

Universidade Federal de Minas Gerais
Instituto de Ciências Biológicas
Programa de Pós-Graduação em Bioinformática

Tiago Bruno Rezende de Castro

**MODULAÇÃO DIFERENCIAL DA EXPRESSÃO GÊNICA EM CAMUNDONGOS
INFECTADOS COM DUAS CEPAS DE *TRYPANOSOMA CRUZI*: uma análise
transcriptômica**

Belo Horizonte

2019

Tiago Bruno Rezende de Castro

**MODULAÇÃO DIFERENCIAL DA EXPRESSÃO GÊNICA EM
CAMUNDONGOS INFECTADOS COM DUAS CEPAS DE
TRYPANOSOMA CRUZI: uma análise transcriptômica**

Tese apresentada ao Programa de Pós-Graduação em Bioinformática da Universidade Federal de Minas Gerais como requisito parcial à obtenção do título de Doutor em Bioinformática

Orientadora: Profa. Dra. Glória Regina Franco

Co-orientadora: Profa. Dra. Andrea Mara Macedo

Belo Horizonte

2019

043

Castro, Tiago Bruno Rezende de.

Modulação diferencial da expressão gênica em camundongos infectados com duas cepas de *Trypanosoma cruzi*: uma análise transcriptômica [manuscrito] / Tiago Bruno Rezende de Castro. – 2019.

101 f. : il. ; 29,5 cm.

Orientadora: Profa. Dra. Glória Regina Franco. Co-orientadora: Profa. Dra. Andrea Mara Macedo.

Tese (doutorado) – Universidade Federal de Minas Gerais, Instituto de Ciências Biológicas. Programa Interunidades de Pós-Graduação em Bioinformática.

1. Bioinformática. 2. *Trypanosoma cruzi*. 3. Doença de Chagas. 4. RNA-Seq. 5. Transcriptoma. I. Franco, Glória Regina. II. Macedo, Andrea Mara. III. Universidade Federal de Minas Gerais. Instituto de Ciências Biológicas. IV. Título.

CDU: 573:004



ATA DA DEFESA DE TESE

Tiago Bruno Rezende de Castro

108/2019
entrada
2º/2015
CPF:
016.314.216-54

Às quatorze horas do dia 05 de julho de 2019, reuniu-se, no Instituto de Ciências Biológicas da UFMG, a Comissão Examinadora de Tese, indicada pelo Colegiado do Programa, para julgar, em exame final, o trabalho intitulado: "Modulação Diferencial da Expressão Gênica Em Camundongos Infectados Com Duas Cepas de Trypanosoma Cruzi: Uma Análise Transcriptômica", requisito para obtenção do grau de Doutor em Bioinformática. Abrindo a sessão, a Presidente da Comissão, Dra. Glória Regina Franco, após dar a conhecer aos presentes o teor das Normas Regulamentares do Trabalho Final, passou a palavra ao candidato, para apresentação de seu trabalho. Seguiu-se a arguição pelos Examinadores, com a respectiva defesa do candidato. Logo após, a Comissão se reuniu, sem a presença do candidato e do público, para julgamento e expedição de resultado final. Foram atribuídas as seguintes indicações:

Prof./Pesq.	Instituição	CPF	Indicação
Dra. Glória Regina Franco	UFMG	623.387.496-34	APROVADO
Dra. Andrea Mara Macedo	UFMG	517.865.166.91	APROVADO
Dr. Marcelo Torres Bozza	UFRJ	98409018722	APROVADO
Dr. Pedro Alexandre Favoretto Galante	Instituto Sírio-Libanês de Ensino e Pesquisa	279450698-60	APROVADO
Dra. Mariana Torquato Quezado de Magalhaes	UFMG	905.464151-72	APROVADO
Dr. João Luis Reis Cunha	UFMG	095-105.936-05	APROVADO

Pelas indicações, o candidato foi considerado: APROVADO
O resultado final foi comunicado publicamente ao candidato pela Presidente da Comissão. Nada mais havendo a tratar, a Presidente encerrou a reunião e lavrou a presente ATA, que será assinada por todos os membros participantes da Comissão Examinadora.
Belo Horizonte, 05 de julho de 2019.

Dra. Glória Regina Franco - Orientadora Glória Regina Franco
Dra. Andrea Mara Macedo - Coorientadora Andrea Mara Macedo
Dr. Marcelo Torres Bozza Marcelo Torres Bozza
Dr. Pedro Alexandre Favoretto Galante Pedro Alexandre Favoretto Galante



Universidade Federal de Minas Gerais
Instituto de Ciências Biológicas
Programa Interunidades de Pós-Graduação em Bioinformática da UFMG

Dra. Mariana Torquato Quezado de Magalhaes

Mariana T. Q. de Magalhães

Dr. João Luís Reis Cunha

João Luís Reis Cunha

AGRADECIMENTOS

Gostaria de agradecer primeiramente a minha orientadora Glória Regina Franco, eterna chefe, por ter me proporcionado a oportunidade de iniciar uma fase importante da minha formação científica no Laboratório de Genética Bioquímica, 7 anos atrás. É uma grande alegria e uma honra poder concretizar este feito ao seu lado, a realização de uma ideia, um sonho, que surgiu durante a minha disciplina de Biologia Molecular na qual a própria chefe administrou: “E se eu pudesse ter esta incrível professora e cientista como minha mentora?” Sua paciência, carinho, companheirismo e inteligência serviram para me definir como a pessoa e o futuro pesquisador que almejo ser. Por ter sempre acreditado no meu potencial, pela enorme força dada e por ter aberto portas a minha frente, o que permitiu o meu próximo passo. Por tudo isso, obrigado chefe! Sua presença em minha vida nunca será esquecida!

Também gostaria de agradecer os professores e chefes do LGB, à minha co-orientadora Prof. Andréa Mara Macedo, ao Prof. Carlos Renato Machado, Prof. Erich Birelli Tahara e Prof Sergio Pena, pela contribuição de todos tanto ao projeto de pesquisa com inúmeras discussões científicas e no auxílio diário no laboratório, mas também pelos ensinamentos pessoais que levaram a contribuição da minha formação pessoa.

Gostaria de fazer um agradecimento a toda secretaria do programa de Bioinformática, principalmente a Sheila e Tiago por todo suporte incrível que foi dado durante esses 4 anos que fui aluno de Doutorado!

Agradeço também aos colegas do LGB pela convivência e companheirismo em todos os momentos alegres e difíceis nesses 7 anos, aos que já se formaram antes de mim e aos que ainda permanecem, boa sorte. A Neuza e a Lorrane obrigado pelo auxílio pessoal e técnico no laboratório. Faço um agradecimento especial ao Willian Prado, e Daniela Chame, pela amizade verdadeira, inquestionável e única, saibam que a convivência, a troca mútua de conhecimento, toda a ajuda nos momentos difíceis e companheirismo jamais poderá ser apagada, obrigado! Agradeço também ao Bruno Repolês, a Natalia Linhares, Daniela de Laet, Héllida Silva, Veronica Santana e a Nayara Toledo pela convivência no laboratório, a amizade de cada dia e por terem me aturado tanto tempo.

