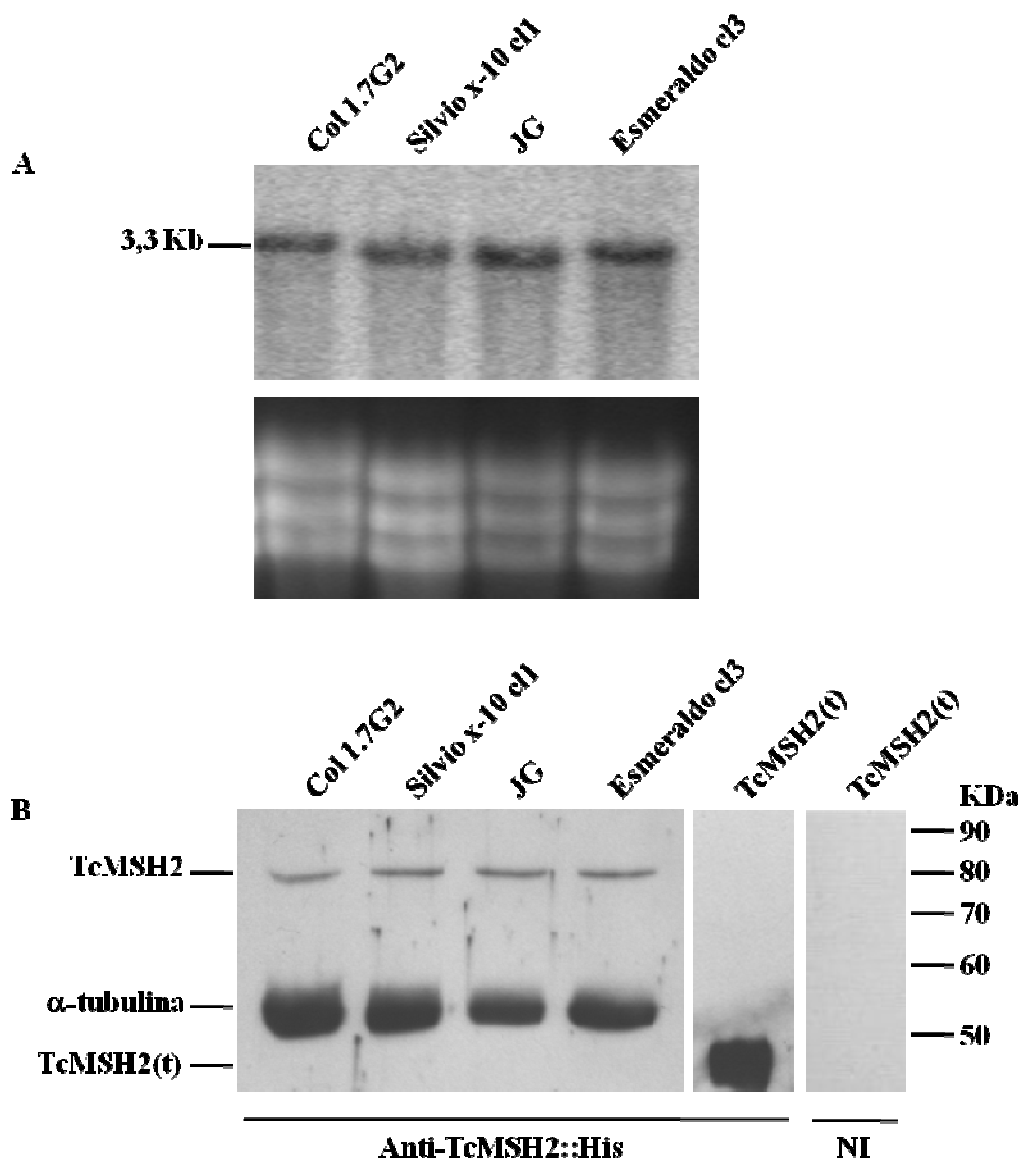


Figura 11: Análise da expressão de TcMSH2 nas diferentes cepas de *T. cruzi*. Em **A**: *Northern* blot utilizando o RNA total de epimastigotas das cepas Col1.7G2, Silvio X-10 cl1, JG e Esmeraldo cl3. As amostras foram hibridizadas com uma sonda correspondente aos primeiros 400 nucleotídeos da sequência codificadora de *Tcmsh2*. A porção inferior da figura mostra a visualização por brometo de etídio das bandas correspondentes ao rRNA, para demonstrar a quantidade de RNA aplicado no gel. Em **B**: Determinação da expressão de TcMSH2 por ensaio de *Western* blot, conforme descrito em Materiais e Métodos. Extratos totais de epimastigotas foram separados em gel de poliacrilamida, transferidos para membranas de nitrocelulose e incubados com soro não imune (NI), anticorpo primário policlonal produzido contra a proteína recombinante TcMSH2::His ou anticorpo monoclonal anti α -tubulina. As amostras aplicadas no gel estão identificadas no topo de cada canaleta.

6. Geração de linhagens heminocaute para o gene *Tcmsh2* no clone CL Brener

Para analisar geneticamente o envolvimento da via de MMR nas diferentes respostas observadas ao dano no DNA de *T. cruzi*, foram construídos vetores para a deleção, por recombinação homóloga, do locus *Tcmsh2*, codificador da principal proteína do MMR, TcMSH2 (Figura 2). Os parasitos foram transfectados com estas construções, com o objetivo de substituir os alelos do locus *Tcmsh2* por genes de resistência aos antibióticos G418 e higromicina B via recombinação homóloga. A deleção de um dos alelos do gene *Tcmsh2* foi obtida alcançada em diferentes cepas do parasito. A Figura 12A mostra a caracterização via PCR dos clones heminocautes (*Tcmsh2*^{+/-}) obtidos, utilizando iniciadores específicos para cada um dos alelos de *Tcmsh2* (B ou C) presentes em CL Brener, conforme mostrado na Figura 12B. Os resultados obtidos com dois clones transfectados indicam que o gene de resistência integrou-se no alelo *Tcmsh2*-B. A deleção do alelo *Tcmsh2*-B foi também confirmada por *Southern* blot (Figuras 12C e 12D) e pela análise dos níveis dos transcritos de *Tcmsh2*, que foram reduzidos para 50% após a deleção de um dos alelos (Figura 12E). Testes em mais de 10 clones *Tcmsh2*^{+/-} mostraram que, em todos os casos houve a deleção do alelo *Tcmsh2*-B (dados não mostrados).

Foram feitas inúmeras tentativas de gerar clones duplo nocaute para o gene *Tcmsh2*, usando uma construção contendo outro gene de resistência flanqueado pelas mesmas regiões homólogas ao alelo *Tcmsh2*-B. Nenhuma delas obteve sucesso. Em alguns experimentos, surgiram clones duplo-resistentes (G418^R e HYG^R), nos quais foram feitas duas rodadas de transfecção dos parasitos com construções diferentes: uma delas contendo o gene de resistência a higromicina flanqueado por regiões homólogas ao alelo *Tcmsh2*-B e a outra, contendo o gene de resistência a G418 flanqueado regiões homólogas ao alelo *Tcmsh2*-C. Entretanto, análises por PCR mostraram que estes clones possuíam uma cópia de *Tcmsh2* (Figura 13). A deleção do alelo *Tcmsh2*-C é possível, pois obtivemos parasitos viáveis transfectados com a construção contendo o gene de resistência a G418 flanqueado regiões homólogas ao alelo *Tcmsh2*-C (Figura 13). Estes resultados nos levaram a postular que este gene poderia ser essencial em *T. cruzi*.

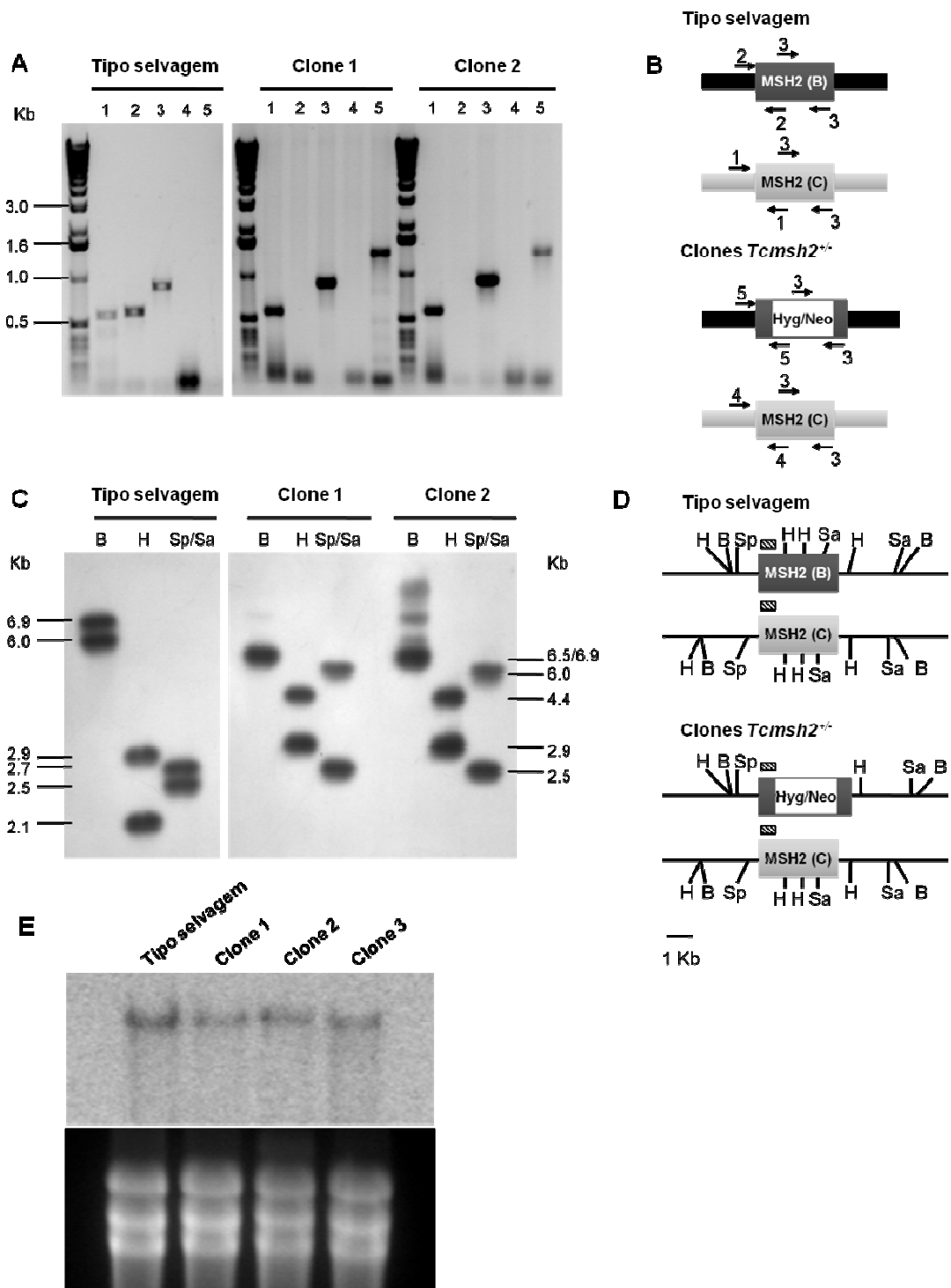


Figura 12: Caracterização dos clones *Tcmsh2*^{+/-} de *T. cruzi*. Em A: gel de agarose de produtos de PCR, obtidos com os iniciadores alelo-C específicos (canaleta 1), alelo-B específicos (canaleta 2) internos aos gene *Tcmsh2* (canaleta 3), iniciador direto alelo-C específico e reverso específico para o gene de resistência (canaleta 4) e iniciador direto alelo-B específico e reverso específico para o gene de resistência

(canaleta 5). Em **B** é mostrado um esquema do locus *Tcmsh2* na cepa CL Brener e os sítios de anelamento de todos os iniciadores usados em **A**. Em **C** é mostrado um *Southern blot* com DNA genômico de epimastigotas de *T. cruzi* (cepa CL Brener) tipo selvagem e de clones *Tcmsh2*^{+/-}. O DNA foi digerido com as endonucleases *Bam*HI (B), *Hind*III (H) e *Sph*I (Sp)/*Sal*I (Sa), hibridizado com a sonda adequada e exposta a um filme de raios-X. Os tamanhos dos fragmentos obtidos após a digestão, bem como a sonda utilizada no experimento (hachurada), está indicada em **D**. Em **E** é mostrado um experimento de *Northern blot*, para verificar o nível de expressão dos transcritos para o gene *Tcmsh2*. O experimento foi realizado com RNA total extraído de epimastigotas em fase exponencial de CL Brener tipo selvagem e clones *Tcmsh2*^{+/-} clones. As amostras foram transferidas para uma membrana de nylon e hibridizada com a mesma sonda indicada em **D**. A porção inferior da figura mostra a visualização por brometo de etídio das bandas correspondentes ao rRNA, para demonstrar a quantidade de RNA aplicado no gel.

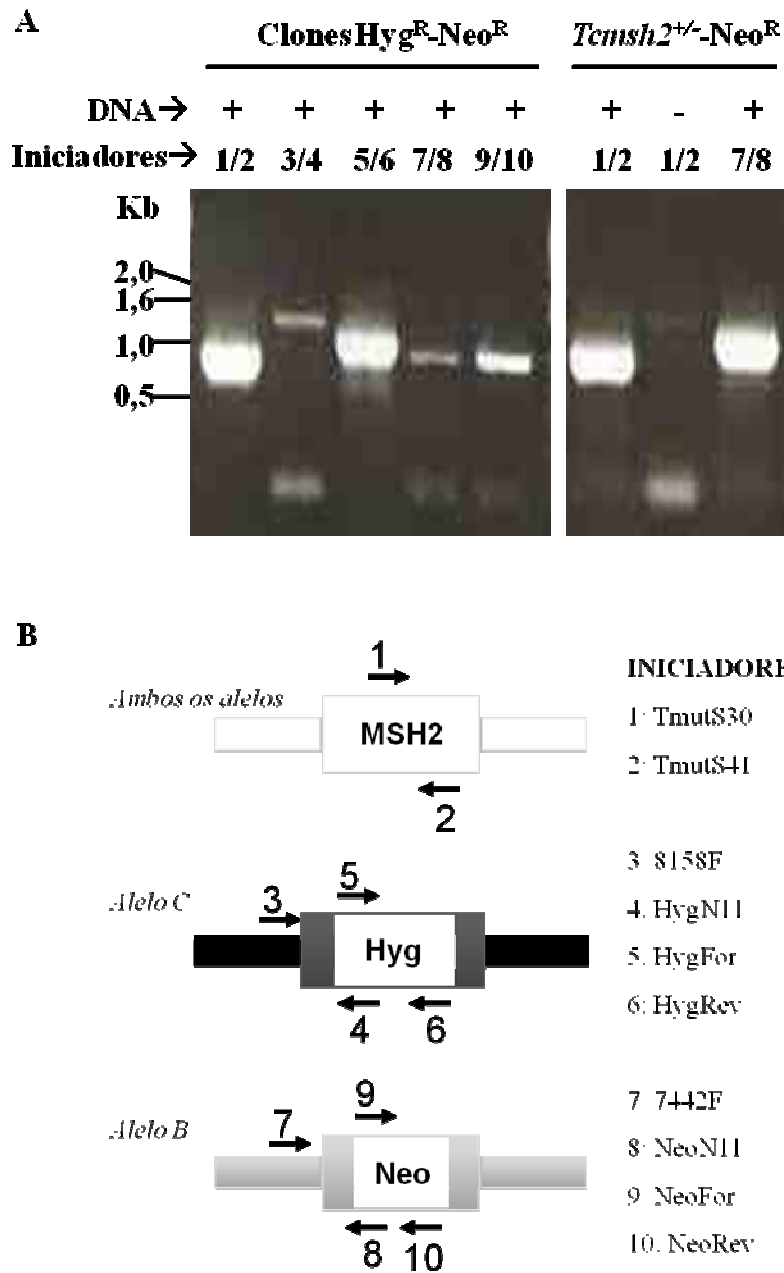


Figura 13: Caracterização dos clones duplo-resistentes e de clones heminocaute para o alelo *Tcmsh2-C* (Neo^R) de *Trypanosoma cruzi*. Em A: gel de agarose de produtos de PCR, obtidos com os iniciadores mostrados em B. Os DNAs usados como molde foram obtidos de clones transfectados com cassetes de resistência à higromicina B e a G418 (clonesHyg^R-Neo^R) ou com o cassete de resistência a G418, flanqueado com regiões específicas para substituição do alelo *Tcmsh2-C* (*Tcmsh2*^{+/-}-Neo^R).

7. Sobrevivência de parasitos CL Brener heminocautes de *Tcmsh2* ao tratamento com agentes genotóxicos

Para analisar o fenótipo resultante da deleção de um dos alelos de *Tcmsh2* em epimastigotas, culturas de epimastigotas do tipo selvagem e clones *Tcmsh2*^{+/-} foram comparadas. Não foram observadas diferenças nas curvas de crescimento entre culturas do tipo selvagem e clones transfectados (dados não mostrados). As taxas de crescimento de parasitos selvagens e clones *Tcmsh2*^{+/-} na presença de diferentes concentrações de cisplatina e de MNNG também foram avaliadas. Como mostrado na Figura 14A, não existem diferenças significativas entre as culturas mutantes após o tratamento com a cisplatina, quando comparadas às culturas selvagens. Entretanto, quando tratados com MNNG, os três clones *Tcmsh2*^{+/-} analisados apresentaram maior susceptibilidade ao tratamento do que parasitos do tipo selvagem (Figura 14B).

O efeito do da deleção de um dos alelos de *Tcmsh2*^{+/-} sobre a resposta a dano oxidativo foi também medido. Parasitos do tipo selvagem e clones *Tcmsh2*^{+/-} foram tratados com H₂O₂ e a taxa de crescimento foi medida após o tratamento, bem como a quantificação de 8-oxoguanina no genoma. Os clones *Tcmsh2*^{+/-} apresentaram uma diminuição na tolerância ao tratamento, quando comparados aos parasitos selvagens (Figura 15A). Observa-se também uma diferença nos níveis basais de acúmulo de 8-oxoG no núcleo e no kDNA das duas culturas analisadas; após o tratamento com H₂O₂, esta diferença se torna mais evidente no kDNA dos três clones mutantes, quando comparados com parasitos do tipo selvagem (Figura 15B). Estes dados sugerem um papel do MMR, ou pelo menos da proteína TcMSH2, na resposta dos parasitos de *T. cruzi* ao dano oxidativo mitocondrial induzido por H₂O₂.

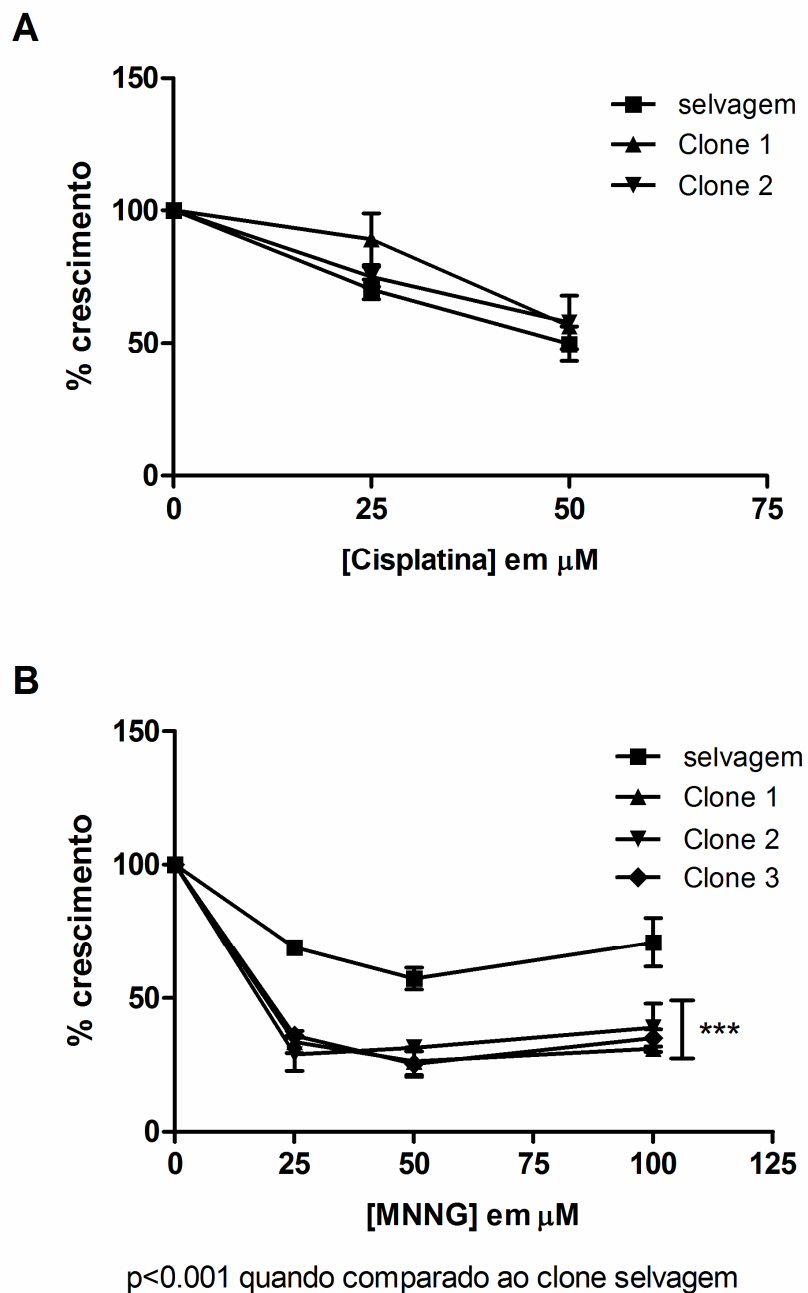


Figura 14: Efeito do tratamento com cisplatina e MNNG no crescimento de clones de CL Brener *Tcmsh2*^{+/-}. Epimastigotas do tipo selvagem e clones *Tcmsh2*^{+/-} em fase exponencial de crescimento foram tratados com 25 e 50 μM de cisplatina (**A**) ou 25, 50 e 100 μM de MNNG (**B**). As células viáveis foram quantificadas após 3 (para o tratamento com MNNG) ou 5 dias (para o tratamento com cisplatina) com eritrosina B. Os resultados são mostrados como médias \pm desvio padrão de 3 experimentos independentes, realizados em triplicata.

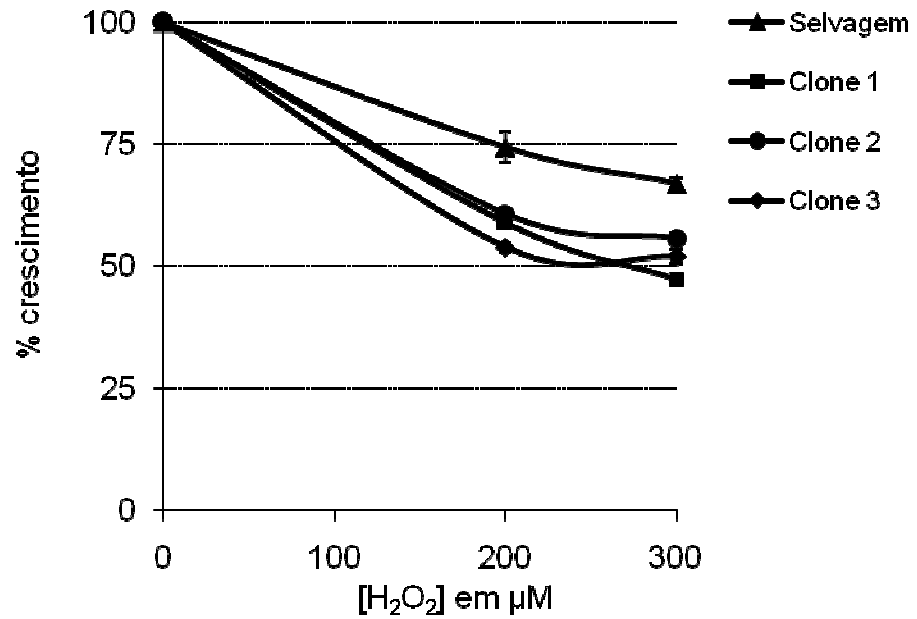
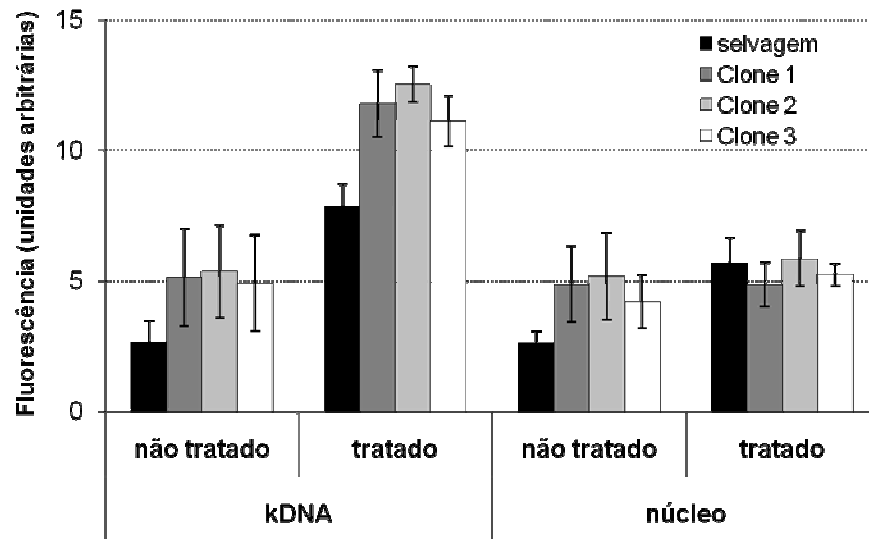
A**B**

Figura 15: Análise da resposta de parasitos CL Brener tipo selvagem e clones *Tcmsh2*^{+/-} ao estresse oxidativo induzido por peróxido de hidrogênio. Em A: Epimastigotas em fase exponencial foram tratados com 200 e 300 μM H₂O₂ conforme descrito em Material e Métodos e as células viáveis foram quantificadas após 5 dias com eritrosina B. Em B: quantificação do acúmulo de 8-oxoguanina em parasitos selvagens e clones *Tcmsh2*^{+/-}.

8. Sobrevivência de parasitos do tipo selvagem e heminocautes para o gene *Tcmsh2* à exposição prolongada com doses subletais de MNNG

A Figura **14B** nos mostra que os clones *Tcmsh2*^{+/-} são mais suscetíveis ao tratamento com MNNG do que parasitos do tipo selvagem. Este fenótipo está em discordância com estudos anteriores, que mostram um aumento na tolerância ao tratamento com MNNG em células mutantes para o gene *msh2* (Bell et al., 2004; Claij and Te Riele, 2002; Humbert et al., 1999). Assim, para investigar se a deleção de um dos alelos de *Tcmsh2* poderia afetar não só o MMR, mas também o metabolismo oxidativo das células, os parasitos do tipo selvagem e clones *Tcmsh2*^{+/-} foram tratados com uma dose sub-letal de MNNG (5 µM) e após um período de 15 dias os parasitos foram coletados e submetidos a ensaios de quantificação do consumo de O₂ e da produção de H₂O₂. Conforme mostrado na Figura **16A**, os parasitos heminocaute para o gene *Tcmsh2* não apresentam diferenças no consumo de O₂ quando comparados aos parasitos do tipo selvagem. Entretanto, há um aumento neste consumo quando as células são mantidas na presença de MNNG. Por outro lado, embora os parasitos heminocaute para o gene *Tcmsh2* apresentem níveis maiores de produção basal de H₂O₂, estes diminuem em relação aos parasitos do tipo selvagem quando submetidos à exposição prolongada ao agente metilante (Figura **16B**). Estes resultados sugerem uma correlação inversa entre o consumo de oxigênio e produção de H₂O₂ nos parasitos analisados e um papel importante da proteína TcMSH2 na resposta ao estresse oxidativo.

9. Localização subcelular da proteína TcMSH2 em *T. cruzi*

Os resultados obtidos até o momento fornecem evidências indicando que a proteína TcMSH2 poderia estar envolvida na resposta ao estresse oxidativo em *T. cruzi*. Além disso, os resultados obtidos nos tratamentos com H₂O₂ (Figura **15**), bem como nas análises do metabolismo oxidativo das células estudadas (Figura **16**) sugerem ainda que esta função possa envolver reparo de DNA mitocondrial no parasito. Para testar esta hipótese, decidimos investigar a localização subcelular da proteína TcMSH2 em *T. cruzi*. Foi construído um vetor contendo a sequência codificadora completa do gene *Tcmsh2*, fusionada ao GFP na sua porção C-terminal

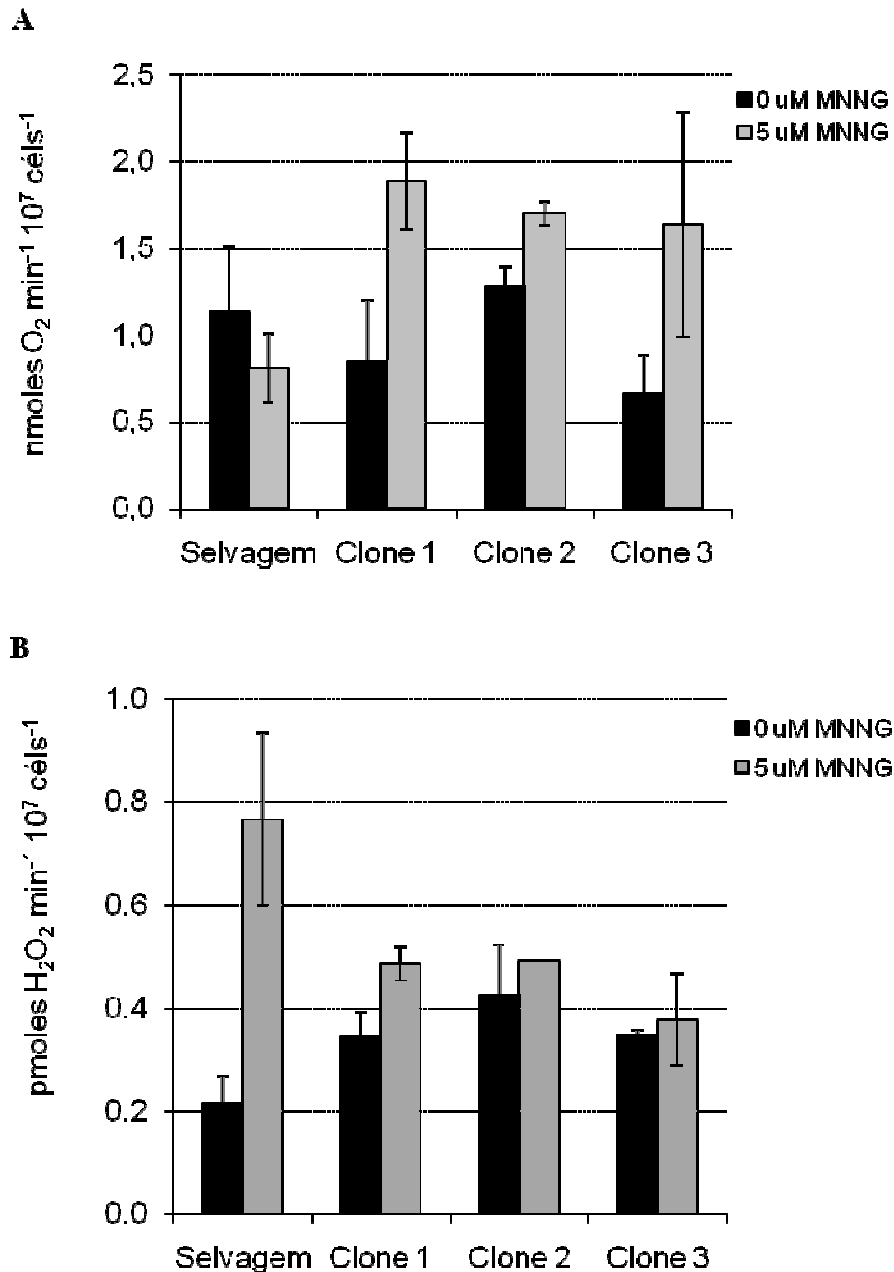


Figura 16: Análise do metabolismo oxidativo dos parasitos CL Brener do tipo selvagem e clones *Tcmsh2*^{+/-} após exposição prolongada a uma dose sub-letal de MNNG. O consumo de oxigênio basal pelos parasitos mantidos por 15 dias na presença ou ausência de 5 μM MNNG é mostrado em **A**. As células foram incubadas em meio de respiração (vide Material e Métodos) e os dados registrados em um oxígrafo. Em **B** é mostrada a produção basal de H₂O₂ pelos parasitos mantidos por 15 dias na presença ou ausência de 5 μM MNNG, medida por fluorimetria, com excitação a 563 nm e emissão a 587 nm. Todos os experimentos foram repetidos três vezes, em duplicata.

(na mesma janela de leitura). Esta construção foi utilizada na transfecção de epimastigotas de *T. cruzi* por eletroporação, que foram visualizados no microscópio 24 horas após a transfecção. Conforme mostrado na Figura **17**, enquanto a expressão da proteína GFP está distribuída homogeneamente em todo o corpo do parasita, a proteína TcMSH2::GFP mostrou uma distribuição predominantemente citoplasmática. Para verificar se este padrão de distribuição não sofreu influência da fusão com GFP, a localização de TcMSH2 foi feita também por imunofluorescência, utilizando anticorpos policlonais gerados contra a proteína recombinante TcMSH2::His. Os resultados obtidos corroboram com o que foi observado nos estudos de localização da proteína em fusão a GFP, indicando novamente a localização citoplasmática da proteína (Figura **18**).

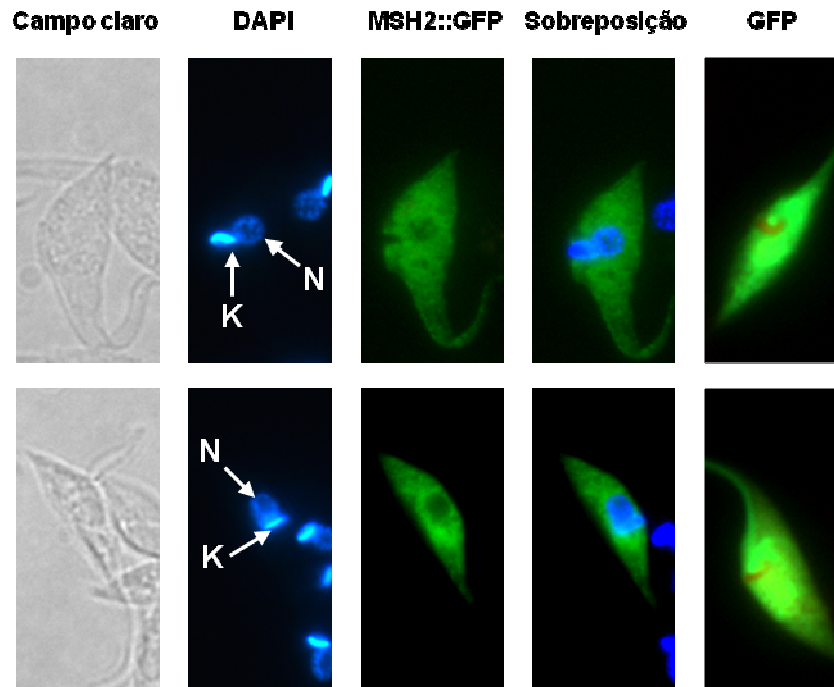


Figura 17: Localização subcelular da proteína TcMSH2::GFP em CL Brener.

Epimastigotas em fase exponencial foram transfectados com os vetores de expressão episomal pTREX-Neo-GFP ou pTREX-Neo-TcMSH2::GFP e após 24 horas de transfecção, os parasitos foram fixados e analisados em microscópio de fluorescência. As proteínas GFP e TcMSH2::GFP são visualizadas como uma fluorescência verde, e os DNAs genômico e do cinetoplasto foram contra-corados com DAPI (em parasitos expressando a proteína TcMSH2::GFP) ou iodeto de propídio (em parasitos expressando somente a proteína GFP). A posição do núcleo (N) e do cinetoplasto (K) está destacada nas imagens com coloração por DAPI.

