

UNIVERSIDADE FEDERAL DE SANTA CATARINA
CENTRO DE CIÊNCIAS BIOLÓGICAS
DEPARTAMENTO DE MICROBIOLOGIA, IMUNOLOGIA E PARASITOLOGIA
CURSO DE CIÊNCIAS BIOLÓGICAS

Beatriz Figueredo

**Análise transcriptômica de *Trypanosoma rangeli* durante a infecção de
*Rhodnius prolixus***

Florianópolis

2023

Beatriz Figueredo

**Análise transcriptômica de *Trypanosoma rangeli* durante a infecção de
*Rhodnius prolixus***

Trabalho de Conclusão de Curso submetido ao curso de Ciências Biológicas do Centro da Universidade Federal de Santa Catarina como requisito parcial para a obtenção do título de Licenciada em Ciências Biológicas.

Orientadora: Profa. Dra. Patrícia Hermes Stoco
Coorientadora: M.Sc. Laryssa Vanessa de Liz

Florianópolis

2023

Figueredo, Beatriz

Análise transcriptômica de *Trypanosoma rangeli* durante a infecção de *Rhodnius prolixus* / Beatriz Figueredo ; orientadora, Patrícia Hermes Stoco, coorientadora, Laryssa Vanessa de Liz, 2023.

76 p.

Trabalho de Conclusão de Curso (graduação) - Universidade Federal de Santa Catarina, Centro de Ciências Biológicas, Graduação em Ciências Biológicas, Florianópolis, 2023.

Inclui referências.

1. Ciências Biológicas. 2. Tripanossomatídeos. 3. RNA-seq. 4. Expressão diferencial. I. Stoco, Patrícia Hermes . II. de Liz, Laryssa Vanessa. III. Universidade Federal de Santa Catarina. Graduação em Ciências Biológicas. IV. Título.

Beatriz Figueredo

Análise transcriptômica de *Trypanosoma rangeli* durante a infecção de *Rhodnius prolixus*

Este Trabalho de Conclusão de Curso foi julgado adequado para obtenção do título de Licenciada e aprovada em sua forma final pelo Curso de Ciências Biológicas.

Florianópolis, 04 de dezembro de 2023.

Insira neste espaço
a assinatura

Coordenação do Curso

Banca examinadora

Insira neste espaço
a assinatura

Profa. Patrícia Hermes Stoco, Dra.

Orientadora

Insira neste espaço
a assinatura

Profa. Franceli Rodrigues Kulcheski, Dra.

UFSC

Insira neste espaço
a assinatura

Prof. Renato Simões Moreira, Dr.

IFSC

Florianópolis, 2023.

AGRADECIMENTOS

Primeiro gostaria de agradecer aos meus pais e minha irmã, que sempre me incentivaram mesmo quando disse que queria cursar biologia e não tinha a mínima ideia do que faria no futuro. Eles vibram a cada pequena conquista, a cada apresentação em congresso e cada passo que dou. Esse amor e apoio não tem preço! E sem eles nada disso seria possível.

Agradeço ao meu amor, Vilmar Benetti Filho! Fazer o TCC em dois meses não foi fácil, mas teu apoio e parceria tornou esse processo mais leve. Obrigada por acreditar em mim em todas as vezes que duvidei.

Agradeço a minha orientadora, Patrícia! Além de todo o suporte, carinho que ela dá a suas meninas da Proto, é incrível ter uma mulher tão forte e uma cientista extraordinária como exemplo.

Agradeço a minha coorientadora, Laryssa! Que basicamente me viu crescer dentro da academia e me acompanha desde a primeira fase da graduação, sempre acreditando em mim e me incentivando. Você é um exemplo de cientista para mim!

Sou muito grata aos meus amigos do Lab, que aguentam os perrengues comigo e não medem esforços para ajudar quem é parceiro. Em especial: a Carol, minha dupla das placas de frangos; Ana Paulinha, ainda bem que meu primeiro IC foi com sua parceria; Eric, que não mede esforços para ajudar os outros; Amábilli, por ser a veterinária mais doce que conheço; e Ana Claudinha, Luiza, Luisa e Helena que também compartilham as angústias, mas também as alegrias.

Aos amigos que conheci por meio do Lab e apesar de não estarmos no mesmo ambiente, ainda fazem parte da minha vida, auxiliam na minha jornada e estão no meu coração. Obrigada Caibe (o mantra “6 é 10” mudou minha vida!), Abadio, Carime, Bruna e Nati!

Aos meus amigos que entraram comigo na graduação e que me acompanharam desde as disciplinas, PPCCs e até nos rolês. Obrigada Gabe, Bia e Letícia!

Agradeço ao meu primeiro orientador, por abrir a porta do Lab para mim e por me passar ensinamentos muito importantes. Obrigada Ed! Eu não teria o mesmo rigor científico, o pensamento crítico e a curiosidade se não tivesse sido orientada por você.

Agradeço ao professor Glauber, que se dispôs e me auxiliou por meio dos PIBICs! Agradeço ao Prof. Mário e Prof. Patrícia Quaresma que contribuíram com minha formação!

Agradeço ao pessoal da Bioinfo que me acolheu e uma parte foi ver as *caipirinhas* em Curitiba comigo. Obrigada Renato, Fernando, Eric, Karin, Taty, Day, Guilherme, Maiara, Bruna e Jaime!

Agradeço a todos que me acompanharam nessa trajetória e me auxiliaram de forma direta ou indireta! Agradeço também a UFSC, ao CNPq e a CAPES, sem o apoio financeiro e estrutural a realização desse trabalho não seria possível!

RESUMO

O *Trypanosoma rangeli* é um parasito hemoflagelado que infecta mamíferos e vetores triatomíneos, especialmente do gênero *Rhodnius*. Apesar de ser considerado não patogênico para os hospedeiros mamíferos, é relatada a ocorrência de coinfeções com *Trypanosoma cruzi*, além de ocorrer reatividade sorológica cruzada em testes imunológicos para doença de Chagas. Para o *T. rangeli* ser transmitido pelo vetor, as formas epimastigotas precisam evadir o trato intestinal, se multiplicar na hemolinfa e posteriormente invadir a glândula salivar, diferenciando-se em tripomastigotas metacíclicas, que são inoculados durante o repasto sanguíneo no hospedeiro mamífero. A fim de compreender melhor as interações parasito-vetor durante a infecção do inseto, este trabalho teve como objetivo avaliar as mudanças na transcrição gênica entre as formas tripomastigotas e epimastigotas de *T. rangeli* durante a infecção em *Rhodnius prolixus*. Para isso, o RNA da hemolinfa e glândula salivar de *R. prolixus* infectado foi extraído e sequenciado. O transcriptoma de *T. rangeli* foi montado utilizando o StringTie, anotado utilizando o AnnotaPipeline e a análise de expressão diferencial foi realizada utilizando o DESeq2, comparando as formas tripomastigotas e epimastigotas. Foram considerados diferencialmente expressos os transcritos com \log_2 fold change ≤ -2 ou ≥ 2 e valor de p-ajustado < 0.01 por FDR. Foram obtidos 2.864 transcritos após remoção das sequências truncadas, observando as sequências aminoacídicas obtidas a partir da primeira janela de leitura. Desses, 2.064 receberam anotação com função conhecida e 800 foram consideradas proteínas hipotéticas. Um total de 1.446 receberam ao menos um termo de ontologia genética (GO), sendo os termos de função molecular mais frequentes a atividade exo-alfa-sialidase e atividade de metaloendopeptidase. 69 transcritos apresentaram aumento dos valores de expressão gênica nas formas tripomastigotas. Destes, 56 foram anotados como proteínas hipotéticas e outros 13 receberam anotação funcional. Considerando os transcritos que apresentaram aumento de expressão pelas formas epimastigotas, dos 34 transcritos identificados, apenas 2 receberam anotações hipotéticas. Destacam-se membros das famílias de proteínas de superfície Trans-sialidase e GP63, que tiveram expressão diferencial em ambas as formas do parasito. Outros transcritos diferencialmente expressos foram cinesinas e nodulinas, relacionadas ao transporte de substâncias, transcetolase e cistationina- γ -liase. A detecção desses transcritos, também encontrados em parasitos patogênicos como *T. cruzi*, corroboram com a importância para a interação parasito-vetor, conseqüentemente, na sobrevivência do parasito e desenvolvimento da infecção no hospedeiro invertebrado.

Palavras-chave: RNA-seq; Tripanossomatídeos; Vetor; Expressão diferencial.

ABSTRACT

Trypanosoma rangeli is a hemoflagellate parasite that infects mammals and triatomine vectors, especially the *Rhodnius* genus. Despite being considered non-pathogenic for mammalian hosts, co-infections with *Trypanosoma cruzi* have been reported, in addition to cross-serological reactivity in immunological tests for Chagas disease. For *T. rangeli* to be transmitted by the vector, the epimastigote forms need to escape the intestinal tract, multiply in the hemolymph, and subsequently invade the salivary gland, differentiating into metacyclic trypomastigotes, which are inoculated during the blood meal in the mammalian host. To better understand parasite-vector interactions during insect infection, this work aimed to evaluate changes in gene transcription between trypomastigotes and epimastigotes forms of *T. rangeli* during infection in *Rhodnius prolixus*. For this, RNA from the hemolymph and salivary gland of infected *R. prolixus* was extracted and sequenced. The *T. rangeli* transcriptome was assembled using StringTie, annotated using AnnotaPipeline and differential expression analysis was performed using DESeq2, comparing the trypomastigote and epimastigote forms. Transcripts with log₂ fold change ≤ -2 or ≥ 2 and adjusted p-value < 0.01 by FDR were considered differentially expressed. 2,864 transcripts were obtained after removing the truncated sequences, observing the amino acid sequences obtained from the first reading window. Of these, 2,064 were annotated with known functions and 800 were considered hypothetical proteins. 1,446 received at least one gene ontology (GO) term, with the most frequent molecular function terms being exo-alpha-sialidase activity and metalloendopeptidase activity. 69 transcripts showed increased gene expression values in trypomastigote forms. Of these, 56 were annotated as hypothetical proteins and another 13 received functional annotation. Considering the transcripts that showed increased expression by epimastigote forms, of the 34 transcripts identified, only 2 received hypothetical annotations. Members of the Trans-sialidase and GP63 surface protein families stand out, which had differential expression in both parasite forms. Other differentially expressed transcripts were kinesins and nodulins, related to the transport of substances, transketolase, and cystathionine- γ -lyase. The detection of these transcripts, also found in pathogenic parasites such as *T. cruzi*, corroborates the importance of the parasite-vector interaction, consequently, in the survival of the parasite and development of the infection in the invertebrate host.

Keywords: RNA-seq; Trypanosomatids; Vector; Differential expression.

LISTA DE FIGURAS

Figura 1 – Ciclo biológico do <i>Trypanosoma rangeli</i> alternando entre seus hospedeiros vertebrados e invertebrados.	17
Figura 2 - Processo de transcrição em tripanossomatídeos	25
Figura 3 - Desenho experimental para análises de montagem, anotação e expressão diferencial de transcritos de <i>T. rangeli</i>	30
Figura 4 - Valores de qualidade phred obtidos pelo programa MultiQC após etapa de controle de qualidade para dados brutos de RNAseq de glândula salivar de <i>R. prolixus</i> infectado com tripomastigotas de <i>T. rangeli</i>	35
Figura 5 - Valores de qualidade phred score obtidos pelo programa MultiQC após etapa de controle de qualidade para dados brutos de RNAseq de hemolinfa de <i>R. prolixus</i> infectado por epimastigotas de <i>T. rangeli</i>	36
Figura 6 – Proporção de genes presentes no transcriptoma de <i>Trypanosoma rangeli</i> quando comparado com ortólogos do filo Euglenozoa.	39
Figura 7 – Termos de ontologia genética mais frequentes entre os transcritos.	41
Figura 8 - Variância entre os grupos Epimastigotas e Tripomastigotas.	44
Figura 9 - Análise comparativa de expressão gênica entre as formas tripomastigotas e epimastigotas de <i>T. rangeli</i>	45
Figura 10 – Frequência de termos de ontologia genética identificados no transcriptoma de <i>T. rangeli</i>	71

LISTA DE TABELAS

Tabela 1 – Características gerais do transcriptoma de <i>T. rangeli</i>	37
Tabela 2 – Níveis de expressão em logFC de transcritos de <i>T. rangeli</i> anotados como genes de transcrição constitutiva.	46
Tabela 3 – Transcritos de <i>T. rangeli</i> com anotação conhecida apresentando aumento de expressão gênica quando comparado formas tripomastigotas contra epimastigotas.	47
Tabela 4 - Transcritos de <i>T. rangeli</i> com anotação conhecida apresentando diminuição de expressão gênica quando comparado formas tripomastigotas contra epimastigotas.	51
Tabela 5 – Totalidade de transcritos identificados exibindo aumento de expressão gênica quando comparado formas tripomastigotas vs. epimastigotas.	72
Tabela 6 - Totalidade de transcritos identificados exibindo diminuição de expressão gênica quando comparado formas tripomastigotas vs. epimastigotas.	75

LISTA DE ABREVIATURAS E SIGLAS

Mb	megabase
kb	kilobase
TS	trans-Sialidase
MASPs	proteínas de superfície associadas à mucina
RHS	<i>retrotransposon Hot Spot</i>
DGF-1	<i>dispersed Gene Family 1</i>
DNA	ácido desoxirribonucleico
kDNA	DNA do cinetoplasto
RNA	ácido ribonucleico
mRNA	RNA mensageiro
SL	<i>spliced leader</i>
RNAP II	RNA polimerase II
pb	pares de base

SUMÁRIO

1	INTRODUÇÃO	14
1.1	O <i>TRYPANOSOMA RANGELI</i>	15
1.2	INTERAÇÕES PARASITO-VETOR	18
1.3	ESTRUTURA GENÔMICA DE TRIPANOSSOMATÍDEOS.....	21
1.3.1	Famílias multigênicas em tripanossomatídeos	22
1.4	A TRANSCRIÇÃO EM TRIPANOSSOMATÍDEOS	24
1.5	O PAPEL DAS CIÊNCIAS ÔMICAS PARA COMPREENSÃO DE EVENTOS BIOLÓGICOS.....	26
2	OBJETIVOS	28
2.1.1	Objetivo Geral	28
2.1.2	Objetivos específicos	28
3	METODOLOGIA	29
3.1	DELINEAMENTO EXPERIMENTAL	29
3.2	OBTENÇÃO DE DADOS	31
3.2.1	Cultivo e diferenciação de Parasitos	31
3.2.2	Manutenção de Triatomíneos	31
3.2.3	Infecção de <i>R. prolixus</i> por <i>T. rangeli</i>	31
3.2.4	Extração de RNA, construção das bibliotecas e sequenciamento	32
3.3	ANÁLISES DE BIOINFORMÁTICA.....	32
3.3.1	Controle de qualidade dos dados de RNA-seq	32
3.3.2	Montagem e anotação do transcriptoma	33
3.3.3	Análise de expressão diferencial	33
4	RESULTADOS E DISCUSSÃO	34
4.1	ANÁLISE DE QUALIDADE DO SEQUENCIAMENTO E DO MAPEAMENTO CONTRA O GENOMA DO VETOR.....	34
4.2	MONTAGEM E ANOTAÇÃO DOS TRANSCRITOS	37
4.3	ANÁLISE DE EXPRESSÃO DIFERENCIAL	43
4.3.1	Genes com expressão gênica aumentada em tripomastigotas	46
4.3.2	Genes com expressão gênica aumentada em epimastigotas	50
5	CONCLUSÕES	54
	REFERÊNCIAS	55

APÊNDICE A – TERMOS DE ONTOLOGIA GENÉTICA IDENTIFICADOS NO TRANSCRIPTOMA DE <i>T. RANGELI</i>.....	70
APÊNDICE B – TRANSCRITOS DE <i>T. RANGELI</i> APRESENTANDO AUMENTO DE EXPRESSÃO GÊNICA QUANDO COMPARADO FORMAS TRIPOMASTIGOTAS CONTRA EPIMASTIGOTAS.	72
APÊNDICE C – TRANSCRITOS DE <i>T. RANGELI</i> APRESENTANDO DIMINUIÇÃO DE EXPRESSÃO GÊNICA QUANDO COMPARADO FORMAS TRIPOMASTIGOTAS CONTRA EPIMASTIGOTAS.....	75

1 INTRODUÇÃO

O *Trypanosoma rangeli* (TEJERA, 1920) é um tripanossomatídeo com ciclo de vida heteroxênico, alternando entre os hospedeiros vertebrados e invertebrados (D'ALESSANDRO, 1976). Diversas espécies de mamíferos, incluindo o ser humano, estão dentre os hospedeiros vertebrados. Apesar da capacidade de infectar os humanos, o parasito é tido como não patogênico (GUHL; VALLEJO, 2003). No entanto, o *T. rangeli* compartilha algumas características com o *Trypanosoma cruzi*, agente etiológico da doença de Chagas, como a distribuição geográfica, os mesmos hospedeiros e vetores (GUHL; VALLEJO, 2003). Por consequência, é comum a identificação de infecções mistas e ainda, reatividade cruzada em testes sorológicos para doença de Chagas, devido ao compartilhamento de proteínas entre esses parasitos (GRISARD *et al.*, 1999; GUHL; VALLEJO, 2003).

Embora o *T. rangeli* e o *T. cruzi* compartilhem os mesmos vetores, há diferenças nos seus ciclos de vida durante a infecção dos triatomíneos, que são seus hospedeiros invertebrados. Enquanto o *T. cruzi* permanece exclusivamente no trato digestivo do inseto, o *T. rangeli* coloniza diferentes tecidos (GUHL; VALLEJO, 2003). O ciclo no vetor inicia no momento em que o inseto realiza a hematofagia em um mamífero infectado, ingerindo junto com o sangue as formas tripomastigotas sanguíneas. No trato intestinal, os parasitos se diferenciam em epimastigotas, que evadem o trato digestivo e alcançam a hemocele, onde o *T. rangeli* se replica intensamente. Em seguida, o parasito invade as glândulas salivares, onde irá se diferenciar para as formas tripomastigotas metacíclicas, a forma infectante do parasito. Durante o próximo repasto sanguíneo, as tripomastigotas metacíclicas serão inoculadas junto com a saliva no hospedeiro mamífero (D'ALESSANDRO, 1976). Curiosamente, não são todos os gêneros de triatomíneos que o *T. rangeli* é capaz de completar seu ciclo de desenvolvimento produzindo formas infectivas, apenas nas espécies do gênero *Rhodnius* (CUBA CUBA, 1998; MARQUEZ *et al.*, 2006).

Para realizar a alternância de hospedeiros, o parasito precisa se adaptar a diferentes ambientes. Dessa forma, altera a expressão de genes que possibilitarão sua sobrevivência em cada hospedeiro. Diferente de outros eucariotos, a expressão gênica em tripanossomatídeos é policistrônica e a regulação dos níveis de expressão acontecem principalmente a nível pós transcricional (CLAYTON, 2002).

Dessa forma, para ampliar a compreensão dos eventos e interações que ocorrem entre o parasito e o hospedeiro, diferentes metodologias têm sido empregadas, incluindo análises de transcriptoma que ajudam a identificar genes fundamentais para o estabelecimento da infecção em cada fase do ciclo de vida do parasito (VERA *et al.*, 2008). Por isso, buscamos analisar o transcriptoma de *T. rangeli* durante a infecção da hemolinfa e glândula salivar de *Rhodnius prolixus* e identificar genes com níveis de expressão distintos dependendo do órgão ou região que o parasito está colonizando.

1.1 O *TRYPANOSOMA RANGELI*

O *T. rangeli* é um parasito hemoflagelado pertencente ao filo Euglenozoa, ordem Kinetoplastida e família Trypanosomatidae (SCHOCH *et al.*, 2020). Na família Trypanosomatidae encontram-se parasitos patogênicos para diversos seres vivos, incluindo plantas e animais (KAUFER *et al.*, 2017). Ganham destaque os parasitos causadores das Doenças Tropicais Negligenciadas, como as Leishmanioses, Tripanossomíase Africana (doença do sono) e doença de Chagas, cujos agentes etiológicos são *Leishmania* spp., o *Trypanosoma brucei*, e o *Trypanosoma cruzi*, respectivamente (WHO, 2023a; WHO, 2023b; WHO, 2023c).