Não poderia deixar de agradecer aos amigos e colegas do laboratório de Imunologia de Mucosas e do laboratório de Dinâmica de Linfócitos da Universidade Rockefeller que tanto contribuíram para minha formação científica e pessoal em tão pouco tempo. Aos meus mentores Daniel Mucida e Gabriel Victora por terem aberto as portas de seus laboratórios e pela enorme contribuição a minha formação como pesquisador. Em especial aos colegas, Angelina Bilate, Ainsley Lockhart, Bernado Reis, Roham Parsa, Fanny Matheis, Paul Müller e Thiago Oliveira, obrigado por tudo que me ensinaram. Obrigado também a Johanne Jacobsen, Ariën Schiepers e Carla Nowosad pela paciência, amizade e todos os ensinamentos.

Agradeço a Mariya London pelo seu apoio e suporte de sempre, por estar sempre ao meu lado e ter me ajudado com paciência e solicitude todas as vezes que precisei no laboratório. Por toda sua compreensão, amizade e carinho!

Em um agradecimento especial, a Maria Cecilia. As amigas verdadeiras não são apagadas e nem esquecidas. Portanto, na luta contra o envelhecimento e a degeneração, preservo aqui nesse documento a grande amiga que você sempre foi, e a sua enorme contribuição para criação desta tese. É uma honra inimaginável saber que mesmo depois de tanta coisa ter acontecido e de tanto tempo ter passado, ainda tenho a oportunidade de trabalhar de perto com você e continuar a poder aprender com você.

Queria fazer um agradecimento especial ao meu pai Antônio Teotônio e a minha mãe Marly Eugênia, por terem sido os fundadores e os meus primeiros orientadores, enfim, os maiores responsáveis por eu estar onde eu estou hoje. Ambos colocaram em mim uma ideia, que me definiu por completo. A minha mãe por ter sempre me ensinado o valor de estudar, e de usar o conhecimento adquirido para sermos pessoas melhores. Sempre me empurrando pra frente e me acordando de manhã para ir à escola. Ao meu pai por ter sido o primeiro e maior mentor do que me tornei. Por ter me ensinado a ser persistente, forte e a não deixar de lutar por meus sonhos. Das curiosidades científicas, seus livros de astronomia e física. Por me instigar com suas histórias e com o seu conhecimento de informática. Com seu telescópio, que durante muitas noites ao observarmos as crateras da Lua, os anéis de Saturno e as manchas de Júpiter, alimentou a minha curiosidade infinita pelo desconhecido. Sementes, que depois de plantadas e cultivadas, se tornaram ideias que por final me definiram como pessoa. Agradeço ao meu

irmão, Matheus Henrique, pelo companheirismo e amizade. Por ter sido alguém cuja enorme inteligência e maturidade me mostraram qual o caminho eu devia seguir. Alguém de imensa admiração que levarei para sempre comigo e estarei sempre me espelhando.

Aos meus tios que tiveram uma grande participação na minha formação. Ao meu tio Eugênio e tia Socorro que me abrigaram em suas casas e custearam todos os meus estudos de pré-vestibular. Muito obrigado por sempre terem acreditado na minha capacidade e nos meus sonhos, pelo carinho e ternura de sempre. A tia Goretti que com seu carinho e doçura infinita sempre esteve por perto me apoiando e auxiliando.

Queria agradecer aos meus eternos amigos de Lagoa da Prata, Ricardo Samimel, Emerson Nydox, Cesar Kalks e Juninho Lenda. Vocês são demais! Amizades de uma vida toda. A convivência de vocês foi essencial e contribuiu muito para a minha formação e, portanto, para esta tese.

Não poderia deixar de agradecer aos amigos de Pará de Minas, em especial Guilherme Paiva, Suzanne Silva, Rodrigo Paiva, Camila Nogueira e Leticia Paiva. Foram muitos anos de convivência, desde o meu ensino médio até hoje, vocês presenciaram cada etapa desta escalada e sempre acreditaram, alimentando meu desejo para nunca desistir. A contribuição de vocês para esta tese não tem como ser paga. Podem ter certeza de que a amizade de vocês nunca será esquecida.

Queria fazer um agradecimento em conjunto a várias pessoas que merecem ser mencionadas, pois fizeram parte da minha vida e estiveram sempre atuando próximas na realização deste sonho. Ao Leandro Barbosa, Lucas Dhom, Barbara Cordeiro, Veronica Rodriguez, Karim Marins e minha prima Anne Caroline.

Por fim agradeço as agencias de fomento que pagaram os custos do projeto, e nossas bolsas, ao CNPq, a CAPES e FAPEMIG.

“Algumas pessoas me dizem:
Como você consegue viver sem saber?
Eu não entendo o que elas querem dizer!
Eu sempre vivi sem saber.
Isso é fácil.
Como você faz pra saber, é o que eu quero saber.”
(Richard Feynman, 1998
The Meaning of It All: Thoughts of a Citizen-Scientist)

RESUMO

O protozoário *Trypanosoma cruzi*, da classe Kinetoplastida, é o parasita causador da doença de Chagas em seres humanos. Há décadas médicos e cientistas têm estudado o mal de Chagas e buscado encontrar tratamentos para a doença, confrontando-se com problemas na criação de estratégias terapêuticas eficientes, que advêm, dentre outras causas, da distribuição tecidual variável dos parasitos em diferentes pacientes durante o estágio crônico, os quais podem afetar o coração ou órgãos do sistema digestivo, como o esôfago e o cólon. Pesquisas realizadas no Laboratório de Genética e Bioquímica da UFMG indicaram que características genéticas do hospedeiro e do *T. cruzi* estariam determinando os tropismos teciduais observados. No entanto, até o momento, não são conhecidas as bases moleculares desse tropismo diferencial, tendo o presente trabalho o objetivo de ampliar a compreensão do mecanismo de estabelecimento diferencial de duas cepas distintas de *T. cruzi*, as cepas Col1.7G2 (*T. cruzi* I) e JG (*T. cruzi* II), no coração e intestino (reto) de camundongos BALB/c na fase aguda da doença. O estudo do transcriptoma em alta profundidade (RNASeq) foi utilizado a fim de avaliar as diferenças de expressão gênica causadas pela infecção com as duas cepas de *T. cruzi* e com a mistura de ambas. Para realização deste estudo, os camundongos foram inoculados com PBS estéril (Controle), Col1.7G2, JG, ou uma mistura de ambas. Os animais foram sacrificados no 15º dia pós-infecção, e os corações e um segmento do reto foram extraídos e preservados em nitrogênio líquido. Posteriormente, o RNA total foi extraído e submetido a sequenciamento de nova geração na plataforma Illumina-Hiseq 2000/2500. As sequencias geradas foram avaliadas segundo sua qualidade e em seguida a abundância de transcritos foi medida pela técnica de pseudoalinhamento com o programa Kallisto e usando o transcriptoma do camundongo. Finalmente, o pacote estatístico Sleuth foi usado para determinar os genes diferencialmente expressos entre os grupos infectados e os controles. As categorias de ontologia gênica, como processos biológicos, foram avaliadas utilizando o pacote TopGO implementado em R. Os dados obtidos mostram que as cepas de *T. cruzi* induzem no hospedeiro perfis de expressão gênica distintos. Na análise de componentes principais, as amostras de corações de camundongos controle e infectados com Col1.7G2 e JG formaram grupos separados. Já os camundongos infectados com a mistura foram agrupados juntamente com JG e

