

UNIVERSIDADE FEDERAL DO PARANÁ

SIBELLE SELL SANTANA

**NOCAUTE DE PROTEÍNAS POSITIVAMENTE REGULADAS NAS FASES
INICIAIS DA METACICLOGÊNESE DE *Trypanosoma cruzi* PROVOCA A
DIMINUIÇÃO DA CAPACIDADE PROLIFERATIVA E DE DIFERENCIAÇÃO DO
PARASITA**

CURITIBA

2014

SIBELLE SELL SANTANA

**NOCAUTE DE PROTEÍNAS POSITIVAMENTE REGULADAS NAS FASES
INICIAIS DA METACICLOGÊNESE DE *Trypanosoma cruzi* PROVOCA A
DIMINUIÇÃO DA CAPACIDADE PROLIFERATIVA E DE DIFERENCIAÇÃO DO
PARASITA**

Monografia apresentada à disciplina
Estágio Supervisionado em Biologia como
requisito parcial à conclusão do Curso de
Ciências Biológicas – Bacharelado, Setor
de Ciências Biológicas, Universidade
Federal do Paraná.

Orientadora: Dra. Lyris Martins Franco de
Godoy
Coorientadora: Dra. Márcia Helena
Mendonça

CURITIBA

2014

PARTICIPANTES

Aluno

Sibelle Sell Santana

Curso de Graduação/Departamento-Universidade/Período

Ciências Biológicas / Universidade Federal do Paraná / 11^o período

Orientador/Unidade/Instituição

Dra. Lyris Martins Franco de Godoy / Instituto Carlos Chagas - ICC/ Fundação Oswaldo Cruz – PR

Área de Conhecimento

Biologia Molecular

Coorientador/Unidade/Instituição

Dra. Márcia Helena Mendonça / Universidade Federal do Paraná – UFPR / Departamento de Biologia Celular

Área de Conhecimento

Biologia Celular

Integração com Outros Projetos/Unidades/Instituições

Caracterização de Proteínas Diferencialmente Expressas nas Fases Iniciais da Metaciclogênese de *Trypanosoma cruzi* / Instituto Carlos Chagas / Fundação Oswaldo Cruz – PR

Dedico este trabalho a meus pais, minhas orientadoras e meus amigos. Através deles compreendi que nossa coragem e força nos tornam capazes de superar nossos maiores medos, acalmar inquietações, superar desafios e delinear a trajetória de muitas conquistas.

AGRADECIMENTOS

Agradeço minha família pelos votos de confiança, amparo e encorajamento nos momentos menos tranquilos e por me mostrarem que, por mais difícil que as situações possam parecer, há sempre um amanhã com novas oportunidades.

Especialmente a meus pais que me ensinaram que nosso melhor momento é agora. O presente é consequência do passado e o futuro virá como resposta ao hoje. Que devemos aprender com nossos erros e que, sempre é hora para recomeçar.

A minha orientadora Dra. Lyris Martins Franco de Godoy pelo seu conhecimento, experiência, tranquilidade, direcionamento, paciência, críticas e força de vontade ao longo do desenvolvimento deste trabalho.

A minha coorientadora Dra. Márcia Helena Mendonça por sua sabedoria, maturidade, dedicação e, por acreditar que podemos resgatar antigos valores e construir um mundo com mais amor e respeito.

Ao Dr. Fabricio Klerynton Marchini por seu conhecimento, apoio e ideias claras e precisas.

A minha companheira de laboratório Aline Castro Rodrigues Lucena por seu conhecimento, experiência, dedicação e paciência durante as discussões, experimentos e redação do trabalho.

Aos meus colegas do Instituto Carlos Chagas, da universidade e de minhas moradias temporárias, que me incentivaram com palavras de otimismo e esperança, que me ensinaram que devemos lutar por algumas coisas. Devemos lutar por aquilo que acreditamos e que nos proporciona crescimento. Alguém me disse certa vez que devemos fechar algumas portas. Não por egoísmo ou orgulho. Mas simplesmente porque elas já não nos levam a lugar algum.

É com muito respeito e gratidão que digo Muito Obrigada.

“Embora ninguém possa voltar atrás e fazer um novo começo qualquer um pode começar agora a fazer um novo fim”.

“Cada dia que amanhece assemelha-se a uma página em branco, na qual gravamos os nossos pensamentos, ações e atitudes. Na essência, cada dia é a preparação de nosso próprio amanhã”.

Francisco Cândido Xavier (1910 – 2002)

RESUMO

A Doença de Chagas, causada pelo protozoário flagelado *Trypanosoma cruzi*, ocorre na América Central e do Sul. Estima-se que 16 a 18 milhões de pessoas estão acometidas por esta patologia, sendo responsável por 21 mil de mortes a cada ano. Até o presente momento não há vacina ou medicamento eficiente para a sua fase crônica. O *T. cruzi* possui um ciclo de vida complexo, com um hospedeiro invertebrado e um hospedeiro mamífero, incluindo o homem. É na porção final do tubo digestivo do inseto que ocorre a metaciclogênese, com a diferenciação do *T. cruzi* de uma forma replicativa e não-infectiva (epimastigota) para uma forma infectiva e não-replicativa (tripomastigota metacíclica). Tal processo envolve alterações na expressão de diferentes genes associados com a aquisição de infectividade e virulência pelo parasita. O presente trabalho objetivou, inicialmente, a construção de cassetes de nocaute gênico, obtenção de parasitas nocaute e realização de ensaios funcionais preliminares (viabilidade celular, capacidade de proliferação e diferenciação) para quatro proteínas diferencialmente expressas durante a metaciclogênese de *T. cruzi*, além da pesquisa em bancos de dados e do uso de ferramentas bioinformáticas para o levantamento de domínios estruturais e informações funcionais disponíveis para as mesmas. Duas proteínas, anotadas como hipotéticas conservadas, não possuem nenhuma indicação funcional; outra tem função putativa de diidrolipoamida desidrogenase; e a última proteína, apesar de anotada como hipotética conservada, possui domínios de ligação e reconhecimento de RNA eucariótico. Análises bioinformáticas indicaram que os genes codificadores das proteínas estudadas seriam de cópia única no genoma, tornando-as passíveis de depleção por nocaute gênico. Para todas as proteínas estudadas, foram construídos cassetes de nocaute gênico para os dois alelos (quatro cassetes contendo NEO e quatro cassetes contendo HIGRO). Foram obtidos parasitas nocaute simples viáveis para as quatro proteínas. Entretanto, em função do tempo viável para execução do projeto, as demais etapas do trabalho foram direcionadas para apenas duas proteínas. Para esses dois nocautes simples foram realizados ensaios funcionais preliminares, os quais apontaram redução na capacidade proliferativa e de diferenciação e; obtidos parasitas duplos transfectantes, os quais também se apresentaram viáveis. Objetivando isolar populações clonais para estudos mais aprofundados, realizou-se o *Cell Sorting* em placas de 96 poços. Para uma das proteínas foram obtidas 61 colônias, enquanto para a segunda foram obtidas 72 colônias. Experimentos iniciais para confirmação do duplo nocaute dos genes de interesse demonstraram que, para a primeira proteína, apesar de confirmada a inserção dos cassetes de nocaute para os dois alelos, houve a amplificação da CDS do gene de interesse e assim, não se verificou a obtenção de nocaute duplo. Já para a segunda proteína foram observadas cinco colônias sem amplificação da CDS do gene de interesse e, assim, acredita-se que houve a obtenção de parasitas duplo nocaute. Em síntese, o presente trabalho resultou na produção de cassetes para nocaute gênico de quatro proteínas possivelmente envolvidas no processo de metaciclogênese de *T. cruzi*, bem como de cepas nocaute do parasita para duas delas, as quais apresentaram diminuição na capacidade de proliferação e de diferenciação. Essas cepas são de grande importância para a caracterização funcional das proteínas.

Palavras-chave: *Trypanosoma cruzi*. Metaciclogênese. Nocaute Gênico. Proteínas. Ensaios Funcionais.

ABSTRACT

Chagas disease, caused by the flagellate protozoan *Trypanosoma cruzi*, occurs in Central and South America and is estimated to affect 16 to 18 million people, accounting for 21 million deaths each year. To date there is no vaccine or effective medicine for its chronic phase. *T. cruzi* has a complex life cycle comprising an invertebrate and a mammal host, including humans. It is in the final portion of the digestive tract of the insect that the metacyclogenesis occurs, with the differentiation of *T. cruzi* from a replicative and non - infective form (epimastigote) to an infective and non - replicative form (metacyclic trypomastigote). This process involves changes in the expression of different genes associated with the acquisition of infectivity and virulence by the parasite. The present study aimed to, initially, obtaining knockout parasites and performing preliminary functional assays (cell viability, proliferation and differentiation capacity) to evaluate the effects of the suppression of four proteins differentially expressed during the metacyclogenesis of *T. cruzi*, in addition to searching in databases and using bioinformatic tools to retrieve structural domains and functional information available for them. Two of the proteins, annotated as hypothetical conserved, have no functional information available; one have a putative function of dihydrolipoamida dehydrogenase; and the last protein, although annotated as hypothetical conserved, bears domains for recognition and binding of eukaryotic RNA. Bioinformatic analyses indicated that the genes encoding for all the proteins studied possess a single copy in the genome, making them susceptible to depletion by gene knockout. For all four proteins studied, gene knockout cassettes were constructed for both alleles (four cassettes containing NEO - four cassettes containing HYGRO) and viable single knockout parasites were obtained. However, as a function due to time for execution of the project, the remaining steps of the work were directed to only two proteins. For these two proteins, preliminary functional tests were performed, which showed reduction in the proliferative and differentiation capacities, and double knockout parasites were obtained, which also showed to be viable. In order to isolate clonal populations for further study, we performed the *Cell Sorting* of the double knockout populations in 96 well plates. For one of the proteins, 61 colonies were obtained; while for the second 72 colonies were obtained. Initial experiments to confirm the double knockout of these genes showed that, for the first protein, although the insertion of the knockout cassettes for the two alleles was confirmed, there was also amplification of the CDS of the gene of interest, indicating that the double knockout wasn't obtained. On the other hand, for the second protein five colonies without amplification of CDS of the gene of interest were observed, and therefore, indicating that the double knockout parasites were successfully obtained. In summary, this work resulted in the production of cassettes for gene knockout of four proteins possibly involved in the process of *T. cruzi* metacyclogenesis, as well as knockout strains of the parasite for two of them, which showed a decrease in the proliferation and differentiation abilities. These strains are of great importance for its functional characterization.

Keywords: *Trypanosoma cruzi*. Metacyclogenesis. Gene knockout. Proteins. Functional assays.

LISTA DE FIGURAS

FIGURA 1 - CICLO DE VIDA DO <i>TRYPANOSOMA CRUZI</i>	18
FIGURA 2 - PROTEÔMICA QUANTITATIVA DA METACICLOGÊNESE DE <i>T. CRUZI</i>	20
FIGURA 3 - PROTEÍNAS REGULADAS NAS FASES INICIAIS DA METACICLOGÊNESE DE <i>TRYPANOSOMA CRUZI</i>	21
FIGURA 4 - ESTRUTURA DO CASSETE DE NOCAUTE GÊNICO.....	30
FIGURA 5 - ALINHAMENTO DOS GENES AGRUPADOS EM CADA SUPERGENE.	42
FIGURA 6 - NÚMERO DE CÓPIAS DOS GENES NO GENOMA DE <i>T. CRUZI</i>	44
FIGURA 7- REPRESENTAÇÃO DA HOMOLOGIA DOS GENES COM A INDICAÇÃO DE SEUS GENES PARÁLOGOS E ORTÓLOGOS.....	47
FIGURA 8 - EXPRESSÃO DA PROTEÍNA SG1321 DURANTE METACICLOGÊNESE DE <i>T. CRUZI</i>	49
FIGURA 9 - EXPRESSÃO DA PROTEÍNA SG2809 DURANTE METACICLOGÊNESE DE <i>T. CRUZI</i>	50
FIGURA 10 - EXPRESSÃO DA PROTEÍNA SG4727 DURANTE METACICLOGÊNESE DE <i>T. CRUZI</i>	51
FIGURA 11- EXPRESSÃO DA PROTEÍNA SG5427 DURANTE METACICLOGÊNESE DE <i>T. CRUZI</i>	52
FIGURA 12 - IDENTIFICAÇÃO DE UM PEPTÍDEO SINAL PARA A PROTEÍNA SG1321	52
FIGURA 13 - ESTRUTURA TÍPICA DO PEPTÍDEO SINAL	54
FIGURA 14 - DOMÍNIOS DESCRITOS PARA A PROTEÍNA SG2809.....	57
FIGURA 15- DOMÍNIOS DESCRITOS PARA A PROTEÍNA SG4727	61
FIGURA 16- <i>POLIMERASE CHAIN REACTION</i> (PCR).....	65
FIGURA 17- AMPLIFICAÇÃO DOS GENES DE RESISTÊNCIA À NEO E À HIGRO	66
FIGURA 18 - AMPLIFICAÇÃO DAS IRS DOS GENES: SG1321, SG2809, SG4727 E SG5427	67
FIGURA 19- CASSETE PARA NOCAUTE GÊNICO APÓS SER FUSIONADO	68
FIGURA 20 - ESQUEMA DA RECOMBINAÇÃO HOMÓLOGA ENTRE CASSETE DE NOCAUTE GÊNICO E GENOMA DE <i>T. CRUZI</i>	69
FIGURA 21- OBTENÇÃO DOS CASSETES DE NOCAUTE GÊNICO CONTENDO OS GENES DE RESISTÊNCIA À NEO E À HIGRO PARA OS GENES SG1321, SG2809, SG4727 E SG5427	70
FIGURA 22- AMPLIFICAÇÃO DOS CASSETES DE NOCAUTE GÊNICO CONTENDO O GENE DE RESISTÊNCIA À NEO PARA OS GENES SG1321, SG2809, SG4727 E SG5427.....	72
FIGURA 23 - AMPLIFICAÇÃO DOS CASSETES DE NOCAUTE GÊNICO CONTENDO O GENE DE RESISTÊNCIA À HIGRO PARA OS GENES SG1321 (A) E SG2809.....	73
FIGURA 24 - CURVA DE CRESCIMENTO DOS PARASITAS SIMPLES NOCAUTE PARA OS GENES SG1321 E SG2809	76
FIGURA 25 - AVALIAÇÃO DA CAPACIDADE DE DIFERENCIAÇÃO DOS PARASITAS SIMPLES NOCAUTE PARA OS GENES SG1321 E SG2809	77

FIGURA 26 - AMPLIFICAÇÃO DOS CASSETES DE RESISTÊNCIA EM PARASITAS DKO PARA OS GENES SG1321 E SG2809.....	79
FIGURA 27 - AMPLIFICAÇÃO DOS GENES DE RESISTÊNCIA EM PARASITAS DKO PARA OS GENES SG1321 E SG2809.....	80
FIGURA 28- CONFIRMAÇÃO DA INTEGRIDADE DOS CASSETES DE RESISTÊNCIA TRANSFECTADOS EM PARASITAS DKO PARA OS GENES SG1321 E SG2809.....	81
FIGURA 29 - AMPLIFICAÇÃO DA CDS DOS GENES ALVO EM POPULAÇÕES CLONAIAS DE PARASITAS DKO PARA OS GENES SG1321 (A) E SG2809 (B)	82
FIGURA SUPLEMENTAR 1 - USO DE FERRAMENTAS BIOINFORMÁTICAS SOBRE INFORMAÇÕES REFERENTES ÀS QUATRO PROTEÍNAS SELECIONADAS.....	95
FIGURA SUPLEMENTAR 2 - CONSTRUÇÃO E AMPLIFICAÇÃO DE CASSETE PARA NOCAUTE GÊNICO.....	95
FIGURA SUPLEMENTAR 3 - OBTENÇÃO DE PARASITAS SIMPLES NOCAUTE. .	96
FIGURA SUPLEMENTAR 4 - OBTENÇÃO DE PARASITAS DUPLO NOCAUTE.....	96
FIGURA SUPLEMENTAR 5 - REALIZAÇÃO DE CURVAS DE CRESCIMENTO DOS PARASITAS SIMPLES NOCAUTE.....	97
FIGURA SUPLEMENTAR 6 - INDUÇÃO DA METACICLOGÊNESE <i>IN VITRO</i>	97

LISTA DE TABELAS

TABELA 1- FAMÍLIAS DE SUPERGENES (SGS) COM SEUS GENES E PROTEÍNAS CORRESPONDENTES.....	39
TABELA SUPLEMENTAR 1 - <i>PRIMERS</i> PARA AMPLIFICAÇÃO DAS IRS À MONTANTE (5') E À JUSANTE (3') DAS SEQUÊNCIAS CODIFICADORAS DAS QUATRO PROTEÍNAS USADOS PARA CONSTRUÇÃO DOS CASSETES DE NOCAUTE GÊNICO E CONFIRMAÇÃO DOS PARASITAS DUPLO NOCAUTE.....	98
TABELA SUPLEMENTAR 2 - <i>PRIMERS</i> DAS SEQUÊNCIAS CODIFICADORAS (CDS) PARA CONFIRMAÇÃO DOS PARASITAS DUPLO NOCAUTE.	99

LISTA DE ABREVIATURAS

cAMP – Monofosfato cíclico de adenosina
CDS – Sequência Codificadora (*Coding Sequence*)
DNA – Ácido Desoxirribonucléico
dKO – Parasita transfectante duplo
dNTPs – Desoxirribonucleotídeos Fosfatados
EDTA – Ácido Etileno Diamino Tetracético
F – *Primer Forward*
G418 – Geneticina
HIGRO – Higromicina fosfotransferase
IBMP – Instituto de Biologia Molecular do Paraná
ICC – Instituto Carlos Chagas
IOC – Instituto Oswaldo Cruz
IRS – Regiões Intergênicas (*Intergenic Regions*)
KO – Parasita transfectante simples
LIT – Liver Infusion Tryptose
LITB – Meio LIT acrescido de extrato de levedura
M.O. – Microscopia Óptica
NEO – Neomicina fosfotransferase
pb – Pares de Bases
PBS – Tampão Fosfato Salino
PCR – Reação em Cadeia da Polimerase (*Polymerase Chain Reaction*)
pHYG2 – Plasmídeo contendo gene de resistência para HIGRO
pNEO2 – Plasmídeo contendo gene de resistência para NEO
R – *Primer Reverse*
RPM – Rotações Por Minuto
Taq – Taq DNA Polimerase
TAU – Triatomine Artificial Urine
TAU3AAG – Triatomine Artificial Urine supplemented with L-proline, L-glutamate, L-aspartate
TBE – Tris/Borato/EDTA
Tris – (Tris)Hidroximetil Aminometano
WT – Wild Type (Tipo Selvagem)

SUMÁRIO

1 INTRODUÇÃO	14
1.1 A DOENÇA DE CHAGAS	14
1.2 CARACTERÍSTICAS CELULARES DE <i>T. CRUZI</i>	15
1.3 METACICLOGÊNESE DE <i>T. CRUZI</i>	18
1.4 PROTEÔMICA PARA O ESTUDO DA REGULAÇÃO DA EXPRESSÃO GÊNICA	19
1.5 REGULAÇÃO DA EXPRESSÃO GÊNICA DURANTE A METACICLOGÊNESE DE <i>T. CRUZI</i>	20
2 OBJETIVOS	22
2.1 OBJETIVO GERAL:	22
2.2 OBJETIVOS ESPECÍFICOS:	22
3 METODOLOGIA	25
3.1 LOCAL DE REALIZAÇÃO DA PESQUISA	25
3.2 MEIOS DE CULTURA E TAMPÕES	25
3.3 META 1: FERRAMENTAS E ANÁLISES BIOINFORMÁTICAS	25
3.4 META 2: CONSTRUÇÃO E AMPLIFICAÇÃO DOS CASSETES DE NOCAUTE	26
3.4.1 Desenho e obtenção dos <i>primers</i> para nocaute	26
3.4.2 Amplificação dos genes de resistência	27
3.4.3 Amplificação das IRS flanqueantes aos genes alvo	28
3.4.4 Purificação dos produtos de PCR	28
3.4.5 Obtenção dos cassetes de nocaute gênico	29
3.4.6 Amplificação dos cassetes de nocaute gênico	30
3.4.7 Confirmação da especificidade e tamanho dos produtos de PCR	31
3.4.8 Quantificação dos cassetes de nocaute gênico	31
3.4.9 Concentração dos cassetes de nocaute gênico	31
3.5 META 3: OBTENÇÃO DE PARASITAS NOCAUTE	32
3.5.1 Cultivo dos parasitas	32
3.5.2 Transfecção dos cassetes de nocaute gênico	32
3.5.3 Seleção de parasitas simples e duplo nocaute	33
3.6 META 4: AVALIAÇÃO DOS EFEITOS DO NOCAUTE GÊNICO	34
3.6.1 Análise da viabilidade celular	34
3.6.2 Análise da capacidade de proliferação	34
3.6.3 Análise da capacidade de diferenciação	35
3.7 META 5: CONFIRMAÇÃO DA OBTENÇÃO DE PARASITAS NOCAUTE	36
3.7.1 Isolamento de populações clonais através de <i>Cell Sorting</i>	36
3.7.2 Extração de DNA genômico	36
3.7.3 Desenho e obtenção dos <i>primers</i> para confirmação dos nocautes	37
3.7.4 PCR de confirmação dos nocautes	37
4 RESULTADOS E DISCUSSÃO	38
4.1 META 1: LEVANTAMENTO DE DADOS PARA CARACTERIZAÇÃO DAS PROTEÍNAS ESTUDADAS	38
4.2 META 2: CONSTRUÇÃO E AMPLIFICAÇÃO DOS CASSETES DE NOCAUTE	64
4.2.1 Amplificação dos genes de resistência	65
4.2.2 Amplificação das IRS flanqueantes aos genes alvo	67
4.2.3 Obtenção dos cassetes de nocaute gênico	68
4.2.4 Amplificação dos cassetes de nocaute gênico	71
4.3 META 3: OBTENÇÃO DOS PARASITAS NOCAUTE SIMPLES E DUPLO	74
4.4 META 4: AVALIAÇÃO DOS EFEITOS DO NOCAUTE GÊNICO	75
4.4.1 Análise da viabilidade celular dos parasitas nocaute	75
4.4.2 Análise da capacidade proliferativa dos parasitas simples nocaute	75
4.4.3 Análise da capacidade de diferenciação dos parasitas simples nocaute	77

4.5 META 5: CONFIRMAÇÃO DA OBTENÇÃO DE PARASITAS NOCAUTE	78
4.5.1 Isolamento de populações clonais através de <i>Cell Sorting</i>	79
4.5.2 PCRs de confirmação dos nocautes	79
5 CONCLUSÃO	84
6 PERSPECTIVAS	87
REFERÊNCIAS	89
APÊNDICES	94

1 INTRODUÇÃO

1.1 A DOENÇA DE CHAGAS

As principais infecções encontradas entre as populações pobres dos continentes africano, asiático e americano são as doenças tropicais negligenciadas. Entre essas enfermidades está a Doença de Chagas, descoberta em 1909 pelo médico e pesquisador Carlos Chagas. Chagas descreveu o *Trypanosoma cruzi*, protozoário flagelado, como o agente etiológico da doença, assim como a sua sintomatologia e o ciclo do parasita no vetor e hospedeiros mamíferos (LIMA *et al.*, 2013). Hoje, em função dos constantes deslocamentos populacionais a doença também ocorre em regiões não endêmicas, tornando importante o desenvolvimento de políticas públicas e investimentos em pesquisa visando o controle da doença. A degradação ambiental tem contribuído para a domicialização do inseto vetor, favorecendo a transmissão da Doença de Chagas.

Segundo dados provenientes da Organização Mundial da Saúde, a Doença de Chagas ocorre principalmente na América Central e do Sul, onde é endêmica em 18 países. Estima-se que 16 a 18 milhões de pessoas estão acometidas por esta enfermidade, a qual é responsável por 21 mil mortes a cada ano. Estima-se que em 2010 um contingente de 10 milhões de pessoas estavam infectadas, e mais de 25 milhões estavam em risco de contrair a doença (WHO, 2012).

A principal forma de transmissão do *T. cruzi* se dá pelas fezes contaminadas do vetor triatomíneo. Normalmente, o inseto popularmente chamado de barbeiro, defeca durante o repasto sanguíneo. A picada provoca coceira e, o indivíduo ao se coçar faz com que o parasita entre na corrente sanguínea pelo local da picada ou; penetre em mucosas como olhos, nariz e boca. Outro mecanismo de transmissão, responsável por 10% dos casos de contaminação, ocorre através da transfusão de sangue. Há ainda a transmissão via congênita, responsável por 5.000 a 18.000 novos casos por ano. Na região da Amazônia têm sido relatados casos de transmissão através de alimentos contaminados, como carne, caldo de cana e açaí. Outro mecanismo de transmissão inclui acidentes de laboratório através da manipulação direta do parasita e de organismos contaminados com o mesmo. E, por

fim, há a transmissão através de transplante de órgãos, como rim, coração e medula óssea (PRATA, 2001).

A Doença de Chagas apresenta duas fases distintas: aguda e crônica. A fase aguda dura de 6 a 12 semanas e a maioria dos casos são assintomáticos. Quando há sintomas, esses incluem: inflamação no local da inoculação (chagoma ou sinal de porta de entrada), edema palpebral unilateral, febre, dor de cabeça, dores musculares, anorexia, vômitos, diarreia, sonolência, apatia e convulsões. Alguns casos agudos podem levar à morte. Posterior à fase aguda, manifesta-se a fase crônica. Essa pode ser indeterminada, quando não há sintomas da enfermidade ou; pode causar sérios danos e até mesmo a morte. Passados de 10 a 30 anos do contágio, as manifestações clínicas típicas da fase crônica são o acometimento de órgãos como coração, esôfago, cólon ou uma combinação desses. A disfunção gastrointestinal desenvolve-se em cerca de 10 a 15% dos pacientes cronicamente infectados. A forma cardíaca é a mais grave e desenvolve-se em cerca de 20 a 30% dos pacientes na fase crônica. A forma cardíaca provoca aneurismas, insuficiência cardíaca, tromboembolismo, e, até mesmo a morte (RASSI JÚNIOR; RASSI; NETO, 2010).

Referente ao tratamento clínico da doença, os medicamentos são eficazes para a fase aguda, sendo utilizados normalmente o Nifurtimox e o Benznidazol. Ambas as drogas provocam severos efeitos colaterais. A maioria dos diagnósticos ocorre na fase crônica e, nessa fase, a cura ocorre apenas em alguns casos, sob tratamento a longo prazo (TALICE *et al.*, 2007).

1.2 CARACTERÍSTICAS CELULARES DE *T. CRUZI*

O *Trypanosoma cruzi* é um protozoário flagelado, da ordem Kinetoplastida, e da família Trypanosomatidae. É um organismo unicelular, eucarionte, com algumas características bastante peculiares (MARTINS *et al.*, 2012).

A membrana plasmática do *T. cruzi* é composta por uma bicamada lipídica com várias proteínas imersas. Além delas, são encontrados fosfolipídios (fosfatidilcolina, fosfatidiletanolamina, fosfatidilinositol, esfingomiéline, fosfatidilserina, cardiolipina, ácido fosfatídico, fosfato de fosfatidilinositol). Na superfície celular há

ainda o glicocálix, importante em processos de sinalização celular (MARTINS *et al.*, 2012).

Abaixo da membrana plasmática encontra-se uma camada de microtúbulos, chamados de microtúbulos subpeliculares. Esta associação entre membrana celular e microtúbulos deve ser a responsável pela rigidez celular e pelas dificuldades em romper a célula através de meios mecânicos. A quantidade de microtúbulos é menor nas regiões anterior e posterior do parasita em comparação com a região entre núcleo e cinetoplasto, região onde está o Complexo de Golgi. Em alguns tripanosomatídeos, há pequenos microfilamentos que conectam os microtúbulos subpeliculares entre si e com a membrana plasmática. Há evidências que a inibição dos microtúbulos de actina induz mudanças na morfologia dos tripomastigotas sanguíneos e inibição do seu movimento (SOUZA, 2002; MARTINS *et al.*, 2012). Assim, parece que os microtúbulos exercem função importante na manutenção da forma do parasita e na sua capacidade de movimentação.

O citoplasma do *T. cruzi* apresenta diferentes organelas, descritas a seguir.

O flagelo é formado por nove pares de microtúbulos e um par central, emergindo na célula da bolsa flagelar, que consiste de uma invaginação na junção entre as membranas plasmática e flagelar. Seu tamanho e posição variam ao longo do desenvolvimento. Quanto ao tamanho, o flagelo pode atingir de 1 μm a 20 μm de comprimento. Em relação à posição, normalmente o flagelo emerge da parte central da células, mas em epimastigotas e em tripomastigostas a sua posição é lateral, sendo esta junção considerada do tipo desmossomo (SOUZA, 2002). A estrutura paraflagelar é uma depressão localizada na região anterior da célula onde o flagelo emerge, envolvida em atividades endocíticas e exocíticas. Em *T. brucei* a estrutura paraflagelar está envolvida em muitos processos celulares, como manutenção da polaridade, divisão celular, evasão do sistema imune e tráfego de proteínas. O citóstomo, invaginação da membrana celular localizada próxima à bolsa flagelar, está envolvido com a absorção de nutrientes. Os reservossomos geralmente estão na parte posterior da célula e desaparecem na metaciclogênese. Eles acumulam proteínas e proteases ingeridas que chegam dentro de vesículas endocíticas (MARTINS *et al.*, 2012). Os glicossomos compartilham algumas funções metabólicas dos peroxissomos de outros organismos, como a defesa contra espécies reativas de oxigênio e participação na β -oxidação de ácidos graxos. Além disso, os glicossomos armazenam enzimas da via glicolítica. O conteúdo dessa estrutura varia entre as

diferentes formas do ciclo de vida do parasita. Parece que, o armazenamento das enzimas da via glicolítica são importantes na adaptação do organismo às diferentes condições nutricionais durante a sua diferenciação e, verifica-se durante o ciclo celular do parasita, a degradação de glicosomos e a síntese de novas populações com um teor enzimático diferente. Estes parecem ser fatores importantes no sucesso adaptativo dos kinetoplastídeos em parasitar várias espécies de hospedeiros (MICHELS; BRINGAUD; HERMAN; HANNAERT, 2006).

Os acidocalcissomos são organelas ácidas. Elas armazenam cálcio e apresentam uma alta concentração de fósforo na forma de pirofosfato (PPi) e polifosfato (poli P). A acidificação dessas organelas é devido à presença de canais iônicos e bombas de prótons. Também estão presentes dentro da organela enzimas de degradação e síntese de PPi e poli P. Assim, os acidocalcissomos tem como função o armazenamento de cálcio e fósforo, além de participar do metabolismo de PPi e Pi, homeostase do cálcio, manutenção do pH intracelular e osmorregulação. Há evidências da importância de Ca^{2+} em processos de sinalização celular para invasão de células hospedeiras pelo parasita, possuindo dessa forma envolvimento na aquisição de infectividade (Docampo; MORENO, 2011).