As cepas do *T. rangeli* podem ser divididas em duas linhagens, KP1 (+) e KP1 (-) (CUERVO *et al.*, 2013; SINCERO *et al.*, 2015; VALLEJO *et al.*, 2002). Isso porque o DNA mitocondrial desses parasitos está organizado numa estrutura chamada de cinetoplasto (kDNA), composto por moléculas de DNA circulares com diferentes tamanhos (maxicírculos e minicírculos) e sequências variáveis entre as espécies (LUKEŠ *et al.*, 2002). Os maxicírculos possuem genes mitocondriais compartilhados entre eucariotos (GERASIMOV *et al.*, 2020), ao passo que os minicírculos são específicos dos tripanossomatídeos e demonstram possuir variações específicas entre espécies e entre cepas (WESTENBERGER *et al.*, 2006). O *T. rangeli*, por exemplo, apresenta três regiões conservadas no minicírculo, denominadas KP1, KP2 e KP3 (CUERVO *et al.*, 2013; SINCERO *et al.*, 2015; VALLEJO *et al.*, 2002). Cepas de *T. rangeli* isoladas de *R. prolixus* possuíam as 3 regiões, sendo denominadas KP1(+). Por outro lado, cepas obtidas de *Rhodnius colombiensis* ou *Panstrongylus megistus* apresentaram apenas as regiões KP2 e

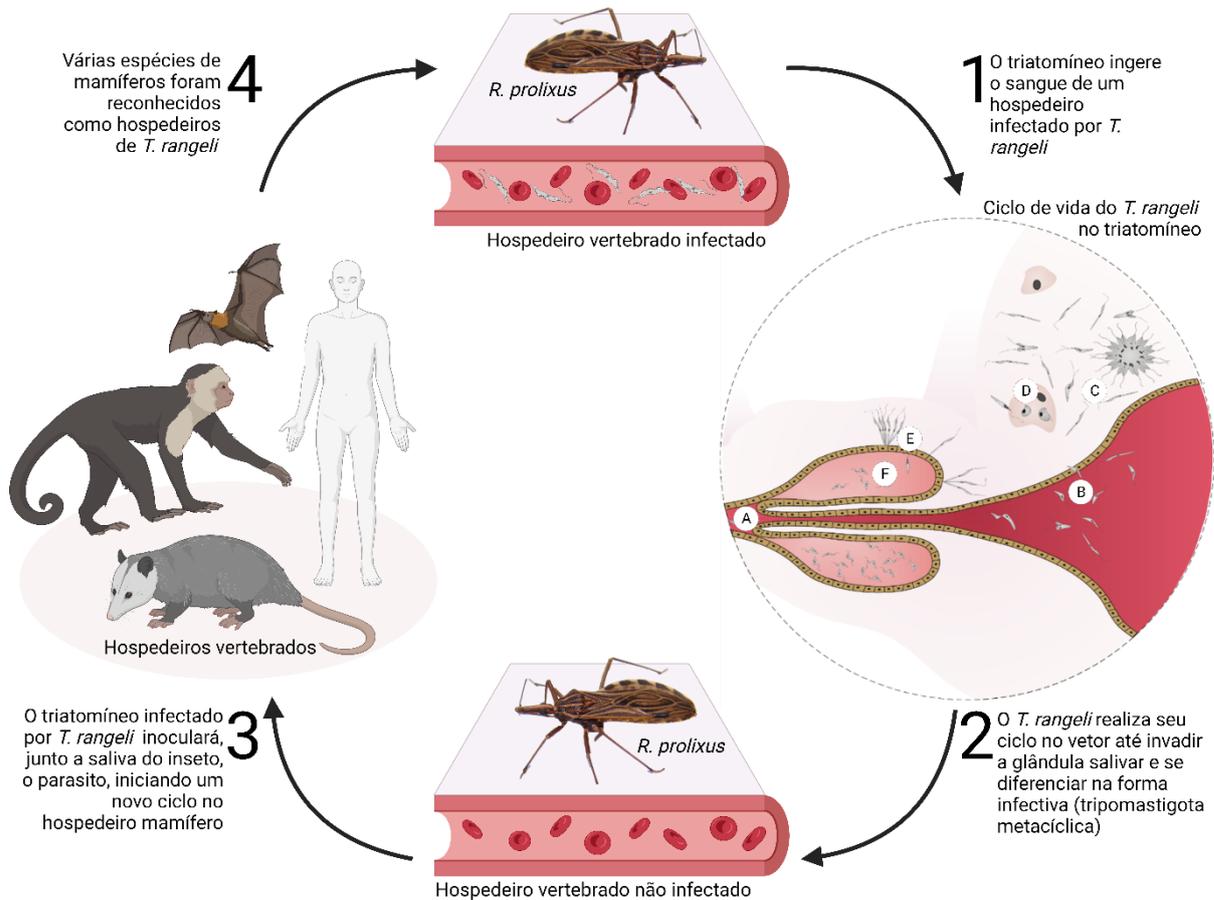
KP3. Dessa forma, essas cepas foram consideradas KP1(-) por não apresentarem fragmentos da região KP1 (VALLEJO *et al.*, 2002).

Uma problemática envolvendo o *T. rangeli* é o compartilhamento dos mesmos hospedeiros vertebrados e invertebrados que o *T. cruzi* (CUBA CUBA, 1998; MAIA DA SILVA *et al.*, 2007; MARQUEZ *et al.*, 2006). O *T. rangeli* é considerado não patogênico para os seus hospedeiros vertebrados, mas devido ao compartilhamento de hospedeiros e a sobreposição da distribuição geográfica com o *T. cruzi*, pode-se encontrar tanto coinfeções, como reatividade cruzada em testes sorológicos (GRISARD *et al.*, 1999; GUHL; VALLEJO, 2003). Uma vez que os dois parasitos compartilham em torno de 60% de sua constituição antigênica solúvel nas formas epimastigotas, o diagnóstico sorológico da doença de Chagas fica comprometido, principalmente na fase crônica da doença (AFCHAIN *et al.*, 1979). Apesar de não ser a forma biológica encontrada nos mamíferos, as formas epimastigotas são amplamente utilizadas nos diagnósticos sorológicos pelo fato de serem facilmente cultivadas, ao contrário das formas tripomastigotas, que apresentam uma redução da reatividade cruzada quando testadas com soros de pacientes chagásicos em ensaios de imunofluorescência indireta (DE MORAES *et al.*, 2008).

O ciclo biológico do *T. rangeli* no hospedeiro mamífero inicia no momento em que o triatomíneo, infectado por *T. rangeli*, realiza a hematofagia. Os parasitos são inoculados junto a saliva do vetor e atingem a corrente sanguínea (Figura 1) (CUBA CUBA, 1998). Entretanto, o curso da infecção não está bem elucidado. Recentemente um estudo realizou a transmissão de *T. rangeli* pela picada de *R. prolixus* em camundongos, identificando a presença de parasitos extracelulares na derme nos primeiros sete dias pós-infecção. Após esse período, não foram encontrados parasitos na pele, mas foram identificados parasitos no sangue circulante e nos tecidos linfoides secundários durante o período de 30 dias pós-infecção. Foram encontrados parasitos nos linfonodos mesentéricos e baço, sendo que no baço foram identificadas formas do parasito semelhantes a amastigotas, sugerindo que o parasito se replica nesse tecido (FERREIRA *et al.*, 2020). Em outro estudo, pela técnica de bioluminescência, foram identificados parasitos até o 13º dia pós infecção, sendo notado um acúmulo de parasitos principalmente no baço e pulmões durante a fase aguda da infecção, além da indução de esplenomegalia nos camundongos, sendo observado um tropismo de *T. rangeli* em tecidos reprodutores masculino (SILVA, 2021). Dessa forma, apesar dos recentes achados, ainda não há

um consenso quanto a capacidade do *T. rangeli* se multiplicar no hospedeiro mamífero e quanto ao local da replicação, que pode ser intracelular ou extracelular (STEINDEL, 1993; ZUÑIGA *et al.*, 1997).

Figura 1 – Ciclo biológico do *Trypanosoma rangeli* alternando entre seus hospedeiros vertebrados e invertebrados.



Legenda: Ciclo biológico do *Trypanosoma rangeli*. Em (1) o triatomíneo, *Rhodnius prolixus*, realiza a hematofagia de um mamífero infectado por *T. rangeli*. Em (2), ocorre o estabelecimento da infecção de *T. rangeli* no vetor. A) O triatomíneo ingere as formas tripomastigotas sanguíneas junto ao sangue do animal infectado. B) O *T. rangeli* chega ao trato digestório do vetor e se diferencia em formas epimastigotas. C) Os epimastigotas invadem a hemocele e se replicam no local. D) Alguns parasitos invadem os hemócitos. E) Os parasitos se aderem, invadem a glândula salivar do triatomíneo e se diferenciam em formas tripomastigotas metacíclicas (forma infectiva). F) Tripomastigotas metacíclicos são inoculados durante a hematofagia. Em (3), o triatomíneo realiza a hematofagia, infectando outros hospedeiros vertebrados. Em (4), o triatomíneo realiza a hematofagia do hospedeiro mamífero infectado, ingerindo as formas tripomastigotas sanguíneas. Fonte: Adaptado de Stoco *et al.*, (2017).

No vetor o ciclo inicia quando o triatomíneo ingere, durante o repasto sanguíneo, as formas tripomastigotas sanguíneas que estão presentes no sangue de um mamífero infectado (Figura 1) (FERREIRA, 2017). Ao chegarem no trato intestinal do inseto, os parasitos diferenciam-se da forma tripomastigota sanguínea para a forma epimastigota. As formas epimastigotas iniciam o processo de replicação e são capazes de transpassar o epitélio intestinal, chegando na cavidade celômica do inseto (D’ALESSANDRO, 1976; HECKER; SCHWARZENBACH; RUDIN, 1990). Na hemocele os parasitos se multiplicam intensamente e então migram e invadem as glândulas salivares, onde acontecerá a metaciclogênese, ocorrendo a diferenciação para a forma tripomastigota metacíclica - a forma infectante do parasito. Durante o próximo repasto sanguíneo, os tripomastigotas metacíclicos serão inoculados junto com a saliva no hospedeiro mamífero (D’ALESSANDRO, 1976; D’ALESSANDRO; SARAIVA, 1992). Apesar do *T. rangeli* ser tido como não patogênico para os hospedeiros mamíferos, para os hospedeiros invertebrados são relatados efeitos patogênicos, como atraso ou inibição da muda, muda apresentando defeitos, aumento da mortalidade (AÑEZ, 1984; BRECHER; WIGGLESWORTH, 1944; FERREIRA *et al.*, 2010; VALLEJO *et al.*, 1986), diminuição da fecundidade e fertilidade, como também atraso na postura de ovos (FELLET *et al.*, 2014).

1.2 INTERAÇÕES PARASITO-VETOR

Os triatomíneos são insetos hematófagos, também denominados de solenofágicos, que retiram o sangue diretamente de arteríolas e vênulas para se alimentarem. Apresentam um aparelho bucal especializado, com um par de mandíbulas e de maxilas justapostas, criando um canal alimentar e um canal salivar (LAVOPIERRE; DICKERSON; GORDON, 1959). O triatomíneo dispõe de um par de glândulas salivares, e cada glândula possui uma estrutura maior e avermelhada e uma estrutura menor, acessória e translúcida (REIS; MEIRELLES; SOARES, 2003). As glândulas formam uma cavidade secretora composta por uma única camada epitelial ligadas a fibras musculares por meio de uma lâmina basal espessa e repleta de traquéias (MEIRELLES *et al.*, 2003). Esses insetos ainda apresentam na composição da saliva, moléculas anestésicas, anticoagulantes e antiplaquetárias, que possibilitam realizar sua alimentação e não serem notados pelo hospedeiro

(RIBEIRO, 1987; RIBEIRO *et al.*, 1998). A liberação da saliva é realizada durante todo o período em que o inseto estiver se alimentando (SOARES *et al.*, 2006). É justamente no momento da hematofagia de um mamífero infectado, em que o *T. rangeli* é ingerido, que a interação parasito-hospedeiro se inicia.

No trato intestinal dos triatomíneos há uma microbiota que desempenha papéis importantes para sua sobrevivência, como o provimento de nutrientes fundamentais, auxílio na digestão, desintoxicação de compostos adquiridos na dieta, mantimento da homeostase da região e de uma resposta imune básica (GARCIA *et al.*, 2010; JING; QI; WANG, 2020; LEULIER; ROYET, 2009). A chegada do *T. rangeli* no trato intestinal acarreta mudanças na microbiota, alterando a resposta imune no intestino de *R. prolixus* (VIEIRA *et al.*, 2015). A atividade antibacteriana no intestino e a expressão de genes que codificam peptídeos antimicrobianos e enzimas que terão funções imunológicas também são remodeladas devido a infecção (VIEIRA *et al.*, 2015). Além desses eventos, no início da infecção, há um aumento do estresse oxidativo no intestino do vetor, sendo considerada a primeira linha de defesa do triatomíneo (COSENTINO-GOMES; ROCCO-MACHADO; MEYER-FERNANDES, 2014).

Após a colonização do trato digestivo, o *T. rangeli* atravessa o epitélio intestinal, alcança a hemolinfa e inicia uma intensa replicação por divisão binária. No início da infecção da hemolinfa, a forma evolutiva do parasito mais numerosa é a forma epimastigota morfologicamente mais curta, que com o passar do tempo se diferencia em epimastigotas longos (AÑEZ, 1983). A infecção da hemolinfa ativa as respostas imunológicas do vetor, como atividade da lisozima, produção de moléculas tripanolíticas, ativação da via da profenoloxidase, produção de superóxido e óxido nítrico, aglutinação da hemolinfa, fagocitose, micro agregação de hemócitos e formação de nódulos (GUARNERI; LORENZO, 2017).

As respostas imunológicas desempenhadas pelo vetor podem variar dependendo da cepa, da forma evolutiva e variantes da mesma forma evolutiva do parasito. Por exemplo, as formas epimastigotas longas ativam menos a resposta imunológica do triatomíneo que ocorre na hemolinfa e assim, não são tão afetadas, enquanto que para os epimastigotas curtos acontece o processo oposto (GOMES *et al.*, 2003; GUARNERI; LORENZO, 2017; WHITTEN *et al.*, 2001). Além disso, as subpopulações de *T. rangeli* caracterizadas como KP1 (-) sofrem danos devido à uma proteína tripanolítica existente na hemolinfa de *R. prolixus*, enquanto cepas

KP1 (+) não são reconhecidas por essa proteína, proporcionando o estabelecimento da infecção (PULIDO; PÉREZ; VALLEJO, 2008). A entrada do parasito nas glândulas salivares é um processo complexo e que não ocorre em todos os triatomíneos. O *T. rangeli* pode infectar alguns gêneros de triatomíneos, como o *Triatoma* e *Panstrongylus*, nos quais é restrito ao tecido intestinal, e o gênero *Rhodnius*, em que além do intestino, o parasito alcança as glândulas salivares (CUBA CUBA, 1998; MAIA DA SILVA *et al.*, 2007; MARQUEZ *et al.*, 2006). A invasão de *T. rangeli* nas glândulas salivares leva em torno de 10 dias após a colonização da hemolinfa (CUBA CUBA, 1974; ELLIS; EVANS; STAMFORD, 1980). Para que ocorra a entrada é necessária a adesão de epimastigotas longos na superfície da glândula salivar, que geralmente ocorrem em aglomerados de parasitos (DOS-SANTOS *et al.*, 2012; ELLIS; EVANS; STAMFORD, 1980). O processo de adesão às glândulas salivares possui uma estrita relação com a atividade de tirosinas fosfatases presentes na superfície do parasito, onde a inibição da enzima causa grande redução na penetração das glândulas (DOS-SANTOS *et al.*, 2012). Além da desfosforilação de resíduos de tirosina presentes na superfície externa das glândulas, ocorre uma interação entre carboidratos e lectinas presentes na glândula salivar e superfície de epimastigotas longos para ocorrer a adesão (BASSERI; TEW; RATCLIFFE, 2002).

Para a entrada do parasito na glândula salivar são formados poros na superfície da glândula por meio da proteína rangelisina (MEIRELLES *et al.*, 2005). Então, ocorre a introdução do *T. rangeli* envolto pela membrana celular, criando vacúolos (ELLIS; EVANS; STAMFORD, 1980). Ao chegar ao lúmen da glândula, o vacúolo é desfeito e os parasitos permanecem aderidos na superfície interna até diferenciarem-se em formas tripomastigotas metacíclicas (HECKER; SCHWARZENBACH; RUDIN, 1990; MEIRELLES *et al.*, 2005). A colonização das glândulas salivares dificulta a alimentação de *R. prolixus*, diminuindo a quantidade de sangue ingerido e a periodicidade das refeições (AÑEZ, 1984). Ainda, ocorre uma diminuição de proteínas totais presentes na saliva, sendo algumas dessas moléculas responsáveis por neutralizarem as respostas hemostáticas liberadas por mamíferos quando ocorre perda de sangue (RIBEIRO *et al.*, 2004). Uma provável razão para a diminuição dessas moléculas seria a utilização delas para a nutrição do próprio parasito (PAIM *et al.*, 2013).

1.3 ESTRUTURA GENÔMICA DE TRIPANOSSOMATÍDEOS

Os tripanossomatídeos são organismos que infectam uma gama de hospedeiros e, em cada ambiente distinto, precisam se adaptar às alterações que ocorrem no local por meio de ajustes na expressão gênica (CLAYTON, 2002). Esses protozoários estão no início da árvore evolutiva dos eucariotos e não regulam a iniciação da transcrição ao nível de genes individuais (CORDON-OBAS *et al.*, 2022). Para entender melhor o funcionamento da regulação da transcrição realizada pelos tripanossomatídeos é preciso olhar inicialmente para a organização e estrutura genômica.

Os tripanossomatídeos são considerados diploides, ainda que sejam identificados casos de aneuploidia e poliploidia entre cepas e até clones da mesma cepa para *T. cruzi* e *Leishmania* spp. (REIS-CUNHA *et al.*, 2015; ROGERS *et al.*, 2011; BLACK *et al.*, 2023). O *T. cruzi* possui 41 pares de cromossomos, com genoma haploide de 55 Mb de tamanho, codificando aproximadamente 12.000 genes (EL-SAYED *et al.*, 2005a; WEATHERLY; BOEHLKE; TARLETON, 2009). O *T. brucei* contém 11 pares de cromossomos e genoma haploide de 25 Mb com 9.068 genes (EL-SAYED *et al.*, 2005a). Por outro lado, a *Leishmania major* tem genoma haploide de 33 Mb, contendo 8.311 genes, organizado em 36 pares de cromossomos (EL-SAYED *et al.*, 2005a). O *T. rangeli* contém genoma estimado de 24 Mb, com 7.613 genes, mas ainda sem número exato de cromossomos definido (STOCO *et al.*, 2014).

Apesar das diferenças em tamanho do genoma e quantidade de cromossomos de cada organismo, muitos genes ortólogos são encontrados nessas espécies, e também há uma conservação da ordem dos genes (sintenia) (DE GAUDENZI *et al.*, 2011). Além disso, os genomas dos tripanossomatídeos tipicamente não possuem íntrons e se organizam por agrupamentos de genes policistrônicos, ou seja, há conglomerados de genes que codificam proteínas ordenadas sequencialmente. Além disso, ainda, são encontradas muitas sequências repetitivas no genoma. O genoma de *T. cruzi*, por exemplo, possui cerca de 50% de seu genoma constituído por sequências repetitivas, ao passo que *L. major* e *T. brucei* apresentam 20 a 25% de repetições no genoma (EL-SAYED *et al.*, 2005a). Por outro lado, *T. rangeli* possui 6% de seu genoma composto por sequências

repetitivas, sendo, portanto, considerado o tripanossomatídeo com menor genoma e com menor conteúdo de repetição. (STOCO *et al.*, 2014).

1.3.1 Famílias multigênicas em tripanossomatídeos

Grande parte das sequências repetitivas no genoma é composta por grandes famílias multigênicas. No caso de *T. cruzi*, estima-se que 50% do genoma do parasito codifica genes das famílias multigênicas (EL-SAYED *et al.*, 2005b). Dentre as famílias multigênicas estão genes que codificam para proteínas de superfície, como as trans-sialidasas (TSs), mucinas, proteínas de superfície associadas à mucina (MASPs), amastinas e GP63, assim como outras famílias como RHS (*Retrotransposon Hot Spot*) e DGF-1 (*Dispersed Gene Family 1*). Estas proteínas desempenham papéis importantes para sobrevivência e estabelecimento da infecção, estando relacionadas ao processo de adesão às células hospedeiras, estratégias de proteção contra o sistema imune e na sinalização celular (PECH-CANUL; MONTEÓN; SOLÍS-OVIEDO, 2017).

Os membros da família TSs estão presentes em toda a superfície do *T. cruzi*, majoritariamente ao longo do flagelo, corpo celular e bolsa flagelar (FREVERT; SCHENKMAN; NUSSENZWEIG, 1992; HERREROS-CABELLO *et al.*, 2020). Sua função mais descrita é a atividade catalítica de trans-sialidase, na qual o parasito, que não sintetiza ácido siálico, cliva e transfere ácidos siálicos dos hospedeiros para moléculas aceptoras, como as mucinas, também localizadas na membrana do parasito (PEREIRA *et al.*, 1980). Ao contrário de *T. cruzi*, as TSs de *T. rangeli* não possuem atividade trans-sialidase, apenas sialidase (AMAYA *et al.*, 2003), ou seja, apenas hidrolisam o ácido siálico, não tendo a capacidade de transferir para um aceptor. Nas formas epimastigotas do parasito, as sialidasas expressas podem ser liberadas em meio de cultura axênica (PEREIRA; MOSS, 1985). Quando comparadas com as sialidasas de *T. cruzi*, essas enzimas apresentam maior taxa de hidrólise do ácido siálico e menor pH ótimo para atividade catalítica (PEREIRA; MOSS, 1985; PONTES-DE-CARVALHO; TOMLINSON; NUSSENZWEIG, 1993). Entretanto, apesar dessas diferenças, quando sequências aminoacídicas correspondentes a essas enzimas de *T. rangeli* e *T. cruzi* são comparadas, há uma similaridade de 70 a 81% (BUSCHIAZZO *et al.*, 1993).