Col1.7G2. Camundongos infectados com Col1.7G2, exibiram grande quantidade de genes superexpressos relacionados à resposta imune do hospedeiro (incluindo vias do sistema imune inato e adaptativo), resposta inflamatória e a patógenos (vírus, bactérias e protozoários). Por outro lado, o grupo infectado com JG mostrou, além de uma quantidade menor de genes de categorias inflamatórias, a subexpressão de uma grande quantidade de genes das proteínas ribossomais (subunidade grande e pequena do ribossomo) e de proteínas que atuam na mitocôndria como enzimas do ciclo do ácido cítrico e da cadeia transportadora de elétrons. Por sua vez, os camundongos infectados com a mistura das cepas exibiram ambos os perfis de expressão gênica. Em contrapartida, o reto desses animais não apresentou grandes alterações de expressão gênica global entre as diferentes cepas no período de infecção estudado. Em conjunto, os dados obtidos permitem compreender melhor a complexa interação parasito-hospedeiro no contexto da doença de Chagas experimental e o efeito que diferentes cepas de *T. cruzi* podem causar na expressão gênica do hospedeiro.

Palavras-chave: *Trypanosoma cruzi*; Doença de Chagas; RNA-Seq; Cepas JG e Col1.7G2; Transcriptoma

ABSTRACT

The protozoan *Trypanosoma cruzi*, from the Kinetoplastida class, is the causative agent of the Chagas disease in humans. For decades, physicians and scientists have been studying the *mal de Chagas* and searching for disease treatments, while also dealing with the creation of efficient therapeutic strategies that deal with the variable tissue distribution of parasites in different patients during the chronic stage, which can affect the heart and other organs of the digestive system, such as the esophagus and colon. Research performed at the Laboratório de Genética e Bioquímica (LGB) of UFMG, indicated that genetic characteristics of the host and of the *T. cruzi* parasites were important to define the tissue tropism observed. However, until the present moment, there was still a lack of understanding of the molecular basis of this differential tissue tropism. Thus, the goal of the present work is to expand the knowledge of this mechanism of differential distribution by distinct *T. cruzi* strains, Col1.7G2 (*T. cruzi* I) and JG (*T. cruzi* II) strains, in the hearts and intestines (rectums) of BALB/c mice in the acute phase of the disease. High throughput RNA sequencing (RNA-Seq) was used to evaluate the expression differences caused by the infection with two strains of *T. cruzi*. To perform this study, mice were inoculated with sterile PBS (Control), Col1.7G2, JG or a mixture of both. The animals were sacrificed on the 15th day post-infection and the hearts and a segment of the rectum were removed and preserved inside liquid nitrogen. Next, the RNA was extracted and submitted to sequencing by the Illumina-Hiseq 2000 platform. The generated reads were evaluated by their quality and later the transcript abundance was measured by a pseudoalignment technique with the software Kallisto using the mouse transcriptome (Gencode release M21). Finally, the statistical package Sleuth was used to determine the differentially expressed genes among the experimental groups and the non-infected controls. The gene ontology categories, biological processes, were evaluated by the package TopGO, implemented in R. The obtained data showed that the *T. cruzi* strains induce distinct expression profiles. In the principal component analysis, heart of control and infected mice formed separated groups. The mice infected with a mixture were grouped in between JG and Col1.7G2. Interestingly, Col1.7G2-infected mice exhibit a high number of upregulated genes related to the immune response of the host (including innate and adaptive pathways), inflammatory response and pathogens (virus, bacteria and protozoans). On

another hand, JG-infected mice showed not only fewer genes related to the inflammatory pathways, but also a strong downregulation of ribosomal protein genes (large and small subunits) and of proteins that have a role inside the mitochondria such as the citric acid cycle and the electron transport chain. Moreover, mice infected with a mixture of Col1.G2 and JG strains showed both gene expression profiles. Conversely, the rectum of these animals did not show substantial gene expression changes in this period of infection. Taken together, the data obtained allow us to better understand the complex host-pathogen interaction in the context of the experimental Chagas disease and the effects that two *T. cruzi* strains may have on the host gene expression.

Keywords: Trypanosoma cruzi; Chagas disease; RNA-Seq; JG e Col1.7G2 Strains; Transcriptome

LISTA DE ABREVIATURAS E SIGLAS

AP-1	Activator Protein 1
CAGE	Cap Analysis Of Gene Expression
CD8	Cluster Of Differentiation 8
CIITA	Class II Major Histocompatibility Complex Transactivator
CoQ	Ubiquinol
DNA	Deoxyribonucleic Acid
DTU	Discrete Typing Units
EST	Expressed Sequence Tag
GDE	Genes Diferencialmente Expressos
GPI	Glicosilfosfatidilinositol
IFN- γ	Interferon Gamma
IL-12	Interleukin 12
IL-6	Interleukin 6
KO	Knockout
LSSP	Low-Stringency Single Specific Primer
MAPK	Mitogen-Activated Protein Kinases
MASP	Mbl-Associated Serine Proteases
MBL	Mannose-Binding Lectin
MHC	Major Histocompatibility Complex
MutT	8-Oxo-Dgtp Diphosphatase
NAD	Nicotinamide Adenine Dinucleotide
NF κ B	Nuclear Factor Kappa-Light-Chain-Enhancer Of Activated B Cells
OG	8-Oxo-7 8-Dihydroguanine
PAMP	Pathogen-Associated Molecular Pattern
PCR	Polymerase Chain Reaction
RFLP	Restriction Fragment Length Polymorphisms
RNA	Ribonucleic Acid
RNA-Seq	Rna Sequencing
ROS	Reactive Oxygen Species
SAGE	Serial Analysis Of Gene Expression
TBX21	T-Box Transcription Factor 21
TLR	Toll-Like Receptor
TNF	Fatores De Necrose Tumoral Alfa
TPM	Transcripts Per Kilobase Million
TRV	Teste Da Razão De Verossimilhança
TS	Transialidase

LISTA DE FIGURAS

Figura 1 - Ilustração da etimologia do nome <i>Trypanosoma cruzi</i>	18
Figura 2 - Morfologia do T. cruzi	19
Figura 3 - Disseminação da doença de Chagas além da América Latina	20
Figura 4 - Ciclo de vida do <i>Trypanosoma cruzi</i>	21
Figura 5 - Prevalência das formas Chagásicas de acordo com a faixa etária	24
Figura 6 - As tecnologias de estudo do transcriptoma no contexto histórico	27
Figura 7 - Resumo da construção de bibliotecas e do sequenciamento pelo método do RNA-Seq	28
Figura 8 - Design experimental	35
Figura 9 - Criação de bibliotecas stranded-specific	40
Figura 10 - Código para criação de um índice para o kallisto e as listas de input para kallisto quant	41
Figura 11 - Código para execução de kallisto e quantificação da abundância de todas as amostras.....	42
Figura 12 - Código em R para execução da análise de detecção de genes diferencialmente expressos usando o pacote Sleuth	44
Figura 13 - Código em R para execução da análise de detecção de enriquecimento de termos de GO usando o pacote topGO	45
Figura 14 - Código em R para criação de diagramas de Venn de três conjuntos, usando o pacote Vennr	48
Figura 15 - Código em R para criação de gráficos de dispersão em formato vulcão (Volcano plot) usando o pacote ggplot2	50
Figura 16 - Design experimental completo	51
Figura 17 - Genotipagem das cepas utilizadas nas infecções	54
Figura 18 - JG e Col1.7G2 diferem quanto a parasitemia e mortalidade de camundongos BALB/c na fase aguda	56
Figura 19 - JG e Col1.7G2 apresentam perfil de expressão gênica distintos no coração	61
Figura 20 - JG e Col1.7G2 modulam a expressão gênica no coração	64
Figura 21 - Teste da razão de verossimilhança e agrupamento de genes	66
Figura 22 - Análise de enriquecimento de processos biológicos	68