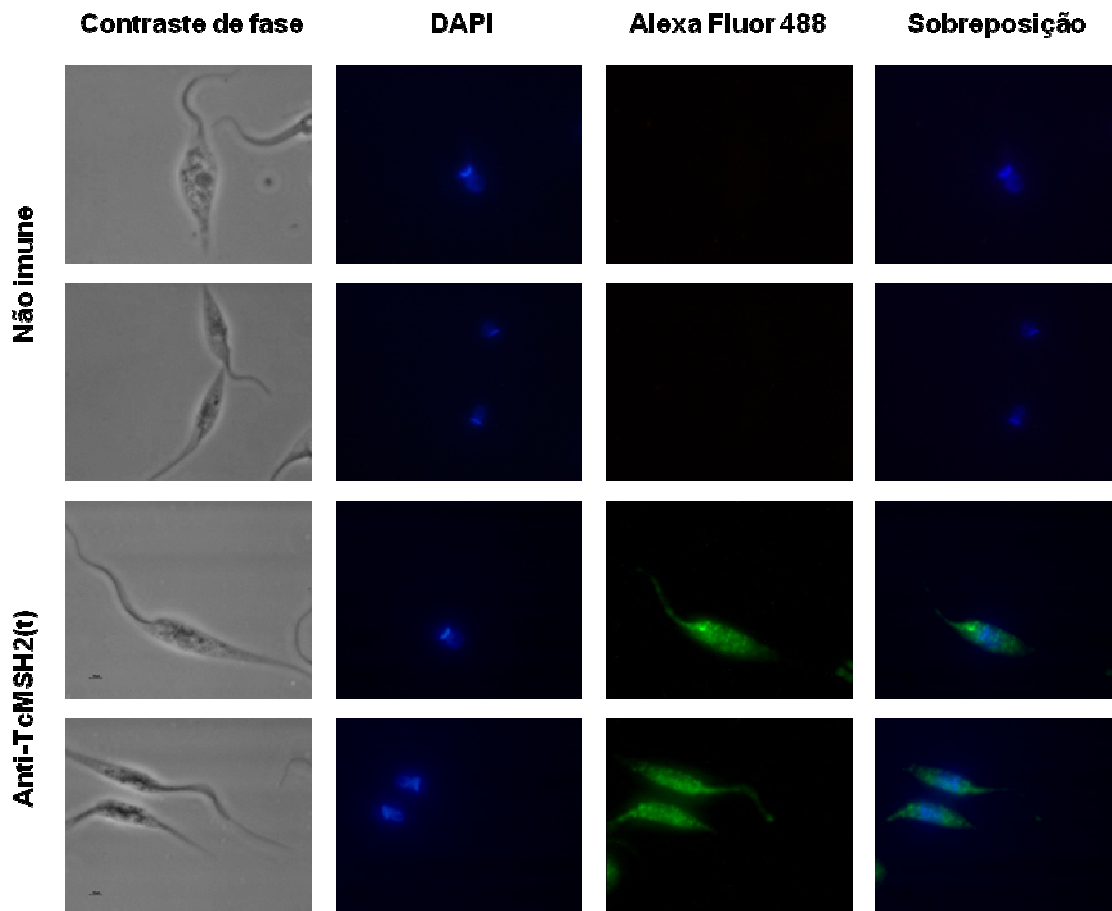


Figura 18: Localização subcelular da proteína TcMSH2 em epimastigotas de *T. cruzi* por imunofluorescência indireta. Na figura, estão indicadas as imagens obtidas por contraste de fase, por coloração com DAPI e por incubação com anticorpos policlonais não imune ou anti-TcMSH2(t). O anticorpo policlonal foi revelado com IgG anti-camundongo conjugado a Alexa Fluor 488. O núcleo e o cinetoplasto foram contra-corados com DAPI.

10. Geração de parasitos superexpressores do gene *Tcmsh2*

O estudo da função de um gene pode ser investigado através de diversas formas. No presente trabalho, procuramos inferir a função do gene *Tcmsh2* no MMR e na resposta ao estresse oxidativo em *T. cruzi* por meio de ensaios fenotípicos em parasitos heminocutes para o gene *Tcmsh2*, bem como analisando sua localização subcelular no parasito. Considerando que a deleção completa do locus não foi possível, outros mecanismos podem ser utilizados, como por exemplo, estratégias para a superexpressão do gene alvo. Assim, como o objetivo de se obter clones superexpressores do gene *Tcmsh2* em *T. cruzi*, foram feitas construções nas quais uma cópia deste gene pudesse ser inserida no locus de β -tubulina, sobre o controle transcricional de regiões regulatórias identificadas através de análises *in silico* [descritas a seguir no item **12** e em (Campos et al., 2008)]. Parasitos do clone CL Brener foram transfectados com este DNA (pROCK-TcMSH2-Neo) e após 30 dias de seleção em meio contendo G418, as culturas foram clonadas. A Figura **19A** mostra a caracterização por PCR de um clone representativo dos parasitos transfectados com o plasmídeo pROCK-TcMSH2-Neo, identificando a presença do gene de resistência a G418. Foram feitos experimentos de *Northern blot* com o objetivo de verificar o nível dos transcritos de *Tcmsh2* nos clones transfectados. Os resultados mostram que houve um pequeno aumento na banda de 3,3 Kb, referente ao transcrito para o gene *Tcmsh2* endógeno; entretanto, nos clones transfectados era esperada uma banda adicional de 3,6 Kb, correspondente aos transcritos para o gene *Tcmsh2* derivados do plasmídeo pROCK-TcMSH2-Neo (Figura **19B**). Experimentos de *Western blot* a partir de extratos protéicos de clones nos quais foi detectada a presença do gene *Neo* (marcador seletivo da transfecção) mostram um pequeno aumento na expressão da banda de 75 KDa nos clones transfectados (Figura **20**). Entretanto, além da banda de peso molecular de aproximadamente 75 KDa, presente na canaleta referente ao parasito controle e nos clones transfectados com o vetor pROCK-TcMSH2-Neo (vista anteriormente em outros experimentos de *Western blot*), é observada uma banda de peso molecular de aproximadamente 100 KDa, ausente no parasito controle.

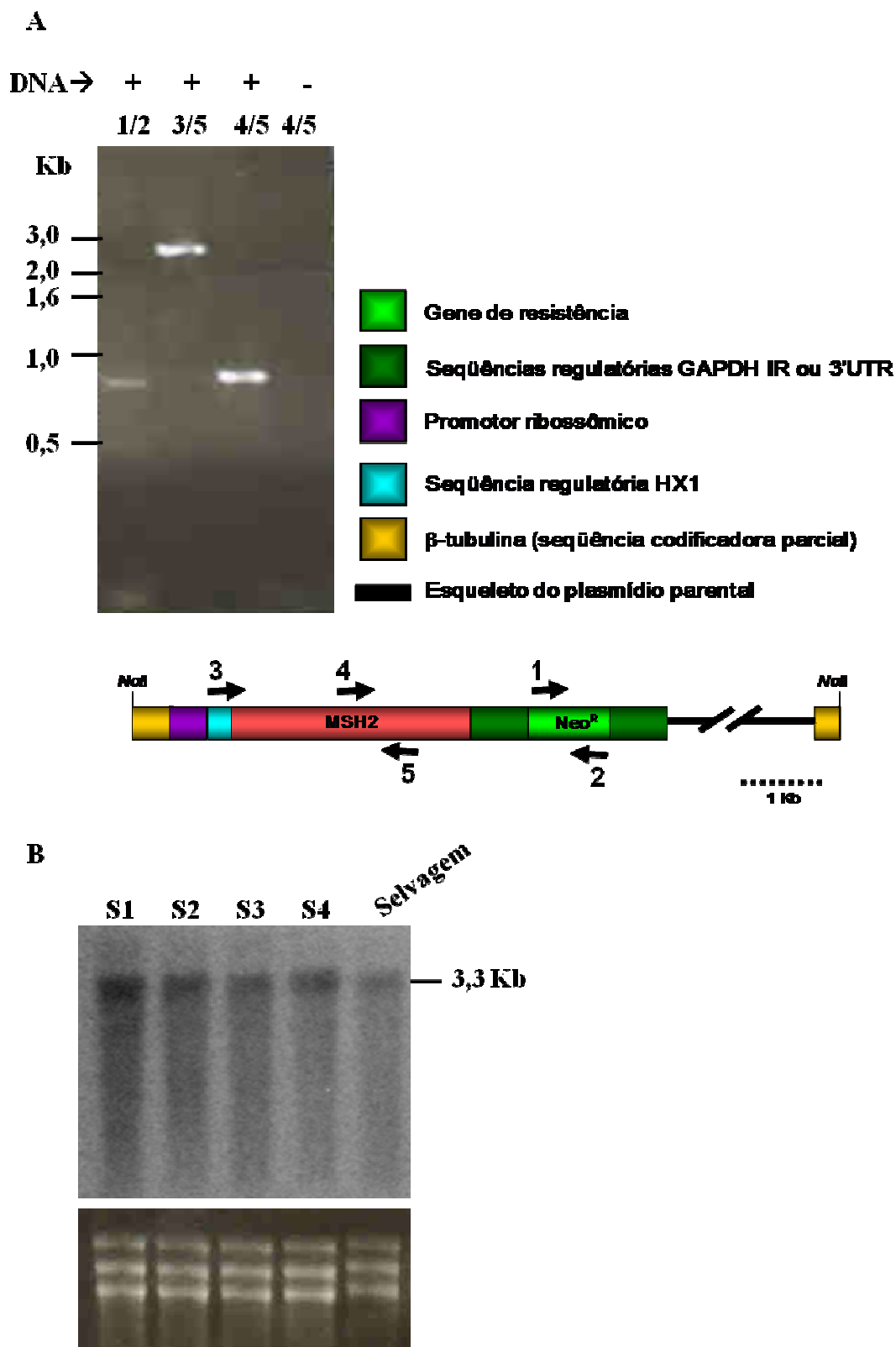


Figura 19: Caracterização dos clones superexpressores do gene *Tcmsh2* de *T. cruzi*. Em **A**: gel de agarose de produtos de PCR, obtidos a partir do DNA de um clone transfectado com o vetor pROCK-TcMSH2-Neo. Os sítios de anelamento dos iniciadores utilizados estão mostrados no esquema acima. Em **B** é mostrado um

experimento de *Northern blot*, para verificar o nível de expressão dos transcritos para o gene *Tcmsh2*. O experimento foi realizado com RNA total extraído de epimastigotas em fase exponencial de CL Brener tipo selvagem e clones *Tcmsh2*^{+/-} clones. As amostras foram transferidas para uma membrana de nylon e hibridizadas com a mesma sonda utilizada na Figura **12E**. A porção inferior da figura mostra a visualização por brometo de etídio das bandas correspondentes ao rRNA, para demonstrar a quantidade de RNA aplicado no gel.

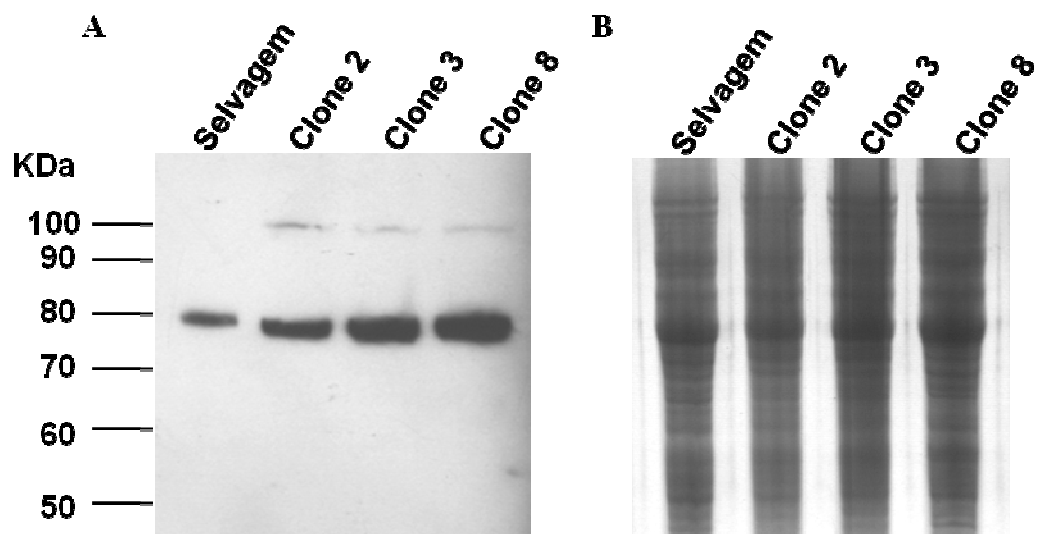


Figura 20: Análise da expressão de TcMSH2 em epimastigotas de CL Brener transfectados com o vetor controle (pROCK-Neo) e com o vetor pROCK-TcMSH2-Neo, por ensaio de Western blot. Os parasitos foram coletados e preparados conforme descrito em Materiais e Métodos. Os extratos (5×10^6 células/canaleta) foram submetidos à eletroforese em gel de poliacrilamida, transferidos para membranas de nitrocelulose e incubados com anticorpo primário policlonal produzido contra a proteína recombinante TcMSH2::His, na diluição 1:500, conforme mostrado em **A**. Em **B** é mostrado um gel de acrilamida a 12,5% corado com comassie, com as mesmas amostras aplicadas no experimento mostrado em **A**. As amostras aplicadas em cada gel estão identificadas no topo das canaletas.

11. Análises dos padrões de processamento do mRNA em *T. cruzi*

Para investigar a influência da proteína TcMSH2 na resposta das diferentes cepas de *T. cruzi* a agentes genotóxicos, bem como na geração de variabilidade genética no parasito, foi inicialmente escolhida a estratégia de deleção do gene *Tcmsh2*, através de experimentos de transfecção. Esta técnica é uma das ferramentas mais utilizadas na manipulação gênica para causar uma alteração fenotípica em tripanossomatídeos. No processo de construção de vetores contendo os genes de resistência a antibióticos, para a deleção dos alelos de *Tcmsh2* por recombinação homóloga, foram adicionadas seqüências regulatórias flanqueadoras, como por exemplo, as regiões intergênica e 3'UTR do gene de GAPDH ou 5' e 3'UTR de TcMSH2. A adição de seqüências derivadas do gene de GAPDH é uma metodologia utilizada por vários autores (DaRocha et al., 2004b; Kelly et al., 1992; Nozaki and Cross, 1995), pois é sabido que nestas regiões estão presentes os sinais que influenciarão a expressão gênica da seqüência codificadora de interesse. Entretanto, nem todos os sinais estão caracterizados, tampouco se os níveis de expressão dos genes adjacentes são dependentes do tipo de sinal presente nestas regiões. Essa é uma questão muito importante e merece ser investigada, considerando que estes sinais são cruciais para se obter a expressão de um gene exógeno em *T. cruzi*. Como descrito anteriormente, uma segunda estratégia utilizada para se estudar o papel do gene *Tcmsh2* foi a super-expressão da proteína através da transfecção desse gene clonado em um vetor que possibilita a expressão estável de genes em forma epimastigotas. A fim de fornecer maiores informações sobre quais são os sinais importantes para a expressão de um gene no parasito, foi realizada uma análise *in silico*, onde se procurou investigar os elementos necessários para o processamento, através de *trans-splicing* e poliadenilação, dos mRNAs em *T. cruzi*.

Como representado na Figura 3, na sessão de Material e Métodos, todos os cDNAs e ESTs disponíveis nos bancos de dados e que continham a seqüência SL e/ou cauda de poli-A foram mapeados no genoma do *T. cruzi*. Paralelamente, foram identificados motivos de polipirimidina, ou poli(Y) presentes nas regiões intergênicas, definidos como seqüências de ≥ 9 pirimidinas, permitindo uma purina interna. Desta maneira foi possível identificar uma série de parâmetros como os tamanhos das regiões UTR assim como as distâncias entre o sítio acceptor da

sequência SL e da cauda poli e os motivos de polipirimidina. As análises mostraram que o tamanho mediano de 5'UTR (região 5' não traduzida) e de 3'UTR (região 5' não traduzida) é de 35 e 137 nucleotídeos, respectivamente. A distribuição dos tamanhos das UTRs é mostrada nas Figuras **21A** e **21B**. Foi observado também que as medianas que correspondem às distâncias entre a poli(Y) e o sítio acceptor de SL ou de poli-A correspondem a 18 e 40 nucleotídeos, respectivamente (Figuras **21C** e **21D**). Foi também medido o tamanho mediano dos motivos ricos em polipirimidinas (poli(Y)s), bem como a sua composição (Figuras **21E** e **21F**). Conforme mostrado na Tabela **2**, os resultados obtidos indicam que o tamanho médio das UTRs é significativamente menor em *T. cruzi* do que em *T. brucei*, bem como as distâncias médias entre os motivos poli(Y) e os sítios de adição de SL e poli(A). Além disso, análises da distribuição das poliYs mostra uma frequência mais pontual na região imediatamente superior ao sítio de adição de SL imediatamente inferior ao sítio de adição de A_(n), sendo que, para este último, são observados dois picos principais de frequência de poliYs (Figuras **22A** e **22B**). O gráfico da frequência de todos os nucleotídeos na região imediatamente inferior ao sítio de adição de A_(n) apresenta dois picos de predominância de timinas, coincidente ao observado para as poliYs (Figura **22C**). O esquema representativo dos padrões de processamento do mRNA em *T. cruzi*, obtido através das análises feitas neste trabalho, pode ser visualizado na Figura **23**.

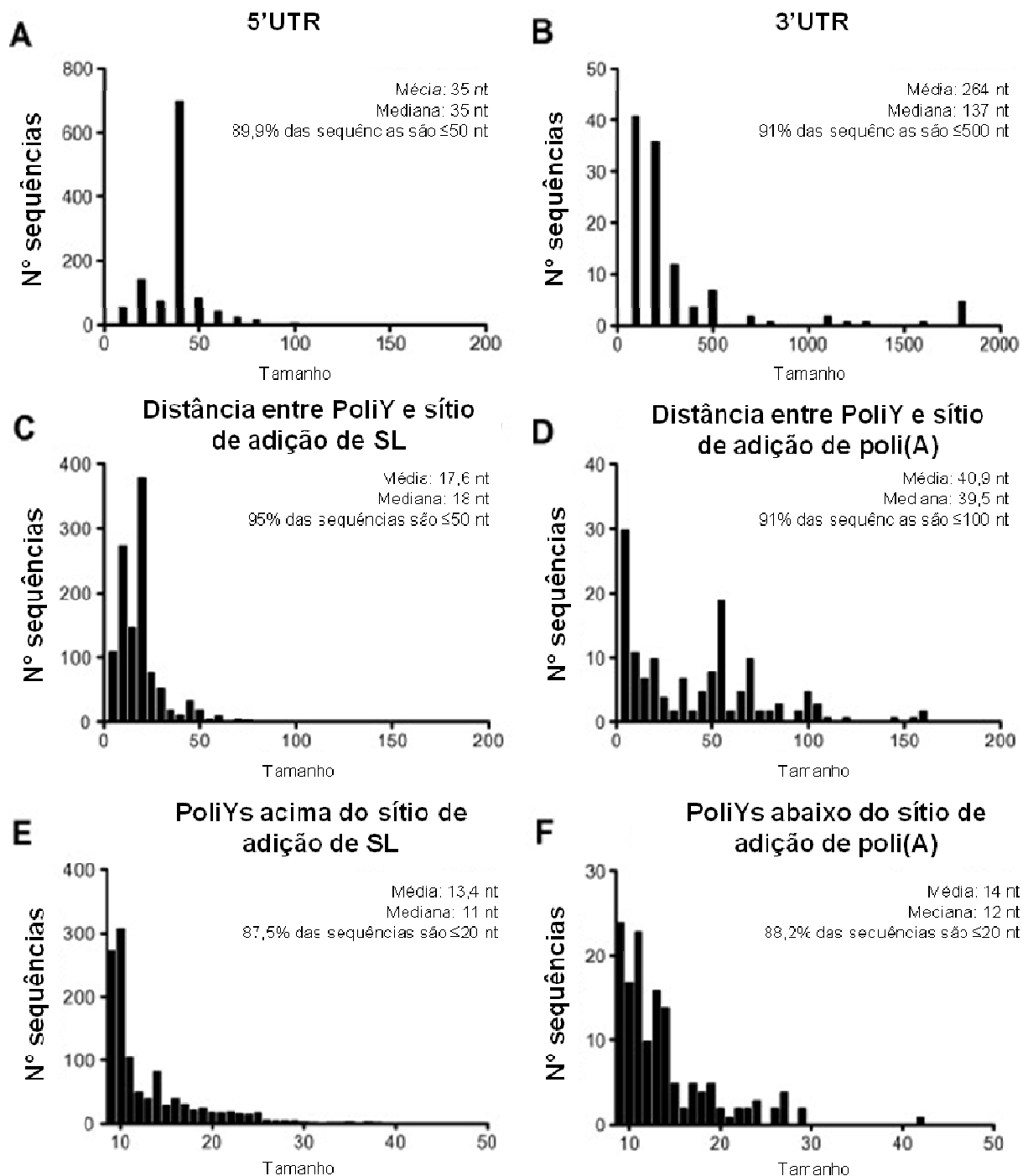


Figura 21: Tamanhos das regiões 5' e 3'UTR (não traduzidas) e suas distâncias em relação aos sinais de processamento dos mRNAs de *Trypanosoma cruzi*. Os histogramas **A** e **B** mostram os tamanhos observados para as 1189 5'UTRs e as 149 3'UTRs, respectivamente; **C** e **D** mostram as distâncias entre o motivo de polipirimidinas (poliY) mais próxima e o sítio acceptor de *Spliced Leader* (SL) imediatamente posterior (**C**) e o acceptor de poli(A) imediatamente anterior (**D**); **E** e **F** mostram os tamanhos observados das poliYs, acima do sítio acceptor de SL (**E**) e abaixo do sítio acceptor de poli(A) (**F**).

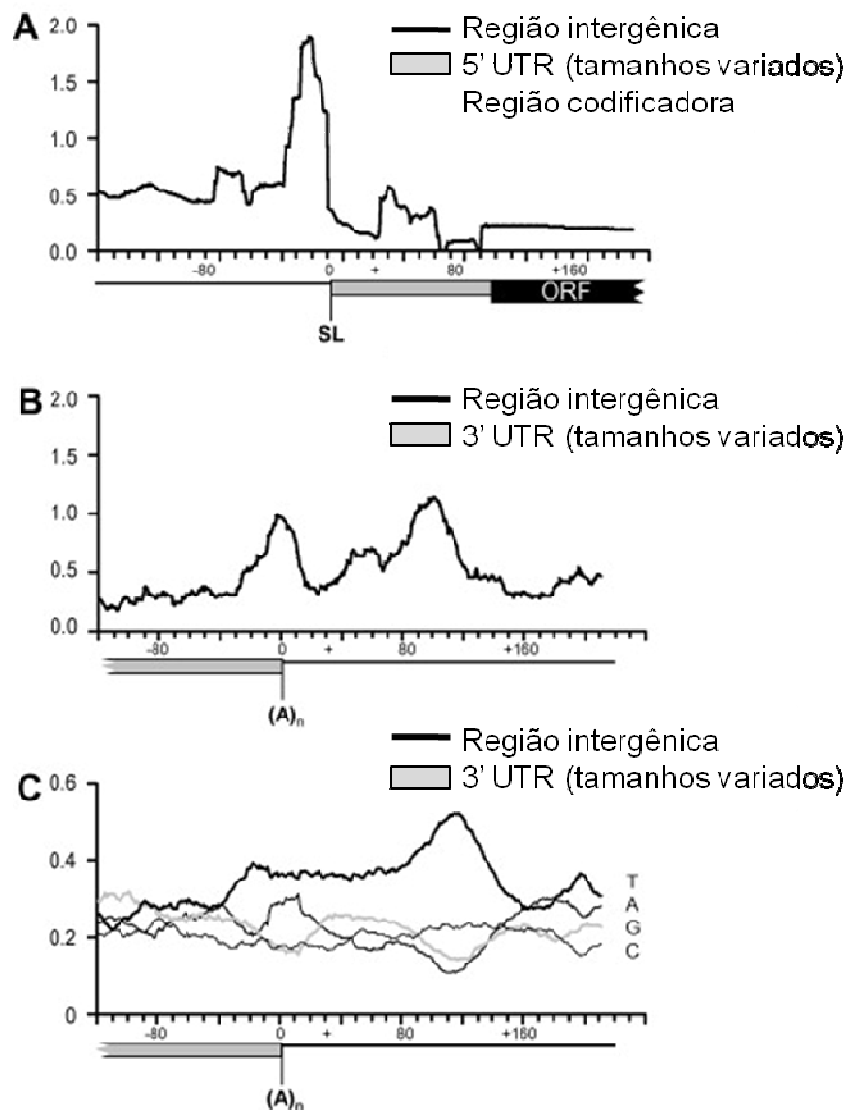


Figura 22: Frequências dos motivos de polipirimidinas (poliYs) e de cada um dos nucleotídeos, nas regiões adjacentes aos sítios de processamento do mRNA de *Trypanosoma cruzi*. Foram analisados os motivos de poliY que flanqueiam os sítios aceptores de *Spliced Leader* (SL) e de poliadenilação ((A)_n) e os resultados são mostrados em **A** e **B**, respectivamente. O mapeamento foi feito a partir de uma janela de busca de 20 nucleotídeos, que era deslocada 1 nucleotídeo para a direita (a partir do sítio aceptor de (A)_n) ou para a esquerda (a partir do sítio aceptor de SL). O n° de poliYs identificadas foi então dividido pelo número total de sequências dentro da área de abrangência da janela de 20 nucleotídeos. Nestas análises, todas as sequências identificadas como pseudogenes ou genes truncados foram excluídas. Em **C** é mostrada a composição de nucleotídeos das regiões que flanqueiam os sítios aceptores de (A)_n.

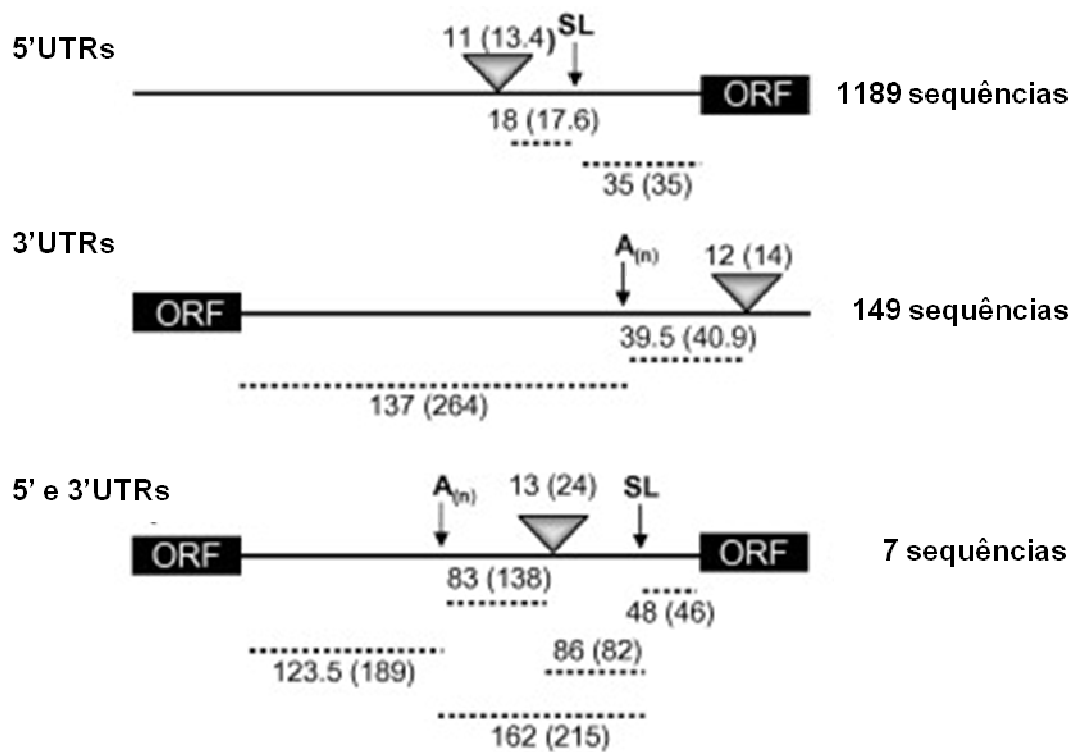


Figura 23: Padrões de processamento do mRNA em *Trypanosoma cruzi*. Foram mapeadas sequências genômicas contendo sítios de adição de *Spliced Leader* (SL) (5' UTRs) e poliadenilação ($A_{(n)}$) (3' UTRs), bem como as sequências que continham ambos os sítios de adição de SL e $A_{(n)}$ derivadas do mesmo cDNA ou EST (5' e 3' UTRs). Números sob as linhas pontilhadas indicam os tamanhos medianos das regiões 5' UTR, 3' UTR e regiões intergênicas. As medianas das distâncias entre o motivo de polipirimidinas (poliYs) e o sítio acceptor de SL, bem como a distância entre o motivo de poliYs ao sítio acceptor de $A_{(n)}$ estão indicadas sobre as linhas pontilhadas. Os triângulos cinza representam tratos de poliYs com números no topo correspondendo aos seus tamanhos medianos. Em cada caso, os valores médios estão entre parênteses. Retângulos pretos simbolizam janelas abertas de leitura (ou *open reading frames*, ORFs); SL com uma seta, sítio acceptor de SL; $A_{(n)}$ com uma seta, sítio acceptor de poliadenilação.

Tabela 2: Comparação entre os padrões de processamento do mRNA entre *Trypanosoma cruzi* e *Trypanosoma brucei*.

Parâmetros	<i>T. cruzi</i>	<i>T. brucei</i>	
Tamanho médio da região codificadora	1457	1511	
Tamanho médio da região inter-codificadora	561	721	
Tamanho médio de 5' UTR	35	68	
Tamanho médio de 3' UTR	304	587	
Média entre a PoliYs e o sítio acceptor de SL	18	24	
Média entre a PoliYs e o sítio acceptor de poli(A)	40	120	
Tamanho médio de PoliYs nas seqüências de 5'UTR	13	13	
Tamanho médio de poliYs nas seqüências de 3'UTR	14	13	
Composição média das poliYs (seqüências 5'UTR)	%T	59,3	59
	%C	33,3	38
	%AG	7,4	3
Composição média das poliYs (seqüências 3'UTR)	%T	66,3	64
	%C	27	34
	%AG	6,6	1

1. O MMR e o papel da proteína TcMSH2

O MMR corrige erros de pareamentos que surgem durante o processo de replicação, bem como contribui para o controle dos níveis endógenos de bases oxidadas presentes no genoma, como por exemplo a 8-oxoG (Slupphaug et al., 2003). No táxon *T. cruzi*, a proteína TcMSH2 (componente central a via de MMR) é codificada por três tipos distintos de alelos que podem ser correlacionados com pelo menos três das linhagens principais de *T. cruzi*. Os níveis distintos de variabilidade genética observados entre as cepas pertencentes a cada haplogrupo de *T. cruzi* (TcMSH2-A, -B e -C) (Augusto-Pinto et al., 2003; de Freitas et al., 2006) nos levaram a investigar a atividade desta proteína componente do MMR e seu envolvimento na resposta de diferentes cepas do parasito a vários tipos de agentes genotóxicos. Nossos resultados levam a duas conclusões principais: (a) os níveis de eficiência do MMR são diferentes entre cepas pertencentes às linhagens *T. cruzi* I e II e o grau de diversidade genética está inversamente correlacionado com os níveis de eficiência de um dos componentes desta via de reparo; e (b) além do seu papel na via de MMR, a proteína MSH2 em *T. cruzi* parece estar envolvida em uma via de proteção do DNA mitocondrial contra o dano oxidativo que parece não envolver outros componentes do MMR (Machado-Silva et al., 2008).

O envolvimento do MMR com o aumento na taxa de mutações (e consequentemente na geração de diversidade) foi inicialmente descrito em estudos que correlacionaram o câncer em humanos com erros na replicação de DNA, nos quais foi observado um aumento na frequência de instabilidade de microssatélites em linhagens celulares derivadas de pacientes com câncer colorretal não polipótico hereditário ou HNPCC (Aaltonen et al., 1993; Ionov et al., 1993). Mutações germinativas em quatro genes envolvidos no MMR (hMLH1, hPMS1, hPMS2 e hMSH2) estão correlacionadas ao HNPCC [revisito por (Peltomaki, 2005)]. Além disso, tem sido relatada a correlação entre perda da função de MMR em células tumorais e a resistência ao tratamento com cisplatina, agentes metilantes e doxorubicina (Aebi et al., 1996; Brown et al., 1997).

O envolvimento do MMR na geração de diversidade genética e na resistência a drogas já foi também descrito em outros organismos patogênicos, incluindo bactérias e outros protozoários parasitos. Estudos em bactérias hipermutáveis de diversas espécies e isoladas de linhagens patogênicas como, por exemplo, *Escherichia coli* e *Salmonella typhimurium*, mostraram que a inativação de *mutS* está relacionada a um fenótipo mutagênico, com um aumento na taxa de mutações para a resistência à rifampicina entre 10 a 1000 vezes (Hall and Henderson-Begg, 2006). O tratamento de culturas de *Toxoplasma gondii* com monesina (uma droga efetiva contra parasitos do filo Apicomplexa) leva ao surgimento de linhagens mutantes para o gene *msh1* (envolvido em reparo mitocondrial). Estas linhagens são também resistentes ao agente alquilante metilnitrosouréia (Garrison and Arrizabalaga, 2009).

Os parasitos do gênero *Plasmodium* são os únicos eucariotos conhecidos que possuem em seu genoma dois genes homólogos de *msh2*. Ao serem comparados os parasitos do tipo selvagem e duplo nocautes para um dos loci de *Pfmsh2* (denominados *Pfmsh2-1* e *Pfmsh2-2*), ambos analisados após a passagem pelo vetor invertebrado (mosquito), um aumento na frequência de parasitos resistentes ao 5-fluoroorotato foi observado nas linhagens mutantes. Como esse aumento foi pequeno, especula-se que a ausência de um fenótipo mais acentuado pode ser decorrente de uma sobreposição de funções entre os dois homólogos de *msh2* em *P. falciparum* (Bethke et al., 2007).

O nosso trabalho mostrou que cepas pertencentes às linhagens *T. cruzi* I e II respondem diferentemente ao tratamento com agentes genotóxicos, tais como cisplatina e MNNG. Os resultados descritos nesta tese, somados a estudos anteriores nos quais diferentes células eucarióticas mutantes para o MMR foram testadas (Stojic et al., 2004), sugerem fortemente que o aumento na tolerância à cisplatina e ao MNNG em parasitos da linhagem *T. cruzi* II está relacionado a um MMR menos eficiente. Além disso, as diferenças observadas entre as cepas quanto à resposta aos agentes genotóxicos não podem ser uma consequência de diferenças nos níveis de expressão do gene *Tcmsh2*, considerando que em todas as cepas analisadas, os níveis dos transcritos e da proteína TcMSH2 foram similares. Assim, pelo fato das cepas de ambas as linhagens apresentarem distintas isoformas da proteína TcMSH2, estabelecemos a hipótese de que os polimorfismos presentes nas sequências de

aminoácidos da proteína TcMSH2 poderiam determinar, pelo menos em parte, as diferenças na resposta aos agentes genotóxicos.

Estudos adicionais realizados pelo nosso grupo corroboram esta hipótese: foram observadas diferenças na atividade ATPásica entre as proteínas recombinantes TcMSH2-A (derivada de Col1.7G2) e TcMSH2-C (derivada de CL Brener) e essa atividade foi inibida na presença de cádmio (Machado-Silva, 2008), conforme observado em *Saccharomyces cerevisiae* (Banerjee and Flores-Rozas, 2005). A atividade de MMR depende da capacidade ATPásica dos heterocomplexos eucarióticos MutS α (formado pelas proteínas MSH2 e MSH6) e MutS β (formado pelas proteínas MSH2 e MSH3). Entretanto, é importante ressaltar que variações na eficiência de outros componentes da via de MMR, bem como de componentes de outras vias de reparo de DNA, podem também estar envolvidos na modulação da eficiência do MMR em diferentes cepas de *T. cruzi*.

Para investigar o papel da proteína TcMSH2 na eficiência do MMR em diferentes linhagens de *T. cruzi*, decidimos gerar parasitos mutantes para o gene *Tcmsh2*. Nosso objetivo inicial era realizar a complementação de parasitos mutantes *Tcmsh2*^{-/-} das linhagens *T. cruzi* I ou II com os genes codificadores das isoformas TcMSH2-A e TcMSH2-C. De maneira inesperada, todas as estratégias utilizadas para a deleção dos dois alelos do locus *Tcmsh2* não obtiveram sucesso, mesmo utilizando construções específicas para cada um desses alelos presentes no genoma de CL Brener. Inicialmente, utilizamos uma sequência referente ao alelo *Tcmsh2*-B como base para as construções. A região correspondente à utilizada na construção dos cassetes para a deleção do gene *Tcmsh2*, quando comparada ao outro alelo (*Tcmsh2*-C), apresenta uma divergência de apenas 3,6%; entretanto, isto pode ter sido o suficiente para impedir a recombinação do fragmento contendo sequências específicas do alelo *Tcmsh2*-B no alelo C. Resultados semelhantes haviam sido observados em *T. brucei* (Bell and McCulloch, 2003). Por outro lado, quando foi utilizada uma construção baseada no alelo *Tcmsh2*-C na transfecção de parasitos do tipo selvagem, foi possível a obtenção de clones *Tcmsh2*^{+/-} para o alelo C (Figura 13A), o que sugere que não existe impedimento para a deleção específica deste alelo. Posteriormente, quando parasitos resistentes a higromicina B (obtidos através da inserção do gene de resistência no alelo B) foram transfectados com a construção C-específica, clones duplo-resistentes – capazes de crescer em meio LIT na presença de

G418 e higromicina B – foram obtidos. Entretanto, análises baseadas em ensaios de PCR, utilizando iniciadores específicos para cada alelo de *Tcmsh2*, mostraram que os clones ainda apresentavam uma cópia do gene de *Tcmsh2*. Esta impossibilidade de gerar duplos nocautes não foi observada em outros eucariotos estudados até então, como o *T. brucei*. Em *T. brucei*, no qual mutantes *Tbmsh2*^{-/-} são viáveis, estes apresentam uma maior taxa de mutações em *loci* de microssatélites, aumentada tolerância ao MNNG e maior frequência de recombinação homóloga (Bell et al., 2004; Bell and McCulloch, 2003). A deleção do gene *Tcmsh2* em *T. cruzi* pode ter levado ao surgimento de linhagens aneuplóides (originadas após ampliações, rearranjos gênicos ou variações na ploidia), com uma ou mais cópias adicionais do gene *Tcmsh2*. Este fenômeno já foi observado em *Leishmania major* após a deleção do gene DHFR-TS (dihidrofolato redutase-timidilato sintase), o que levou os autores à hipótese de que o gene DHFR-TS em *L. major* possa ser essencial (Cruz et al., 1993; Gueiros-Filho and Beverley, 1996). Neste contexto, a viabilidade de clones de *T. cruzi* duplo-resistentes (a G418 e higromicina B) só teria sido possível se os parasitos alterassem o número de cópias do gene *Tcmsh2*, mantendo uma ou mais cópias adicionais.