Apresenta uma única mitocôndria tubular, com paredes e cristas irregulares, delimitada por uma dupla membrana. A mitocôndria possui uma íntima relação com o cinetoplasto em função da continuidade das membranas de ambas as organelas. O cinetoplasto é uma estrutura peculiar da ordem Kinetoplastida. Trata-se de uma estrutura próxima à base do flagelo, que contém altas concentrações de DNA extranuclear, conhecido por kDNA (*Kinetoplast DNA*). Trata-se de uma rede complexa com várias moléculas concatenadas de DNA circular, formada por algumas dezenas de maxicírculos e milhares de minicírculos. Os maxicírculos contém genes para rRNAs e para proteínas da cadeia respiratória; os minicírculos codificam RNAs-guia, que atuam sobre os transcritos dos maxicírculos (MARTINS *et al.*, 2012; SOUZA, 2002).

Atualmente são reconhecidas “seis unidades discretas de anotação” (DTUs) de *T. cruzi*, identificadas como TcI-TcVI, e tem sido proposto um sétimo grupo, Tcbat, baseados nas características genéticas de cada cepa. Há variabilidade entre as DTUs em termos de virulência, infectividade, progressão da doença, susceptibilidade às drogas, distribuição geográfica, ciclos de transmissão, assim como espécies hospedeiras do parasita (PEREZ; LYMBERY; THOMPSON, 2014).

1.3 METACICLOGÊNESE DE *T. CRUZI*

O *T. cruzi* apresenta um ciclo de vida complexo, envolvendo hospedeiros mamíferos e insetos, e quatro formas morfológicas bem definidas: amastigota, epimastigota, tripomastigota sanguíneo e tripomastigota metacíclico (Figura 1).

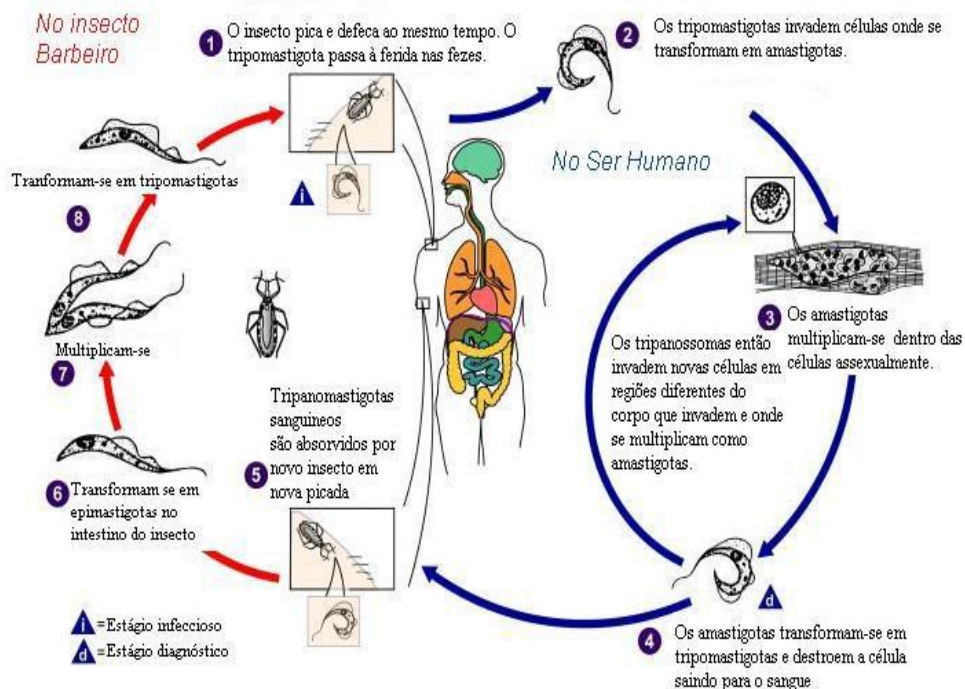


Figura 1 - Ciclo de vida do *Trypanosoma cruzi* (<http://www.dpd.cdc.gov/dpdx/HTML/TrypanosomiasisAmerican.htm>).

Dentro desse ciclo de vida, a metaciclogênese ocorre no intestino do inseto vetor e consistiu o processo através do qual o *T. cruzi* se diferencia de uma forma não infectiva e replicativa (epimastigotas) para uma forma infectiva e não replicativa (tripomastigota metacíclico), o que envolve alterações na expressão de diferentes genes associados com a aquisição de infectividade e virulência pelo parasita. Nas porções finais do intestino do vetor triatomíneo, formas replicativas epimastigotas sofrem um estresse nutricional, aderindo à parede do intestino do vetor e se diferenciando em formas infectivas tripomastigotas metacíclicas. A contaminação de um hospedeiro vertebrado, incluindo o homem, acontece no momento do repasto sanguíneo do inseto. Os parasitas passam à corrente sanguínea pelo local da picada

ou através de mucosas, infectando diferentes tipos celulares. Dentro das células do hospedeiro, os parasitas se diferenciam em formas replicativas amastigotas. Após muitos períodos de replicação celular, as formas amastigotas passam a se diferenciar nas formas infectivas tripomastigotas sanguíneas, capazes de infectar outras células do hospedeiro e novamente se diferenciar em amastigotas reiniciando o ciclo dentro da célula do hospedeiro ou; capaz de infectar outro vetor triatomíneo. No estômago do inseto infectado inicia-se a transformação para formas epimastigotas, seguida da metaciclogênese com a diferenciação em formas metacíclicas (TALICE *et al.*, 2007).

A metaciclogênese pode ser induzida *in vitro* com a incubação de formas epimastigotas em condições quimicamente semelhantes à urina do inseto vetor. Brevemente, a indução inicia-se após um estresse nutricional, onde as formas epimastigotas são incubadas em meio TAU (*Triatomine Artificial Urine*). Após o estresse nutricional, os parasitas são transferidos para o meio de diferenciação (TAU3AAG). Esse meio induz à adesão do parasita ao substrato, conforme o observado no trato digestivo do hospedeiro invertebrado, ocorrendo assim a diferenciação de formas epimastigotas em tripomastigotas metacíclicas (revisto por FERREIRA *et al.*, 2008).

1.4 PROTEÔMICA PARA O ESTUDO DA REGULAÇÃO DA EXPRESSÃO GÊNICA

A partir da publicação do genoma de *T. cruzi* em 2005 (LESCURE *et al.*, 2010), vários estudos proteômicos passaram a ser desenvolvidos visando uma compreensão mais apurada em relação aos produtos codificados pelos genes do parasita, assim como a sua expressão e regulação em diferentes contextos biológicos. Segundo Mattick (2003), acredita-se que as proteínas exerçam grande influência no entendimento de processos complexos e dinâmicos relacionados aos genes. A análise proteômica é definida de acordo com Pennington & Dunn (2001); Campos *et al.* (2006), como um conjunto de metodologias analíticas empregadas para caracterizar (quali e quantitativamente) as proteínas expressas por um genoma, em conjunto denominadas proteoma. A proteômica permite detectar e localizar modificações pós-traducionais; caracterizar interações proteína-ligante; e

caracterizar proteínas íntegras. É uma área do conhecimento que agrega a química, a biologia, e a informática, utilizando como ferramentas para interpretação dos dados a espectrometria de massa (MS) e a bioinformática.

Assim, a análise proteômica tem se mostrado um instrumento essencial nos estudos de expressão gênica em nível de proteínas, fornecendo informações complementares àquelas obtidas nas análises de expressão em nível transcricional.

1.5 REGULAÇÃO DA EXPRESSÃO GÊNICA DURANTE A METACICLOGÊNESE DE *T. CRUZI*

Recentemente, pesquisadores do Instituto Carlos Chagas (ICC), unidade regional da Fundação Oswaldo Cruz (FIOCRUZ) Paraná, realizaram um estudo de proteômica quantitativa para o estudo de cinco diferentes formas de diferenciação do *T. cruzi* durante a metaciclogênese (Figura 2).

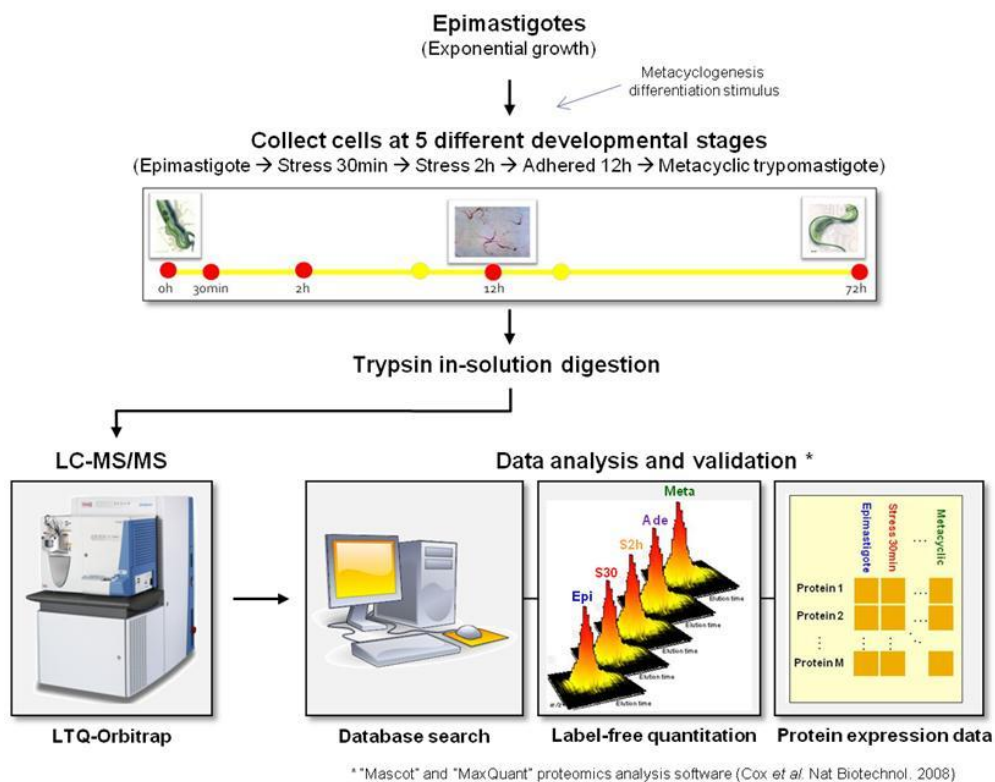


Figura 2 - Proteômica quantitativa da metaciclogênese de *T. cruzi* (de GODOY *et al.*, 2012).

Neste estudo foi realizada a indução da metaciclogênese *in vitro*. A identificação e quantificação proteica foi feita através do equipamento LTQ-Orbitrap, com a aplicação de espectrometria de massa (MS) de alta precisão. As amostras foram lisadas para extração de proteínas e digeridas em solução com tripsina. Após, as amostras foram analisadas por LC-MS/MS no LTQ-Orbitrap. A identificação e quantificação dos peptídeos foi feita com auxílio do programa MaxQuant. Foi realizada a análise do perfil proteômico de parasitas em suas formas iniciais de diferenciação, formas intermediárias e formas tardias, com a identificação de cerca de três mil proteínas ao longo de todo o processo. No referido trabalho os dados foram coletados após a indução nos tempos específicos de 30 minutos, duas horas e doze horas (Figura 3). Nessas fases iniciais da metaciclogênese foram identificadas oito proteínas diferencialmente expressas, com a identificação de três proteínas hipotéticas e conservadas em estresse trinta minutos e estresse duas horas, sendo elas reguladas positivamente. Em aderido doze horas foram identificadas cinco proteínas, sendo quatro delas positivamente reguladas e uma com regulação negativa. Dessas cinco proteínas, duas são hipotéticas e conservadas, uma delas é uma glicoproteína de adesão flagelar, uma proteína é um transportador folato/pteridina e, uma dihidrolipoamida desidrogenase.

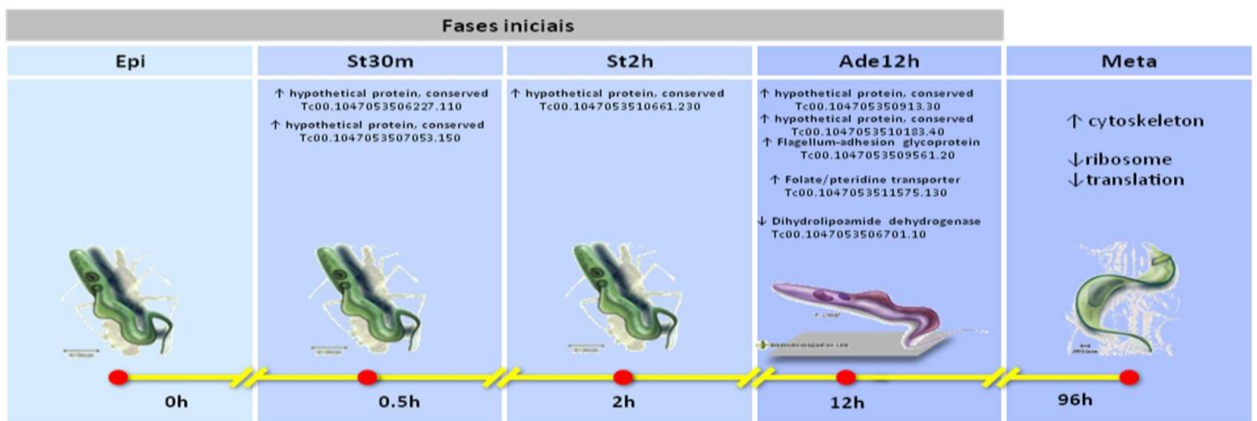


Figura 3 - Proteínas reguladas nas fases iniciais da metaciclogênese de *Trypanosoma cruzi*. Epi = Epimastigotas; St30m = Parasitas após 30 minutos de indução (estresse nutricional); St2h = Parasitas após 2 horas de indução (estresse nutricional); Ade12h = Parasitas aderidos após 12 horas de indução (estresse nutricional); Meta = Tripomastigotas metacíclicos (de GODOY *et al.*, 2012).

Estas proteínas diferencialmente expressas são candidatas relevantes para o estudo funcional, uma vez que podem trazer novas informações sobre a biologia do *T. cruzi*, ao serem alvos potencialmente envolvidos na diferenciação e aquisição de infectividade pelo parasita. Desta forma, podem ser utilizadas para o desenvolvimento de novas drogas, estratégias terapêuticas e indicadores de prognóstico clínico da Doença de Chagas.

2 OBJETIVOS

2.1 OBJETIVO GERAL:

O trabalho objetiva caracterizar efeitos de nocaute gênico de quatro proteínas diferencialmente expressas nas fases iniciais da metaciclologênese de *Trypanosoma cruzi*, em termos de viabilidade, capacidade proliferativa e de diferenciação celular:

Proteína 1 (SG1321) - Tc00.1047053506227.110 (*Hypothetical protein, conserved*);

Proteína 2 (SG2809) - Tc00.1047053510661.230 (*Hypothetical protein, conserved*);

Proteína 3 (SG4727) - Tc00.1047053506701.10 (*dihydrolipoamide dehydrogenase, putative*);

Proteína 4 (SG5427) - Tc00.1047053510183.40 (*Hypothetical protein, conserved*).

2.2 OBJETIVOS ESPECÍFICOS:

META 1: Realizar, através do uso de ferramentas bioinformáticas e pesquisa em bancos de dados, o levantamento de dados para caracterização das quatro proteínas, tais como:

- Alinhamento dos genes;
- Função, peso e comprimento molecular;
- Localização cromossomal;

- Número de cópias da CDS preditas no genoma de *T. cruzi*;
- Expressão proteica durante a metaciclogênese *in vitro* de *T. cruzi*;
- Identificação de domínios estruturais e funcionais;
- Ontologia Gênica;
- Homologia Gênica;
- Participação em vias metabólicas.

META 2: Construir e amplificar os cassetes de nocaute gênico

- Desenhar os *primers* dos genes de resistência NEO e HIGRO; e das IRS à montante (5') e à jusante (3') da CDS de cada gene selecionado;
- Construir, por PCR de fusão, os cassetes para nocaute de cada gene selecionado, contendo o gene de resistência à NEO;
- Purificar os produtos de PCR;
- Construir, por PCR de fusão, os cassetes para nocaute de cada gene selecionado, contendo o gene de resistência à HIGRO;
- Amplificar, por PCR, os cassetes de nocaute (NEO e HIGRO);
- Quantificar os cassetes de nocaute gênico;
- Concentrar os cassetes de nocaute gênico.

META 3: Obter parasitas nocaute simples e duplo

- Cultivar formas epimastigotas de *T. cruzi*;
- Transfectar os cassetes de nocaute gênico;
- Selecionar parasitas nocaute simples e duplo.

META 4: Confirmar a obtenção de parasitas nocaute

- Obter populações clonais de parasitas nocaute – *Cell Sorting*;
- Extrair DNA das populações clonais;
- Realizar PCR de confirmação dos nocautes.

META 5: Avaliar os efeitos do nocaute gênico

- Analisar a viabilidade celular dos parasitas simples nocaute;
- Analisar a capacidade de proliferação dos parasitas simples nocaute (curva de crescimento);
- Analisar a capacidade de diferenciação dos parasitas simples nocaute (contagem diferencial de epimastigotas e tripomastigotas metacíclicos).

3 METODOLOGIA

3.1 LOCAL DE REALIZAÇÃO DA PESQUISA

A pesquisa foi desenvolvida no Laboratório de Genômica Funcional do Instituto Carlos Chagas, Fundação Oswaldo Cruz – Paraná (ICC/FIOCRUZ-PR), situado no Parque Industrial do Instituto de Tecnologia do Paraná, na Cidade Industrial de Curitiba (CIC).

3.2 MEIOS DE CULTURA E TAMPÕES

Tampão Fosfato: $\text{NaH}_2\text{PO}_4 \cdot \text{H}_2\text{O}$ 0.2 M, Na_2HPO_4 0.2 M;

Tampão PBS (1x): NaCl 137 mM, KCl 2.7 mM, $\text{Na}_2\text{HPO}_4 \cdot 7\text{H}_2\text{O}$ 4.3 mM, KH_2PO_4 1.5 mM;

Meio LITB: Infuso de fígado 0.5%, NaCl 75.3 mM, KCl 5.4 mM, Glicose monohidratada 10 mM, Bacto-Triptose 0.5%, Na_2HPO_4 56.4 mM, Hemina 0.0025%, Soro Fetal Bovino 10%, 15g/L de extrato de levedura;

Meio TAU: NaCl 190 mM, KCl 17 mM, $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$ 2 mM, $\text{MgCl}_2 \cdot 6\text{H}_2\text{O}$ 2 mM, Tampão fosfato 8 mM, NaHCO_3 0.035%;

Meio TAU3AAG: Meio TAU suplementado com: Glicose 10 mM, Ácido L-aspártico 2 mM, Ácido L-glutâmico 50 mM, L-prolina 10 mM.

3.3 META 1: FERRAMENTAS E ANÁLISES BIOINFORMÁTICAS

Ferramentas bioinformáticas foram usadas para o levantamento de diversos dados estruturais e funcionais referentes às quatro proteínas, conforme ilustrado na Figura Suplementar 1 (em apêndice). As informações estruturais e funcionais

referentes às quatro proteínas foram obtidas a partir dos seguintes bancos de dados:

- TriTrypDB (<http://tritrypdb.org/tritrypdb/>);
- TrypanosOmics (<http://omics.icc.fiocruz.br/welcome>);
- Pfam (<http://pfam.sanger.ac.uk/>);
- InterPROScan (<http://www.ebi.ac.uk/interpro/>);
- TDR Targets DATABASE (<http://tdrtargets.org/documentation>).

A classificação dos genes em Supergenes foi obtida utilizando-se a ferramenta TrypanosOmics. As informações referentes à posição nucleotídica, localização cromossomal, peso molecular, comprimento, homologia gênica, participação dos genes em vias metabólicas, ontologia gênica e domínios proteicos foram obtidas utilizando-se as ferramentas TriTrypDB e TDR Targets.

O gráfico de predição do número de cópias gênicas foi obtido utilizando-se a ferramenta TcruziKO, previamente desenvolvido pelo Laboratório de Bioinformática do ICC/FIOCRUZ-PR (LORUSSO, 2013).

As sequências proteicas usadas para alinhamento foram obtidas utilizando-se a ferramenta TriTrypDB. Os alinhamentos de sequência foram realizados utilizando-se a ferramenta ClustalÔmega (<http://www.ebi.ac.uk/Tools/msa/clustalo/>).

3.4 META 2: CONSTRUÇÃO E AMPLIFICAÇÃO DOS CASSETES DE NOCAUTE

A construção e amplificação dos cassetes de nocaute gênico foram desenvolvidas de acordo com algumas metodologias anotadas no desenho experimental ilustrado na Figura Suplementar 2 (em apêndice).

3.4.1 Desenho e obtenção dos *primers* para nocaute

Para a construção dos cassetes de nocaute gênico foram desenhados *primers* para a amplificação dos genes de resistência, amplificação das regiões intergênicas à montante (5') e à jusante (3') das sequências codificadoras de cada proteína alvo e dos cassetes de resistência a antibióticos NEO e HIGRO. Para tanto,

foi utilizado o software de predição de iniciadores para nocaute gênico denominado TcruziKO, previamente desenvolvido pelo Laboratório de Bioinformática do ICC/FIOCRUZ-PR (LORUSSO, 2013), o qual além de predizer a sequência adequada para cada *primer*, adiciona pequenas sequências repetitivas complementares às extremidades dos mesmos, necessárias para a realização da técnica de PCR de fusão. Os *primers* para nocaute foram sintetizados e adquiridos do fornecedor comercial The Midland certified reagent Company Inc. of Midlands, Texas, USA. A identificação e a sequência de cada *primer* estão anotadas na Tabela Suplementar 1 (em apêndice).

3.4.2 Amplificação dos genes de resistência

Os genes de resistência à NEO e à HIGRO, usados nos cassetes de nocaute gênico como marcadores de seleção, foram amplificados a partir dos plasmídeos pNEO2 e pHYG2, respectivamente, cedidos pelo Dr. Stênio Perdigão Fragoso, do Laboratório de Biologia Molecular de Tripanossomatídeos, do Instituto Carlos Chagas. Estes plasmídeos foram previamente construídos a partir de inserções por clonagem clássica no vetor pBluescript, que resultaram em plasmídeos contendo a região codificante de um gene de resistência a antibiótico flanqueado por duas sequências distintas obtidas de regiões não-codificantes do genoma de *T. cruzi*. Estas regiões não-codificantes do genoma de *T. cruzi* estão situadas entre os genes que codificam as proteínas enolase (TcCLB.511529.90) e KAP3 (TcCLB.511529.80), denominado “IR KAP3”, e entre duas cópias do gene que codifica a enzima GAPDH (TcCLB.506943.60 e TcCLB.506943.50), denominado “IR GAPDH”. Para a amplificação dos genes de resistência foi utilizado um *primer* que se anela à extremidade 5’ (IR KAP3) e um *primer* que se anela à extremidade 3’ (IR GAPDH) do cassete, denominados de K7-F e K7-R, respectivamente. As reações de amplificação foram feitas para o volume de 100 µL, utilizando por reação os seguintes reagentes: 15 ng do plasmídeo (DNA-molde), 1 µL da enzima Accuprime™Pfx DNAPolimerase, tampão Accuprime™Pfx Reaction Mix na concentração final de 1x, 2 µL de cada *primer* (na concentração final de 0,5 µM), e volume de água MilliQ suficiente para completar 100 µL de reação. Com exceção do

DNA-molde e dos *primers*, os demais reagentes foram obtidos do fabricante Invitrogen. As reações de PCR para amplificação dos genes de resistência tiveram três etapas: (1) desnaturação a 95°C por 2 minutos; (2) 35 ciclos de amplificação – 95°C por 20 segundos, 55°C por 30 segundos e 68°C por 2 minutos e; (3) extensão a 68°C por 7 minutos.

3.4.3 Amplificação das IRS flanqueantes aos genes alvo

As reações de amplificação das IRS à montante (5') e à jusante (3') das sequências codificadoras dos genes alvo foram preparadas para o volume de 50 µL cada, utilizando por reação os seguintes reagentes: 42 ng de DNA genômico de formas epimastigotas de *T. cruzi*, cepa Dm 28c (DNA-molde), 1,5 µM de cloreto de magnésio, 0,4 µL da Taq DNA Polimerase (0,4 U/ µL de reação) produzida pelo IBMP, tampão da enzima sem adição de magnésio na concentração final de 1x, 0,5 µM de solução contendo dNTPs, *primers* (2 µM cada), e volume de água MilliQ suficiente para completar um volume final de 50 µL. Com exceção do DNA-molde, da enzima Taq IBMP e dos *primers*, os demais reagentes foram obtidos da fabricante Invitrogen. A amplificação das intergênicas referentes às proteínas 1, 2 e 3 foi realizada com as seguintes etapas: (1) desnaturação a 94°C por 3 minutos; (2) 35 ciclos de amplificação – 94°C por 20 segundos, 55°C por 20 segundos e 72°C por 30 segundos e; (3) extensão a 72°C por 5 minutos. O programa para amplificação das intergênicas referentes à proteína 4 teve as seguintes etapas: (1) desnaturação a 94°C por 3 minutos; (2) 35 ciclos de amplificação – 94°C por 20 segundos, 50°C por 20 segundos e 72°C por 30 segundos e; (3) extensão a 72°C por 5 minutos.

3.4.4 Purificação dos produtos de PCR

Para otimizar as reações de PCR de fusão, os produtos de PCR dos genes de resistência e das intergênicas dos genes alvo foram previamente digeridos com a enzima Exonuclease I de *Escherichia coli* (New England Biolab), que degrada

moléculas de DNA fita simples a partir da extremidade 3' livre. A digestão com Exonuclease I foi utilizada como alternativa à purificação tradicional do produto de PCR desejado. O intuito foi remover fragmentos pequenos de DNA sintetizados e os *primers* não-anelados que poderiam interferir negativamente nas reações de amplificação posteriores à amplificação das intergênicas e dos genes de resistência. A reação consistiu em adicionar 1 µL de Exonuclease I (20 U/µL) para cada 10 µL de produto de PCR. Os tubos de digestão foram incubados em ThermoMixer por 37°C durante 1 hora para reação de digestão e, depois, por 80°C durante 30 minutos em ThermoMix para a inativação da enzima.

3.4.5 Obtenção dos cassetes de nocaute gênico

Os cassetes de nocaute gênico em *T. cruzi* foram construídos usando a metodologia de PCR de fusão (KUWAYAMA *et al.*, 2002), a qual consistiu na fusão a partir da complementaridade existente entre as intergênicas amplificadas e o gene de resistência. O desenho dos respectivos *primers* foi feito de modo a permitir a fusão das estruturas. Assim, os cassetes são formados pelas intergênicas do gene alvo flanqueando um gene de resistência. A fusão aconteceu entre a IR5' do gene alvo com IR KAP3 do cassete e entre a IR3' do gene alvo com IR GAPDH do cassete (Figura 4). Para cada um dos quatro genes foi feita a fusão das respectivas intergênicas com NEO e com HIGRO. Em cada reação de fusão foram preparadas reações de 50 µL e utilizados os seguintes reagentes: 25 ng do gene de resistência digerido, 0,5 µL de Taq DNA polimerase do IBMP (0,5 U/ µL de reação), 5 µL do tampão 10x PCR *Buffer* minus Mg (na concentração final de 1x), 1,5 µL de 50 mM de cloreto de magnésio, 4 µL de dNTPs (na concentração final de 0,8 µM), 7,5 µL de cada intergênica digerida, e volume de Água MilliQ necessária para completar o volume de 50 µL de reação. Com exceção dos genes de resistência e da Taq IBMP, os demais reagentes foram obtidos da fabricante Invitrogen. O programa para a reação de fusão teve três etapas: (1) desnaturação a 94°C por 2 minutos; (2) 20 ciclos de amplificação – 94°C por 20 segundos, 50°C por 30 segundos e 72°C por 3 minutos e; (3) extensão a 72°C por 8 minutos. A especificidade e o tamanho das bandas foram verificados em eletroforese de DNA em gel de agarose 1% conforme

já descrito. Foram construídos oito cassetes de nocaute gênico, sendo quatro cassetes de NEO e quatro cassetes de HIGRO.

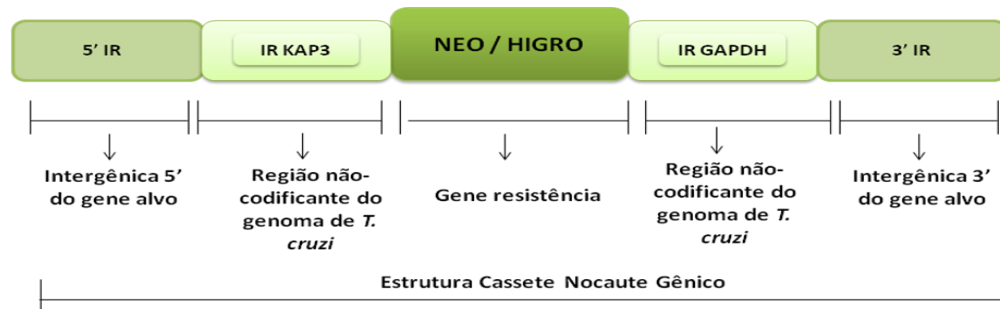


Figura 4 - Estrutura do cassete de nocaute gênico.

3.4.6 Amplificação dos cassetes de nocaute gênico

Após a obtenção dos cassetes de NEO e HIGRO para os quatro genes, foi necessário amplificá-los para conseguir a massa suficiente para a posterior transfecção em *T. cruzi*. Para tal, o produto da PCR de fusão (cassete) foi usado como molde. Foram preparados tubos de 100 μ L e utilizados os seguintes reagentes: 1 μ L do produto de fusão (cassete), 0,5 μ L de Taq IBMP (5u/ μ L), 5 μ L do tampão 10x PCR *Buffer* minus Mg, 4 μ L de dNTPs (10 μ M), 1,5 mM de cloreto de magnésio, 1 μ L de cada *primer*, e volume de Água MilliQ suficiente para completar o volume de 100 μ L de reação. Foram utilizados nesta reação o *primers forward* da IR5' e o *primer reverse* da IR3'. Com exceção do DNA-molde, da enzima Taq IBMP e dos *primers*, os outros reagentes foram obtidos da fabricante Invitrogen. O programa para a amplificação do produto de fusão teve três etapas: (1) desnaturação a 94°C por 2 minutos; (2) 20 ciclos de amplificação – 94°C por 20 segundos, 50°C por 30 segundos e 72°C por 3 minutos e; (3) extensão a 72°C por 8 minutos.