Os integrantes dessa família podem ser divididos em oito grupos, dependendo dos motivos e atividades específicas que a proteína apresenta

(FREITAS *et al.*, 2011). Apenas TSs do grupo I têm atividade catalítica trans-sialidase (SCHENKMAN *et al.*, 1991). No grupo II de TSs consta algumas glicoproteínas envolvidas na ligação e invasão da célula hospedeira (MATTOS *et al.*, 2014), e membros do grupo III estão relacionados a inativação das vias clássicas e alternativas do sistema complemento (BEUCHER; NORRIS, 2008). Curiosamente, *T. rangeli*, assim como *T. cruzi*, é o único parasito que apresenta pelo menos um membro em cada grupo TSs, enquanto *T. brucei* apresenta apenas membros do grupo I (HERREROS-CABELLO *et al.*, 2020).

As mucinas são uma grande família de genes em *T. cruzi*, sendo a proteína mais expressa na membrana do parasito (CALLEJAS-HERNÁNDEZ *et al.*, 2018). No *T. rangeli* houve uma redução do número de membros dessa família (STOCO *et al.*, 2014). As mucinas são caracterizadas principalmente por serem aceptores de ácido siálico (SCHENKMAN *et al.*, 1993), por participar no processo de adesão e invasão de células hospedeira e pela proteção do parasito contra os mecanismos de defesa dos hospedeiros (BUSCAGLIA *et al.*, 2006).

As MASPs recebem este nome por estarem dispostas entre grupos de genes de TSs e mucinas (BARTHOLOMEU *et al.*, 2009). Esses genes podem ser encontrados superexpressos em formas infectantes do *T. cruzi*, como as formas tripomastigotas metacíclicas, sendo importantes no processo de invasão celular (DOS SANTOS *et al.*, 2012). Assim como as famílias TS e mucinas, as MASPs têm número de cópias gênicas reduzidas no genoma de *T. rangeli* comparativamente ao *T. cruzi* (STOCO *et al.*, 2014).

A família GP63 é composta por metaloproteases dependentes de zinco, que estão envolvidas no processo de infecção celular e patogenicidade, sendo encontradas em formas amastigotas de *T. cruzi* e promastigotas de *Leishmania* spp. (CUEVAS; CAZZULO; SÁNCHEZ, 2003; SANTOS; BRANQUINHA; D'AVILA-LEVY, 2006). Interessantemente, são encontrados semelhantes números de cópias de GP63 entre *T. cruzi* e *T. rangeli*, apesar do papel em *T. rangeli* ser desconhecido (BRADWELL *et al.*, 2018; STOCO *et al.*, 2014).

A família DGF-1 é considerada a quinta maior família multigênica de *T. cruzi* (EL-SAYED *et al.*, 2005b). Por meio de estudos *in silico*, genes DGF-1 foram identificados codificando proteínas localizadas na superfície celular (WINCKER; MURTO-DOVALES; GOLDENBERG, 1992), contudo, proteínas DGF-1 foram localizadas intracelularmente em formas tripomastigotas e amastigotas de *T. cruzi*,

demonstrando que a localização dessas proteínas não é de superfície, como sugerido inicialmente (LANDER *et al.*, 2010). Seu papel ainda não é bem elucidado, entretanto, as proteínas DGF-1 podem estar correlacionadas a adesão dos parasitos às proteínas da matriz extracelular de células hospedeiras (KAWASHITA *et al.*, 2009).

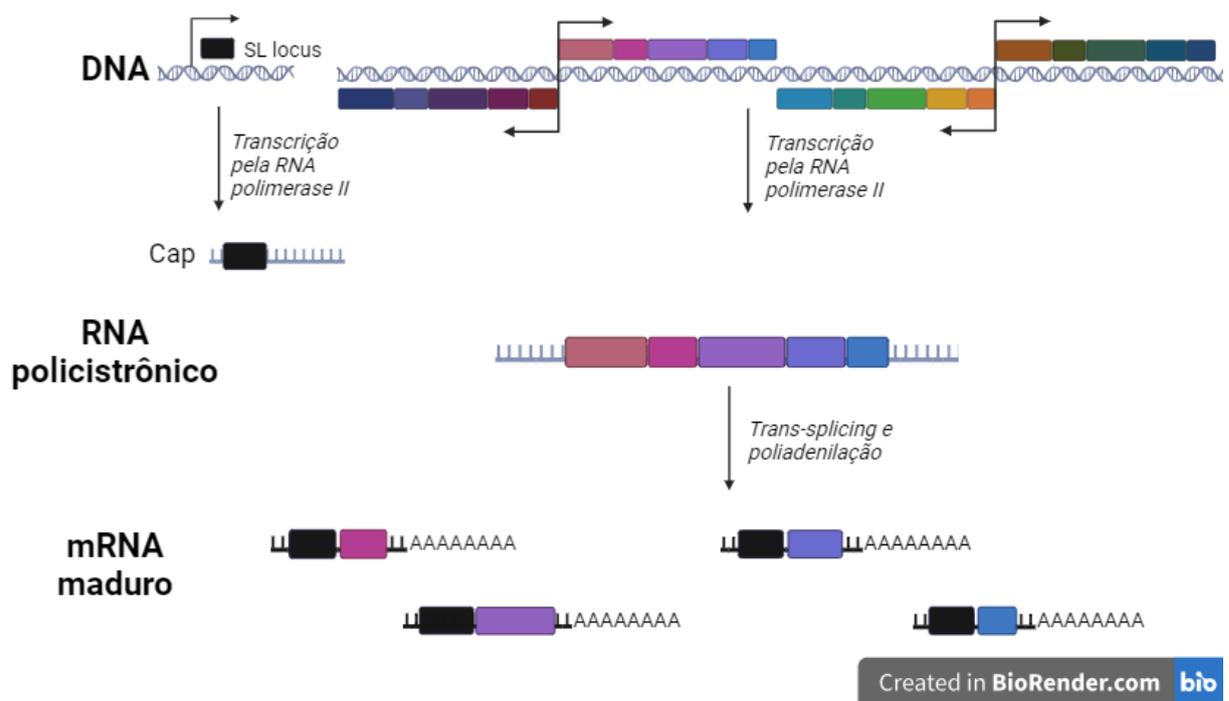
A família RHS é caracterizada por possuir uma região favorável para a inserção de retrotransposon dentro do gene RHS. Dessa forma, quando ocorre a inserção de retrotransposons, forma-se pseudogenes RHS (BERNARDO *et al.*, 2020). Com função ainda desconhecida, estima-se que os membros dessa família estão relacionados no controle de expansão dos retroelementos, da estrutura da cromatina e na expressão gênica em tripanossomatídeos (BERNARDO *et al.*, 2020; BRINGAUD *et al.*, 2002). Curiosamente, são encontrados maior número de cópias gênicas da família RHS em *T. rangeli* do que em *T. cruzi* (STOCO *et al.*, 2014).

1.4 A TRANSCRIÇÃO EM TRIPANOSSOMATÍDEOS

Os tripanossomatídeos, em seus ciclos de vida, precisam diferenciar-se em diferentes formas biológicas, dependendo do hospedeiro e do órgão que está parasitando no hospedeiro. Para cada momento haverá RNAs e proteínas chave que irão contribuir para o sucesso da infecção, portanto, é necessário que ocorra a transcrição desses genes e a regulação de expressão. Curiosamente, a transcrição realizada por tripanossomatídeos é policistrônica, ou seja, são transcritos diversos genes dispostos sequencialmente na mesma fita de DNA, formando um longo RNA com inúmeras sequências codificadoras de proteínas. Contudo, os genes que estão agrupados, podem não ter funções semelhantes ou relacionáveis (HERREROS-CABELLO *et al.*, 2020). Para *T. cruzi* os agrupamentos de genes variam de 30 a 500 kb, sendo separados por regiões de troca de fita divergentes (cromatina aberta) ou convergentes (cromatina condensada). Por conseguinte, centenas de genes podem ser transcritos conjuntamente pela RNA polimerase II (RNAP II) (HERREROS-CABELLO *et al.*, 2020). Recentemente foram identificados em *T. brucei* dois promotores centrais da RNAP II localizados em regiões com unidades de transcrição policistrônica divergente, cada um direcionando a transcrição unidirecional (CORDON-OBRAS *et al.*, 2022).

Concomitantemente à transcrição dos RNAs policistrônicos, a RNAP II transcreve genes do *spliced leader* (SL), também denominado de mini-éxon. A próxima etapa na maturação do mRNA acontece por meio do trans-splicing e poliadenilação. Em tripanossomatídeos, o trans-splicing compreende a etapa de incorporação do SL (~40 nucleotídeos) na porção 5' de cada transcrito, fornecendo uma estrutura *cap* (DE GAUDENZI *et al.*, 2011). A introdução do *cap*-SL garante estabilidade ao mRNA e promove a separação dos mRNA do RNA policistrônico, possibilitando a poliadenilação na região 3' do mRNA (Figura 2).

Figura 2 - Processo de transcrição em tripanossomatídeos



Legenda: Representação da transcrição e produção do mRNA maduro. O *Spliced leader* (SL), assim como o RNA policistrônico é transcrito por meio da RNA polimerase II. O RNA policistrônico passa pelo processo de *trans-splicing*, com a inserção do Cap, em seguida passa por clivagem e poliadenilação, transformando-se em vários mRNAs monocistrônicos maduros. Os quadros coloridos representam genes hipotéticos. Fonte: elaborado pela autora. Criado no BioRender.com.

Com a ausência de promotores clássicos da RNAP II que regulariam o início da transcrição, como ocorre em outros eucariotos (CLAYTON, 2002), a regulação dos níveis de expressão acontece majoritariamente a nível pós-transcricional. Entretanto, são identificadas expressões diferenciais durante o ciclo de vida de tripanossomatídeos (MINNING *et al.*, 2009; RASTROJO *et al.*, 2013). Dentre as

principais estratégias de regulação de expressão gênica, estão, mecanismos que proporcionam a estabilidade ou instabilidade dos transcritos, a duplicação de genes, regulação de histonas e a eficiência na tradução (PALENCHAR; BELLOFATTO, 2006; THOMAS *et al.*, 2009).

1.5 O PAPEL DAS CIÊNCIAS ÔMICAS PARA COMPREENSÃO DE EVENTOS BIOLÓGICOS

As tecnologias de sequenciamento de DNA surgiram na década de 70 (MAXAM; GILBERT, 1977; SANGER; NICKLEN; COULSON, 1977) e desde essa época, houve um grande desenvolvimento de tecnologias que foram e são essenciais para o entendimento de eventos biológicos. Essa evolução tecnológica permitiu o aperfeiçoamento do sequenciamento, aumentando o volume de dados obtidos, reduzindo o tempo e o custo (VAN DIJK *et al.*, 2014). O sequenciamento genômico permitiu desenvolver o entendimento e relações entre genótipo e fenótipo (KILPINEN; BARRETT, 2013), genômica agrícola (POLAND *et al.*, 2012), diagnóstico clínico (KINGSMORE; SAUNDERS, 2011), além do monitoramento e identificação de cepas patogênicas e doenças infecciosas (LIPKIN, 2013).

A vigilância e sequenciamento genômico de agentes etiológicos têm sido imprescindíveis para o combate às Doenças Tropicais Negligenciadas. O sequenciamento tanto dos agentes infecciosos quanto de seus vetores viabilizou o mapeamento genético de sequências de interesse, esclarecendo a filogenia, filogeografia, genética de populações, ecologia molecular (BRAS; GUERREIRO; HARDY, 2012) e identificação de perfis de resistência a medicamentos (MITCHELL; SIMNER, 2019). Isso refletiu em um aumento de trabalhos mais aplicados e de genômica funcional, incluindo os transcriptomas.

No contexto patógeno-hospedeiro, a abordagem transcriptômica tem auxiliado na elucidação de aspectos fisiológicos e de interação no contexto da infecção. Além disso, a metodologia de sequenciamento em larga escala de RNA (RNA-Seq), utilizada para gerar o transcriptoma, pode ser utilizada sem dados genômicos disponibilizados previamente, sendo útil principalmente quando organismos não modelos estão sendo estudados (VERA *et al.*, 2008). Esta metodologia é denominada *de novo* (do latim do zero) e não requer anotações genômicas completas, podendo auxiliar na identificação de novos genes, não identificados nas

anotações iniciais dos genomas de interesse (DILLON *et al.*, 2015). Como no caso do transcriptoma da cepa Sylvio X10/1 de *T. cruzi*, no qual dados de RNA-Seq foram utilizados para melhorar a anotação do genoma prévio, revelando que 79,95% do genoma corresponde a sequências codificadoras, ao passo que a análise genômica tinha estabelecido apenas 37,73% (CALLEJAS-HERNÁNDEZ *et al.*, 2019). Ademais, o RNA-seq proporciona uma quantificação mais precisa dos níveis de expressão, que posteriormente, podem ser confirmados por meio de qPCR (NAGALAKSHMI *et al.*, 2008).

Essa estratégia faz-se favorável quando se trabalha com organismos não modelos, como o *T. rangeli* e o *R. prolixus*, e quando não há um genoma contíguo disponível (VERA *et al.*, 2008). Especialmente porque as interações vetor-parasito são muito mais complexas do que as simuladas *in silico*. Portanto, o entendimento a nível molecular dos fenômenos que acontecem com o parasito durante a infecção do vetor é capaz de elucidar aspectos biológicos e moleculares, em especial, em tecidos relevantes para o estabelecimento da infecção, desenvolvimento e diferenciação do *T. rangeli* no hospedeiro invertebrado.

2 OBJETIVOS

2.1.1 Objetivo Geral

Avaliar as mudanças na transcrição gênica entre diferentes formas biológicas de *Trypanosoma rangeli* durante a infecção em *Rhodnius prolixus*.

2.1.2 Objetivos específicos

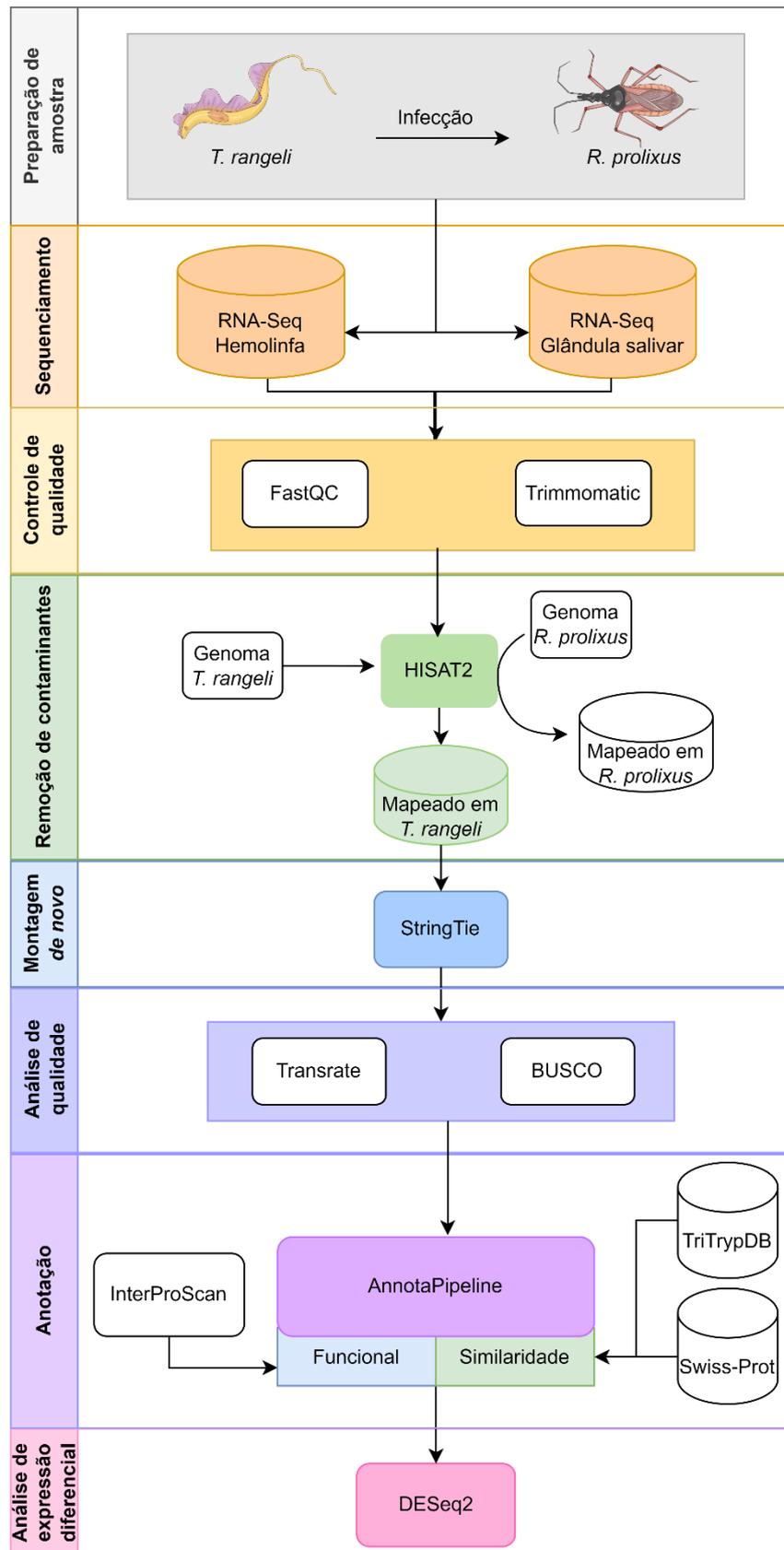
- Obter as sequências e quantificar a expressão dos transcritos de *T. rangeli* nas formas epimastigotas e tripomastigotas metacíclicas;
- Anotar e identificar termos de ontologia genética dos transcritos de *T. rangeli*;
- Identificar os transcritos diferencialmente expressos nas formas tripomastigotas metacíclicas e epimastigotas de *T. rangeli*.

3 METODOLOGIA

3.1 DELINEAMENTO EXPERIMENTAL

O delineamento experimental das análises executadas, bem como os programas utilizados estão resumidos na Figura 3. São representadas na imagem as etapas: (i) obtenção das amostras, (ii) sequenciamento da hemolinfa e glândula salivar de *R. prolixus* infectado por *T. rangeli*, (iii) controle de qualidade dos dados brutos obtidos do sequenciamento, (iv) remoção de contaminantes, (v) montagem do transcriptoma, (vi) análise de qualidade do transcriptoma, (vii) anotação dos transcritos obtidos, e (viii) análise de expressão diferencial.

Figura 3 - Desenho experimental para análises de montagem, anotação e expressão diferencial de transcritos de *T. rangeli*.



3.2 OBTENÇÃO DE DADOS

A metodologia descrita a seguir até a etapa de sequenciamento das bibliotecas foi realizada previamente ao desenvolvimento desse trabalho, mas será descrita por ser fundamental para a compreensão das etapas subsequentes do projeto.

3.2.1 Cultivo e diferenciação de parasitos

Formas epimastigotas de *T. rangeli* da cepa Choachí, linhagem KP1(+) foram cultivadas em meio LIT (do inglês *Liver Infusion Tryptose*) suplementado com 10% de soro bovino fetal (Cultilab), 50U/ml de penicilina e 50 µg/ml de estreptomicina à 27,5°C por passagens semanais. Estas cepas foram submetidas a passagens periódicas em camundongos e encontram-se criopreservadas no Laboratório de Protozoologia (MIP/CCB/UFSC).

3.2.2 Manutenção de triatomíneos

As colônias de *R. prolixus* foram mantidas no insetário do Laboratório de Protozoologia, em estufa BOD à 25°C e a umidade relativa de 75%. Semanalmente, as colônias foram alimentadas naturalmente diretamente com camundongos ou ratos, que são sedados e anestesiados previamente com Cetamina (100 mg/kg) e Xilazina (10 mg/kg). Os animais utilizados na alimentação dos barbeiros são fornecidos pelo Biotério Central da Universidade Federal de Santa Catarina e são manipulados de acordo com as normas do Conselho Nacional de Controle de Experimentação Animal (CONCEA). Esse procedimento tem aprovação na CEUA/UFSC sob o protocolo 4195070618.

3.2.3 Infecção de *R. prolixus* por *T. rangeli*

Para realizar a infecção dos insetos foram utilizadas ninfas de terceiro a quinto estágio. Os epimastigotas mantidos em cultivo foram centrifugados e tiveram seu meio de cultura removido, em seguida, foi adicionado PBS (Tampão salina-fosfato) pH 7,4, ajustando a concentração para 10⁶ parasitos/ml. A infecção foi realizada por via intracelômica com cerca de 5 µl da suspensão de parasitos. Foram infectados 150 triatomíneos para coleta da hemolinfa e 150 para obtenção das

glândulas salivares, sendo distribuídos 50 animais para cada réplica/biblioteca. O desenvolvimento da infecção foi monitorado durante 30 dias e após esse período, a hemolinfa e glândulas salivares foram coletadas para posterior extração de RNA.

3.2.4 Extração de RNA, construção das bibliotecas e sequenciamento

O RNA das amostras de hemolinfa e glândulas salivares (todas as amostras tinham triplicatas biológicas) foi extraído utilizando o kit PureLink™ RNA (Invitrogen) seguindo o protocolo estabelecido pelo fabricante. A qualidade e concentração dos RNAs obtidos foi verificada utilizando Bioanalyzer 2100 (Agilent Technologies). As amostras foram encaminhadas para o Instituto Karolinska, na Suécia, para a construção das bibliotecas e sequenciamento.

O DNA complementar (cDNA) foi sintetizado a partir de 500 ng de RNA total, e as bibliotecas transcriptômicas construídas de acordo com o protocolo TruSeq Stranded mRNA (Illumina). O sequenciamento foi realizado utilizando-se a plataforma Illumina HiSeq2500, método de fragmentos de leitura pareados, com tamanho de 150 pares de base (pb).