Para investigar se o gene *Tcmsh2* em *T. cruzi* é essencial, experimentos estão em andamento com o objetivo de gerar linhagens de *Tcmsh2*^{-/-}, através de uma etapa de transfecção intermediária entre a deleção do primeiro e segundo alelos. Para isso, foi utilizado o vetor pROCK-GFP-Neo (DaRocha et al., 2004b). Este vetor contém um cassete de expressão flanqueado por sequências de β -tubulina e o gene *Tcmsh2* pode ser inserido no lugar de *GFP* por dupla digestão com as endonucleases *XbaI/XhoI*. Para facilitar a recombinação homóloga das sequências de β -tubulina, o plasmídeo é linearizado com *NotI*. Se a hipótese de que o gene *Tcmsh2* é essencial estiver correta, a expressão da proteína TcMSH2 a partir de um outro locus nos permitirá a deleção dos dois alelos no locus original. Entretanto, o sucesso desta estratégia é dependente da geração de parasitos viáveis resistentes a três drogas diferentes e, para este fim, vetores portando um terceiro gene de resistência (a puromicina) estão em fase de teste. De forma inesperada, o cultivo de epimastigotas na presença das três drogas (higromicina B, G418 e puromicina) gerou parasitos com morfologia alterada, o que indica que mais testes serão necessários para se averiguar quais concentrações podem ser usadas quando as três drogas são usadas

simultaneamente. Paralelamente, estamos construindo plasmídeos para a deleção de outros genes da via de MMR em *T. cruzi* como *Tcmsh3*, *Tcmsh6* e *Tcmlh1*. Se for obtida a deleção completa destes *loci*, esses resultados indicarão que é possível a deleção de outros genes da via de MMR e que parasitos *Tcmsh2*^{-/-} não são viáveis pelo fato deste gene ser essencial.

A superexpressão da proteína TcMSH2 em *T. cruzi* foi outra estratégia adotada neste trabalho. Entretanto, análises de *northern* e *western* blots indicaram que as tentativas de expressão da proteína TcMSH2 no parasito transfectado não obtiveram sucesso. O sistema de superexpressão adotado com base no vetor pROCK (DaRocha et al., 2004b) não foi eficaz, pois nos experimentos de *Northern* blot foram observadas somente bandas de tamanho aproximado de 3,3 Kb nos clones transfectados, similares em tamanho ao mRNA endógeno de *Tcmsh2* (Augusto-Pinto et al., 2001). Como o mRNA gerado a partir da sequência clonada no vetor de expressão pROCK deveria ter ~3,6 Kb, podemos concluir que, apesar de terem sido gerados parasitos resistentes a G418, transcritos para o gene *Tcmsh2* presente no vetor pROCK podem não ser estáveis. Nos experimentos de *Western* blot a banda correspondente a 75 KDa apresenta níveis de expressão similares nos parasitos transfectados com o vetor pROCK-TcMSH2-Neo e nos parasitos controle. Nos parasitos transfectados, uma segunda banda de tamanho aproximado de 100 KDa pode ser visualizada, possivelmente como resultado da expressão do gene *Tcmsh2* sob o controle transcricional do promotor ribossomal presente no vetor. Duas bandas de tamanhos similares foram observadas em estudos anteriores, nos quais foram utilizados extratos protéicos de formas sanguíneas de *T. brucei* e anticorpos anti-*Tbms2* (Machado-Silva et al., 2008). No nosso caso, entretanto, a banda de 100 KDa não foi visível em todos os experimentos de *western* blot, realizados com extratos protéicos de formas epimastigotas de *T. cruzi* transfectados. É possível que o anticorpo policlonal anti-TcMSH2::His esteja reconhecendo uma banda de 75 KDa que não corresponda à proteína TcMSH2 e, para confirmar os resultados dos experimentos de *Western* blot, novos anticorpos estão sendo gerados. Existe também a possibilidade da proteína TcMSH2 ser estável somente na forma de heterodímeros como, por exemplo, nos complexos MutS α (MSH2-MSH6) e MutS β (MSH2-MSH3), pois foi demonstrado em humanos que as proteínas MSH3 e MSH6 se estabilizam apenas quando estão associadas à proteína MSH2 (Marra et al., 1998). A

proteína TcMSH2 pode ainda sofrer uma clivagem pós-traducional, gerando uma isoforma alternativa de tamanho menor. O gene *Tcmsh2* em *T. cruzi* codifica uma proteína putativa de 106 KDa. É possível que esta isoforma seja produzida para exercer funções de resposta aos danos causados por agentes genotóxicos ou ser clivada, para gerar a isoforma de 75 KDa que exerce funções de resposta ao dano oxidativo. Em resposta aos danos gerados pelos agentes genotóxicos, o parasito poderia ter a expressão aumentada de uma ou mais isoformas da proteína TcMSH2 e este efeito seria visualizado em ensaios de *western blot*. Extratos protéicos totais de parasitas tratados com estes agentes genotóxicos poderiam ser submetidos a ensaios de imunoprecipitação e espectrometria de massa, para a identificação das bandas obtidas. Este experimento seria relevante não só para investigar se a banda obtida pela imunoprecipitação corresponde à proteína TcMSH2, como também verificar qual(is) fragmento(s) da proteína estaria(m) sendo clivada(s) para se obter a isoforma de 75 KDa.

A disponibilidade de mutantes *Tbmsh2*^{-/-} nos permitiu testar se a proteína MSH2 de *T. cruzi* (TcMSH2) era capaz de complementar a deficiência do MMR em *T. brucei* (Machado-Silva et al., 2008). Embora tenha sido possível a expressão de TcMSH2 em mutantes *Tbmsh2*^{-/-}, a complementação do MMR – testada através de ensaios de resistência ao MNNG e de instabilidade de microssatélites – não foi alcançada. Entretanto, a expressão de TcMSH2 em mutantes *Tbmsh2*^{-/-} resultou na reversão da sensibilidade ao estresse oxidativo induzido por H₂O₂, fenótipo observado não somente em mutantes duplo nocautes de *msh2* em *T. brucei*, mas também em heminocautes de *Tcmsh2* em *T. cruzi* descritos neste trabalho. Embora linhagens de *T. brucei Tbmsh2*^{-/-} sejam viáveis, cerca de 2% das células mutantes são discinetoplásticas (com algum tipo de alteração na razão núcleo/cinetoplasto) mesmo na ausência de tratamento com H₂O₂ (Campos et al., 2009, submetido). Estes dados, obtidos pelo grupo do Dr. Richard McCulloch (Wellcome Centre for Molecular Biology, Glasgow, Reino Unido) que fazem parte do artigo submetido à publicação na revista *PLoS Pathogens*, indicam que a proteína MSH2 pode exercer funções MMR-independentes e pode estar envolvida na resposta ao dano oxidativo mitocondrial em *T. cruzi* e em *T. brucei*.

É possível que, em função do metabolismo em *T. cruzi* ser dependente da fosforilação oxidativa (Bringaud et al., 2006), de maneira similar ao observado na

forma procíclica (a forma presente no vetor invertebrado, a mosca tsé-tsé) de *T. brucei*, a deleção do alelo de *Tcmsh2* em epimastigotas acarretaria em um acúmulo maior de dano oxidativo endógeno no DNA. Esta seria, portanto, uma explicação para a viabilidade de formas sanguíneas de *T. brucei* mutantes para o alelo *Tbmsh2*, pois os parasitos nesta forma são dependentes principalmente da glicólise para a produção de ATP (Besteiro et al., 2005). Uma forma de se verificar esta hipótese seria através da deleção do gene *Tcmsh2* em formas procíclicas de *T. brucei*, que são metabolicamente mais semelhantes às formas epimastigotas de *T. cruzi* (Bringaud et al., 2006).

Os resultados combinados dos ensaios de complementação em *T. brucei*, da inviabilidade de mutantes *Tcmsh2* em *T. cruzi* e dos fenótipos comuns observados nos mutantes heminocaute *Tcmsh2*^{+/-} e o duplo nocaute *Tbmsh2*^{-/-} sugerem funções adicionais da proteína MSH2, envolvidas na manutenção da integridade mitocondrial. Tais funções podem ser independentes de outros componentes do MMR, com base em dois resultados importantes: a deleção do alelo *Tbmlh1* em *T. brucei* não resultou em um aumento na sensibilidade ao H_2O_2 (Machado-Silva et al., 2008), mas a deleção de *Tbmsh2* em *T. brucei* leva a um aumento no número de células discinetoplásticas (Campos et al., submetido). Estes resultados são consistentes com os dados obtidos pela Dra. Carolina Furtado (Torres-Silva, 2009), nos quais foi mostrado que a incubação de parasitos com cádmio aboliu as diferenças entre as cepas na resposta ao tratamento com cisplatina e MNNG, mas não com H_2O_2 . Sabe-se que o cádmio suprime a atividade de complexos MutS α e MutS β , o que provavelmente diminui a interação dos mesmos com os heterodímeros contendo a proteína MLH1 (Clark and Kunkel, 2004). Os ensaios na presença de cádmio indicam que a resposta ao estresse oxidativo pode envolver proteínas da via de MMR, ou pelo menos a proteína MSH2, mas não a via completa.

Por outro lado, o aumento na sensibilidade ao MNNG nos heminocautes é um resultado inesperado: é possível que, para responder ao dano causado pelo MNNG, TcMSH2 seja “sequestrada” para a sua função de MMR, o que comprometeria a resposta ao estresse oxidativo? As taxas de crescimento de mutantes *T. cruzi* *Tcmsh2*^{+/-} e *T. brucei* *Tbmsh2*^{+/-} frente ao tratamento com MNNG se mostram inversas: em *T. brucei*, os parasitos heminocautes apresentam um pequeno aumento na tolerância ao MNNG (Bell et al., 2004). Novamente, a dependência de formas

epimastigotas de *T. cruzi* em obter energia da fosforilação oxidativa, ao contrário das formas sanguíneas de *T. brucei*, poderia explicar a diferença no fenótipo observado, pois epimastigotas de *T. cruzi* estariam mais expostas ao estresse endógeno do que as formas sanguíneas de *T. brucei* (Bringaud et al., 2006). É importante salientar novamente que os mutantes de *T. brucei* foram gerados na forma sanguínea, o que significa que formas procíclicas mutantes de *T. brucei* *Tbmsh2* tratadas com MNNG poderiam apresentar um fenótipo similar ao que foi observado em formas epimastigotas de *T. cruzi* mutantes para o gene *Tcmsh2*.

Essa hipótese é compatível com os resultados dos experimentos de cultivo de epimastigotas na presença de uma dose sub-letal de MNNG por 15 dias, que mostraram que os clones *Tcmsh2*^{+/-} consumiram mais O₂ e produziram menos H₂O₂ do que os parasitos do tipo selvagem. Podemos especular que a dose usada de MNNG torna as células apenas mais suscetíveis a erros de pareamento decorrentes de transversões G-T, gerando assim um aumento lento e constante dos substratos clássicos para o MMR e conseqüentemente, recrutando a proteína TcMSH2. Neste caso é possível que, em *T. cruzi*, a proteína TcMSH2 seja recrutada para funções de MMR, restando uma quantidade menor de moléculas da proteína para lidar com o dano oxidativo endógeno e permitindo o aumento das espécies reativas de oxigênio (ROS) via cadeia transportadora de elétrons. Para evitar que o aumento nos níveis de ROS comprometa a viabilidade da célula, outras vias em cascata de resposta ao dano oxidativo podem ser ativadas, sinalizando o deslocamento de mais O₂ comoceptor final de elétrons. Conseqüentemente, a diminuição nos níveis de ROS leva à diminuição de um dos seus subprodutos, o H₂O₂.

Os mutantes de *T. cruzi* *Tcmsh2*^{+/-} mostraram ser mais sensíveis ao dano oxidativo induzido por H₂O₂ do que parasitos do tipo selvagem, além de apresentarem um aumento no acúmulo de 8-oxoG, particularmente no kDNA. Em *Escherichia coli*, três enzimas diferentes, que constituem o sistema GO, cooperam simultaneamente para prevenir a mutagênese espontânea causada pela 8-oxoG: Fpg, MutY e MutT (Michaels and Miller, 1992). Em *T. cruzi* não foram descritos ortólogos da proteína MutT (responsável pela hidrólise de 8-oxoGTP a 8-oxoGMP do pool de nucleotídeos). No parasito, foram identificadas seqüências ortólogas a MutY e FPG, que removem a adenina pareada a 8-oxoG ou a 8-oxoG pareada com citosina (El-Sayed et al., 2005a). Se a MutT está ausente em *T. cruzi*, a formação de

8-oxoG no genoma do parasito torna-se um evento altamente mutagênico. Neste contexto, pode-se especular que o envolvimento de TcMSH2 na resposta ao dano oxidativo seria crucial, de forma que a perda de um único alelo teria um impacto significativo para a sobrevivência do parasito. A natureza de tal papel de TcMSH2 no reparo mitocondrial ainda é indefinida, mas é possível que a proteína exerça o papel da proteína MSH1, envolvida no reparo mitocondrial em leveduras, visto que um gene ortólogo a *msh1* em tripanossomatídeos não foi identificado (El-Sayed et al., 2005a). A proteína MSH1 não apresenta distribuição filogenética ampla, tendo sido descrita até o momento somente em leveduras (Pogorzala et al., 2009) e *T. gondii* (Garrison and Arrizabalaga, 2009). Além disso, o parasito *T. cruzi* não possui um gene ortólogo da proteína YB1, proteína descrita como integrante do MMR mitocondrial em humanos (de Souza-Pinto et al., 2009). A localização citoplasmática da proteína TcMSH2 não exclui a possibilidade de ela ser recrutada ao núcleo para desempenhar funções de reparo mitocondrial, pois experimentos de co-precipitação em células humanas mostraram que o complexo MutS α (MSH2-MSH6) é pré-formado no citoplasma e a translocação para o núcleo é dependente de danos por alquilação e fosforilação das proteínas do complexo (Christmann et al., 2002). É possível que a proteína TcMSH2 seja translocada para a mitocôndria após o dano oxidativo, para desempenhar funções de reparo de DNA mitocondrial. O envolvimento de MSH2 no controle do dano oxidativo no DNA foi previamente descrito em outros tipos celulares, tais como fibroblastos de embriões de camundongo, nos quais foi detectado um aumento nos níveis de 8-oxoG após a inativação do gene (Russo et al., 2009). Têm sido descritas também funções adicionais para as proteínas do MMR MSH2, MSH3 e PMS2, incluindo anelamento de fita simples e recombinação homóloga, bloqueio anti-recombinação, participação na sinalização de danos no DNA para vias apoptóticas, bem como mutagênese sítio-específica durante hipermutação somática e recombinação para mudança de classe de imunoglobulinas [revisito por (Jun et al., 2006)]. Estas funções estão limitadas a atividades no núcleo, o que significa que o papel mitocondrial visto em nosso trabalho para MSH2 não foi identificado até então em nenhum outro organismo.

O dano oxidativo pode ter um efeito significativo sobre as taxas de mutação de *T. cruzi*, considerando ser este um parasito intracelular. A geração de espécies reativas de oxigênio é um dos principais mecanismos pelos quais fagócitos controlam

a infecção pelo parasito. As enzimas antioxidantes, tais como a Fe-SOD (superóxido dismutase mitocondrial dependente de ferro), TcMPX e TcCPX (triparredoxina peroxidases mitocondrial e citoplasmática, respectivamente), e TcAPX (hemoperoxidase ascorbato-dependente) têm sua expressão aumentada durante a transformação de epimastigotas não infectivos derivados no inseto para as formas infectivas tripomastigotas metacíclicas (Mielniczki-Pereira et al., 2007), provavelmente como uma pré-adaptação das formas metacíclicas para lidar com as espécies reativas de oxigênio e de nitrogênio geradas pelo macrófago durante a infecção de *T. cruzi* no hospedeiro vertebrado (Mielniczki-Pereira et al., 2007). Nosso trabalho apresenta evidências que demonstram o envolvimento da proteína MSH2 na resposta ao dano oxidativo em tripanossomatídeos. É possível que, na ausência de um gene ortólogo da catalase em *T. cruzi* e em *T. brucei* (El-Sayed et al., 2005b), a proteína MSH2 teria um papel crucial na prevenção dos efeitos mutagênicos do dano oxidativo ao DNA, juntamente com outras vias de reparo já previamente descritas (Russo et al., 2009; Russo et al., 2007), e principalmente no reparo do DNA mitocondrial (Pogorzala et al., 2009). Além disso, a proteína TcMSH2 poderia exercer funções de sinalização para vias apoptóticas e/ou interagir com outras proteínas envolvidas em diferentes vias de reparo de DNA ou de resposta ao dano oxidativo em *T. cruzi*.

Estudos comparativos da taxa de mutação de cepas pertencentes a diferentes linhagens de *T. cruzi* (antes e depois do estresse oxidativo induzido por H₂O₂) poderão fornecer informações importantes sobre o papel do MMR – e mais precisamente, da proteína TcMSH2 – na geração de variabilidade genética. Neste sentido, foi desenvolvida pelo nosso grupo (Dra. Carolina Furtado, Tese de Doutorado) uma metodologia para análise da frequência de mutação em *T. cruzi*, baseada na reversão de uma mutação presente no gene de resistência a G418, inserido no genoma do parasita. Os dados obtidos nos ensaios preliminares de detecção de mutantes revertentes resistentes a G418 indicam a viabilidade da metodologia proposta para análise da frequência de mutação em *T. cruzi*. Entretanto, ele apresenta algumas limitações: o sistema não é capaz de detectar a ocorrência de mutações neutras ou de mutações que resultem em perda de funcionalidade do gene repórter (mutações negativas), detectando somente aquelas que geram alterações fenotípicas, resultando em uma subestimação da taxa de mutação. Além disso, o

ensaio não é capaz de determinar com precisão a taxa de mutação genômica do parasito, visto que a frequência de mutações pode variar ao longo do genoma.

Em resumo, os dados apresentados nesta Tese indicam que a maior variabilidade genética verificada em cepas de *T. cruzi* II pode estar diretamente relacionada à atividade do gene *Tcmsh2*, seja na sua função no MMR, seja na sua função na resposta ao estresse oxidativo. Se a isoforma de *Tcmsh2* presente nas cepas da linhagem *T. cruzi* II apresenta menor eficiência quando comparada à isoforma presente nas cepas de *T. cruzi* I, isso pode implicar em uma menor eficiência no MMR, o que levaria a uma maior taxa de mutação e, conseqüentemente, uma maior variabilidade genética. O papel de TcMSH2 na resposta ao estresse oxidativo também poderia influenciar a geração de variabilidade genética nesse organismo. Se os mecanismos de resposta ao dano oxidativo sequestrarem a proteína TcMSH2 para exercer essa função na mitocôndria, as células que apresentarem um maior nível de estresse oxidativo poderiam ter uma menor quantidade de TcMSH2 disponível para atuar no MMR, o que tornaria essa via menos eficiente nessas células. Os dados apresentados nesta Tese mostram que as cepas do *T. cruzi* II apresentam um acúmulo maior de dano no DNA, o que poderia estar sequestrando o TcMSH2 de sua função do MMR. Assim, um MMR menos eficiente levaria a uma maior taxa de mutação e maior geração de variabilidade genética. Ambas as hipóteses estão sumarizadas na Figura 24 e é importante ressaltar que elas não se excluem: os dois mecanismos de modulação da eficiência do MMR podem agir conjuntamente. Outros estudos são necessários para avaliar o papel de TcMSH2 no reconhecimento, sinalização e possíveis interações com outras vias de reparo de DNA. Entre os experimentos em andamento, destacamos a determinação da taxa de mutação em diferentes cepas de *T. cruzi*, a geração de linhagens duplo nocaute para *Tcmsh6*, *Tcmsh3* e *Tcmlh1* (que codificam outros componentes da via de MMR), bem como a realização de ensaios de atividade de MMR *in vitro* nas diferentes cepas do parasito.

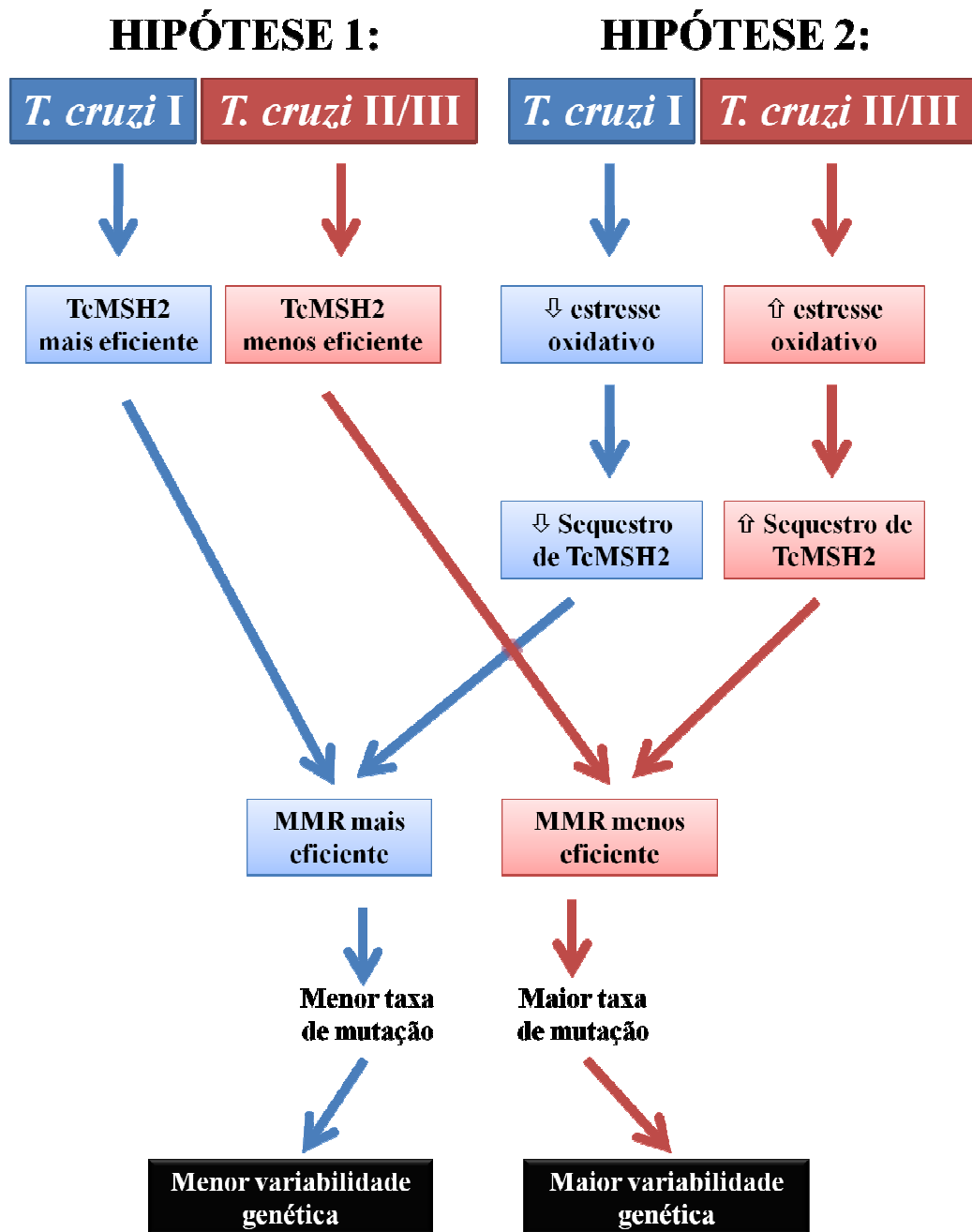


Figura 24: Correlação entre as isoformas da proteína TcMSH2, a eficiência do MMR e a variabilidade genética nas diferentes linhagens de *T. cruzi*. As cepas pertencentes à linhagem *T. cruzi* I podem apresentar um MMR mais eficiente quando comparados às cepas pertencentes às linhagens II e III. A eficiência da proteína TcMSH2 em processos MMR-dependentes (Hipótese 1) e MMR-independentes (Hipótese 2) pode afetar diretamente as taxas de mutação, influenciando na variabilidade genética observada entre as linhagens de *T. cruzi*.

2. Padrões de processamento do mRNA em *Trypanosoma cruzi*

Durante o ciclo celular do *T. cruzi*, o parasito sofre contínuas mudanças de um meio ambiente para outro e este processo requer a expressão diferencial de muitas proteínas para permitir sua rápida adaptação. Ao contrário de outros organismos, os tripanossomatídeos não regulam a expressão gênica no nível da iniciação da transcrição. A transcrição nestes organismos é policistrônica e o principal ponto de regulação da expressão gênica ocorre no nível pós-traducional [revisado por (Clayton, 2002)].

A expressão gênica em tripanossomatídeos requer a clivagem e o processamento de transcritos policistrônicos, com conseqüente geração de mRNAs monocistrônicos. Estes eventos de clivagem são acoplados à adição de uma seqüência *Spliced Leader sequence* (SL) na extremidade 5' e da poliadenilação da extremidade 3' de cada mRNA. Ambos os eventos são dependentes de seqüências regulatórias ricas em polipirimidinas (poliYs), localizadas nas regiões intercodificadoras [revisado por (Teixeira and daRocha, 2003)]. Em nosso trabalho, analisamos os padrões de processamento do mRNA em *T. cruzi* e os nossos resultados forneceram informações importantes no que concerne à distribuição, composição e tamanho de regiões regulatórias (poliYs). Nossas análises utilizando cDNAs (DNAs complementares) e ESTs (do inglês *expressed sequence tags*) identificaram o tamanho mediano de regiões 5' e 3'UTR (não traduzidas) de 35 e 137 nt, respectivamente (Figuras **21A** e **21B**), sendo que a distribuição dos tamanhos das regiões 3'UTR é mais ampla do que a observada para as seqüências correspondentes à região 5'UTR. Estes resultados corroboram com diversos estudos anteriores, nos quais é mostrado que a seqüências regulatórias estão presentes principalmente na região 3'UTR (Coughlin et al., 2000; da Silva et al., 2006; Di Noia et al., 2000; Weston et al., 1999). Grande parte dos elementos regulatórios analisados nestes trabalhos estão envolvidos com o controle da estabilidade do mRNA, o principal mecanismo responsável pela expressão gênica diferencial em *T. cruzi* (Haile and Papadopoulou, 2007).

Foi feita a busca por poliYs antes de sítios aceptores de *Spliced Leader* (SL) e depois dos sítios aceptores de poliadenilação ($A_{(n)}$), usando o algoritmo *Fuzznuc* (pacote EMBOSS, plataforma Linux). Foi previamente relatado que um dos

elementos essenciais no controle tanto da adição de SL quanto de $A_{(n)}$ em *T. brucei* é constituído de 9 ou 10 resíduos de pirimidinas (Matthews et al., 1994; Siegel et al., 2005). Nossos resultados, obtidos por buscas de poliYs constituídas de pelo menos 9 resíduos (permitindo uma purina interna) mostraram que as distâncias medianas entre as poliYs mais próximas dos sítios aceptores de SL e $A_{(n)}$ foram de 18 e 40 nt, respectivamente (Figuras **21C** e **21D**). Além disso, foi observado que 95% das distâncias analisadas entre poliYs e sítios aceptores de SL foram menores que 50 nt, e que entre poliYs e sítios aceptores de $A_{(n)}$, as distâncias foram menores que 100 nt em 91% dos casos. Estes dados estão de acordo com a observação de que o 5' UTR possui tamanho médio menor que 3' UTR e conseqüentemente, a distância entre a poliY mais próxima e o sítio acceptor de SL é também menor. A comparação entre os resultados obtidos neste trabalho (sumarizados na Figura **23**) e os descritos em *T. brucei* (Benz et al., 2005), bem como as informações reveladas no seqüenciamento genômico de *T. cruzi* (Berriman et al., 2005; El-Sayed et al., 2005a) e *T. brucei* (Berriman et al., 2005), estão de acordo com as diferenças de tamanho de genoma e de densidade de genes entre tripanossomatídeos (El-Sayed et al., 2005b) (Tabela **2**). A análise comparativa dos genomas dos três tripanossomatídeos revela que o genoma de *T. cruzi* é o mais compacto, com os menores tamanhos médios de regiões codificadoras e intercodificadoras entre os três parasitos (Berriman et al., 2005; El-Sayed et al., 2005a; Ivens et al., 2005).

A distribuição das poliYs ao longo do genoma do parasito não é homogênea; ao contrário, análises utilizando seqüências localizadas acima de regiões codificadoras de proteínas mostraram uma freqüência reduzida destes motivos, quando comparadas às porções intercodificadoras (Figuras **22A** e **22B**). Estas análises mostram também que há duas regiões ricas em pirimidinas, próximas aos sítios de adição de $A_{(n)}$, com o primeiro pico quase se sobrepondo com o sítio de $A_{(n)}$ e um segundo pico aproximadamente 100 nt após o final da região 3' do transcrito. Além disso, a composição de nucleotídeos destas duas regiões ricas em pirimidinas é um pouco diferente, conforme mostrado na Figura **22C**. Embora ambas as regiões sejam compostas principalmente de timinas, após o primeiro pico há uma região com alta concentração de adeninas, que é o nt mais frequentemente utilizado para a adição de $A_{(n)}$ (dados não mostrados). Por outro lado, o segundo pico de poliYs é predominantemente composto de timinas, com a supressão significativa de guaninas

e adeninas naquela região. É possível que em *T. cruzi* a região rica em poliYs mais próxima à região de adição de $A_{(n)}$ possa não corresponder ao motivo de poliYs utilizado na adição de SL no gene imediatamente inferior. Assim, a maquinaria de *splicing* poderia reconhecer dois motivos distintos de poliYs, um para o *trans-splicing* e outro para a poliadenilação. Essas possibilidades poderiam ser testadas através da construção de um vetor com dois genes repórteres em *tandem*, separados por três tipos de sequências regulatórias: a primeira, contendo um único motivo de poliYs próximo ao primeiro gene repórter; a segunda, contendo um único motivo de poliYs próximo ao segundo gene repórter; e a terceira, contendo dois motivos de poliYs entre os dois genes. Os motivos de poliYs poderiam também variar na sua sequência nucleotídica, de forma que a influência da composição das poliYs no mecanismo de adição da sequência *Spliced Leader* e da poliadenilação pudesse ser avaliada.

Os requerimentos exatos de poliYs e, mais importante, se o tamanho e a composição dos mesmos poderia influenciar o *trans-splicing* e a poliadenilação em *T. cruzi*, são informações que precisam ser testadas experimentalmente. Através de um sistema de transfecção composto de dois genes repórteres, foi recentemente relatado um estudo sistemático de sequências necessárias para o *trans-splicing* em *T. brucei* (Siegel et al., 2005). Neste trabalho, foi demonstrada uma clara correlação entre a expressão de luciferase e tamanho de poliYs, com um tamanho ótimo de 25 nt, contendo uridinas e citosinas, além da importância da posição da poliY em relação ao sítio acceptor de SL (o dinucleotídeo AG). Análises da região HX1 de *T. cruzi* (uma sequência altamente eficiente para o processamento do pré-mRNA) também mostraram que o primeiro dinucleotídeo AG abaixo do poliY é sempre o selecionado para o *trans-splicing* e que adeninas como sítio de adição de $A_{(n)}$ estão sempre muito próximas ao poliY (Ben-Dov et al., 2005). Testes de mutagênese dentro da região intergênica que separa o gene associado à calmodulina-ubiquitina do gene de fusão à ubiquitina em *T. cruzi* suportam o modelo de seleção por varredura do sítio de processamento da região 3' UTR em *T. cruzi* (Hummel et al., 2000). Com o objetivo de melhorar vetores para a expressão de genes exógenos em *T. cruzi*, nosso grupo comparou 5 diferentes regiões intergênicas inseridas imediatamente acima do gene de luciferase, através de transfecção transiente de epimastigotas (DaRocha et al., 2004b). Neste estudo, foram testadas as regiões intergênicas de

quatro genes constitutivos e uma região referente à amastina, um gene altamente expresso em amastigotas do parasito. Os motivos de poliYs presentes nas regiões intergênicas utilizadas neste trabalho poderiam ser reconhecidas por PTBs (do inglês *polypyrimidine tract binding*), proteínas multifuncionais envolvidas no *splicing* e na regulação da estabilidade do mRNA (Stern et al., 2009). Entretanto, com exceção da sequência de amastina, somente pequenas diferenças foram observadas quando as outras sequências são comparadas, indicando que, pelo menos para os genes analisados, as diferenças no tamanho de poliYs, sua composição ou localização em relação ao sítio acceptor de SL não afetam a expressão destes genes em epimastigotas de *T. cruzi* (DaRocha et al., 2004b). É interessante notar que os quatro genes constitutivos analisados apresentam longas sequências de poliYs acima da região acceptora de SL com capacidade de preencher os requerimentos mínimos para um eficiente evento de processamento do mRNA.

A comparação entre os resultados obtidos em *T. cruzi* com os padrões de processamento observados em outros tripanossomatídeos, como por exemplo, *Leishmania*, ressalta as características existentes entre os membros desta família: em *L. infantum*, os motivos de poliYs são quase duas vezes mais numerosos nos 50 primeiros nucleotídeos acima do sítio de A_(n) (Smith et al., 2008), enquanto que em *T. cruzi*, este motivo é quase coincidente com o sítio de A_(n); a adição de A_(n) ocorre preferencialmente em um resíduo de adenosina dentro das regiões 3' UTR, visto que em tripanossomatídeos, não foi detectado um sítio canônico AAUAAA dentro destas regiões (Benz et al., 2005; Campos et al., 2008; Smith et al., 2008); similar ao observado em *T. cruzi* e *T. brucei*, foi também identificado em *L. infantum* um segundo motivo de poliYs abaixo do sítio de adição de A_(n), possivelmente com um papel central no controle da maquinaria de *trans*-*splicing*; no genoma de *T. cruzi*, a região rica em polipirimidinas está localizada a 40 nt do sítio de A_(n), em *T. brucei* e em *Leishmania*, ela se encontra a 120 e 300-600 nt, respectivamente. A menor distância entre o sítio de A_(n) e o motivo de poliY em *T. cruzi* está em concordância com uma maior compactação de seu genoma, quando comparada à dos outros dois tripanossomatídeos (El-Sayed et al., 2005b).

Algoritmos foram desenvolvidos para a predição dos sítios de processamento do mRNA em tripanossomatídeos, baseados nos padrões observados em *T. brucei* e em *Leishmania* (Benz et al., 2005; Smith et al., 2008). As observações feitas em

nosso estudo (Campos et al., 2008) poderiam ser utilizadas para o refinamento destes algoritmos, de forma que estes seriam também capazes de prever os produtos derivados do processamento do mRNA em *T. cruzi*. A identificação experimental de um grande número de UTRs em mRNAs de *T. cruzi*, para os quais não existem ESTs ou cDNAs disponíveis, é um estratégia dispendiosa. Assim sendo, a predição *in silico* de UTRs em *T. cruzi*, a partir de sequências genômicas, permitirá uma rápida identificação de um grande número de supostos elementos regulatórios em *cis*, cujo papel na modulação da expressão gênica poderia ser validado experimentalmente.

As informações obtidas neste trabalho podem ser utilizadas para a construção de vetores contendo genes repórteres sob o controle transcricional de diferentes sequências regulatórias como, por exemplo, polipirimidinas de tamanhos e composições diferentes. Os resultados obtidos poderiam contribuir significativamente para estudos de aperfeiçoamento de vetores comumente usados na manipulação da expressão gênica do parasito e obter uma expressão mais eficiente de genes exógenos em *T. cruzi*.

- ✓ Gerar linhagens duplo nocaute do gene *Tcmsh2* em *Trypanosoma cruzi*, através da inserção de uma cópia adicional do gene no locus de β -tubulina entre o primeiro e o segundo ciclo de transfecção (linhagens *Tcmsh2*^{+/-});
- ✓ Investigar as causas da ineficácia em superexpressar o gene *Tcmsh2* em parasitos transfectados com o vetor pROCK-TcMSH2-Neo;
- ✓ Comparar os níveis de acúmulo de 8-oxoguanina nas linhagens superexpressoras do gene *Tcmsh2*;
- ✓ Realizar ensaios de localização subcelular de TcMSH2 no parasito após o tratamento com diversos agentes genotóxicos;
- ✓ Determinar a taxa de mutação em diferentes cepas de *T. cruzi* (Col1.7G2, Silvio X-10 cl1, JG, Esmeraldo cl3 e CL Brener);
- ✓ Gerar linhagens duplo nocaute para outros genes da via de MMR em *Trypanosoma cruzi*, tais como *Tcmsh3*, *Tcmsh6* e *Tcmlh1*;
- ✓ Caracterizar as linhagens *Tcmsh3*^{-/-}, *Tcmsh6*^{-/-} e *Tcmlh1*^{-/-} e verificar o seu fenótipo frente a agentes genotóxicos e ao estresse oxidativo induzido por H₂O₂;
- ✓ Realizar ensaios de atividade de MMR *in vitro*.