3.4.7 Confirmação da especificidade e tamanho dos produtos de PCR

A confirmação da especificidade e tamanho das regiões amplificadas foi realizada em gel de agarose horizontal 1%. As amostras e o padrão de massa molecular (1 Kb Plus\Gibco - Invitrogen; 1 Kb DNA Ladder\Gibco - Invitrogen; 1 Kb\Promega) foram diluídos em tampão contendo glicerol na concentração 5x e aplicados em gel de agarose, preparado em tampão TBE (Tris/Borato/EDTA) na concentração 0,5x. A corrida eletroforética foi realizada no mesmo tampão, mantida de 90 a 115 volts, durante tempo variável. O DNA foi corado com solução de brometo de etídio 0,5 µg/mL por aproximadamente 15 minutos, lavado em água por 5 minutos e visualizado sob luz ultravioleta em trans-iluminador. O perfil eletroforético foi registrado em um sistema de vídeo documentação.

3.4.8 Quantificação dos cassetes de nocaute gênico

A quantificação dos produtos de PCR foi feita na plataforma Qubit™ Fluorometer (Invitrogen™) utilizando o kit Quant-it™ dsDNA BR Assay. Para a quantificação, foram usados 5 µL de cada amostra, 1 µL do reagente e 194 µL do tampão, sendo preparado tubo mix quando eram quantificadas duas ou mais amostras. Os tubos, específicos para a quantificação, foram mantidos no escuro, à temperatura ambiente, 2 minutos antes da quantificação. Foram quantificados os cassetes de nocaute gênico de NEO e HIGRO, uma vez que os valores de suas massas são importantes para cálculo de massa de cassete utilizada na transfecção em *T. cruzi*.

3.4.9 Concentração dos cassetes de nocaute gênico

Para transfecção, os cassetes de nocaute gênico amplificados foram concentrados em um volume máximo de 50 µL. Essa concentração foi obtida através

de ultrafiltração, feita em colunas Amicon Ultra-0,5 mL 30K, da Millipore®, seguindo instruções do fabricante. A primeira centrifugação foi realizada em 14.000 xg por cinco minutos, sendo descartado o volume acumulado no tubo coletor. A seguir, a coluna foi invertida e colocada em um tubo coletor limpo, sendo realizada a segunda centrifugação a 1.000 xg por 2 minutos. Nesta segunda centrifugação, os cassetes concentrados estavam na porção eluída.

3.5 META 3: OBTENÇÃO DE PARASITAS NOCAUTE

A obtenção de parasitas simples e duplo nocaute foi realizada segundo os desenhos experimentais ilustrados na Figura Suplementar 3 e 4, respectivamente (em apêndice).

3.5.1 Cultivo dos parasitas

Formas epimastigotas de *T. cruzi* cepa Dm 28c, foram mantidas a 28°C em meio LIT. Os parasitas em fase logarítmica de crescimento foram obtidos a partir de repiques a cada três dias, partindo de um inóculo de 1×10^6 parasitas/mL de meio de cultura.

3.5.2 Transfecção dos cassetes de nocaute gênico

A transfecção dos cassetes de nocaute gênico se deu através da técnica de eletroporação, usando o equipamento Gene Pulser® II Apparatus e cubetas estéreis de eletroporação Gene Pulser/Micro Pulser® 0,2 cm, ambos da Bio-Rad. Para obtenção dos parasitas simples nocaute (KO), foi feita a transfecção dos cassetes contendo o gene de resistência à NEO. Para tanto, inicialmente as células foram centrifugadas a 6.000 xg por 5 minutos a 4°C. O sobrenadante foi descartado por

inversão. Adicionou-se 50 mL de PBS (tampão fosfato salino) estéril e as células foram centrifugadas a 6.000 xg por 5 minutos a 4 °C. O sobrenadante foi descartado e as células foram ressuspensas e lavadas em quantidade suficiente de tampão de eletroporação estéril para obter a concentração final de 5×10^8 céls/mL. Para cada gene foram usadas cubetas 0.2 previamente resfriadas a 4°C, onde foram colocados 400 µL da suspensão de parasitas, contendo cerca de 2×10^8 células, e mais 50 µg de cassete de nocaute. As cubetas foram encaixadas no aparelho de eletroporação, com os parâmetros configurados em 450V / 500µF, e dois pulsos seguidos foram dados em cada uma. Após, as células foram imediatamente colocadas em uma garrafinha de 25 cm³ contendo 10 mL de meio LIT+penicilina e levada para estufa de CO₂ a 28°C por 24 horas. Os parasitas devidamente selecionados com NEO foram submetidos à nova transfecção (seguindo o mesmo protocolo acima descrito), mas desta vez usando-se os cassetes de nocaute gênico com HIGRO para obtenção de parasitas duplo nocaute (dKO). Como controle negativo foi feita a eletroporação sem adição de nenhum cassete durante os dois procedimentos: na obtenção de parasitas simples nocaute; e obtenção de parasitas duplo nocaute.

3.5.3 Seleção de parasitas simples e duplo nocaute

Para seleção dos nocautes simples (com o marcador de seleção NEO), o antibiótico G418 foi adicionado à cultura 24 horas após a transfecção, na concentração final de 500 µg/mL. O primeiro repique foi feito quatro dias após a transfecção, com 2,5 mL da cultura em 7,5 mL de meio LIT estéril e com a mesma concentração de droga. Depois, foram feitos repiques semanais. Os parasitas foram deixados na estufa de CO₂ a 28°C até a constatação de que o controle negativo utilizado tivesse morrido, pois parasitas que tivessem internalizado o cassete de nocaute gênico contendo o marcador de seleção sobreviveriam. A seleção dos parasitas duplo nocaute se deu com o cultivo dos mesmos em meio contendo os antibióticos G418 e HIGRO, ambos na concentração final de 500 µg/mL. Foram selecionados os sobreviventes e realizados repiques a partir do quarto dia de transfecção, com a adição de 2,5 mL de cultura, 7,5 mL de meio LIT estéril e ambos os antibióticos para a concentração final de 500 µg/mL. Também foram utilizados

parasitas selvagens como controle negativo para os parasitas duplo nocaute. E, novamente, os parasitas foram deixados na estufa de CO₂ a 28°C até a constatação de que o controle negativo utilizado tivesse morrido. A adição dos antibióticos de seleção aconteceu 24 horas após a transfecção, uma vez que muitas células morrem naturalmente com esta técnica. Assim, este intervalo de tempo entre a eletroporação e a adição do antibiótico é necessário para certeza da morte celular em função da ausência dos genes de resistência.

3.6 META 4: AVALIAÇÃO DOS EFEITOS DO NOCAUTE GÊNICO

A avaliação dos efeitos do nocaute gênico foi realizada a partir da análise da viabilidade celular, da capacidade de proliferação e da capacidade de diferenciação dos parasitas.

3.6.1 Análise da viabilidade celular

A análise da viabilidade celular foi feita visualmente ao microscópio óptico, baseada na morfologia e na capacidade de crescimento dos parasitas. O crescimento celular foi entendido como a capacidade de sobrevivência do *T. cruzi* na presença dos antibióticos de seleção e, como condição necessária para se ter a viabilidade celular.

3.6.2 Análise da capacidade de proliferação

A análise da capacidade de proliferação dos parasitas foi realizada com base na construção de uma curva de crescimento, segundo o desenho experimental ilustrado na Figura Suplementar 5 (em apêndice). Para obter culturas “frescas” e em condições “mais homogêneas”, foi inicialmente feito o repique das células, com a

adição de meio LIT, parasitas simples nocaute e adição de G418. Os parasitas foram mantidos em meio LIT, com densidade celular de 1×10^6 céls/mL a 28°C. A contagem do número de parasitas foi realizada a cada 24 horas, por sete dias consecutivos. Para cada gene em estudo manteve-se uma cultura de parasitas simples nocaute. Como controle utilizou-se uma cultura de formas epimastigotas cepa Dm 28c de *T. cruzi*. A contagem foi efetuada em contador automático de células Beckamn/Coulter e em Câmara de Neubauer, sendo realizada média das leituras.

3.6.3 Análise da capacidade de diferenciação

A análise da capacidade de diferenciação dos parasitas foi realizada com base na indução da metaciclogênese *in vitro*, conforme ilustrado na Figura Suplementar 6 (em apêndice). Para indução da metaciclogênese *in vitro*, 5×10^7 células de epimastigotas foram inicialmente centrifugadas em tubos Falcon a 3.000 xg por 5 minutos a 10°C. Para indução do estresse, descartou-se o sobrenadante e as células foram ressuspensas em 100 µL de meio TAU pré-aquecido a 28°C, na concentração final de 5×10^8 células/mL. Procedeu-se à incubação dos tubos a 28°C por 2 horas. Após duas horas de estresse, os parasitas foram ressuspensos e transferidos para garrafas contendo 9,9 mL de meio enriquecido TAU3AAG. As garrafas foram deitadas e incubadas a 28°C para permitir a adesão dos parasitas. Como controle empregou-se culturas de parasitas selvagens cepa Dm 28c de *T. cruzi*. A análise da capacidade de diferenciação foi feita a partir de observações visuais em microscopia óptica, de acordo com alterações na morfologia do parasita. Foi realizada a contagem diferencial das formas epimastigotas e tripomastigotas metacíclicos após 72h e 96h da indução da metaciclogênese. A contagem foi realizada em triplicata e feita média entre elas. Os dados coletados foram utilizados para compor gráficos comparativos. Mantiveram-se três culturas ao longo da metaciclogênese, sendo uma cultura selvagem e uma cultura para cada parasita simples nocaute.

3.7 META 5: CONFIRMAÇÃO DA OBTENÇÃO DE PARASITAS NOCAUTE

3.7.1 Isolamento de populações clonais através de *Cell Sorting*

Visando a obtenção de populações clonais dos parasitas nocaute, adotou-se como metodologia o *Cell Sorting*, utilizando o citômetro BD FACSAria II, que utiliza parâmetros como a densidade celular, a granulometria e o diâmetro das células para selecionar a população para o *Sorting*. Neste estudo utilizou-se uma população celular com alta densidade, além de certa homogeneidade entre a granulometria e diâmetro das células. A partir de uma cultura em fase exponencial de crescimento, entre 5×10^7 e 1×10^8 parasitas foram centrifugados a 3.000 xg por 10 minutos. O *pellet* foi ressuspensão em 2 mL de PBS estéril na concentração 1x. Seguiu-se uma segunda centrifugação a 3.000 xg por 10 minutos. O sobrenadante foi descartado, os parasitas foram ressuspensos em tampão de lavagem estéril, a uma concentração final de 2×10^7 células/mL. Processou-se o *Cell Sorting* de uma célula por poço, em placas de 96 poços contendo 100 μ L de meio LIT com penicilina, ampicilina 10 μ g/mL, gentamicina 10 μ g/mL e os antibióticos de seleção G418 e HIGRO. As culturas foram incubadas a 28°C e 5% CO₂, por duas a três semanas, tempo necessário para a turbidez do meio, indicando crescimento do clone isolado. O crescimento dos parasitas foi acompanhado por microscopia óptica. As culturas com crescimento visível foram preparadas para a realização da PCR de confirmação dos parasitas nocaute.

3.7.2 Extração de DNA genômico

Para a extração de DNA foram realizados dois protocolos. O primeiro foi feito com o KitDNeasy Plant Maxi Kit, da Qiagen®, de acordo com as instruções do fabricante. A extração foi feita a partir de culturas celulares de oito dias, com 5×10^7 células. As células foram centrifugadas a 5000 xg por 5 minutos, ressuspensas em 1 mL de PBS 1x, centrifugadas novamente a 5.000 xg por 5 minutos, ressuspensas

em 400 µL de AP1 (reagente de lise celular). A seguir promoveu-se o choque térmico com as células incubadas a 65°C por 10 minutos, adicionados 130 µL de AP2 (reagente de precipitação), seguido da centrifugação das células a 14.000 rpm por 5 minutos. Como reagente de ligação à coluna foram usados 675 µL de AP3 (*binding reagent*) e como tampão de lavagem 500 µL do *Buffer AW*. Para eluição foram empregados 50 µL do *Buffer AE*. Em algumas etapas adotou-se uma metodologia usada por alguns usuários do nosso laboratório de pesquisa, cujo procedimento se mostrou mais rápido que o anterior. A metodologia consistiu na extração de 10⁷ células, sendo adicionados 50 µL de água ultra pura a 50 µL de cultura. As amostras foram incubadas em Thermomixer a 98°C por 15 minutos (para lisar as células), e centrifugadas em velocidade máxima por 1 minuto (o DNA permanece no sobrenadante).

3.7.3 Desenho e obtenção dos *primers* para confirmação dos nocautes

Para a PCR de confirmação utilizaram-se algumas combinações de *primers*, conforme detalhado a seguir. Os *primers* das regiões intergênicas dos genes alvos e *primers* das regiões intergênicas dos cassetes de resistência foram desenhados através do programa TcruziKO e adquiridos de fornecedor comercial conforme detalhado no item 3.4.1 da Metodologia. Os *primers* para as CDS dos genes alvo e genes de resistência foram desenhados com auxílio do programa DNASTAR, ferramenta *Primer SelectTM*, versão 5.06, do fabricante LASERGENE. Para a obtenção das sequências nucleotídicas empregou-se o *site* TriTrypDB com acesso da aba *Tools, Sequence Retrieval*. A identificação e a sequência de cada *primer* para confirmação dos nocautes estão anotadas na Tabela Suplementar 2 (em apêndice).

3.7.4 PCR de confirmação dos nocautes

A metodologia da PCR de confirmação consistiu em uma série de reações com o emprego de diferentes combinações de *primers*. Foi feita a PCR das CDS

usando os *primers* das CDS dos genes alvo, os *primers* dos genes de resistência; a PCR do cassete de nocaute usando os *primers* das IRS do cassete, e a combinação de *primers* das IRS do cassete com os *primers* dos genes de resistência; a PCR das CDS com suas IRS flanqueantes usando os *primers* das IRS. A reação foi preparada para tubos com volume final de 15 μL . Para cada reação, a concentração final de cada *primer* foi de 5 pmol, PCR *Buffer* minus Mg 1x, 2,5 mM MgCl_2 , 0,2 mM dNTPs mix, 1u Taq DNA polimerase, 1 μL de DNA-molde, 5,6 μL de água para completar o volume de 15 μL por reação de PCR. Com exceção do DNA-molde, dos *primers* e da Taq polimerase, os outros reagentes foram adquiridos da fabricante Invitrogen. O programa para a amplificação envolveu três etapas: (1) desnaturação a 94°C por 8 minutos; (2) 40 ciclos de amplificação – 94°C por 45 segundos, 56°C por 45 segundos e 72°C por 4 minutos e; (3) extensão a 72°C por 7 minutos. Os produtos amplificados tiveram sua especificidade e tamanhos das bandas verificados em eletroforese de DNA em gel de agarose 1%.

4 RESULTADOS E DISCUSSÃO

4.1 META 1: LEVANTAMENTO DE DADOS PARA CARACTERIZAÇÃO DAS PROTEÍNAS ESTUDADAS

Atualmente o acesso a bancos de dados e o uso de ferramentas bioinformáticas é uma necessidade frente ao volume de dados biológicos existentes; à rápida geração de novas informações e ao processamento das mesmas para execução de análises e estudos mais apurados. A bioinformática é usada para análise de dados genéticos, bioquímicos e de biologia molecular e hoje é um recurso imprescindível para a análise de dados em biologia molecular (PROSDOCIMI, 2007).

Os bancos de dados acessados e as ferramentas computacionais utilizadas estão descritas a seguir.

Avaliação do número de cópias gênicas

Primeiramente foi realizada uma avaliação a cerca do número de cópias gênicas dos quatro genes, para tomar conhecimento se esses eram passíveis de depleção por nocaute. Quanto maior o número de cópias no genoma, mais trabalhoso torna-se o nocaute. Primeiramente, realizou-se uma busca para as quatro proteínas de interesse no *site* de pesquisa TrypanosOmics, desenvolvido previamente por pesquisadores de bioinformática de nosso grupo. Nesse banco de dados, os genes de *T. cruzi* estão agrupados em famílias denominadas Supergenes (SGs), de acordo com as suas similaridades funcionais e estruturais, como domínios, funções, localizações, estrutura molecular, etc. Todas as proteínas estudadas estão contidas em supergenes contendo apenas dois integrantes (provavelmente genes alelos), uma vez que para os genes de cada Supergene a posição nucleotídica é muito semelhante, a localização cromossomal; o peso molecular e o comprimento são idênticos. Os Supergenes, seus genes e proteínas correspondentes podem ser vistos na Tabela 1, a seguir.

Tabela 1- Famílias de Supergenes (SGs) com seus genes e proteínas correspondentes. As proteínas sublinhadas correspondem àquelas selecionadas para estudo neste trabalho.

Supergenes (SGs)	Genes	Proteínas Reguladas	Localização Cromossomal	Posição Nucleotídica	Peso Molecular (Da)	Comprimento (aa)
SG1321 (Proteína 1)	TcCLB.506227.110	<u>Tc00.1047053506227.110</u>	TcChr14-S	175035 - 175802	28904	255
	TcCLB.511847.30	Tc00.1047053511847.30	TcChr14-P	174440 - 175207	28975	255
SG2809 (Proteína 2)	TcCLB.510661.230	<u>Tc00.1047053510661.230</u>	TcChr39-P	1419578 - 1421323	63211	581
	TcCLB.510741.40	Tc00.1047053510741.40	TcChr39-S	1419736 - 1421481	63140	581
SG4727 (Proteína 3)	TcCLB.506701.10	<u>Tc00.1047053506701.10</u>	TcChr36-P	456104 - 457636	56609	510
	TcCLB.509379.10	Tc00.1047053509379.10	TcChr36-S	455649 - 457181	56674	510
SG5427 (Proteína 4)	TcCLB.510183.40	<u>Tc00.1047053510183.40</u>	TcChr37-P	1278484 - 1279449	36454	321
	TcCLB.503637.10	Tc00.1047053503637.10	TcChr37-S	1278484 - 1279449	36438	321

* Da = Dalton; ** aa = aminoácido

<http://omics.icc.fiocruz.br/welcome>

<http://tritrypdb.org/tritrypdb/>

Assim, as proteínas de interesse desse estudo estão detalhadas a seguir: Tc00.1047053506227.110 (SG1321), Tc00.1047053510661.230 (SG2809), Tc00.1047053506701.10 (SG4727) e Tc00.1047053510183.40 (SG5427). Ao longo desse trabalho as proteínas estudadas foram referenciadas em nível de Supergenes (SGs) para facilitar a identificação das mesmas.

O alinhamento dos genes para cada Supergene é outro indicativo de tratar-se de alelos do mesmo gene em função da similaridade do número e dos tipos de aminoácidos entre as sequências (Figura 5).

A

```

Gene:TcCLB.506227.110      MRLRTVRCPAAVSTAATVKMTRAFSEARSNYDGTAAAPPWPAPGEKPTYPSALSEFFLQRP 60
Gene:TcCLB.511847.30      MRLRTVRCPAAVSTAATVKMTRAFSEARSNYDGTAAAPPWPAPGEKPKYPSALSEFFLQRP 60
*****

Gene:TcCLB.506227.110      RMRKTHIEWMFFHGHGGRPGVYQTREIADFEFADGTPASMSGRRFAFKHHQDHLVQLI 120
Gene:TcCLB.511847.30      RMRKTHIEWMFFHGHGGRPGVYQTREIADFEFADGTPASMSGRRFAFKHHQDHLVQLI 120
*****

Gene:TcCLB.506227.110      RAGAAVERKAARGLLPRVPGTPEQRNWDPAIPLFLEDVDEQGRPPPSVGRSICSRVIDER 180
Gene:TcCLB.511847.30      RAGAAVERKAARGLLPRVPGTPEQRNWDPAIPLFLEDVDEQGRPPPSVDRSICSRVIDER 180
*****

Gene:TcCLB.506227.110      MASSTQTPNELANRHEGETLEANTMFESYDPQAFVSDHIKLRDERRPYWSRRRWALTDNF 240
Gene:TcCLB.511847.30      MASSTQTPNELANRHEGETLEANTMFESYDPQAFVSDHIKLRDEKRPYWSRRRWALTDNF 240
*****

Gene:TcCLB.506227.110      LVPKSPKPKNTIKDE          255
Gene:TcCLB.511847.30      LVPKSPKPKNTIKDE          255
*****

```

B

Gene: TcCLB.510661.230 MGKTLGLSSLELKALSVPSLRFLLVGEYDVLLCVTEVVGKRRSCTIRFPVLGINFQKLE 60
 Gene: TcCLB.510741.40 MGKTLGLSSLELKALSVPSLRFLLVGEYDVLLCVTEVVGKRRSCTIRFPVLGINFQKLE 60

Gene: TcCLB.510661.230 LCAERFGLCCTQNEHSHLLSGTAIIPPLRYIDYLPALHAEELLGCLGVKPVTPNPKRYGA 120
 Gene: TcCLB.510741.40 LCAERFGLCCTQNEHSHLLSGTAIIPPLRYIDYLPALHAEELLSCLGVKPVTPNPKRYGA 120

Gene: TcCLB.510661.230 SVHDFLASFIDYSGYNTDASLADVHGTGFYLTLDHGLSTATHSHLVKISFHKGCSLEAL 180
 Gene: TcCLB.510741.40 SVHDFLASFIDYSGYNTDASLADVHGTGFYLTLDHGLSTATHSHLVKISFHKGCSLEAL 180

Gene: TcCLB.510661.230 REFEEFSGWTTFDIRYEDVDDKLAGGVVMPVSVQDACDLFEKYEDESDDEGGLLPFLMR 240
 Gene: TcCLB.510741.40 REFEEFSGWSTFDIRYEDVDDKLAGGVVMPVSVQDACDLFEKYEDESDDEGGLLPFLMR 240

Gene: TcCLB.510661.230 PVSPLTSVPSVEDSALKRLRDHQWGEIEKFWERFPKWKNCVVVSNLYESASLEEALALFK 300
 Gene: TcCLB.510741.40 PVSPLTSVPSVEDSALKRLRDHQWGEIEKFWERFPKWKNCVVVSNLYESASLEEALALFK 300

Gene: TcCLB.510661.230 DLTVTASSLVEDNSPARRRRVFTFATSDEAKTALFVDGKSTKPKTLRVQVSPPYLNE 360
 Gene: TcCLB.510741.40 DLTVTASSLVEDNSPARRRRVFTFATSDEAKIALFVDGKSTKPKTLRVQVSPPYLNE 360

Gene: TcCLB.510661.230 RGRVLESRPPSLSEPPGQQPATEKRPVSLGSPAHAAREKRDAAASNNPEASTPSGRNRSSK 420
 Gene: TcCLB.510741.40 RGRVLESRPPSLSEPPGQQPATETRPVSLGSPAHAAREKRDAAASNNPEASTPSGRDRSSK 420

Gene: TcCLB.510661.230 KAGDDKNWNAFVPKSRTPIQDAASAPVPGAAAAAREGGDQKKTATPPKQSQRSPIIVASTM 480
 Gene: TcCLB.510741.40 KAGDDKNWNAFVPTSRPTQEAASAPVTGAAAAAREGGDQKKTATPPKKPQCSPIIVASTM 480

Gene: TcCLB.510661.230 NANAREFVPSSTYQASTTSPEYYHSSIPMAYLDENYQPPPCYPAVALGPVGMARPMGVV 540
 Gene: TcCLB.510741.40 NANAREFVPSSTYQASTTSPEYYHSSIPMAYLDENYQPPPCYPAVALGPVGMVRPMGVV 540

Gene: TcCLB.510661.230 PRSLRSGAAPPYSSPPPYSMVSGQSMHMVQPPPPPPSS 581
 Gene: TcCLB.510741.40 PRSLRSGAAPPYSSPPPYSMVSGQSMQMAQPPPPPPSS 581

C

Gene: TcCLB.506701.10 MLHRSHFWGLKKVDVCVIGGGPAGIATALRAVDYKKKACIVEANRIGGADLWNGALQSKT 60
 Gene: TcCLB.509379.10 MLHRSYFWGLKKVDVCVIGGGPAGIATALRAVDYKKKACIVEANRIGGADLWNGALQSKT 60

Gene: TcCLB.506701.10 LWEMSKFVRLMTGQTASRFMKPLMPLPIDSSNAQKAIKNASELRYAQIVHQLKIVDIDV 120
 Gene: TcCLB.509379.10 LWEMSKFVRLMTGQTASRFMKPLMPLPIDSSNAQKAIKNASELRYTQIVHQLKIVDIDV 120

Gene: TcCLB.506701.10 IEGYGSFMSQSVVDVQLASGGTERVEADYFVIACGAQPRKHVSVTWGDKVVFTSDDIMNR 180
 Gene: TcCLB.509379.10 IEGYGSFMSPHSVDVQLASGGTERVEADYFVIACGAQPRKHVSVTWGDKVVFTSDDIMNR 180

Gene: TcCLB.506701.10 AFPKSIIVIGAGVIGCFEASIFANFGMTKVNVEIKSNRILPMEDDDVAEFVQKLEENKGV 240
 Gene: TcCLB.509379.10 AFPKSIIVIGAGVIGCFEASIFANFGMTKVNVEIKSNRILPMEDDDVAEFVQKLEENKGV 240

Gene: TcCLB.506701.10 TFHHHSALESNKVKNGQFHYTLRNLIDNKLTHIVDNALVSI GREPNLSGLGLENIGVGI 300
 Gene: TcCLB.509379.10 TFHHHSALESNKVKNGQFHYTLRNLIDNKLTHIVDNALVSI GREPNLSGLGLENIGVGL 300

Gene: TcCLB.506701.10 QRGKLERDQFNVRVPHKHIIYACGDVATKFALVNVGELGRSCIDHIYHPHPEGEFVQRLD 360
 Gene: TcCLB.509379.10 QRGKLERDQFNVRVPHKHIIYACGDVATKFALVNVGELGRSCIDHIYHPHPEGEFVQRLD 360

Gene: TcCLB.506701.10 NLSTIMFLDQEVAAVGLNEQQCRKQNIAYKMARYGYEFVARAVAMGNTRGFVKLIVTNDR 420
 Gene: TcCLB.509379.10 NLSTIMFLDQEVAAVGLNEQQCRKQNIAYKMARYGYEFVARAVAMGNTRGFVKLIVTNDR 420

Gene: TcCLB.506701.10 EMQVLGVRAIGPHASSVIELASLAIHNKETLYNLSELRTAYPTVTQGFQECVNMLLGTSL 480
 Gene: TcCLB.509379.10 EMQVLGVRAIGPHASSVIELASLAIHNKETLYNLSELRTAYPTVTQGFQECVNMLLGTSL 480

Gene: TcCLB.506701.10 LKPNVFPQLVLQEWSPPNFDRGRAYQGDR 510
 Gene: TcCLB.509379.10 LKPNVFPQLVLQEWSPPNFDRGRAYQGDR 510

D

```

Gene:TcCLB.510183.40      MLTPYQLELVEQAFHSLAPKGGDALPRIDEIFRVFDVSMHPRVKEGVMSPGAVLEMLAYQ 60
Gene:TcCLB.503637.10    MLTPYQLELVEQAFHSLAPKGGDALPRIDEIFRVFDVSMHPRVKEGVMSPGAVREMLAYQ 60
*****

Gene:TcCLB.510183.40      FGEESEKEQNGVTFDVFMRFHERMAMEAANECVSDTEGFLTDTIRGVWRLHELLEPTFIRP 120
Gene:TcCLB.503637.10    FGEESEKEQNGVTFDVFMRFHERMAMEAVNECVSDTEGFLTDTIKGVWRLHELLEPTFIRP 120
*****

Gene:TcCLB.510183.40      VIPIEGFPSNIYATQLMSLVWPDESKPPRSYVLRVHGVVQPLFSRGDLPELPGFFAYP 180
Gene:TcCLB.503637.10    AIPIEGFPSNIYATQLMSLVWPDESKPPRSYVLRVHGVVQPLFSRGDLPELPGFFAYP 180
*****

Gene:TcCLB.510183.40      AELAGMKLIEVPVQISMQRWLDVWEYAEQKYVAVPGIISARVDLETVPEYLRKLIIEPK 240
Gene:TcCLB.503637.10    AELAGMKLIEVPVQISMQRWLDVWEYAEQKYVAVPGIISARVDLETVPEYLRKLIIEPK 240
*****

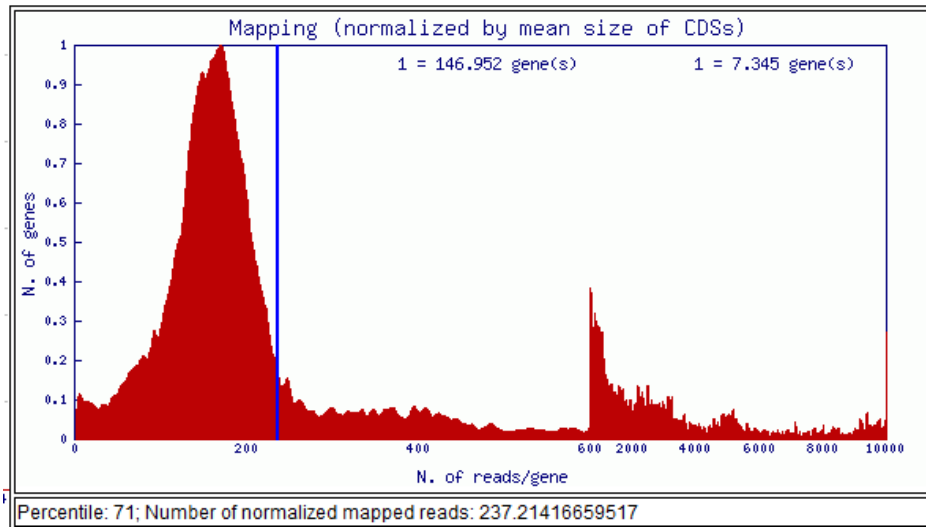
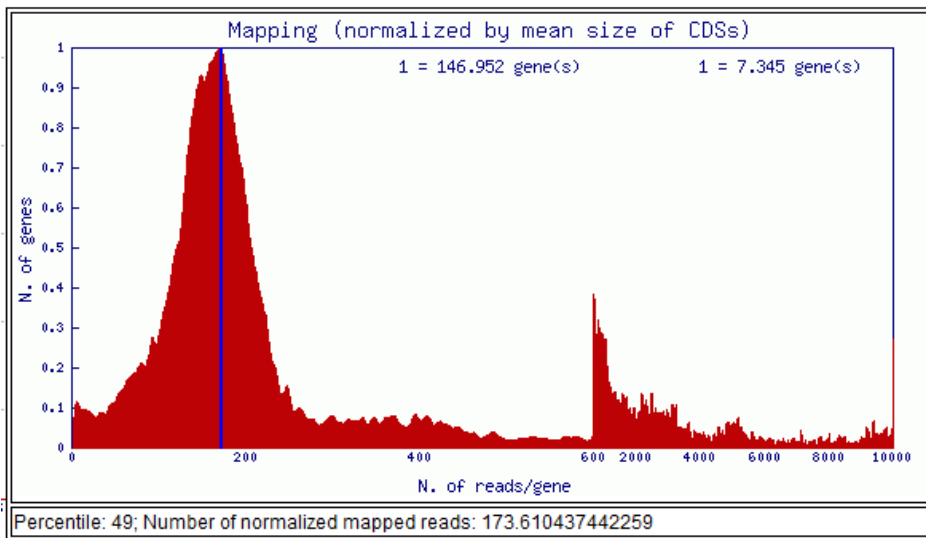
Gene:TcCLB.510183.40      VVKQLPSVQFVSTIKSLNPMYKRISDGYGYGITEEIKRVHRWKCETHDGTACGLIYHGRV 300
Gene:TcCLB.503637.10    VVKQLPSVQFVSTIKALNPMYKRISDGYGYGIPEEVKRVHRWKCETHDGTACGLIYHGRV 300
*****

Gene:TcCLB.510183.40      GKFTKKLYGQAQSLAASGINL 321
Gene:TcCLB.503637.10    GKFTKKLYGQAQSLAASGINL 321
*****

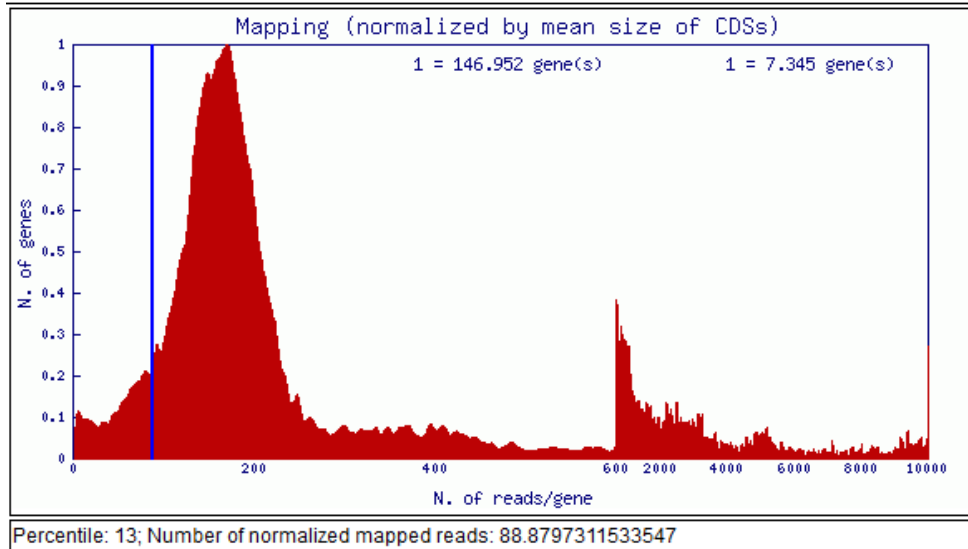
```

Figura 5 - Alinhamento dos genes agrupados em cada Supergene. (A) SG1321; (B) SG2809; (C) SG4727; (D) SG5427 (<http://www.ebi.ac.uk/Tools/msa/clustalw2/>). As setas em preto indicam as diferenças no alinhamento.