3.3 ANÁLISES DE BIOINFORMÁTICA

3.3.1 Controle de qualidade dos dados de RNA-seq

Para a análise dos transcriptomas, primeiramente foi realizada a avaliação de qualidade dos fragmentos de leitura obtidos no sequenciamento utilizando o programa FastQC v0.11.8 (ANDREWS, 2010) e MultiQC v1.11 (EWELS *et al.*, 2016). Para remover quaisquer sequências de adaptadores Illumina e aparar bases de baixa qualidade, utilizando como parâmetro o índice de qualidade phred abaixo de 30, foi utilizada a ferramenta Trimmomatic v0.39 (BOLGER; LOHSE; USADEL, 2014).

Em seguida, para realizar a montagem do transcriptoma de *T. rangeli* foi realizada uma etapa de remoção de fragmentos de leitura contaminantes. Para isso foi realizado um alinhamento dos fragmentos de leitura de RNA-Seq contra os dados genômicos de *R. prolixus* (GCA_000181055.3) através do programa HISAT2 v2.2.1 (KIM *et al.*, 2019). Todos os fragmentos de leitura que mapearam nesse genoma foram removidos das análises utilizando a ferramenta SAMtools v1.7 (LI *et al.*, 2009). Em seguida, os fragmentos de leitura foram mapeados contra o genoma de *T.*

rangeli cepa SC58 (genoma montado pelo nosso grupo de pesquisa, dados não publicados), sendo os fragmentos de leitura que mapearam separados e ordenados utilizando pelo programa SAMtools v1.7 para serem utilizados na montagem do transcriptoma.

3.3.2 Montagem e anotação do transcriptoma

A montagem do transcriptoma foi desempenhada pelo programa StringTie v2.2.1 (PERTEA *et al.*, 2015). Primeiro foi feita a montagem *de novo* individualmente das réplicas, tanto das amostras de hemolinfa, quanto de glândula salivar. Após as montagens individuais, foi realizada a junção dos transcriptomas individuais em um só, o que denominamos de montagem geral. O StringTie v2.2.1 proporciona a possibilidade de utilizar um arquivo de anotações gênicas e características genômicas para otimizar a montagem geral. Então, nessa etapa foi realizada uma montagem geral com a referência de *T. rangeli* cepa SC58. Após a montagem, foi realizada a análise de qualidade pelo programa Transrate v1.0.3 (SMITH-UNNA *et al.*, 2016) para obtenção de métricas como N50, N90, percentual de mapeamento dos fragmentos de leitura contra o transcriptoma montado, e pelo BUSCO v5.5.0 (MANNI *et al.*, 2021) para avaliar a representatividade do transcriptoma montado buscando ortólogos universais.

A etapa de anotação dos transcritos foi realizada pelo AnnotaPipeline v1.0 (MAIA *et al.*, 2022). Resumidamente, as sequências nucleotídicas dos transcritos foram traduzidas na primeira janela de leitura para sequências aminoacídicas por meio do programa EMBOSS v6.6.0 (RICE; LONGDEN; BLEASBY, 2000). Foram selecionadas as sequências proteicas não redundantes e, então, foi realizada análise de similaridade através do programa BLASTp (CAMACHO *et al.*, 2009). Os bancos de dados utilizados para análise de similaridade foram o SwissProt (*release* 2023_04) (BAIROCH, 1996) e TriTrypDB v62 (AMOS *et al.*, 2022). Em seguida, o AnnotaPipeline realizou uma etapa de anotação funcional com o programa InterProScan v5.54-96 (JONES *et al.*, 2014).

3.3.3 Análise de expressão diferencial

Os transcritos que obtiveram sequências aminoacídicas truncadas foram retirados desta análise, e em seguida, foi utilizado o modo de estimativa de

expressão disponibilizado pelo programa StringTie v2.2.1. A análise de expressão diferencial foi realizada utilizando o DESeq2 v1.40.2 (LOVE; HUBER; ANDERS, 2014), comparando a expressão do grupo tripomastigota em relação ao grupo epimastigota de *T. rangeli*. Para tal, foram considerados como transcritos diferencialmente expressos aqueles que apresentaram \log_2 Fold Change (\log_2FC) ≥ 2 ou ≤ -2 e valor de p-ajustado $<0,01$.

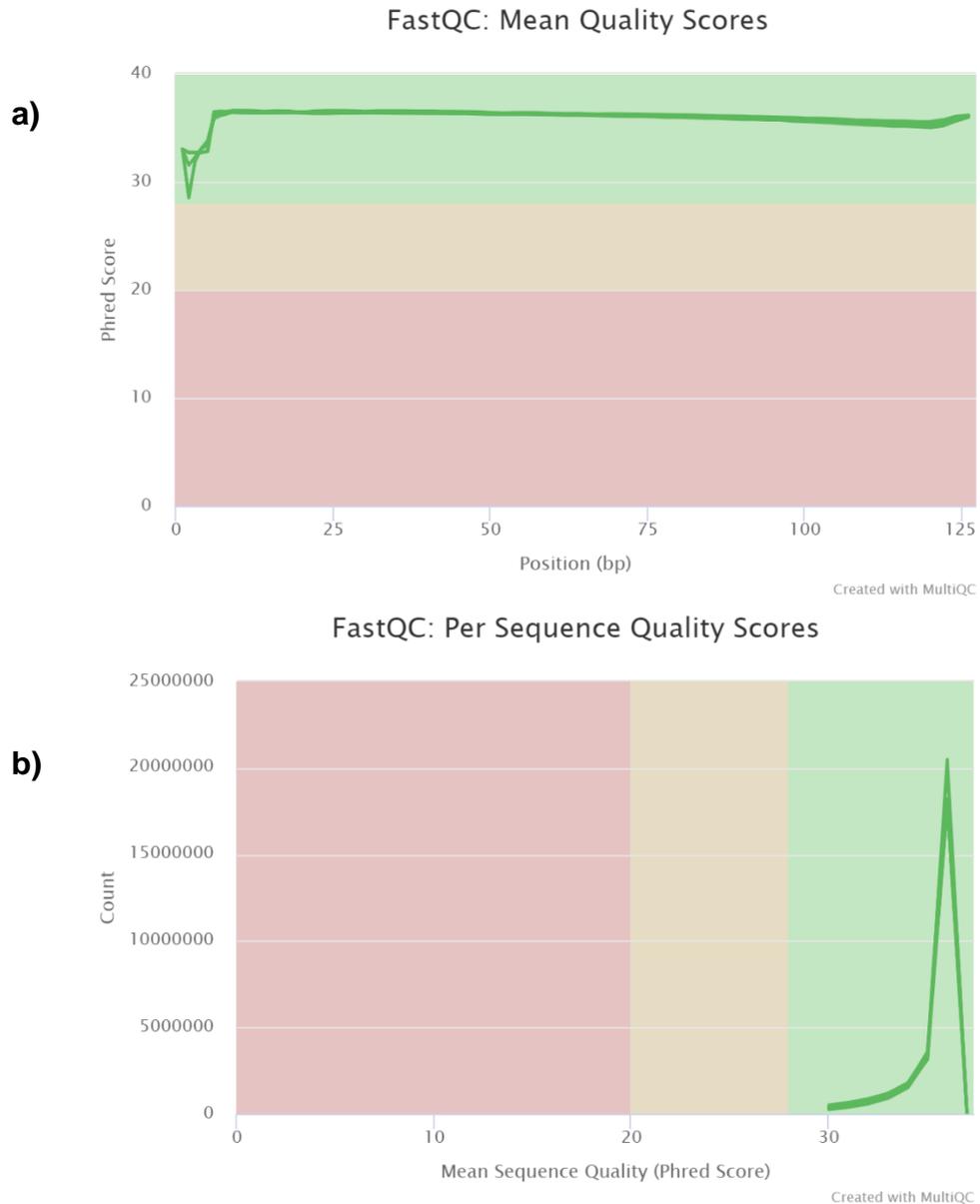
4 RESULTADOS E DISCUSSÃO

4.1 ANÁLISE DE QUALIDADE DO SEQUENCIAMENTO E DO MAPEAMENTO CONTRA O GENOMA DO VETOR

A partir do sequenciamento foi obtida uma média de 29,4 milhões de fragmentos de leitura para as amostras de glândula salivar de *R. prolixus*, com tamanho médio de 126 pb. Para as amostras originadas a partir da hemolinfa do inseto vetor, foram obtidos 23,8 milhões de fragmentos de leitura em média, com tamanho médio de 126 pb. Os dados brutos de sequenciamento passaram por análise de qualidade.

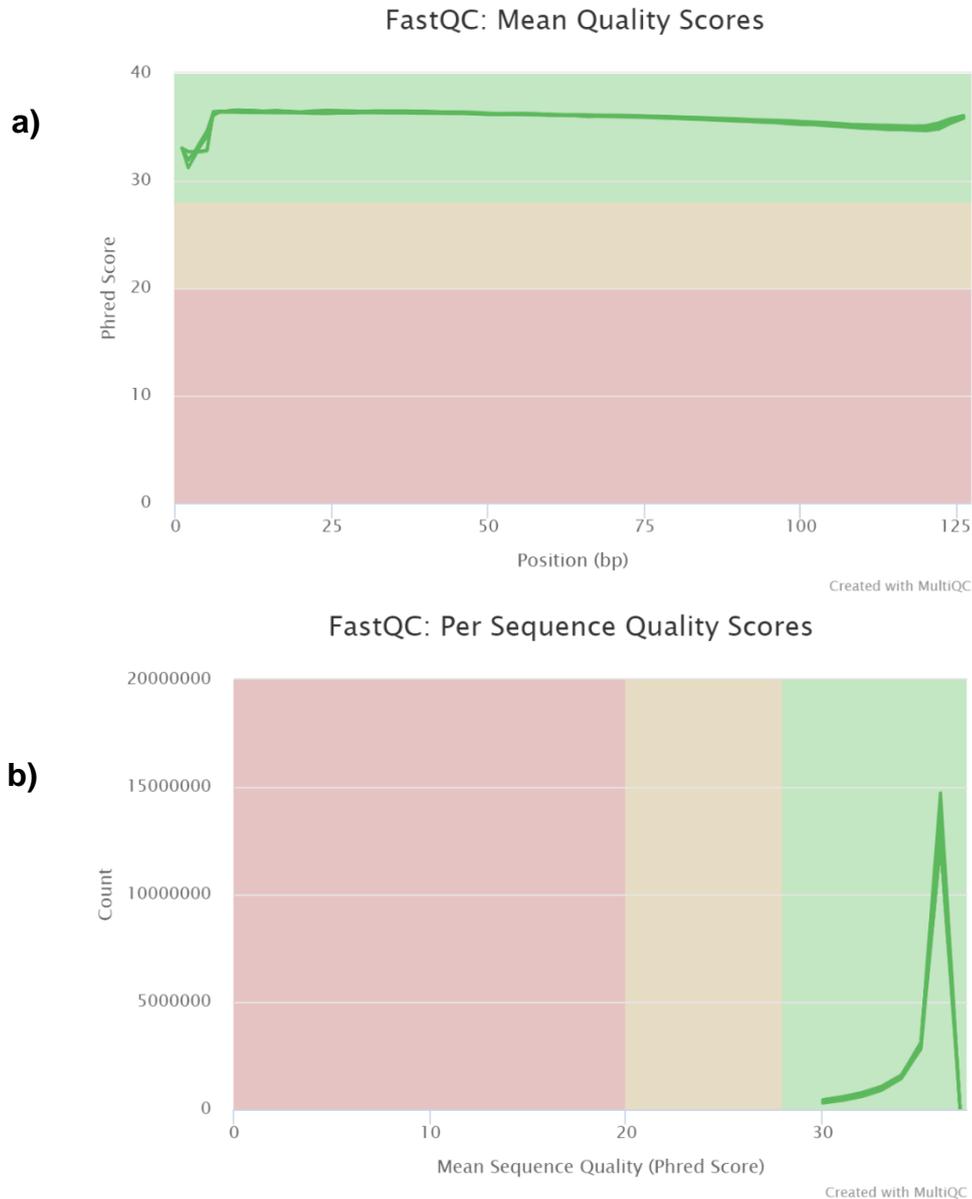
Posteriormente, foi realizada a retirada dos adaptadores utilizados no sequenciamento e remoção das bases de baixa qualidade. Em seguida, a qualidade dos dados foi reavaliada, por conseguinte, as amostras de glândula salivar obtiveram qualidade *phred* acima de 30 (Figura 4) e as amostras de hemolinfa apresentaram o mesmo padrão de qualidade (Figura 5).

Figura 4 - Valores de qualidade *phred* obtidos pelo programa MultiQC após etapa de controle de qualidade para dados brutos de RNAseq de glândula salivar de *R. prolixus* infectado com tripomastigotas de *T. rangeli*.



Legenda: As linhas verdes correspondem às médias atribuídas para cada réplica. A) Qualidade média por base sequenciada em cada réplica. No eixo Y temos a qualidade por escala phred (score: 10: 90% de acurácia; 20: 99% de acurácia; 30: 99,9% de acurácia). B) Qualidade média por sequência. No eixo X temos a qualidade por escala phred e no eixo Y temos as contagens de leituras. Fonte: elaborado pela autora.

Figura 5 - Valores de qualidade *phred* score obtidos pelo programa MultiQC após etapa de controle de qualidade para dados brutos de RNAseq de hemolinfa de *R. prolixus* infectado por epimastigotas de *T. rangeli*.



Legenda: As linhas verdes correspondem às médias atribuídas para cada réplica. A) Qualidade média por base sequenciada em cada réplica. No eixo Y temos a qualidade por escala phred (score: 10: 90% de acurácia; 20: 99% de acurácia; 30: 99,9% de acurácia). B) Qualidade média por sequência. No eixo X temos a qualidade por escala phred e no eixo Y temos as contagens de leituras. Fonte: elaborado pela autora.

Como o objetivo foi a análise do transcriptoma do *T. rangeli*, após a etapa de controle de qualidade foi realizado o mapeamento dos fragmentos de leitura contra o genoma do *R. prolixus* para retirada de sequências do vetor. A média de fragmentos

de leitura mapeados nas réplicas das bibliotecas de hemolinfa foi de 26,51% e nas bibliotecas de glândula salivar foi de 79,07%. Essa diferença na porcentagem das médias entre os tecidos pode ser respaldada pelo ciclo do parasito e pela quantidade de células do hospedeiro presente no tecido. O *T. rangeli* realiza a intensa replicação na hemolinfa (D'ALESSANDRO, 1976), local onde apenas hemócitos em pequena quantidade são encontrados. Em contrapartida, a glândula salivar de *R. prolixus* é composta por células epiteliais conectadas por fibras musculares por meio de uma lâmina basal espessa (BORELLA MARFIL ANHÊ *et al.*, 2021). À vista disso, as amostras da hemolinfa apresentam uma proporção maior de parasitos em relação a células do hospedeiro, e conseqüentemente mais RNA do parasito em comparação com o do vetor, causando essa redução na porcentagem de fragmentos de leitura que mapearam no genoma do vetor.

4.2 MONTAGEM E ANOTAÇÃO DOS TRANSCRITOS

A montagem do transcriptoma realizada pelo StringTie v2.2.1 resultou num transcriptoma com 15.199 transcritos, sendo identificados os 9.102 genes presentes no genoma de referência de *T. rangeli* SC58 e outros 6.097 novos transcritos. Os transcritos apresentaram tamanho médio de 1.812 pb, não estando muito distante do tamanho médio das regiões codificantes (CDS) do genoma do *T. rangeli* (STOCO *et al.*, 2014). O percentual do conteúdo GC dos transcritos foi de 54,52%, assemelhando-se aos 54,27% de GC das CDSs obtido por Stoco *et al.* (2014). Quanto aos *contigs* do transcriptoma montado, temos que o valor de N50 foi de 2.476 pb, no qual 50% dos transcritos montados possuem o tamanho mínimo de 2.476 pb, ao passo que para N90 (90%) o valor foi 923 pb. Ambas as métricas indicam que os transcritos montados não estão fragmentados e não apresentam tamanho inferior ao tamanho médio das CDS do genoma do *T. rangeli* (STOCO *et al.*, 2014).

Tabela 1 – Características gerais do transcriptoma de *T. rangeli*.

Nº total de transcritos	15.199
Nº de transcritos com equivalência a referência	9.102
Nº de novos transcritos	6.097

Tamanho médio dos transcritos (pb)	1.812
Porcentagem de GC	54,52%
N50 (pb)	2.476
N90 (pb)	923

Em relação à proporção de fragmentos de leituras mapeados nos transcritos montados, considerando que esses fragmentos de leituras foram os utilizados para a montagem, temos que 64,45% dos pares de leituras mapearam no transcriptoma. De forma que 58,74% dos pares de fragmento de leitura estão alinhando na mesma região, em um mesmo *contig* na montagem, orientados de forma esperada para sequenciamento Illumina (um par de encontro ao outro), métricas classificadas como bons mapeamentos de acordo com o Transrate. Em suma, 91,14% dos fragmentos de leitura mapeados no transcriptoma, foram classificados como bons mapeamentos.

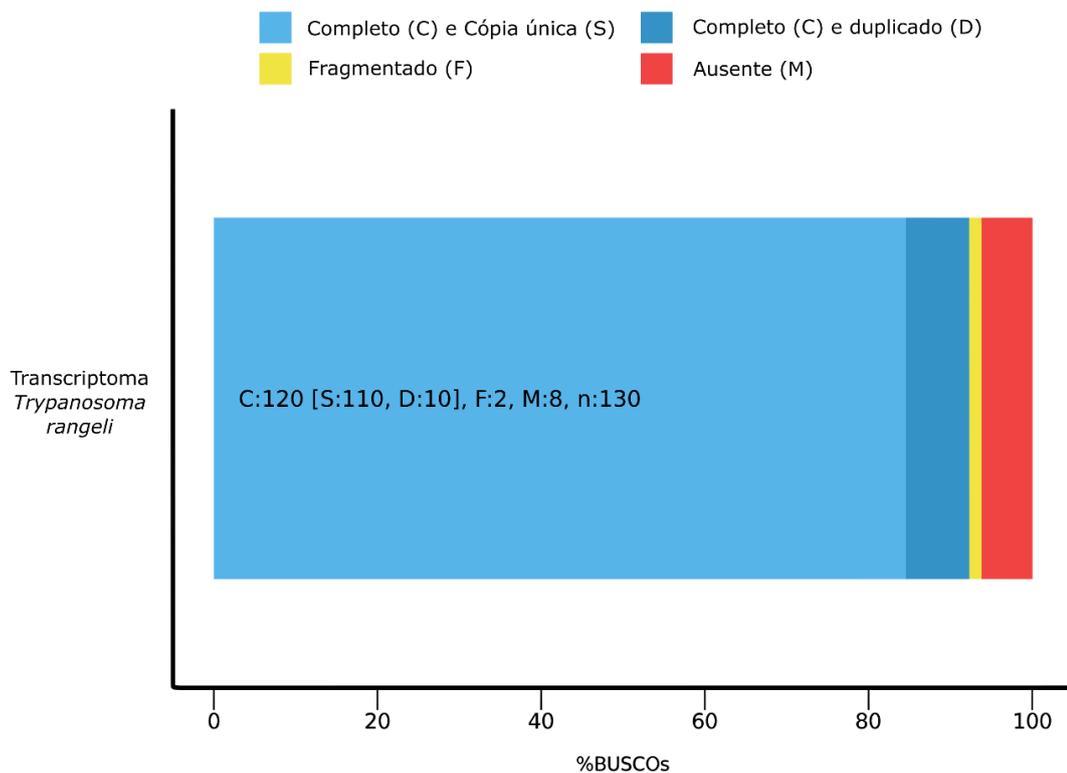
No entanto, todos os 15.199 *contigs* apresentaram uma baixa profundidade, com média de leitura por base menor que 10 (SMITH-UNNA *et al.*, 2016). Um fator que pode ter contribuído para a baixa profundidade foi a alta porcentagem de mapeamento das leituras das amostras de glândula salivar no genoma de *R. prolixus* na etapa de remoção de contaminantes, atingindo uma média de alinhamento maior do que 70% do total de fragmentos de leitura. Ao passo que, uma abordagem que poderia auxiliar na obtenção de *contigs* com maior profundidade é a utilização da opção que define a profundidade mínima necessária para as transcrições previstas, durante a etapa de montagem do transcriptoma.

Acerca da comparação do transcriptoma com as regiões codificantes do genoma (CDS) utilizado como referência (dados não publicados), o programa Transrate alinha os *contigs* com cada transcrito da referência utilizando o *Conditional Reciprocal Best BLAST* (CRBB), realizando alinhamentos bidirecionais (montagem → referência e referência → montagem) e selecionando os melhores resultados em ambos alinhamentos. Portanto, obtivemos uma cobertura da referência de 92,16%. Este resultado demonstra que foram montados grande parte dos transcritos já descritos no genoma do parasito.

No que se refere à composição do transcriptoma montado, buscamos ortólogos do filo Euglenozoa selecionados pelo programa BUSCO. Dos 130 genes

presentes no conjunto de dados de Euglenozoa, identificamos 120 genes (92%) dos ortólogos completos, 2 genes (1,5%) fragmentados e 8 (6,2%) genes ortólogos não foram identificados no transcriptoma (Figura 6). Por consequência, o transcriptoma apresenta uma completude necessária para ser considerada uma montagem de qualidade, tendo acima de 80% de correspondência dos genes ortólogos (RAGHAVAN *et al.*, 2022).