- Aaltonen, L.A., Peltomaki, P., Leach, F.S., Sistonen, P., Pylkkanen, L., Mecklin, J.P., Jarvinen, H., Powell, S.M., Jen, J., Hamilton, S.R., *et al.* (1993). Clues to the pathogenesis of familial colorectal cancer. *Science* 260, 812-816.
- Aebi, S., Kurdi-Haidar, B., Gordon, R., Cenni, B., Zheng, H., Fink, D., Christen, R.D., Boland, C.R., Koi, M., Fishel, R., *et al.* (1996). Loss of DNA mismatch repair in acquired resistance to cisplatin. *Cancer Res* 56, 3087-3090.
- Akiyama, Y., Sato, H., Yamada, T., Nagasaki, H., Tsuchiya, A., Abe, R., and Yuasa, Y. (1997). Germ-line mutation of the hMSH6/GTBP gene in an atypical hereditary nonpolyposis colorectal cancer kindred. *Cancer Res* 57, 3920-3923.
- Allaoui, A., Francois, C., Zemzoumi, K., Guilvard, E., and Ouaiissi, A. (1999). Intracellular growth and metacyclogenesis defects in *Trypanosoma cruzi* carrying a targeted deletion of a Tc52 protein-encoding allele. *Mol Microbiol* 32, 1273-1286.
- Allen, D.J., Makhov, A., Grilley, M., Taylor, J., Thresher, R., Modrich, P., and Griffith, J.D. (1997). MutS mediates heteroduplex loop formation by a translocation mechanism. *EMBO J* 16, 4467-4476.
- Andrade, L.O., Machado, C.R., Chiari, E., Pena, S.D., and Macedo, A.M. (1999). Differential tissue distribution of diverse clones of *Trypanosoma cruzi* in infected mice. *Mol Biochem Parasitol* 100, 163-172.
- Apweiler, R., Attwood, T.K., Bairoch, A., Bateman, A., Birney, E., Biswas, M., Bucher, P., Cerutti, L., Corpet, F., Croning, M.D., *et al.* (2000). InterPro--an integrated documentation resource for protein families, domains and functional sites. *Bioinformatics* 16, 1145-1150.
- Augusto-Pinto, L., Bartholomeu, D.C., Teixeira, S.M., Pena, S.D., and Machado, C.R. (2001). Molecular cloning and characterization of the DNA mismatch repair gene class 2 from the *Trypanosoma cruzi*. *Gene* 272, 323-333.
- Augusto-Pinto, L., Teixeira, S.M., Pena, S.D., and Machado, C.R. (2003). Single-nucleotide polymorphisms of the *Trypanosoma cruzi* MSH2 gene support the existence of three phylogenetic lineages presenting differences in mismatch-repair efficiency. *Genetics* 164, 117-126.
- Aust, A.E., and Eveleigh, J.F. (1999). Mechanisms of DNA oxidation. *Proc Soc Exp Biol Med* 222, 246-252.
- Azuaje, F.J., Ramirez, J.L., and Da Silveira, J.F. (2007). In silico, biologically-inspired modelling of genomic variation generation in surface proteins of *Trypanosoma cruzi*. *Kinetoplastid Biol Dis* 6, 6.
- Bairoch, A. (1991). PROSITE: a dictionary of sites and patterns in proteins. *Nucleic Acids Res* 19 Suppl, 2241-2245.
- Banerjee, S., and Flores-Rozas, H. (2005). Cadmium inhibits mismatch repair by blocking the ATPase activity of the MSH2-MSH6 complex. *Nucleic Acids Res* 33, 1410-1419.
- Barrett, M.P., Burchmore, R.J., Stich, A., Lazzari, J.O., Frasch, A.C., Cazzulo, J.J., and Krishna, S. (2003). The trypanosomiasis. *Lancet* 362, 1469-1480.
- Beardsley, D.I., Kim, W.J., and Brown, K.D. (2005). N-methyl-N'-nitro-N-nitrosoguanidine activates cell-cycle arrest through distinct mechanisms activated in a dose-dependent manner. *Mol Pharmacol* 68, 1049-1060.

- Bell, J.S., Harvey, T.I., Sims, A.M., and McCulloch, R. (2004). Characterization of components of the mismatch repair machinery in *Trypanosoma brucei*. *Mol Microbiol* *51*, 159-173.
- Bell, J.S., and McCulloch, R. (2003). Mismatch repair regulates homologous recombination, but has little influence on antigenic variation, in *Trypanosoma brucei*. *J Biol Chem* *278*, 45182-45188.
- Bellofatto, V., and Cross, G.A. (1989). Expression of a bacterial gene in a trypanosomatid protozoan. *Science* *244*, 1167-1169.
- Ben-Dov, C.P., Levin, M.J., and Vazquez, M.P. (2005). Analysis of the highly efficient pre-mRNA processing region HX1 of *Trypanosoma cruzi*. *Mol Biochem Parasitol* *140*, 97-105.
- Benz, C., Nilsson, D., Andersson, B., Clayton, C., and Guilbride, D.L. (2005). Messenger RNA processing sites in *Trypanosoma brucei*. *Mol Biochem Parasitol* *143*, 125-134.
- Berriman, M., Ghedin, E., Hertz-Fowler, C., Blandin, G., Renaud, H., Bartholomeu, D.C., Lennard, N.J., Caler, E., Hamlin, N.E., Haas, B., *et al.* (2005). The genome of the African trypanosome *Trypanosoma brucei*. *Science* *309*, 416-422.
- Besteiro, S., Barrett, M.P., Riviere, L., and Bringaud, F. (2005). Energy generation in insect stages of *Trypanosoma brucei*: metabolism in flux. *Trends Parasitol* *21*, 185-191.
- Bethke, L., Thomas, S., Walker, K., Lakhia, R., Rangarajan, R., and Wirth, D. (2007). The role of DNA mismatch repair in generating genetic diversity and drug resistance in malaria parasites. *Mol Biochem Parasitol* *155*, 18-25.
- Borst, P., Rottenberg, S., and Jonkers, J. (2008). How do real tumors become resistant to cisplatin? *Cell Cycle* *7*, 1353-1359.
- Branch, P., Aquilina, G., Bignami, M., and Karran, P. (1993). Defective mismatch binding and a mutator phenotype in cells tolerant to DNA damage. *Nature* *362*, 652-654.
- Branche, C., Ochaya, S., Aslund, L., and Andersson, B. (2006). Comparative karyotyping as a tool for genome structure analysis of *Trypanosoma cruzi*. *Mol Biochem Parasitol* *147*, 30-38.
- Brener, Z., Andrade, Z.A., and Barral-Neto, M. (2000). *Trypanosoma cruzi* e Doença de Chagas. (Rio de Janeiro, Guanabara Koogan).
- Bringaud, F., Riviere, L., and Coustou, V. (2006). Energy metabolism of trypanosomatids: adaptation to available carbon sources. *Mol Biochem Parasitol* *149*, 1-9.
- Briones, M.R., Souto, R.P., Stolf, B.S., and Zingales, B. (1999). The evolution of two *Trypanosoma cruzi* subgroups inferred from rRNA genes can be correlated with the interchange of American mammalian faunas in the Cenozoic and has implications to pathogenicity and host specificity. *Mol Biochem Parasitol* *104*, 219-232.
- Brisse, S., Barnabe, C., and Tibayrenc, M. (1998). *Trypanosoma cruzi*: How Many Relevant Phylogenetic Subdivisions are There? *Parasitol Today* *14*, 178-179.
- Brisse, S., Barnabe, C., and Tibayrenc, M. (2000). Identification of six *Trypanosoma cruzi* phylogenetic lineages by random amplified polymorphic DNA and multilocus enzyme electrophoresis. *Int J Parasitol* *30*, 35-44.
- Brown, R., Hirst, G.L., Gallagher, W.M., McIlwrath, A.J., Margison, G.P., van der Zee, A.G., and Anthoney, D.A. (1997). hMLH1 expression and cellular responses of ovarian tumour cells to treatment with cytotoxic anticancer agents. *Oncogene* *15*, 45-52.

Buermeyer, A.B., Deschenes, S.M., Baker, S.M., and Liskay, R.M. (1999). Mammalian DNA mismatch repair. *Annu Rev Genet* 33, 533-564.

Buscaglia, C.A., Campo, V.A., Frasch, A.C., and Di Noia, J.M. (2006). Trypanosoma cruzi surface mucins: host-dependent coat diversity. *Nat Rev Microbiol* 4, 229-236.

Buscaglia, C.A., and Di Noia, J.M. (2003). Trypanosoma cruzi clonal diversity and the epidemiology of Chagas' disease. *Microbes Infect* 5, 419-427.

Caler, E.V., Vaena de Avalos, S., Haynes, P.A., Andrews, N.W., and Burleigh, B.A. (1998). Oligopeptidase B-dependent signaling mediates host cell invasion by Trypanosoma cruzi. *EMBO J* 17, 4975-4986.

Campbell, D.A., Thomas, S., and Sturm, N.R. (2003). Transcription in kinetoplastid protozoa: why be normal? *Microbes Infect* 5, 1231-1240.

Campos, P.C., Bartholomeu, D.C., DaRocha, W.D., Cerqueira, G.C., and Teixeira, S.M. (2008). Sequences involved in mRNA processing in Trypanosoma cruzi. *Int J Parasitol* 38, 1383-1389.

Cejka, P., and Jiricny, J. (2008). Interplay of DNA repair pathways controls methylation damage toxicity in Saccharomyces cerevisiae. *Genetics* 179, 1835-1844.

Cerqueira, G.C., Bartholomeu, D.C., DaRocha, W.D., Hou, L., Freitas-Silva, D.M., Machado, C.R., El-Sayed, N.M., and Teixeira, S.M. (2008). Sequence diversity and evolution of multigene families in Trypanosoma cruzi. *Mol Biochem Parasitol* 157, 65-72.

Cerqueira, G.C., DaRocha, W.D., Campos, P.C., Zouain, C.S., and Teixeira, S.M. (2005). Analysis of expressed sequence tags from Trypanosoma cruzi amastigotes. *Mem Inst Oswaldo Cruz* 100, 385-389.

Christmann, M., Tomicic, M.T., and Kaina, B. (2002). Phosphorylation of mismatch repair proteins MSH2 and MSH6 affecting MutSalphamismatch-binding activity. *Nucleic Acids Res* 30, 1959-1966.

Chung, S.H., and Swindle, J. (1997). Trypanosoma cruzi: circularization of linear DNA fragments prior to integration during generation of stable transformants. *Mol Biochem Parasitol* 89, 321-325.

Claij, N., and Te Riele, H. (2002). Methylation tolerance in mismatch repair proficient cells with low MSH2 protein level. *Oncogene* 21, 2873-2879.

Clark, A.B., Cook, M.E., Tran, H.T., Gordenin, D.A., Resnick, M.A., and Kunkel, T.A. (1999). Functional analysis of human MutSalpham and MutSbetacomplexes in yeast. *Nucleic Acids Res* 27, 736-742.

Clark, A.B., and Kunkel, T.A. (2004). Cadmium inhibits the functions of eukaryotic MutS complexes. *J Biol Chem* 279, 53903-53906.

Clayton, C.E. (1999). Genetic manipulation of kinetoplastida. *Parasitol Today* 15, 372-378.

Clayton, C.E. (2002). Life without transcriptional control? From fly to man and back again. *EMBO J* 21, 1881-1888.

Cooper, R., de Jesus, A.R., and Cross, G.A. (1993). Deletion of an immunodominant Trypanosoma cruzi surface glycoprotein disrupts flagellum-cell adhesion. *J Cell Biol* 122, 149-156.

Coughlin, B.C., Teixeira, S.M., Kirchhoff, L.V., and Donelson, J.E. (2000). Amastin mRNA abundance in Trypanosoma cruzi is controlled by a 3'-untranslated region position-dependent cis-element and an untranslated region-binding protein. *J Biol Chem* 275, 12051-12060.

- Cribb, P., Tapia, E., Diosque, P., and Serra, E. (2004). Spliced leader RNA gene promoter sequence heterogeneity in CL-Brener *Trypanosoma cruzi* reference strain. *Infect Genet Evol* 4, 153-157.
- Cruz, A.K., Titus, R., and Beverley, S.M. (1993). Plasticity in chromosome number and testing of essential genes in *Leishmania* by targeting. *Proc Natl Acad Sci U S A* 90, 1599-1603.
- da Silva, R.A., Bartholomeu, D.C., and Teixeira, S.M. (2006). Control mechanisms of tubulin gene expression in *Trypanosoma cruzi*. *Int J Parasitol* 36, 87-96.
- DaRocha, W.D., Otsu, K., Teixeira, S.M., and Donelson, J.E. (2004a). Tests of cytoplasmic RNA interference (RNAi) and construction of a tetracycline-inducible T7 promoter system in *Trypanosoma cruzi*. *Mol Biochem Parasitol* 133, 175-186.
- DaRocha, W.D., Silva, R.A., Bartholomeu, D.C., Pires, S.F., Freitas, J.M., Macedo, A.M., Vazquez, M.P., Levin, M.J., and Teixeira, S.M. (2004b). Expression of exogenous genes in *Trypanosoma cruzi*: improving vectors and electroporation protocols. *Parasitol Res* 92, 113-120.
- de Boer, J.G. (2002). Polymorphisms in DNA repair and environmental interactions. *Mutat Res* 509, 201-210.
- de Freitas, J.M., Augusto-Pinto, L., Pimenta, J.R., Bastos-Rodrigues, L., Goncalves, V.F., Teixeira, S.M., Chiari, E., Junqueira, A.C., Fernandes, O., Macedo, A.M., *et al.* (2006). Ancestral genomes, sex, and the population structure of *Trypanosoma cruzi*. *PLoS Pathog* 2, e24.
- de Souza-Pinto, N.C., Mason, P.A., Hashiguchi, K., Weissman, L., Tian, J., Guay, D., Lebel, M., Stevensner, T.V., Rasmussen, L.J., and Bohr, V.A. (2009). Novel DNA mismatch-repair activity involving YB-1 in human mitochondria. *DNA Repair (Amst)* 8, 704-719.
- de Wind, N., and Hays, J.B. (2001). Mismatch repair: praying for genome stability. *Curr Biol* 11, R545-548.
- DeFedericis, H.C., Patrzyc, H.B., Rajecki, M.J., Budzinski, E.E., Iijima, H., Dawidzik, J.B., Evans, M.S., Greene, K.F., and Box, H.C. (2006). Singlet oxygen-induced DNA damage. *Radiat Res* 165, 445-451.
- Devera, R., Illarramendi, X., Montoya-Araujo, R., Pirmez, C., Fernandes, O., and Coura, J.R. (2002). [Biodemes of *Trypanosoma cruzi* strains isolated from humans from three endemic areas in Minas Gerais State]. *Rev Soc Bras Med Trop* 35, 323-330.
- Di Noia, J.M., Buscaglia, C.A., De Marchi, C.R., Almeida, I.C., and Frasch, A.C. (2002). A *Trypanosoma cruzi* small surface molecule provides the first immunological evidence that Chagas' disease is due to a single parasite lineage. *J Exp Med* 195, 401-413.
- Di Noia, J.M., D'Orso, I., Sanchez, D.O., and Frasch, A.C. (2000). AU-rich elements in the 3'-untranslated region of a new mucin-type gene family of *Trypanosoma cruzi* confers mRNA instability and modulates translation efficiency. *J Biol Chem* 275, 10218-10227.
- dos Santos, W.G., and Buck, G.A. (2000). Simultaneous stable expression of neomycin phosphotransferase and green fluorescence protein genes in *Trypanosoma cruzi*. *J Parasitol* 86, 1281-1288.
- Eisen, J.A. (1998). A phylogenomic study of the MutS family of proteins. *Nucleic Acids Res* 26, 4291-4300.

el-Sayed, N.M., Alarcon, C.M., Beck, J.C., Sheffield, V.C., and Donelson, J.E. (1995). cDNA expressed sequence tags of *Trypanosoma brucei rhodesiense* provide new insights into the biology of the parasite. *Mol Biochem Parasitol* 73, 75-90.

El-Sayed, N.M., Myler, P.J., Bartholomeu, D.C., Nilsson, D., Aggarwal, G., Tran, A.N., Ghedin, E., Worthey, E.A., Delcher, A.L., Blandin, G., *et al.* (2005a). The genome sequence of *Trypanosoma cruzi*, etiologic agent of Chagas disease. *Science* 309, 409-415.

El-Sayed, N.M., Myler, P.J., Blandin, G., Berriman, M., Crabtree, J., Aggarwal, G., Caler, E., Renauld, H., Worthey, E.A., Hertz-Fowler, C., *et al.* (2005b). Comparative genomics of trypanosomatid parasitic protozoa. *Science* 309, 404-409.

Elena, S.F., and Lenski, R.E. (2003). Evolution experiments with microorganisms: the dynamics and genetic bases of adaptation. *Nat Rev Genet* 4, 457-469.

Feng, W.Y., Lee, E.H., and Hays, J.B. (1991). Recombinagenic processing of UV-light photoproducts in nonreplicating phage DNA by the *Escherichia coli* methyl-directed mismatch repair system. *Genetics* 129, 1007-1020.

Fire, A., Xu, S., Montgomery, M.K., Kostas, S.A., Driver, S.E., and Mello, C.C. (1998). Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* 391, 806-811.

Freitas, J.M., Lages-Silva, E., Crema, E., Pena, S.D., and Macedo, A.M. (2005). Real time PCR strategy for the identification of major lineages of *Trypanosoma cruzi* directly in chronically infected human tissues. *Int J Parasitol* 35, 411-417.

Garrison, E.M., and Arrizabalaga, G. (2009). Disruption of a mitochondrial MutS DNA repair enzyme homologue confers drug resistance in the parasite *Toxoplasma gondii*. *Mol Microbiol* 72, 425-441.

Gill, E.E., and Fast, N.M. (2007). Stripped-down DNA repair in a highly reduced parasite. *BMC Mol Biol* 8, 24.

Gueiros-Filho, F.J., and Beverley, S.M. (1996). Selection against the dihydrofolate reductase-thymidylate synthase (DHFR-TS) locus as a probe of genetic alterations in *Leishmania major*. *Mol Cell Biol* 16, 5655-5663.

Haile, S., and Papadopoulou, B. (2007). Developmental regulation of gene expression in trypanosomatid parasitic protozoa. *Curr Opin Microbiol* 10, 569-577.

Hall, L.M., and Henderson-Begg, S.K. (2006). Hypermutable bacteria isolated from humans--a critical analysis. *Microbiology* 152, 2505-2514.

Hariharan, S., Ajioka, J., and Swindle, J. (1993). Stable transformation of *Trypanosoma cruzi*: inactivation of the PUB12.5 polyubiquitin gene by targeted gene disruption. *Mol Biochem Parasitol* 57, 15-30.

Horton, R.M., Cai, Z.L., Ho, S.N., and Pease, L.R. (1990). Gene splicing by overlap extension: tailor-made genes using the polymerase chain reaction. *Biotechniques* 8, 528-535.

Hsieh, P., and Yamane, K. (2008). DNA mismatch repair: molecular mechanism, cancer, and ageing. *Mech Ageing Dev* 129, 391-407.

Humbert, O., Fiumicino, S., Aquilina, G., Branch, P., Oda, S., Zijno, A., Karran, P., and Bignami, M. (1999). Mismatch repair and differential sensitivity of mouse and human cells to methylating agents. *Carcinogenesis* 20, 205-214.

Hummel, H.S., Gillespie, R.D., and Swindle, J. (2000). Mutational analysis of 3' splice site selection during trans-splicing. *J Biol Chem* 275, 35522-35531.

Ionov, Y., Peinado, M.A., Malkhosyan, S., Shibata, D., and Perucho, M. (1993). Ubiquitous somatic mutations in simple repeated sequences reveal a new mechanism for colonic carcinogenesis. *Nature* 363, 558-561.

Ivens, A.C., Peacock, C.S., Worthey, E.A., Murphy, L., Aggarwal, G., Berriman, M., Sisk, E., Rajandream, M.A., Adlem, E., Aert, R., *et al.* (2005). The genome of the kinetoplastid parasite, *Leishmania major*. *Science* *309*, 436-442.

Iyer, R.R., Pluciennik, A., Burdett, V., and Modrich, P.L. (2006). DNA mismatch repair: functions and mechanisms. *Chem Rev* *106*, 302-323.

Jin, Y.H., Clark, A.B., Slebos, R.J., Al-Refai, H., Taylor, J.A., Kunkel, T.A., Resnick, M.A., and Gordenin, D.A. (2003). Cadmium is a mutagen that acts by inhibiting mismatch repair. *Nat Genet* *34*, 326-329.

Jun, S.H., Kim, T.G., and Ban, C. (2006). DNA mismatch repair system. Classical and fresh roles. *FEBS J* *273*, 1609-1619.

Junqueira, A.C., Degraeve, W., and Brandao, A. (2005). Minicircle organization and diversity in *Trypanosoma cruzi* populations. *Trends Parasitol* *21*, 270-272.

Kawashita, S.Y., Sanson, G.F., Fernandes, O., Zingales, B., and Briones, M.R. (2001). Maximum-likelihood divergence date estimates based on rRNA gene sequences suggest two scenarios of *Trypanosoma cruzi* intraspecific evolution. *Mol Biol Evol* *18*, 2250-2259.

Kelly, J.M., Ward, H.M., Miles, M.A., and Kendall, G. (1992). A shuttle vector which facilitates the expression of transfected genes in *Trypanosoma cruzi* and *Leishmania*. *Nucleic Acids Res* *20*, 3963-3969.

Kruger, S., Plaschke, J., Pistorius, S., Jeske, B., Haas, S., Kramer, H., Hinterseher, I., Bier, A., Kreuz, F.R., Theissig, F., *et al.* (2002). Seven novel MLH1 and MSH2 germline mutations in hereditary nonpolyposis colorectal cancer. *Hum Mutat* *19*, 82.

Lahue, R.S., Au, K.G., and Modrich, P. (1989). DNA mismatch correction in a defined system. *Science* *245*, 160-164.

Li, G., Mitchell, D.L., Ho, V.C., Reed, J.C., and Tron, V.A. (1996). Decreased DNA repair but normal apoptosis in ultraviolet-irradiated skin of p53-transgenic mice. *Am J Pathol* *148*, 1113-1123.

Liang, X.H., Haritan, A., Uliel, S., and Michaeli, S. (2003). trans and cis splicing in trypanosomatids: mechanism, factors, and regulation. *Eukaryot Cell* *2*, 830-840.

Llewellyn, M.S., Miles, M.A., Carrasco, H.J., Lewis, M.D., Yeo, M., Vargas, J., Torrico, F., Diosque, P., Valente, V., Valente, S.A., *et al.* (2009). Genome-scale multilocus microsatellite typing of *Trypanosoma cruzi* discrete typing unit I reveals phylogeographic structure and specific genotypes linked to human infection. *PLoS Pathog* *5*, e1000410.

Machado-Silva, A. (2008). Reparo de DNA em tripanossomatídeos: análises in vitro e in vivo e envolvimento na resposta a stress oxidativo. Tese de Doutorado. In Departamento de Bioquímica e Imunologia (Belo Horizonte, Universidade Federal de Minas Gerais), pp. 154.

Machado-Silva, A., Teixeira, S.M., Franco, G.R., Macedo, A.M., Pena, S.D., McCulloch, R., and Machado, C.R. (2008). Mismatch repair in *Trypanosoma brucei*: heterologous expression of MSH2 from *Trypanosoma cruzi* provides new insights into the response to oxidative damage. *Gene* *411*, 19-26.

Machado, C.A., and Ayala, F.J. (2002). Sequence variation in the dihydrofolate reductase-thymidylate synthase (DHFR-TS) and trypanothione reductase (TR) genes of *Trypanosoma cruzi*. *Mol Biochem Parasitol* *121*, 33-47.

Machado, C.R., Augusto-Pinto, L., McCulloch, R., and Teixeira, S.M. (2006). DNA metabolism and genetic diversity in Trypanosomes. *Mutat Res* *612*, 40-57.

MacRae, J.I., Obado, S.O., Turnock, D.C., Roper, J.R., Kierans, M., Kelly, J.M., and Ferguson, M.A. (2006). The suppression of galactose metabolism in *Trypanosoma*

cruzi epimastigotes causes changes in cell surface molecular architecture and cell morphology. *Mol Biochem Parasitol* 147, 126-136.

Mair, G., Shi, H., Li, H., Djikeng, A., Aviles, H.O., Bishop, J.R., Falcone, F.H., Gavrilescu, C., Montgomery, J.L., Santori, M.I., *et al.* (2000). A new twist in trypanosome RNA metabolism: cis-splicing of pre-mRNA. *RNA* 6, 163-169.

Manning-Cela, R., Cortes, A., Gonzalez-Rey, E., Van Voorhis, W.C., Swindle, J., and Gonzalez, A. (2001). LYT1 protein is required for efficient in vitro infection by *Trypanosoma cruzi*. *Infect Immun* 69, 3916-3923.

Marra, G., Iaccarino, I., Lettieri, T., Roscilli, G., Delmastro, P., and Jiricny, J. (1998). Mismatch repair deficiency associated with overexpression of the MSH3 gene. *Proc Natl Acad Sci U S A* 95, 8568-8573.

Marti, T.M., Kunz, C., and Fleck, O. (2002). DNA mismatch repair and mutation avoidance pathways. *J Cell Physiol* 191, 28-41.

Mathieu-Daude, F., Bosseno, M.F., Garzon, E., Lelievre, J., Sereno, D., Ouaisi, A., and Breniere, S.F. (2007). Sequence diversity and differential expression of Tc52 immuno-regulatory protein in *Trypanosoma cruzi*: potential implications in the biological variability of strains. *Parasitol Res* 101, 1355-1363.

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

Mena, A., Smith, E.E., Burns, J.L., Speert, D.P., Moskowitz, S.M., Perez, J.L., and Oliver, A. (2008). Genetic adaptation of *Pseudomonas aeruginosa* to the airways of cystic fibrosis patients is catalyzed by hypermutation. *J Bacteriol* 190, 7910-7917.

Michaels, M.L., and Miller, J.H. (1992). The GO system protects organisms from the mutagenic effect of the spontaneous lesion 8-hydroxyguanine (7,8-dihydro-8-oxoguanine). *J Bacteriol* 174, 6321-6325.

Mielniczki-Pereira, A.A., Chiavegatto, C.M., Lopez, J.A., Colli, W., Alves, M.J., and Gadelha, F.R. (2007). *Trypanosoma cruzi* strains, Tulahuen 2 and Y, besides the difference in resistance to oxidative stress, display differential glucose-6-phosphate and 6-phosphogluconate dehydrogenases activities. *Acta Trop* 101, 54-60.

Miles, M.A., Lanham, S.M., de Souza, A.A., and Pova, M. (1980). Further enzymic characters of *Trypanosoma cruzi* and their evaluation for strain identification. *Trans R Soc Trop Med Hyg* 74, 221-237.

Miles, M.A., Souza, A., Pova, M., Shaw, J.J., Lainson, R., and Toye, P.J. (1978). Isozymic heterogeneity of *Trypanosoma cruzi* in the first autochthonous patients with Chagas' disease in Amazonian Brazil. *Nature* 272, 819-821.

Momen, H. (1999). Taxonomy of *Trypanosoma cruzi*: a commentary on characterization and nomenclature. *Mem Inst Oswaldo Cruz* 94 *Suppl 1*, 181-184.

Morel, C., Chiari, E., Camargo, E.P., Mattei, D.M., Romanha, A.J., and Simpson, L. (1980). Strains and clones of *Trypanosoma cruzi* can be characterized by pattern of restriction endonuclease products of kinetoplast DNA minicircles. *Proc Natl Acad Sci U S A* 77, 6810-6814.

Muller-Koch, Y., Kopp, R., Lohse, P., Baretton, G., Stoetzer, A., Aust, D., Daum, J., Kerker, B., Gross, M., Dietmeier, W., *et al.* (2001). Sixteen rare sequence variants of the hMLH1 and hMSH2 genes found in a cohort of 254 suspected HNPCC (hereditary non-polyposis colorectal cancer) patients: mutations or polymorphisms? *Eur J Med Res* 6, 473-482.

Ngo, H., Tschudi, C., Gull, K., and Ullu, E. (1998). Double-stranded RNA induces mRNA degradation in *Trypanosoma brucei*. *Proc Natl Acad Sci U S A* 95, 14687-14692.

Nozaki, T., and Cross, G.A. (1995). Effects of 3' untranslated and intergenic regions on gene expression in *Trypanosoma cruzi*. *Mol Biochem Parasitol* 75, 55-67.

Otsu, K., Donelson, J.E., and Kirchhoff, L.V. (1993). Interruption of a *Trypanosoma cruzi* gene encoding a protein containing 14-amino acid repeats by targeted insertion of the neomycin phosphotransferase gene. *Mol Biochem Parasitol* 57, 317-330.

Pais, F.S., DaRocha, W.D., Almeida, R.M., Leclercq, S.Y., Penido, M.L., Frago, S.P., Bartholomeu, D.C., Gazzinelli, R.T., and Teixeira, S.M. (2008). Molecular characterization of ribonucleoproteic antigens containing repeated amino acid sequences from *Trypanosoma cruzi*. *Microbes Infect* 10, 716-725.

Palfi, Z., Schimanski, B., Gunzl, A., Lucke, S., and Bindereif, A. (2005). U1 small nuclear RNP from *Trypanosoma brucei*: a minimal U1 snRNA with unusual protein components. *Nucleic Acids Res* 33, 2493-2503.

Peacock, C.S., Seeger, K., Harris, D., Murphy, L., Ruiz, J.C., Quail, M.A., Peters, N., Adlem, E., Tivey, A., Aslett, M., *et al.* (2007). Comparative genomic analysis of three *Leishmania* species that cause diverse human disease. *Nat Genet* 39, 839-847.

Peltomaki, P. (2005). Lynch syndrome genes. *Fam Cancer* 4, 227-232.

Pogorzala, L.A., Mookerjee, S., and Sia, E.A. (2009). Evidence That Msh1p Plays Multiple Roles in Mitochondrial Base Excision Repair. *Genetics*.

Read, T.D., Peterson, S.N., Tourasse, N., Baillie, L.W., Paulsen, I.T., Nelson, K.E., Tettelin, H., Fouts, D.E., Eisen, J.A., Gill, S.R., *et al.* (2003). The genome sequence of *Bacillus anthracis* Ames and comparison to closely related bacteria. *Nature* 423, 81-86.

Reenan, R.A., and Kolodner, R.D. (1992a). Characterization of insertion mutations in the *Saccharomyces cerevisiae* MSH1 and MSH2 genes: evidence for separate mitochondrial and nuclear functions. *Genetics* 132, 975-985.

Reenan, R.A., and Kolodner, R.D. (1992b). Isolation and characterization of two *Saccharomyces cerevisiae* genes encoding homologs of the bacterial HexA and MutS mismatch repair proteins. *Genetics* 132, 963-973.

Robinson, K.A., and Beverley, S.M. (2003). Improvements in transfection efficiency and tests of RNA interference (RNAi) approaches in the protozoan parasite *Leishmania*. *Mol Biochem Parasitol* 128, 217-228.

Rodrigues, M.M., de Alencar, B.C., Claser, C., Tzelepis, F., Silveira, E.L., Haolla, F.A., Dominguez, M.R., and Vasconcelos, J.R. (2009). Swimming against the current: genetic vaccination against *Trypanosoma cruzi* infection in mice. *Mem Inst Oswaldo Cruz* 104, 281-287.

Rodrigues Coura, J., and de Castro, S.L. (2002). A critical review on Chagas disease chemotherapy. *Mem Inst Oswaldo Cruz* 97, 3-24.

Russo, M.T., De Luca, G., Casorelli, I., Degan, P., Molatore, S., Barone, F., Mazzei, F., Pannellini, T., Musiani, P., and Bignami, M. (2009). Role of MUTYH and MSH2 in the control of oxidative DNA damage, genetic instability, and tumorigenesis. *Cancer Res* 69, 4372-4379.

Russo, M.T., De Luca, G., Degan, P., and Bignami, M. (2007). Different DNA repair strategies to combat the threat from 8-oxoguanine. *Mutat Res* 614, 69-76.

Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989). *Molecular Cloning - A Laboratory Manual*, 2nd Edition, Vol 1 (New York, Cold Spring Harbour Laboratory Press).

- Shibutani, S., Takeshita, M., and Grollman, A.P. (1991). Insertion of specific bases during DNA synthesis past the oxidation-damaged base 8-oxodG. *Nature* 349, 431-434.
- Siegel, T.N., Tan, K.S., and Cross, G.A. (2005). Systematic study of sequence motifs for RNA trans splicing in *Trypanosoma brucei*. *Mol Cell Biol* 25, 9586-9594.
- Slupphaug, G., Kavli, B., and Krokan, H.E. (2003). The interacting pathways for prevention and repair of oxidative DNA damage. *Mutat Res* 531, 231-251.
- Smith, M., Blanchette, M., and Papadopoulou, B. (2008). Improving the prediction of mRNA extremities in the parasitic protozoan *Leishmania*. *BMC Bioinformatics* 9, 158.
- Sonnhammer, E.L., Eddy, S.R., and Durbin, R. (1997). Pfam: a comprehensive database of protein domain families based on seed alignments. *Proteins* 28, 405-420.
- Souto, R.P., Fernandes, O., Macedo, A.M., Campbell, D.A., and Zingales, B. (1996). DNA markers define two major phylogenetic lineages of *Trypanosoma cruzi*. *Mol Biochem Parasitol* 83, 141-152.
- Souto, R.P., and Zingales, B. (1993). Sensitive detection and strain classification of *Trypanosoma cruzi* by amplification of a ribosomal RNA sequence. *Mol Biochem Parasitol* 62, 45-52.
- Stern, M.Z., Gupta, S.K., Salmon-Divon, M., Haham, T., Barda, O., Levi, S., Wachtel, C., Nilsen, T.W., and Michaeli, S. (2009). Multiple roles for polypyrimidine tract binding (PTB) proteins in trypanosome RNA metabolism. *RNA* 15, 648-665.
- Stevens, J.R., Noyes, H.A., Schofield, C.J., and Gibson, W. (2001). The molecular evolution of Trypanosomatidae. *Adv Parasitol* 48, 1-56.
- Stojic, L., Brun, R., and Jiricny, J. (2004). Mismatch repair and DNA damage signalling. *DNA Repair (Amst)* 3, 1091-1101.
- Struthers, L., Patel, R., Clark, J., and Thomas, S. (1998). Direct detection of 8-oxodeoxyguanosine and 8-oxoguanine by avidin and its analogues. *Anal Biochem* 255, 20-31.
- Stuart, J.A., and Brown, M.F. (2006). Mitochondrial DNA maintenance and bioenergetics. *Biochim Biophys Acta* 1757, 79-89.
- Tamura, K., Dudley, J., Nei, M., and Kumar, S. (2007). MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol Biol Evol* 24, 1596-1599.
- Teixeira, A.R., Nascimento, R.J., and Sturm, N.R. (2006). Evolution and pathology in chagas disease--a review. *Mem Inst Oswaldo Cruz* 101, 463-491.
- Teixeira, S.M., and daRocha, W.D. (2003). Control of gene expression and genetic manipulation in the Trypanosomatidae. *Genet Mol Res* 2, 148-158.
- Teixeira, S.M., Kirchhoff, L.V., and Donelson, J.E. (1995). Post-transcriptional elements regulating expression of mRNAs from the amastin/tuzin gene cluster of *Trypanosoma cruzi*. *J Biol Chem* 270, 22586-22594.
- Thompson, J.D., Higgins, D.G., and Gibson, T.J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* 22, 4673-4680.
- Tibayrenc, M. (1995). Population genetics of parasitic protozoa and other microorganisms. *Adv Parasitol* 36, 47-115.

- Tibayrenc, M. (2003). Genetic subdivisions within *Trypanosoma cruzi* (Discrete Typing Units) and their relevance for molecular epidemiology and experimental evolution. *Kinetoplastid Biol Dis* 2, 12.
- Tibayrenc, M., Ward, P., Moya, A., and Ayala, F.J. (1986). Natural populations of *Trypanosoma cruzi*, the agent of Chagas disease, have a complex multiclonal structure. *Proc Natl Acad Sci U S A* 83, 115-119.
- Tornier, C., Bessone, S., Varlet, I., Rudolph, C., Darmon, M., and Fleck, O. (2001). Requirement for Msh6, but not for Swi4 (Msh3), in Msh2-dependent repair of base-base mismatches and mononucleotide loops in *Schizosaccharomyces pombe*. *Genetics* 158, 65-75.
- Torres-Silva, C.F. (2009). Reparo de DNA em dois patógenos humanos: caracterização do gene IMP4 de *Schistosoma mansoni* e estudos acerca do MMR, sistema GO e taxa de mutação em *Trypanosoma cruzi*. In Departamento de Bioquímica e Imunologia (Belo Horizonte, Universidade Federal de Minas Gerais).
- Trojan, J., Zeuzem, S., Randolph, A., Hemmerle, C., Brieger, A., Raedle, J., Plotz, G., Jiricny, J., and Marra, G. (2002). Functional analysis of hMLH1 variants and HNPCC-related mutations using a human expression system. *Gastroenterology* 122, 211-219.
- Ullu, E., Tschudi, C., and Chakraborty, T. (2004). RNA interference in protozoan parasites. *Cell Microbiol* 6, 509-519.
- Urbina, J.A., and Docampo, R. (2003). Specific chemotherapy of Chagas disease: controversies and advances. *Trends Parasitol* 19, 495-501.
- Vanhamme, L., and Pays, E. (1995). Control of gene expression in trypanosomes. *Microbiol Rev* 59, 223-240.
- Vazquez, M.P., and Levin, M.J. (1999). Functional analysis of the intergenic regions of TcP2beta gene loci allowed the construction of an improved *Trypanosoma cruzi* expression vector. *Gene* 239, 217-225.
- Wang, G., Alamuri, P., Humayun, M.Z., Taylor, D.E., and Maier, R.J. (2005). The *Helicobacter pylori* MutS protein confers protection from oxidative DNA damage. *Mol Microbiol* 58, 166-176.
- Weatherly, D.B., Boehlke, C., and Tarleton, R.L. (2009). Chromosome level assembly of the hybrid *Trypanosoma cruzi* genome. *BMC Genomics* 10, 255.
- Westenberger, S.J., Barnabe, C., Campbell, D.A., and Sturm, N.R. (2005). Two hybridization events define the population structure of *Trypanosoma cruzi*. *Genetics* 171, 527-543.
- Weston, D., La Flamme, A.C., and 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* 102, 53-66.
- Yamada, M., O'Regan, E., Brown, R., and Karran, P. (1997). Selective recognition of a cisplatin-DNA adduct by human mismatch repair proteins. *Nucleic Acids Res* 25, 491-496.
- Zingales, B., Souto, R.P., Mangia, R.H., Lisboa, C.V., Campbell, D.A., Coura, J.R., Jansen, A., and Fernandes, O. (1998). Molecular epidemiology of American trypanosomiasis in Brazil based on dimorphisms of rRNA and mini-exon gene sequences. *Int J Parasitol* 28, 105-112.
- Zingales, B., Stolf, B.S., Souto, R.P., Fernandes, O., and Briones, M.R. (1999). Epidemiology, biochemistry and evolution of *Trypanosoma cruzi* lineages based on ribosomal RNA sequences. *Mem Inst Oswaldo Cruz* 94 Suppl 1, 159-164.