Além disso, o gráfico do número de *reads*/número de genes é outro indicativo dos genes em estudo serem cópia única. Genes que possuem até cerca de 200 *reads* são cópia única. A linha em azul representa a posição do gene em estudo. Os grandes picos correspondem a maioria dos genes de *T. cruzi*, que são cópia única no genoma. Os gráficos do número de *reads*/número de genes consta a seguir (Figura 6).

A**B**

C



D

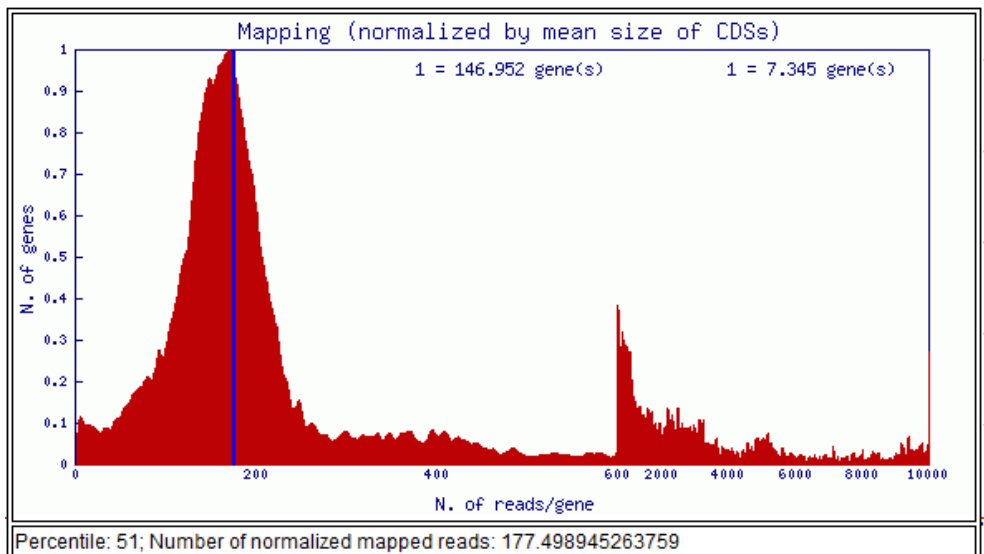


Figura 6 - Número de cópias dos genes no genoma de *T. cruzi*. O eixo das abcissas representa o número de *reads* por gene, e o eixo das ordenadas representa o número de genes. A linha em azul indica a posição do gene em estudo. (A) SG1321; (B) SG2809; (C) SG4727; (D) SG5427 (<http://omics.icc.fiocruz.br/welcome>).

Os dados de expressão na proteômica foram coincidentes para os dois genes de cada Supergene, sendo outro indicativo de tratar-se de alelos do mesmo gene. Além destas, outras informações foram checadas nos bancos de dados Omics.icc.fiocruz e TritrypDB.org para cada gene dos Supergenes e o resultado foi

idêntico ou muito semelhante para ambos os genes, indicando provavelmente a condição de alelos.

Homologia Gênica

A homologia gênica indica a ancestralidade gênica em comum entre organismos. A exclusividade da sequência em uma espécie indica função relacionada à biologia particular do organismo, em algum mecanismo não existente em outras espécies. Por sua vez, a não exclusividade indica importância e participação em mecanismos biológicos existentes em mais de um grupo. A homologia possui importância para a predição de função, principalmente quando a sequência gênica ocorre em outros grupos. Para tanto, é necessária a análise do grau de similaridade, visto que genes homólogos podem acumular, ao longo do tempo, variações que os tornam tão diferentes geneticamente que passam a ter diferentes funções. Os quatro genes analisados são conservados entre outros grupos de tripanosomatídeos, indicando a importância dos mesmos. A análise da homologia gênica foi realizada em banco de dados exclusivo de tripanosomatídeos. Seria interessante a análise da homologia com outros grupos de organismos como, por exemplo, com a espécie humana. Isto seria interessante para alguns propósitos, como o desenvolvimento de fármacos específicos contra algumas proteínas exclusivas de *T. cruzi*. A descrição da homologia dos genes em estudo encontra-se a seguir (Figura 7).

A

Orthologs and Paralogs within TriTrypDB [Hide](#)

Gene	Organism	Product	is syntenic	has comments
LbrM.13.0130	Leishmania braziliensis MHOM/BR/75/M2904	hypothetical protein, conserved	yes	no
LdBPK_130140.1	Leishmania donovani BPK282A1	hypothetical protein, conserved	yes	no
LinJ.13.0140	Leishmania infantum JPCM5	hypothetical protein, conserved	yes	no
LmjF.13.0140	Leishmania major strain Friedlin	hypothetical protein, conserved	yes	no
LmxM.13.0140	Leishmania mexicana MHOM/GT/2001/U1103	hypothetical protein, conserved	yes	no
LtaP13.0130	Leishmania tarentolae Parrot-Tarll	Hypothetical protein, conserved	yes	no
Tb427tmp.02.2250	Trypanosoma brucei Lister strain 427	hypothetical protein, conserved	yes	no
Tb927.11.4650	Trypanosoma brucei TREU927	hypothetical protein, conserved	yes	yes
Tbg972.11.5290	Trypanosoma brucei gambiense DAL972	hypothetical protein, conserved	yes	no
TclL3000.11.4730	Trypanosoma congolense IL3000	hypothetical protein, conserved	yes	no
TclL3000.11.4740	Trypanosoma congolense IL3000	unspecified product	yes	no
TcCLB.506227.120	Trypanosoma cruzi CL Brener Esmeraldo-like	hypothetical protein, conserved	no	no
TcCLB.511847.20	Trypanosoma cruzi CL Brener Non-Esmeraldo-like	hypothetical protein, conserved	yes	no
TcCLB.511847.30	Trypanosoma cruzi CL Brener Non-Esmeraldo-like	hypothetical protein, conserved	yes	no
TCSYLVI0_002174	Trypanosoma cruzi Sylvio X10/1	hypothetical protein	yes	no
TCSYLVI0_002175	Trypanosoma cruzi Sylvio X10/1	hypothetical protein	yes	no
Tc_MARK_1385	Trypanosoma cruzi marinkellei strain B7	hypothetical protein, conserved	yes	no
Tc_MARK_1386	Trypanosoma cruzi marinkellei strain B7	hypothetical protein, conserved	yes	no
TVY486_1104740	Trypanosoma vivax Y486	hypothetical protein, conserved (fragment)	yes	no
TVY486_1104750	Trypanosoma vivax Y486	hypothetical protein, conserved	yes	no

B

Orthologs and Paralogs within TriTrypDB [Hide](#)

Gene	Organism	Product	is syntenic	has comments
LbrM.34.3110	Leishmania braziliensis MHOM/BR/75/M2904	hypothetical protein, conserved	yes	no
LdBPK_353250.1	Leishmania donovani BPK282A1	hypothetical protein, conserved	yes	no
LinJ.35.3250	Leishmania infantum JPCM5	hypothetical protein, conserved	yes	no
LmjF.35.3200	Leishmania major strain Friedlin	RNA-binding protein, putative	yes	no
LmxM.34.3200	Leishmania mexicana MHOM/GT/2001/U1103	hypothetical protein, conserved	yes	no
LtaP35.3230	Leishmania tarentolae Parrot-Tarll	Hypothetical protein, conserved	yes	no
Tb927.9.12360	Trypanosoma brucei TREU927	RNA-binding protein, putative (RBP35)	yes	no
Tbg972.9.7550	Trypanosoma brucei gambiense DAL972	hypothetical protein, conserved	yes	no
TclL3000_9_5150	Trypanosoma congolense IL3000	unspecified product	yes	no
TcCLB.510661.230	Trypanosoma cruzi CL Brener Non-Esmeraldo-like	hypothetical protein, conserved	yes	no
TCSYLVI0_005823	Trypanosoma cruzi Sylvio X10/1	hypothetical protein	yes	no
Tc_MARK_4451	Trypanosoma cruzi marinkellei strain B7	hypothetical protein, conserved	yes	no
TVY486_0905890	Trypanosoma vivax Y486	hypothetical protein, conserved	yes	no

C

Orthologs and Paralogs within TriTrypDB [Hide](#)

Gene	Organism	Product	is syntenic	has comments
LbrM.31.0230	Leishmania braziliensis MHOM/BR/75/M2904	dihydroipoamide dehydrogenase	yes	no
LdBPK_291950.1	Leishmania donovani BPK282A1	dihydroipoamide dehydrogenase, putative	yes	no
LinJ.29.1950	Leishmania infantum JPCM5	dihydroipoamide dehydrogenase, putative (GCVL-1)	yes	no
LmjF.29.1830	Leishmania major strain Friedlin	dihydroipoamide dehydrogenase, putative (GCVL-1)	yes	no
LmxM.08_29.1830	Leishmania mexicana MHOM/GT/2001/U1103	dihydroipoamide dehydrogenase, putative	yes	no
LtaP29.1940	Leishmania tarentolae Parrot-Tarll	Dihydroipoamide dehydrogenase;dihydroipoamide dehydrogenase, putative	yes	no
Tb427.03.4390	Trypanosoma brucei Lister strain 427	dihydroipoamide dehydrogenase, putative (GCVL-1)	yes	no
Tb927.3.4390	Trypanosoma brucei TREU927	dihydroipoamide dehydrogenase, putative (GCVL-1)	yes	no
Tbg972.3.4790	Trypanosoma brucei gambiense DAL972	dihydroipoamide dehydrogenase, putative	yes	no
TcCLB.506701.10	Trypanosoma cruzi CL Brener Non-Esmeraldo-like	dihydroipoamide dehydrogenase, putative (GCVL-1)	yes	no
TCSYLVI0_007650	Trypanosoma cruzi Sylvio X10/1	dihydroipoamide dehydrogenase, putative	yes	no
Tc_MARK_3138	Trypanosoma cruzi marinkellei strain B7	dihydroipoamide dehydrogenase, putative	yes	no

D

Orthologs and Paralogs within TriTrypDB [Hide](#)

Gene	Organism	Product	is syntenic	has comments
LbrM.35.2610	Leishmania braziliensis MHOM/BR/75/M2904	hypothetical protein, conserved	yes	no
LdBPK_362530.1	Leishmania donovani BPK282A1	hypothetical protein, conserved	yes	no
LinJ.36.2530	Leishmania infantum JPCM5	hypothetical protein, conserved	yes	no
LmjF.36.2400	Leishmania major strain Friedlin	hypothetical protein, conserved	yes	no
LmxM.36.2400	Leishmania mexicana MHOM/GT/2001/U1103	hypothetical protein, conserved	yes	no
LtaP36.2470	Leishmania tarentolae Parrot-Tarll	Hypothetical protein, conserved	yes	no
Tb427.10.6920	Trypanosoma brucei Lister strain 427	hypothetical protein	yes	no
Tb427.10.6960	Trypanosoma brucei Lister strain 427	hypothetical protein, conserved	yes	no
Tb10.v4.0248	Trypanosoma brucei TREU927	hypothetical protein,chrX additional, unordered contigs	no	no
Tb927.10.6920	Trypanosoma brucei TREU927	hypothetical protein	yes	no
Tb927.10.6960	Trypanosoma brucei TREU927	hypothetical protein, conserved	yes	no
Tbg972.10.8470	Trypanosoma brucei gambiense DAL972	hypothetical protein, conserved	yes	no
TcCLB.503637.10	Trypanosoma cruzi CL Brener Esmeraldo-like	hypothetical protein, conserved	yes	no
TCSYLVI0_007534	Trypanosoma cruzi Sylvio X10/1	hypothetical protein	yes	no
Tc_MARK_3747	Trypanosoma cruzi marinkellei strain B7	hypothetical protein, conserved	yes	no
TvY486_1006860	Trypanosoma vivax Y486	hypothetical protein, conserved	yes	no

Figura 7- Representação da homologia dos genes com a indicação de seus genes parálogos e ortólogos (<http://tritrypdb.org/tritrypdb/>). (A) SG1321; (B) SG2809; (C) SG4727; (D) SG5427.

Expressão proteica durante a metaciclogênese *in vitro* de *T. cruzi*

A expressão proteica das quatro proteínas diferencialmente expressas nas fases iniciais da metaciclogênese de *T. cruzi* foi obtida no trabalho de proteômica quantitativa desenvolvida por pesquisadores do Instituto Carlos Chagas, cuja regulação proteica durante a metaciclogênese foi representada em gráficos. Esses gráficos relacionam tempos específicos após a indução da metaciclogênese e a quantificação proteica. Referente a eles, tem-se: *epi* = forma epimastigota; *meta* = forma tripomastigota metacíclica; *St30m* = estresse nutricional após 30 minutos da indução; *St2h* = estresse nutricional após 2 horas de indução; *Ade12h* = parasitas aderidos após 12 horas de indução (de GODOY *et al.*, 2012). Além do estresse nutricional, a condição de adesão ao substrato pelo parasita, um dos momentos monitorados no estudo acima citado, é um processo essencial para a diferenciação. A adesão e o estresse nutricional envolvem a participação de ácidos graxos, cAMP e adenilato ciclase (SANTOS; SILVA; YOSHIDA, SILVEIRA, 2013). A expressão proteica diferenciada ao longo da metaciclogênese pode indicar participação em processos de sinalização celular. Os tripomastigotas metacíclicos são a primeira forma de contato com as células do hospedeiro vertebrado, e isto pode indicar a existência de funções invasivas especializadas, em comparação aos tripomastigotas sanguíneos. Estudos moleculares referentes à capacidade invasiva do *T. cruzi* tem identificado um grande número de moléculas envolvidas em interações parasita-hospedeiro e detecção de sinais bioquímicos que facilitam a entrada do parasita nas células (TALICE *et al.*, 2007).

SG1321 - Tc00.1047053506227.110

A proteína apresenta um aumento transiente em estresse 30 minutos. Esta regulação positiva pode ser um indicativo da participação em processos de sinalização celular, uma vez que este aumento seja uma condição necessária para, supostamente, o recrutamento de outras proteínas (Figura 8).

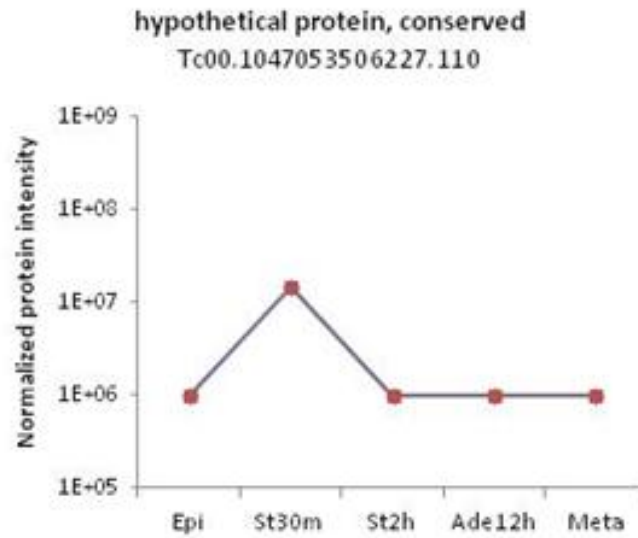


Figura 8 - Expressão da proteína SG1321 durante metaciclogênese de *T. cruzi* (de GODOY *et al.*, 2012).

SG2809 - Tc00.1047053510661.230

A proteína apresenta um aumento transiente em estresse duas horas. Esta regulação positiva, a exemplo da proteína anterior, pode ser um indicativo de participação em processos de sinalização celular (Figura 9).

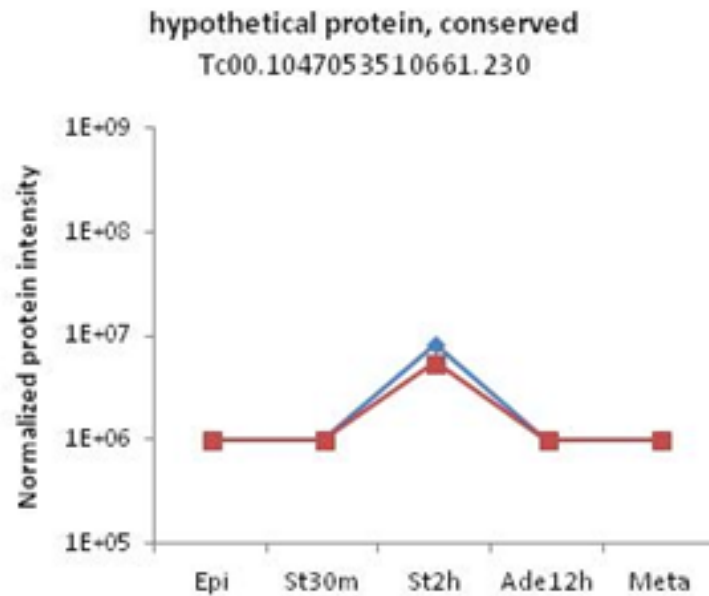


Figura 9 - Expressão da proteína SG2809 durante metaciclogênese de *T. cruzi* (de GODOY *et al.*, 2012).

SG4727 - Tc00.1047053506701.10

A proteína é regulada negativamente em aderido 12 horas. Apesar de ser uma diidrolipoamida desidrogenase, enzima mitocondrial envolvida na respiração celular e cujos detalhes serão discutidos em tópico referente aos domínios proteicos, essa sofre uma diminuição da sua expressão, em oposição à ideia de necessidade de energia celular em processos de diferenciação celular. Além de possivelmente atuar em processos para geração de energia celular, pode ser necessária a diminuição da expressão desta proteína como sinalização para uma segunda proteína que seja crucial para o processo de adesão do parasita. A adesão é um processo imprescindível para a diferenciação em formas metacíclicas, e envolve proteínas flagelares do parasita que promovem a adesão do flagelo à parede do intestino do inseto vetor ou; na metaciclogênese *in vitro*, a adesão ocorre no fundo das garrafas de cultivo, sendo por este motivo que as garrafas permanecem deitadas após o acréscimo do meio TAU3AAG às culturas (Figura 10).

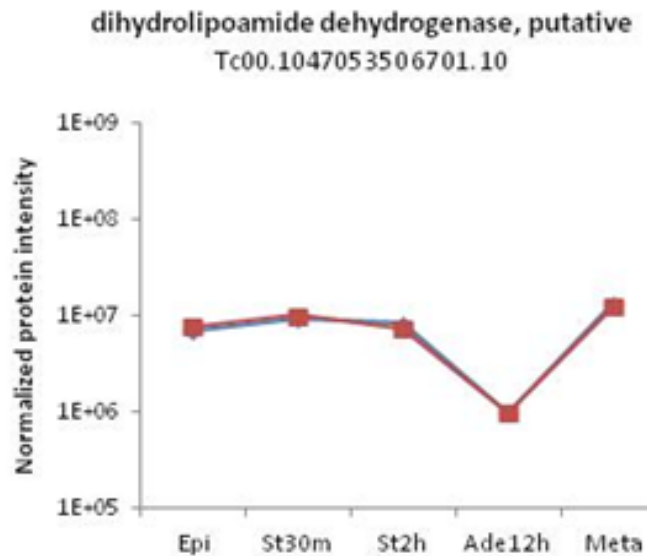


Figura 10 - Expressão da proteína SG4727 durante metaciclogênese de *T. cruzi* (de GODOY et al., 2012).

SG5427 - Tc00.1047053510183.40

A proteína apresenta um aumento gradativo na sua expressão durante toda a metaciclogênese. Isto pode ser um indicativo de sua importância e envolvimento em processos celulares existentes ao longo de toda a diferenciação celular. Pelo fato desta proteína ser regulada positivamente ao longo de toda a metaciclogênese, esta torna-se uma candidata interessante para o desenvolvimento de possíveis fármacos contra a Doença de Chagas (Figura 11).

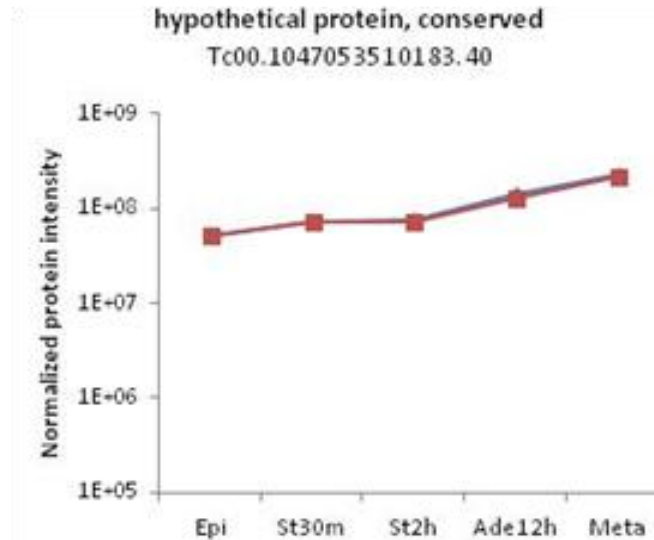


Figura 11- Expressão da proteína SG5427 durante metaciclogênese de *T. cruzi* (de GODOY *et al.*, 2012)

Domínios proteicos

SG1321 - Tc00.1047053506227.110 (hypothetical protein)

Não há informações referentes a domínios proteicos funcionais, assim, está anotada como proteína hipotética. No entanto, essa proteína apresenta em sua sequência um peptídeo sinal (Figura 12), que pode ser indicativo de sua localização subcelular associada a organelas membranosas.

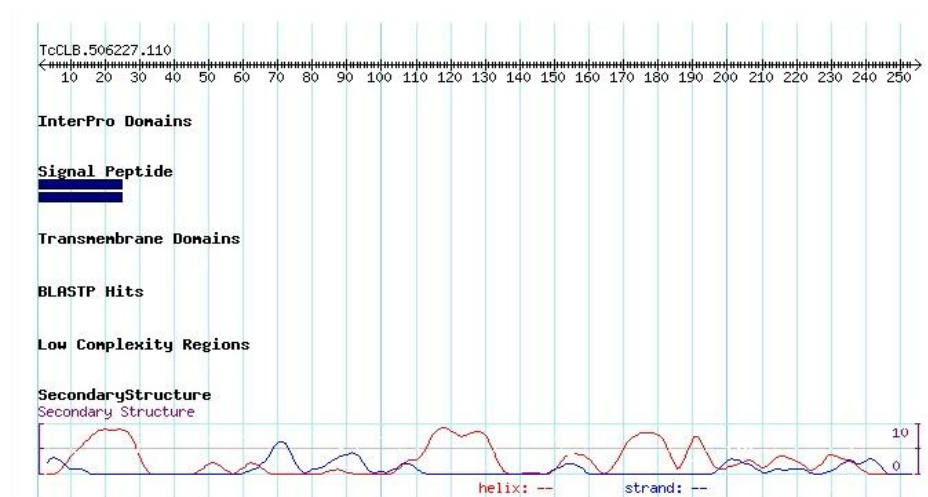


Figura 12 - Identificação de um peptídeo sinal para a proteína SG1321 (<http://tritrypdb.org/tritrypdb/>).

A maioria das proteínas é sintetizada no citoplasma. Mas, dependendo da sua função, elas podem atuar em diferentes locais celulares, ou ainda ser exportadas para o meio extracelular. Assim, é necessário que elas sejam transportadas corretamente para locais celulares específicos para desempenhar sua função (NETO, 2012).

O transporte de proteínas a partir do citoplasma se dá por diferentes vias: 1) no transporte mediado, as proteínas movimentam-se entre o citosol e o núcleo por meio dos complexos de poro nuclear no envelope nuclear; 2) no transporte transmembrana, proteínas translocadoras transportam as proteínas específicas através da membrana a partir do citosol para as mitocôndrias, peroxissomos, plastídeos ou retículo endoplasmático, onde a proteína transportada geralmente desdobra-se para passar pelo transportador; 3) no transporte vesicular, as proteínas são deslocadas por meio de vesículas, como no caso do transporte de proteínas entre retículo endoplasmático e Complexo de Golgi (ALBERTS *et al.*, 2010).

O endereçamento de proteínas é dependente de sinais intrínsecos encontrados nas sequências proteicas. Os sinais de endereçamento nas proteínas transportadas devem ser reconhecidos por proteínas receptoras que as guiam, por exemplo, através do poro nuclear; através de alguma membrana pela qual deva ser translocada; ou na formação de vesículas para seu transporte. Dentre estes sinais intrínsecos de endereçamento, o mais conhecido e estudado é o peptídeo sinal, responsável pela translocação de proteínas através da membrana plasmática (em procariotos) ou através da membrana do retículo endoplasmático rugoso (RER) (em eucariotos). Cada peptídeo sinal especifica um destino particular na célula (NETO, 2012; ALBERTS *et al.*, 2010).

Proteínas que possuem peptídeo sinal são translocadas para o lúmen do RER, de onde poderá seguir para compartimentos como: Complexo de Golgi, retículo endoplasmático, lisossomos, vacúolo digestivo, vesículas secretórias (proteínas que são destinadas ao meio extracelular) e diferentes membranas plasmáticas (TASHIRO, 1983).

O peptídeo sinal consiste em uma sequência de 15-60 resíduos de aminoácidos, encontrada normalmente na porção N-terminal, sendo formada por uma porção central hidrofóbica flanqueada na sua extremidade N-terminal por uma região polar com carga positiva (Figura 13). Alguns peptídeos sinais podem ser

extensões internas de aminoácidos, as quais permanecem como parte da proteína, sendo chamados de região sinal (NETO, 2012; ALBERTS *et al.*, 2010).

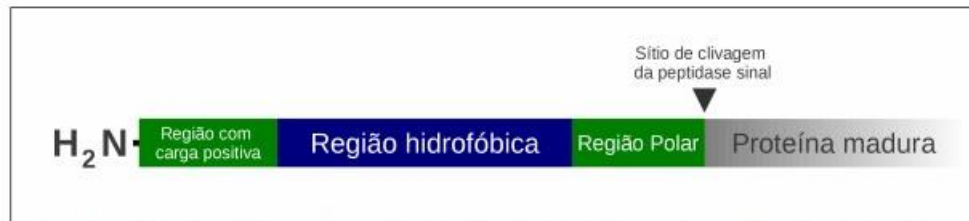


Figura 13 - Estrutura típica do peptídeo sinal. O peptídeo sinal geralmente se localiza na extremidade N-terminal de proteínas pré-processadas, sendo formado por uma região polar com carga positiva na sua extremidade N-terminal, uma região hidrofóbica e uma região polar que determina o seu sítio de clivagem (NETO, 2012).

Não foram encontradas informações referentes a domínios proteicos; Ontologia Gênica e participação em vias metabólicas e; apesar de tratar-se de uma proteína conservada com grande número de genes parálogos e ortólogos, estes têm como produtos proteínas também hipotéticas. Desta maneira, é difícil prever qual a sua função. Apesar disto, a identificação de um peptídeo sinal é relevante uma vez que o seu endereçamento para algum compartimento celular é um indicativo do local onde desempenha a sua função. Esta informação, associada a outros dados, podem ser usados para predição da sua função, que neste caso está possivelmente envolvida no processo de diferenciação celular. A função de um gene não engloba apenas as interações moleculares diretas nas quais o seu produto proteico se envolve, mas também está relacionada ao momento e ao local subcelular nos quais estas interações ocorrem. Além disto, uma proteína que apresenta um peptídeo sinal é uma candidata interessante para o desenvolvimento de fármacos, visto que as proteínas que são secretadas ou que estejam expostas na superfície celular são alvos mais acessíveis para ação de vacinas e drogas (GOODSWEN; KENNEDY; ELLIS, 2012). Portanto, a identificação de um peptídeo sinal torna esta proteína uma interessante candidata para a realização de estudos futuros, principalmente visando o desenvolvimento de drogas e vacinas.