Figura 23 - Rede de interação entre PBs significantes em Col1.7G2, JG e mistura	70
Figura 24 - Heatmap de vias metabólicas e imunologias do KEGG	72
Figura 25 - Análise do sequenciamento total do reto de camundongos infectados com JG, Col1.7G2 e a mistura	74
Figura 26 - Modelo proposto para a distribuição diferencial de JG e Col1.7G2 em uma infecção mista	84

LISTA DE TABELAS

Tabela 1 - Integridade do RNA total do coração enviado ao sequenciamento	57
Tabela 2 - Integridade do RNA total do reto enviado ao sequenciamento.....	58
Tabela 3 - Tamanho e qualidade das bibliotecas	58

SUMÁRIO

1. INTRODUÇÃO	19
1.1 TRYPANOSOMA CRUZI E A DOENÇA DE CHAGAS	20
1.2 EVOLUÇÃO DA DOENÇA DE CHAGAS.....	24
1.3 O ESTUDO DA DISTRIBUIÇÃO DIFERENCIAL DO <i>T. CRUZI</i>	26
1.4 TECNOLOGIAS DE SEQUENCIAMENTO DE NOVA GERAÇÃO E O ESTUDO DO TRANSCRIPTOMA	28
1.5 O ESTUDO DO TRANSCRIPTOMA DO <i>T. CRUZI</i> E DO SEU HOSPEDEIRO MAMÍFERO INFECTADO.	31
2. JUSTIFICATIVA	33
3. OBJETIVOS.....	35
4. MATERIAIS E MÉTODOS.....	37
4.1 ANIMAIS EXPERIMENTAIS	38
4.2 AQUISIÇÃO E GENOTIPAGEM DOS PARASITOS.....	39
4.3 OBTENÇÃO DO RNA TOTAL	39
4.3.1 CONTROLE DE QUALIDADE.....	40
4.4 SEQUENCIAMENTO DAS AMOSTRAS	41
4.4.1 SEQUENCIAMENTO DE CORAÇÕES.....	41
4.4.2 SEQUENCIAMENTO DE RETOS.....	41
4.5 PIPELINE PARA ANÁLISE DOS DADOS DE RNA-SEQ GERADOS NA PLATAFORMA ILLUMINA HISEQ 2000/2500.....	43
4.5.1 ANÁLISE DE QUALIDADE	43
4.5.2 QUANTIFICAÇÃO DA ABUNDÂNCIA DE TRANSCRITOS	44
4.5.3 NORMALIZAÇÃO E ANÁLISE.....	45
4.5.4 ANÁLISES DE CATEGORIAS FUNCIONAIS.....	48
4.5.5 ANÁLISES DE REDES DE PROCESSOS BIOLÓGICOS	50
4.5.6 CONSTRUÇÃO DO DIAGRAMA DE VENN	51
4.5.7 CONSTRUÇÃO DE GRÁFICOS DE VULCÃO.....	51
4.5.8 DETECÇÃO DE SEQUENCIAS DE PARASITO <i>T. CRUZI</i>	53
4.5.7 PIPELINE RESUMIDO	54
5. RESULTADOS	55
5.1 INFECÇÕES DE CAMUNDONGOS BALB/c COM JG E COL1.7G2.....	56
5.1.1 TIPAGEM DAS CEPAS UTILIZADAS NAS INFECÇÕES.....	56
5.1.2 MORTALIDADE E PARASITEMIA DOS CAMUNDONGOS INFECTADOS ATÉ 15 E 60 DIAS	58
5.2 ESTUDO DO TRANSCRIPTOMA DE CORAÇÕES INFECTADOS POR DIFERENTES CEPAS DE <i>T. CRUZI</i>	60
5.2.1 ANÁLISE QUANTITATIVA E INTEGRIDADE DO RNA TOTAL.....	60
5.2.2 CONTROLE DE QUALIDADE DAS BIBLIOTECAS.....	61
5.2.3 PERFIL DIFERENCIAL DE INFECÇÃO NO CORAÇÃO POR JG, COL1.7G2 E POR MISTURA	63
5.2.5 GENES DIFERENCIALMENTE EXPRESSOS PELA INFECÇÃO EM RELAÇÃO AO CONTROLE NÃO INFECTADO.....	66
5.2.5 DIFERENÇAS GLOBAIS DE EXPRESSÃO GÊNICA POR COL1.7G2, JG E A MISTURA	68

5.2.6 PROCESSOS BIOLÓGICOS ENRIQUECIDOS PELA INFECÇÃO POR JG E COL1.7G2.....	70
5.2.7 ANÁLISE DE REDES DE INTERAÇÃO ENTRE PROCESSOS BIOLÓGICOS	72
5.2.8 VISUALIZAÇÃO DE GENES DIFERENCIALMENTE EXPRESSOS EM VIAS METABÓLICAS DO KEGG	74
5.3 ESTUDO DO TRANSCRIPTOMA DE RETOS INFECTADOS POR DIFERENTES CEPAS DE <i>T. CRUZI</i>.....	76
6. DISCUSSÃO	78
7. CONCLUSÕES	86
8. REFERÊNCIAS BIBLIOGRÁFICAS	89
9. ANEXOS	99

1. INTRODUÇÃO

1.1 *Trypanosoma cruzi* e a doença de Chagas

O parasito *Trypanosoma cruzi*, do grego *trypanon*, broca, e *soma*, corpo (Figura 1), agente etiológico da doença de Chagas, é um protozoário flagelado pertencente à ordem Trypanosomatida, classe Kinetoplastea. O nome Kinetoplastea está relacionado à presença de uma estrutura característica, denominada cinetoplasto (Figura 2), uma massa densa de DNA extranuclear contida na única mitocôndria presente no citoplasma dos organismos dessa classe (DESCHAMPS; LARA; MARANDE; LOPEZ-GARCIA *et al.*, 2011). Outra peculiaridade interessante presente no grupo é a compartimentalização da glicólise em organelas denominadas glicossomos, estruturas evolutivamente relacionadas aos peroxissomos (GABALDON; GINGER; MICHELS, 2016). O *T. cruzi* é um organismo primariamente diploide e utiliza o mecanismo de fissão binária como principal estratégia reprodutiva (ZINGALES; MILES; CAMPBELL; TIBAYRENC *et al.*, 2012). Além disso, a espécie possui alta diversidade intraespecífica, e um grande número de cepas do parasito já foi identificada. Alguns estudos sugerem que o tipo de cepa tenha influência nas consequências da infecção pelo *T. cruzi* e na progressão da doença de Chagas (CAMPBELL; WESTENBERGER; STURM, 2004).