1. Sequências gênicas dos alelos *Tcmsh2-B* e *Tcmsh2-C* e os sítios de anelamento dos iniciadores utilizados neste trabalho.
2. Artigo publicado no jornal *Memórias do Instituto Oswaldo Cruz* em 2005.
3. Artigo publicado no jornal *International Journal for Parasitology* em 2008.
4. Artigo publicado no jornal *Molecular and Biochemical Parasitology* em 2011.

Lócus *Tcmsh2* alelo B (selvagem)

TCACGAGTCATGTTGTTACTGTCTTCATGGCGAACATTCTCTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTGACCTCATCGAAGGGTGGGCGAGCGGCGGG
GAAGAAAAAAAATCCTACTGATGAACGGGACGTGTCTGTGCTGCAGGCGTTTACAGCCGTGGGGGGCGACGACAAGGCGTCCGTCCGTCTCTTTTTCGCGGGGTGGCTCTACTCCAGGGT
CTACGTCCTGGGCTCGTGGGCCGGATTTCATTGCGCGGGAGTACATCCGTCCTACTGCGGTGCTGAAGAACTGGGGCGGAGTGGATGTGGTGGTTCGTGAACGACAGTATTGCGAAGGAAGT
CATCCGTGATTTGTCCTTTCGCGGCGTGGTGCCTCTGTTCGAGTACTACGAACGACCACCCGGTGGTGGGACGTACGCATGTGTACAGCGTGGCTCACCGGGAACATTACCGACTTTGAGGC
AACGCTTTTTTTTCGTTTTGAAGAGGCGGAAATCCAGTTGATGGCAACTGGAGCAATTTGTTCTCGAGAACAAATAACAAAGGTGGCAACGCGGGTAAATGCATGATTTTCAGGCGGGCAAGGAAT
CAGGGTGGGATTTGCCGCACCTGAACACAACCTCTTCGGACCCCTGACTTACGCGGAGTATCTTGACACGCTGCAGCTGACCAGTCTTGATGCCTTGATGGCGCAATGCAATCTGAAGGAGCT
TATCTATTTGGACAAGAGCCGATGCAGAGGTGGCAGCAATAGTGATAGGGGAGGAGAGCCGAGGTGTTGCCGGCGAAGCGGGCGAGAAGAAGAGTCTTACGGGCGGTGAAACAGTT
GTGTGAGCGTGC AAAGATTACTTACCCTGAATTGGGACACAATGGCGCTGCTCCATCACATGATTCTCCGCAGAAACCCCTACAGCTGCAGGGAAGACGAAAGGCAACGTGACC AAAG
CGATTTTCTCAGTGCCTTGAAGACATTTCTTCGTGTGCCAGAGGACCGGCTTGCCTTTTCCAATTGTCCATTGGCGTCTCGAGCGATTGAATATCTTGTCTCTAATAATTATTGATAACTT
TGATGCAACGAATCACCGTGCCTTCTACTTAAAGCATAACAATCCATCCACCTTTATGAAACTGGACACCGCGGCAATACAGGCATTGCACATTATACACCAGAAACCAGAAGCGCGCGG
TTCCCTTCCCACATCTGTCTATTTCGTGGCTGAACCGCTGCGTTACCGGCATGGGCTCTCGGATGATGCGGCAGTGGTTGCTTTCAGCCGCTGCGCAACGCAGAGGACATCAATCAGCGTCT
TTCCCTTGTAGAATTGATGGTGAAGACTCGATTTTGC GGATGCTCTTCTTTTCGAGGTGCTGCGATGTTGTGGTGACATGGATCGGTTAAATCGCAAGCTTCAACGACGTAGCATCGC
TCTCAAAGATCTTCAGTCCATTTCTGACGTTTGTCAACACCATTCCGCGGGCCGTCAGGTGTTGAGGACCCACCAAGGCGGGCGAAATGACAAACTGCTGATGGATGAATACATCGCTCC
ATTTGGAGGACATTAACGAGCATTTCTCAAACCTCCGCATTTCTTATCACAGCCACTGTTGACCTCTCCGACGAAAACACTACTCGCATCAACCCCTGAATTTGATGACGAACTGATGGA
GGAGGAACAACGAAAAGCGTGGTCAAGGCCATTGAAAGCGAACACCAACCGTGATGAAAGTGTATGGATGGACAGAGAAGCAACTGAAGTGTGAGTACCATACGACATACGGTTATGT
CTTTTCGTGTAACGCGAAAAGAAGACCAACAGGTGCGCACAAGCAAGGAAC TAATCACGGTGAGCACTTCCAAGGACGGTGTAAAGATTTGTTTCAGAACGACTTTCTTCACTTAGTGAGCA
ATATAAAGGCATTCGAAAGGTATATGATGTGCGCCAGCAAGATCTTAAGCAGAAGCTTGTAAAGCACCCTGTCACCTTATCTTCCAGTACTTGTATGATGCAAAGGAGTTAATGCGGCGCT
GGATGTTTTTGTGGCCTGGGCAACGGTGGTAAGGGATTTCTCCACATCCCATGGTTTCGTCCAACGATTAGGACACCTGAGACAGAGGAGGAACAAGAGGGCAATAAAAAGTCTCATCACGCT
TCTTAATGTTTCGTCACCCGCTTGTGTAACCTCCGTCAGCCAGTGTACACACCCAACACGTTACGTCCTAACGGACGATGCCAATGCGCTTATTATAACGGGACCCAAATGGGAGGCAATC
AACTTTTATGCGGAGCGTGGGAATATGTGTTGTGCTGGCACAGGCAGGCTGCTTTGTGCCGGCCGATTCGGCGGATGTGGTAACAAGGGACCGCGTAATGTGTGCTGTTGGAGCCACGG
TCACCTTGACAAGGAGTGTGACGTTTATGGTGGAGATGTTGGAGTCTGCTGCCATTTTAAACGCTGCCACGCGAGATTCACTTGCCATTATTGATGAACTTGGCCGCGGAACGTCTAC
TTACGATGGGTTTTGGTTTTGGCTTGGGCCATTGCACAGGAGGTGGCGGTAAGGCTCGTTCTGCGCTTCTTTTTTTCTACCCACTTCCATGAGATGACACAGCTAGCAGAGCATCACACCAA
TGTGAGGAATGCGCACTTTGGCGCCGAAGTCAACACAGTTCGAGGGGACACTGCGCTTTTTCTTACCGGCTGGAGCCTGGGCGGTGGTTCGCAGTTACGGTCTGTATGTGGCCAGCTGGC
CAACCTCCCAGAGGAAGTTGTGCAATCTGCCAAGTGC AAGGCGGCAGGTTGGAAACCTTTGAGCAGGACGAGGGGCAAAAAGGACTACACAACCTATCTCTTCTGCTTCTTCAGAGAT
CACCGAACGACTTGCTTACTATGCAAAATGCATACGTGGGCTAAATGACACCAAATTTGGCGGCTGACAATGAGGAGGCGAGTCCGCGGCTGCGTCTTGAATCCAGCAAGACTCTCTTAT
ATCTTCTTTGCTT TAA

Legenda:

Em azul (letras brancas) – iniciador 8158 direto

Em verde – códon inicial de *Tmsh2*

Em preto (letras brancas) – sonda usada nos experimentos de *Southern* e *Northern blot*

Em amarelo – iniciador 8158 reverso

Em ciano – iniciador TmutS30 direto

Em cinza – iniciador TmutS41 reverso

Em vermelho – códon de terminação de *Tcmsh2*

Locus *Tcmsh2* alelo C (selvagem)

GTGGAATGGGTTGCTCGTGCTTGTTCCTTTTTCTTCTTTTTCTTTTGACCTCAGCGAAGGGTGGGCGAGCGGCGGGGAAAAAAAAAAAAAAAAAGGAAGAAAAATGACTGATGAACGGGACG
TGTCGTGCTGCAGGCGTTTTACAGCCGTGGGGGGCGACGACAAGGCGTGCGTCCGTCTCTTTTTCGCGGGGTGGCTCTACTCCGGGTGCTACGTCTTGGGCTCGTGGGCGGATTCATTG
CGCGGGAGTACATCCGGTCCACCGCGGTGTTGAAGAACTGGGGCGGAGTGGATGTGGTGGTCTGTAACGACAGTATTGCGAAGGAAGTCATCCGTGATTTGTCTTCTGCGGGCGTGGTGCC
CTGTGAGTACTACGAACGACCACCGGTGGCGGGACGTACGCATGTGTACAGCGCGGCTCACCGGGGAACATTACCGACTTTGAGGCAACGCTTTTTTCGTTTGAAGAGGCGGAAATCC
AGTTGATGGCAACTGGAGCAATTTGTTCTTGAGAAAAATAACACAAGTGGCAACGCGGGTAAATGCATGATTTTCAGGCGGGCAAGGAATCAGGGTGGGATTTGCCGCACTGAACACAACCTC
TTCGGACCCCTGATTTACGCGGAGTATCTTGACACGCTGCAGCTGACCAGTCTTGATGCCCTTGATGGCGCAGTGAATCTGAAGGAGCTTATCTATTTGGACAGGAGCCGATGCAGAAATTA
GCAGCAATAGTAACAAGGGAGGAGGAGCCGAGGTGTTGCCGGCGGAAGCGGGCGAGAAGAAGAGTCGTTACGGGCGGTGAAACAGTTGTGTGAGCGTGCAACATTACTTACCGTGAAT
TGGGACACAATGGCGCTGCTCCATCACATGATTTCTCCGCAGAAACCCCTACAGCCGAGGGAAGACGAGAGGGCAACGTGACCAAAGGCGATTTTCTCAGTGCTCTTGAAGACATTCTTC
GTGTGCCAGAGGACCGGCTTGCCCTTTCTAACGTGCCATTGGCGTCTCGAGCGATTGAATATCTTGTCTCTAATATTATTGATAACTTTGACGCAACGAATCACCGTGCTTTCTACTTAA
AGCATAAATTTTCATCCACCTTTATGAAACTGGACACCGCGCAATACAGGCATTGCACATTATCCACCAGAAACCAGAAGCGCGGGTCCCTTCCCACATCCGCTCTATTCGTGGCTGA
ACCGCTGCGTTACCGGCATGGGCTCTCGGATGATGCGGCAGTGGTTGCTTTCAGCCGCTGCGCAACGCAGAGGAAATCAATCAGCGCTTTTCCCTTGTAGAAATTGATGGTGGAAAGACTCGA
TTTTGCGGGATGCTCTTCTTTTCGCAAGTGTACGATGTTGTGGTACATGGATCGTTTTAAATCGCAAGCTTCAACGACGTAGCATCGCTCTCAAAGATCTTCAGTCCATTCGACGTTTG
TCAACACCATTTCCGCGGGCCGTCAGGTGTTGAGGACCCACCAAGGCGGGCGAAATGACAACTGCTGATGGATGAATACATCGCCCCATTGGAGGACATTAACGAGCATTTCTCAAACC
TCCGCATCCTTATCACAGCAACTGTTGACCTCTCCGACGAAAACACTACTCGCATCAACCCCTGAATTTGATGACGAACTGATGGAACCTAGAAGAACAAACGAAAAAGCGTGGTCAAGGCCA
TTGAAAGCGAACACCAACGCGTGATGAAAGTGTATGGATGGACAGAAAAGCAACTGAAGTGTGAGTACCATACGACATACGGTTATGTCTTTCGTTGTAACGCGAAAAAGAACCAACAGG
TGCGCACAAGCAAGGAACATAATTACGGTGAGCACTTCCAAGGACGGTGTAAAGATTTGTTTTCAGAACGACTTTCTTCACTTAGTGAGCAATATAAAGGCATTCGAAAGGTATATGATGTGC
GTCAGCAAGATCTTAAGCAGAAGCTTGTAAAGCACCCTGCTCATTATCTGCCAGTACTTGATGATGCAAAGGAGTTGATTTGCGGCGCTGGATGTTTTTGTGGCCTGGGCAACGGTGGTAA
GGGATTTCTCCACATCCCATGGTTGCTCCAACGATTAGGACACCTGAGACAGAGGAGGAACAAGAGGGCAATAAAAGTCTTATCACGCTTCTTAATGTTTCGTCACCCGCTTGTGAACTCC
GTCAGCCAGTGTACACACCAACACGTTACATCTAACGGACGATGCCAATGCGCTTATTATAACGGGACCAACATGGGAGGCAATCAACTTTTATGCGGAGCGTGGGAATATGTGTTG
TGCTGGCACAGGCAGGCTGCTTTGTGCCGGCCGATTCGGCGGATATGGTAACAAGGGACGCCGTATGTGTGCTGTTGGAGCCACGGATCACCTTGCACAAGGGGTGTCGACGTTTATGG
TGGAGATGTTGGAATCTGCTGCCATTTTAAACGCTGCTACGCGAGATTACCTTGCCATTATTGATGAACTTGGCCGCGGAACGTCTACTTACGATGGGTTTGGTTTGGCTTGGGCCATTG
CACAGGAGGTGGCGTAAAGGCTCGTTCTGCGCTTCTTTTTCTACCCACTTCCATGAGATGACACAGCTAGCAGAGCATCACACCAATGTGAGGAATGCGCACTTTGGCGCCGAAAGTCA
ACACGGTTCGAGGGGACACTGCGCTTTTTCTTACCGACTGGAGCCTGGGCCGTGTGGTTCGAGTTACGGTCTGTATGTGGCCAGCTGGCCAACCTCCAGAGGAAGTTGTGCAATCTGCTA
AGTGCAAGGCGGCAGGGTTGAAACCTTTGAGCAGGACGAGGGGCAAAAAGGGACTACGCAACCATCTCTTCTGCTTCTTTCAGAGATCACCGAACGACTTGTCTACTATGCAAAATGCA
TACGTGGGCTAAATGACACCAAATGGCGGCTGGCAATGAGGAGGCGAGTCAGCGGCTGCGTCTTGAATCCAGCAAGACTCTCTCATATCGTATTTGCTTAA

Legenda:

Em rosa – iniciador 7442 direto

Em verde – códon inicial de *Tmsh2*

Em preto (letras brancas) – sonda usada nos experimentos de *Southern* e *Northern* blot

Em amarelo – iniciador 7442 reverso

Em ciano – iniciador TmutS30 direto

Em cinza – iniciador TmutS41 reverso

Em vermelho – códon de terminação de *Tcmsh2*

Locus *Tcmsh2* alelo C (gene Neo^R integrado)

GTGGAATGGGTTGCTCGTGC TGTTC TTTTTTCTTCTTTTCTTTTGACCTCAGCGAAGGGTGGGCGAGCGGCGGGGAAAAAAAAAAAAAAAAAGGAAGAAAAATATC ACTGATGAACGGGACG
TGCTGTGCTGCAGGCGTTTACAGCCGTGGGGGGCGACGACAAGGCGTGCCTCCGTCCTTTTTCGCGGGGTGGCTCTACTCCGGGGTGTACGTCCTGGGCTCGTGGGCCGGATTCAATTG
CGCGGGAGTACATCCGCTCCACCGCGGTGTGAAGAACTGGGGCGGAGTGGATGTGGTGGTCTGTAACGACAGTATTCGGAAGGAAGTCATCCGTGATTTGCTTTCTGCGGCGTGGTGCCT
CTGTGAGTACTACGAACGACCACCCGGTGGCGGGACGTACGCATGTGTACAGCGCGGCTCACCGGGGAACAATTACCGACTTTGAGGCAACGCTTTTTTCTTTGAGAGGGCGGAAATCC
AGTTGATGGCAACTGGAGCAATTGTTCTTTGAGCCATTTACGACTCCAAGGCAACGCTGCAGAACAACCTGCCGAAGGAGCGCCGTTCTTCAAGATTTGTGTCTGTGGTACGACAACGAGTG
GGGATAC TCCACCGCGTGGTGGACCTTGTACGCCACATGGCC TCGAAGGATCGTTTCGGCAAGGTTGTAGGCGTGGCGATGACTTCAGGTC TTTCTTTTGGCAATAGGGATCTTATAATA
CACGATGCGTGTCCCGTGTATGATCGTTACC GGTTGCCACGATCCAATTGACACAGCGTCAAGAGCAAAAACAATTTTACTTTTCCCTTTAAGGACAACAACAAAAAATATATAACTTT
TTTTTCTTTTTTTTTTTTTTTTTTGA AATTATATTTATGGTCATCTTTGGGAAACAAAAAGCAGCAATTTAATGATGCGGAAGGATGAGTGAATAAATGTTAATCAATGTACGAGGATTTG
GGGTATTGCAAGGAAAATGTAGATGATTTAATTGGGTGTGTGATGCAGCTTGTGGTAATTTTTTGCTCACTTCCCTTTTTTGCCACATCTTTTTAGTTTTTCTGCTTTTCTTTCCCATTA
TCCACTTTCCTCTCTTTTCCCAGTTTCTGACGAATGCAGAAAGTGATATTTTTACTTTTGAAGCCATCTACCAACAACAATTACATTGAACAGAATTTGGGATTGCGAATTAATTCCT
GCAGCCCGCGTGTGGCTCGAACACCGAGCGACCTGCAGCCAATATGGGATCGGCCATTGAACAAGATGGA TGGCAGCGAGGTTCTCCGGCCGCTTGGGTGGAGAGGCTATTCGGCTA
TGACTGGGCACAACAGACAATCGGCTGCTCTGATGCCGCGGTGTTCGGCTGTGACGCGAGGGGCGCCCGTTCTTTTTGTCAAGACCGACTGTCCGGTGCCCTGAATGAACTGCAGGA
CGAGGCAGCGCGGCTATCGTGGCTGGCCACGACGGGCGTTCCTTGGCGAGCTGTGCTCGACGTTGTCTGACGAAGCGGAAGGGACTGGCTGCTATTTGGGCGAAGTGC CGGGGCGAGGATCT
CCTGTCTCTCACCTTGTCTCCGCGAGAAAGTATCCATCATGGCTGATGCAATGCGGCGGCTGCATACGCTTGATCCGGCTACCTGCCCATTCGACCACCAAGCGAAACATCGCATCGA
GCGAGCACGTACTCGGATGGAAGCCGGTCTTGTGATCAGGATGATCTGGACGAAGAGCATCAGGGGCTCGCGCCAGCCGAACTGTTCCGACGGCTCAAGCGCGCATGCCGACGGCGA
GGATCTCGTGTGACCCATGGCGATGCCGTGCTTGGCGAATATCATGGTGGAAAATGGCCGCTTTTCTGGATTTCATCGACTGTGGCCGGCTGGGTGTGGCGGACCGCTATCAGGACATAGC
GTTGGCTACCCGTGATATTGCTGAAGAGCTTGGCGGCGAATGGGCTGACCGCTTCCCTCGTGTCTTTACGGTATCGCCGCTCCCGATTGCGAGCGCATCGCCTTCTATCGCCTTCTTGACGA
GTTCTTCTGA GGGGATCGATCCGGAACAACCTGCCGAAGGAGCGCCGCTTCTTCAAGATTTGTGTCTGTGGTACGACAACGAGTGGGGATACTCCCACCGCGTGGTGGACCTTGTACGCCAC
ATGGCTCGAAGGATCGTTCCGCAAGGTTGTAGGCGTGGCGATGACTTCAGGTC TTTCTTTTGGCAATAGGGATCTTATAATACAGATGCGTGTCCCGTGTATGATCGTTACC GGTCG
CCACGATCCAAGTGACACAGTGTCAAGAGCAAAAACAATTTTACTTTTCCCTTTAAGGACAACAACAAAAAATATATAACTTTTTTTTTTCTTTTTTTTTTTTTTGA AATTATATTTATGGTC
ATCTTTGGGAAACAAAAAGCAGCAATTTAATGATGCGGAAGGATGAGTGAATAAATGTTAATCAATGTACGAGGATTTGGGGTATTGCAAGGAAAATGTAGATGATTTAGTTGGGTGTA
TGATGCAGCTTGTGGTAATTTTTTACAAC TTTGCGTTATTGGAAGATTTTGGTTGTTTCAAGTTCTTTTTGTTTATAACTGACTGTTTTAATTTAGTGTGATTTTTTTTTTTTTAAGGATTT
AAGCTGTTTTTCTTTTTTTTTTCTTTTTCTTCACTGATGGTGTTTAAGGTGTAGTTTAGTGTGTTGTGCATTTATGCGTCGTAGTGTGTTGTTCTTTGGACTTTTTCGAACAACTG GGCA
TTATAGCCCACCTGTGTATATGCCGCTAGAGTATGCTAGCAGAGCATCACACCAATGTGAGGAATGCGCAC TTTGGCGCCGAAGTCAACACGGTTCGAGGGGACACTGCGCTTTTCTTACC
GACTGGAGCCTGGGCGGTGTGGTTCGAGTTACGGTCTGTATGTGGCCAGCTGGCCAACCTCCAGAGGAAGTTGTGCAATCTGCTAAGTGCAAGGCGGAGGGTTGGAAACCTTTGAGC
AGGACGAGGGGCAAAAAAGGACTACGCAACCATCTCTTCTGCTTCTT CAGAGATCACCGAACGACTTGTCTACTATGCAAAATGCATACGTGGGCTAAATGACACCAAATTTGGCGGCTG
GCAATGAGGAGGCGAGT CAGCGCTGCGTCTTGA AATCCAGCAAGACTCTCTCATATCGTATTTGCTT TAA

Legenda:

- Em rosa – iniciador 7442 direto
- Em verde – códon inicial de *Tmsh2*
- Em preto (letras brancas) – sonda usada nos experimentos de *Southern* e *Northern blot*
- Em amarelo (letras azuis) – iniciador NeoN11 reverso
- Em vermelho – códon de terminação de *Tcmsh2*

Analysis of expressed sequence tags from *Trypanosoma cruzi* amastigotes

Gustavo C Cerqueira⁺⁺, Wanderson D DaRocha⁺⁺, Priscila C Campos, Cláudia S Zouain, Santuza MR Teixeira⁺

Departamento de Bioquímica e Imunologia, ICB, Universidade Federal de Minas Gerais, Av. Antônio Carlos 6627, 31270-010 Belo Horizonte, MG, Brasil

A total of 880 expressed sequence tags (EST) originated from clones randomly selected from a Trypanosoma cruzi amastigote cDNA library have been analyzed. Of these, 40% (355 ESTs) have been identified by similarity to sequences in public databases and classified according to functional categorization of their putative products. About 11% of the mRNAs expressed in amastigotes are related to the translational machinery, and a large number of them (9% of the total number of clones in the library) encode ribosomal proteins. A comparative analysis with a previous study, where clones from the same library were selected using sera from patients with Chagas disease, revealed that ribosomal proteins also represent the largest class of antigen coding genes expressed in amastigotes (54% of all immunoselected clones). However, although more than thirty classes of ribosomal proteins were identified by EST analysis, the results of the immunoscreening indicated that only a particular subset of them contains major antigenic determinants recognized by antibodies from Chagas disease patients.

Key words: *Trypanosoma cruzi* - gene expression - amastigotes - antigens

Chagas disease, caused by the protozoan *Trypanosoma cruzi*, is a parasitic disease, which affects about 16 million people in Latin America. Although vectorial and transfusional transmission in most countries is now under control, chronic *T. cruzi* infections still result in serious morbidity and mortality, mainly due to heart tissue damage observed in a significant percentage of chronically infected individuals (WHO, www.who.ch).

T. cruzi has three main developmental stages. Epimastigotes multiply extracellularly in the midgut of the reduviid bug vector and migrate to the hindgut, where they differentiate into non-dividing metacyclic trypomastigotes, that are excreted with the insect feces. Infective trypomastigotes, inoculated into the mammalian host through mucous membranes and skin cuts, invade a number of cell types and differentiate into amastigotes, which multiply intracellularly in the cytoplasm of the parasitized cell. After a certain number of cell divisions, amastigotes transform into trypomastigotes that enter the circulatory system when the host cell lyses. Bloodstream trypomastigotes invade new host cells and can be transmitted to a reduviid bug during a blood meal where they differentiate again into epimastigotes and complete the life cycle (Brener 1973).

Knowledge of the gene expression pattern of each developmental stage is a valuable contribution to the study of the parasite biology and the pathogenesis of Chagas disease. Expressed sequence tags (EST) analysis is a powerful approach that rapidly expands the knowledge not only towards gene discovery but also regarding stage specific gene expression profiles. EST data has also been useful in the final steps of genome annotation and gene prediction of whole genome sequencing projects. Beginning in 1994, the *T. cruzi* EST sequencing project has generated more than 10,000 ESTs, mainly derived from a normalized cDNA library constructed with RNA from epimastigote cultures (Brandão et al. 1997, Verdun et al. 1998, Porcel et al. 2000) and, more recently, from trypomastigotes and amastigotes of the CL Brener strain (Agüero et al. 2004) (<http://cruzi.cgb.ki.se/TcruziEST/index.html>). After an intense effort initiated by an International Consortium of three sequencing centers (The Institute for Genomic Research, TIGR, Seattle Biomedical Research Institute, and Karolinska Institute), the complete genome sequence of the 60 Mb haploid genome of the CL Brener clone is expected to be published shortly.

We have been interested in studying genes expressed in amastigotes, particularly the ones encoding amastigote antigens, which may be more relevant to the chronic chagasic pathology and are likely to be more useful as a source for antigenic targets for serodiagnosis and vaccine development (Matsumoto et al. 1993, Low et al. 1998, Boscardin et al. 2003). To generate more information regarding the amastigote gene expression profile and to identify new candidate targets for vaccine studies, we isolated cDNA clones derived from an amastigote library immunoselected with sera from patients with Chagas disease (DaRocha et al. 2002). Here we compared the results obtained in this previous study with the analysis of 880 EST originated from randomly chosen clones from the same cDNA library as well as with other *T. cruzi* EST

Financial support: CNPq

⁺Corresponding author and fellowship from Capes. E-mail: santuzat@mono.icb.ufmg.br

⁺⁺Both author have contributed equally to the work.

The sequence data reported herein has been deposited in GenBank under accession numbers listed at http://www.icb.ufmg.br/~nage/aest/supplemental_table.html.

Received 10 February 2005

Accepted 19 May 2005

datasets. Based on sequence comparisons, we report the putative functions of several novel *T. cruzi* genes expressed in amastigotes. The EST analyses also showed that clones encoding ribosomal proteins, which have been previously identified as one of the main targets for the humoral immune response of patients with Chagas disease (DaRocha et al. 2002), also constitute the largest class of cDNAs identified in the amastigote library.

MATERIALS AND METHODS

cDNA library - The unidirectional amastigote cDNA library was constructed in the bacteriophage λ -ZAPII according to the procedures indicated in the cDNA synthesis and in vitro phage packaging kit (Invitrogen), as described by Teixeira et al. (1994). Amastigote cells of the Tulahuén strain were obtained from cultures of RA786 cell line culture two weeks after infection and purified by centrifugation in metrizamide gradient also described by Teixeira et al. (1994). Based on rRNA and microsatellite typing, the Tulahuén strain belongs to the same *T. cruzi* II lineage of the clone CL Brener that was chosen for the genome project (SMR Teixeira, unpublished). In vivo excision protocol using the *Escherichia coli* strain XL-1 Blue was conducted with an aliquot of phage supernatant according to the cDNA library manual (Invitrogen). Transformed bacteria were plated in medium containing X-gal and about 1200 individual colonies, identified as β -galactosidase negative, were randomly picked, inoculated into 96 deep well microtiter plates and submitted to double-strand plasmid DNA preparation according to previously published protocols (Vasconcelos et al. 2003).

Nucleotide sequencing - Plasmid DNA was sequenced in 96-well microplates with "DYEnamic™ ET dye terminator cycle sequencing (MegaBACE™)" and the MegaBACE 1000 capillary sequencer (Amersham Pharmacia Biotech). Single-pass sequencing reactions were performed on each template with M13 Forward and Reverse primers, and sequences longer than 100 bases were further analyzed.

Sequence analyses - The chromatogram produced for each clone was analyzed by a software developed by us, named AEST (stands for *Análise de ESTs de Tripanosomatídeos*, in portuguese). AEST is a web-based application, which automates the processing of ESTs through the integration of several softwares that are routinely used for DNA sequence analyzes. AEST executes a pipeline consisting of the following softwares: (1) phred, for base-calling processing; (2) cross_match, for screening vector sequences; (3) BLASTCLUST, for clustering sequences; and (4) BLASTX and BLASTN for searching similar sequences in the GenBank database (Altschul et al. 1990, Erwing et al. 1998). In the current EST analysis, phred was executed with trim parameter turned on, trim value 0.18; cross_match penalty was adjusted to -2 and minscore to 14. BLASTX and BLASTN queries were made against GenBank nonredundant database with the default parameter, except from the low complexity filter, which was turned off. BLASTCLUST executions were performed with the following parameters: -L 0.7 (length coverage), -S 96 (score density), and -W 40 (word size). After automated

analysis, the resulting sequences were manually inspected with the following objectives: to assign, if possible, a putative product according to the results of the similarity search (BLASTX and BLASTN) and to remove from further analyses low quality sequences or sequence contaminants, such as ribosomal RNA sequences. Each sequence with a positive identification product was classified according to its probable function in one of the following functional categories: cell cycle, RNA, and DNA synthesis; metabolism; protein destination; translation machinery; structural and cytoskeletal; signal transduction; transporter and cellular trafficking; and those sequences that do not fit in none of the previous categories were grouped in the "others" class. All 880 ESTs were submitted to the dbEST database but can also be accessed through the supplemental table available at http://www.icb.ufmg.br/~nage/aest/supplemental_table.html.

RESULTS AND DISCUSSION

Single-pass sequencing was performed from either the 5' or 3'-end of 1152 cDNA clones randomly picked from an unidirectional, non-normalized cDNA library constructed with poly-A (+) RNA isolated from *T. cruzi* amastigotes. After removal of low-quality sequences and vector sequences, a total of 880 ESTs were submitted to blast searches against the GenBank nucleotide database and assembled into clusters. The average length of reliable sequence was 510 nt. The *T. cruzi* SL sequence was detected in 19 of the selected clones and a poly-A tail was found in 379 of them. A total of 151 sequences (17% of the clones analyzed) were assembled into 53 clusters, whereas 729 ESTs represent singletons. The largest cluster contains 18 ESTs, and 12 clusters were found to contain more than 5 ESTs (see below). This clustering analysis indicates that the redundancy of the library estimated from the percentage of clones assembled into clusters is relatively low, considering the levels of redundancy reported by Porcel et al. (2000) (54%) and Verdun et al. (1998) (28%) for cDNAs libraries constructed with RNA derived from *T. cruzi* epimastigotes.

Almost 60% of the sequences (525 ESTs) did not show any significant similarities to sequences deposited in GenBank. It is thus possible that, even with a small-scale sequencing survey, a significant number of novel *T. cruzi* genes can be identified. Similarly to other reports of EST analyses 0.4% of the clones contain ribosomal RNA sequences, but a relatively large proportion (3%) contains known *T. cruzi* repetitive elements. The repetitive sequences identified were mainly SIRE repetitive elements (Vazquez et al. 1994), which have also been reported in previous analyses of epimastigote libraries. Stable transcripts derived from repeated elements in *T. cruzi* have been identified by several groups. It has been speculated that they may be involved with reshaping of the genome, or that they may affect the expression of neighboring genes by influencing *trans*-splicing reactions (Vazquez et al. 1994).

Fig. 1 shows the distribution among eight functional categories of all ESTs to which a putative identity could be assigned using Blast analyses. Similar to the EST distribution obtained in previous EST studies (Verdun et al.

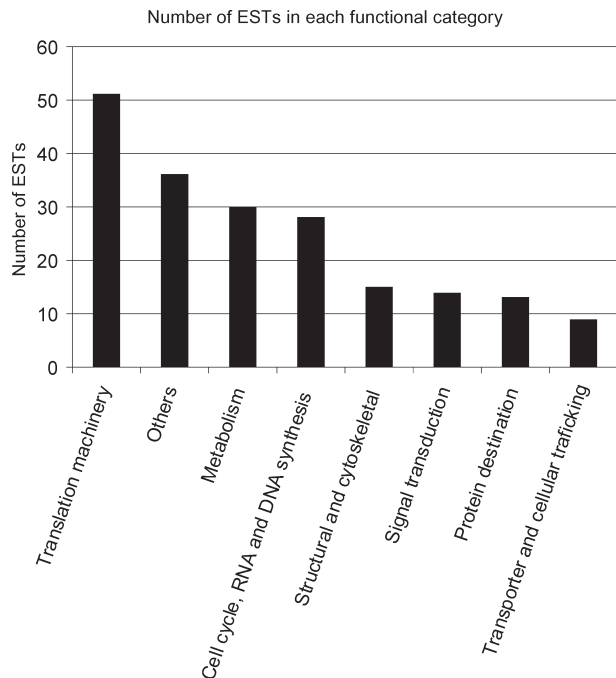


Fig. 1: functional distribution of amastigote expressed sequence tags (ESTs). The graph includes only non-redundant ESTs and shows the number of ESTs classified in each of the eight functional categories, according to the Blast search analysis. The category "others" comprises ESTs whose products could not be classified into any of these categories and includes detoxification enzymes, such as trypanodioxin and membrane and cell surface proteins. In the "protein destination" category we included proteases, proteasomes, chaperonins, and heat shock proteins. The full list of the identified ESTs and a detailed functional classification can be found at http://www.icb.ufmg.br/~nage/aest/supplemental_table.html

1998, Porcel et al. 2000, with epimastigote cDNA libraries and Agüero et al. 2004, with an amastigote cDNA library), a large fraction of mRNAs encoding components of the translation machinery is found in amastigotes. Because the two amastigote libraries were not normalized, we expected a significant number of cDNAs to correspond to abundant transcripts present in this stage of the parasite. In addition to ribosomal proteins, cDNAs encoding histones are among the most abundant sequences in the amastigote libraries. ESTs encoding histone H3 constitute the largest cluster we identified (2% of the total number of clones, Table) and were also detected as highly abundant cDNAs in the amastigote library described by Agüero et al. (2004) (corresponding to 1.2% of the total number of clones). Histone H3 was identified as the 2nd and 3rd most frequently identified EST in the two epimastigote cDNA libraries (1 and 0.6%, in the libraries described by Porcel and Verdun's group, respectively) as well as an abundant sequence in the trypanomastigote cDNA library reported by Agüero et al. (2004). It should be noted that, in contrast to other eukaryotes, histone transcripts from trypanosomatids are poly-adenylated, and thus, will be part of the cDNA population generated from oligo-dT primed reverse transcription reactions. Such high frequency of ESTs encoding histones in the epimastigotes

and amastigotes libraries is in accordance with northern blot analyses performed by Recinos et al. (2001) who shows that steady state histone mRNA levels are much higher in dividing epimastigotes and amastigotes, as compared to non-dividing trypanomastigotes. It is however surprising that the ESTs derived from the trypanomastigote library described by Agüero et al. (2004) also showed a relatively high proportion of histone sequences.

Besides histones and ribosomal proteins, sequences homologous to cytochromes, and heat-shock proteins are among the largest clusters identified in the amastigote EST dataset (Table). We also identified a sizeable number of clones derived from large *T. cruzi* gene families such as mucins and trans-sialidase (GP85), known to be expressed in amastigotes (Acosta-Serrano et al. 2001, Frasch 2001). As expected, we identified the amastigote-specific amastin mRNA as one of the six most abundant cDNA clusters in the amastigote library. This frequency is consistent with northern blot results showing a 50-fold increase of the levels of amastin mRNA expression in amastigotes, as compared to the other two forms of the parasite (Teixeira et al. 1994). A complete list of all ESTs is available at http://www.icb.ufmg.br/~nage/aest/supplemental_table.html.