Figura 6 – Proporção de genes presentes no transcriptoma de *Trypanosoma rangeli* quando comparado com ortólogos do filo Euglenozoa.



Legenda: No eixo X temos a porcentagem de BUSCOs (do inglês, *Benchmarking Universal Single-Copy Orthologs*). Na barra, a cor azul clara representa a porcentagem de ortólogos identificados dentre os transcritos do *T. rangeli* que correspondem a genes completos (C) e de cópia única (S), a cor azul escura indica a porcentagem dos genes completos e duplicados (D), a cor amarela retrata a porcentagem dos genes fragmentados (F) e o vermelho equivale à porcentagem dos genes ausentes. Os valores dentro da barra indicam a quantidade de genes em cada categoria (C, S, D, F, M) e (n) representa a quantidade de ortólogos dentro do filo Euglenozoa.

Após a avaliação da qualidade do transcriptoma, foi iniciada a etapa de anotação dos transcritos. Para a anotação funcional foi necessário realizar a

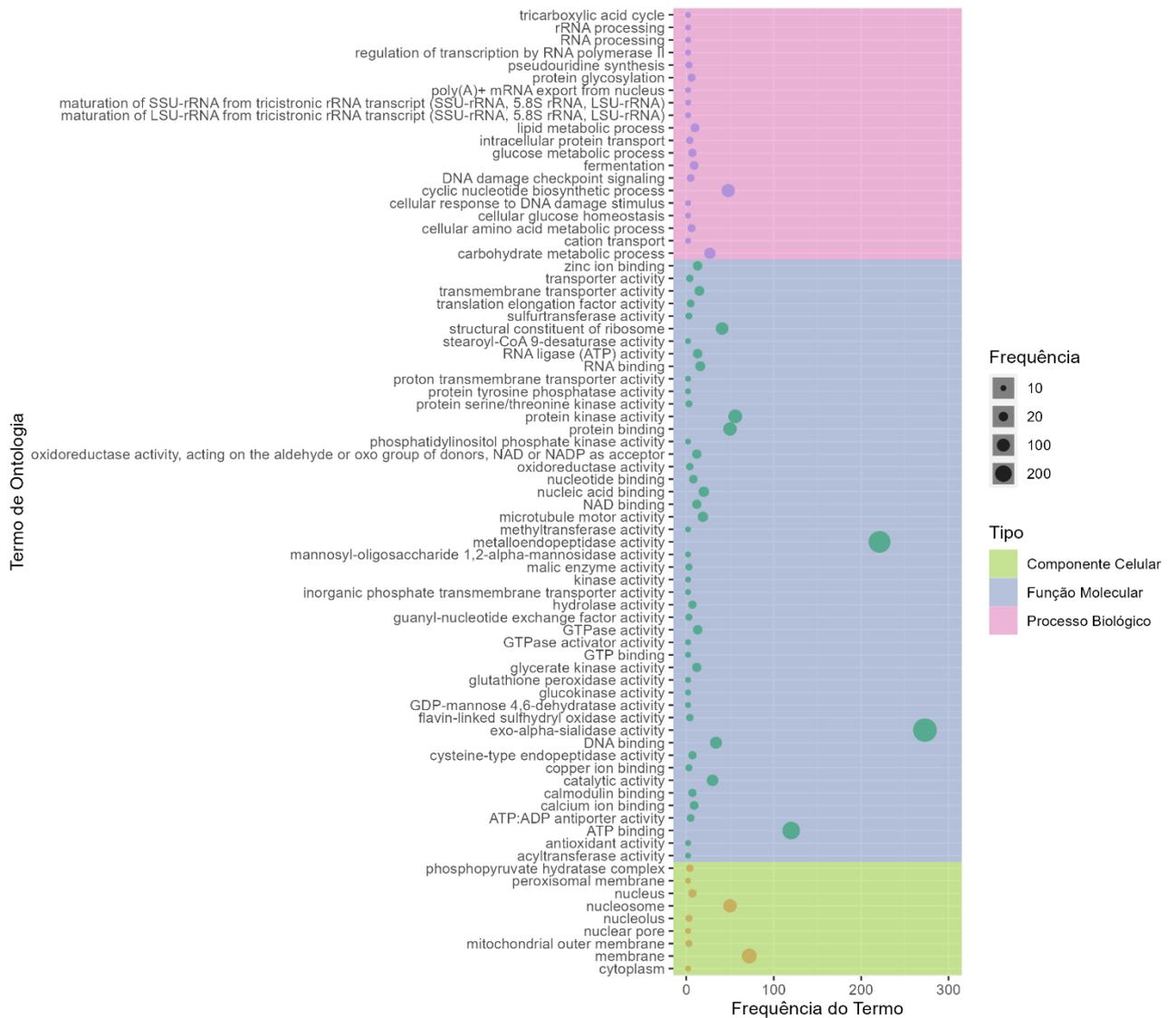
tradução das sequências nucleotídicas para sequências aminoacídicas e, em seguida, retirar as sequências truncadas, ou seja, as sequências aminoacídicas preditas no qual o códon de parada não estava no fim da sequência, mas no meio da proteína predita. Essa seleção de sequências aminoacídicas não truncadas foi necessária, pois a *pipeline* utilizada para realizar a anotação funcional não aceita as sequências truncadas. Dessa forma, quando retiradas as sequências truncadas, o total de 15.199 transcritos reduziu para 2.864, dos quais 2.359 são transcritos que obtiveram equivalência com a referência e 505 foram consideradas novas isoformas.

A quantidade de sequências truncadas encontradas está relacionada a um erro metodológico na etapa de tradução dos transcritos. De forma que os transcritos foram traduzidos apenas na primeira janela de leitura, ao invés de realizar a tradução nas seis janelas. Portanto, um grande número de transcritos que foram identificados com sequência truncada certamente não teria um códon de parada no meio da sequência se traduzidos em diferentes janelas de leitura. Todavia, a presença de sequências truncadas também pode ser um indicativo de que houve uma montagem incorreta, acarretando na montagem de quimeras. À vista disso, são necessárias mais análises para avaliar esse conjunto de dados.

Quanto à anotação dos transcritos com tradução correta, 800 foram consideradas proteínas hipotéticas e 2.064 receberam anotação específica. O AnnotaPipeline considera uma sequência aminoacídica como proteína hipotética quando há similaridade com sequências sem anotação específica disponíveis nos bancos de dados. Com relação a anotação funcional, das 2.864 proteínas preditas, 1.446 receberam termo de ontologia genética (GO), sendo que 82 foram caracterizadas como proteínas hipotéticas e também possuem GO.

A partir dessa etapa inicial de anotação, olhamos para os GOs mais frequentes e que obtiveram mais de uma contagem (Figura 7). Todos os termos identificados com contagem maior ou igual a um podem ser consultados no APÊNDICE A. Podemos notar que na categoria componente celular, há maior frequência nas categorias membrana e nucleossomo. As proteínas de superfície estão relacionadas aos processos de adesão celular, proteção contra o sistema imune do hospedeiro e conseqüentemente, no sucesso da infecção (PECH-CANUL; MONTEÓN; SOLÍS-OVIEDO, 2017). À vista disso, a maior frequência de proteínas na membrana condiz com a importância dessas proteínas para a sobrevivência e adaptação do parasito aos diferentes tecidos do vetor.

Figura 7 – Termos de ontologia genética mais frequentes entre os transcritos.



Legenda: No eixo Y temos os termos de ontologia genética, no eixo X temos a frequência do termo, ou seja, números de ocorrências de cada termo. Os termos estão divididos em três categorias: Componente celular – verde; Função molecular – azul; Processo biológico – rosa. O tamanho dos pontos refere-se à frequência do termo. Para essa análise foram considerados apenas os GOs com contagem maior que 1.

As proteínas presentes no nucleossomo estão basicamente relacionadas aos processos de replicação do DNA e transcrição. Os nucleossomos são dinâmicos, capazes de deslizar ao longo do DNA e se desmontar. Essa dinamicidade permite que o DNA esteja acessível às proteínas, incluindo reguladores da cromatina e mecanismos de transcrição e replicação (LAI; PUGH, 2017). No que concerne aos

termos referentes aos processos biológicos, observa-se uma maior frequência de transcritos classificados em processo biossintético de nucleotídeos cíclicos e processo metabólico de carboidratos. Dentre os nucleotídeos cíclicos em tripanossomatídeos, o AMP cíclico (cAMP) ganha destaque por estar envolvido em diversos processos celulares e desenvolvimento do ciclo de vida dos parasitos. O cAMP é sintetizado através da ação da adenilato ciclase, e em *Leishmania* spp., *T. cruzi* e *T. brucei* a sinalização de cAMP está relacionada no processo de diferenciação celular do parasito, tanto da forma promastigota para amastigota, no caso da *Leishmania* spp., quanto de epimastigotas para tripomastigotas em *T. cruzi* e das formas tripomastigotas sanguíneas longas para as curtas em *T. brucei*. Para esses parasitos, houve um aumento dos níveis de cAMP no estágio de diferenciação celular (SEEBECK; SCHAUB; JOHNER, 2004), portanto, a transcrição de genes envolvidos na síntese de cAMP podem estar implicados na manutenção de níveis elevados de cAMP.

O processo de metabolização de carboidratos é um processo necessário a todos os eucariotos para produzir energia e assim, sobreviver. A geração de energia é essencial para os processos de replicação, transcrição, tradução, transdução de sinal, entre outros. Dessa forma, durante o ciclo de vida, dependendo do hospedeiro e forma evolutiva, o parasito adapta a forma de obtenção de energia, seja pela via glicolítica, gliconeogênica e das pentoses-fosfato, obtendo energia de variadas fontes, como: sacarídeos, aminoácidos ou lipídeos (MICHELS *et al.*, 2021). Portanto, a alta frequência de termos de GO no processo de metabolização de carboidratos era esperado, considerando que estamos analisando transcritos de *T. rangeli* obtidos da infecção de *R. prolixus*, que estavam em processos de intensa replicação, diferenciação celular e adaptação as respostas imunológicas do vetor.

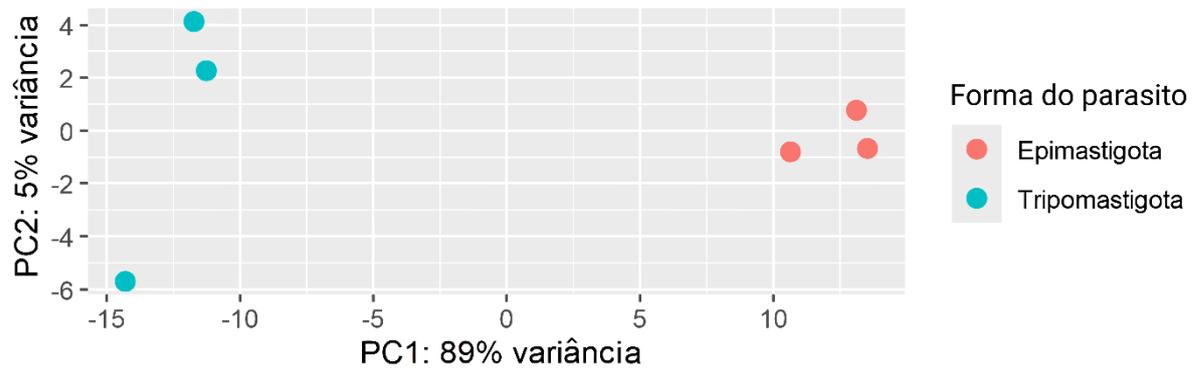
Em relação aos termos de função molecular mais frequentes, identificamos: atividade exo-alfa-sialidase e atividade de metaloendopeptidase. A atividade exo-alfa-sialidase representa ação realizada pelas sialidases, sendo definida como a catálise da hidrólise de determinadas ligações alfa-glicosídicas de resíduos siálicos em compostos oligossacarídeos, glicoproteínas, glicolipídios, entre outros (ROSENBERG; SCHENGRUND, 1976). Enquanto a atividade metaloendopeptidase refere-se ao processo de hidrólise de ligações alfa-peptídicas em uma cadeia polipeptídica realizado por enzimas endopeptidases dependentes da ligação de um metal. As metaloendopeptidases GP63 ganham destaque por apresentarem funções

relevantes durante o ciclo de vida de parasitos como *Leishmania* spp., onde no hospedeiro mamífero, as GP63 foram relatadas nos processos de evasão da lise mediada pelo complemento, facilitação da fagocitose por macrófagos, migração através da matriz extracelular, inibição de células natural *killer*, promoção da sobrevivência de amastigotas intracelulares, entre outros (D'AVILA-LEVY *et al.*, 2014). No entanto, essas proteínas também têm sido descritas desempenhando papéis durante a infecção do inseto vetor, incluindo a atividade proteolítica para aquisição de nutriente, migração através da matriz celular proteica de insetos e degradação das defesas dos insetos e o processo de adesão, onde a introdução de quelantes de metais prejudicaram a adesão de *T. cruzi* ao intestino de *R. prolixus* (D'AVILA-LEVY *et al.*, 2014; REBELLO *et al.*, 2019).

4.3 ANÁLISE DE EXPRESSÃO DIFERENCIAL

A análise de expressão diferencial foi desempenhada pelo DESeq2 comparando as amostras glândula salivar vs. hemolinfa, ou seja, formas tripomastigotas vs. epimastigotas. No primeiro momento, observamos a qualidade das réplicas por agrupamento de amostras por meio da Análise de componentes principais (PCA, do inglês, *principal component analysis*) gerada a partir da transformação por *vst* (do inglês, *variance stabilizing transformation*) das contagens normalizadas dos transcritos oriundos dos parasitos. Pode ser observado na figura 8 distintos agrupamentos, um contendo as réplicas das amostras de formas tripomastigotas, e o outro com as amostras de epimastigotas, mostrando similaridade entre as réplicas. De maneira oposta, há um distanciamento dos grupos epimastigotas e tripomastigotas, sendo observada maior porcentagem de variância entre os grupos.

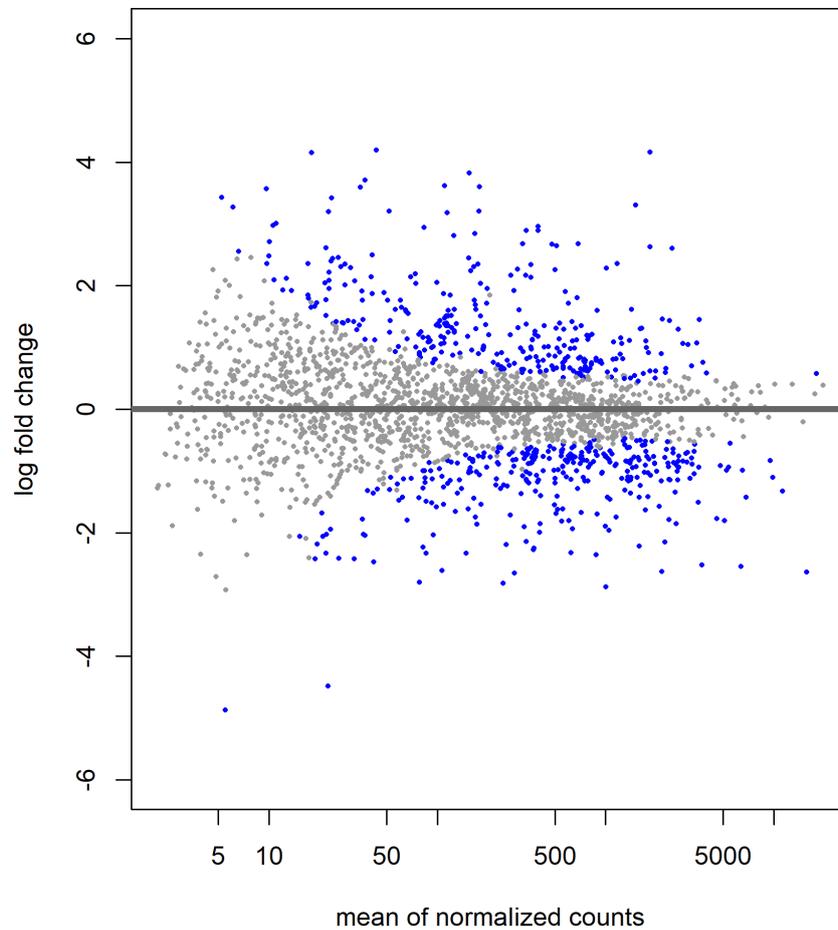
Figura 8 - Variância entre os grupos Epimastigotas e Tripomastigotas.



Legenda: Análise de Componentes Principais das Amostras (PCA) realizada a partir dos dados de contagem normalizados das réplicas utilizadas para montagem do transcriptoma. Os pontos vermelhos correspondem às amostras retiradas da hemolinfa do vetor, onde são encontradas formas epimastigotas do *T. rangeli*, e as azuis retiradas da glândula salivar do vetor, onde localizam-se as formas tripomastigotas do parasito.

A partir dos 2.864 transcritos, a análise de expressão diferencial resultou em 638 transcritos com valor de p-ajustado < 0.01 (Figura 9). Ao total obtivemos 103 transcritos com expressão diferencial, ou seja, valores de expressão $\log_{2}FC \geq 2$ ou ≤ -2 .

Figura 9 - Análise comparativa de expressão gênica entre as formas tripomastigotas e epimastigotas de *T. rangeli*.



Legenda: Níveis de expressão representados por \log *Fold Change* pela média das contagens normalizadas. Os pontos azuis representam os transcritos com valor de p-ajustado < 0.01 . Os pontos cinzas são os transcritos com valor de p-ajustado > 0.01 .

Para avaliar a representatividade dos 638 transcritos com valor de p-ajustado < 0.01 , procuramos por genes de transcrição constitutiva como GAPDH (Gliceraldeído-3-fosfato desidrogenase), nucleosídeo difosfato quinase (NDK), glicose-6-fosfato desidrogenase (G6PD), proteína ribossômica 60S e alguns genes de transcrição constitutiva de cópia única como: actina, triosefosfato isomerase, tripanotona redutase e a proteína ribossômica 40S S4. Desses genes, encontramos apenas 4 transcritos com anotação pertencente a subunidade ribossômica 60S e um transcrito que compõe a subunidade ribossômica 40S e seus respectivos níveis de expressão (Tabela 2).

Tabela 2 – Níveis de expressão em logFC de transcritos de *T. rangeli* anotados como genes de transcrição constitutiva.

Transcrito	Nível de Expressão (logFC)
Proteína ribossômica 40S S21	-0.57
Proteína de subunidade ribossômica grande P2-A	0.72
Proteína L31 da subunidade ribossômica 60S	0.61
Proteína L31 da subunidade ribossômica 60S	0.58
Proteína ribossômica 60S L29	1.08
Proteína ribossômica 60S L29	0.65

Os transcritos identificados que se esperava transcrição constitutiva de fato não apresentaram um aumento de expressão gênica que caracteriza expressão diferencial. Esse fato era esperado, pois genes de transcrição constitutiva apresentam níveis de expressão constantes entre as formas do parasito. No entanto, a não identificação de genes de transcrição constitutiva como GAPDH pode ser em virtude de a tradução dos transcritos ter sido realizada apenas na primeira janela de leitura, gerando uma grande quantidade de sequências truncadas que foram retiradas das análises posteriores. Conseqüentemente, esses transcritos com sequência aminoacídica predita truncada não tiveram os níveis de expressão analisados.

4.3.1 Genes com expressão gênica aumentada em tripomastigotas

Os genes podem ter tanto um aumento da transcrição quanto a diminuição da transcrição, tornando-os diferencialmente expressos quando comparamos condições. Dessa forma, neste trabalho, consideramos os transcritos que obtiveram níveis de expressão maior que 2 logFC como regulados positivamente e com aumento de expressão gênica. Os transcritos que obtiveram níveis de expressão menor que -2 logFC foram caracterizados como regulados negativamente e com diminuição de expressão gênica.

As formas tripomastigotas quando comparadas com as epimastigotas apresentaram 69 transcritos com aumento de expressão gênica. Desses, 56 transcritos receberam anotação de proteínas hipotéticas e apenas 13 receberam

anotação funcional. A quantidade de transcritos caracterizados como proteínas hipotéticas dificulta o entendimento dos eventos biológicos que estão ocorrendo com o parasito na glândula salivar do vetor. A anotação de transcritos como proteínas hipotéticas é prevista, considerando que *T. rangeli* não é um organismo modelo e que cerca de 65% dos genes de *T. rangeli* codificam proteínas hipotéticas (STOCO *et al.*, 2014). Esse cenário não é visto apenas em *T. rangeli*, mesmo para tripanossomatídeos modelos como o *T. brucei* e a *L. amazonensis*, 60% e 63% dos genes preditos são anotados como hipotéticos, respectivamente (SIEGEL *et al.*, 2011; TSCHOEKE *et al.*, 2014). No entanto, dessas 56 proteínas hipotéticas reguladas positivamente, 9 não obtiveram nenhuma correspondência com sequências homólogas dos bancos de dados. Enquanto as outras 47 proteínas hipotéticas obtiveram correspondência com sequências disponíveis nos bancos de dados que não possuem anotação conhecida, ou seja, também estão anotadas como proteínas hipotéticas. Sendo assim, esses transcritos que não tiveram semelhança com nenhuma proteína do banco de dados podem indicar a presença de quimeras e montagem incorreta de fragmentos de leitura. Quanto à localização predita das proteínas que esses transcritos codificariam, foram identificadas proteínas nucleares, extracelulares, mitocondriais, da membrana plasmática, do retículo endoplasmático e do citosol.