SG2809 - Tc00.1047053510661.230 (hypothetical protein)

A proteína SG2809, apesar de anotada como hipotética, possui descritos dois domínios de reconhecimento de RNA (RRM, *RNA-recognition motif*) (Figura 14). Os domínios RRM também são chamados de RBD (*RNA binding domain*). As proteínas que possuem este domínio são denominadas proteínas hnRNP (*heterogeneous nuclear ribonucleoproteins*) ou simplesmente RBPs (ribonucleoproteínas). O domínio RRM é o mais abundante e o mais estudado domínio de ligação ao RNA em vertebrados superiores, apesar de também ser encontrado em procariotos e em vírus. Foram identificadas 85 proteínas contendo um domínio RRM em procariotos e, seis proteínas com este domínio em vírus. Entre os eucariotos, um total de 6.056 motivos RRM foram identificados em 3.541 diferentes proteínas. Em humanos, 497 proteínas contendo ao menos um domínio RRM foram encontradas. Estudos bioquímicos para caracterização molecular foram realizados e sabe-se que este domínio é formado por duas cadeias α -hélices relacionadas a quatro folhas β -pregueadas. É considerado um domínio pouco extenso, com cerca de 90 aminoácidos contendo uma sequência central conservada de oito resíduos, além de estruturas aromáticas e positivamente carregadas. Estas regiões conservadas são chamadas de sequência consenso e são conhecidas por RNP 1 e RNP 2. O sítio de ligação do domínio aos alvos tem tamanho variável para cada proteína (de 2 a 8 nucleotídeos), e geralmente localizam-se nos domínios RNP1 e RNP2 (CLÉRY; BLATTER; ALLAIN, 2008; MARIS; DOMINGUEZ; ALLAIN, 2005; DREYFUSS; KIM; KATAOKA, 2002).

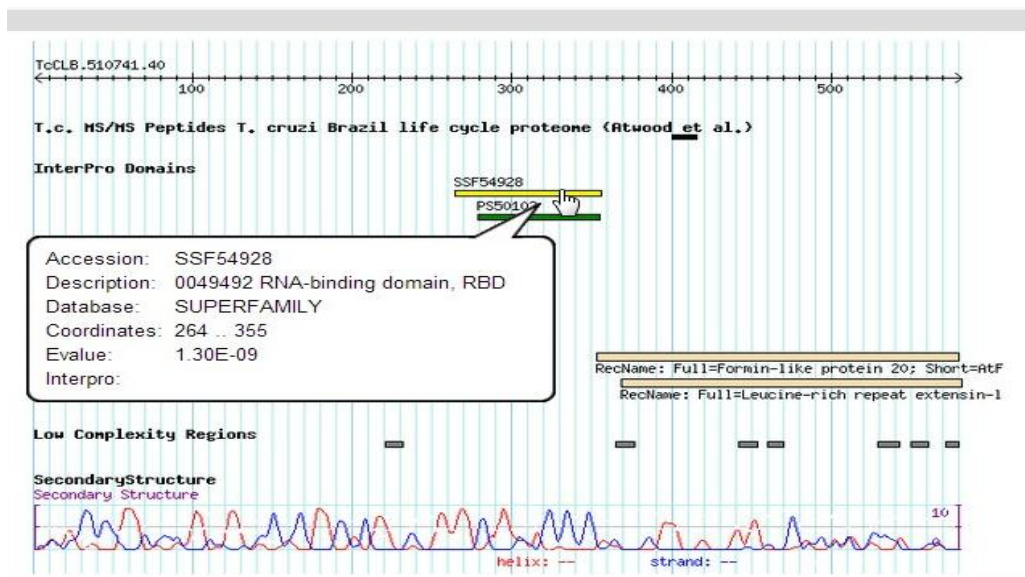
As proteínas que contém o domínio RRM estão envolvidas em mecanismos pós-transcricionais de regulação da expressão gênica, que englobam o processamento do mRNA e rRNA, a exportação nuclear do RNA, localização subcelular, *trans-splicing*, adição de mini-éxon e estabilidade destes transcritos. Nos últimos dez anos, estudos bioquímicos e estruturais demonstraram que este domínio não está envolvido apenas com o reconhecimento de RNA/DNA, mas também participa em interações proteína-proteína, onde a sua função e os parceiros de interação não são totalmente conhecidos (CLÉRY; BLATTER; ALLAIN, 2008).

A maioria das proteínas contém múltiplos domínios RRM que aumentam a afinidade e especificidade de ligação em comparação com as proteínas com único

domínio. Em algumas proteínas, na presença de RNA, os dois domínios se arranjam quase paralelamente para acomodar, entre elas, a molécula de RNA, a partir da qual há o recrutamento de outras proteínas para o *splicing*. Há ainda outros casos de arranjo, onde os domínios se arranjam em posição anti-paralela, provocando o encaixe da molécula de RNA que passa a formar alças e, desta maneira, favorecendo a transcrição de alguns éxons na obtenção do RNA maduro (CLÉRY; BLATTER; ALLAIN, 2008).

Os mecanismos de regulação da expressão gênica em tripanosomatídeos ocorre predominantemente por mecanismos pós-transcricionais. Os mRNAs de *T. cruzi* interagem com diferentes fatores proteicos em uma complexa maquinaria de regulação gênica pós-transcricional, que determina os níveis dos produtos proteicos de acordo com as demandas do parasita em cada estágio do ciclo de vida. O mRNA é transcrito em unidades policistrônicas de pré-mRNAs, que são processadas por *trans-splicing* e reações de poliadenilação. O *trans-splicing* consiste no capeamento na região 5' de uma sequência de 39 nucleotídeos (*capped-spliced leader* – SL) e a adição da cauda poli-A na extremidade 3' dos transcritos, gerando mRNAs maduros monocistrônicos que se acumulam em diferentes níveis no citoplasma (ARAÚJO; TEIXEIRA, 2011).

A



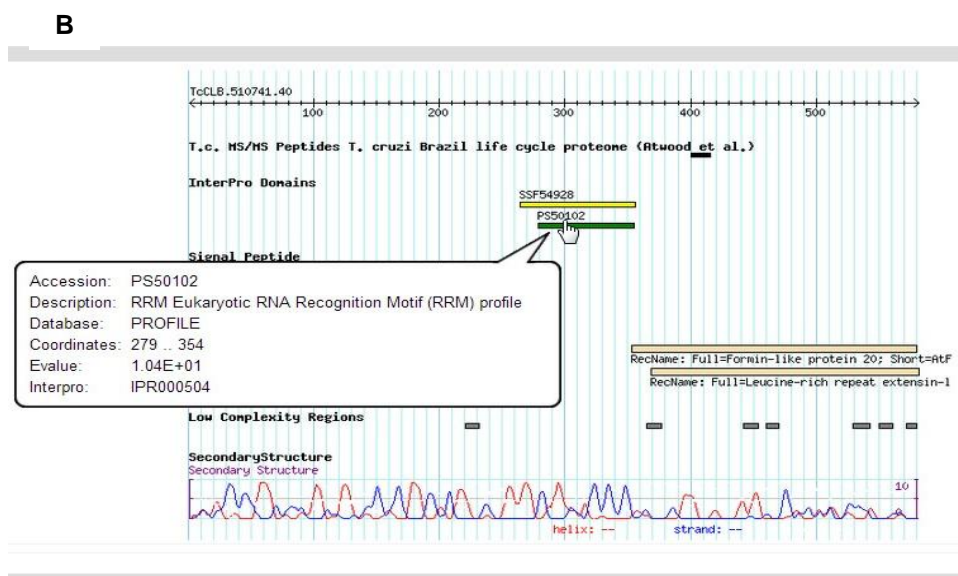


Figura 14 - Domínios descritos para a proteína SG2809 (<http://tritypdb.org/tritypdb/>).

As RBPs assumem um papel central nos mecanismos pós-transcricionais de regulação da expressão gênica, modulando diversas etapas do metabolismo do RNA. O gene codificador dessa proteína possui um ortólogo em *T. brucei* (Tb09.211.3370), o qual é similar a RBP35. Em fungos filamentosos, a proteína RBP35 está associada ao processamento de RNA e à origem de infectividade (de GODOY *et al.*, 2012).

Referente à função da proteína SG2809, não foram encontradas informações referentes à participação em vias metabólicas. Porém, a presença destes dois domínios RRM parece indicar o envolvimento da proteína em mecanismos pós-transcricionais de regulação da expressão gênica, sendo este modelo de regulação predominante em tripanosomatídeos. A Ontologia Gênica aponta para participação no processo de ligação a ácidos nucleicos; enquanto a Homologia Gênica aponta genes parálogos e ortólogos que possuem como produtos algumas proteínas hipotéticas e outras de ligação ao RNA. Além desta função de reconhecimento e ligação a RNA e DNA, o domínio possui função de interação entre proteínas. Essas proteínas podem ser parceiras de interação, cuja elevação do nível celular de uma delas concorra positivamente para a ocupação dos sítios ativos da outra proteína. Desta maneira, esta proteína regulada positivamente em estresse duas horas também pode participar de processos de sinalização celular, onde sua expressão aumentada ativa outras proteínas envolvidas com a metaciclo gênese; além de essas

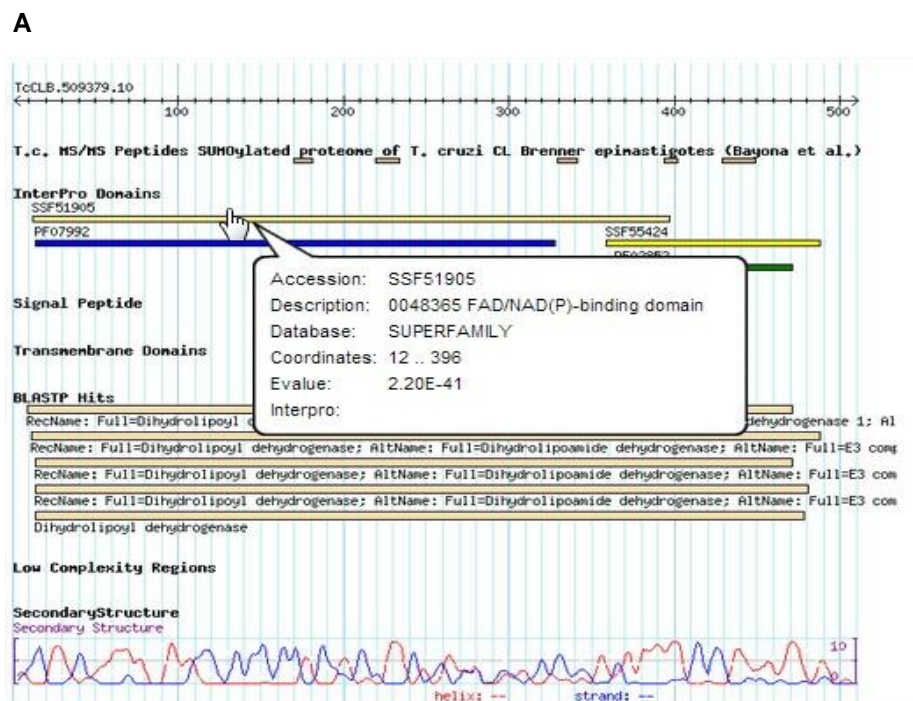
proteínas poderem estar envolvidas com a aquisição de infectividade para células de mamíferos.

SG4727 - Tc00.1047053506701.10 (dihidrolipoamide dehydrogenase, putative)

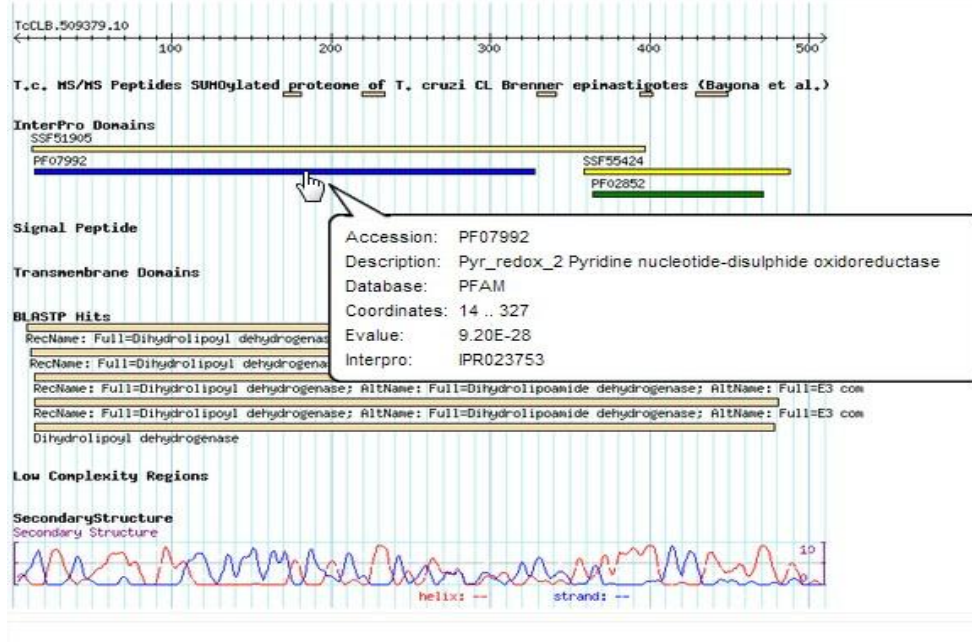
A proteína SG4727 é uma diidrolipoamida desidrogenase (TcLipDH, LipDH ou LADH). A Homologia Gênica indica tratar-se de uma proteína conservada com a existência de genes parálogos e ortólogos, tendo como produtos proteínas com função de diidrolipoamida desidrogenase. A LipDH consiste em uma subunidade de quatro diferentes complexos enzimáticos mitocondriais. Como parte da piruvato desidrogenase (PDH) e 2-oxoglutarato desidrogenase (KDH) complexos, a flavoenzima está envolvida com o metabolismo energético da célula. Como uma subunidade da cadeia-lateral da desidrogenase cetoácida (BCKDH), ela participa da degradação de cadeias-laterais de aminoácidos (BCAAs). Como componente do complexo da clivagem da glicina (GCC), a LipDH participa da formação de metilenotetrahidrofolato (ROLDÁN; COMINI; CRISPO; SIEGEL, 2011). Ela é encontrada em todos os organismos aeróbicos e em alguns organismos anaeróbicos (SANTOS, 2008). Referente à localização subcelular, há evidências em outros tripanosomatídeos, como em *T. brucei*, que ao longo do ciclo de vida do organismo a LipDH se localiza na mitocôndria ou na membrana plasmática (ROLDÁN; COMINI; CRISPO; SIEGEL, 2011).

Referente ao metabolismo energético da célula, a diidrolipoamida desidrogenase faz parte do complexo da piruvato desidrogenase, enzima que promove a descarboxilação oxidativa, um processo irreversível de oxidação no qual o grupo carboxila é removido do piruvato na forma de uma molécula de CO₂ e os dois carbonos remanescentes tornam-se o grupo acetil do acetil-CoA. A diidrolipoamida desidrogenase catalisa a regeneração da forma dissulfeto (oxidado) do lipoato, os elétrons passam para FAD, e depois para o NAD⁺ (LEHNINGER; NELSON; COX, 2006). A LipDH é uma dissulfeto oxidorreductase, catalisa o NAD(P)H- dependente do dissulfeto como substrato, mantendo o meio intracelular reduzido. No complexo mitocondrial formado por piruvato desidrogenase e 2-oxoglutarato desidrogenase a enzima catalisa a oxidação da diidrolipoamida

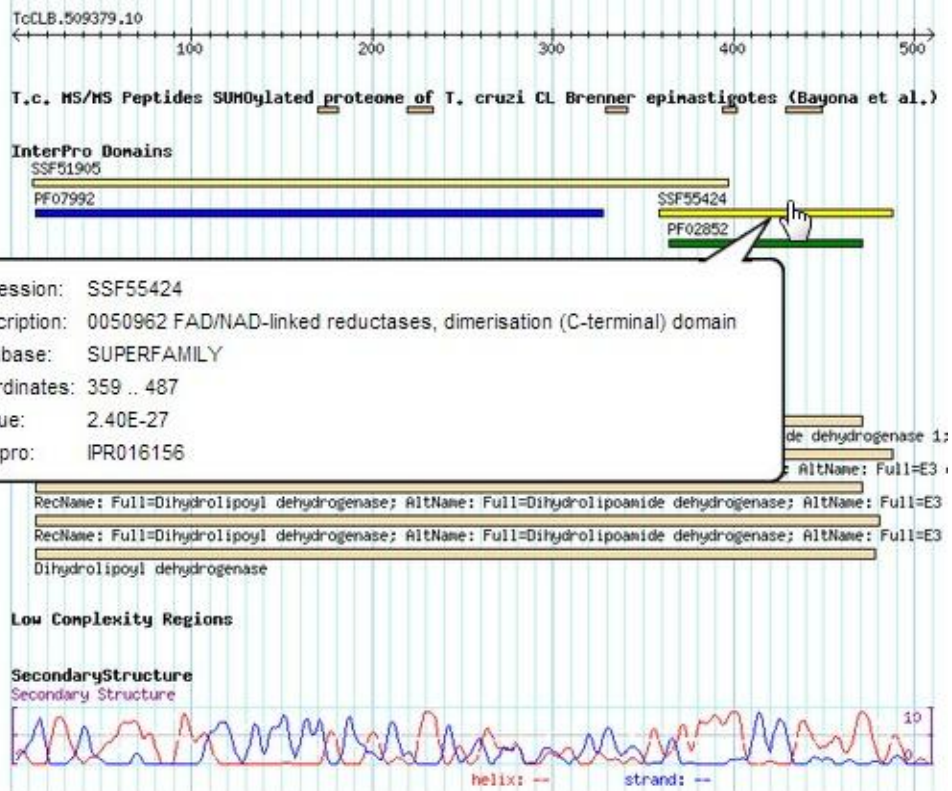
(dihidrolipoamida + NAD^+ \rightarrow lipoamida + $\text{NADH} + \text{H}^+$), utilizando FAD como cofator. As flavoenzimas LipDH, tripanotona redutase e glutatona redutase podem realizar uma oxidação não-específica lenta de NADPH com oxigênio molecular. O consumo de NADPH também diminui a proteção contra radicais livres (de GODOY *et al.*, 2012; SANTOS, 2008). A Ontologia Gênica corrobora essas informações, ao apontar como função da proteína a ligação à FAD e homeostase redox celular, além da participação no transporte de elétrons, processos de oxidação-redução, atividade dissulfeto-oxidorreductase, e ligação à flavina-adenina. Ainda em conformidade com esses dados, em relação à participação em vias metabólicas, está anotado o envolvimento na glicólise/gliconeogênese, no ciclo do ácido cítrico, e no metabolismo do piruvato (TDR Targets). Foram descritos quatro domínios para a proteína SG4727 relacionados à participação no metabolismo energético celular (Figura 15). Trata-se de um domínio de ligação FAD/NAD (SSF51905) (A), um domínio de piridina nucleotídeo dissulfeto oxidorreductase (PF07992) (B), um domínio de dimerização, de redução associado à FAD/NAD (SSF55424) (C), um domínio de dimerização, de piridina nucleotídeo dissulfeto oxidorreductase (PF02852) (D).



B



C



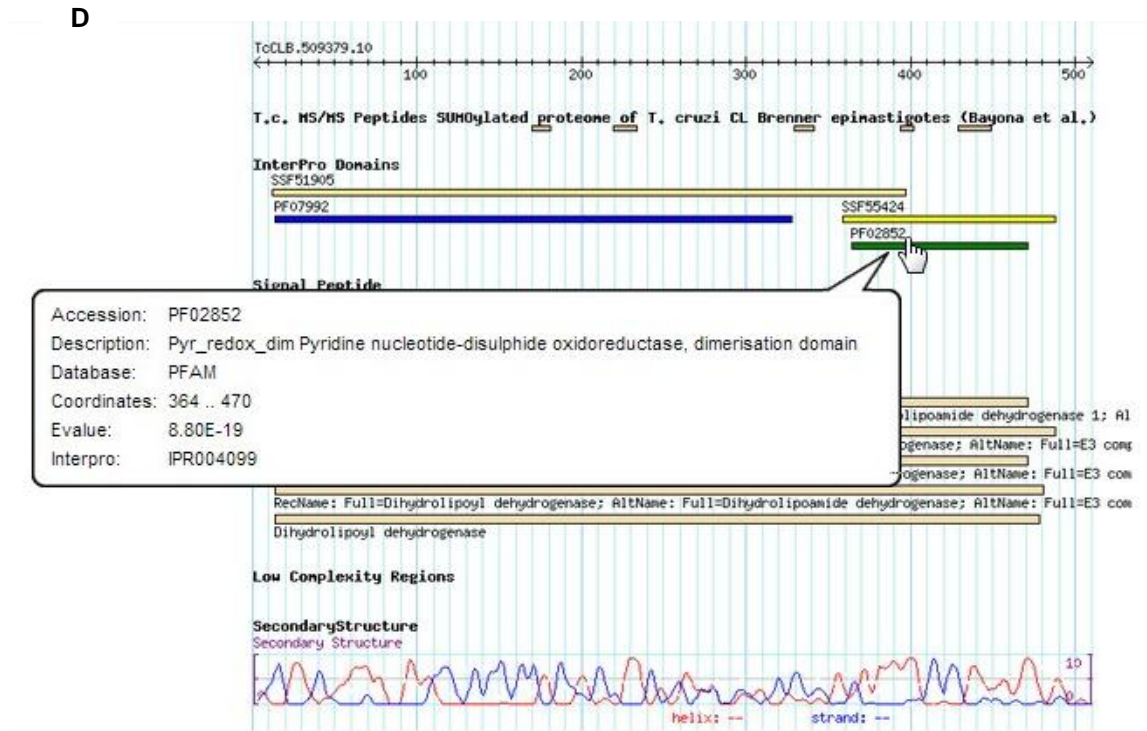


Figura 15- Domínios descritos para a proteína SG4727 (<http://tritrypdb.org/tritrypdb/>).

A LipDH também promove a redução de nitrofuranos, gerando radicais ânion superóxido e/ou peróxido de hidrogênio. É nesta via metabólica que são produzidos precursores importantes de bases nitrogenadas e de alguns aminoácidos (de GODOY *et al.*, 2012; SANTOS, 2008). Há dados referentes à participação da proteína em alguns processos metabólicos que estão em conformidade com esses dados, como o envolvimento no metabolismo da alanina, aspartato, glutamato, serina, leucina e degradação da isoleucina (TDR Targets).

Os Kinetoplastida são conhecidos pela maior sensibilidade ao estresse oxidativo em comparação com seus hospedeiros, devido à ausência da glutathione peroxidase e da catalase (SCHONECK *et al.*, 1997). Em função desta sensibilidade, eles tiveram que desenvolver sistemas antioxidantes específicos. Em contraste com os outros organismos, eles perderam a glutathione reductase (em procariotos e demais eucariotos esta enzima é responsável pela redução do meio intracelular) e adquiriram a tripanotona reductase (TR), cujo substrato é o dissulfeto espermidina glutathionil (CHEMIN *et al.*, 2001). Em função da sensibilidade ao estresse oxidativo, é importante que os Kinetoplastida possuam o metabolismo antioxidante do tiol (LOHRER; SIEGEL, 1990).

As drogas usadas no tratamento da Doença de Chagas atuam na manutenção do ciclo redox para a consequente produção de espécies reduzidas de oxigênio (LOHRER; SIEGEL, 1990). Há diferenças estruturais entre a LipDH do parasita e do hospedeiro, tornando a enzima um bom alvo para a quimioterapia. Além disto, ela é capaz de sequestrar radicais livres do citoplasma e da membrana plasmática, assim, a sua inibição pode interferir no mecanismo antioxidante do parasita. E, a LipDH reduz naftoquinonas e derivados de nitrofuranos em um elétron, contribuindo para a produção de radicais ânion superóxido (O_2^-) e peróxido de hidrogênio (H_2O_2) (SANTOS, 2008; ROLDÁN; COMINI; CRISPO; SIEGEL, 2011).

Assim, referente à função, acredita-se que a proteína SG4727 tenha grande importância celular por ser uma diidrolipoamida desidrogenase, enzima envolvida no metabolismo energético celular, com participação na homeostase redox celular e; envolvimento no metabolismo de precursores de bases nitrogenadas e de aminoácidos, fundamental para os processos de duplicação e transcrição gênica; assim como o processo de tradução. A Ontologia Gênica corrobora a função energética ao anotar a proteína como sendo de ligação à FAD e participação em mecanismos de oxirredução. A Homologia Gênica indica tratar-se de uma proteína conservada, sendo em alguns grupos também identificada como uma diidrolipoamida desidrogenase. As funções já descritas estão anotadas nas vias metabólicas nas quais a proteína participa, como na glicólise/gliconeogênese, ciclo do ácido cítrico, metabolismo do piruvato e metabolismo de aminoácidos. E, a enzima é um alvo interessante para o desenvolvimento de quimioterápicos visando a sua inibição, uma vez que há diferenças estruturais entre a enzima do parasita e de seus hospedeiros, além da sua inibição poder interferir no mecanismo antioxidante do parasita. Fato que chamou a atenção foi a sua expressão diminuída em aderido 12 horas, sendo que no momento faltam informações mais precisas que possam explicar essa regulação negativa.

SG5427 - Tc00.1047053510183.40 (hypothetical protein, conserved)

A proteína SG5427 é uma proteína conservada. Conforme a Homologia Gênica, a proteína possui genes parálogos e ortólogos, tendo como produtos, na

maioria dos casos, proteínas hipotéticas e conservadas. Sua função é desconhecida, visto que não foram encontradas informações referentes a domínios proteicos, Ontologia Gênica e participação em vias metabólicas. Assim, esta proteína é dita hipotética e conservada.

4.2 META 2: CONSTRUÇÃO E AMPLIFICAÇÃO DOS CASSETES DE NOCAUTE

Foram adotadas algumas técnicas em biologia celular e molecular para a construção e amplificação dos cassetes de nocaute gênico. Para a construção dos cassetes de nocaute gênico foi primeiramente necessária a amplificação dos genes de resistência e das IRS de cada gene alvo. Os cassetes de nocaute gênico foram construídos por PCR de fusão. E, os cassetes de nocaute também foram amplificados por PCR. A confirmação da especificidade e do tamanho dos fragmentos amplificados foi realizada através de eletroforese em gel de agarose.

Em relação à amplificação dos genes de resistência, das IRS e dos cassetes de nocaute, foi realizada a PCR convencional (Figura 16). A *Polimerase Chain Reaction* (PCR) – Reação em Cadeia da Polimerase - consiste na rápida replicação do DNA *in vitro*. A técnica baseia-se em ciclos de síntese de DNA, onde cada ciclo consiste em três etapas: 1) desnaturação da dupla fita de DNA; 2) anelamento dos oligonucleotídeos iniciadores (*primers*); e 3) extensão dos *primers*. Os reagentes necessários para uma reação de PCR são: tampão, nucleotídeos, *primers*, enzima (Taq DNA polimerase) e ácido nucleico de interesse. Os *primers* são cadeias curtas de nucleotídeos complementares à região inicial (*primer forward*) e à região final (*primer reverse*) da sequência alvo e delimitam a região a ser amplificada. Em uma mesma reação, *primers* e sequência alvo (a qual deve estar na forma desnaturada, ou seja, fita simples) se unem por complementaridade entre suas bases. Quando o complexo é formado, *primer* e sequência alvo, a enzima DNA polimerase se liga ao grupo OH da extremidade do carbono C3' do *primer* (extremidade 3' OH) para dar início ao processo de síntese. Assim, a DNA polimerase utiliza a sequência alvo como molde e, de acordo com a complementaridade entre as bases, adiciona novos nucleotídeos aos *primers*. Este processo recebe o nome de extensão. Ao final da extensão, uma nova molécula, complementar à sequência molde, é produzida (REZENDE *et al*, 2002; GRIFFITHS *et al.*, 2002)

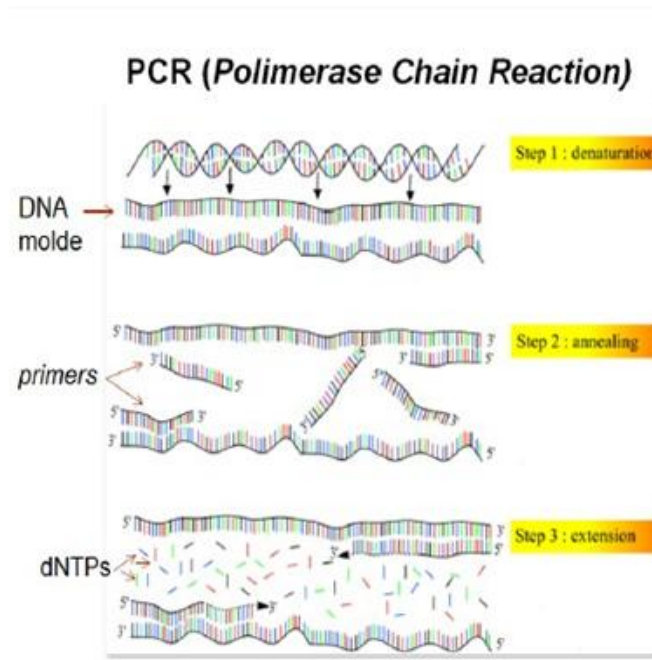


Figura 16- Polimerase Chain Reaction (PCR) (adaptado: ANDY VIERSTRAETE, 1999).

A confirmação do tamanho e especificidade dos fragmentos de DNA foi realizada por eletroforese em gel de agarose. Trata-se de uma técnica de separação de partículas de acordo com a carga elétrica e peso molecular. Moléculas de menor peso molecular tem mais facilidade de migrar pelo gel do que as moléculas de maior peso. A eletroforese permite a separação de RNA, DNA, proteínas e enzimas, pela migração destas em um gel durante a aplicação de um potencial elétrico. A distância que o fragmento percorre é comparada com um padrão de peso molecular que tem fragmentos de tamanhos conhecidos (GRUMANN, 2009).

4.2.1 Amplificação dos genes de resistência

Os genes de resistência a antibióticos foram utilizados como marcadores de seleção nos parasitas nocauteados. Assim, esses foram amplificados para a posterior construção dos cassetes de nocaute gênico. Os genes de resistência à NEO e à HIGRO foram amplificados com sucesso (Figura 17).