TABLE

Most abundant amastigote expressed sequence tags (ESTs)	
Putative product	Number of ESTs
Histone H3	18
Ribosomal protein S9	15
Cytochrome-c oxidase chain III	13
Heat shock protein 90	8
Cytochrome b	7
GP85	7
Histone H2A	6
Amastin	6
Ribosomal protein S12	5
Elongation factor 1-alpha	5
Mucins	5
Casein kinase 1 isoform 2	5

The high frequency of ribosomal protein cDNAs (9% of the total number of cDNAs) is in agreement with the results of the amastigote EST survey reported by Agüero et al. (2004), who found an even larger percentage (17.4%) of total ESTs generated from the CL Brener amastigote library encoding ribosomal proteins. This proportion is also similar in the data set from the two epimastigote cDNA libraries (Verdun et al. 1998, Porcel et al. 2000), thus indicating that both forms of *T. cruzi* depend on the maintenance of the protein synthesis apparatus characteristic of rapidly dividing cells. In line with that, microarray studies of gene expression during trypanomastigote (a non-dividing stage) to amastigote differentiation also identified several genes encoding ribosomal proteins as a group of amastigote up-regulated genes (Minning et al. 2003). The most abundant mRNAs encoding ribosomal proteins expressed in amastigotes are S9 and S12. A cluster encoding the *T. cruzi* homologue of the S12 ribosomal protein,

which may present and additional extra-ribosomal function in the cell (Coetzee et al. 1994), was also identified as one of the most predominant cDNAs of the amastigote library described by Agüero et al. (2004).

Such high prevalence of clones encoding ribosomal proteins was also evident in our previous survey of the same amastigote cDNA library, using an immunoscreening protocol. Ribosomal proteins accounted for the large majority (54%) of all clones that react with a pool of sera from Chagas disease patients (DaRocha et al. 2002). Previous literature reports have indicated that ribosomal proteins constitute one of the main targets of the humoral immune response against infections not only by *T. cruzi* but also by other parasites such as *Leishmania chagasi* and *Plasmodium falciparum* (Levin et al. 1989, Skeiky et al. 1994, Chatterjee et al. 2000). This is in contrast to histone proteins, which are recognized as a prominent antigen in *Leishmania* and *Plasmodium falciparum* infections but not by *T. cruzi* infected individuals (Requena et al. 2000). Interestingly, similar to ribosomal proteins, histones also form nucleoprotein complexes and are also encoded by abundant mRNAs in amastigotes. Ribosomal P proteins are one of the most extensively studied *T. cruzi* antigens because antibodies directed against them, found in sera from patients with chronic Chagas heart disease, cross react with host proteins. Unlike other eukaryotes, *T. cruzi* possesses four types of P protein genes TcP0, TcP1, TcP2 α , and TcP2 β , each of which is present in multiple copies in the genome (Skeiky et al. 1994). The epitopes recognized by antibodies present in sera from Chagas disease patients, located at the C-terminal region of the *T. cruzi* ribosomal P proteins, share sequence homology with human P proteins and β -adrenergic receptors (Levin et al. 1989, Kaplan et al. 1997). As shown in Fig. 2, in the immunoscreening survey, we found four classes of ribosomal proteins reacting with sera from Chagas disease patients, among them, the previously characterized *T. cruzi* ribosomal P proteins TcP0 and TcP2 β . In addition, we iden-

tified the orthologues of two eukaryotic ribosomal proteins L19 and L7a, named TcRpL19 and TcRpL7a, as highly immunogenic. Thus, in contrast to the limited number of ribosomal sequences showing antigenic reactions, the EST data indicated that a much larger variety of ribosomal protein mRNAs are highly expressed in amastigotes, since ESTs encoding a total of 34 classes of ribosomal proteins were identified. While TcRpL19 and TcRpL7a are encoded by abundant mRNAs, corresponding, each, to 5% of the total number of ESTs containing ribosomal protein sequences, cDNAs encoding TcP0 and TcP2 β were not detected in our EST survey.

One way to explain why TcRpL7a and TcRpL19 are highly immunogenic is the fact that they contain repeated amino acid motifs. The repetitive Ala, Lys, and Pro-rich motifs are similar, but not identical, and are positioned at the N-terminal (TcRpL7a) or the C-terminal (TcRpL19) of the polypeptide chains (DaRocha et al. 2002). Interestingly, these repetitive sequences are not found in the orthologous genes of any other organism including *T. brucei* and *Leishmania*. Further studies using recombinant versions of the full-length protein and truncated forms containing only the repetitive and the non-repetitive portions of TcRpL7a, have demonstrated that the repetitive sequence is responsible for the reactivity of this antigen with sera from Chagas disease patients (Pais et al. 2002). It will be interesting to determine how the presence of this repetitive sequence affects the three-dimensional structure of these proteins as well as the overall structure of the ribosome. A search for repetitive motifs in other *T. cruzi* ribosomal proteins using the available CL Brener genomic database (www.genedb.org) showed that amino acids repeats are also present in ribosomal proteins S2, S10, and L23a, but only L23a has a repetitive motif similar to TcRpL7a and TcRpL19. Amino acid repeats have been recognized as a recurrent theme since the earlier studies characterizing *T. cruzi* and other parasite antigens, but their role still remains a subject of intense speculation

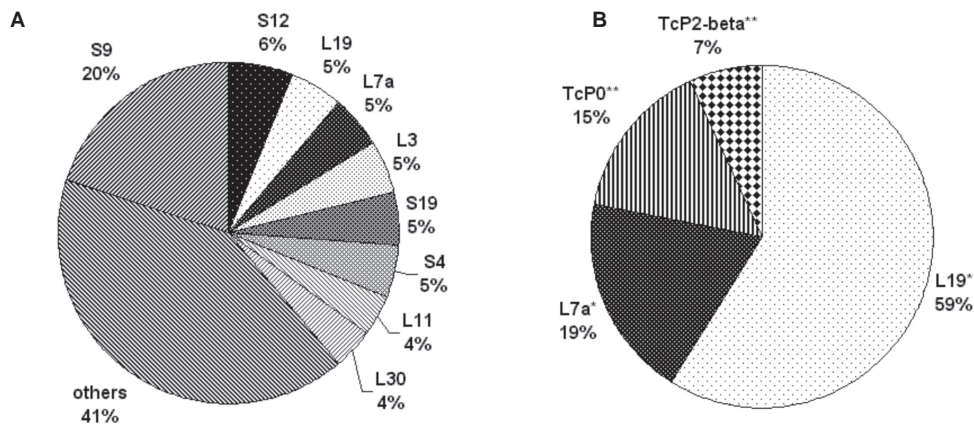


Fig. 2: comparison of expressed sequence tag (EST) and immunoscreening approaches for the identification of cDNAs encoding ribosomal proteins in the amastigote library. cDNA clones containing sequences homologous to ribosomal proteins that were identified by the EST analysis (A) and immunoscreening the amastigote cDNA library with sera from Chagas disease (B), were stratified based on their frequency. "Others" stands for cDNAs encoding ribosomal proteins identified as ESTs with a frequency lower than 2%. The values correspond to the percentage of each clone regarding the total number of ribosomal proteins. Single asterisks highlight clones present in both groups. Double asterisks correspond to cDNAs identified only by serum reactivity. Ribosomal proteins account for 9 and 54% of the total number of clones obtained by EST and immunoscreening protocols, respectively.

(Hoft et al. 1989). We are currently investigating whether the repetitive portion of L23a also makes this protein a member of the group of ribosomal antigens recognized by antibodies present in human chagasic sera.

Given that the completion of the sequencing of the *T. cruzi* genome, conducted by The International *T. cruzi* Sequencing Consortium, is within sight, all the information about the estimated 10,000 genes harbored by this parasite will soon be available. Together with other EST studies, this amastigote EST survey represents an additional contribution to future studies on the characterization of genes that may have a potential application towards the development of new methods to control Chagas disease.

ACKNOWLEDGMENTS

To Rosiane A Silva and Dr Daniella C Bartholomeu for discussions and comments on the manuscript.

REFERENCES

- Acosta-Serrano A, Almeida IC, Freitas-Junior L, Yoshida N, Schenkman S 2001. The mucin-like super-family of *Trypanosoma cruzi*: structure and functional roles. *Mol Biochem Parasitol* 114: 143-150.
- Agüero F, Abdellah KB, Tekiel V, Sánchez DO, González A 2004. Generation and analysis of expressed sequence tags from *Trypanosoma cruzi* trypomastigotes and amastigotes cDNA libraries. *Mol Biochem Parasitol* 136: 221-225.
- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ 1990. Basic local alignment search tool. *J Mol Biol* 215: 403-410.
- Boscardin SB, Kinoshita SS, Fujimura AE, Rodrigues MM 2003. Immunization with cDNA expressed by amastigotes of *Trypanosoma cruzi* elicits protective immune response against experimental infection. *Infect Immun* 71: 2744-2757.
- Brandão A, Urmenyi T, Rondinelli E, Gonzalez A, De Miranda AB, Degraeve W 1997. Identification of transcribed sequences (ESTs) in the *Trypanosoma cruzi* genome project. *Mem Inst Oswaldo Cruz* 92: 863-866.
- Brener Z 1973. Biology of *Trypanosoma cruzi*. *Ann Rev Microbiol* 27: 347-382
- Chatterjee S, Singh S, Sohoni R, Kattige V, Deshpande C, Chiplunkar S, Kumar N, Sharma S 2000. Characterization of domains of the phosphoriboprotein P0 of *Plasmodium falciparum*. *Mol Biochem Parasitol* 107: 143-154.
- Coetzee T, Herschlag D, Belfort M 1994. *Escherichia coli* proteins, including ribosomal protein S12, facilitates in vitro splicing of phage T4 introns by acting as RNA chaperones. *Gene Dev* 8: 11575-11565.
- DaRocha WD, Bartholomeu DC, Macedo CD, Horta MF, Cunha-Neto E, Donelson JE, Teixeira SM 2002. Characterization of cDNA clones encoding ribonucleoprotein antigens expressed in *Trypanosoma cruzi* amastigotes. *Parasitol Res* 88: 292-300.
- Ewing B, Hillier L, Wendl MCM, Green P 1998. Base-calling of automated sequencer traces using phred. I. Accuracy assessment. *Genome Res* 8: 175-185.
- Frasch ACC 2001. Functional diversity in the Trans-sialidase and Mucin families in *Trypanosoma cruzi*. *Parasitol Today* 16: 282-285.
- Hoft DF, Kim KS, Otsu K, Moser DR, Yost WJ, Blumin JH, Donelson JE, Kirchhoff LV 1989. *Trypanosoma cruzi* expresses diverse repetitive protein antigens. *Infect Immun* 57: 1959-1964.
- Kaplan D, Ferrari I, Bergami PL, Mahler E, Levitus G, Chiale P, Hoebeke J, Van Regenmortel MHV, Levin MJ 1997. Antibodies to ribosomal P proteins of *Trypanosoma cruzi* in Chagas disease possess functional autoreactivity with heart tissue and differ from anti-P autoantibodies in lupus. *Proc Nat Acad Sci USA* 94: 10301-10306.
- Levin MJ, Mesri E, Benarous R, Levitus G, Schijman A, Levy-Yeyati P, Chiale PA, Ruiz AM, Kahn A, Rosenbaum MB 1989. Identification of major *Trypanosoma cruzi* antigenic determinants in chronic Chagas' heart disease. *Am J Trop Med Hyg* 41: 530-538.
- Low HP, Santos MAM, Wizel B, Tarleton RL 1998. Amastigote surface proteins of *Trypanosoma cruzi* are targets for CD8⁺ CTL. *J Immunol* 160: 1817-1823.
- Matsumoto TK, Hoshino-Schimizu S, Nakamura PM, Andrade HF, Umezawa ES 1993. High resolution of *Trypanosoma cruzi* amastigote antigen in serodiagnosis of different clinical forms of Chagas disease. *J Clin Microbiol* 31: 1486-1492.
- Minning TA, Bua J, Garcia GA, McGraw RA, Tarleton RL 2003. Microarray profiling of gene expression during trypomastigote to amastigote transition in *Trypanosoma cruzi*. *Mol Biochem Parasitol* 131: 55-64.
- Pais FS, DaRocha WD, Campos P, Cunha-Neto E, Teixeira SMR 2002. Human chagasic responses against *T. cruzi* recombinant antigens derived from ribonucleoprotein complexes. Proceedings of the XXIX Annual Meeting on Basic Research in Chagas Disease Caxambu, MG.
- Porcel BM, Tran AN, Tammi M, Nyarady Z, Rydaker M, Urmenyi TP, Rondinelli E, Pettersson U, Andersson B, Aslund L 2000. Gene survey of the pathogenic protozoan *Trypanosoma cruzi*. *Genome Res* 10: 1103-1107.
- Recinos RF, Kirchhoff LV, Donelson JE 2001. Cell cycle expression of histone genes in *Trypanosoma cruzi*. *Mol Biochem Parasitol* 113: 215-222.
- Requena JM, Alonso CA, Soto M 2000. Evolutionary conserved proteins as prominent immunogens during *Leishmania* infections. *Parasitol Today* 16: 246-250.
- Skeiky YA, Benson DR, Elwasila M, Badaro R, Burns JM JR, Reed SG 1994. Antigens shared by *Leishmania* species and *Trypanosoma cruzi*: immunological comparison of the acidic ribosomal P0 proteins. *Infect Immun* 62: 1643-1651.
- Teixeira SMR, Russell DG, Kirchhoff LV, Donelson JE 1994. A differentially expressed gene family encoding "amastin", a surface glycoprotein of *Trypanosoma cruzi* amastigotes. *J Biol Chem* 269: 20509-20516.
- Vasconcelos AT, The Brazilian National Genome Project Consortium 2003. The complete genome sequence of *Chromobacterium violaceum* reveals remarkable and exploitable bacterial adaptability. *Proc Nat Acad Sci USA* 100: 11660-11665.
- Vazquez MP, Schijman AG, Levin MJ 1994. A short interspersed repetitive element provides a new 3' acceptor site for trans-splicing in certain ribosomal P2 beta protein genes of *Trypanosoma cruzi*. *Mol Biochem Parasitol* 64: 327-336.
- Verdun RE, Di Paolo N, Urmenyi TP, Rondinelli E, Frasch AC, Sanchez DO 1998. Gene discovery through expressed sequence Tag sequencing in *Trypanosoma cruzi*. *Infect Immun* 66: 5393-5398.

Rapid Communication

Sequences involved in mRNA processing in *Trypanosoma cruzi*

Priscila C. Campos^{a,1}, Daniella C. Bartholomeu^{b,1}, Wanderson D. DaRocha^{a,2},
Gustavo C. Cerqueira^{a,3}, Santuza M.R. Teixeira^{a,*}

^a Departamento de Bioquímica e Imunologia, ICB, Universidade Federal de Minas Gerais, Belo Horizonte, MG, Brazil

^b Departamento de Parasitologia, Universidade Federal de Minas Gerais, 31270-901 Belo Horizonte, MG, Brazil

Received 25 April 2008; received in revised form 27 June 2008; accepted 8 July 2008

Abstract

Gene expression in Trypanosomatids requires processing of polycistronic transcripts to generate monocistronic mRNAs by cleavage events that are coupled to the addition of a Spliced Leader sequence (SL) at the 5'-end and a poly(A) tail at the 3'-end of each mRNA. Here we investigate the sequence requirements involved in *Trypanosoma cruzi* mRNA processing by mapping all available expressed sequence tags and cDNAs containing poly(A) tail and/or SL to genomic intergenic regions. Amongst other parameters, we determined that the median lengths of 5' untranslated region (UTR) and 3'UTR sequences are 35 and 264 nucleotides, respectively; and that the median distance between SL addition sites and a polypyrimidine motif is 18 nucleotides, whereas the median distance between poly(A) addition sites and the closest polypyrimidine-rich sequence is 40 nucleotides.

© 2008 Australian Society for Parasitology Inc. Published by Elsevier Ltd. All rights reserved.

Keywords: *Trypanosoma cruzi*; mRNA processing; trans-splicing; Polypyrimidine; Untranslated regions

Gene expression has remarkable differences in prokaryotes and eukaryotes. Distinct promoter sequences which are recognised by specific RNA polymerases, differences in the level of complexity of the transcription machinery and the requirement of RNA processing events are major aspects that differentiate prokaryotic and eukaryotic gene transcription (Struhl, 1999). Trypanosomatids belong to the Kinetoplastida order, a group of early branched eukaryotic microorganisms that includes human pathogens such as *Trypanosoma brucei* (the causative agent of sleeping sickness or African trypanosomiasis), *Trypanosoma cruzi*

(the causative agent of Chagas disease) and *Leishmania* spp. (the causative agent of different forms of leishmaniasis). In addition to their medical relevance, these parasites display a number of distinctive features regarding mechanisms controlling gene expression that are not common in other eukaryotes, such as polycistronic transcription, trans-splicing processing of pre-mRNA, extensive mitochondrial RNA editing and transcription of protein coding genes carried by RNA polymerase I (Donelson et al., 1999). Most notable is trans-splicing, an mRNA processing mechanism that requires two cleavage events to occur before and after each coding region present in the polycistronic precursor in order to generate mature, monocistronic mRNAs. Cleavages within the pre-mRNA are coupled to two RNA-processing reactions: trans-splicing of a small capped RNA of 39–41 nucleotides (nt), the spliced leader RNA (SL-RNA), added to the 5'-terminus of all known protein-encoding RNAs, and 3'-end polyadenylation (LeBowitz et al., 1993; Matthews et al., 1994; Liang et al., 2003). Both events are dependent on polypyrimidine-rich motifs (poly(Y)) located within the intergenic regions.

* Corresponding author. Address: Departamento de Bioquímica e Imunologia, ICB, Universidade Federal de Minas Gerais Av. Antônio Carlos 6627, 31270-010, Belo Horizonte, MG, Brasil. Tel.: +55 31 3499 2665; fax: +55 31 3499 2614.

E-mail address: santuzat@icb.ufmg.br (S.M.R. Teixeira).

¹ Both authors have contributed equally to this work.

² Present address: Departamento de Bioquímica e Biologia Molecular, Universidade Federal do Paraná, Curitiba, PR, Brazil.

³ Present address: Department of Cell Biology and Molecular Genetics, University of Maryland, MD, USA.

Whilst no canonical poly(A) addition signal has been identified, only AG dinucleotides situated downstream from a poly(Y) motif are used as SL acceptor sites (LeBowitz et al., 1993; Matthews et al., 1994; Schürch et al., 1994; Benz et al., 2005; Siegel et al., 2005). Because trypanosomatid protein-coding genes are often organised as large directional gene clusters or polycistronic transcription units (Myler et al., 1999), the mechanisms controlling gene expression in these organisms depend almost exclusively on regulatory pathways acting at post-transcriptional levels. Indeed, numerous studies including the complete sequences of the *T. brucei*, *T. cruzi* and *Leishmania major* genomes, known as the Tri-Tryps, reveal not only a total lack of evidence for differential regulation of RNA polymerase II transcription of individual protein-coding genes, but also no identifiable RNA polymerase II promoter consensus sequences (Berriman et al., 2005; El-Sayed et al., 2005a,b; Ivens et al., 2005). This lack of transcription initiation control implies that knowledge of post-transcriptional processes, such as mRNA processing and stabilization, is crucial for the understanding of gene expression in these organisms (for a recent review, see Haile and Papadopoulou, 2007).

Although the Tri-Tryp genome sequencing project (Berriman et al., 2005; El-Sayed et al., 2005a; Ivens et al., 2005) provided all the relevant information required for studies of elements involved in gene expression, a genome-wide analysis of RNA processing signals was only performed for *T. brucei* and *Leishmania*. A global genome screening for RNA processing elements reported by Benz et al. (2005) revealed that, in *T. brucei*, *trans*-splicing reactions generally occur at the first AG dinucleotide after a eight to 25 nt poly(Y), generating mRNAs with 5' untranslated regions (UTRs) with a median length of 68 nt. These authors also showed that polyadenylation occurs at a position with one or more A residues located between 80 and 140 nt from a downstream poly(Y) motif, resulting in 3'UTRs as short as 21 nt or up to 5040 nt long, with a median 3'UTR length of 348 nt. These results were then used to train an algorithm that enabled prediction of *trans*-splicing and polyadenylation sites for most protein coding genes in *T. brucei*.

Although comparative analyses revealed that the three trypanosomatid genomes share 6158 orthologue clusters of protein-encoding genes and have large syntenic blocks, they differ in several aspects not only regarding their life cycle, but also in relation to molecular mechanisms controlling gene expression, such as switching of variant surface glycoprotein (VSG) genes and mRNA transcription by RNA polymerase I (El-Sayed et al., 2005b). Therefore, we conducted a similar whole genome screening process to identify the sequence requirements involved in mRNA processing in *T. cruzi*. Using DNA sequences derived from the *T. cruzi* genome (El-Sayed et al., 2005a) and all available expressed sequence tags (ESTs) and cDNAs containing poly(A) tail and/or SL, we uncovered a pattern for mRNA processing in *T. cruzi* and compared the elements involved in RNA processing events in *T. cruzi* and

T. brucei. To this aim, we obtained all available EST and cDNA sequences from TcruziDB (<http://www.tcruzi.org>) and NCBI (<http://www.ncbi.nlm.nih.gov>) and selected those containing at their ends the last 10 nt from the SL (TACTATATTG) or eight consecutive As or Ts. The reverse complement of sequences containing poly(T) was used. Based on these criteria, a total of 1910 and 1449 ESTs/cDNAs containing the SL and poly(A) tail, respectively, were identified. After excising the SL sequence and/or the poly-A tail from the ESTs/cDNAs, we mapped the two groups of sequences onto a *T. cruzi* inter-coding sequence (inter-CDS) database using the BLASTN algorithm. The *T. cruzi* inter-CDS database was constructed using the parasite genome sequence deposited in GenBank (El-Sayed et al., 2005a). To ensure that UTRs smaller than the BLASTN seed length of 11 nt would be identified, we included the first and last 10 nucleotides of each coding region in the inter-CDS dataset. Due to the low quality of the ESTs inherent in such single pass sequences, we used stringent criteria to map the *T. cruzi* ESTs to the intergenic regions, which require matches with at least 90% identity. The results were further filtered as follows. For SL addition site mapping, we only considered those matches that satisfied the following criteria: (i) the first position of the cDNA/EST that matched the intergenic region must immediately follow the SL sequence; (ii) the last position of the inter-CDS region must match the EST/cDNA or the entire length of the cDNA/EST sequence must match the inter-CDS. For mapping the poly-A addition site, we only considered sequences: (i) containing at least 100 nt; (ii) in which the last position of the cDNA/EST that matched the intergenic region immediately preceded the poly(A) tail; and (iii) the first position of the inter-CDS region matched the cDNA/EST or the entire length of the cDNA/EST sequence matched the inter-CDS.

We were able to identify 1189 and 149 regions in the *T. cruzi* genome containing genes that matched the SL plus and poly-A plus ESTs/cDNAs, respectively, and therefore identified the corresponding 5' and 3'UTRs. The 5'UTRs presented a median length of 35 nt (mean = 35), whereas the 3'UTRs has a median length of 137 nt (mean = 264). Almost 90% of the 5'UTRs have less than 50 nt and 91% of the 3'UTRs analysed had less than 500 nt. As shown in Fig. 1A and B, the size distribution of 3'UTRs shows a broader range than 5'UTRs. The larger average size of the 3'UTRs is in agreement with several studies showing the presence of regulatory sequences downstream to coding sequences (Weston et al., 1999; Coughlin et al., 2000; Di Noia et al., 2000; Silva et al., 2006). Most of the regulatory elements analysed to date are responsible for controlling mRNA stability, which is a major mechanism involved in differential gene expression in *T. cruzi* (Haile and Papadopoulou, 2007). It is noteworthy that long UTRs may be less numerous or less easy to clone, therefore, these may be under-represented in the EST or cDNA databases. Consequently there could be a bias towards mRNAs with shorter lengths that we were able to map onto the genome.

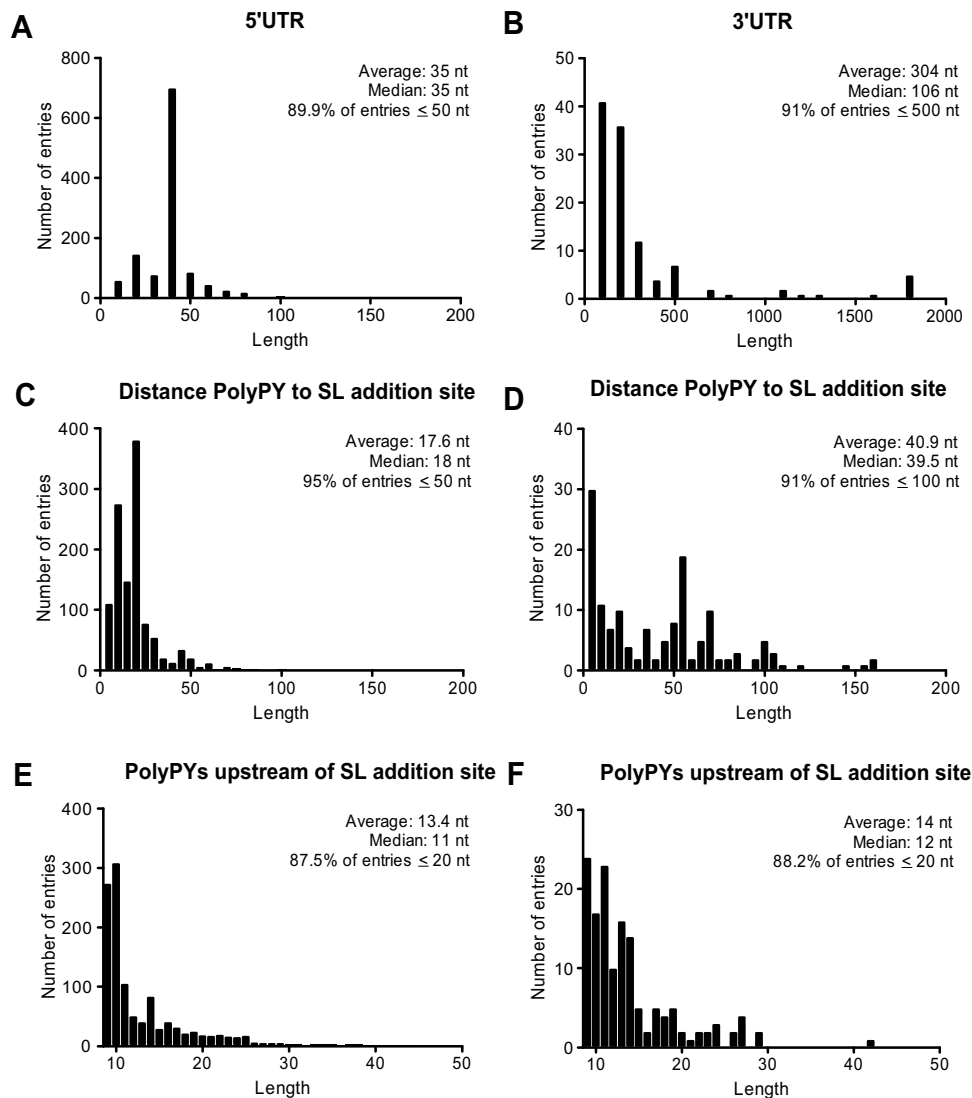


Fig. 1. Size range of 5' and 3' untranslated regions (UTRs) and distances of processing signals in *Trypanosoma cruzi* mRNAs. Histograms (A) and (B) show the length distribution of 1189 and 149 mapped 5' and 3' UTRs, respectively; (C) and (D) show the distances between the closest polypyrimidine to the downstream Spliced Leader (SL) acceptor site (C) and to the upstream poly(A) addition site (D); (E) and (F) show the lengths of polypyrimidine motifs located upstream from the SL addition site (E) and downstream from the poly(A) addition site (F).

We next searched for poly(Y) tracts upstream and downstream of the spliced-leader and polyadenylation acceptor sites, respectively, using the Fuzznuc algorithm (EMBOSS package). As previously reported, a nine to 10 nt poly(Y) motif is an essential element controlling both *trans*-splicing and polyadenylation in *T. brucei* (Matthews et al., 1994; Siegel et al., 2005). In our analysis, we searched for poly(Y) motifs containing at least nine residues allowing one internal purine. The poly(Y) tract most likely used in mRNA processing was defined as the one closest to the SL or poly(A) acceptor sites. Using this dataset, we analysed the distances between the poly(Y) and the SL or poly(A) acceptor sites as well as the poly(Y) length and composition. As shown in Fig. 1C and D, the median distances between SL addition sites and the first poly(Y) motif and between polyadenylation sites and a poly(Y) were 18 nt (mean = 17.6) and 39.5 nt (mean = 40.9), respec-

tively. Approximately 95% of SL addition site/poly(Y) distances analysed were less than 50 nt whereas approximately 91% of the distances between poly(A) addition sites and poly(Y) were smaller than 100 nt. Thus, in agreement with the smaller average sizes of 5'UTR, the distance between the closest poly(Y) motif and SL acceptor site is significantly smaller than the distance between poly(Y) and poly(A) addition sites. The median lengths of the closest poly(Y) tract found upstream from the SL addition site and downstream from the polyadenylation sites were 11 and 12 nt, respectively. Fig. 2 shows a schematic representation of the mRNA processing signals present in the *T. cruzi* genome, according to our analyses.

A comparative analysis between our results with those obtained for *T. brucei* by Benz et al. (2005) revealed striking differences that are in accordance with previous findings related to average gene size and coding sequence

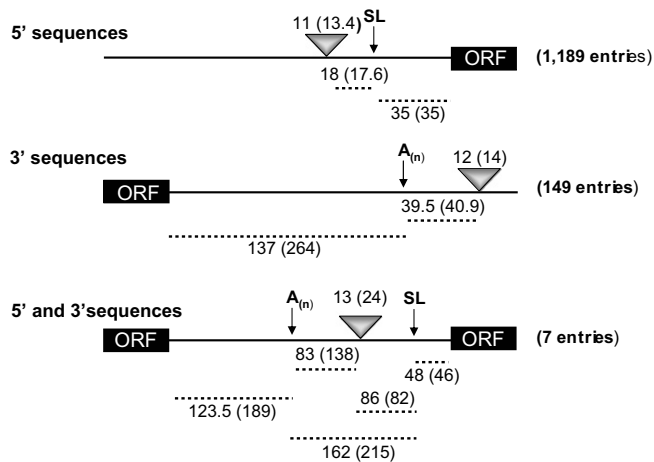


Fig. 2. Patterns of pre-mRNA processing in *Trypanosoma cruzi*. Genomic sequences containing Spliced Leader (SL) addition sites (5' sequences) and polyadenylation sites (3' sequences) as well as a few entries containing both *trans*-splicing and polyadenylation sites derived from the same cDNA or expressed sequence tag (5' and 3' sequences) were mapped in the *T. cruzi* genome. Numbers under hatched lines indicate the median lengths of 5' untranslated region (UTR), 3' UTR and intergenic regions. The median distances from the polypyrimidine tract to the SL addition site or to the polyadenylation site are indicated above the hatched lines. Gray triangles represent poly(Y) tracts with numbers on the top indicating their median sizes. In each case, the mean values are shown within parentheses. Black boxes denote open reading frames (ORFs); SL with an arrow, spliced-leader addition site; (A)_n with an arrow, poly(A) addition site.

Table 1
Comparison between *Trypanosoma cruzi* and *Trypanosoma brucei* mRNA processing patterns

Parameters	<i>T. cruzi</i>	<i>T. brucei</i>	
CDS size ^a	1457	1511	
Inter-CDS size ^a	561	721	
5'UTR length ^a	35	68	
3'UTR length ^a	264	587	
Distance poly(PY) to SL addition site ^a	18	24	
Distance poly(PY) to poly(A) addition site ^a	41 ^b	120	
Poly(PY) length upstream of SL addition site ^a	13	13	
Poly(PY) length downstream of poly(A) addition site ^a	14	13	
Nucleotide content of Poly(PY)s upstream of the SL addition site	T (%)	59.3	59
	C (%)	33.3	38
	AG (%)	7.4	3
Nucleotide content of Poly(PY)s downstream of the poly(A) tract	T (%)	66.3	64
	C (%)	27.0	34
	AG (%)	6.6	1

Nucleotide numbers of various regions in both genomes and the percentages of each nucleotide present in poly(Y) motifs located upstream from Spliced Leader (SL) addition sites and downstream from poly(A) addition sites were obtained from data derived from our analyses and from the work published by Benz et al. (2005) and El-Sayed et al. (2005a). CDS, coding sequence; UTR, untranslated region; Poly(PY), polypyrimidine motifs; poly(A), polyadenylation; SL, Spliced Leader.

^a Mean values in nucleotides.

^b Distance between the poly(A) addition site and the closest poly(Y) motif.

density (El-Sayed et al., 2005b). As shown in Table 1, the reported results indicated that the average size of 5' and 3'UTR is significantly smaller in *T. cruzi* than in *T. brucei*,

as well as the mean distances between poly(Y) motifs and SL and poly(A) addition sites. This is in agreement with the comparative analysis of the Tri-Tryp genomes, which indicates that the *T. cruzi* genome is the most compact, with a smaller average size of coding regions and inter-coding regions compared with *T. brucei* and *L. major* (Berri-man et al., 2005; El-Sayed et al., 2005a; Ivens et al., 2005).

We also determined the poly(Y) motif distribution in the regions flanking the SL and polyadenylation addition sites (Fig. 3). A non-random distribution of these motifs in the vicinity of both processing sites was observed. As shown in Fig. 3A, the polypyrimidine tracts are preferentially located within 15–20 nt upstream from the SL addition site with a marked decrease in their occurrence within 5'UTRs and downstream coding sequences. Since the average length of the 5'UTR regions is quite short (<50 nt), it

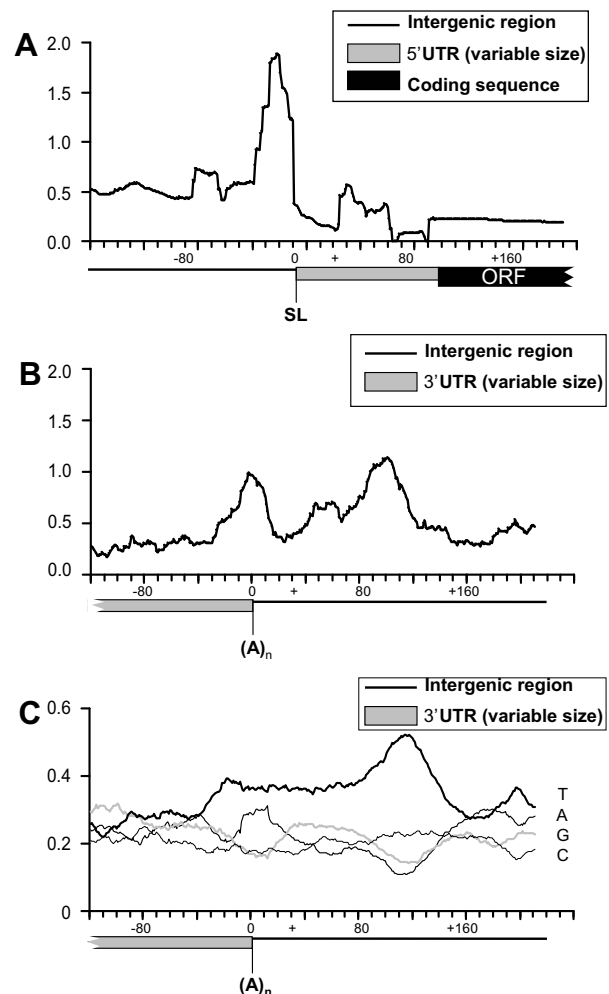


Fig. 3. Polypyrimidine and nucleotide frequencies in the vicinity of the *Trypanosoma cruzi* mapped processing sites. Poly(Y) motifs flanking the spliced-leader addition site (A) and polyadenylation site (B). The total number of poly(Y) motifs identified over a 20 nucleotide sliding window, with increment of one nucleotide was divided by the total number of sequences that cover each given window. For the analysis of the distribution of poly(Y) motifs within coding regions (A) all the *T. cruzi* genes excluding pseudogenes and truncated genes were analysed. (C) Nucleotide composition flanking genomic poly(A) addition sites.

would be tempting to speculate that the suppression of poly(Y) in the 5'UTR and downstream coding sequence may avoid the erroneous use of a branch point for *trans*-splicing inside coding regions. The distribution of the poly(Y) motifs in the vicinity of the poly(A) addition sites shows two pyrimidine rich-regions, with a first peak almost coincident with the polyadenylation site and a second peak approximately 100 nt downstream the 3' end of the transcript. Interestingly, the nt composition of these two poly(Y) rich regions is slightly different (Fig. 3C). Although both regions are composed mainly of thymines, the first peak is followed by a region with an increase in the content of adenines, which is the nt most frequently used for the addition of the poly(A) tail (data not shown). In contrast, the second poly(Y) rich region is predominantly composed of thymines with guanines and adenines significantly suppressed in that region. It is possible that, in *T. cruzi*, the poly(Y) rich region that is closest to the poly(A) may not correspond to the poly(Y) motif that lies upstream from the SL addition site. Therefore, the splicing machinery may recognise two distinct poly(Y) tracts, one for *trans*-splicing reactions and another for polyadenylation.