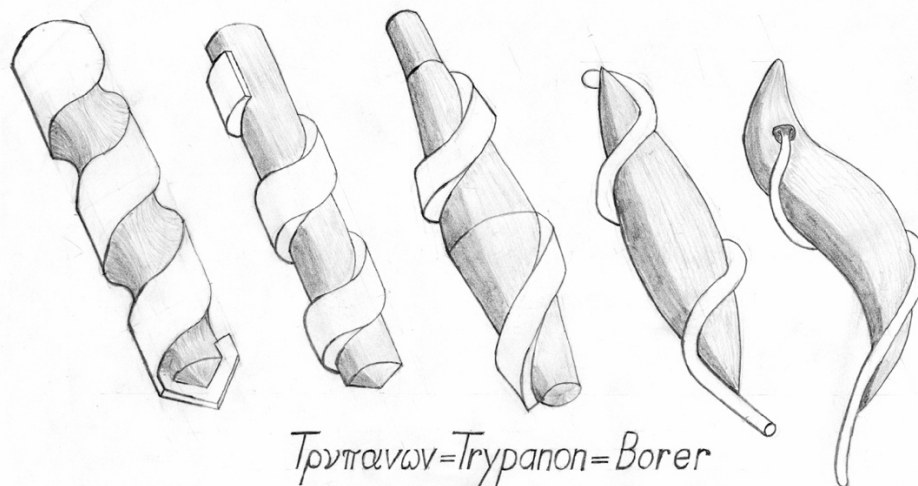


Figura 1: Ilustração da etimologia do nome *Trypanosoma cruzi* Ilustração gentilmente cedida por André Schneider – University of Bern/ Zürich

Inúmeras metodologias, como análises de polimorfismos de tamanho de fragmentos de restrição (RFLP – *Restriction Fragment Length Polymorphisms*), impressões digitais de DNA, LSSP-PCR (*Low Stringency Single Primer PCR*), perfis

de microssatélites, polimorfismos nos genes de rRNA, mini-éxon e mitocondriais, dentre outras, demonstraram quão vastas podem ser as variações genéticas do protozoário (BURGOS; ALTCHER; BISIO; DUFFY *et al.*, 2007; DE FREITAS; AUGUSTO-PINTO; PIMENTA; BASTOS-RODRIGUES *et al.*, 2006; SOUTO; FERNANDES; MACEDO; CAMPBELL *et al.*, 1996; ZINGALES; MILES; CAMPBELL; TIBAYRENC *et al.*, 2011). Em função da multiplicidade de classificações que existiam na literatura, em 2009 convencionou-se a subdivisão da espécie *T. cruzi* em seis DTUs (*Discrete Typing Units*): TcI, TcII, TcIII, TcIV, TcV e TcVI, além de uma sétima, que foi postulada em seguida (MARCILI; LIMA; CAVAZZANA; JUNQUEIRA *et al.*, 2009; ZINGALES; ANDRADE; BRIONES; CAMPBELL *et al.*, 2009). As DTUs são definidas como grupos de isolados que podem ser identificados por um conjunto de marcadores moleculares (TIBAYRENC, 1998) e constituem unidades confiáveis para estudos epidemiológicos e evolutivos. É importante salientar que cepas inseridas em uma mesma DTU devem ser vistas como famílias de clones relacionados e não como clones únicos (ZINGALES; MILES; CAMPBELL; TIBAYRENC *et al.*, 2012).

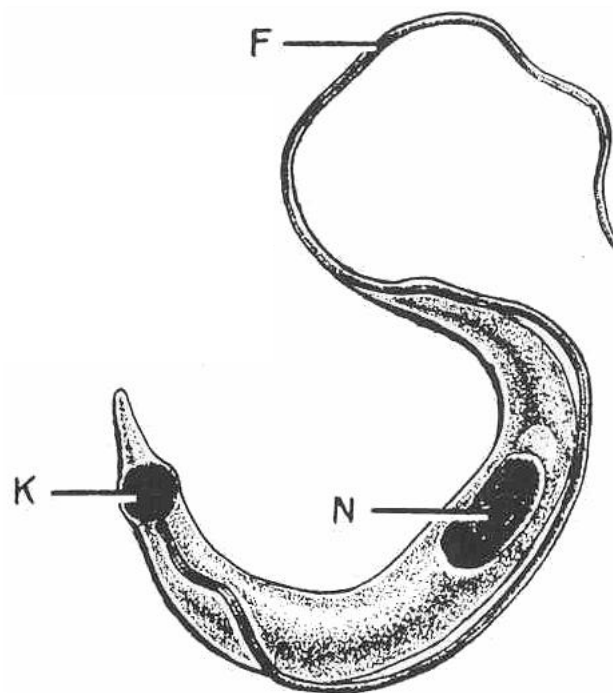


Figura 2: Morfologia do *T. cruzi* Ilustração da forma tripomastigota do *T. cruzi*, destacando o cinetoplasto (K), o núcleo (N) e o flagelo (F), Adaptado de: Parasitology Protozoology and Helminthology.

A doença de Chagas, ou Tripanossomíase Americana, foi inicialmente descrita por Carlos Chagas, pesquisador do Instituto Oswaldo Cruz, que descreveu a patologia em sua totalidade, caracterizando o vetor, o agente etiológico, os hospedeiros, bem como os aspectos clínicos e epidemiológicos (CHAGAS, 1909). Sendo considerada uma doença tropical negligenciada, a infecção humana pelo *T. cruzi* é endêmica nos países das Américas Central e do Sul, onde é primariamente associada à existência natural do inseto transmissor. Todavia, a globalização e, conseqüentemente, os eventos migratórios frequentes têm propiciado o surgimento de casos da doença de Chagas além dos países endêmicos (Figura 3), como no Canadá, Europa, Japão e Austrália (COURA; VIÑAS, 2010).

Graças a diversas medidas de controle vetorial adotadas, a Organização Pan-Americana da Saúde (*Pan American Health Organization, PAHO*) certificou o fim da transmissão da doença por vetores domésticos em diversos países das Américas do Sul e Central (SCHOFIELD *et al.*, 2006). Apesar dos avanços alcançados nas últimas décadas em relação ao controle da transmissão vetorial e transfusional do mal de Chagas, dados recentes estimam que 6 a 7 milhões de indivíduos ainda estejam infectados pelo *T. cruzi*, e 25 milhões permaneçam em risco de contraírem a infecção em 21 países endêmicos (CONNERS; VINETZ; WEEKS; BROUWER, 2016; RASSI; RASSI; MARIN-NETO, 2010; WHO, 2017).



Figura 3: Disseminação da doença de Chagas além da América Latina. Adaptado de (COURA; VIÑAS, 2010).