Sobre os transcritos que receberam anotação funcional, foram observadas proteínas de superfície ToIT, trans-sialidases do Grupo I, S-metil-5'-tioadenosina fosforilase, protease de superfície GP63, proteína quinase e proteína semelhante à cinesina. Os transcritos com anotações repetidas possuem média dos níveis de expressão representada (Tabela 3), caso a frequência da anotação seja maior ou igual a 3, é representado o desvio padrão entre parênteses junto a média dos níveis de expressão. A listagem de todos os transcritos com expressão gênica regulada positivamente e sua localização predita pode ser consultada no APÊNDICE B.

Tabela 3 –Transcritos de *T. rangeli* com anotação conhecida apresentando aumento de expressão gênica quando comparado formas tripomastigotas contra epimastigotas.

Transcrito	Nível de Expressão (logFC)	Média dos níveis de expressão (logFC) ^(a)
Proteína de superfície ToIT	3.29	

Protease de superfície GP63		2.29 (0.26)
Proteína semelhante à cinesina	2.34	
Trans-sialidase, Grupo I		2.44 (0.28)
Proteína quinase	2.16	
S-metil-5'-tioadenosina fosforilase		2.29

^a Média (desvio padrão).

As proteínas das famílias multigênicas de membrana como Trans-sialidase e GP63 são compostas por um grande número de genes, conseqüentemente, são encontrados vários membros dessas famílias sendo transcritos. No caso da família GP63, ela está relacionada à virulência e patogenicidade em parasitos patogênicos, como o *T. cruzi* (CUEVAS; CAZZULO; SÁNCHEZ, 2003). No entanto, membros da família GP63 foram relatados em parasitos não patogênicos como *Phytomonas* sp., *Herpetomonas* sp., *Leptomonas* sp. e *Crithidia* sp. (SANTOS; BRANQUINHA; D'AVILA-LEVY, 2006), como também já foi relatada a transcrição de GP63 em *T. rangeli* (GRISARD *et al.*, 2010). Em *T. cruzi*, foi demonstrado a influência das GP63 na interação parasito-vetor, no qual, o uso de quelantes de metais divalentes dificultaram o processo de adesão do parasito ao intestino médio de *R. prolixus*, além disso, esses parasitos isolados do intestino colonizado expressavam GP63 em grandes níveis (REBELLO *et al.*, 2019). Dessa forma, é possível que genes da família GP63 em *T. rangeli* desempenhem papéis importantes para o sucesso da infecção no vetor e as interações que ocorrem no processo de infecção.

Com relação aos membros da família Trans-sialidase, foram identificadas proteínas pertencentes ao grupo I, que incluem as proteínas que possuem atividade catalítica de sialidase. No *T. rangeli* essas proteínas possuem dois domínios *Asp box* (SxDxGxTW), um domínio VTV comum (VTVxNVxLYNR), e uma âncora GPI (glicosilfosfatidilinositol) (HERREROS-CABELLO *et al.*, 2020). Contudo, são necessárias mais análises para avaliar se esses transcritos possuem os domínios característicos de TS do grupo I. Ademais, apesar de *T. rangeli* não possuir proteínas TS com atividade trans-sialidase, apenas sialidase (AMAYA *et al.*, 2003), devido à metodologia de anotação, esses transcritos receberam anotação de trans-sialidase devido à similaridade com as sequências de TS disponíveis no banco de dados. Além disso, as TS em *T. cruzi* estão estritamente relacionadas a patogenicidade e infectividade (NARDY *et al.*, 2016), entretanto, como *T. rangeli* não

apresenta patogenicidade contra o hospedeiro mamífero, essas proteínas podem desempenhar papéis diferentes, mas ainda sendo relevantes para seu ciclo de vida, principalmente considerando a regulação positiva da transcrição desses genes.

A proteína ToIT, também uma proteína de membrana, recebe esse nome por apresentar homologia com proteínas ToIA de bactérias (MORGAN; ORTIZ; RIORDAN, 2014). Nesses organismos, a proteína ToIA fornece integridade estrutural à célula, e dessa forma, devido a semelhança, estima-se que a proteína ToIT desempenhe função semelhante em tripanossomatídeos (QUANQUIN *et al.*, 1999). Em *T. cruzi*, a proteína ToIT foi descrita localizando-se exclusivamente no flagelo do parasito de formas tripomastigotas (SABORIO *et al.*, 1990; LOBO *et al.*, 2019). Recentemente, foi confirmada a expressão de ToIT por RT-qPCR em *T. cruzi*, mostrando uma diferença significativa nos níveis de expressão quando comparado às formas tripomastigotas e epimastigotas (LOBO *et al.*, 2019), reforçando o resultado encontrado neste trabalho.

As proteínas quinases são essenciais para diversos processos como controle do ciclo celular, divisão celular, adaptação ao estresse e diferenciação celular (CAYLA *et al.*, 2022). Algumas proteínas quinases têm sido descritas como fundamentais para a diferenciação celular de *T. brucei*, como o gene MKK1 (MAP kinase kinase 1), que se mostrou essencial para a transmissão de *T. brucei* por *Glossina morsitans*. Quando o gene MKK1 foi deletado, as formas epimastigotas replicativas do parasito não conseguiram se diferenciar em formas tripomastigotas metacíclicas e assim, colonizar a glândula salivar do vetor (MORAND *et al.*, 2012). Como na glândula salivar de *R. prolixus* também ocorre a diferenciação de *T. rangeli* na forma tripomastigota metacíclicas (D'ALESSANDRO, 1976; D'ALESSANDRO; SARAIVA, 1992), a proteína quinase regulada positivamente pode exercer papel na diferenciação, no entanto, para averiguar qual quinase está sendo expressão por tripomastigotas de *T. rangeli* são necessárias mais análises.

Outras proteínas que desempenham papéis essenciais à célula são as cinesinas. São proteínas motoras baseadas em microtúbulos e realizam funções como: transporte intracelular de vesículas, organelas e complexos proteicos. As cinesinas formam uma superfamília com 15 subfamílias e possuem um domínio motor conservado (GILBERT; GUZIK-LENDRUM; RAYMENT, 2018). No caso dos kinetoplastídeos, existem cinesinas específicas dessa classe, mas os organismos ainda podem apresentar uma expansão desses genes, como é o caso de *T. brucei*.

Nesse parasito, foi descoberto que a proteína cinesina KIN-E é fundamental para a diferenciação celular do parasito, sendo responsável pelo transporte de uma proteína flagelar que promove o posicionamento do flagelo e a alteração de estruturas associadas ao flagelo para que os parasitos assumam a forma tripomastigotas (AN; LI, 2018). Assim, o aumento da expressão gênica de uma cinesina pode estar relacionado ao transporte de proteínas essenciais para a diferenciação em formas tripomastigotas, no entanto, são necessárias mais análises para afirmar esses eventos.

A S-metil-5'-tioadenosina fosforilase é uma enzima que tem sido alvo de drogas contra tripanossomíases, isso porque está envolvida no metabolismo de purinas e na via de resgate da metionina, principalmente em razão de que parasitos como os tripanossomatídeos não possuem as vias para biossíntese *de novo* de purinas (ABID *et al.*, 2017; EL KOUNI, 2003). Em um estudo analisando o secretoma de *T. cruzi*, foi identificado a excreção de S-metil-5'-tioadenosina fosforilase por formas tripomastigotas do parasito, mostrando que além de serem transcritos, esse gene também é traduzido e seu produto excretado por formas tripomastigotas (MACHADO, 2013).

4.3.2 Genes com expressão gênica aumentada em epimastigotas

Quando observamos os transcritos com expressão gênica regulada negativamente na comparação tripomastigotas vs. epimastigotas, na verdade estamos visualizando os transcritos que estão tendo aumento de expressão gênica nas formas epimastigotas. Dos 638 transcritos com valor de p-ajustado < 0.01, 34 transcritos apresentaram diminuição de expressão, desses, apenas 2 transcritos foram anotados como proteínas hipotéticas, um perfil bem diferente do visualizado no grupo de transcritos regulados positivamente. Dentre as anotações funcionais encontramos: cistationina- γ -liase, sialidases, transcetolase, tirosina aminotransferase, nodulinas-like, glicoproteína de superfície de 85 kDa, exo-alfa-sialidase, proteína reguladora do complemento, cinesina, transportador ABC e ATPase transportadora de cálcio. Os transcritos com anotações repetidas possuem média dos níveis de expressão representada (Tabela 4), caso a frequência da anotação seja maior ou igual a 3, é representado o desvio padrão entre parênteses junto a média dos níveis de expressão. A listagem de todos os transcritos com

expressão gênica regulada negativamente e sua localização predita pode ser consultada no APÊNDICE C.

Tabela 4 - Transcritos de *T. rangeli* com anotação conhecida apresentando diminuição de expressão gênica quando comparado formas tripomastigotas contra epimastigotas.

Transcrito	Nível de Expressão (logFC)	Média dos níveis de expressão (logFC) ^(a)
Cistationina-γ-liase		-2.71
Sialidase		-2.75 (0.68)
Transcetolase	-2.51	
Tirosina aminotransferase	-2.35	
Nodulinas-like		-2.18 (0.05)
85 kDa antígeno de superfície		-2.63 (1.00)
Proteína reguladora do complemento	-2.32	
Cinesina II 95 kDa		-2.08 (0.09)
transportador ABC	-2.42	
Metionina gama-liase	-2.32	
ATPase transportadora de cálcio	-2.03	

^a Média (desvio padrão).

Assim como nas formas tripomastigotas, as trans-sialidases também possuem um aumento de expressão nas formas epimastigotas, indicando possivelmente que proteínas distintas dessa família multigênica desempenham funções diferentes e importantes em mais de um estágio de vida de *T. rangeli*. A transcrição de TSs em epimastigotas já foi descrita em *T. cruzi*, sendo que essas proteínas apresentaram grande similaridade com as TSs de tripomastigotas. No entanto, a porção 3'UTR desses genes diferem entre TSs de epimastigotas e TSs de tripomastigotas, como também estão relacionadas na modulação da expressão. Em *T. cruzi*, a região 3'UTR de TSs de tripomastigotas e epimastigotas promoveram a redução da expressão de um gene repórter em epimastigotas, apesar da porção 3'UTR de epimastigota ter sido mais eficiente que a de tripomastigotas. Ao passo que a porção

3'UTR de tripomastigota levou ao aumento da expressão do gene repórter nos estágios de mamíferos (JÄGER; MUIÁ; CAMPETELLA, 2008).

Outra família identificada é a gp85/TS, em que as glicoproteínas de superfície de 85 kDa estão inseridas. Essa subfamília pertence ao grupo II da superfamília trans-sialidase e está entre as proteínas que não possuem atividade catalítica trans-sialidase/sialidase. Curiosamente, essas proteínas são descritas em formas de *T. cruzi* em estágio de vertebrados, nas formas tripomastigotas e amastigotas intracelulares, estando relacionadas a adesão da superfície celular do hospedeiro e invasão celular (ALVES; COLLI, 2007). Em *T. rangeli*, foi identificada a transcrição de genes gp85 por epimastigotas por meio de ensaios de Northern blot (AÑEZ-ROJAS *et al.*, 2005). Além disso, foram confirmadas a tradução e a localização dessas proteínas na superfície celular em epimastigotas de *T. rangeli* obtidas da hemolinfa de triatomíneos (PEÑA *et al.*, 2009). Acredita-se que apesar do *T. rangeli* não ser patogênico para os hospedeiros mamíferos, a regulação positiva da transcrição desses genes pode ser fundamental para a capacidade do parasito invadir e colonizar diferentes tecidos, podendo ter influência nos efeitos patológicos observados no vetor. Em *T. cruzi*, a ausência em epimastigota pode ser justificada pela colonização apenas do trato digestivo do vetor, sem nenhuma invasão celular.

Assim como as cinesinas encontradas reguladas positivamente em tripomastigotas, em epimastigotas também encontramos cinesinas com um aumento de expressão gênica. Como discutido anteriormente, essas proteínas têm papel importante no transporte de proteínas e outras estruturas. Outros transcritos relacionados ao transporte são os identificados como proteínas semelhantes à nodulina. Essas proteínas foram descritas primeiramente em plantas e estão relacionadas ao transporte de nutrientes, solutos e aminoácidos (DENANCÉ; SZUREK; NOËL, 2014; LEGOCKI; VERMA, 1980). A proteína semelhante a nodulina já foi descrita sendo encontrada no proteoma de epimastigotas de *T. cruzi* e assim como neste trabalho, sua localização foi predita como presente na membrana plasmática (FERELLA *et al.*, 2008).

É na hemolinfa que acontecem as principais respostas imunológicas do vetor, dessa forma, observamos transcritos que estão relacionados às vias de proteção, como é o caso da proteína reguladora do complemento (CRP). Em *T. cruzi*, essa proteína confere proteção à morte do parasito mediada pelo complemento. A CRP se liga nos componentes C3b e C4d do complemento humano e inibe a ativação da

cascata do complemento (BEUCHER *et al.*, 2003). No entanto, epimastigotas de *T. cruzi* não possuem níveis detectáveis de CRP (NORRIS, 1998). Porém, o ciclo de *T. cruzi* no vetor é exclusivo ao trato digestório, enquanto o *T. rangeli* coloniza a hemolinfa e se replica intensamente nesse tecido, além de estar exposto nesse tecido a uma intensa resposta imune formada pelos hemócitos do hospedeiro invertebrado (GUARNERI; LORENZO, 2017). Dessa forma, devido às diferenças do ciclo de vida no hospedeiro invertebrado, o aumento da expressão de CRP pode justificar a sobrevivência do parasito a resposta imune do inseto.

Outra enzima importante é a transcetolase, presente no ramo não oxidativo da via das pentoses fosfato. Essa via está envolvida na biossíntese de nucleotídeos, NADPH e na defesa contra o estresse oxidativo (WOOD, 1986). Em particular, para *Leishmania* sp. e *T. cruzi*, quando as células foram expostas ao estresse oxidativo, as enzimas dessa via tiveram aumento de expressão gênica, conferindo defesa contra o estresse oxidativo (MAUGERI *et al.*, 2003; MAUGERI; CAZZULO, 2004). Portanto, a regulação positiva do gene da transcetolase encontrada em epimastigotas pode estar envolvida na proteção do parasito contra os estresses oxidativos que ele enfrenta na hemolinfa do vetor.

Uma outra estratégia importante para regular a homeostase redox é a ação da tripanotiona (FAIRLAMB *et al.*, 1985). No entanto, esse tiol é formado por outras moléculas como a cisteína, sendo o componente que confere a capacidade de doar e receber elétrons das espécies reativas de oxigênio (NOZAKI; ALI; TOKORO, 2005). É na via de biossíntese de transulfuração reversa (RTS) que a enzima cistationina- γ -liase realiza a conversão do precursor em cisteína. Além disso, a cisteína está envolvida em diversos processos celulares, como: estrutura, estabilidade e regulação da atividade catalítica de inúmeras proteínas. Dessa forma, o aumento da transcrição gênica da cistationina- γ -liase em epimastigotas pode ser uma estratégia para fornecer cisteína em todas as vias necessárias (NOZAKI *et al.*, 2001).

5 CONCLUSÕES

A análise de RNA-Seq de *T. rangeli* permitiu a obtenção e identificação de transcritos e seus níveis de expressão auxiliando o entendimento de alguns eventos que estão ocorrendo durante a infecção da glândula salivar e hemolinfa de *R. prolixus*. Os parasitos apesar de terem transcrição constitutiva e policistrônica, apresentam regulação da expressão gênica de uma maneira que é possível identificar expressão diferencial de alguns genes. Apesar de analisar uma das seis janelas de leitura, foi possível a identificação de transcritos pertencentes a famílias multigênicas diferencialmente expressas. Destacando as famílias Trans-sialidase e GP63 presentes tanto nas formas tripomastigotas e epimastigotas, além de outras proteínas como gp85 também presentes em tripanossomatídeos patogênicos como *T. cruzi*. A transcrição e expressão diferencial destes genes por *T. rangeli* reforça a importância destes para a sobrevivência e desenvolvimento do parasito no hospedeiro invertebrado.

Ainda, como perspectiva, a análise complementar do transcriptoma por meio da tradução dos transcritos nas seis janelas de leitura possibilitará a redução de sequências truncadas e a identificação de um número maior de transcritos diferencialmente expressos. Ademais, a identificação de genes que estão sendo transcritos pelo *R. prolixus* e seus níveis de expressão gênica podem auxiliar na compreensão dos eventos e interações que ocorrem durante a infecção da hemolinfa e glândula salivar por *T. rangeli*.

REFERÊNCIAS

ABID, Hela et al. Leishmania infantum 5'-Methylthioadenosine Phosphorylase presents relevant structural divergence to constitute a potential drug target. *BMC Structural Biology*, v. 17, n. 1, p. 9, 2017.

AFCHAIN, D. et al. Antigenic Make-Up of *Trypanosoma cruzi* Culture Forms: Identification of a Specific Component. *The Journal of Parasitology*, v. 65, n. 4, p. 507, 1979.

ALVES, M.J.M.; COLLI, W. *Trypanosoma cruzi*: Adhesion to the host cell and intracellular survival. *IUBMB Life*, v. 59, n. 4–5, p. 274–279, 2007.

AMAYA, M. F. et al. The High Resolution Structures of Free and Inhibitor-bound *Trypanosoma rangeli* Sialidase and its Comparison with *T.cruzi* Trans-sialidase. *Journal of Molecular Biology*, v. 325, n. 4, p. 773–784, jan. 2003.

AMOS, Beatrice et al. VEuPathDB: the eukaryotic pathogen, vector and host bioinformatics resource center. *Nucleic Acids Research*, v. 50, n. D1, p. D898–D911, 2022.

AN, Tai; LI, Ziyin. An orphan kinesin controls trypanosome morphology transitions by targeting FLAM3 to the flagellum. *PLOS Pathogens*, v. 14, n. 5, p. e1007101, 2018.

ANDREWS, S. **FastQC: a quality control tool for high throughput sequence data.** Disponível em: <<https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>>.

AÑEZ, N. Studies on *Trypanosoma rangeli* Tejera 1920: VI. Developmental pattern in the haemolymph of *Rhodnius prolixus*. *Memórias do Instituto Oswaldo Cruz*, v. 78, n. 4, p. 413–419, 1983.

AÑEZ, N. Studies on *Trypanosoma rangeli* Tejera, 1920. VII--Its effect on the survival of infected triatomine bugs. *Memórias do Instituto Oswaldo Cruz*, v. 79, n. 2, p. 249–255, 1984.

AÑEZ-ROJAS, Néstor et al. *Trypanosoma rangeli* expresses a gene of the group II trans-sialidase superfamily. *Molecular and Biochemical Parasitology*, v. 142, n. 1, p. 133–136, 2005.

BAIROCH, A. The SWISS-PROT protein sequence data bank and its new supplement TREMBL. *Nucleic Acids Research*, v. 24, n. 1, p. 21–25, 1996.

BARTHOLOMEU, Daniella C. *et al.* Genomic organization and expression profile of the mucin-associated surface protein (masp) family of the human pathogen *Trypanosoma cruzi*. **Nucleic Acids Research**, v. 37, n. 10, p. 3407–3417, 2009.

BASSERI, H.R.; TEW, I.F.; RATCLIFFE, N.A. Identification and distribution of carbohydrate moieties on the salivary glands of *Rhodnius prolixus* and their possible involvement in attachment/invasion by *Trypanosoma rangeli*. **Experimental Parasitology**, v. 100, n. 4, p. 226–234, 2002.

BERNARDO, Werica P. *et al.* Genomic Organization and Generation of Genetic Variability in the RHS (Retrotransposon Hot Spot) Protein Multigene Family in *Trypanosoma cruzi*. **Genes**, v. 11, n. 9, p. 1085, 2020.

BEUCHER, Margaret *et al.* Expression and purification of functional, recombinant *Trypanosoma cruzi* complement regulatory protein. **Protein Expression and Purification**, v. 27, n. 1, p. 19–26, 2003.

BEUCHER, M.; NORRIS, K. A. Sequence Diversity of the *Trypanosoma cruzi* Complement Regulatory Protein Family. **Infection and Immunity**, v. 76, n. 2, p. 750–758, 2008.

BLACK, Jennifer A. *et al.* Life in plastic, it's fantastic! How Leishmania exploit genome instability to shape gene expression. **Frontiers in Cellular and Infection Microbiology**, v. 13, 2023.

BOLGER, Anthony M.; LOHSE, Marc; USADEL, Bjoern. Trimmomatic: a flexible trimmer for Illumina sequence data. **Bioinformatics**, v. 30, n. 15, p. 2114–2120, 2014.

BORELLA MARFIL ANHÊ, Ana Carolina *et al.* Microanatomical and secretory characterization of the salivary gland of the *Rhodnius prolixus* (Hemiptera, Reduviidae, Triatominae), a main vector of Chagas disease. **Open Biology**, v. 11, n. 6, 2021.

BRADWELL, Katie R *et al.* Genomic comparison of *Trypanosoma conorhini* and *Trypanosoma rangeli* to *Trypanosoma cruzi* strains of high and low virulence. **BMC Genomics**, v. 19, n. 1, p. 770, 2018.

BRAS, Jose; GUERREIRO, Rita; HARDY, John. Use of next-generation sequencing and other whole-genome strategies to dissect neurological disease. **Nature Reviews Neuroscience**, v. 13, n. 7, p. 453–464, 2012.