The exact sequence requirements of poly(Y) tracts and, most importantly, whether their size and composition could influence *trans*-splicing and polyadenylation reactions in *T. cruzi*, need to be tested experimentally. Using a dual reporter transfection system, Siegel et al. (2005) conducted a systematic study of sequence motifs for RNA *trans*-splicing in *T. brucei* and demonstrated a clear correlation between luciferase expression and poly(Y) tract length with an optimum poly(Y) tract length of 25 nt containing uridines interspersed with cytosines. These authors also showed that the position of the poly(Y) tract relative to the SL acceptor site (AG dinucleotide) is important. Consistent with the scanning model for 3' splice acceptor site selection, analyses of the highly efficient pre-mRNA processing region HX1 of *T. cruzi* also showed that the first AG dinucleotide downstream from the branch point/poly(Y) tract was always selected for *trans*-splicing and that adenines used as putative branch points are always very close to the poly(Y) tract (Ben-Dov et al., 2005). Mutational analyses within the intergenic region separating the calmodulin–ubiquitin associated gene and the ubiquitin fusion gene of *T. cruzi* also support the scanning model for 3' splice site selection (Hummel et al., 2000). Aiming to improve vectors for expression of foreign genes in *T. cruzi*, we compared five different intergenic regions placed upstream from the luciferase reporter gene in transiently transfected epimastigotes (DaRocha et al., 2004). In this study, intergenic regions from four constitutive genes highly expressed in epimastigotes and from amastin, an amastigote-specific gene, were analysed. With the exception of the amastin sequences, we observed only small differences amongst the other sequences, indicating that the differences in the poly(Y) tract length, their composition or localisation regarding the SL acceptor site do not affect the expression of these genes in *T. cruzi* epimastigotes

(DaRocha et al., 2004). It is worth noting that all four constitutive genes analysed possess long poly(Y) tracts upstream the spliced-leader acceptor site that may fulfil the minimal requirement for an efficient processing event.

Recently, Smith et al. (2008) reported the development of an algorithm to predict *trans*-splicing and poly(A) addition signals in *Leishmania* mRNAs. Besides the metrics regarding the length of UTRs defined by mapping ESTs on the genome, the algorithm also incorporates pyrimidine content of query sequences into a scoring function. Further, to improve the prediction of the polyadenylation sites, the method compares query sequences with polyadenylated ESTs using position-specific scanning matrices (PSSMs). Although differences in the sequence features surrounding the *T. cruzi* and *Leishmania* polyadenylation sites can be observed, a few common patterns emerge from these studies. Firstly, the presence of a polypyrimidine region enriched with thymines located at the vicinity of the polyadenylation site. In *Leishmania infantum*, thymines are almost twice as numerous around 50 bases upstream of the poly(A) site, whilst in *T. cruzi*, as mentioned above, this motif is almost coincident to the poly-A addition site. Smith et al. (2008) proposed that in *Leishmania* this poly(Y)-rich region may play a role in the 3'UTR cleavage site, since its exclusion from scanning matrices reduces sensitivity of the method at close range. The presence of a poly(Y) peak in a similar region in *T. cruzi* may reinforce this hypothesis. Second, these authors detected an increase in adenine content within close range of the mapped poly(A) addition site. In trypanosomatids, no canonical AAUAAA is found within their 3'UTRs, but poly(A) addition usually occurs at an adenosine residue (Smith et al., 2008; Benz et al., 2005, and this work, 85% of the *T. cruzi* mapped poly(A) addition site). Third, similar to *T. cruzi* and *T. brucei* a second poly(Y) motif downstream of the polyadenylation site, which is believed to play a central role in, driving the polyadenylation machinery, was also identified. Whilst in the *T. cruzi* genome a polypyrimidine-rich region is located 40 nt from the poly(A) site, in *T. brucei* it is at 120 nt and in *Leishmania* it is found at 300–600 nt. Again, the shorter distance between the poly(A) addition site and the poly(Y) tract in *T. cruzi* is in agreement with the greater compaction of its genome compared with the other two Tri-trypan genomes (El-Sayed et al., 2005b).

Algorithms developed to predict processing sites in Trypanosomatids, based on the *T. brucei* and *L. major* pattern of mRNA processing, have been created (Benz et al., 2005; Smith et al., 2008). Sequence features surrounding the *T. cruzi* processing sites identified in this study could be used to fine tune these algorithms so that they could also be used to predict the *T. cruzi* mRNA processing products. Experimental identification of a large number of UTRs in *T. cruzi* mRNAs, for which no EST/cDNA is currently available, is a time-consuming approach. Therefore, in silico prediction of *T. cruzi* UTRs from genomic sequences will allow a rapid identification of a large number of putative *cis*-regulatory elements whose role in the modulation of gene

expression could be experimentally validated. In addition, the information provided in this work may also contribute to studies of the optimisation of vectors currently being used to manipulate gene expression in this parasite.

Acknowledgements

This work was supported by grants from CNPq and FAPEMIG, Brazil, by UNICEF/UNDP/World Bank WHO for Research and Training in Tropical Diseases and by an International Research Scholars Grant from the Howard Hughes Medical Institute.

References

- Benz, C., Nilsson, D., Andersson, B., Clayton, C., Guilbride, D.L., 2005. Messenger RNA processing sites in *Trypanosoma brucei*. *Molecular and Biochemical Parasitology* 143, 125–134.
- Ben-Dov, C.P., Levin, M.J., Vázquez, M.P., 2005. Analysis of the highly efficient pre-mRNA processing region HX1 of *Trypanosoma cruzi*. *Molecular and Biochemical Parasitology* 140 (1), 97–105.
- Berriman, M., Ghedin, E., Hertz-Fowler, C., Blandin, G., Renaud, H., Bartholomeu, D.C., Lennard, N.J., Caler, E., Hamlin, N.E., Haas, B., Bohme, U., Hannick, L., Aslett, M.A., Shallom, J., Marcello, L., Hou, L., Wickstead, B., Alsmark, U.C., Arrowsmith, C., Atkin, R.J., Barron, A.J., Bringaud, F., Brooks, K., Carrington, M., Cherevach, I., Chillingworth, T.J., Churcher, C., Clark, L.N., Corton, C.H., Cronin, A., Davies, R.M., Doggett, J., Djikeng, A., Feldblyum, T., Field, M.C., Fraser, A., Goodhead, I., Hance, Z., Harper, D., Harris, B.R., Hauser, H., Hostetler, J., Ivens, A., Jagels, K., Johnson, D., Johnson, J., Jones, K., Kerhornou, A.X., Koo, H., Larke, N., Landfear, S., Larkin, C., Leech, V., Line, A., Lord, A., Macleod, A., Mooney, P.J., Moule, S., Martin, D.M., Morgan, G.W., Mungall, K., Norbertczak, H., Ormond, D., Pai, G., Peacock, C.S., Peterson, J., Quail, M.A., Rabbinowitsch, E., Rajandream, M.A., Reitter, C., Salzberg, S.L., Sanders, M., Schobel, S., Sharp, S., Simmonds, M., Simpson, A.J., Tallon, L., Turner, C.M., Tait, A., Tivey, A.R., Van Aken, S., Walker, D., Wanless, D., Wang, S., White, B., White, O., Whitehead, S., Woodward, J., Wortman, J., Adams, M.D., Embley, T.M., Gull, K., Ullu, E., Barry, J.D., Fairlamb, A.H., Opperdoes, F., Barrell, B.G., Donelson, J.E., Hall, N., Fraser, C.M., Melville, S.E., El-Sayed, N.M., 2005. The genome of the African trypanosome *Trypanosoma brucei*. *Science* 309, 416–422.
- Coughlin, B.C., Teixeira, S.M.R., Kirchoff, L.V., Donelson, J.E., 2000. Amastin mRNA abundance in *Trypanosoma cruzi* is controlled by a 3'-untranslated region position-dependent *cis*-element and an untranslated region-binding protein. *The Journal of Biological Chemistry* 275, 12051–12060.
- DaRocha, W.D., Silva, R.A., Bartholomeu, D.C., Pires, S.F., Freitas, J.M., Macedo, A., Vazquez, M.P., Levin, M.J., Teixeira, S.M.R., 2004. Expression of exogenous genes in *Trypanosoma cruzi*: improving vectors and electroporation protocols. *Parasitology Research* 92, 113–120.
- Di Noia, J.M., D'Orso, I., Sánchez, D.O., Frasch, A.C., 2000. AU-rich elements in the 3'-untranslated region of a new mucin-type gene family of *Trypanosoma cruzi* confers mRNA instability and modulates translation efficiency. *The Journal of Biological Chemistry* 275 (14), 10218–10227.
- Donelson, J.E., Gardner, M.J., El-Sayed, N.M., 1999. More surprises from Kinetoplastida. *Proceeding of the National Academy of Sciences USA* 96 (6), 2579–2581.
- El-Sayed, N.M., Myler, P.J., Bartholomeu, D.C., Nilsson, D., Aggarwal, G., Tran, A.N., Ghedin, E., Worthey, E.A., Delcher, A.L., Blandin, G., Westenberger, S.J., Caler, E., Cerqueira, G.C., Branche, C., Haas, B., Anupama, A., Arner, E., Aslund, L., Attipoe, P., Bontempi, E., Bringaud, F., Burton, P., Cadag, E., Campbell, D.A., Carrington, M., Crabtree, J., Darban, H., da Silveira, J.F., de Jong, P., Edwards, K., Englund, P.T., Fazelina, G., Feldblyum, T., Ferella, M., Frasch, A.C., Gull, K., Horn, D., Hou, L., Huang, Y., Kindlund, E., Klingbeil, M., Kluge, S., Koo, H., Lacerda, D., Levin, M.J., Lorenzi, H., Louie, T., Machado, C.R., McCulloch, R., McKenna, A., Mizuno, Y., Mottram, J.C., Nelson, S., Ochaya, S., Osoegawa, K., Pai, G., Parsons, M., Pentony, M., Pettersson, U., Pop, M., Ramirez, J.L., Rinta, J., Robertson, L., Salzberg, S.L., Sanchez, D.O., Seyler, A., Sharma, R., Shetty, J., Simpson, A.J., Sisk, E., Tammi, M.T., Tarleton, R., Teixeira, S., Van Aken, S., Vogt, C., Ward, P.N., Wickstead, B., Wortman, J., White, O., Fraser, C.M., Stuart, K.D., Andersson, B., 2005a. The genome sequence of *Trypanosoma cruzi*, etiologic agent of Chagas' disease. *Science* 309, 409–415.
- El-Sayed, N.M., Myler, P.J., Blandin, G., Berriman, M., Crabtree, J., Aggarwal, G., Caler, E., Renaud, H., Worthey, E.A., Hertz-Fowler, C., Ghedin, E., Peacock, C., Bartholomeu, D.C., Haas, B.J., Tran, A.N., Wortman, J.R., Alsmark, U.C., Angiuoli, S., Anupama, A., Badger, J., Bringaud, F., Cadag, E., Carlton, J.M., Cerqueira, G.C., Creasy, T., Delcher, A.L., Djikeng, A., Embley, T.M., Hauser, C., Ivens, A.C., Kummerfeld, S.K., Pereira-Leal, J.B., Nilsson, D., Peterson, J., Salzberg, S.L., Shallom, J., Silva, J.C., Sundaram, J., Westenberger, S., White, O., Melville, S.E., Donelson, J.E., Andersson, B., Stuart, K.D., Hall, N., 2005b. Comparative genomics of *Trypanosomatid* parasitic protozoa. *Science* 309, 404–409.
- Haile, S., Papadopoulos, B., 2007. Developmental regulation of gene expression in trypanosomatid parasitic protozoa. *Current Opinion in Microbiology* 10 (6), 569–577.
- Hummel, H.S., Gillespie, R.D., Swindle, J., 2000. Mutational analysis of 3' splice site selection during *trans*-splicing. *The Journal of Biological Chemistry* 275 (45), 35522–35531.
- Ivens, A.C., Peacock, C.S., Worthey, E.A., Murphy, L., Aggarwal, G., Berriman, M., Sisk, E., Rajandream, M.A., Adlem, E., Aert, R., Anupama, A., Apostolou, Z., Attipoe, P., Bason, N., Bauser, C., Beck, A., Beverley, S.M., Bianchetti, G., Borzym, K., Bothe, G., Bruschi, C.V., Collins, M., Cadag, E., Ciarloni, L., Clayton, C., Coulson, R.M., Cronin, A., Cruz, A.K., Davies, R.M., De Gaudenzi, J., Dobson, D.E., Duesterhoeft, A., Fazelina, G., Fosker, N., Frasch, A.C., Fraser, A., Fuchs, M., Gabel, C., Goble, A., Goffeau, A., Harris, D., Hertz-Fowler, C., Hilbert, H., Horn, D., Huang, Y., Klages, S., Knights, A., Kube, M., Larke, N., Litvin, L., Lord, A., Louie, T., Marra, M., Masuy, D., Matthews, K., Michaeli, S., Mottram, J.C., Muller-Auer, S., Munden, H., Nelson, S., Norbertczak, H., Oliver, K., O'Neil, S., Pentony, M., Pohl, T.M., Price, C., Purnelle, B., Quail, M.A., Rabbinowitsch, E., Reinhardt, R., Rieger, M., Rinta, J., Robben, J., Robertson, L., Ruiz, J.C., Rutter, S., Saunders, D., Schafer, M., Schein, J., Schwartz, D.C., Seeger, K., Seyler, A., Sharp, S., Shin, H., Sivam, D., Squares, R., Squares, S., Tosato, V., Vogt, C., Volckaert, G., Wambutt, R., Warren, T., Wedler, H., Woodward, J., Zhou, S., Zimmermann, W., Smith, D.F., Blackwell, J.M., Stuart, K.D., Barrell, B., Myler, P.J., 2005. The genome of the kinetoplastid parasite, *Leishmania major*. *Science* 309, 436–442.
- LeBowitz, J.H., Smith, H.Q., Rusche, L., Beverley, S.M., 1993. Coupling of poly(A) site selection and *trans*-splicing in *Leishmania*. *Genes and Development* 7, 996–1007.
- Liang, X.H., Haritan, A., Uliel, S., Michaeli, S., 2003. *trans* and *cis* splicing in trypanosomatids: mechanism, factors, and regulation. *Eukaryotic Cell* 2 (5), 830–840.
- 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 and Development* 8 (4), 491–501.
- Myler, P.J., Audleman, L., deVos, T., Hixson, G., Kiser, P., Lemley, C., Magness, C., Rickel, E., Sisk, E., Sunkin, S., Swartzell, S., Westlake, T., Bastien, P., Fu, G., Ivens, A., Stuart, K., 1999. *Leishmania major* Friedlin chromosome 1 has an unusual distribution of protein-coding genes. *Proceedings of the National Academy of Sciences USA* 96 (6), 2902–2906.

- Schürch, N., Hehl, A., Vassella, E., Braun, R., Roditi, I., 1994. Accurate polyadenylation of procyclin mRNAs in *Trypanosoma brucei* is determined by pyrimidine-rich elements in the intergenic regions. *Molecular Cell Biology* 14 (6), 3668–3675.
- Siegel, T.N., Tan, K.S.W., Cross, G.A., 2005. Systematic study of sequence motifs for RNA *trans*-splicing in *Trypanosoma brucei*. *Molecular and Cellular Biology* 25 (21), 9586–9594.
- Silva, R.A., Bartholomeu, D.C., Teixeira, S.M., 2006. Control mechanisms of tubulin gene expression in *Trypanosoma cruzi*. *International Journal for Parasitology* 36 (1), 87–96.
- Smith, M.A., Blanchette, M., Papadopoulou, B., 2008. Improving the prediction of mRNA extremities in the parasitic protozoan *Leishmania*. *BMC Bioinformatics* 9 (1), 158, [Epub ahead of print].
- Struhl, K., 1999. Fundamentally different logic of gene regulation in eukaryotes and prokaryotes. *Cell* 98 (1), 1–4.
- 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. *Molecular and Biochemical Parasitology* 102 (1), 53–66.



Trypanosoma cruzi MSH2: Functional analyses on different parasite strains provide evidences for a role on the oxidative stress response[☆]

Priscila C. Campos^a, Viviane G. Silva^a, Carolina Furtado^a, Alice Machado-Silva^a, Wanderson D. DaRocha^a, Eduardo F. Peloso^b, Fernanda R. Gadelha^b, Marisa H.G. Medeiros^c, Gustavo de Carvalho Lana^d, Ying Chen^e, Rebecca L. Barnes^e, Danielle Gomes Passos-Silva^a, Richard McCulloch^e, Carlos Renato Machado^a, Santuza M.R. Teixeira^{a,*}

^a Departamento de Bioquímica e Imunologia, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Belo Horizonte, MG, Brazil

^b Departamento de Bioquímica, Instituto de Biologia, Universidade Estadual de Campinas, Campinas, SP, Brazil

^c Departamento de Bioquímica, Instituto de Química, Universidade de São Paulo, SP, MG, Brazil

^d Departamento de Estatística, Instituto de Ciências Exatas, Universidade Federal de Minas Gerais, Brazil

^e Wellcome Centre for Molecular Parasitology, University of Glasgow, Glasgow, UK

ARTICLE INFO

Article history:

Received 23 April 2010

Received in revised form 24 October 2010

Accepted 1 November 2010

Available online 10 November 2010

Keywords:

Trypanosoma cruzi

DNA repair

MSH2

Oxidative stress

Trypanosoma brucei

ABSTRACT

Components of the DNA mismatch repair (MMR) pathway are major players in processes known to generate genetic diversity, such as mutagenesis and DNA recombination. *Trypanosoma cruzi*, the protozoan parasite that causes Chagas disease has a highly heterogeneous population, composed of a pool of strains with distinct characteristics. Studies with a number of molecular markers identified up to six groups in the *T. cruzi* population, which showed distinct levels of genetic variability. To investigate the molecular basis for such differences, we analyzed the *T. cruzi* MSH2 gene, which encodes a key component of MMR, and showed the existence of distinct isoforms of this protein. Here we compared cell survival rates after exposure to genotoxic agents and levels of oxidative stress-induced DNA in different parasite strains. Analyses of *msh2* mutants in both *T. cruzi* and *T. brucei* were also used to investigate the role of *Tcmsh2* in the response to various DNA damaging agents. The results suggest that the distinct MSH2 isoforms have differences in their activity. More importantly, they also indicate that, in addition to its role in MMR, TcMSH2 acts in the parasite response to oxidative stress through a novel mitochondrial function that may be conserved in *T. brucei*.

© 2010 Elsevier B.V. All rights reserved.

1. Introduction

Trypanosoma cruzi is a protozoan parasite of great medical importance, since it causes Chagas' disease (or American Trypanosomiasis), a malady endemic throughout Latin America, with an estimated 50,000 deaths/year and 100 million people at risk (<http://who.int/tdr/diseases/Chagas>). *T. cruzi* belongs to the order Kinetoplastida, which is characterized by the presence of one flagellum and a single mitochondrion in which is situated the kinetoplast, a specialized DNA-containing organelle. Though separated by many millions of years of evolution, *T. cruzi* is related to another kinetoplastid parasite, *T. brucei* [1], which continues to afflict the economy and health of the people of sub-Saharan Africa.

T. cruzi has a heterogeneous population composed by a pool of strains that circulate in the domestic and sylvatic cycles involving humans, insect vectors and animal reservoirs. Based on studies with a number of molecular markers, the taxon was divided initially in two well-defined groups, so-called *T. cruzi* I and *T. cruzi* II lineages [2]. More recent studies proposed the existence of six discrete typing units named *T. cruzi* I to VI [3]. Although the *T. cruzi* population is predominantly clonal, a few hybrid lineages have been identified, including the clone CL Brener, chosen as the reference strain for the genome project [4].

Several reports highlighted the differential genetic diversity between *T. cruzi* I and II lineages and, more importantly, the intra-group variability observed within *T. cruzi* II. Phylogeny studies of DHFR-TS and TR sequences from 31 strains showed that all sequences from *T. cruzi* I strains are monophyletic, whereas sequences from *T. cruzi* II strains are paraphyletic and fall into 3 clades [5]. Analysis of *T. cruzi* multi-gene families showed that paralogous sequences encoding amastin, an amastigote surface glycoprotein, and TcAG48, an antigenic RNA binding protein, present higher intragenomic variability in *T. cruzi* II and CL Brener strains,

[☆] Sequence data have been submitted to GenBank database under the accession number GQ869654.

* Corresponding author at: Departamento de Bioquímica e Imunologia, ICB, Universidade Federal de Minas Gerais, Av. Antônio Carlos, 6627, Pampulha, Belo Horizonte, MG 31270-901, Brazil. Tel.: +55 31 3409 2665; fax: +55 31 3409 2614.

E-mail address: santuzat@icb.ufmg.br (S.M.R. Teixeira).

when compared with *T. cruzi* I strains [6]. Likewise, whereas sequences encoding the Tc52 immuno-regulatory protein are homogeneous in strains belonging to the *T. cruzi* I lineage, they show higher sequence polymorphisms in *T. cruzi* II strains [7].

Components of DNA repair pathways are key players in the processes that result in changes in genetic variability within any cell population. Germline mutations in DNA mismatch repair (MMR) genes are associated with susceptibility to hereditary nonpolyposis colorectal human cancer (HNPCC) [8]. Mutator populations of bacteria, yeast and mammalian cells with defects in mismatch repair have also been described in isolates showing increased frequency of drug resistance [9–11]. Post-replicative DNA mismatch repair promotes genetic stability by repairing DNA replication errors, inhibiting recombination between non-identical DNA sequences and participating in responses to DNA damage induced by genotoxic agents, such as *N*-methyl-*N'*-nitro-*N*-nitrosoguanidine (MNNG) and cisplatin [12,13]. MMR components are also involved in the response to oxidative lesions in DNA such as 7,8-dihydro-8-oxoguanine (8oxoG) [14]. The initial steps of MMR are accomplished by heterodimers of MSH proteins, named MSH1 to MSH7 in eukaryotes. In *Saccharomyces cerevisiae*, the MSH2/MSH6 complex, or MutS α , recognizes nuclear base–base mismatches and small (1–2 bp) insertion–deletion (InDel) loops, while MSH2/MSH3 heterodimers, or MutS β , recognizes a range of small and larger loop-outs. *S. cerevisiae* MSH1 protein is involved in mitochondrial genome stability [15].

TcMSH2 was the first component of the MMR machinery described in *T. cruzi* [16]. Analyses of the *TcMSH2* single copy gene present in different strains of the parasite demonstrated the existence of three isoforms of this protein, named MSH2A, B and C, which correlate with the division in three phylogenetic *T. cruzi* lineages [17]. A role of MMR in creating differences in genetic variability among *T. cruzi* strains was initially unveiled by studies showing that strains representative of *T. cruzi* II displayed allelic variation of nuclear microsatellite *loci* induced by hydrogen peroxide, the same not occurring in a strain belonging to the *T. cruzi* I lineage [18]. Moreover, parasites from MSH2 haplogroups B and C were more resistant to cisplatin treatment, as previously reported in other MMR-deficient cells [13]. These studies suggest that, at least under genotoxic stress conditions, strains belonging to the *T. cruzi* I lineage (presenting the A isoform of *TcMSH2* gene) have a more efficient MMR activity than *T. cruzi* II strains [18]. They also support the hypothesis that the distinct genetic diversity found in the two *T. cruzi* lineages might be consequence of different levels of MMR efficiency. Therefore, to better understand the mechanisms involved in generating genetic diversity in *T. cruzi* and, more precisely, the role of the MMR pathway, we investigated the response of DNA damaging agents in strains representative of the two main lineages and the activities of different isoforms of *T. cruzi* MSH2 proteins. Using gene deletion analysis to examine the response to oxidative damage, we uncovered a novel function for MSH2, related to mitochondrial DNA repair, which is conserved in both *T. cruzi* and *T. brucei*.

2. Materials and methods

2.1. Strains and growth conditions

Five cloned *T. cruzi* strains (Col1.7G2, Silvio X-10 c11, Esmeraldo c13, JG and CL Brener) were used. Genotyping of each strain using several markers demonstrated that Col1.7G2, Silvio X-10c11 belong to *T. cruzi* I lineage whereas Esmeraldo c13 and JG are representatives of *T. cruzi* II lineage [19]. CL Brener is a hybrid strain (*T. cruzi* II/III), reference for the genome sequencing project [4] and it was selected for the knockout experiments. Epimastigotes were

maintained as described [20]. *T. brucei* wild type and *MSH2* deletion mutants were grown as previously described [21].

2.2. *TcMSH2* amino acid sequence analysis and northern blot assays

TcMSH2 sequences from CL Brener were retrieved from GeneDB (accession numbers Tc00.1047053509643.80 and Tc00.1047053507711.320), whereas the Col1.7G2 sequence was obtained from a cloned PCR amplified fragment as described [22]. Deduced amino acid sequences were aligned using ClustalW version 1.8 and all protein sequences were run against the Pfam [23], Prosite [24] and InterPro [25] databases. For Northern blot analysis, total RNA (25 μ g) was separated on formaldehyde agarose gels, blotted onto nylon membranes, cross-linked through UV irradiation and probed with [α -³²P] labeled *TcMSH2* as previously described [19]. The probe was labeled according to the Megaprime DNA labeling (GE Healthcare) protocol and the signals were quantified using the ImageJ program (<http://rsbweb.nih.gov/ij/>).

2.3. Treatment of the strains with genotoxic agents

Cultures with 10^7 parasites per mL were incubated in 24-well plates with different concentrations of MNNG (*N*-methyl-*N'*-nitro-*N*-nitrosoguanidine) (provided by Dr. Álvaro Augusto da Costa Leitão, Instituto de Biofísica Carlos Chagas Filho, RJ), cisplatin (Quiral Química do Brasil S/A), or H₂O₂ (Merck), in the presence or absence of 3 μ M of the MMR inhibitor cadmium chloride, as indicated. Previous studies have shown a wide range for the H₂O₂ IC₅₀ using various *T. cruzi* strains, varying from 98 to 190 μ M [26,27]. After incubation for 3 or 5 days, cell densities were measured with a haematocytometer using Erythrosin B exclusion.

2.4. Measurement of 8-oxoguanine accumulation

Two different protocols were used to assess the 8-oxoG accumulation in *T. cruzi* DNA strains. In the *in situ* experiment, a protocol adapted from Struthers et al. [28] was used. Epimastigotes were incubated in the presence of 200 or 300 μ M hydrogen peroxide for 20 min at 28 °C, washed twice with PBS and fixed with 4% paraformaldehyde. Aliquots (20 μ L) of the cell suspension were distributed into wells of 8-wells chambered-slides. After 1 h of incubation at 4 °C, cells were permeabilized with 0.2% Triton X-100, treated with 100 μ g/mL RNase A and incubated with FITC-conjugated avidin (5 μ g/mL final concentration) for 1 h at room temperature in the absence of light. After washing with PBS and mounted with a solution of 9:1 Glycerol:Tris–HCl, pH 9.0, the slides were visualized under a fluorescence microscope in a 100 \times oil immersion. Pre-incubation of FITC-conjugated avidin with 0.5 mM of 8oxodG results in 80% decrease in the fluorescence signal, whereas pre-incubation with dGTP has no significant effect on parasite labeling. Fluorescence intensities were averaged with the ImageJ program and plotted as fluorescence arbitrary units (average fluorescence intensity measured in 100 cells after subtracting the average background intensity). Background signals were measured in 100 fields, randomly chosen on the slides.

The 8-oxoG accumulation was also assessed by HPLC-electrochemical detection. Cells (10^9 /mL) were treated with 20 mM H₂O₂ for 1 h at 28 °C, washed with PBS, and the DNA was isolated by the chaotropic NaI method [29] in the presence of 0.1 mM desferrioxamine. DNA samples (100 μ g) were treated with nuclease P₁ and alkaline phosphatase and analyzed by HPLC. Samples (100 μ g) of digested DNA were injected into the HPLC/electrochemical detection system consisting of a Shimadzu model LC-10AD pump connected to a Luna C₁₈ (Phenomenex, Torrance, CA, USA) reverse-phase column (250 mm \times 4.6 mm ID, particle size 5 μ m). The flow

rate of the isocratic eluent (50 mM potassium phosphate buffer, pH 5.5, and 8% methanol) was 1 mL/min. Coulometric detection was obtained with a Coulochem II detector (ESA, Chemsford, MA, USA). The potentials of the two electrodes were set at 120 and 280 mV. Elution of unmodified nucleosides was monitored simultaneously with a Shimadzu SPD-10A UV detector set at 254 nm. A Shimadzu Class-LC10 1.6 software was used to calculate the peak areas. The molar ratio of 8-oxodG to dG in each DNA sample was determined based on coulometric detection at 280 mV for 8-oxodG and on absorbance at 254 nm for dG in each injection.

2.5. Analysis of nuclear and mitochondrial genomes in *T. brucei*

T. brucei cultures were treated with H₂O₂ as described previously [22]. Paraformaldehyde-fixed cells were spotted onto microscope slides mounted with vectashield containing DAPI (4,6-diamidino-2-phenylindole) (Vector Laboratories Inc.). DNA configurations of the cells were analyzed, blind, by 2 researchers independently, and >500 cells were counted for each data point. Microscopic analysis was performed using an Axioskop 2 microscope (Zeiss) and images obtained using Openlab software (Improvision).

2.6. Analysis of oxidative metabolism

Epimastigotes were incubated for 18 h at 28 °C in a 96-well plate in the presence of 20 µL of 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT) (5.0 mg/mL MTT in PBS pH 7.6 as stock solution) and the final volume adjusted to 200 µL with culture medium. The reaction was stopped by the addition of 1.0 mL 0.04 N hydrochloric acid in 2-propanol. After 30 min, the lysate was vigorously mixed until completely homogenized. The supernatant was separated by centrifugation at 200 × g for 10 min and its absorbance was measured at 570 nm. NADPH production by the pentose phosphate pathway was measured as previously described [27].

2.7. Plasmid constructions and parasite transfections

The full-length *TcMSH2* gene was amplified from genomic DNA (CL Brener strain). The PCR was performed using primers forward and reverse carrying XbaI and NotI restriction sites, respectively, and amplicons were inserted into TOPO pCR4 (Invitrogen) generating the plasmid TOPOTcMSH2. Fragments containing the resistance genes hygromycin phosphotransferase (HYG) or neomycin phosphotransferase (NEO) were obtained after digestion of the pROCK-GFP-HYG or pROCK-GFP-NEO [30] with XhoI/NheI and ligated to TOPOTcMSH2, digested with the same enzymes. The resultant plasmids presenting each one of the resistance genes flanked by N and C-terminal fragments of *TcMSH2* coding sequence were used to delete the *TcMSH2* alleles by homologous recombination. Before the transfection, both plasmids (TOPO N-Neo-C and TOPO N-Hyg-C) were digested with XbaI and NotI and the respective 1.7 Kb and 1.9 Kb fragments were gel-purified. Approximately 50 µg of gel-purified fragments was used to transfect epimastigotes by electroporation according to DaRocha et al. [30]. In order to generate *TcMSH2* knockouts, fragments containing 5' and 3' sequences of different alleles were also generated by gene SOEing [31]. *TcMSH2* upstream and downstream regions (approximately 500 pb each) were PCR-amplified and fused flanking a PCR-generated cassette containing the *HYG* gene. Transfection of epimastigotes was performed by electroporation with 50 µg DNA as described previously [30]. Twenty-four hours after transfection, 250 µg/mL of Hygromycin B or G418 was added to the cultures and selected population was obtained approximately 30 days after transfection.

Clones were obtained by limiting dilution or by plating on semi-solid agar plates, after 30 days of incubation at 28 °C.

2.8. Recombinant protein expression

The sub-fragment encoding the N-terminal region (amino acids 1–442) of the *TcMSH2* protein was obtained from digestion of the plasmid TOPO2.1 containing the *TcMSH2* gene and cloned directionally into the pET21a+ vector (Novagen), to generate a truncated His-tagged recombinant protein *TcMSH2(t)*. The partially purified recombinant protein obtained from isopropylthiogalactoside induced bacterial cultures was electrophoresed on a 10% SDS-polyacrylamide gel. The 48 kDa corresponding band was excised and used to raise polyclonal antibodies in BALB/c mice by subcutaneous inoculation with Freund's adjuvant. Animals were boosted twice at 2 weeks intervals and bled after 6 weeks.

We used a MBP (maltose binding protein) gene fusion system to sub-clone and express a full-length *TcMSH2B* gene, PCR-amplified from genomic DNA using the following primers flanked by EcoRI and XbaI restriction sites, respectively. The amplicon was inserted into TOPO pCR4, which was digested with EcoRI and XbaI and the fragment was inserted into pMal-c2G (New England Biolabs). The recombinant protein obtained from isopropylthiogalactoside-induced bacterial cultures at 28 °C was purified by affinity chromatography on amylose columns.

2.9. ATPase assays

ATPase assays were carried out as described [32], using 2 pmol of purified MSH2::MBP or MBP alone (total volume 20 µL). Aliquots (2 µL) were removed at different time intervals (0 and 2 h) and the reactions were terminated by the addition of an equal volume of formamide. Samples were kept on ice before separation on a 20% polyacrylamide gel. Gels were exposed to a photographic film for 1 h at room temperature and the bands corresponding to ATP were quantified using a Phosphoimager Typhoon 8600 (Amersham). To test the effect of cadmium chloride on the ATPase activity, the reagent was added to a final concentration of 10 µM.

2.10. Statistical analyses

The results shown in this work were from triplicate determinations and represent three independent experiments performed by identical methods. The unpaired *t* test was used to determine the statistical significance (*p* value) of differences at each pair of groups. All analyses were performed using a *p* value of 0.05 with independent samples with unknown and different variances. For multiple comparisons, the one-way analysis of variance (ANOVA) was used, followed by Tukey's test.

3. Results

3.1. *T. cruzi* strains display differences in the susceptibility to alkylating agents and hydrogen peroxide

Cisplatin, a drug widely used in the treatment of cancer, forms intra and inter-strand Platinum–DNA cross-links, which kill the cell if inadequately repaired. *N*-methyl-*N*'-nitro-*N*-nitrosoguanidine (MNNG) is an alkylating agent that induces methylation in the O⁶ position of guanine [13]. MMR-proficient cells are more susceptible to both drugs since MMR proteins may act as DNA damage signal transducers to downstream effectors, participating in the initiation of cell death [33]. We investigated the effect of cisplatin and MNNG on different *T. cruzi* strains by incubation with increasing drug concentrations and determining cell viability after 5 days. The two strains that belong to the *T. cruzi* I group, Silvio X-10 c11

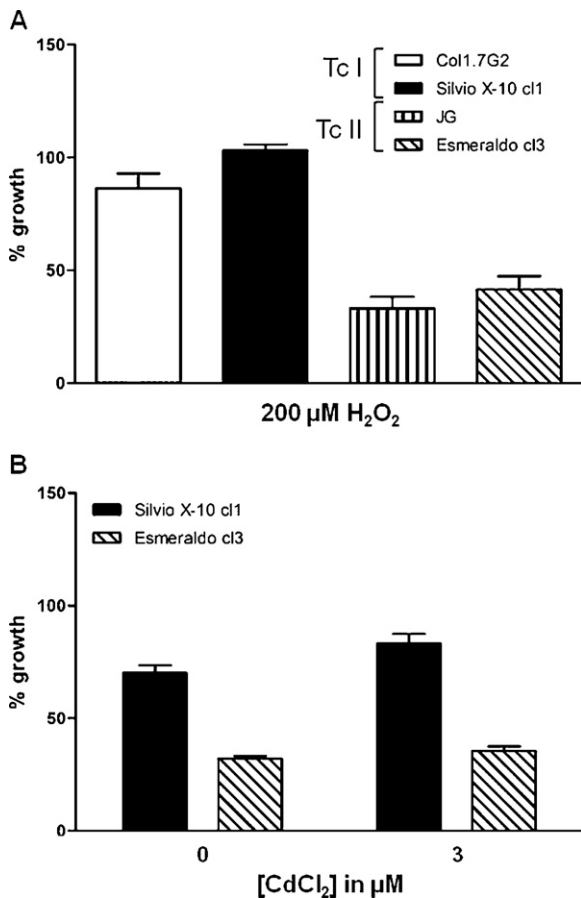


Fig. 1. Effect of H₂O₂-induced oxidative stress on the growth of *Trypanosoma cruzi* strains in the presence or absence of cadmium. Exponential cultures of epimastigotes were grown in LIT media containing (A) 200 μM H₂O₂ or (B) 200 μM H₂O₂ in the presence or absence of 3 μM Cd²⁺ for 5 days (B). Viable cells were determined with erythrosine B. Results are shown as mean ± SD of three independent experiments performed in triplicate.

and Col1.7G2, were more sensitive to treatment with both drug when compared to Esmeraldo c13 and JG strains (which are *T. cruzi* II strains). The results shown in Supplementary Fig. 1A and C extend our previous observations [18] to a wider range of strains, confirming that there is a clear distinction between strains from *T. cruzi* I and II groups regarding the response to both DNA damaging agents. At cisplatin concentrations of 20 and 30 μM and MNNG concentrations of 25 and 50 μM, the differences between *T. cruzi* I and *T. cruzi* II strains are statistically significant ($p \leq 0.05$). The mechanisms underlying the difference in the sensitivity to these drugs involve MMR, since, when we repeated the treatment in the presence of cadmium, a potent inhibitor of ATP binding and hydrolysis activities of MSH2-MSH6 complex [34], the differences previously observed between the strains were abolished (Supplementary Fig. 1B and D).

Oxidative DNA damage, which can arise from endogenous metabolism, represents a particularly significant threat for an intracellular parasite. To further compare the response of *T. cruzi* I and II parasites to genotoxic agents, we investigated the ability of different *T. cruzi* strains to proliferate in the presence of hydrogen peroxide (H₂O₂). Base lesions arising from oxidative DNA damage, such as 8-oxoG, if not removed by base-excision repair (BER), are recognized by MMR [14]. To test the response of *T. cruzi* strains to oxidative damage, 200 μM of H₂O₂ was added to epimastigote cultures and cell viability was determined after 5 days. As shown in Fig. 1A, the two groups of strains exhibited distinct responses to H₂O₂ ($p \leq 0.05$). In contrast to the results observed with cisplatin

and MNNG treatments, *T. cruzi* I strains (Col1.7G2 and Silvio X-10 c1) were more resistant to hydrogen peroxide treatment than *T. cruzi* II strains (JG and Esmeraldo c13). These differences cannot be attributed to differences in the redox potential of the strains analyzed, as indicated by the results of comparative analysis of the oxidative metabolism of the strains using the MTT reduction assay (Supplementary Fig. 2A) and measuring NADPH production (Supplementary Fig. 2B). Also, in contrast to the response to cisplatin and MNNG treatments, cadmium did not abolish the differential responses of the *T. cruzi* strains to H₂O₂. As shown in Fig. 1B, when parasites were incubated with hydrogen peroxide in the presence of this MMR inhibitor, survival rates of the Silvio X-10 (*T. cruzi* I strain) remain higher than the Esmeraldo strain (*T. cruzi* II) ($p \leq 0.05$).