O *T. cruzi* é um parasito digenético e apresenta um complexo ciclo de vida, que envolve alternância entre hospedeiros vertebrados e invertebrados. Nas áreas

endêmicas, onde a transmissão vetorial é a mais frequente, o ciclo da doença de Chagas se inicia quando um barbeiro hematófago da família Reduviidae, como o *Rhodinus prolixus* ou o *Triatoma infestans*, durante seu repasto sanguíneo, ingere formas tripomastigotas do sangue do hospedeiro vertebrado. A partir daí, elas transformam-se em formas epimastigotas dentro do intestino do inseto e se replicam sucessivas vezes por fissão binária. No reto do barbeiro, tornam-se tripomastigotas metacíclicas, que são eliminadas nas fezes e podem infectar um novo hospedeiro vertebrado. Uma vez na circulação sanguínea de mamíferos, as tripomastigotas podem penetrar nas células nucleadas de diversos tecidos. Os parasitos conseguem sobrepujar as defesas naturais da célula e evadir os mecanismos de digestão lisossomal. Uma vez no citoplasma, os parasitos tornam-se amastigotas. Uma única amastigota pode replicar até 9 vezes, chegando a produzir 500 novos parasitos (TAN; ANDREWS, 2002). Com o aumento de seu número, as amastigotas intracelulares se transformam novamente na forma tripomastigota para, então, lisarem a célula e entrarem novamente na corrente sanguínea (Figura 4). A partir daí, as tripomastigotas podem infectar novas células no mesmo tecido ou em outros, ou serem ingeridas por novos vetores, dando início a um novo ciclo (DE SOUZA, 2002).

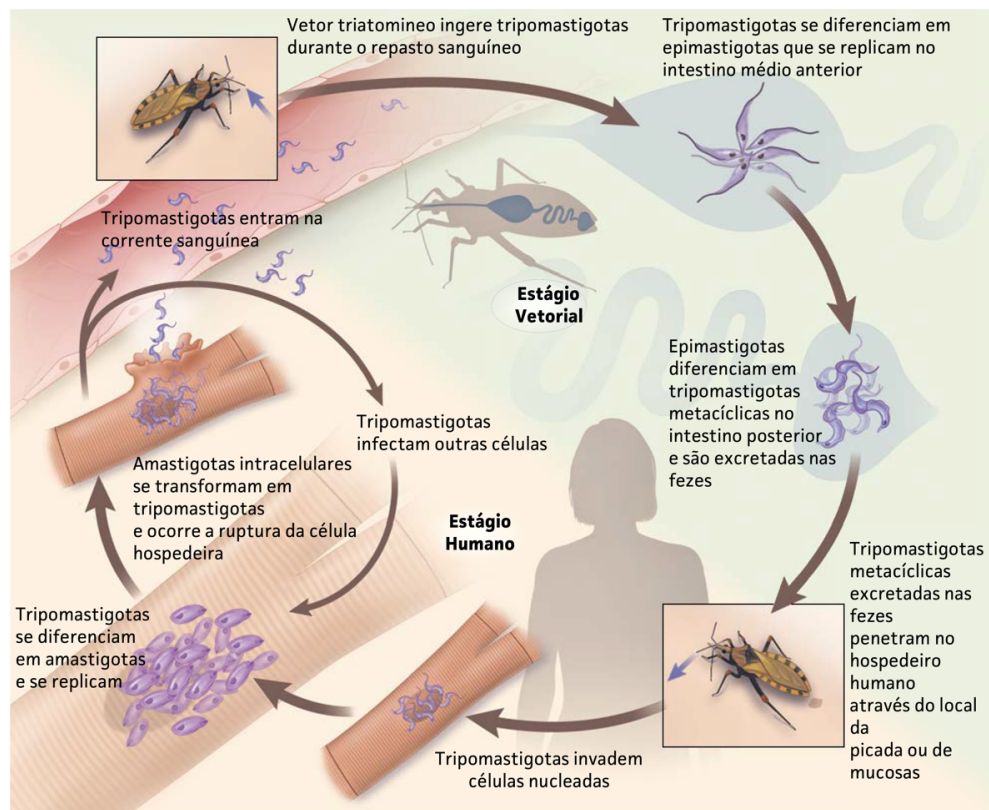


Figura 4: Ciclo de vida do *Trypanosoma cruzi*. Adaptado de: (LONGO; BERN, 2015).

Algumas formas de infecção não vetoriais ganham maior importância no cenário médico internacional atualmente, em especial: transfusão sanguínea, transplante de órgãos, via oral e de forma congênita. Nos países não-endêmicos, a principal forma de transmissão da doença é por transfusão de sangue, pois neles a realização de testes sanguíneos contra o *T. cruzi* é pouco comum (DUTRA; ROCHA; TEIXEIRA, 2005).

1.2 Evolução da doença de Chagas

Após a infecção, a doença de Chagas humana apresenta uma fase inicial aguda, de intensidade e duração variáveis, na qual ocorre exacerbada multiplicação do parasito e manifestações clínicas como febre, aumento do volume dos linfonodos, hepatomegalia e náusea. Na grande maioria dos indivíduos a etapa é assintomática, ou há um desenvolvimento de sintomas leves, frequentemente confundidos com uma gripe (RASSI; RASSI; DE REZENDE, 2012). Crianças podem, entretanto, desenvolver sintomas mais severos depois de alguns dias de incubação, incluindo um edema periorbital (sinal de Romaña), conjuntivite, febre, náusea, vômito, diarreia, linfadenopatia, hepatoesplenomegalia, erupções cutâneas, anorexia, dentre outros (DIAS; LARANJA; MIRANDA; NOBREGA, 1956). Nesse período, uma mistura de células inflamatórias polimorfonucleares e mononucleares é encontrada junto a pseudocistos, contendo amastigotas espalhadas entre fibras musculares. Ao mesmo tempo, células tripomastigotas metacíclicas são encontradas no sangue periférico, sendo facilmente observáveis em microscópio óptico, o que facilita o diagnóstico nesse momento (LONGO; BERN, 2015).

Estudos imunológicos das últimas décadas têm ajudado a dissecar as bases moleculares da infecção por *T. cruzi*. O reconhecimento de padrões Moleculares Associados a Patógenos (PAMPs) advindos de um microrganismo invasor é realizado principalmente por macrófagos residentes e células dendríticas. Dentre os fatores do sistema imune inato mais estudados no contexto da doença de Chagas experimental podemos citar o reconhecimento de PAMPs por receptores do tipo Toll9 e Toll2 (TLRs), que se ligam, respectivamente, ao DNA não metilado do parasito e a âncoras de Glicosilfosfatidilinositol (GPI), amplamente distribuídas na sua membrana

(CAMPOS; ALMEIDA; TAKEUCHI; AKIRA *et al.*, 2001; MOGENSEN, 2009). Estes TLRs tiveram sua importância avaliada em infecções por *T. cruzi* em camundongos *knockout* (KO), que demonstraram susceptibilidade significativamente maior à infecção quando comparados a animais WT (BAFICA; SANTIAGO; GOLDSZMID; ROPERT *et al.*, 2006; GRAVINA; ANTONELLI; GAZZINELLI; ROPERT, 2013). Após a sinalização de TLRs, uma ampla cascata de sinalização mediada por células inflamatórias, citocinas, quimiocinas e fatores de transcrição é determinante para o curso da fase aguda (MOGENSEN, 2009). Por exemplo, a expressão de fatores de transcrição como AP-1 e NFκB é responsável por aumentar a expressão de citocinas e quimiocinas pró-inflamatórias, que terão papel na diferenciação de células do sistema imune inato e adaptativo, no aumento da produção de fagócitos na medula óssea, e no recrutamento de fagócitos para os sítios de infecção (GILMORE, 1999).