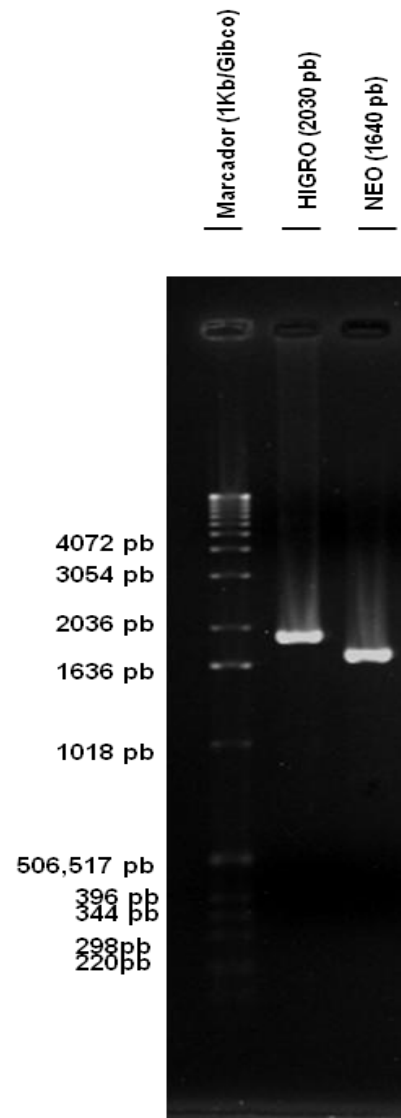


Figura 17- Amplificação dos genes de resistência à NEO e à HIGRO

A reação foi altamente específica para ambas as resistências, pois apresentou especificidade e tamanho esperado para os dois fragmentos (1640 pb para NEO e 2030 pb para HIGRO) não havendo necessidade de otimização da reação. Os reagentes e o programa de PCR utilizado para a amplificação das resistências estão descritos no item 3.4.2 da Metodologia.

4.2.2 Amplificação das IRS flanqueantes aos genes alvo

As IRS flanqueantes aos quatro genes alvo foram amplificadas, uma vez que estas sequências foram posteriormente utilizadas para a construção dos cassetes de nocaute gênico. As IRS foram amplificadas de forma satisfatória para os quatro genes (Figura 18).

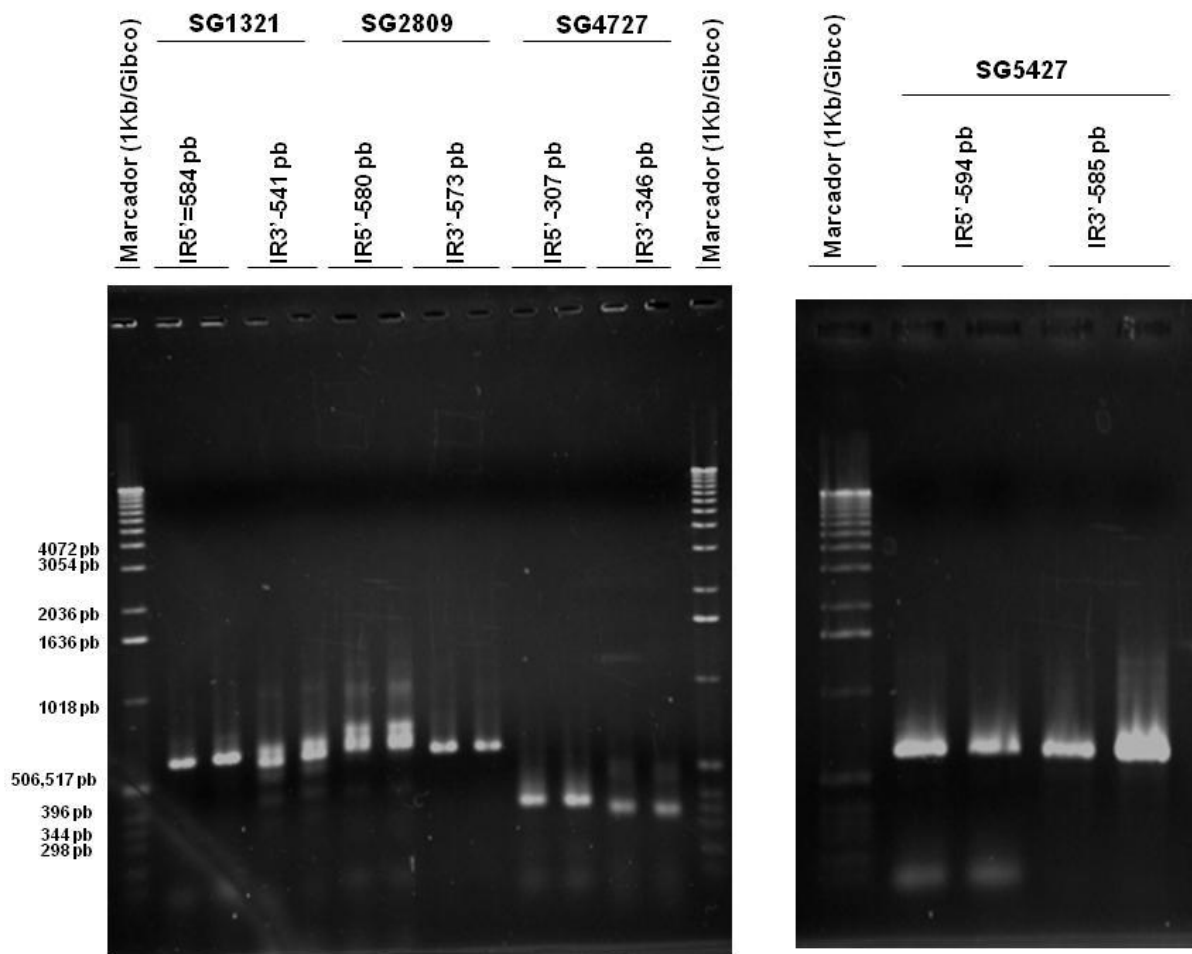


Figura 18 - Amplificação das IRS dos genes: SG1321, SG2809, SG4727 e SG5427

Observa-se que para alguns genes a amplificação das IRS foi altamente específica, como no caso de SG4727 e SG5427. Para os genes SG1321 e SG2809, a reação não foi tão específica. Não foi realizada a otimização desta PCR visto que, teoricamente, na PCR de fusão ocorre a fusão do cassete de resistência às

intergênicas do gene alvo. Os *primers* das IRS e os *primers* internos dos cassetes de resistência foram desenhados para haver complementaridade entre estes fragmentos e assim, construir o cassete de nocaute gênico. E, mesmo que houvesse algum problema nesta construção, teoricamente a inserção do cassete no genoma de *T. cruzi* ocorre por recombinação das IRS flanqueadoras do gene alvo existentes no cassete e no genoma do parasita. Estas razões nos desviaram da tentativa de otimizar a amplificação das intergênicas. Os reagentes e o programa de PCR utilizado para a amplificação das IRS estão descritos no item 3.4.3 da Metodologia.

4.2.3 Obtenção dos cassetes de nocaute gênico

O nocaute gênico, seguido da análise dos fenótipos resultantes, é uma estratégia poderosa para esclarecer a função de um determinado gene. Modificar genes escolhidos, substituindo-os por fragmentos de material genético externo ou inativando-os para que as proteínas deles derivadas não sejam produzidas, é hoje uma das estratégias mais promissoras da pesquisa biológica. Nocautear um gene é provocar nele uma alteração específica que leva a sua inativação ou, removê-lo completamente substituindo-o ou não por um gene de interesse (KELLY *et al.*, 1992; TAYLOR; KELLY, 2006). O nocaute gênico se dá a partir da transfecção de cassetes de nocaute. Estes podem ser construídos por PCR de fusão, que é um método rápido, relativamente barato e oferece a liberdade de escolha entre uma variedade de marcadores selecionáveis, permitindo construir cassetes precisos de qualquer estrutura (AMBERG; BOTSTEIN; BEASLEY, 1995). Um cassete para nocaute gênico está ilustrado na Figura 19.



Figura 19- Cassete para nocaute gênico após ser fusionado (Adaptado de LORUSSO, 2013).

O cassete de nocaute para transfecção contém as regiões intergênicas (*intergenic regions*, IRS) à montante (*upstream*, 5') e à jusante (*downstream*, 3') das sequências codificadoras (*coding sequences*, CDS) do gene alvo, flanqueando um gene de resistência a antibiótico (e.g. neomicina, higromicina, etc.). Com a transfecção ocorre a homologia das IRS do cassete com as IRS do gene alvo e assim, a substituição da CDS pelo gene de resistência. Durante a recombinação, o reconhecimento das extremidades é tão preciso que nenhuma alteração ocorre na sequência das bases nucleotídicas nos sítios de ligação (KELLY *et al.*, 1992; TAYLOR; KELLY, 2006) (Figura 20).

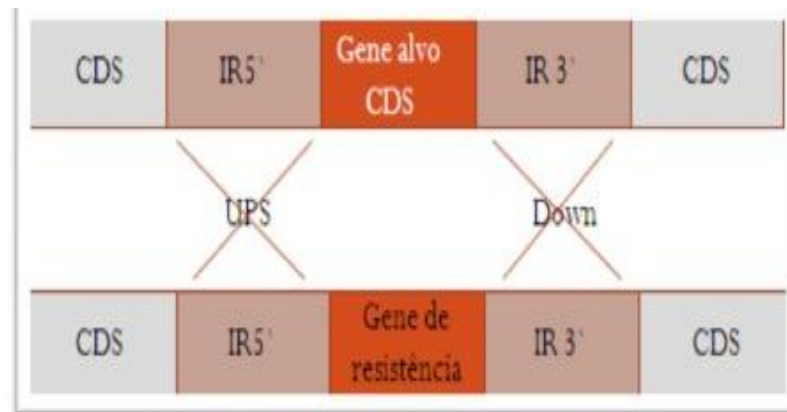


Figura 20 - Esquema da recombinação homóloga entre cassete de nocaute gênico e genoma de *T. cruzi* (SANTOS, 2012).

Foram obtidos os cassetes de nocaute gênico contendo os genes de resistência à NEO e o gene de resistência à HIGRO para os quatro genes, totalizando oito cassetes de nocaute gênico (Figura 21).

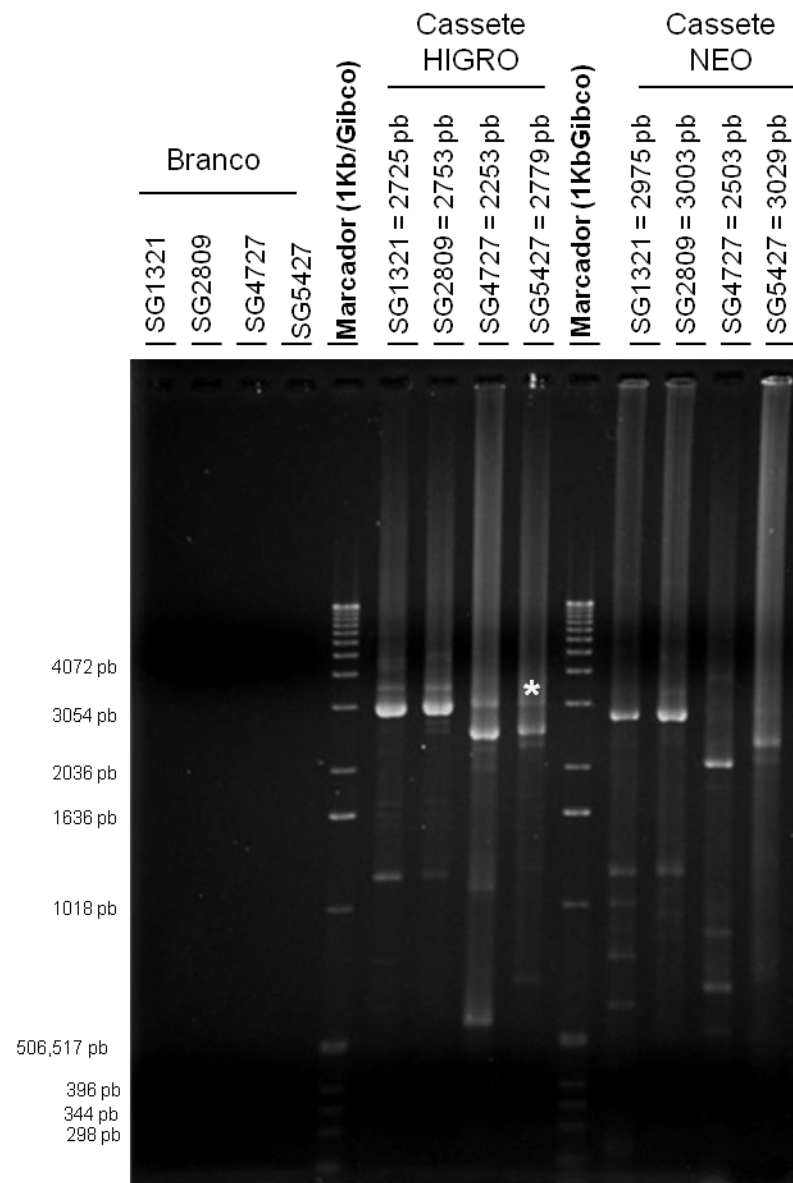


Figura 21- Obtenção dos cassetes de nocaute gênico contendo os genes de resistência à NEO e à HIGRO para os genes SG1321, SG2809, SG4727 e SG5427.

O tubo branco, preparado para cada gene, corresponde à adição de todos os reagentes da PCR com exceção dos *primers* e teve o intuito de verificar a ausência de contaminação dos mesmos. E, verificou-se que os *primers* estavam livres de contaminação. A reação de fusão não foi tão específica, visto que os cassetes de nocaute gênico contendo os genes de resistência à NEO e à HIGRO apresentaram mais de uma banda. Tentativas de otimização foram realizadas para eliminar as bandas inespecíficas, mas sem sucesso. Com exceção do cassete contendo o gene de resistência à HIGRO para o gene SG5427, marcado na Figura 20 com um

asterisco, todos os demais cassetes de nocaute apresentaram a banda de interesse como sendo a mais intensa. De maneira análoga ao discutido sobre a amplificação das IRS, não foram realizadas mais tentativas para otimizar a reação, uma vez que apenas os fragmentos de DNA contendo as IRS flanqueadoras do gene alvo recombinariam com as IRS do genoma de *T. cruzi*. E, além disto, a amplificação dos cassetes de nocaute gênico usariam os *primers* das IRS (*primer forward* da IR5' e *primer reverse* da IR3'), assim, apenas os fragmentos contendo as duas IRS em cada cassete amplificariam. Os reagentes e o programa de PCR utilizado para a construção dos cassetes estão descritos no item 3.4.5 da Metodologia.

4.2.4 Amplificação dos cassetes de nocaute gênico

Os cassetes de nocaute gênico foram amplificados visando a obtenção de massa necessária para a transfecção em *T. cruzi*. Foram amplificados com sucesso os cassetes contendo o gene de resistência à NEO para os genes SG1321, SG2809, SG4727 e SG5427 (Figura 22). E, entre os cassetes contendo o gene de resistência à HIGRO foi obtida a amplificação para os genes SG1321 e SG2809 (Figura 23).

O produto de amplificação dos cassetes de nocaute gênico contendo o gene de resistência à NEO e à HIGRO apresentou mais de uma banda em todos os casos, sendo a banda mais intensa correspondente ao cassete de nocaute gênico. As amplificações com bandas de cassetes um pouco menos intensas, como o cassete contendo o gene de resistência à NEO para o gene SG4727, necessitaram de maior número de reações para obter a massa de 50 µg de cassete, indicada para a transfecção. Algumas tentativas de otimização foram realizadas para eliminar as bandas inespecíficas, mas não houve sucesso. Não houve a preocupação em purificar a banda do gel ou dispendar mais tempo nas tentativas de otimização, visto que teoricamente a inserção do cassete de nocaute gênico ocorre através de recombinação homóloga com as IRS do gene alvo no genoma do parasita. Os reagentes e o programa de PCR utilizado para a amplificação dos cassetes estão descritos no item 3.4.6 da Metodologia.

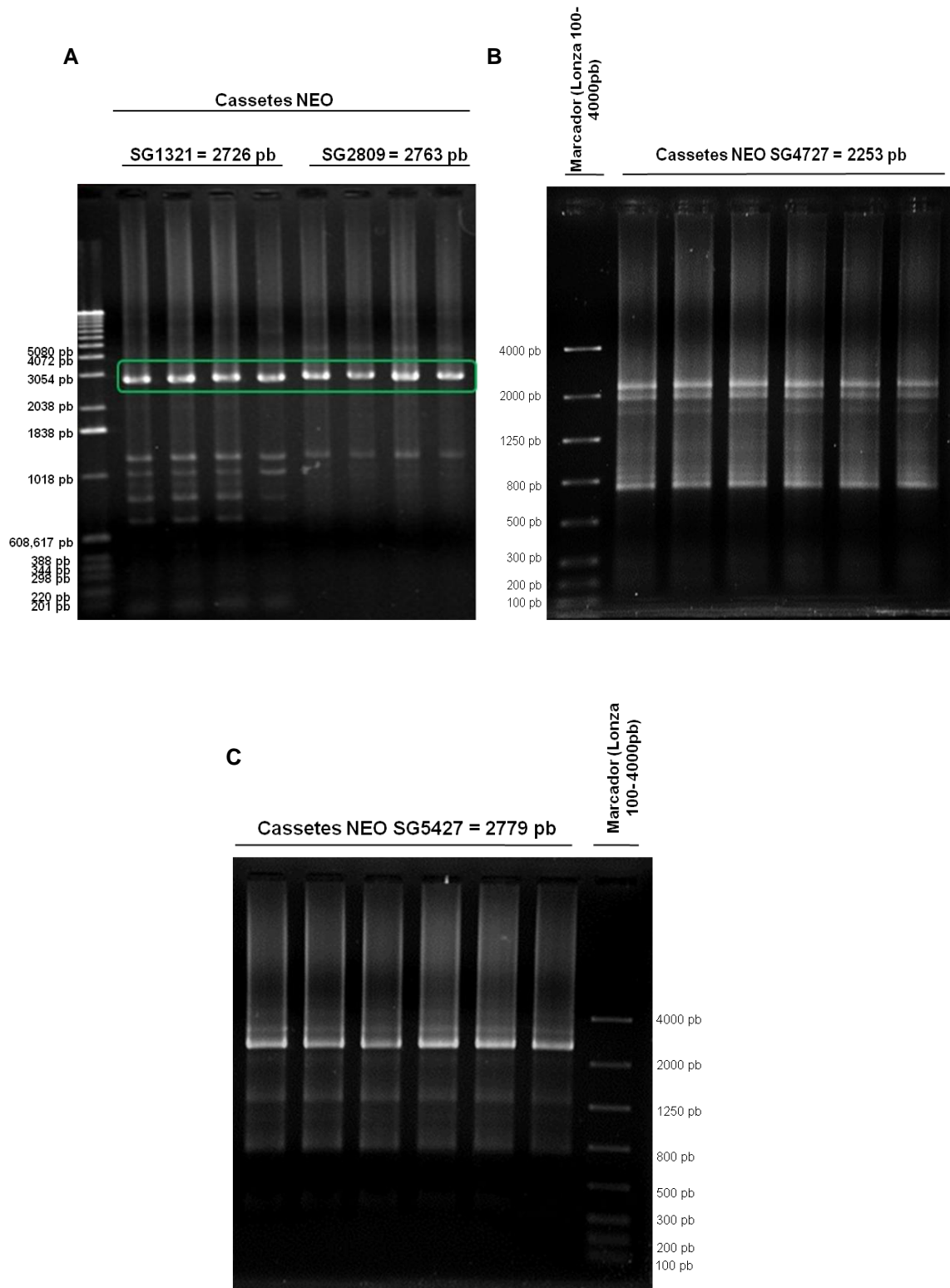


Figura 22- Amplificação dos cassetes de nocaute gênico contendo o gene de resistência à NEO para os genes SG1321, SG2809, SG4727 e SG5427. Em (A) está a amplificação dos cassetes de nocaute para os genes SG1321 e SG2809. Em (B) está a amplificação dos cassetes de nocaute para o gene SG4727 e, em (C) está a amplificação dos cassetes de nocaute para o gene SG5427.

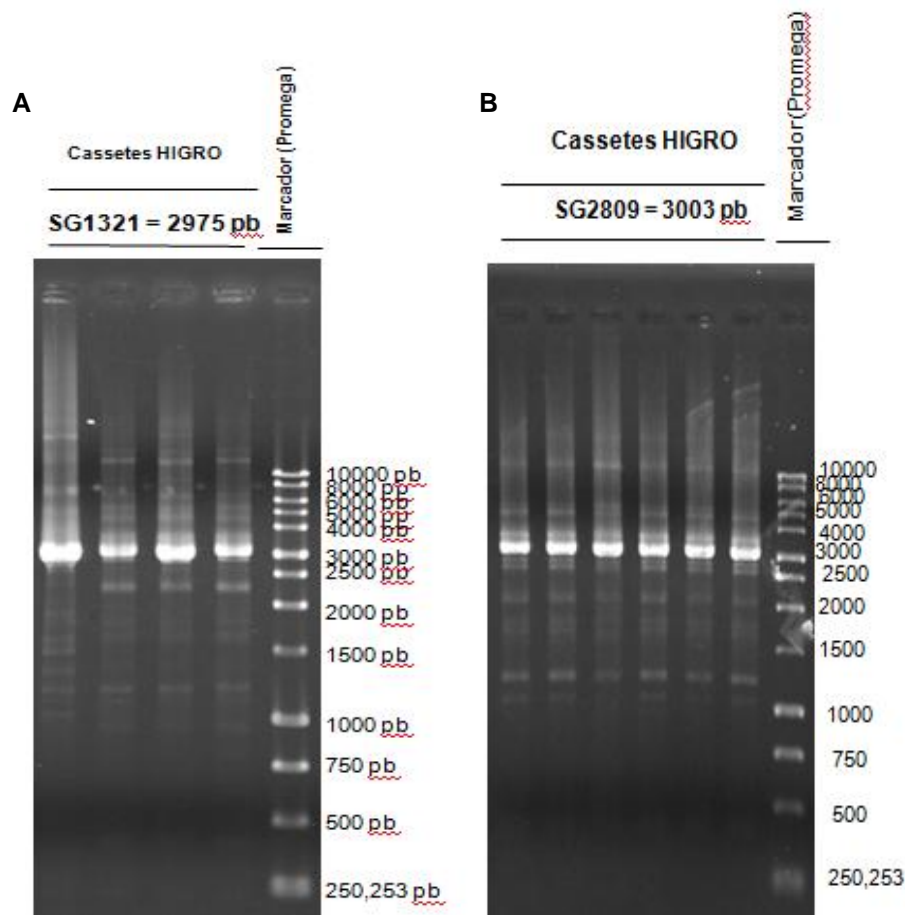


Figura 23 - Amplificação dos cassetes de nocaute gênico contendo o gene de resistência à HIGRO para os genes SG1321 (A) e SG2809.

Após a amplificação dos cassetes, esses foram quantificados para obtenção de massa para transfecção em *T. cruzi*.

A quantificação foi realizada na plataforma Qubit™ Fluorometer (Invitrogen™) utilizando o kit Quant-it™ dsDNA BR Assay, o qual detecta concentrações de DNA entre 0,1 e 100 ng/μL, de acordo com as instruções do fabricante. Esta técnica de quantificação apresenta alta sensibilidade e acurácia. Este aparelho é ideal para quantificar DNA por fluorometria, que é uma técnica denominada espectrofluorometria, um tipo de espectroscopia eletromagnética a qual analisa a fluorescência de uma amostra. Isto envolve o resultado da absorção de energia radiante e emissão de parte desta energia na forma de luz de baixa energia, mas não necessariamente luz visível. A emissão de luz se dá pela excitação de elétrons nas moléculas de certos compostos, normalmente luz ultravioleta. O kit conta com reagente e tampão. O reagente serve para emitir luz a ser contabilizada

pelo aparelho Qubit. A quantificação deve ser realizada na temperatura ambiente e os tubos, a serem quantificados, protegidos da luz para evitar a degradação do reagente.

4.3 META 3: OBTENÇÃO DOS PARASITAS NOCAUTE SIMPLES E DUPLO

Inicialmente, foram construídos os cassetes de nocaute gênico com NEO para todos os genes estudados, os quais foram transfectados em formas epimastigotas cepa Dm 28c de *T. cruzi* selvagens e resultaram na obtenção de cepas simples nocaute em todos os casos. Em média, os parasitas sobreviventes foram selecionados ao final de 4 semanas. Como controle manteve-se uma cultura selvagem, cuja morte se deu em torno de 21 dias.

Em paralelo, foram construídos os cassetes de nocaute gênico com HIGRO para os genes SG1321 e SG2809. Os parasitas simples nocaute com NEO para esses genes, previamente selecionados, foram usados para transfecção dos cassetes correspondentes com HIGRO e, assim, obtidos parasitas duplo nocaute. Após 24 horas foram adicionados os antibióticos de seleção NEO e HIGRO. Mais uma vez, os parasitas sobreviventes foram selecionados ao final de 4 semanas e a morte dos controles se deu em torno de 21 dias.

Os cassetes com HIGRO para os genes SG4727 e SG5427 foram construídos, mas ocorreram problemas em sua amplificação para obter a massa necessária para transfecção. Várias tentativas de otimização foram efetuadas, mas sem sucesso. Isso inviabilizou a obtenção de parasitas duplo nocaute para estes dois genes. Assim, nessa etapa foram obtidos parasitas simples e duplo nocaute para os genes SG1321 e SG2809 e parasitas simples nocaute com NEO para os genes SG4727 e SG5427.

4.4 META 4: AVALIAÇÃO DOS EFEITOS DO NOCAUTE GÊNICO

A avaliação dos efeitos do nocaute gênico foi realizada com parasitas simples nocaute contendo o gene de resistência à NEO para os genes SG1321 e SG2809, enquanto ocorria a seleção dos duplos nocaute. O intuito foi verificar se a supressão de apenas um alelo do gene provocaria alguma modificação na viabilidade celular; ou na capacidade proliferativa e de diferenciação dos parasitas, o que seria interessante para indicar a importância dessas proteínas para o parasita, particularmente caso os parasitas duplo nocaute se mostrassem inviáveis.

4.4.1 Análise da viabilidade celular dos parasitas nocaute

Os parasitas simples nocaute para todas as quatro proteínas se mostraram viáveis ao sobreviver mesmo com a supressão de um alelo do gene. As análises foram feitas a partir de seleção dos parasitas em meio contendo o antibiótico de seleção e observações dos parasitas em microscópio óptico. O mesmo aconteceu para os parasitas duplo nocaute para os genes SG1321 e SG2809, o que nos levou a crer que nenhuma das quatro proteínas é essencial à sobrevivência do parasita. No entanto, foram encontrados problemas na confirmação do nocaute do gene SG1321, conforme será discutido a seguir, sugerindo que essa hipótese precisa ser confirmada com outros experimentos.

4.4.2 Análise da capacidade proliferativa dos parasitas simples nocaute

Para ambas as proteínas, as curvas de crescimento indicaram leve diferença entre parasitas simples nocaute e parasitas selvagens, com leve redução na capacidade proliferativa. Esse resultado parece indicar que a supressão de um dos

alelos nos genes SG1321 e SG2809 é suficiente para afetar negativamente a capacidade proliferativa dos parasitas (Figura 24).

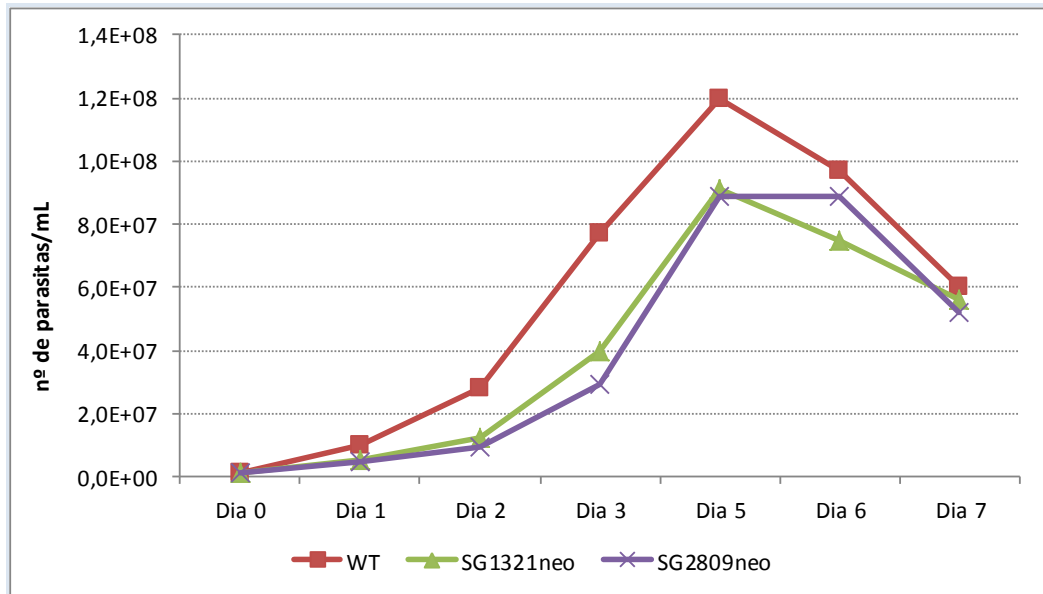


Figura 24 - Curva de crescimento dos parasitas simples nocaute para os genes SG1321 e SG2809.

Apesar desta leve diferença na capacidade proliferativa entre parasitas selvagens e simples nocaute, ambos atingem a fase estacionária ao mesmo tempo, por volta do quinto dia, embora não com a mesma densidade populacional, a partir do qual a população começa a decair.

O resultado é promissor visto se tratar da supressão de apenas um alelo para cada um dos genes. No entanto, há necessidade da repetição do experimento com maior número de réplicas biológicas para estes parasitas simples nocaute, além da realização com parasitas duplo nocaute para confirmar, de fato, a diminuição ou, supressão da capacidade proliferativa quando da supressão total do gene.

4.4.3 Análise da capacidade de diferenciação dos parasitas simples nocaute

Os parasitas simples nocaute apresentaram menor capacidade de diferenciação em tripomastigotas metacíclicos em comparação aos parasitas selvagens. Em 72 h houve uma sobreposição na variação dos dados entre os selvagens e os parasitas simples nocaute para o gene SG2809. Em 96 h esta sobreposição não ocorreu, sendo mais evidente a diminuição na capacidade de diferenciação para ambos os genes (Figura 25). Em acréscimo, os parasitas simples nocaute pareceram ter menor capacidade de adesão, pois foram encontradas mais células no sobrenadante dessas culturas em relação ao selvagem. Este dado é importante e possui relação com a diminuição da capacidade de diferenciação, uma vez que a adesão é uma das etapas necessárias durante a diferenciação de formas epimastigotas em tripomastigotas metacíclicos.