BRECHER, G.; WIGGLESWORTH, V.B. The transmission of actinomyces rhodnii Erikson in *Rhodnius prolixus* stål (hemiptera) and its influence on the growth of the host. **Parasitology**, v. 35, n. 4, p. 220–224, 1944.

BRINGAUD, Frédéric *et al.* A New, Expressed Multigene Family Containing a Hot Spot for Insertion of Retroelements Is Associated with Polymorphic Subtelomeric Regions of *Trypanosoma brucei*. **Eukaryotic Cell**, v. 1, n. 1, p. 137–151, 2002.

BUSCAGLIA, Carlos A. *et al.* *Trypanosoma cruzi* surface mucins: host-dependent coat diversity. **Nature Reviews Microbiology**, v. 4, n. 3, p. 229–236, 2006.

BUSCHIAZZO, Alejandro *et al.* Sequence of a *Trypanosoma rangeli* gene closely related to *Trypanosoma cruzi* trans-sialidase. **Molecular and Biochemical Parasitology**, v. 62, n. 1, p. 115–116, 1993.

CALLEJAS-HERNÁNDEZ, Francisco *et al.* Analysis of mRNA processing at whole transcriptome level, transcriptomic profile and genome sequence refinement of *Trypanosoma cruzi*. **Scientific Reports**, v. 9, n. 1, p. 17376, 2019.

CALLEJAS-HERNÁNDEZ, Francisco *et al.* Genomic assemblies of newly sequenced *Trypanosoma cruzi* strains reveal new genomic expansion and greater complexity. **Scientific Reports**, v. 8, n. 1, p. 14631, 2018.

CAMACHO, Christiam *et al.* BLAST+: architecture and applications. **BMC Bioinformatics**, v. 10, n. 1, p. 421, 2009.

CAYLA, Mathieu *et al.* Distinguishing functions of trypanosomatid protein kinases. **Trends in Parasitology**, v. 38, n. 11, p. 950–961, 2022.

CLAYTON, C. E. Life without transcriptional control? From fly to man and back again. **The EMBO Journal**, v. 21, n. 8, p. 1881–1888, 2002.

CORDON-OBAS, Carlos *et al.* Identification of sequence-specific promoters driving polycistronic transcription initiation by RNA polymerase II in trypanosomes. **Cell Reports**, v. 38, n. 2, p. 110221, 2022.

COSENTINO-GOMES, Daniela; ROCCO-MACHADO, Nathália; MEYER-FERNANDES, José Roberto. *Rhodnius prolixus*: Modulation of antioxidant defenses by *Trypanosoma rangeli*. **Experimental Parasitology**, v. 145, p. 118–124, 2014.

CUBA CUBA, César A. Revisión de los aspectos biológicos y diagnósticos del *Trypanosoma* (Herpetosoma) *rangeli*. **Revista da Sociedade Brasileira de Medicina Tropical**, v. 31, n. 2, p. 207–220, 1998.

CUBA CUBA, C.A. Study of a Peruvian strain of *Trypanosoma rangeli*. II. Observations on the experimental infection of *Rhodnius ecuadoriensis* | ESTUDO DE UMA CEPA PERUANA DE *TRYPANOSOMA RANGELI*. II. OBSERVACOES SOBRE A INFECCAO EXPERIMENTAL DE *RHODNIUS ECUADORIENSIS*. **Revista do Instituto de Medicina Tropical de Sao Paulo**, v. 16, n. 1, p. 19–27, 1974.

CUERVO, Claudia *et al.* Sequence polymorphism in the *Trypanosoma rangeli* HSP70 coding genes allows typing of the parasite KP1(+) and KP1(-) groups. **Experimental Parasitology**, v. 133, n. 4, p. 447–453, 2013.

CUEVAS, Ileana C.; CAZZULO, Juan J.; SÁNCHEZ, Daniel O. gp63 Homologues in *Trypanosoma cruzi*: Surface Antigens with Metalloprotease Activity and a Possible Role in Host Cell Infection. **Infection and Immunity**, v. 71, n. 10, p. 5739–5749, 2003.

D’ALESSANDRO, A. Biology of *Trypanosoma* (Herpetosoma) *rangeli* Tejera 1920. **Biology of the kinetoplastida**, v. 1, p. 328–393, 1976.

D’ALESSANDRO, A; SARAIVA, N.G. *Trypanosoma rangeli* In JP Kreier. **Parasitic Protozoa**, v. 2, p. 1–54, 1992.

D’AVILA-LEVY, Claudia M. *et al.* GP63 Function in the Interaction of Trypanosomatids with the Invertebrate Host: Facts and Prospects. 2014. p. 253–270.

DE GAUDENZI, Javier G. *et al.* Gene expression regulation in trypanosomatids. **Essays in Biochemistry**, v. 51, p. 31–46, 2011.

DE MORAES, Milene H *et al.* Different serological cross-reactivity of *Trypanosoma rangeli* forms in *Trypanosoma cruzi*-infected patients sera. **Parasites & Vectors**, v. 1, n. 1, p. 20, 2008.

DENANCÉ, Nicolas; SZUREK, Boris; NOËL, Laurent D. Emerging Functions of Nodulin-Like Proteins in Non-Nodulating Plant Species. **Plant and Cell Physiology**, v. 55, n. 3, p. 469–474, 2014.

DILLON, Laura A. L. *et al.* Transcriptomic profiling of gene expression and RNA processing during *Leishmania major* differentiation. **Nucleic Acids Research**, v. 43, n. 14, p. 6799–6813, 2015.

DOS SANTOS, Sara Lopes *et al.* The MASP Family of *Trypanosoma cruzi*: Changes in Gene Expression and Antigenic Profile during the Acute Phase of Experimental Infection. **PLoS Neglected Tropical Diseases**, v. 6, n. 8, p. e1779, 2012.

DOS-SANTOS, André L.A. *et al.* Interaction between *Trypanosoma rangeli* and the *Rhodnius prolixus* salivary gland depends on the phosphotyrosine ecto-phosphatase activity of the parasite. **International Journal for Parasitology**, v. 42, n. 9, p. 819–827, 2012. Disponível em: Acesso em: 22 set. 2023.

EL KOUNI, Mahmoud H. Potential chemotherapeutic targets in the purine metabolism of parasites. **Pharmacology & Therapeutics**, v. 99, n. 3, p. 283–309, 2003.

ELLIS, D.S.; EVANS, D.A.; STAMFORD, S. The penetration of the salivary glands of *Rhodnius prolixus* by *Trypanosoma rangeli*. **Zeitschrift für Parasitenkunde Parasitology Research**, v. 62, n. 1, p. 63–74, 1980.

EL-SAYED, Najib M. *et al.* Comparative Genomics of Trypanosomatid Parasitic Protozoa. **Science**, v. 309, n. 5733, p. 404–409, 2005a.

EL-SAYED, Najib M. *et al.* The Genome Sequence of *Trypanosoma cruzi*, Etiologic Agent of Chagas Disease. **Science**, v. 309, n. 5733, p. 409–415, 2005b.

EWELS, Philip *et al.* MultiQC: summarize analysis results for multiple tools and samples in a single report. **Bioinformatics**, v. 32, n. 19, p. 3047–3048, 2016.

FAIRLAMB, Alan H. *et al.* Trypanothione: A Novel Bis(glutathionyl)spermidine Cofactor for Glutathione Reductase in Trypanosomatids. **Science**, v. 227, n. 4693, p. 1485–1487, 1985.

FELLET, M.R. *et al.* Effects of infection by *Trypanosoma cruzi* and *Trypanosoma rangeli* on the reproductive performance of the vector *Rhodnius prolixus*. **PLoS ONE**, v. 9, n. 8, 2014.

FERELLA, Marcela *et al.* Proteomics in *Trypanosoma cruzi* - localization of novel proteins to various organelles. **PROTEOMICS**, v. 8, n. 13, p. 2735–2749, 2008.

FERREIRA, L.L. *et al.* A standardizable protocol for infection of *Rhodnius prolixus* with *Trypanosoma rangeli*, which mimics natural infections and reveals physiological effects of infection upon the insect. **Journal of Invertebrate Pathology**, v. 105, n. 1, p. 91–97, 2010.

FERREIRA, L. L. **Explorando o ciclo de transmissão do *Trypanosoma rangeli*: avaliação do papel das aves na circulação do parasito em triatomíneos, e dos eventos que ocorrem na fase inicial da infecção em camundongos.** 2017. - Centro de Pesquisas René Rachou, Fundação Oswaldo Cruz, Belo Horizonte, 2017.

FERREIRA, Luciana de Lima *et al.* New features on the survival of human-infective *Trypanosoma rangeli* in a murine model: Parasite accumulation is observed in lymphoid organs. **PLOS Neglected Tropical Diseases**, v. 14, n. 12, p. e0009015, 2020.

FREITAS, Leandro M. *et al.* Genomic Analyses, Gene Expression and Antigenic Profile of the Trans-Sialidase Superfamily of *Trypanosoma cruzi* Reveal an Undetected Level of Complexity. **PLoS ONE**, v. 6, n. 10, p. e25914, 2011.

FREVERT, U; SCHENKMAN, S; NUSSENZWEIG, V. Stage-specific expression and intracellular shedding of the cell surface trans-sialidase of *Trypanosoma cruzi*. **Infection and Immunity**, v. 60, n. 6, p. 2349–2360, 1992.

GARCIA, Eloi S. *et al.* Interactions between intestinal compounds of triatomines and *Trypanosoma cruzi*. **Trends in Parasitology**, v. 26, n. 10, p. 499–505, 2010.

GERASIMOV, Evgeny S. *et al.* Common Structural Patterns in the Maxicircle Divergent Region of Trypanosomatidae. **Pathogens**, v. 9, n. 2, p. 100, 2020.

GILBERT, Susan P.; GUZIK-LENDRUM, Stephanie; RAYMENT, Ivan. Kinesin-2 motors: Kinetics and biophysics. **Journal of Biological Chemistry**, v. 293, n. 12, p. 4510–4518, 2018.

GOMES, S.A.O *et al.* Suppression of the prophenoloxidase system in *Rhodnius prolixus* orally infected with *Trypanosoma rangeli*. **Journal of Insect Physiology**, v. 49, n. 9, p. 829–837, 2003.

GRISARD, Edmundo C *et al.* Characterization of *Trypanosoma rangeli* Strains Isolated in Central and South America: an Overview. **Memórias do Instituto Oswaldo Cruz**, v. 94, n. 2, p. 203–209, 1999.

GRISARD, Edmundo C. *et al.* Transcriptomic analyses of the avirulent protozoan parasite *Trypanosoma rangeli*. **Molecular and Biochemical Parasitology**, v. 174, n. 1, p. 18–25, 2010.

GUARNERI, Alessandra Aparecida; LORENZO, Marcelo Gustavo. Triatomine physiology in the context of trypanosome infection. **Journal of Insect Physiology**, v. 97, p. 66–76, 2017.

GUHL, Felipe; VALLEJO, Gustavo Adolfo. *Trypanosoma* (Herpetosoma) *rangeli* Tejera, 1920: an updated review. **Memórias do Instituto Oswaldo Cruz**, v. 98, n. 4, p. 435–442, 2003.

HECKER, H.; SCHWARZENBACH, M.; RUDIN, W. Development and interactions of *Trypanosoma rangeli* in and with the reduviid bug *Rhodnius prolixus*. **Parasitology Research**, v. 76, n. 4, p. 311–318, 1990.

HERREROS-CABELLO, Alfonso *et al.* *Trypanosoma cruzi* Genome: Organization, Multi-Gene Families, Transcription, and Biological Implications. **Genes**, v. 11, n. 10, p. 1196, 2020.

JÄGER, Adriana V.; MUIÁ, Romina P.; CAMPETELLA, Oscar. Stage-specific expression of *Trypanosoma cruzi* trans-sialidase involves highly conserved 3' untranslated regions. **FEMS Microbiology Letters**, v. 283, n. 2, p. 182–188, 2008.

JING, Tian-Zhong; QI, Feng-Hui; WANG, Zhi-Ying. Most dominant roles of insect gut bacteria: digestion, detoxification, or essential nutrient provision?. **Microbiome**, v. 8, n. 1, p. 38, 2020.

JONES, Philip *et al.* InterProScan 5: genome-scale protein function classification. **Bioinformatics**, v. 30, n. 9, p. 1236–1240, 2014.

KAUFER, Alexa *et al.* The evolution of trypanosomatid taxonomy. **Parasites & Vectors**, v. 10, n. 1, p. 287, 2017.

KAWASHITA, Silvia Y. *et al.* Homology, paralogy and function of DGF-1, a highly dispersed *Trypanosoma cruzi* specific gene family and its implications for information entropy of its encoded proteins. **Molecular and Biochemical Parasitology**, v. 165, n. 1, p. 19–31, 2009. Disponível em: Acesso em: 1 out. 2023.

KILPINEN, H.; BARRETT, J.C. How next-generation sequencing is transforming complex disease genetics. **Trends in Genetics**, v. 29, n. 1, p. 23–30, 2013.

KIM, Daehwan *et al.* Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. **Nature Biotechnology**, v. 37, n. 8, p. 907–915, 2019.

KINGSMORE, S.F.; SAUNDERS, C.J. Deep sequencing of patient genomes for disease diagnosis: When will it become routine?. **Science Translational Medicine**, v. 3, n. 87, 2011.

LAI, William K. M.; PUGH, B. Franklin. Understanding nucleosome dynamics and their links to gene expression and DNA replication. **Nature Reviews Molecular Cell Biology**, v. 18, n. 9, p. 548–562, 2017.

LANDER, Noelia *et al.* Localization and Developmental Regulation of a Dispersed Gene Family 1 Protein in *Trypanosoma cruzi*. **Infection and Immunity**, v. 78, n. 1, p. 231–240, 2010.

LAVOIEPIERRE, M M; DICKERSON, G; GORDON, R M. Studies on the methods of feeding of blood-sucking arthropods. I. The manner in which triatomine bugs obtain their blood-meal, as observed in the tissues of the living rodent, with some remarks on the effects of the bite on human volunteers. **Annals of tropical medicine and parasitology**, v. 53, p. 235–50, 1959.

LEGOCKI, Roman P.; VERMA, Desh Pal S. Identification of “nodule-specific” host proteins (nodulins) involved in the development of Rhizobium-Legume symbiosis. **Cell**, v. 20, n. 1, p. 153–163, 1980.

LEULIER, François; ROYET, Julien. Maintaining immune homeostasis in fly gut. **Nature Immunology**, v. 10, n. 9, p. 936–938, 2009.

LI, Heng *et al.* The Sequence Alignment/Map format and SAMtools. **Bioinformatics**, v. 25, n. 16, p. 2078–2079, 2009.

LIPKIN, W.I. The changing face of pathogen discovery and surveillance. **Nature Reviews Microbiology**, v. 11, n. 2, p. 133–141, 2013.

LOBO, Maite *et al.* Molecular and antigenic characterization of *Trypanosoma cruzi* ToIT proteins. **PLOS Neglected Tropical Diseases**, v. 13, n. 3, p. e0007245, 2019.

LOVE, Michael I; HUBER, Wolfgang; ANDERS, Simon. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. **Genome Biology**, v. 15, n. 12, p. 550, 2014.

LUKEŠ, Julius *et al.* Kinetoplast DNA Network: Evolution of an Improbable Structure. **Eukaryotic Cell**, v. 1, n. 4, p. 495–502, 2002.

MACHADO, Mara Olimpia. **Análise do secretoma/excretoma da forma tripomastigota de *Trypanosoma cruzi***. 2013. 1–110 f. - Universidade de Brasília, Brasília, 2013.

MAIA, Guilherme Augusto *et al.* AnnotaPipeline: An integrated tool to annotate eukaryotic proteins using multi-omics data. **Frontiers in Genetics**, v. 13, 2022.

MAIA DA SILVA, F. *et al.* Comparative phylogeography of *Trypanosoma rangeli* and Rhodnius (Hemiptera: Reduviidae) supports a long coexistence of parasite lineages and their sympatric vectors. **Molecular Ecology**, v. 16, n. 16, p. 3361–3373, 2007.

MANNI, Mosè *et al.* BUSCO: Assessing Genomic Data Quality and Beyond. **Current Protocols**, v. 1, n. 12, 2021.

MARQUEZ, Daniela De Stefani *et al.* Susceptibility of Different Triatomine Species to *Trypanosoma rangeli* Experimental Infection. **Vector-Borne and Zoonotic Diseases**, v. 6, n. 1, p. 50–56, 2006.

MATTOS, Eliciane C. *et al.* The Gp85 Surface Glycoproteins from *Trypanosoma cruzi*. 2014. p. 151–180.

MAUGERI, D.A. *et al.* Pentose phosphate metabolism in *Leishmania mexicana*. **Molecular and Biochemical Parasitology**, v. 130, n. 2, p. 117–125, 2003.

MAUGERI, D.A.; CAZZULO, J.J. The pentose phosphate pathway in *Trypanosoma cruzi*. **FEMS Microbiology Letters**, v. 234, n. 1, p. 117–123, 2004.

MAXAM, A M; GILBERT, W. A new method for sequencing DNA. **Proceedings of the National Academy of Sciences**, v. 74, n. 2, p. 560–564, 1977.

MEIRELLES, R.M.S. *et al.* Penetration of the salivary glands of *Rhodnius domesticus* Neiva & Pinto, 1923 (Hemiptera: Reduviidae) by *Trypanosoma rangeli* Tejera, 1920 (Protozoa: Kinetoplastida). **Parasitology Research**, v. 97, n. 4, p. 259–269, 2005.

MEIRELLES, R.M.S. *et al.* Ultrastructure of the salivary glands of *Rhodnius domesticus* Neiva & Pinto, 1923 (Hemiptera: Reduviidae). **Journal of Submicroscopic Cytology and Pathology**, v. 35, n. 2, p. 199–207, 2003.

MICHELS, Paul A.M. *et al.* Carbohydrate metabolism in trypanosomatids: New insights revealing novel complexity, diversity and species-unique features. **Experimental Parasitology**, v. 224, p. 108102, 2021.

MINNING, Todd A *et al.* The steady-state transcriptome of the four major life-cycle stages of *Trypanosoma cruzi*. **BMC Genomics**, v. 10, n. 1, p. 370, 2009.

MITCHELL, Stephanie L.; SIMNER, Patricia J. Next-Generation Sequencing in Clinical Microbiology. **Clinics in Laboratory Medicine**, v. 39, n. 3, p. 405–418, 2019.

MORAND, Sabine *et al.* MAP kinase kinase 1 (MKK1) is essential for transmission of *Trypanosoma brucei* by *Glossina morsitans*. **Molecular and Biochemical Parasitology**, v. 186, n. 1, p. 73–76, 2012.

MORGAN, Jason K.; ORTIZ, Jose A.; RIORDAN, James T. The role for TolA in enterohemorrhagic *Escherichia coli* pathogenesis and virulence gene transcription. **Microbial Pathogenesis**, v. 77, p. 42–52, 2014.

NAGALAKSHMI, Ugrappa *et al.* The Transcriptional Landscape of the Yeast Genome Defined by RNA Sequencing. **Science**, v. 320, n. 5881, p. 1344–1349, 2008.

NARDY, Ana F. F. R. *et al.* Role of *Trypanosoma cruzi* Trans-sialidase on the Escape from Host Immune Surveillance. **Frontiers in Microbiology**, v. 7, 2016.

NORRIS, K.A. Stable transfection of *Trypanosoma cruzi* epimastigotes with the trypomastigote-specific complement regulatory protein cDNA confers complement resistance. **Infection and Immunity**, v. 66, n. 6, p. 2460–2465, 1998.

NOZAKI, Tomoyoshi *et al.* Characterization of Transsulfuration and Cysteine Biosynthetic Pathways in the Protozoan Hemoflagellate, *Trypanosoma cruzi*. **Journal of Biological Chemistry**, v. 276, n. 9, p. 6516–6523, 2001.

NOZAKI, Tomoyoshi; ALI, Vahab; TOKORO, Masaharu. Sulfur-Containing Amino Acid Metabolism in Parasitic Protozoa. 2005. p. 1–99.

PAIM, R.M.M. *et al.* The interaction between *Trypanosoma rangeli* and the nitrophorins in the salivary glands of the triatomine *Rhodnius prolixus* (Hemiptera; Reduviidae). **Insect Biochemistry and Molecular Biology**, v. 43, n. 3, p. 229–236, 2013.

PALENCHAR, Jennifer B.; BELLOFATTO, Vivian. Gene transcription in trypanosomes. **Molecular and Biochemical Parasitology**, v. 146, n. 2, p. 135–141, 2006.

PECH-CANUL, Ángel de la Cruz; MONTEÓN, Victor; SOLÍS-OVIEDO, Rosa-Lidia. A Brief View of the Surface Membrane Proteins from *Trypanosoma cruzi*. **Journal of Parasitology Research**, v. 2017, p. 1–13, 2017.

PEÑA, C.P. *et al.* Molecular analysis of surface glycoprotein multigene family TrGP expressed on the plasma membrane of *Trypanosoma rangeli* epimastigotes forms. **Acta Tropica**, v. 111, n. 3, p. 255–262, 2009.

PEREIRA, M E *et al.* Lectin receptors as markers for *Trypanosoma cruzi*. Developmental stages and a study of the interaction of wheat germ agglutinin with sialic acid residues on epimastigote cells. **The Journal of experimental medicine**, v. 152, n. 5, p. 1375–1392, 1980.