A notável capacidade de evasão ao sistema imune do *T. cruzi* pode ser decisiva para seu estabelecimento no hospedeiro (CARDOSO; REIS-CUNHA; BARTHOLOMEU, 2015). Padilha *et al.*, demonstraram que a capacidade de evasão do *T. cruzi* relaciona-se a uma insuficiente ativação de TLRs no início da infecção. O pico do número de linfócitos CD8+ parasito-específicos é observado com 3 semanas após a infecção inicial, tempo consideravelmente maior do que o da maioria das infecções causadas por bactérias e vírus. Esse pico é significativamente antecipado quando camundongos são tratados com agonistas de Toll2 e Toll9 (PADILLA; SIMPSON; TARLETON, 2009). Caso os camundongos sobrevivam, são capazes de desenvolver uma resposta imune adaptativa potente e efetiva que elimina a maioria dos parasitos. Todavia, a eliminação não é completa, e ninhos de amastigotas podem permanecer por longos períodos dentro de cardiomiócitos, gânglios entéricos do músculo esquelético e tecido adiposo (PEREZ; LYMBERY; THOMPSON, 2014). Estudos recentes sugerem que formas dormentes não replicativas podem explicar a permanência prolongada do *T. cruzi* em diversos tecidos durante a fase crônica (SANCHEZ-VALDEZ; PADILLA; WANG; ORR *et al.*, 2018). Além disso, cepas diferentes podem promover respostas imunes notavelmente distintas, o que contribui para a variabilidade da doença na fase aguda e, conseqüentemente, na fase crônica (RODRIGUES, C.; VALADARES, H.; FRANCISCO, A.; ARANTES, J. *et al.*, 2010).

A maioria dos indivíduos sobrevive à fase aguda e evolui para a fase crônica da doença. O nível de sequelas ocorridas durante a fase aguda, como a destruição de gânglios nervosos, tanto no sistema digestivo quanto no coração, já foi reconhecido

como uma possível explicação para as diferentes sobrevivências de pacientes durante a fase crônica (KÖBERLE, 1970). Nessa fase, não é mais possível verificar formas do parasito circulantes no sangue. Além do mais, grande parte dos indivíduos sorologicamente positivos para *T. cruzi*, não apresenta manifestações clínicas importantes, caracterizando a forma assintomática ou indeterminada da doença. Por motivos ainda não conhecidos, a forma mais frequente e principal manifestação clínica da doença de Chagas é de pacientes que evoluem para a cardiopatia chagásica crônica. A forma digestiva, representada por alterações estruturais e funcionais do intestino e esôfago, é menos frequente. Já a forma neural, descrita por alguns autores, é rara e manifesta-se principalmente em pacientes imunossuprimidos (MASOCHA; KRISTENSSON, 2012; NETO; GRUBER; ZINGALES; KALIL, 1995). As prevalências de cada forma chagásica variam de acordo a idade do paciente (COURA, R. J.; ANUNZIATO, N.; WILLCOX, H. P. F., 1983). Interessantemente, é cada vez menos frequente observar pacientes assintomáticos com idade avançada (Figura 5).

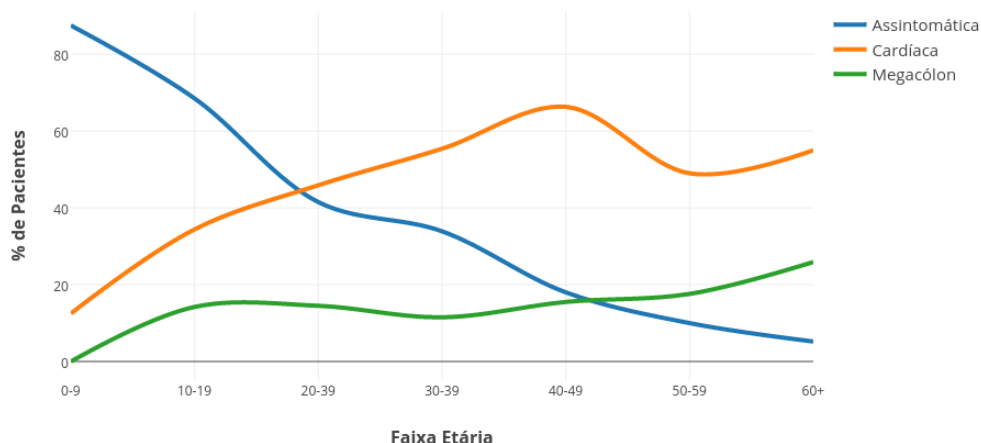


Figura 5: Prevalência das formas Chagásicas de acordo com a faixa etária. Adaptado de: (COURA, J. R.; ANUNZIATO, N.; WILLCOX, H. P. F., 1983).

1.3 O estudo da distribuição diferencial do *T. cruzi*

Desde os estudos iniciais sobre a doença de Chagas, pesquisadores têm chamado a atenção para uma distribuição diferencial do parasito em tecidos específicos (CAMPOS, 1927; VIANNA, 1911). Entretanto, apesar do enorme avanço científico adquirido em mais de um século de estudo, as causas da grande variabilidade das formas encontradas em pacientes chagásicos de diferentes regiões, idades e também em animais experimentais segue ainda pouco compreendida, assim

como o chamado tropismo da doença de Chagas. (ANDRADE; CARVALHO; FIGUEIRA, 1970; BRENER, 1965; COURA, R. J.; ANUNZIATO, N.; WILLCOX, H. P. F., 1983; KÖBERLE, 1970; MACEDO; OLIVEIRA; PENA, 2002; MARTINS; CASTANHO; DA ROSA; DA SILVA *et al.*, 2003; MELO; BRENER, 1978). É interessante ressaltar que o termo tropismo, do grego *virada/giro*, tem sido usado para conceitos diferentes por muitos autores (HUSSON, 1968). O termo foi cunhado em 1835 pelo botânico Augustin Candolle para descrever o movimento dos caules de plantas em direção a uma fonte de luz. Somente em 1918, quando o pesquisador Jacques Loeb do instituto de pesquisas médicas da Rockefeller publicou seu livro intitulado: *Forced movements, tropisms, and animal conduct*, o termo tropismo passou a ser usado para descrever um movimento direto e involuntário em direção a uma fonte de energia, fosse química, térmica ou luminosa, não apenas de plantas, mas inclusive de animais 'inferiores' e 'superiores' (GUSSIN, 1963). O conceito foi amplamente criticado pela comunidade científica, não apenas pela sua visão simplista do comportamento biológico, mas também por se limitar a explicar poucas situações em organismos inferiores. Além disso, mesmo quando era capaz de explicar os fenômenos observados, ignorava que estruturas celulares e moleculares com um contexto evolutivo, processos estocásticos e fisiológicos estavam envolvidos (BUDDENBROCK, 1916). Apesar da 'teoria do tropismo' não ter ganhado a amplitude que Loeb almejou, o termo tem desde então sido usado na microbiologia para descrever uma afinidade intrínseca do patógeno para algum determinado tipo celular ou tecido, em que o micróbio seja capaz de sobreviver e se reproduzir (MCFADDEN; MOHAMED; RAHMAN; BARTEE, 2009).