To assess more directly the effect of H₂O₂, we examined the accumulation of 8-oxoG in the genome of different *T. cruzi* strains before and after H₂O₂ treatment using avidin-conjugated FITC. Avidin is shown to bind with high specificity to 8-oxoG and has been used to detect oxidative DNA damage in different cell types [28]. As shown in Fig. 2A, strains belonging to *T. cruzi* I lineage (Col1.7G2 and Silvio X-10 c1) present lower levels of the oxidized base than *T. cruzi* II strains (JG and Esmeraldo c13). This difference ($p \leq 0.05$) is due to an increased accumulation of 8-oxoG in the kDNA of *T. cruzi* II strains after 20 min in the presence of 200 μM H₂O₂ (Fig. 2B). An increased accumulation of 8-oxoG in the DNA of a *T. cruzi* II strain (*T. cruzi* II) compared to a *T. cruzi* I (Silvio X-10) was also observed in experiments using HPLC/electrochemical detection analysis. Because of its high sensitivity and low saturation values, this method detected a much higher difference (145-fold) in 8-oxoG levels (Supplementary Fig. 3).

To verify whether the differences in the response to DNA damaging drugs could be due to differences in the expression of MSH2 among strains, we performed northern and Western blot analyses. As shown in Supplementary Fig. 4, no obvious differences the expression of *TcMSH2* mRNA or protein in all strains analyzed were found.

3.2. *TcMSH2* from *T. cruzi* I strains displays higher ATPase activity

To directly investigate the role of *TcMSH2* in the differential response of parasite strains to genotoxic treatments, we performed sequence analyses as well as *in vitro* assays with two recombinant isoforms of this protein. Whereas Col1.7G2 is a homozygous *T. cruzi* I strain, presenting two identical *TcMSH2A* alleles, CL Brener is a hybrid cloned strain and presents two *TcMSH2* alleles corresponding to MSH2-haplogroups B or C [18]. The alignment of the complete amino acid sequences of *TcMSH2* B and C derived from CL Brener, and the *TcMSH2* A from Col1.7G2, identified all common domains and motifs present in MSH2 from other organisms, as described for the *Thermus aquaticus* MutS [16]. A total of 29 amino acid substitutions, concentrated in domains I and II (mismatch recognition domain and connector domain) and domain V (which contains the Walker ATPase motif and the helix-turn-helix motif involved in the protein dimerization), was identified among the three MSH2 haplogroups, 6 of them corresponding to non-conservative substitutions (Supplementary Fig. 5).

MSH2 are ATPases and, while ATP binding seems not be required for initial recognition of mismatched substrate, ATP hydrolysis is involved in the subsequent steps of MMR [35,36]. Using *in vitro* assays, we compared ATP hydrolysis activities of recombinant *TcMSH2* derived from *T. cruzi* I and II strains. As shown in Fig. 3A and C, recombinant *TcMSH2* A, the isoform present in *T. cruzi* I strains shows higher levels of ATP hydrolysis after 2 h incubation when compared with the *TcMSH2* B isoform, present in *T. cruzi* II strains. In Fig. 3B, it is also shown that cadmium has an inhibitory effect on the *in vitro* *TcMSH2* ATPase activity.

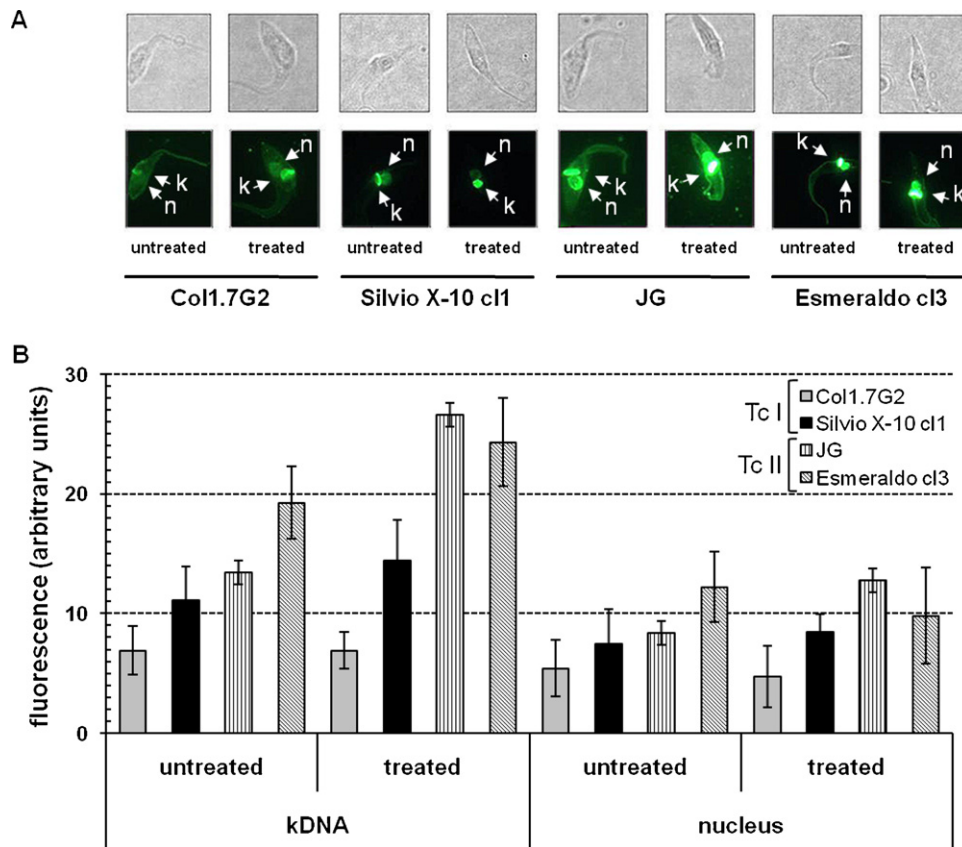


Fig. 2. Accumulation of 8-oxoguanine in different *T. cruzi* strains. Exponential cultures of epimastigotes (early log phase) were treated with 200 μ M H_2O_2 for 20 min and fixed with 4% paraformaldehyde. After transferred to chamber glass slides they were incubated with 5 μ g/mL avidin-FITC for 1 h. Slides were visualized in a bright field or by fluorescence (A) in a 100 \times oil immersion and fluorescence intensity was averaged with the ImageJ program (<http://rsbweb.nih.gov/ij/>) and plotted as fluorescence arbitrary units, as shown in B. The positions of nucleus (n) and the kinetoplast (k) are shown. Each bar represents the mean \pm SD of 100 cells analyzed.

3.3. *TcMSH2* single knockout clones are more sensitive to oxidative stress

To examine genetically the involvement of the MMR pathway in the differential response of *T. cruzi* to DNA damage, we attempted

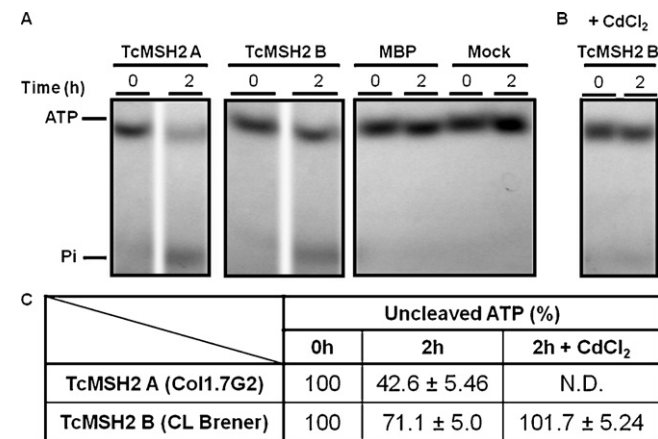


Fig. 3. ATPase activity of MSH2 isoforms A (from Col1.7G2) and B (from CL Brener). Affinity purified TcMSH2::MBP proteins were incubated with [γ -³²P]ATP for 2 h at 37 $^{\circ}$ C. Negative controls were carried out in the absence of protein or in the presence of MBP only. The position of ATP and inorganic phosphate (Pi) are indicated (A). The effect of cadmium chloride (CdCl₂) on the ATPase activity of TcMSH2B was evaluated by adding 10 μ M CdCl₂ to the reaction buffer prior to the assay (B) Radioactive bands corresponding to ATP were quantified using the Storm Phosphorimager (GE-HealthCare). ND: not determined (C).

to generate null mutants for the *TcMSH2* gene. Using plasmid constructs aimed at homologous recombination with the *TcMSH2* gene, single allele deletions of *TcMSH2* ($\Delta TcmsH2::HYG/TcMSH2$ parasites) could readily be achieved in different parasite strains. However after several attempts, including transfections with constructs that are specific for both CL Brener alleles, we were not able to generate *TcMSH2* null mutants. As shown in **Supplementary Fig. 6A**, DNA amplifications with primers specific for *TcMSH2* alleles B and C, which are present in the CL Brener genome (see **Supplementary Fig. 6B**) and DNA isolated from two transfected cloned cell lines showed that the resistance gene was integrated at the *TcMSH2-B* allele. Disruption of the *TcMSH2-B* allele was also confirmed by Southern blot (**Supplementary Fig. 6C and D**) and by the analysis of *TcMSH2* transcript levels, which were reduced to about 50% after the single-allele deletion (**Supplementary Fig. 6E**). Analysis of more than 10 distinct $\Delta TcmsH2::HYG$ clones showed that, in every case, deletion of the *TcMSH2-B* allele had occurred (data not shown).

To investigate the phenotype of *TcMSH2* mutants, epimastigote culture of wild type and $\Delta TcmsH2::HYG/TcMSH2$ clones were compared. No significant difference of the growth curves of wild type and mutant cells were observed (data not shown). We also compared the growth of wild type and $\Delta TcmsH2::HYG/TcMSH2$ clones in the presence of increasing concentrations of cisplatin and found no differences (**Fig. 4A**). In contrast, when treated with 200 or 300 μ M H_2O_2 , all three cloned $\Delta TcmsH2::HYG/TcMSH2$ mutants presented a significantly lower tolerance ($p \leq 0.05$) than wild type cells (**Fig. 4B**). The effect of H_2O_2 -induced oxidative damage was also observed by measuring the levels of 8-oxoG in the genome of wild type and $\Delta TcmsH2::HYG/TcMSH2$ cells (**Fig. 4C**). Although we observed no

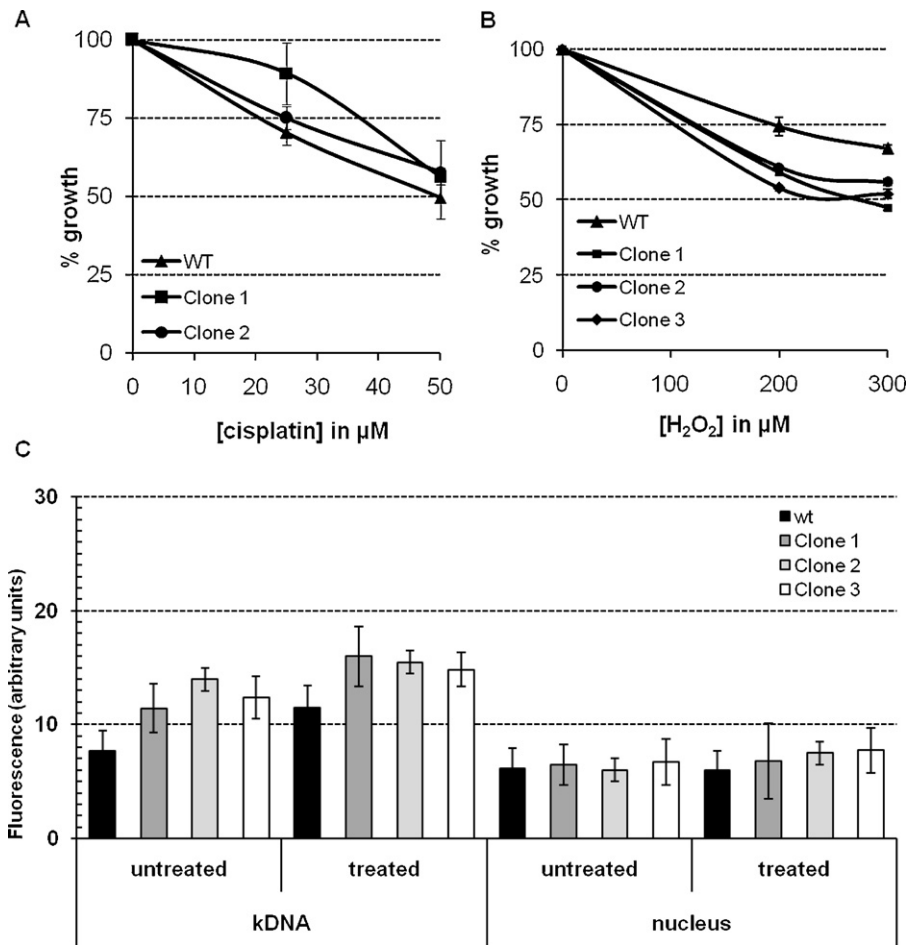


Fig. 4. Survival of $\Delta Tcmsh2::HYG/TcMSH2$ clones in the presence of cisplatin and H_2O_2 and accumulation of 8-oxoguanine after oxidative stress damage. Wild type and $\Delta Tcmsh2::HYG/TcMSH2$ clones were treated in early log phase with 25 and 50 μM cisplatin (A) or 200 and 300 μM H_2O_2 (B). After 5 days, viable cell numbers were determined using erythrosine B dye exclusion. Accumulation of 8-oxoG in the nucleus and kinetoplast DNA in wild type and $\Delta Tcmsh2::HYG/TcMSH2$ mutants was measured before and after treatment with 300 μM H_2O_2 (C). Each bar represents the mean \pm SD of 100 cells analyzed. Statistical differences ($p \leq 0.05$) were observed between clone 2 and wt cells in the analysis of kDNA both in untreated and treated parasites.

significant differences in the levels of 8-oxoG present in the nuclei of wild type and mutant strains before or after treatment with H_2O_2 , increased amounts of the oxidized base were observed in the kDNA of $\Delta Tcmsh2::HYG/TcMSH2$ mutant cell lines compared to wild type parasites ($p \leq 0.05$).

3.4. *Trypanosoma brucei* MSH2 mutants display loss of mitochondrial DNA due to oxidative damage

In contrast to *T. cruzi*, MSH2 null mutants have been generated in *T. brucei* and are viable, at least in the bloodstream (mammalian-infective) stage of this parasite [21]. Similar to *T. cruzi* $\Delta Tcmsh2::HYG/TcMSH2$ clones, *T. brucei* MSH2 null mutants ($\Delta Tbmsh2::BSD/Tbmsh2::PUR$) are more sensitive to H_2O_2 than wild type [21]. To ask if this phenotype is also a manifestation of mitochondrial DNA damage, *Tbmsh2* null mutants and wild type parasites were grown in increasing concentrations of H_2O_2 for 72 h and the DNA content of the cells analyzed microscopically after staining with DAPI. Kinetoplast (K) and nuclear (N) DNA synthesis and segregation occur at distinct times during the *T. brucei* cell cycle [37], meaning that parasites with 1N1K, 1N2K and 2N2K contents are found, normally representing the majority (~98%) of cell types in wild type cultures; the small minority of cells with differing N–K ratios ('others') arise due to errors in DNA replication, segregation or mitosis. Here, we found that, even before H_2O_2 treatment,

Tbmsh2 mutants displayed an elevation in the proportion of aberrant cells. Whereas the very small numbers of these aberrant cells in the wild type were essentially equally distributed between 0N1K and 2N1K types (Fig. 5A), virtually all were characterized by a loss of kDNA (predominantly having a 1N0K configuration) in the *Tbmsh2* null mutants (Fig. 5A–C). Treatment of *Tbmsh2* mutants with H_2O_2 accentuated this loss of kDNA to an extent that was substantially more pronounced ($p \leq 0.05$) than in wild type cells (Fig. 5A). These data suggest that in *T. brucei*, MSH2 deficiency, before and after H_2O_2 -induced damage, results in accelerated loss of mitochondrial DNA, consistent with a function for MSH2 in mitochondrial genome repair or protection.

4. Discussion

MMR corrects mismatches generated during DNA replication. It also contributes to the control of endogenous levels of oxidized DNA bases present in the genome such as 8-oxoG [14]. In the *T. cruzi* population, MSH2, a central component of MMR, is encoded by three distinct alleles which can be correlated with the three major *T. cruzi* lineages [17,18]. Because distinct levels of genetic variability have been described among parasite strains belonging to each MSH2 haplogroup, we investigated the activity of this MMR protein and its involvement in the response of different parasite strains to various DNA damaging agents. In so doing, we reveal two things.

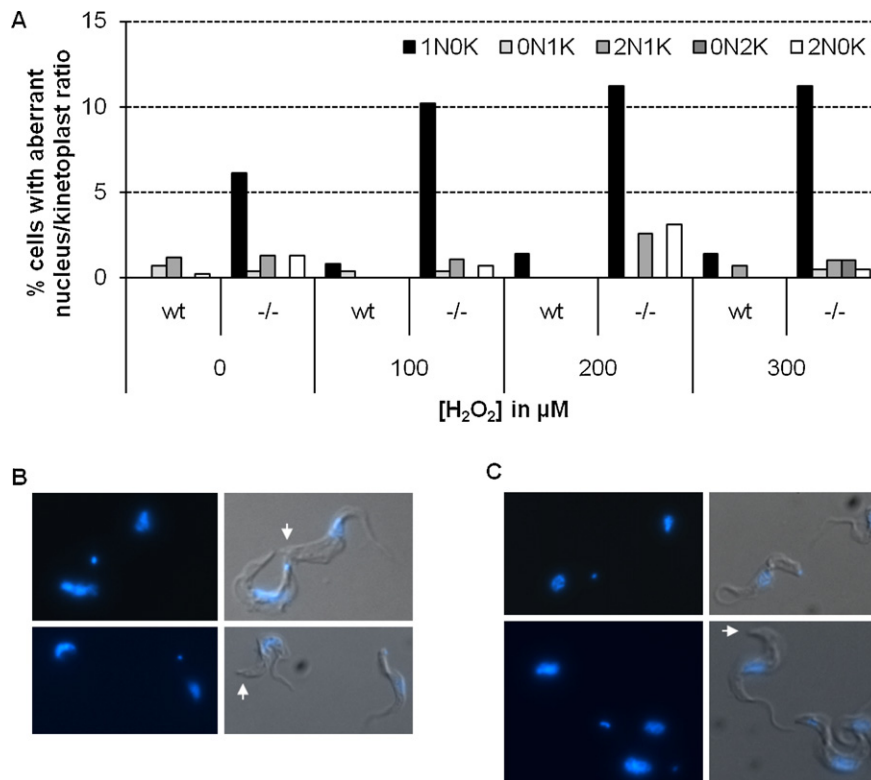


Fig. 5. Effect of MSH2 deletion on the DNA content of *T. brucei* before and after hydrogen peroxide exposure. (A) Quantitation of the nuclear (N) and kinetoplast (K) DNA content of *T. brucei* wild type and $\Delta Tcmsh2::BSD/Tcmsh2::PUR$ mutants after 72 h growth in the presence or absence of increasing concentrations of H₂O₂. DNA in the cells was visualized by DAPI staining; >500 cells were counted in each sample. Only aberrant cells that differ from the expected N/K ratios (1N1K, 1N2K or 2N2K) are shown, as a percentage of the total population. (B, C) Examples of $\Delta Tcmsh2::BSD/Tcmsh2::PUR$ cells lacking detectable kDNA are shown after 72 h growth in 100 μ M H₂O₂ (B) or without H₂O₂ treatment (C); in all cases the cells are shown both as a DAPI-stain image or as a merge of DAPI and differential interference contrast images and arrows denote the cells lacking kDNA, which are shown beside 1N1K cells for comparison.

First, the level of MMR activity is distinct between *T. cruzi* lineage I and II strains. Second, we obtained evidences indicating that MSH2 in both *T. cruzi* and *T. brucei* provides a hitherto unseen function in protecting mitochondrial DNA from oxidative damage.

The involvement of MMR in generating genetic diversity and drug resistance has been described in several organisms, including pathogenic bacteria and other protozoan parasites. Analyses of hypermutable bacteria of different species isolated from human clinical isolates have shown that *mutS* inactivation is implicated in the mutator phenotype [38]. In *Toxoplasma gondii*, disruption of mitochondrial *MSH1* was found in parasites selected for monensin-resistance, a drug proven to be effective against apicomplexan parasites. As expected, the *MSH1* mutants were also resistant to the alkylating agent methylnitrosourea [39]. *Plasmodium* species are the only eukaryotes that encode two *MSH2* homologs. Disruption of one the *MSH2* loci in *Plasmodium falciparum* resulted in a slightly higher frequency of parasites resistant to 5-Fluoroorotate when mosquito-passaged mutants and wild type parasites were compared, suggesting that the function of the two *PfMSH2* genes may overlap [40]. Here we show that *T. cruzi* I and II strains respond differently when treated with genotoxic agents, such as cisplatin and MNNG. Because this difference was abolished when MMR activity was inhibited by cadmium, we suggest that the differential response among strains to these DNA damaging agents is due to differential activity of one of more components of the MMR. In agreement with previous studies with MMR mutants in different eukaryotic cells [13], the observed increased tolerance to cisplatin and MNNG in *T. cruzi* II parasites can be inferred to result from less efficient MMR. This is also consistent with previous results showing that *T. cruzi* II strains have increased nuclear genetic variability when compared with strains belonging to *T.*

cruzi I lineage, an observation that may have major epidemiological significance [6]. The differences observed between strains in the response to genotoxic agents cannot be associated with different levels of expression of *TcMSH2*, since, in all strains analyzed, the levels of *TcMSH2* mRNA and protein were found to be similar. However, because *T. cruzi* I and *T. cruzi* II strains present distinct *TcMSH2* isoforms, we proposed that sequence polymorphisms in *TcMSH2* may be at least partially responsible for the differences in the response to these DNA damaging agents, an assumption that was corroborated by the results showing differences in ATPase activities from recombinant proteins MSH2A and MSH2B. Notwithstanding, variations in the efficiency of other components of MMR, as well as from other DNA repair pathways are also likely to be involved in modulating mutation rates in the *T. cruzi* population. Attempts to determine whether differences in mutation rates can be observed among distinct strains are underway.

To further investigate the role of *TcMSH2* protein, we attempted to generate null mutants for the *TcMSH2* gene. Unexpectedly, we were unable to delete both alleles. This inability to generate *MSH2* null mutants is in sharp contrast with most other eukaryotes, including *T. brucei* bloodstream forms. Bloodstream forms of *TbMSH2* null mutants are viable, present an increased rate of sequence variation of nuclear microsatellite loci and increased tolerance to MNNG, as well as an increased frequency of homologous recombination [21,41]. The availability of *TbMSH2* null mutants allowed us to test whether the *TcMSH2* gene can complement MMR deficiency in *T. brucei* [22]. Although we were able to demonstrate heterologous expression of *T. cruzi* MSH2, MMR complementation, as assayed by resistance to MNNG and microsatellite instability, was not achieved. However, expression of *TcMSH2* in *TbMSH2* null

mutants resulted in reversion of the sensitivity to H₂O₂-induced oxidative stress. The observation described here that this mutation gives rise to a significant percentage of the dyskinetoplastic cells in *T. brucei*, even in the absence of induced oxidative stress, is striking. In contrast to *T. cruzi* epimastigote metabolism, which is largely dependent on oxidative phosphorylation and generates endogenous oxidative damage, *T. brucei* bloodstream forms rely exclusively on glycolysis for ATP production, which may explain why *TbMSH2* null mutants are viable.

Taken together, the results of the complementation assays in *T. brucei*, the non-viability of *TcMSH2* mutants, and the common phenotypes observed in the *TcMSH2* single knockout and *TbMSH2* null mutants suggest that MSH2 provides a function related to mitochondrial genome integrity in addition to its role in nuclear MMR. Such additional functions may be independent of at least some other components of MMR, since *TbMLH1* null mutants do not present increased sensitivity to H₂O₂ [22]. This finding is consistent with the observation here that incubation of *T. cruzi* cells with cadmium abolished the differences observed between strains in response to treatment with cisplatin and MNNG, but not with H₂O₂:cadmium is known to suppress the ATPase activity of MSH2/MSH6 complexes [33], which is likely to impair interaction with MLH1-containing heterodimers [36]. In both *Trypanosoma* species MSH2 abundance in *MSH2* mutants appears sufficient for nuclear MMR function. However, the *Tcmsh2* mutants showed higher sensitivity to H₂O₂-induced oxidative stress than wild type cells and increased accumulation of 8-oxoG, particularly in the kDNA. It can be speculated that the involvement of *TcMSH2* in the response to oxidative stress is so relevant to mitochondrial function that a loss of a single allele would have a significant impact. The nature of such a mitochondrial role for MSH2 is unknown. Perhaps the protein compensates for the lack of MSH1, a mitochondrial DNA repair protein, for which an orthologous gene could not be identified in the *T. cruzi* genome [4]. A role of MSH2 in the control of oxidative DNA damage has been previously described in other cell types, such as mouse embryo fibroblasts in which increased levels of DNA 8-oxoG were detected after inactivation of the *MSH2* gene [42]. Additional functions of MSH2, MSH3, and PMS2 mismatch repair proteins have also been described, including homologous recombination, anti-recombination, DNA damage signaling, apoptosis, as well as site-specific mutagenesis during immunoglobulin somatic hypermutation and class switch recombination (reviewed in [43]). Nevertheless, these functions have been limited to nuclear activities, and the mitochondrial role we see here for MSH2 appears not to have been described thus far in any other organism.

Being an intracellular parasite, oxidative damage may have a significant effect on mutation rates in *T. cruzi*. In addition of MMR, other DNA repair pathways such as Base Excision Repair (BER) are certainly involved in the parasite response to oxidative damage in DNA. We and others have characterized two DNA polymerase beta from *T. cruzi* and showed that both enzymes co-localize with the parasite kinetoplast [44]. *T. cruzi* has a complex antioxidant defense system that includes enzymes like Fe-SOD (mitochondrial iron-containing superoxide dismutase), TcMPX and TcCPX (mitochondrial and cytoplasmic trypanedoxin peroxidases, respectively), and TcAPX (ascorbate-dependent haemoperoxidase), which are upregulated during transformation of the insect-derived non-infective epimastigotes into the infective metacyclic trypomastigote [45]. The evidences presented here indicating a novel role of MSH2 in *T. cruzi* and *T. brucei* in the complex mechanisms responsible for the oxidative stress response require further investigation, particularly regarding the study of MSH2 sub-cellular localization and interactions with other proteins. Whether this new role of MSH2 is part of the driving force behind the differential genetic variability observed within the *T. cruzi* population, it also remains to be investigated.

Acknowledgements

This work is supported by funds from Conselho Nacional de Desenvolvimento Científico e Tecnológico – CNPq, Fundação de Amparo a Pesquisa do Estado de Minas Gerais – FAPEMIG (Brazil) and the Howard Hughes Medical Institute – HHMI. The work from MHGM was funded by Fundação de Amparo à Pesquisa do Estado de São Paulo – FAPESP (Brazil), and the INCT de Processos Redox em Biomedicina – Redoxoma (Brazil). RMcC thanks the Medical Research Council and Wellcome Trust for funding. We thank Osmar F. Gomes for help with 8-oxoG analyses and Glynn Forsythe for help with microscopy.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.molbiopara.2010.11.001.

References

- [1] Machado CR, Augusto-Pinto L, McCulloch R, Teixeira SM. DNA metabolism and genetic diversity in *Trypanosoma* species. *Mutat Res* 2006;612:40–57.
- [2] Momen H. Taxonomy of *Trypanosoma cruzi*: a commentary on characterization and nomenclature. *Mem Inst Oswaldo Cruz* 1999;94(Suppl 1):181–4.
- [3] Zingales B, Andrade SG, Briones MR, et al. A new consensus for *Trypanosoma cruzi* intraspecific nomenclature: second revision meeting recommends TcI to TcVI. *Mem Inst Oswaldo Cruz* 2009;104:1051–4.
- [4] El-Sayed NM, Myler PJ, Bartholomeu DC, et al. The genome sequence of *Trypanosoma cruzi*, etiologic agent of Chagas disease. *Science* 2005;309:409–15.
- [5] Machado CA, Ayala FJ. Sequence variation in the dihydrofolate reductase–thymidylate synthase (DHFR–TS) and trypanothione reductase (TR) genes of *Trypanosoma cruzi*. *Mol Biochem Parasitol* 2002;121:33–47.
- [6] Cerqueira GC, Bartholomeu DC, DaRocha WD, et al. Sequence diversity and evolution of multigene families in *Trypanosoma cruzi*. *Mol Biochem Parasitol* 2008;157:65–72.
- [7] Mathieu-Daude F, Bosseno MF, Garzon E, et al. Sequence diversity and differential expression of Tc52 immuno-regulatory protein in *Trypanosoma cruzi*: potential implications in the biological variability of strains. *Parasitol Res* 2007;101:1355–63.
- [8] Trojan J, Zeuzem S, Randolph A, et al. Functional analysis of hMLH1 variants and HNPCC-related mutations using a human expression system. *Gastroenterology* 2002;122:211–9.
- [9] Branch P, Aquilina G, Bignami M, Karran P. Defective mismatch binding and a mutator phenotype in cells tolerant to DNA damage. *Nature* 1993;362:652–4.
- [10] Clark AB, Cook ME, Tran HT, Gordenin DA, Resnick MA, Kunkel TA. Functional analysis of human MutSalpha and MutSbeta complexes in yeast. *Nucleic Acids Res* 1999;27:736–42.
- [11] Mena A, Smith EE, Burns JL, et al. Genetic adaptation of *Pseudomonas aeruginosa* to the airways of cystic fibrosis patients is catalyzed by hypermutation. *J Bacteriol* 2008;190:7910–7.
- [12] Cejka P, Jiricny J. Interplay of DNA repair pathways controls methylation damage toxicity in *Saccharomyces cerevisiae*. *Genetics* 2008;179:1835–44.
- [13] Stojic L, Brun R, Jiricny J. Mismatch repair and DNA damage signalling. *DNA Repair (Amst)* 2004;3:1091–101.
- [14] Slupphaug G, Kavli B, Krokan HE. The interacting pathways for prevention and repair of oxidative DNA damage. *Mutat Res* 2003;531:231–51.
- [15] Pogorzala LA, Mookerjee S, Sia EA. Evidence that Msh1p plays multiple roles in mitochondrial base excision repair. *Genetics* 2009.
- [16] Augusto-Pinto L, Bartholomeu DC, Teixeira SM, Pena SD, Machado CR. Molecular cloning and characterization of the DNA mismatch repair gene class 2 from the *Trypanosoma cruzi*. *Gene* 2001;272:323–33.
- [17] de Freitas JM, Augusto-Pinto L, Pimenta JR, et al. Ancestral genomes, sex, and the population structure of *Trypanosoma cruzi*. *PLoS Pathog* 2006;2:e24.
- [18] Augusto-Pinto L, Teixeira SM, Pena SD, Machado CR. Single-nucleotide polymorphisms of the *Trypanosoma cruzi* MSH2 gene support the existence of three phylogenetic lineages presenting differences in mismatch-repair efficiency. *Genetics* 2003;164:117–26.
- [19] Andrade LO, Machado CR, Chiari E, Pena SD, Macedo AM. Differential tissue distribution of diverse clones of *Trypanosoma cruzi* in infected mice. *Mol Biochem Parasitol* 1999;100:163–72.
- [20] Teixeira SM, Kirchhoff LV, Donelson JE. Post-transcriptional elements regulating expression of mRNAs from the amastin/tuzin gene cluster of *Trypanosoma cruzi*. *J Biol Chem* 1995;270:22586–94.
- [21] Bell JS, Harvey TI, Sims AM, McCulloch R. Characterization of components of the mismatch repair machinery in *Trypanosoma brucei*. *Mol Microbiol* 2004;51:159–73.
- [22] Machado-Silva A, Teixeira SM, Franco GR, et al. Mismatch repair in *Trypanosoma brucei*: heterologous expression of MSH2 from *Trypanosoma cruzi* provides new insights into the response to oxidative damage. *Gene* 2008;411:19–26.

- [23] Sonnhammer EL, Eddy SR, Durbin R. Pfam: a comprehensive database of protein domain families based on seed alignments. *Proteins* 1997;28:405–20.
- [24] Bairoch A. PROSITE: a dictionary of sites and patterns in proteins. *Nucleic Acids Res* 1991;19(Suppl):2241–5.
- [25] Apweiler R, Attwood TK, Bairoch A, et al. InterPro – an integrated documentation resource for protein families, domains and functional sites. *Bioinformatics* 2000;16:1145–50.
- [26] Finzi JK, Chiavegatto CW, Corat KF, et al. *Trypanosoma cruzi* response to the oxidative stress generated by hydrogen peroxide. *Mol Biochem Parasitol* 2004;133:37–43.
- [27] Mielniczki-Pereira AA, Chiavegatto CM, Lopez JA, Colli W, Alves MJ, Gadelha FR. *Trypanosoma cruzi* strains. Tulahuen 2 and Y, besides the difference in resistance to oxidative stress, display differential glucose-6-phosphate and 6-phosphogluconate dehydrogenases activities. *Acta Trop* 2007;101:54–60.
- [28] Struthers L, Patel R, Clark J, Thomas S. Direct detection of 8-oxodeoxyguanosine and 8-oxoguanine by avidin and its analogues. *Anal Biochem* 1998;255:20–31.
- [29] Nakae D, Mizumoto Y, Kobayashi E, Noguchi O, Konishi Y. Improved genomic/nuclear DNA extraction for 8-hydroxydeoxyguanosine analysis of small amounts of rat liver tissue. *Cancer Lett* 1995;97:233–9.
- [30] DaRocha WD, Silva RA, Bartholomeu DC, et al. Expression of exogenous genes in *Trypanosoma cruzi*: improving vectors and electroporation protocols. *Parasitol Res* 2004;92:113–20.
- [31] Horton RM, Cai ZL, Ho SN, Pease LR. Gene splicing by overlap extension: tailor-made genes using the polymerase chain reaction. *Biotechniques* 1990;8:528–35.
- [32] Hughes MJ, Jiricny J. The purification of a human mismatch-binding protein and identification of its associated ATPase and helicase activities. *J Biol Chem* 1992;267:23876–82.
- [33] Topping RP, Wilkinson JC, Scarpinato KD. Mismatch repair protein deficiency compromises cisplatin-induced apoptotic signaling. *J Biol Chem* 2009;284:14029–39.
- [34] Jin YH, Clark AB, Slebos RJ, et al. Cadmium is a mutagen that acts by inhibiting mismatch repair. *Nat Genet* 2003;34:326–9.
- [35] Drotschmann K, Yang W, Kunkel TA. Evidence for sequential action of two ATPase active sites in yeast Msh2–Msh6. *DNA Repair (Amst)* 2002;1:743–53.
- [36] Lin DP, Wang Y, Scherer SJ, et al. An Msh2 point mutation uncouples DNA mismatch repair and apoptosis. *Cancer Res* 2004;64:517–22.
- [37] Hammarton TC, Monnerat S, Mottram JC. Cytokinesis in trypanosomatids. *Curr Opin Microbiol* 2007;10:520–7.
- [38] Hall LM, Henderson-Begg SK. Hypermutable bacteria isolated from humans – a critical analysis. *Microbiology* 2006;152:2505–14.
- [39] Garrison EM, Arrizabalaga G. Disruption of a mitochondrial MutS DNA repair enzyme homologue confers drug resistance in the parasite *Toxoplasma gondii*. *Mol Microbiol* 2009;72:425–41.
- [40] Bethke L, Thomas S, Walker K, Lakhia R, Rangarajan R, Wirth D. The role of DNA mismatch repair in generating genetic diversity and drug resistance in malaria parasites. *Mol Biochem Parasitol* 2007;155:18–25.
- [41] Bell JS, McCulloch R. Mismatch repair regulates homologous recombination, but has little influence on antigenic variation, in *Trypanosoma brucei*. *J Biol Chem* 2003;278:45182–8.
- [42] Russo MT, De Luca G, Casorelli I, et al. Role of MUTYH and MSH2 in the control of oxidative DNA damage, genetic instability, and tumorigenesis. *Cancer Res* 2009;69:4372–9.
- [43] Jun SH, Kim TG, Ban C. DNA mismatch repair system. Classical and fresh roles. *FEBS J* 2006;273:1609–19.
- [44] Lopes Dde O, Schamber-Reis BL, Regis-da-Silva CG, et al. Biochemical studies with DNA polymerase beta and DNA polymerase beta-PAK of *Trypanosoma cruzi* suggest the involvement of these proteins in mitochondrial DNA maintenance. *DNA Repair (Amst)* 2008;7:1882–92.
- [45] Piacenza L, Zago MP, Peluffo G, Alvarez MN, Basombrio MA, Radi R. Enzymes of the antioxidant network as novel determiners of *Trypanosoma cruzi* virulence. *Int J Parasitol* 2009;39:1455–64.