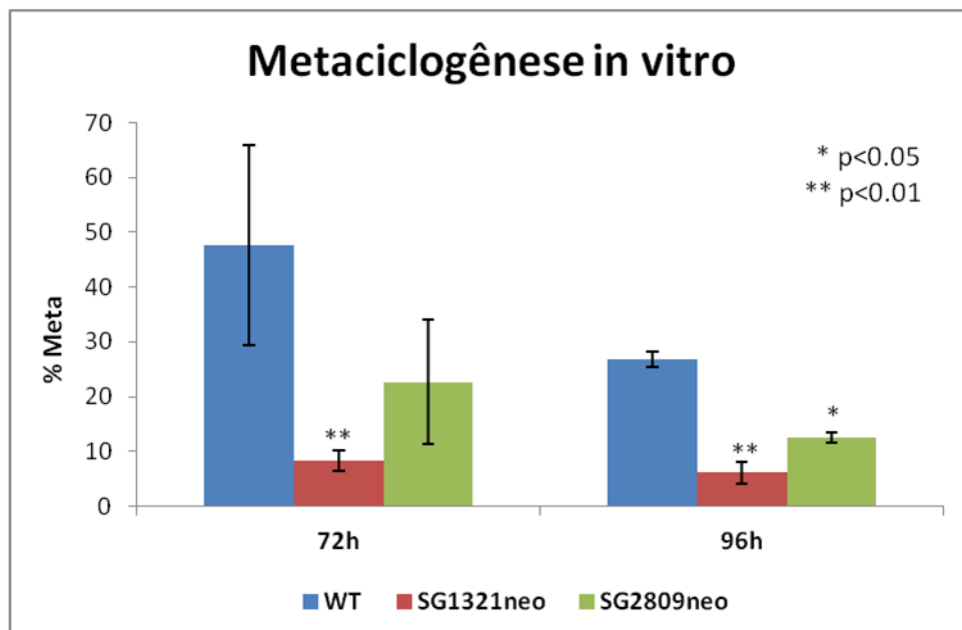


Figura 25 - Avaliação da capacidade de diferenciação dos parasitas simples nocaute para os genes SG1321 e SG2809. Os dados foram analisados estatisticamente através do teste *t-Student* e são representativos de três experimentos independentes.

Como verificado anteriormente para os genes SG1321 e SG2809, o resultado é igualmente promissor visto também se tratar da supressão de apenas um alelo para cada um dos genes. Há igualmente necessidade da repetição do experimento para estes parasitas simples nocaute, além da realização com parasitas duplo nocaute para confirmar, de fato, a diminuição ou, supressão da capacidade diferenciação quando da supressão total do gene.

4.5 META 5: CONFIRMAÇÃO DA OBTENÇÃO DE PARASITAS NOCAUTE

Os cassetes de nocaute gênico, em teoria, se inserem no genoma de *T. cruzi* por recombinação homóloga das regiões intergênicas dos cassetes com as intergênicas dos genes alvo no genoma, promovendo a inserção do gene de resistência no lugar da região codificadora e consequente deleção do gene a ser estudado.

Foi realizada a transfecção dos cassetes com NEO, selecionados os parasitas a partir de culturas contendo o respectivo antibiótico existente no cassete, e obtidos inicialmente parasitas simples nocaute. Estes foram transfectados com o segundo cassete de nocaute, selecionados em culturas contendo ambos os antibióticos e, assim, foram obtidos teoricamente parasitas duplo nocaute. Mas, na prática, o fato dos parasitas duplo nocaute terem sido selecionados em meio contendo NEO e HIGRO indica apenas que houve a incorporação dos cassetes contendo as duas resistências pelos parasitas, mas não garante a introdução dos cassetes nos locais desejados com a deleção dos genes de interesse. Além da possível incorporação do cassete em local diferente da CDS, há ainda a possibilidade da CDS não ser cópia única no genoma de *T. cruzi* (contrariando a predição realizada) ou ter sido duplicada no momento da tentativa de depleção, o que torna necessária a confirmação da obtenção dos nocautes gênicos.

Para confirmar a ocorrência de duplos nocautes, inicialmente alguns testes foram realizados com as populações mistas de parasitas transfectantes. No entanto, réplicas geraram resultados contraditórios e, por isso, adotou-se a metodologia do *Cell Sorting* para obtenção de populações clonais de parasitas duplo nocaute para os genes SG1321 e SG2809.

4.5.1 Isolamento de populações clonais através de *Cell Sorting*

Objetivando isolar populações clonais para estudos mais aprofundados, foi feito o *Cell Sorting* em placas de 96 poços. Para o dKO da proteína SG1321, foram obtidas 61 colônias, enquanto para o dKO da proteína SG2809, foram obtidas 72 populações clonais.

4.5.2 PCRs de confirmação dos nocautes

Primeiramente, foram realizados testes para confirmar a inserção dos genes de resistência no genoma dos parasitas nocaute. Para tanto, utilizaram-se os *primers forward e reverse* para os cassetes de resistência, assim como para os genes de resistência em si. Conforme esperado, os parasitas duplo nocaute para os genes SG1321 e SG2809 apresentaram as duas bandas específicas correspondentes aos cassetes de resistência contendo os genes de resistência à NEO e à HIGRO (Figuras 26 e 27), demonstrando que os parasitas testados possuíam no genoma ambos os cassetes de resistência.

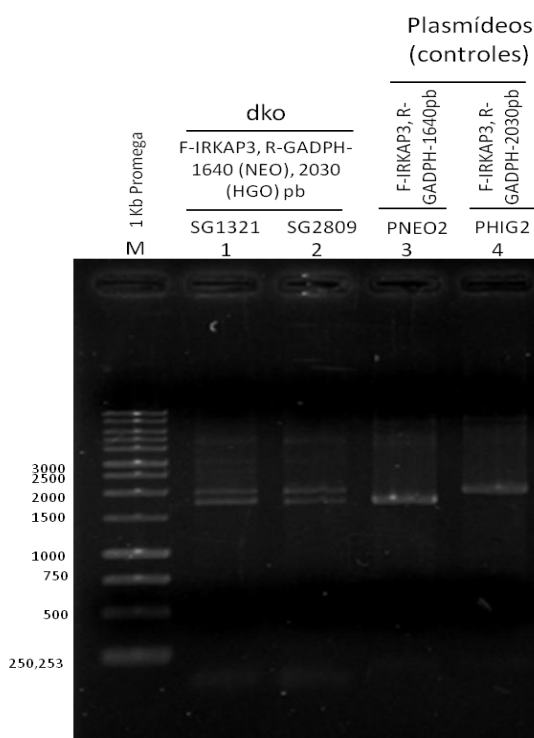


Figura 26 - Amplificação dos cassetes de resistência em parasitas dKO para os genes SG1321 e SG2809. Foram utilizados os *primers* dos cassetes de resistência, sendo F – *primer forward* e R – *primer reverse*. Como controle utilizaram-se os plasmídeos usados como DNA-molde para amplificação dos cassetes de resistência.

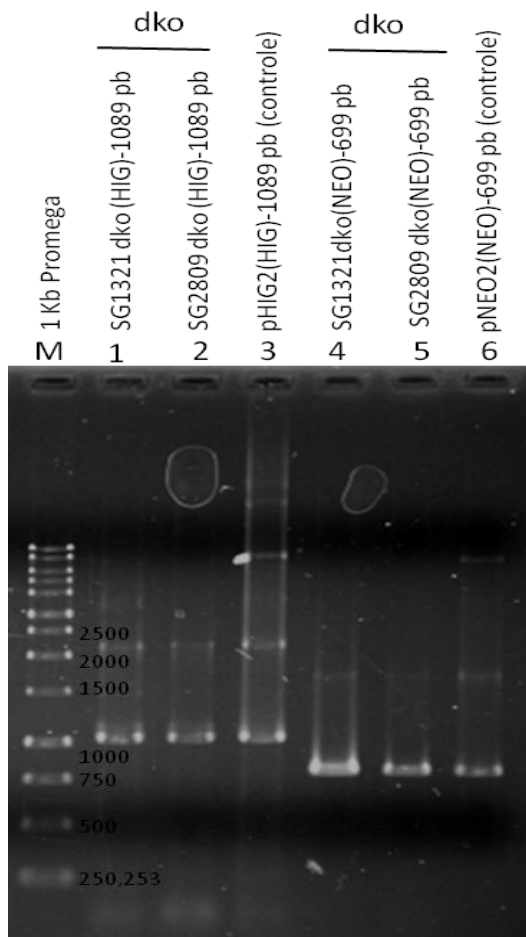


Figura 27 - Amplificação dos genes de resistência em parasitas dKO para os genes SG1321 e SG2809. Foram utilizados os *primers* dos genes de resistência, que estão indicados entre parentêses. Como controle empregaram-se os plasmídeos usados como DNA-molde para amplificação dos genes de resistência.

Ainda nesse contexto, em outro teste realizado com culturas mistas de parasitas duplo nocaute utilizaram-se os *primers* das IRS dos genes alvo associados com os *primers* dos genes de resistência (Figura 28). Os cassetes foram usados como controle para indicar a manutenção da integridade dos mesmos pós transfecção. Para ambos os genes houve a amplificação das bandas esperadas, tanto para os parasitas duplo nocaute quanto para as amostras utilizando os cassetes de nocaute gênico amplificados como molde. Este teste indica que a integridade dos cassetes foi mantida após a transfecção para os parasitas.

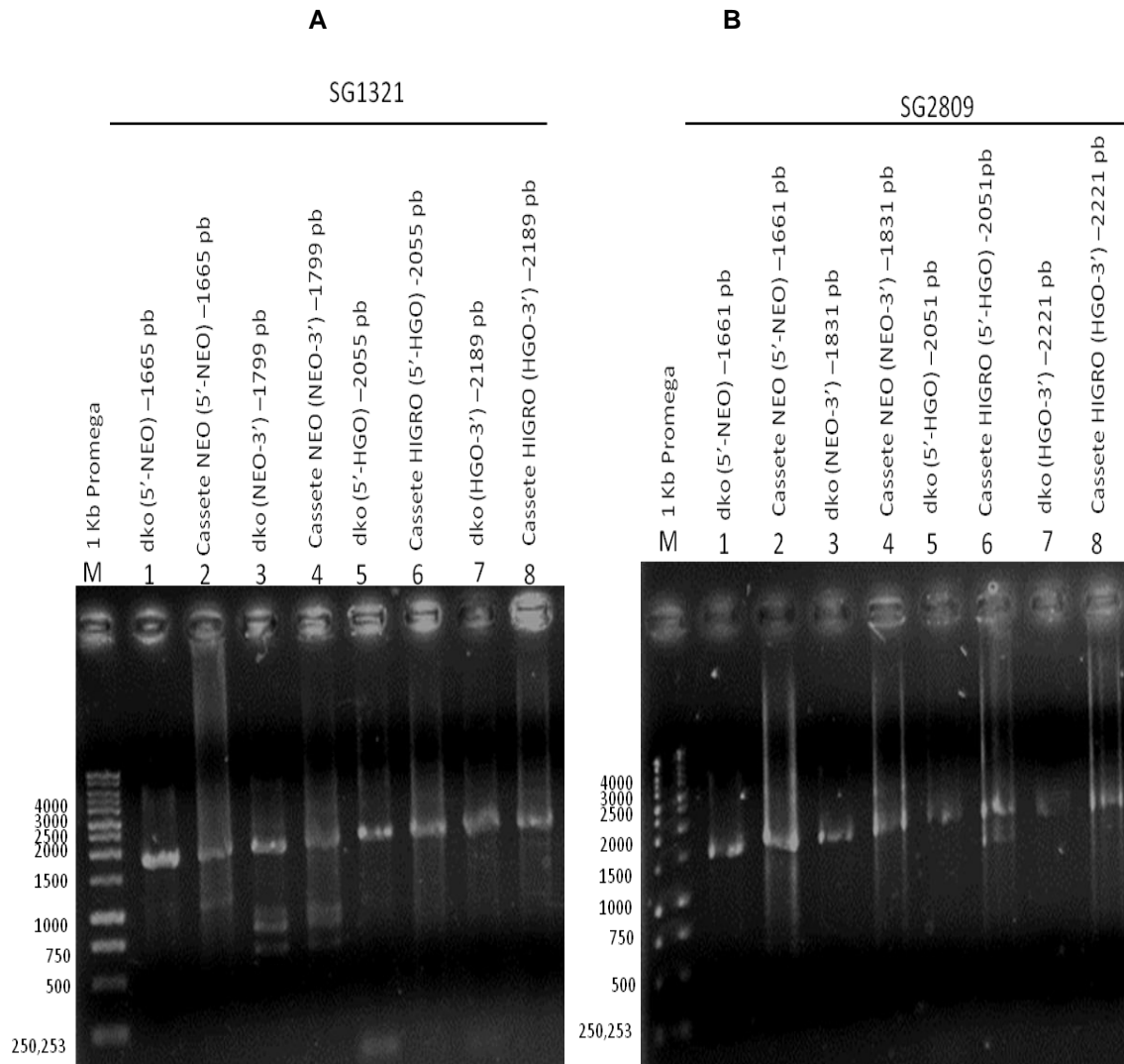


Figura 28- Confirmação da integridade dos cassetes de resistência transfectados em parasitas dKO para os genes SG1321 e SG2809. Foram utilizadas combinações de *primers* das IRS dos genes alvo com os genes de resistência. Os *primers* utilizados estão entre parênteses. O poço **M** é o marcador utilizado. Como DNA-molde controle foram usados os cassetes amplificados contendo os genes de resistência à NEO e os cassetes contendo os genes de resistência à HIGRO.

Após estes testes, realizou-se o *Cell Sorting* dos parasitas duplo nocaute para os genes SG1321 e SG2809. Foram selecionadas as colônias de parasitas nas posições A1e A2 da placa de *Cell Sorting*, para cada um dos genes e foram repetidos os testes anteriormente realizados, com adição do teste usando os *primers* das CDS (Figura 29). Sendo que, como teoricamente as CDS foram substituídas pelas resistências à NEO e à HIGRO, não era esperada amplificação das CDS para as proteínas em parasitas dKO. Entretanto, para os parasitas dKO para o gene SG1321 houve amplificação da CDS nas duas colônias (em destaque em A). Já para os parasitas dKO para o gene SG2809 não houve amplificação da CDS em nenhuma das duas colônias.

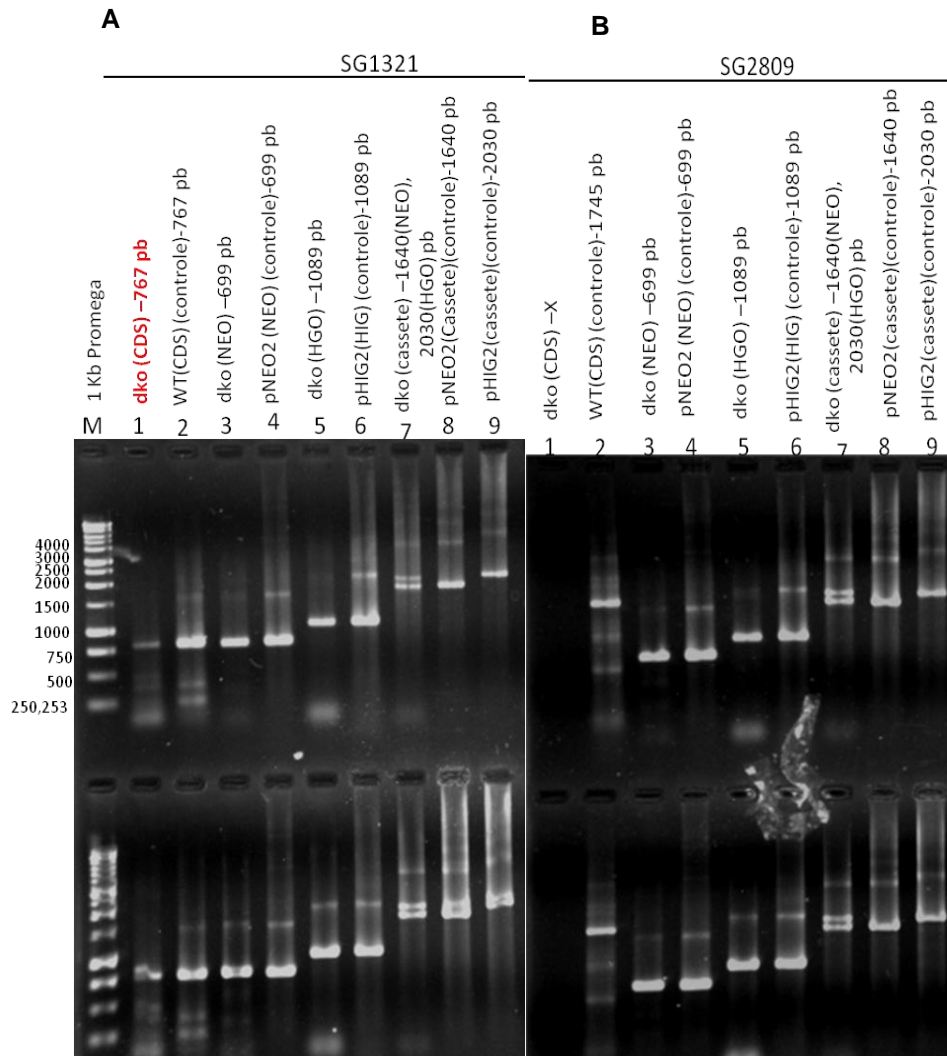


Figura 29 - Amplificação da CDS dos genes alvo em populações clonais de parasitas dKO para os genes SG1321 (A) e SG2809 (B). Na porção superior de cada gel está a colônia A1 e na porção inferior a colônia A2 da placa de Cell Sorting. Foram utilizados os *primers* das CDS cujo controle foi o parasita selvagem, *primers* dos genes de resistência cujos controles foram os plasmídeos, e *primers* dos cassetes de seleção cujos controles foram os plasmídeos. Observar em (A) a amplificação da CDS para as duas colônias (destaque na cor vermelha). Em (B) não aconteceu a amplificação da CDS em nenhuma das colônias. Os demais testes realizados tiveram os resultados esperados, conforme já tinham sido obtidos em testes anteriores.

Para confirmar o resultado anterior, realizou-se a amplificação das CDS para todas as populações de parasitas dKO clonais obtidas para ambos os genes.

Para o gene SG1321, as 61 populações clonais obtidas por *Cell Sorting* foram testadas. O experimento com as mesmas colônias foi repetido em cinco ocasiões, sendo que nas duas últimas as amostras foram preparadas em duplicatas. Para o gene SG1321, não foi confirmada a existência de parasitas duplo nocaute das colônias testadas. Para cada amostra houve a amplificação da CDS do gene alvo em pelo menos um dos testes. Em alguns casos observou-se a divergência entre as

duplicatas de uma mesma amostra, implicando não ter sido possível confirmar a depleção dos genes de interesse nesses parasitas. Tal fato pode ter decorrido, por exemplo, pela CDS (contrariando a predição realizada) não ser de cópia única no genoma do *T. cruzi*, pela inserção dos cassetes em local diferente da CDS do referido gene ou, ainda, o parasita pode ter feito mais cópias da CDS após a transfecção dos cassetes. No entanto, como foi observada variação de fenótipo – com a diminuição da capacidade proliferativa e de diferenciação nos parasitas simples nocaute para a proteína SG1321, essa continua sendo de interesse para estudo e esses experimentos deverão ser repetidos para confirmar a deleção do primeiro alelo da proteína e na tentativa de obter novos nocautes duplos.

Já para o gene SG2809, 72 populações clonais obtidas por *Cell Sorting* se mostraram viáveis para testes. Entre elas, cinco não tiveram amplificação da CDS em nenhum dos testes. Estas colônias que não tiveram amplificação foram testadas no mínimo três vezes, sendo pelo menos uma vez em duplicata. Assim, tudo parece indicar que nestas cinco colônias houve realmente a obtenção de parasitas duplo nocaute.

5 CONCLUSÃO

As quatro proteínas analisadas não parecem ser essenciais à sobrevivência do *T. cruzi*, uma vez que os parasitas simples nocaute com resistência à NEO e os parasitas duplo nocaute com resistência à NEO e HIGRO se mostraram viáveis. Apesar disto, as proteínas para os genes SG1321 e SG2809 parecem ser relevantes no processo de proliferação e diferenciação celular, uma vez que os parasitas simples nocaute com NEO sofreram redução na capacidade proliferativa e de diferenciação. Os dados observados sugerem que a supressão de apenas um dos alelos destes genes já provoca alterações na biologia do parasita.

Embora os resultados mostrem-se bastante promissores, há que se considerar que os ensaios funcionais foram realizados com parasitas simples nocaute. Oportuna será a repetição dos ensaios para efeitos do nocaute gênico com os parasitas simples nocaute e a realização com os parasitas duplo nocaute.

Igualmente verifica-se a necessidade da realização do *Cell Sorting* e confirmação prévia de parasitas duplo nocaute para os genes. A repetição dos experimentos deverá ser realizada com estas populações clonais, visto que para o gene SG1321, por exemplo, apesar de confirmada a inserção dos cassetes de nocaute para os dois alelos, houve a amplificação da CDS. Isto parece indicar que não foram obtidos parasitas duplo nocaute para o gene. Mas, apesar disso, se observou alteração do fenótipo. Esse resultado coloca em dúvida se, contrariando nossas predições e análises, a CDS não seja cópia única ou, se os cassetes se inseriram em uma região desconhecida do genoma; ou ainda, a possibilidade do parasita possuir algum mecanismo de compensação, como por exemplo, a duplicação da CDS em função do nocaute gênico. Já para o gene SG2809, foram obtidas cinco populações clonais sem amplificação da CDS. Isto parece indicar que foram obtidos parasitas duplo nocaute para o gene. Os ensaios referentes aos efeitos do nocaute gênico foram realizados com populações mistas. No entanto, teria sido mais interessante a realização dos mesmos com populações geneticamente idênticas de parasitas, para que os resultados obtidos de fato representassem os organismos analisados como um todo, sem variação entre eles.

Com exceção de uma proteína que apresentou um aumento gradativo ao longo da metaciclogênese, as demais proteínas foram reguladas de forma positiva

ou negativa em momentos específicos do processo. Esse aumento ou diminuição na expressão proteica pode ser indício do envolvimento das mesmas em processos de sinalização celular. A proteína SG2809 possui dois domínios RRM, assim, acredita-se que ela esteja relacionada a mecanismos pós-transcricionais de regulação da expressão gênica. O domínio RRM, além de indicar reconhecimento e ligação à molécula de ácido nucleico, também é indicativo de interação proteína-proteína. Desta maneira, é viável a realização de ensaios moleculares para a determinação de parceiros de interação. Estes ensaios podem somar informações relevantes no processo de caracterização dessas proteínas.

Há necessidade de mais informações referentes a domínios proteicos, uma vez que não foram encontrados dados sobre domínios para as proteínas SG1321 e SG5427. Para a proteína SG1321 foi identificado um peptídeo sinal. Isto a torna uma candidata interessante para o desenvolvimento de fármacos, visto que as proteínas que são secretadas ou que estejam expostas na superfície celular são alvos mais acessíveis para ação de drogas e vacinas. Além disto, a identificação desse peptídeo sinal é um dado relevante para predição da função do gene, uma vez que a função proteica não está relacionada apenas a sua função predita, mas também ao momento e ao local subcelular onde atua. Por isso, é viável a realização de ensaios moleculares para a determinação da localização subcelular do polipeptídeo.

A proteína SG4727 é uma diidrolipoamida desidrogenase, envolvida em mecanismos energéticos celulares e em reações de oxirredução celular. Durante o ciclo de vida de *T. brucei*, a sua localização celular pode ser mitocondrial ou fortemente associada à membrana plasmática, podendo estar localizada na superfície celular. Essa proteína também é uma candidata interessante para o desenvolvimento de fármacos visando a sua inibição, uma vez que há diferenças estruturais entre a enzima do parasita e de seus hospedeiros, da sua inibição poder interferir no mecanismo antioxidante do parasita, além de ser um alvo mais fácil de drogas quando está na superfície celular. Há necessidade de mais informações referentes à conservação da proteína em relação a outras espécies, como por exemplo, a conservação na espécie humana. Este tipo de informação é fundamental para análise da viabilidade de fármacos contra proteínas específicas do parasita.

Os dados obtidos nesse trabalho são promissores e preservam as proteínas como candidatas interessantes para o desenvolvimento de trabalhos futuros visando

a sua caracterização funcional, a partir da realização de outros ensaios moleculares e funcionais, além da exploração de outras ferramentas bioinformáticas.

6 PERSPECTIVAS

1. Referente à confirmação da obtenção de parasitas duplo nocaute para os genes SG1321 e SG2809:

- Repetir o *Cell Sorting*, realizar a PCR de confirmação com os *primers* das CDS, realizar *Southern Blotting* e, realizar PCR de confirmação usando *primers* dos genes à montante e à jusante do gene alvo;

2. Referente aos efeitos do nocaute gênico para os genes SG1321 e SG2809:

- Repetir os ensaios referentes à viabilidade celular, à capacidade de proliferação e à capacidade de diferenciação para os parasitas simples nocaute e, realizá-los com os parasitas duplo nocaute;

3. Referente aos ensaios moleculares e funcionais para os genes SG1321 e SG2809:

- Realizar ensaios sobre morfologia celular a partir da análise de lâminas de microscopia óptica, ensaios sobre parceiros de interação a partir de *Pull Down*, ensaios sobre localização subcelular a partir de clonagem, e ensaios sobre a capacidade infectiva a partir da infecção de camundongos com o parasita nocauteado;

4. Referente ao levantamento de dados e uso de ferramentas bioinformáticas:

- Realizar análises sobre Homologia Gênica em relação a outros grupos além dos tripanosomatídeos;

5. Referente aos genes SG4727 e SG5427, sendo que parasitas simples nocaute contendo o gene de resistência à NEO já foram obtidos e cassetes de nocaute gênico contendo o gene de resistência à HIGRO já foram construídos:

- Amplificar os cassetes de nocaute gênico contendo o gene de resistência à HIGRO, fazer a transfecção, seleção e confirmação da obtenção de parasitas duplo nocaute, assim como realizar a análise dos efeitos do nocaute gênico para parasitas simples nocaute e, na sequência, avaliar os efeitos do nocaute gênico para os parasitas duplo nocaute;
- Realizar os ensaios moleculares e funcionais – como análise da morfologia celular, da capacidade infectiva em células de mamíferos, sobre os parceiros de interação, sobre a localização subcelular;
- Fazer levantamento de dados e uso de ferramentas bioinformáticas referentes à Homologia Gênica em relação a outros grupos além dos tripanosomatídeos.

REFERÊNCIAS

ALBERTS, B.; BRAY, D.; LEWIS, J.; RAFF, M.; ROBERTS, K.; WATSON, J. D. **Biologia molecular da célula**. 5. Ed. Porto Alegre: Artmed, 2010.

AMBERG, D. C.; BOTSTEIN, D.; BEASLEY, E. M.: Precise Gene Disruption in *Saccharomyces cerevisiae* by Double Fusion Polymerase Chain Reaction. **Yeast**, vol. 11, n. 13, p. 1275-80, Oct. 1995.

ANDY VIERSTRAETE. **Molecular Tools and Infectious Disease Epidemiology**. Ed. Elsevier, 1999, USA.

ARAÚJO, P. R.; TEIXEIRA, S. M.: Regulatory elements involved in the post-transcriptional control of stage-specific gene expression in *Trypanosoma cruzi* - A Review. **Mem Inst Oswaldo Cruz**, v. 106, n. 3, p. 257-66. May, 2011.

CAMPOS, A. C. F.; BARBOSA, D. V. C.; LOBO, F.; MELO, R. C. M. de. Apostila de Bioinformática. Disponível em: <http://www.4shared.com/get/V_0zg7iz/apostila_de_bioinformatica_200.html>. Acesso em 30 de janeiro de 2013.

CDC – Centers for Disease Control and Prevention. Disponível em: <<http://www.cdc.gov/dpdx/trypanosomiasisAmerican/index.html>>. Acesso em 10 de fevereiro de 2014.

CHEMIN, L. S.; BUISINE, E.; YARDLEY, V.; KOHLER, S.; DEBREU, M. A.; LANDRY, V.; SERGHERAERT, C.; CROFT, S. L.; SIEGEL, R. L. K.; CHARVET, E. D. 2- and 3-Substituted 1,4-Naphthoquinone Derivatives as Subversive Substrates of Trypanothione Reductase and Lipoamide Dehydrogenase from *Trypanosoma cruzi*: Synthesis and Correlation between Redox Cycling Activities and in Vitro Cytotoxicity. **J. Med. Chem.** V. 15, n. 44, p. 548-65. Feb. 2001.

CLÉRY, A.; BLATTER, M.; ALLAIN, F. H.-T.: RNA recognition motifs: boring? Not quite. **Current Opinion in Structural Biology**, v. 18, n.3, 290-8, Jun. 2008.

de GODOY, L. M. F.; MARCHINI, F. K.; PAVONI, D. P.; RAMPAZZO, R. C. P.; PROBST, C. M.; GOLDENBERG, S.; KRIEGER, M. A.: Quantitative proteomics of *Trypanosoma cruzi* Metaciclogenesis. Research Article submitted for publication **Proteomics**, vol. 12, n. 17, p. 2694-2703. Aug. 2012.

DOCAMPO, R.; MORENO, S. N. J. Acidocalcisomes. **Cell Calcium**. V. 50, n. 2, p. 113-19. Aug. 2011.

DREYFUSS, G.; KIM, V. N.; KATAOKA, N.: Messenger-RNA-Binding proteins and the messages they carry. **Nature Reviews Molecular Cell Biology**, V. 3, n. 3, p. 195-205. Mar, 2002.

FERREIRA, L. R.; DOSSIN, F. M.; RAMOS, T.C.; FREYMULLER, E.; SCHENKMAN, S. Active transcription and ultrastructural changes during *Trypanosoma cruzi* metacyclogenesis. **An Acad Bras Cienc**. V. 80, n. 1, p. 157-66. Mar, 2008.

GOODSWEN, S. J.; KENNEDY, P. J.; ELLIS, J. T. A guide to in silico vaccine discovery for eukaryotic pathogens. **Briefings in bioinformatics**, v. 14, n. 6, p. 753-74. Oct. 2012.

GRIFFITHS, A. J. F.; WESSLER, S. R.; LEWONTIN, R. C.; CARROLL, S. B. **Introdução à Genética**. 7. ed. Rio de Janeiro: Guanabara. Koogan, 2002.

GRUMANN, C. Eletroforese. Disponível em: <http://genetica.ufcspa.edu.br/biomedic/conteudo/genetica_molecular/eletroforese.pdf>. Acesso em 04 de dezembro de 2013.