PEREIRA, Miercio E.A.; MOSS, Deborah. Neuraminidase activity in *Trypanosoma rangeli*. **Molecular and Biochemical Parasitology**, v. 15, n. 1, p. 95–103, 1985.

PERTEA, Mihaela *et al.* StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. **Nature Biotechnology** 2015 **33:3**, v. 33, n. 3, p. 290–295, 2015. Disponível em: <https://www.nature.com/articles/nbt.3122>. Acesso em: 19 abr. 2023.

POLAND, J.A. *et al.* Development of high-density genetic maps for barley and wheat using a novel two-enzyme genotyping-by-sequencing approach. **PLoS ONE**, v. 7, n. 2, 2012.

PONTES-DE-CARVALHO, Lain C.; TOMLINSON, Stephen; NUSSENZWEIG, Victor. *Trypanosoma rangeli* sialidase lacks trans-sialidase activity. **Molecular and Biochemical Parasitology**, v. 62, n. 1, p. 19–25, 1993.

PULIDO, Ximena Carolina; PÉREZ, Gerardo; VALLEJO, Gustavo Adolfo. Preliminary characterization of a *Rhodnius prolixus* hemolymph trypanolytic protein, this being a determinant of *Trypanosoma rangeli* KP1(+) and KP1(-) subpopulations' vectorial ability. **Memórias do Instituto Oswaldo Cruz**, v. 103, n. 2, p. 172–179, 2008.

QUANQUIN, Natalie M. *et al.* Immunization of Mice with a TolA-Like Surface Protein of *Trypanosoma cruzi* Generates CD4 + T-Cell-Dependent Parasiticidal Activity. **Infection and Immunity**, v. 67, n. 9, p. 4603–4612, 1999.

RAGHAVAN, Venket *et al.* A simple guide to de novo transcriptome assembly and annotation. **Briefings in Bioinformatics**, v. 2022, n. 2, p. 1–30, Disponível em: <https://doi.org/10.1093/bib/bbab563>. Acesso em: 17 abr. 2023.

RASTROJO, Alberto *et al.* The transcriptome of *Leishmania major* in the axenic promastigote stage: transcript annotation and relative expression levels by RNA-seq. **BMC Genomics**, v. 14, n. 1, p. 223, 2013.

RATCLIFFE, N.A. *et al.* *Trypanosoma cruzi* and Erythrocyte Agglutinins: A Comparative Study of Occurrence and Properties in the Gut and Hemolymph of *Rhodnius prolixus*. **Experimental Parasitology**, v. 83, n. 1, p. 83–93, 1996.

REBELLO, Karina M. *et al.* Participation of *Trypanosoma cruzi* gp63 molecules on the interaction with *Rhodnius prolixus*. **Parasitology**, v. 146, n. 8, p. 1075–1082, 2019.

REIS, Marcos M; MEIRELLES, Rosane M.S; SOARES, Maurilio J. Fine structure of the salivary glands of *Triatoma infestans* (Hemiptera: Reduviidae). **Tissue and Cell**, v. 35, n. 5, p. 393–400, 2003.

REIS-CUNHA, João Luís *et al.* Chromosomal copy number variation reveals differential levels of genomic plasticity in distinct *Trypanosoma cruzi* strains. **BMC Genomics**, v. 16, n. 1, p. 499, 2015.

RIBEIRO, J.M.C. *et al.* Exploring the sialome of the blood-sucking bug *Rhodnius prolixus*. **Insect Biochemistry and Molecular Biology**, v. 34, n. 1, p. 61–79, 2004.

RIBEIRO, J M C. Role of Saliva in Blood-Feeding by Arthropods. **Annual Review of Entomology**, v. 32, n. 1, p. 463–478, 1987.

RIBEIRO, José M. C. *et al.* Role of Salivary Antihemostatic Components in Blood Feeding by Triatomine Bugs (Heteroptera). **Journal of Medical Entomology**, v. 35, n. 4, p. 599–610, 1998.

RICE, Peter; LONGDEN, Ian; BLEASBY, Alan. EMBOSS: The European Molecular Biology Open Software Suite. **Trends in Genetics**, v. 16, n. 6, p. 276–277, 2000.

ROGERS, Matthew B. *et al.* Chromosome and gene copy number variation allow major structural change between species and strains of *Leishmania*. **Genome Research**, v. 21, n. 12, p. 2129–2142, 2011.

ROSENBERG, Abraham; SCHENGRUND, Cara-Lynne. Sialidases. *In*: BIOLOGICAL ROLES OF SIALIC ACID. Boston, MA: Springer US, 1976. p. 295–359.

SABORIO, Jose L. *et al.* *Trypanosoma cruzi*: Identification of a surface antigen restricted to the flagellar region of the infective form of the parasite. **Experimental Parasitology**, v. 70, n. 4, p. 411–418, 1990.

SANGER, F.; NICKLEN, S.; COULSON, A.R. DNA sequencing with chain-terminating inhibitors. **Proceedings of the National Academy of Sciences of the United States of America**, v. 74, n. 12, p. 5463–5467, 1977.

SANTOS, André L.S.; BRANQUINHA, Marta H.; D'AVILA-LEVY, Claudia M. The ubiquitous gp63-like metalloprotease from lower trypanosomatids: in the search for a function. **Anais da Academia Brasileira de Ciências**, v. 78, n. 4, p. 687–714, 2006.

SCHENKMAN, Sergio *et al.* A novel cell surface trans-sialidase of *Trypanosoma cruzi* generates a stage-specific epitope required for invasion of mammalian cells. **Cell**, v. 65, n. 7, p. 1117–1125, 1991.

SCHENKMAN, Sergio *et al.* Mucin-like glycoproteins linked to the membrane by glycosylphosphatidylinositol anchor are the major acceptors of sialic acid in a reaction catalyzed by trans-sialidase in metacyclic forms of *Trypanosoma cruzi*. **Molecular and Biochemical Parasitology**, v. 59, n. 2, p. 293–303, 1993.

SCHOCH, Conrad L *et al.* NCBI Taxonomy: a comprehensive update on curation, resources and tools. **Database**, v. 2020, 2020.

SEEBECK, T.; SCHAUB, R.; JOHNER, A. cAMP signalling in the kinetoplastid protozoa. **Current Molecular Medicine**, v. 4, n. 6, p. 585–599, 2004.

SIEGEL, T. Nicolai *et al.* Gene expression in *Trypanosoma brucei*: lessons from high-throughput RNA sequencing. **Trends in Parasitology**, v. 27, n. 10, p. 434–441, 2011.

SILVA, Adriana Corrêa da. **Avaliação do tropismo tissular de *Trypanosoma rangeli* em mamíferos e caracterização das glicoproteínas mucinas-like nessa espécie.** 2021. 1–216 f. - Universidade Federal de Santa Catarina, Florianópolis, 2021.

SINCERO, Thaís Cristine Marques *et al.* *Trypanosoma rangeli* displays a clonal population structure, revealing a subdivision of KP1(–) strains and the ancestry of the Amazonian group. **International Journal for Parasitology**, v. 45, n. 4, p. 225–235, 2015.

SMITH-UNNA, Richard *et al.* TransRate: reference-free quality assessment of de novo transcriptome assemblies. **Genome Research**, v. 26, n. 8, p. 1134–1144, 2016.

SOARES, Adriana Coelho *et al.* Salivation pattern of *Rhodnius prolixus* (Reduviidae; Triatominae) in mouse skin. **Journal of Insect Physiology**, v. 52, n. 5, p. 468–472, 2006.

STEINDEL, M. **Caracterização de Cepas de *Trypanosoma cruzi* e *Trypanosoma rangeli* Isoladas de Reservatórios e Vetores Silvestres Naturalmente Infectados de Santa Catarina.** 1993. - Universidade Federal de Minas Gerais, Belo Horizonte, 1993.

STOCO, Patrícia Hermes *et al.* Genome of the Avirulent Human-Infective Trypanosome—*Trypanosoma rangeli*. **PLoS Neglected Tropical Diseases**, v. 8, n. 9, p. e3176, 2014.

STOCO, Patricia H. *et al.* Other Major Trypanosomiasis. *In*: ARTHROPOD BORNE DISEASES. Cham: Springer International Publishing, 2017. p. 299–324.

TEJERA, Enrique. Un nouveau flagellé de *Rhodnius prolixus*, Trypanosoma (ou Crithidia) rangeli. **Bull Soc Path Exot**, v. 13, p. 527–530, 1920.

THOMAS, Sean *et al.* Histone acetylations mark origins of polycistronic transcription in *Leishmania major*. **BMC Genomics**, v. 10, n. 1, p. 152, 2009.

TSCHOEKE, Diogo A. *et al.* The Comparative Genomics and Phylogenomics of *Leishmania Amazonensis* Parasite. **Evolutionary Bioinformatics**, v. 10, p. EBO.S13759, 2014.

VALLEJO, Gustavo A *et al.* kDNA markers define two major *Trypanosoma rangeli* lineages in Latin-America. **Acta Tropica**, v. 81, n. 1, p. 77–82, 2002.

VALLEJO, G.A. *et al.* Laboratory maintenance of Trypanosoma (Herpetosoma) rangeli Tejera, 1920 | Mantenimiento en el laboratorio de Trypanosoma (Herpetosoma) rangeli Tejera, 1920. **Revista de biología tropical**, v. 34, n. 1, p. 75–81, 1986.

VAN DIJK, Erwin L. *et al.* Ten years of next-generation sequencing technology. **Trends in Genetics**, v. 30, n. 9, p. 418–426, 2014.

VERA, J. CRISTOBAL *et al.* Rapid transcriptome characterization for a nonmodel organism using 454 pyrosequencing. **Molecular Ecology**, v. 17, n. 7, p. 1636–1647, 2008.

VIEIRA, Cecilia S *et al.* *Rhodnius prolixus* interaction with *Trypanosoma rangeli*: modulation of the immune system and microbiota population. **Parasites & Vectors**, v. 8, n. 1, p. 135, 2015.

WEATHERLY, D Brent; BOEHLKE, Courtney; TARLETON, Rick L. Chromosome level assembly of the hybrid *Trypanosoma cruzi* genome. **BMC Genomics**, v. 10, n. 1, p. 255, 2009.

WESTENBERGER, Scott J *et al.* *Trypanosoma cruzi* mitochondrial maxicircles display species- and strain-specific variation and a conserved element in the non-coding region. **BMC Genomics**, v. 7, n. 1, p. 60, 2006.

WHITTEN, M.M.A. *et al.* Role of Superoxide and Reactive Nitrogen Intermediates in *Rhodnius prolixus* (Reduviidae)/*Trypanosoma rangeli* Interactions. **Experimental Parasitology**, v. 98, n. 1, p. 44–57, 2001.

WHO. **Chagas disease (also known as American trypanosomiasis)**. Disponível em: [https://www.who.int/news-room/fact-sheets/detail/chagas-disease-\(american-trypanosomiasis\)](https://www.who.int/news-room/fact-sheets/detail/chagas-disease-(american-trypanosomiasis)). Acesso em: 17 set. 2023a.

WHO. **Leishmaniasis**. Disponível em: <https://www.who.int/news-room/fact-sheets/detail/leishmaniasis>. Acesso em: 17 set. 2023b.

WHO. **Trypanosomiasis, human African (sleeping sickness)**. Disponível em: [https://www.who.int/news-room/fact-sheets/detail/trypanosomiasis-human-african-\(sleeping-sickness\)](https://www.who.int/news-room/fact-sheets/detail/trypanosomiasis-human-african-(sleeping-sickness)). Acesso em: 17 set. 2023c.

WINCKER, Patrick; MURTO-DOVALES, Ana Cristina; GOLDENBERG, Samuel. Nucleotide sequence of a representative member of a *Trypanosoma cruzi* dispersed gene family. **Molecular and Biochemical Parasitology**, v. 55, n. 1–2, p. 217–220, 1992. Disponível em: Acesso em: 1 out. 2023.

WOOD, T. Physiological functions of the pentose phosphate pathway. **Cell biochemistry and function**, v. 4, n. 4, p. 241–247, 1986.

ZUÑIGA, C. *et al.* *Trypanosoma rangeli*: increase in virulence with inocula of different origins in the experimental infection in mice. **Parasitology Research**, v. 83, n. 8, p. 797–800, 1997.

APÊNDICE A – Termos de ontologia genética identificados no transcriptoma de *T. rangeli*.

A figura 10 exibe a frequência de todos os termos de ontologia genética com contagem maior ou igual a 1, identificados por meio da análise de anotação funcional. São ilustradas as frequências dos GOs em três categorias: componente celular – verde; função molecular – azul; e processo biológico – rosa. O tamanho dos pontos refere-se à frequência do termo.

Figura 10 – Frequência de termos de ontologia genética identificados no transcriptoma de *T. rangeli*



APÊNDICE B – Transcritos de *T. rangeli* apresentando aumento de expressão gênica quando comparado formas tripomastigotas contra epimastigotas.

Na Tabela 5 são listados todos os transcritos apresentando expressão diferencial com regulação positiva da transcrição quando comparado formas tripomastigotas contra epimastigotas de *T. rangeli*. São descritos os IDs, ou seja, o código de identidade de cada transcrito conferido pelo programa StringTie no momento da montagem do transcriptoma, os níveis de expressão diferencial obtidos pelo DESeq2, a anotação atribuída realizada por anotação por similaridade e a localização celular predita.

Tabela 4 – Totalidade de transcritos identificados exibindo aumento de expressão gênica quando comparado formas tripomastigotas vs. epimastigotas.

ID	Nível de expressão gênica (logFC)	Anotação	Localização predita
MSTRG.57	4.16	hypothetical protein	Núcleo
MSTRG.6998	3.30	surface protein ToIT	Extracelular
MSTRG.59	3.83	hypothetical protein	Extracelular
MSTRG.12	3.20	hypothetical protein	Mitocôndria
MSTRG.640	3.60	hypothetical protein	Mitocôndria
g7253	3.18	hypothetical protein	Mitocôndria
MSTRG.5312	3.62	hypothetical protein	Extracelular
MSTRG.6742	4.20	hypothetical protein	Núcleo
g8352	3.20	hypothetical protein	Citosol
MSTRG.8931	3.71	hypothetical protein	Mitocôndria
MSTRG.8917	3.60	hypothetical protein	Mitocôndria
MSTRG.2039	3.42	hypothetical protein	Núcleo
g5035	4.15	hypothetical protein	Membrana plasmática
MSTRG.9638	3.20	hypothetical protein	Mitocôndria
g7994	3.57	hypothetical protein	Inconclusivo
MSTRG.8928	3.01	hypothetical protein	Mitocôndria
MSTRG.2	3.27	hypothetical protein	Mitocôndria
g8083	3.42	hypothetical protein	Mitocôndria
MSTRG.5313	2.68	hypothetical protein	Núcleo
MSTRG.9633	2.95	hypothetical protein	Extracelular
MSTRG.8900	2.89	hypothetical protein	Núcleo
MSTRG.9595	2.65	hypothetical protein	Núcleo
MSTRG.665	2.68	trans-sialidase, Group I, putative	Mitocôndria
MSTRG.8370	2.60	hypothetical protein	Citosol
MSTRG.3156	2.26	hypothetical protein	Citosol
MSTRG.3577	2.84	hypothetical protein	Retículo endoplasmático

MSTRG.8926	2.63	trans-sialidase, Group I, putative	Peroxisomo
MSTRG.645	2.28	hypothetical protein	Núcleo
MSTRG.3986	2.35	hypothetical protein	Extracelular
MSTRG.8742	2.34	hypothetical protein	Núcleo
MSTRG.3644	2.16	protein kinase	Extracelular
MSTRG.3175	2.89	hypothetical protein	Núcleo
MSTRG.8449	2.27	hypothetical protein	Extracelular
MSTRG.3645	2.94	hypothetical protein	Mitocôndria
MSTRG.2955	2.24	S-methyl-5'- thioadenosine phosphorylase	Citosol
MSTRG.2957	2.35	S-methyl-5'- thioadenosine phosphorylase	Citosol
MSTRG.7990	2.67	hypothetical protein	Extracelular
MSTRG.3648	2.44	hypothetical protein	Extracelular
MSTRG.2856	2.13	hypothetical protein	Extracelular
MSTRG.62	2.81	hypothetical protein	Extracelular
MSTRG.8906	2.06	hypothetical protein	Extracelular
MSTRG.4353	2.19	hypothetical protein	Mitocôndria
MSTRG.9800	2.03	hypothetical protein	Membrana plasmática
MSTRG.8957	2.14	hypothetical protein	Retículo endoplasmático
MSTRG.9797	2.14	hypothetical protein	Membrana plasmática
MSTRG.3172	2.50	hypothetical protein	Extracelular
g8758	2.45	hypothetical protein	Núcleo
MSTRG.8024	2.08	hypothetical protein	Mitocôndria
MSTRG.54	2.44	hypothetical protein	Mitocôndria
MSTRG.2983	2.04	surface protease GP63	Membrana plasmática
g7681	2.40	trans-sialidase, Group I, putative	Núcleo
MSTRG.56	2.61	hypothetical protein	Núcleo
g7500	2.31	hypothetical protein	Extracelular
MSTRG.4496	2.01	hypothetical protein	Extracelular
g9068	2.21	hypothetical protein	Extracelular
g8690	2.08	hypothetical protein	Inconclusivo
MSTRG.1124	2.29	hypothetical protein	Extracelular
g9021	2.05	trans-sialidase, Group I, putative	Núcleo
MSTRG.1376	2.36	hypothetical protein	Extracelular
MSTRG.7210	2.30	hypothetical protein	Extracelular
MSTRG.8809	2.97	hypothetical protein	Extracelular
g8677	2.35	putative kinesin-like protein	Extracelular
g9078	2.71	hypothetical protein	Inconclusivo
MSTRG.6543	2.17	hypothetical protein	Núcleo
MSTRG.6228	2.48	surface protease GP63	Extracelular
g1005	2.09	surface protease GP63, putative	Extracelular

MSTRG.1374	2.12	hypothetical protein	Extracelular
MSTRG.6227	2.56	surface protease GP63, putative	Extracelular
MSTRG.45	2.36	hypothetical protein	Extracelular

APÊNDICE C – Transcritos de *T. rangeli* apresentando diminuição de expressão gênica quando comparado formas tripomastigotas contra epimastigotas.

Na Tabela 6 são listados todos os transcritos apresentando expressão diferencial com regulação negativa da transcrição quando comparado formas tripomastigotas contra epimastigotas de *T. rangeli*. São descritos os IDs, ou seja, o código de identidade de cada transcrito conferido pelo programa StringTie no momento da montagem do transcriptoma, os níveis de expressão diferencial obtidos pelo DESeq2, a anotação atribuída realizada por anotação por similaridade e a localização celular predita.

Tabela 5 - Totalidade de transcritos identificados exibindo diminuição de expressão gênica quando comparado formas tripomastigotas vs. epimastigotas.

ID	Nível de expressão gênica (logFC)	Anotação	Localização predita
MSTRG.216	-2.55	Cystathionine gamma-lyase	Extracelular
MSTRG.8333	-2.63	hypothetical protein	Núcleo
MSTRG.6761	-2.88	Cystathionine gamma-lyase	Extracelular
MSTRG.7268	-2.63	Sialidase domain-containing protein	Citosol
g2250	-2.52	Transketolase 1	Citosol
MSTRG.8525	-2.35	Tyrosine aminotransferase	Inconclusivo
MSTRG.1620	-2.15	Nodulin-like putative	Membrana plasmática
MSTRG.3653	-2.27	trans-sialidase	Citosol
MSTRG.5167	-2.22	85 kDa surface antigen	Mitocôndria
MSTRG.2206	-2.82	trans-sialidase	Extracelular
MSTRG.3632	-2.25	Nodulin-like	Membrana plasmática
MSTRG.3542	-2.32	Sialidase	Extracelular
MSTRG.4105	-2.14	Nodulin-like	Membrana plasmática
MSTRG.6662	-2.19	Nodulin-like	Membrana plasmática
MSTRG.8471	-2.66	exo-alpha-sialidase	Núcleo
MSTRG.4390	-2.33	complement regulatory protein	Citosol
MSTRG.673	-2.61	trans-sialidase	Citosol
g8744	-2.80	trans-sialidase	Citosol
MSTRG.6760	-2.33	L-methionine gamma-lyase	Mitocôndria
g5936	-2.03	Kinesin motor domain-containing protein	Citosol
MSTRG.8498	-2.23	Kinesin-II 95 kDa subunit	Citosol
MSTRG.1339	-2.48	hypothetical protein	Extracelular
MSTRG.3328	-4.48	trans-sialidase	Extracelular
g7611	-2.42	ABC transporter	Citosol

g5937	-2.03	Kinesin-II 95 kDa subunit	Citosol
MSTRG.5579	-2.42	85 kDa surface antigen	Núcleo
g5938	-2.04	Kinesin-II 85 kDa subunit	Citosol
MSTRG.2978	-2.33	85 kDa surface antigen	Núcleo
MSTRG.2238	-2.42	85 kDa surface antigen	Núcleo
MSTRG.6639	-2.06	85 kDa surface antigen	Núcleo
MSTRG.5578	-2.18	trans-sialidase	Núcleo
MSTRG.3544	-2.03	Calcium-transporting ATPase 2	Membrana plasmática
MSTRG.5469	-4.87	85 kDa surface antigen	Citosol
MSTRG.6668	-2.06	85 kDa surface antigen	Mitocôndria