KELLY, J. M.; WARD, H. M.; MILES, M. A.; KENDALL, G.: A shuttle vector which facilitates the expression of transfected genes in *Trypanosoma cruzi* and Leishmania. **Nucleic Acids Res**, v. 20, n. 15, pg. 3963-9. Aug. 1992.

KUWAYAMA, H.; OBARA, S.; MORIO, T.; KUHARA, M. K.; TANAKA, Y.: A novel PCR-mediated method for one-tube generation of a gene disruption construct. **Biotechnology Letters**, v. 24, n. 16, p. 1307-12. Aug. 2002.

LEHNINGER, A. L.; NELSON, D. L.; COX, M. M. **Lehninger Princípios de Bioquímica**. 4. ed. São Paulo: Sarvier, 2006.

LESCURE, F. X.; LOUP, G. L.; FREILIJ, H.; DEVELOUX, M.; PARIS, L.; BRUTUS, L.; PIALOUX, G. Chagas disease: changes in knowledge and management. **Lancet Infect Dis**. V. 10, n. 8, p. 556-70. Aug. 2010.

LIMA, A. C.; VÁZQUEZ, M. C. G.; MORALES, O. R.; PACHECO, L. B.; ENCINA, J. L. R.; LÓPEZ, P. A. R.; FONSECA, M. A. R. Chagas disease (American trypanosomiasis) in Mexico: An Update. **Acta Tropica**. V. 127, n. 2, p. 126-35. Apr. 2013.

LORUSSO, A. B. **Desenvolvimento de um sistema semi-automatizado para nocautes gênicos em *Trypanosoma cruzi***. 132 f. Dissertação de Mestrado (Pós-Graduação em Biociências e Biotecnologia) - Instituto Carlos Chagas, Curitiba, 2013.

LOHRER, H.; SIEGEL, L. K. Purification and characterization of lipoamide dehydrogenase from *Trypanosoma cruzi*. **Eur. J. Biochem.** V. 194, n. 3, p. 863-9. Dec. 1990.

MARIS, C.; DOMINGUEZ, C.; ALLAIN, F., H.-T.: The RNA recognition motif, a plastic RNA-binding platform to regulate post-transcriptional gene expression. **The FEBS Journal**, v. 272, n. 9, p. 272-9. May. 2005.

MARTINS, A. V.; GOMES, A. P.; MENDONÇA, E. G.; FIETTO, J. L. R.; SANTANA, L. A.; OLIVEIRA, M. G. A.; GELLER, M.; SANTOS, R. F.; VITORINO, R. R.; BATISTA, R. S. Biology of *Trypanosoma cruzi*: An update. **Infectio**. V. 16, n. 1, p. 45-58. Fev. 2012.

MATTICK, J. S. Challenging the dogma: the hidden layer of non-protein-coding RNAs in complex organisms. **BioEssays**, v.25, n. 10, p.930-9. Out. 2003.

NETO, A. M. de: **Aprimoramento da anotação N-terminal de proteínas através da predição de peptídeo sinal em proteínas ortólogas e desenvolvimento de uma ferramenta automática para a identificação de grupos ortólogos contendo erros de anotação**. Tese (Doutorado pelo Programa de Pós-graduação em Ciências da Saúde) - Fundação Oswaldo Cruz – FIOCRUZ. Ministério da Saúde. Centro de Pesquisas René Rachou, Belo Horizonte. 2012. Disponível em: <<http://www.arca.fiocruz.br/handle/icict/6517>>. Acesso em 01 de março de 2014.

MICHELS, P. A. M.; BRINGAUD, F.; HERMAN, M.; HANNAERT, V. Metabolic functions of glycosomes in trypanosomatids. **Biochimica et Biophysica Acta – Molecular Cell Research**. V. 1763, n. 12, p. 1463-77. Dec. 2006.

PENNINGTON, S. R.; DUNN, M. J. **Proteomics: from protein sequence to function**. New York: Springer-Verlag e BIOS Scientific Publishers, 2001.

PEREZ, C. J.; LYMBERY, A. L.; THOMPSON, R. C. A. Chagas disease: the challenge of polyparasitism? **Cell Press**. V. 30, n. 4, p. 176-82. Apr. 2014.

PRATA, A. Clinical and epidemiological aspects of Chagas disease. **Lancet Infect Dis**. V. 1. N. 2, p. 92-100. Sep. 2001.

PROSDOCIMI, F. Curso on-line Introdução à bioinformática. Disponível em: <http://www.iq.usp.br/setubal/bmc/2013/FProsdocimi07_CursoBioinfo.pdf>. Acesso em 12 de outubro de 2013.

RASSI JÚNIOR, A.; RASSI, A.; NETO, J, A. M. Chagas Disease. **Lancet**. V. 4, n. 376, p. 1388-402. Apr. 2010.

REZENDE, P. R.; ALVES, G. B.; PEREIRA, L. M. C. M.; VALE, T. J. L.; PIMENTA, A. S. C.; LEMOS, J. A. R. Sensibilidade da técnica de reação em cadeia da polimerase para HIV-1 em relação à técnica de ensaio imunoenzimático. **Rev. bras. hematol. hemoter**. V. 24, n. 1, p. 25-8. Jan. 2002.

ROLDÁN, A.; COMINI, M. A.; CRISPO, M.; SIEGEL, R. L. K. Lipoamide desidrogenase is essential for both bloodstream and procyclic *Trypanosoma brucei*. **Molecular Microbiology**. V. 81, n. 3, p. 623-39. Jul. 2011.

SANTOS, E. B.; SILVA, N. L. C.; YOSHIDA, N; SILVEIRA, J. F. Expression and cellular trafficking of GP82 and GP90 glycoproteins during *Trypanosoma cruzi* metacyclogenesis. **Parasites & Vectors**. V. 6, n. 127. May. 2013.

SANTOS, F. R. C. C. dos. **Produção de Ferramentas Moleculares para Nocaute de Genes Envolvidos na Metaciclógênese de *Trypanosoma cruzi***. 45 p. Monografia (Bacharelado em Biomedicina). Faculdades Pequeno Príncipe, Curitiba, 2012.

SANTOS, P. F. **Caracterização dos genes que codificam a enzima dihidrolipoamida desidrogenase (TcLIPDH) e a proteína transportadora de hexoses (TcrHT1) em cepas do *Trypanosoma cruzi* sensíveis e resistentes ao benzonidazol**. Dissertação (Dissertação apresentada ao Programa de Pós-graduação em Ciências da Saúde) – Fundação Oswaldo Cruz – FIOCRUZ. Ministério da Saúde. Centro de Pesquisas René Rachou, Belo Horizonte. 2008.

SCHONECK, R.; MULOT, O. B.; NUMRICH, P.; OUAISSI, M.A.; SIEGEL, R. L. K. Cloning, sequencing and functional expression of dihydrolipoamide dehydrogenase from the human pathogen *Trypanosoma cruzi*. **Eur. J. Biochem**. V. 243, n. 3, p. 739-47. Feb. 1997.

TALICE, A. P.; GOES, V. M.; ARRAMBIDE, N.; AVILA, A. R.; DURAN, R.; CORREA, A.; DALLAGIOVANNA, B.; CAYOTA, A.; KRIEGER, M.; GOLDENBERG, S.; ROBELLO, C. Proteomic analysis of metacyclic trypomastigotes undergoing *Trypanosoma cruzi* metacyclogenesis. **J. Mass Spectrom**. V. 42, n. 11, p. 1422-1432. Nov. 2007.

TASHIRO, Y. Subcellular compartments and protein topogenesis. **Cell structure and function**, v. 8, n. 2, p. 91-107. Jun. 1983.

SOUZA, W. De. Basic cell biology of *Trypanosoma cruzi*. **Curr Pharm Des**. V. 8, n. 4, p. 269-85. 2002.

TAYLOR, M. C.; KELLY, J. M. pTcINDEX: a stable tetracycline-regulated expression vector for *Trypanosoma cruzi*, **BMC Biotechnology**, vol. 6, n. 32. 2006.

WHO - World Health Organization. Research Priorities for Chagas Disease, Human African Trypanosomiasis and Leishmaniasis. Who Technical Report Series, n. 975. 2012. Disponível em: <http://apps.who.int/iris/bitstream/10665/77472/1/WHO_TRS_975_eng.pdf>. Acesso em 17 de abril de 2013.

APÊNDICES

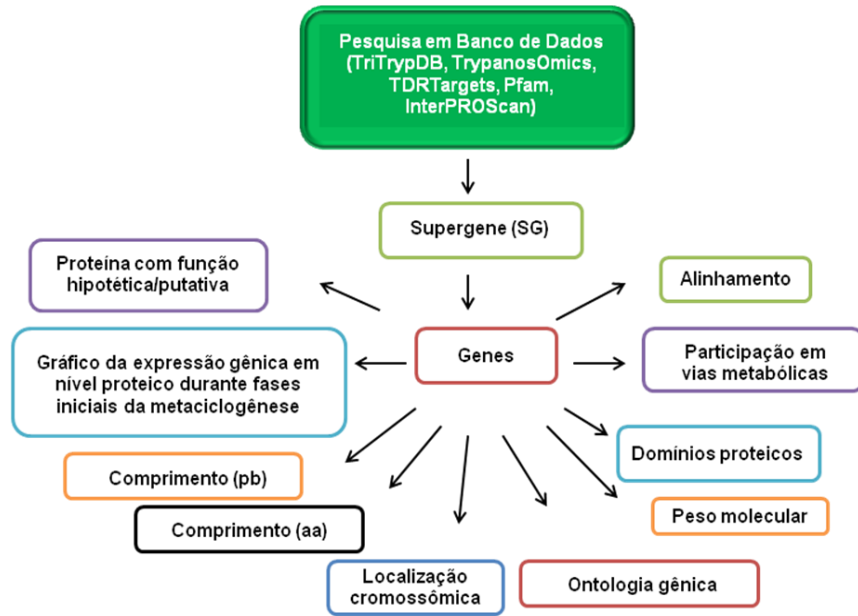


Figura Suplementar 1 - Uso de ferramentas bioinformáticas sobre informações referentes às quatro proteínas selecionadas.

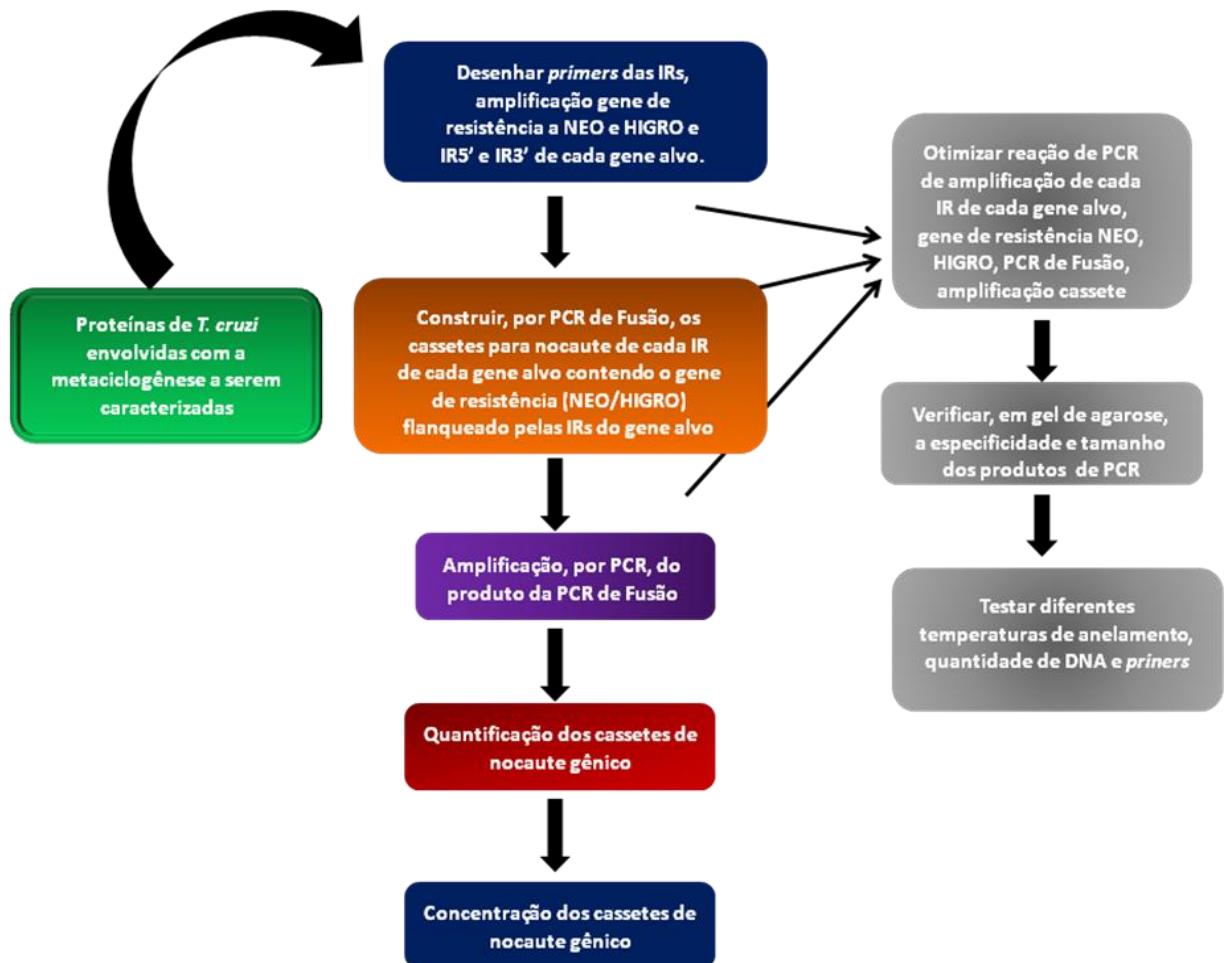


Figura Suplementar 2 - Construção e amplificação de cassette para nocaute gênico.

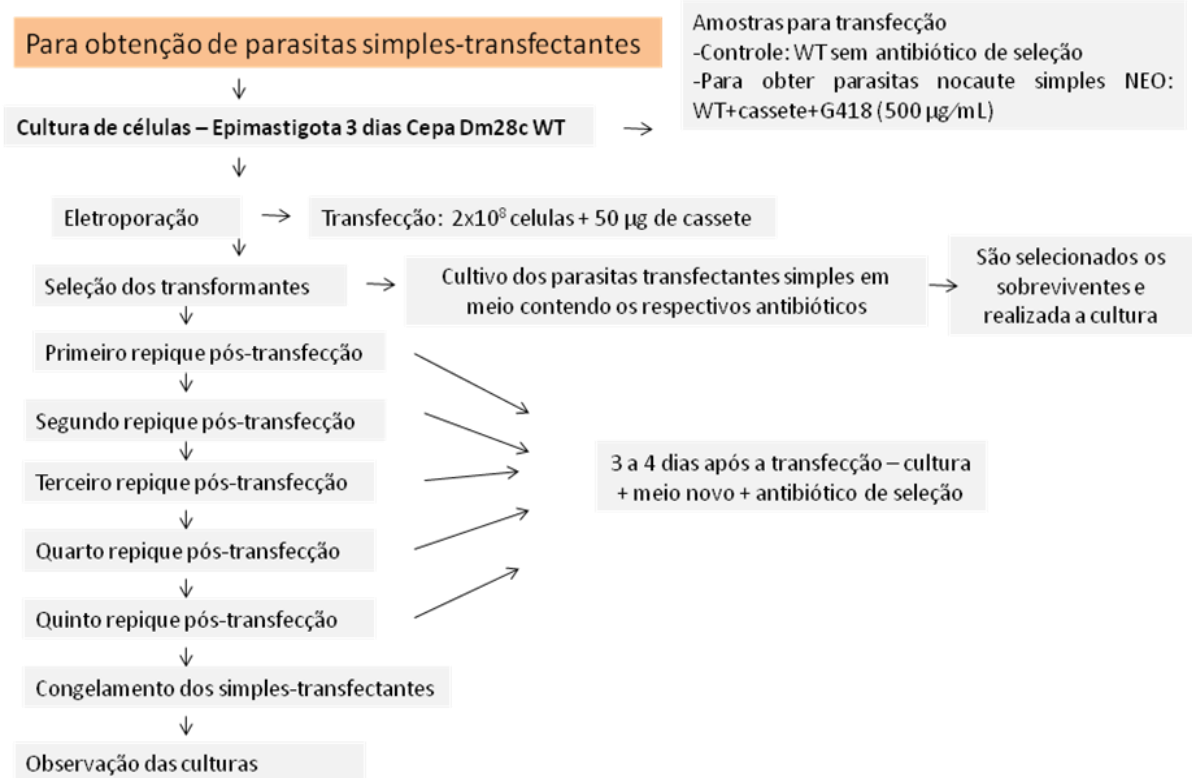


Figura Suplementar 3 - Obtenção de parasitas simples nocaute.

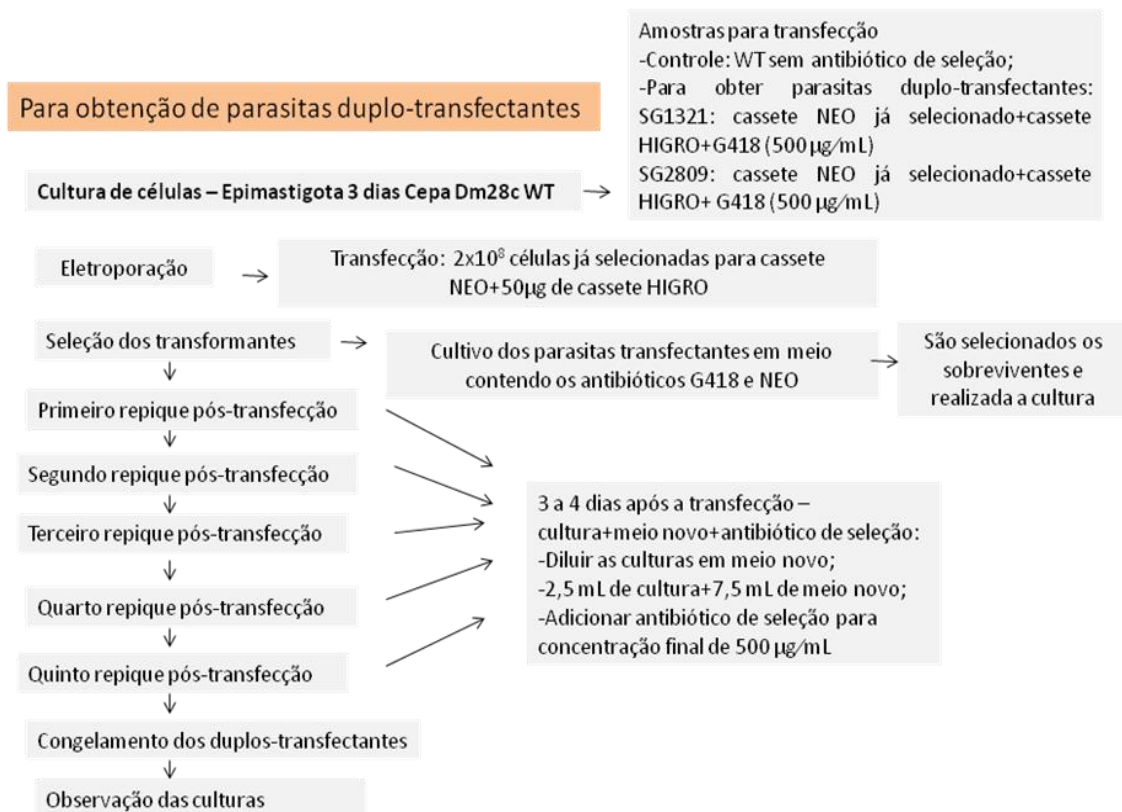


Figura Suplementar 4 - Obtenção de parasitas duplo nocaute.

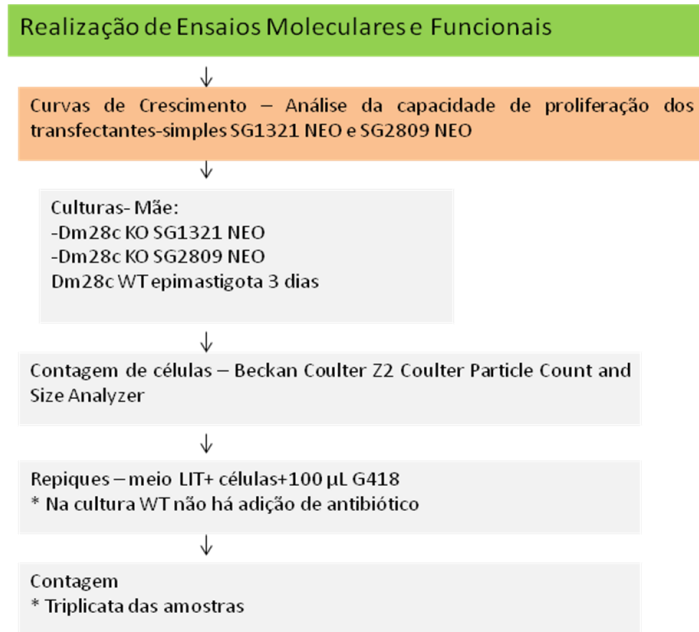


Figura Suplementar 5 - Realização de curvas de crescimento dos parasitas simples nocaute.

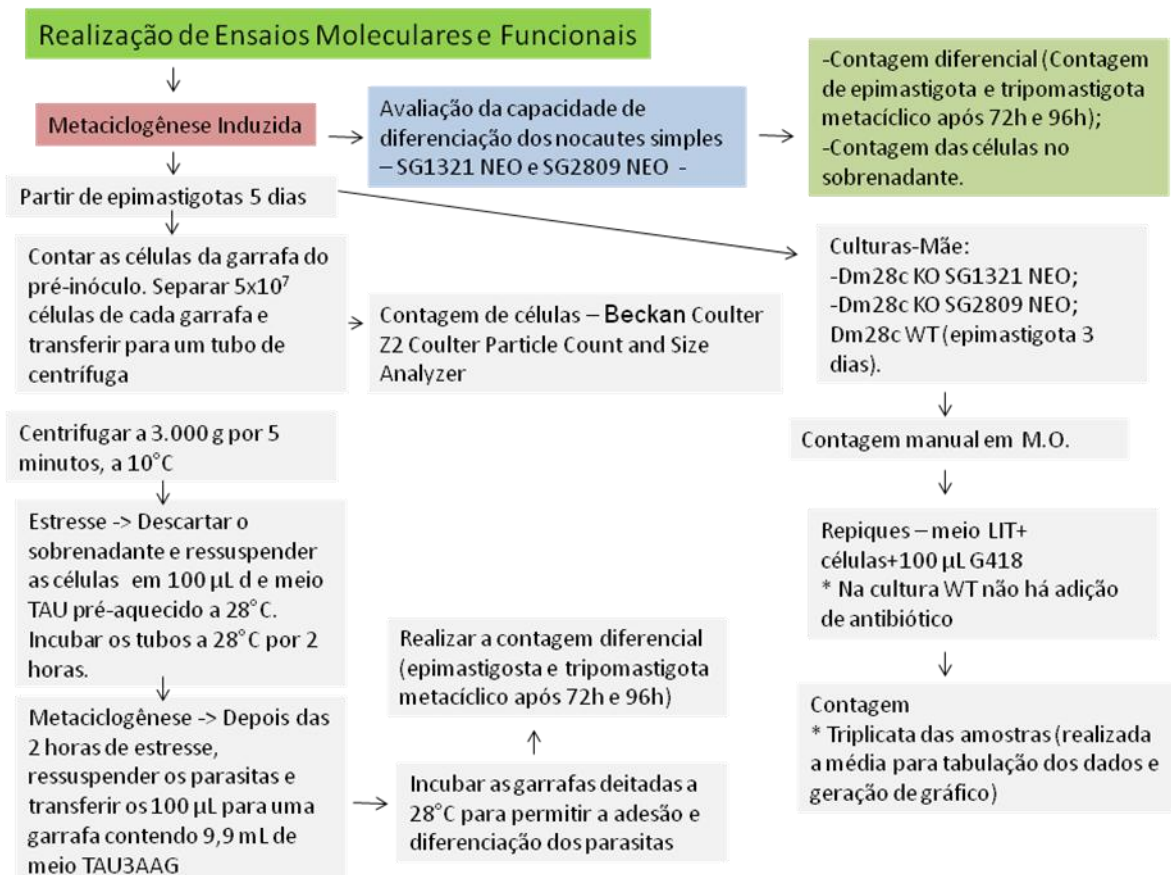


Figura Suplementar 6 - Indução da metaciclôgênese *in vitro*.

Tabela Suplementar 1 - Primers para amplificação das IRS à montante (5') e à jusante (3') das sequências codificadoras das quatro proteínas usados para construção dos cassetes de nocaute gênico e confirmação dos parasitas duplo nocaute.

Proteína	Primer	Primer ID	Sequencia	Amplicon
1	LG_01	Tc00.1047053511847.30_UPS_F	GGGGGGGGACCTGCCGTGTTTTATCT	584
	LG_02	Tc00.1047053511847.30_UPS_R	AGTGCGGAAAGAGAAATGGACGGTGCATTTATGGTCCTGGTGTGTTGA	
	LG_03	Tc00.1047053511847.30_DOWN_F	ACTTGAACAGAATTTAACATGCCCATCAAGTGGACGAACAGGGCAGAC	541
	LG_04	Tc00.1047053511847.30_DOWN_R	GGGGTGGGGGGTAACCCCAACAAGCA	
2	LG_09	Tc00.1047053510741.40_UPS_F	GGGGGGGGTCCACACCTTCTGCGTACAG	580
	LG_10	Tc00.1047053510741.40_UPS_R	AGTGCGGAAAGAGAAATGGACGGTGCATTTGCAGATAATAAAAAATACTGGCGGG	
	LG_11	Tc00.1047053510741.40_DOWN_F	ACATTGAACAGAATTTAACATGCCCATCAAGCAAAGAGGGAAAAGAAGC	573
	LG_12	Tc00.1047053510741.40_DOWN_R	GGGGGGGGCTTCCTTCGTGCTGAATTA	
3	LG_13	Tc00.1047053506701.10_UPS_F	GGGGGGGGGTGACGGCAGAAAGGAGA	307
	LG_14	Tc00.1047053506701.10_UPS_R	AGTGCGGAAAGAGAAATGGACGGTCAATTTAGGTCCGCTCCTCCAATTC	
	LG_15	Tc00.1047053506701.10_DOWN_F	ACATTGAACAGAATTTAACATGCCCATCAATTCTGTTGGCACAAGGAAAA	346
	LG_16	Tc00.1047053506701.10_DOWN_R	GGGGGGGGTTTCTTATTATCTGCCCTCACTCC	
4	LG_25	Tc00.1047053510183.40_UPS_F	GGGGGGGTGAAAATCAACAACAAGAAGTGG	594
	LG_26	Tc00.1047053510183.40_UPS_R	AGTGCGGAAAGAGAAATGGACGGTGCATTTTCTTCTGTTATTTCACTCTCCC	
	LG_27	Tc00.1047053510183.40_DOWN_F	ACATTGAACAGAATTTAACATGCCCATGCCATCAAGTCGATCTAGAGACCGTGCC	585
	LG_28	Tc00.1047053510183.40_DOWN_R	GGGGGGGGCCCTAACCCTGCCAAAA	

Tabela Suplementar 2 - *Primers* das sequências codificadoras (CDS) para confirmação dos parasitas duplo nocaute.

Proteína	Gene	Nome <i>primer</i>	Sequência	qtde nuc
SG1321 (1)	TcCLB506227.110	FMJA_6227_110_IR5	5'AAGTGGCTCTCGGGACCTG3'	20
		FMJA_6227_110_IR5	5'GGTGTGCTAAAGTGGCGACTTTCTAC3'	26
		FMJA_6227_110_IR3	5'AAAAAATTATTATTTTTATTTTTTTGGGTGCC3'	35
		FMJA_6227_110_IR3	5'GCGCTGCCCCACCC3'	15
		FMJA_6227_110_F	5'GGGGACAAGTTTGTACAAAAAGCAGGCTTCATGCGTCTTCGCACGGTCC3'	48
		FMJA_6227_110_R	5'GGGGACCACCTTTGTACAAGAAAGCTGGGTCCCTCATCCTTGATGGTGTCTTCGG3'	54
SG2809 (2)	TcCLB510661.230	FMJA_0661_230_IR5	5'ATATATTTCTCATCCACAGGAGCGTCG3'	27
		FMJA_0661_230_IR5	5'GCTGCTCTGAATTGCGCAAAAC3'	22
		FMJA_0661_230_IR3	5'TCCCCACGGAAGTGTGAGGC3'	20
		FMJA_0661_230_IR3	5'CGACAAATAAAAAACACACAAAAAAACAAC3'	31
		FMJA_0661_230_F	5'GGGGACAAGTTTGTACAAAAAGCAGGCTTCATGGGGAAAAACACTAGGCTTAAGTTCCG3'	56
		FMJA_0661_230_R	5'GGGGACCACCTTTGTACAAGAAAGCTGGGTCCGAGGACGGTGGAGGTGGAG3'	49
SG4727 (3)	TcCLB.506701.10	FMJA_SG4727_CDS	5'GGGGACAAGTTTGTACAAAAAGCAGGCTTCATGTTACATCGTAGCCACTTTTG3'	54
		FMJA_SG4727_CDS	5'GGGGACCACCTTTGTACAAGAAAGCTGGGTCTATTCTATCTCCTTGATATGCTCTACC3'	59
		FMJA_SG4727_IR5_F	5'GCGGAAGTTTCGTCAAGT3'	18
		FMJA_SG4727_IR5_R	5'TATCTCTGATTATGAACTGAAAAA3'	24
		FMJA_SG4727_IR3_F	5'GACGAATTTGGATATATATAGATATATGTGTAT3'	33
		FMJA_SG4727_IR3_R	5'CTTATTATCTGCCCTCACTCCT3'	22
SG5427 (4)	TcCLB.510183.40	FMJA_SG5427_CDS	5'GGGGACAAGTTTGTACAAAAAGCAGGCTTCATGTTGACGCCCTATCAGCTGGAACTC3'	58
		FMJA_SG5427_CDS	5'GGGGACCACCTTTGTACAAGAAAGCTGGGTCTAGATTGATGCCACTTGCCAGCCAGG3'	57
		FMJA_SG5427_IR5_F	5'CTGCTGTAAATACAAATGAGGAAGGAC3'	27
		FMJA_SG5427_IR5_R	5'AGTCCACACAATCCTCCTCCG3'	21
		FMJA_SG5427_IR3_F	5'TGTAAATTTTCAATAAGAAGAA3'	22
		FMJA_SG5427_IR3_R	5'CCTAACCTGCCAAAAC3'	17

* As proteínas estão simplificada e identificadas a nível de supergene (SG), segundo agrupamento realizado no site TrypanosOmics, conforme detalhado no item 4.1 dos Resultados e Discussão.