Entender os mecanismos e interações entre o patógeno e seu hospedeiro é fundamental para o entendimento da progressão da doença e o desenvolvimento de novos tratamentos (MCCALL; SIQUEIRA-NETO; MCKERROW, 2016). Em 1999, Andrade *et al.* mostraram que, ao se infectar camundongos BALB/c isogênicos com duas cepas diferentes de *T. cruzi*, a saber, Col1.7G2 (*T. cruzi* I) e JG (*T. cruzi* II), cada cepa apresentava um tropismo diferenciado. Durante a fase aguda, o DNA do parasito era encontrado em todos os tecidos estudados (coração, cérebro, reto, diafragma, esôfago e sangue), e não houve padrão claro de tropismo quando as cepas foram infectadas isoladamente, inclusive na fase crônica. Entretanto, quando as cepas eram infectadas na forma de uma mistura artificial, observava-se que a cepa JG adquiria um tropismo preferencial para o coração, enquanto que Col1.G2 era encontrada em

outros tecidos, incluindo o reto, demonstrando-se assim o papel das infecções mistas no desenvolvimento da doença (ANDRADE; MACHADO; CHIARI; PENA *et al.*, 1999). Posteriormente, o mesmo estudo foi aplicado em diferentes linhagens de camundongo, e além da reprodução dos dados anteriores em animais BALB/c, resultados semelhantes foram obtidos para camundongos DBA/6J. Por outro lado, foi observado um tropismo bem distinto para camundongos C57BL/6J e SWISS. A observação de que camundongos apresentando distribuição tecidual semelhante, possuíam o mesmo haplótipo de MHC levou Freitas *et al.* a avaliarem a relevância desta região cromossômica usando camundongos congênicos, isto é, camundongos que possuem haplótipo de MHC de uma linhagem, mas *background* genético de outra (ROGNER; AVNER, 2003), e demonstrarem que era possível alterar o tropismo tecidual apenas com a variação dessa porção cromossômica. Além disso, Vago *et al.* usaram assinaturas do kDNA geradas por PCR de baixa estringência (LSSP-PCR) em humanos na fase crônica para estudar as populações de cepas nos tecidos infectados (ADAMS, 1996; VAGO; ANDRADE; LEITE; D'AVILA REIS *et al.*, 2000; VAGO; MACEDO; OLIVEIRA; ANDRADE *et al.*, 1996) e observaram que, em um mesmo indivíduo, havia cepas de *T. cruzi* distintas em tecidos diferentes. A partir daí originou-se o modelo histotrópico clonal, que buscava correlacionar as diferentes formas chagásicas crônicas com os diferentes clones de *T. cruzi* existentes na natureza (MACEDO; PENA, 1998).

Os estudos realizados então se concentraram na descrição da distribuição diferencial do *T. cruzi* e na avaliação patológica da doença em diversas condições experimentais e em humanos infectados. Também foram avaliadas as diferenças sistêmicas na produção de citocinas e de linfócitos, comparando-se diferentes cepas de *T. cruzi* em camundongos. Entretanto, a real extensão das diferenças de expressão gênica e das respostas do coração não era conhecida. Para tal fim, a utilização das novas técnicas de sequenciamento de RNA representa uma estratégia barata e eficiente para a realização do estudo.

1.4 Tecnologias de sequenciamento de nova geração e o estudo do transcriptoma

O estudo do transcriptoma, ou seja, o conjunto de transcritos expressos em uma determinada célula, tecido, órgão ou organismo tem sido realizado muito antes da existência da própria palavra, cunhada em 1997 (VELCULESCU; ZHANG; ZHOU;

VOGELSTEIN *et al.*, 1997). O método de sequenciamento por Sanger, predominante desde os anos 80, foi responsável por tornar possível o estudo da expressão gênica baseado na identificação de cDNAs a partir do sequenciamento parcial dessas moléculas, gerando as chamadas etiquetas de sequências expressas (ESTs) (ADAMS; KELLEY; GOCAYNE; DUBNICK *et al.*, 1991). A metodologia das ESTs foi então adaptada para ganhar maior *throughput* e permitir a quantificação das etiquetas, métodos denominados *Serial Analysis of Gene Expression (SAGE)* e *Cap Analysis of Gene Expression (CAGE)* (ADAMS, 1996). Ao mesmo tempo que surgiram as técnicas de SAGE/CAGE, as ESTs já depositadas em bancos de dados públicos forneceram a informação necessária para a criação dos primeiros *chips* de microarranjos, que dominaram o mercado e os estudos transcriptômicos praticamente por uma década (BARBULOVIC-NAD; LUCENTE; SUN; ZHANG *et al.*, 2006).

Recentemente, tem-se presenciado enorme progresso nos estudos transcriptômicos, devido ao surgimento das atuais tecnologias de sequenciamento em massa de RNA (RNA-Seq), principalmente o método de sequenciamento por síntese (Illumina, San Diego, CA). Em razão disso, a utilização de microarranjos tem ficado defasada (Figura 5) em muitas áreas da pesquisa básica (SU; FANG; HONG; SHI *et al.*, 2014). Além da sua alta capacidade de *throughput* e baixo custo, a grande vantagem do RNA-Seq está em detectar sequências não previamente conhecidas, permitindo a descoberta de um mundo novo de tipos de RNAs codificadores e não codificadores de proteína, a detecção de isoformas de transcritos expressas em baixas quantidades derivadas de *splicing* alternativo, mutações em genes, fusão de transcritos, dentre outros (LOWE, ROHAN; SHIRLEY, NEIL; BLEACKLEY, MARK; DOLAN, STEPHEN *et al.*, 2017; OZSOLAK; MILOS, 2011).

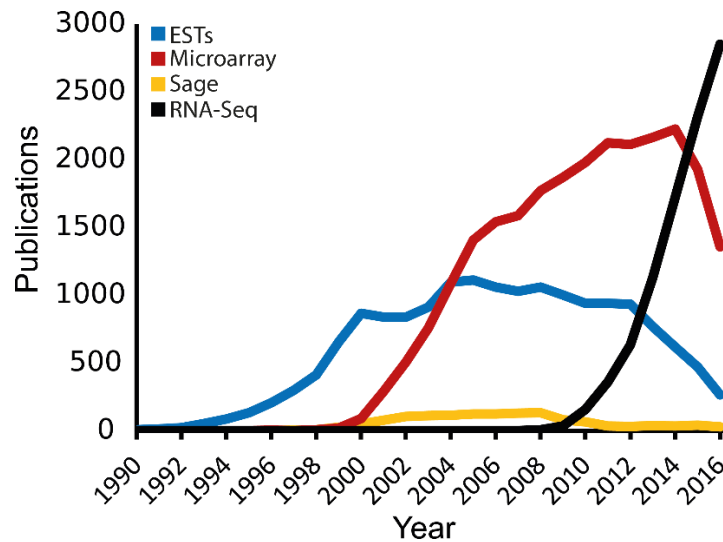


Figura 6: As tecnologias de estudo do transcriptoma no contexto histórico.

Adaptado de: (LOWE, R.; SHIRLEY, N.; BLEACKLEY, M.; DOLAN, S. *et al.*, 2017).

Para o sequenciamento por RNA-Seq existem estratégias para captura apenas das moléculas de RNA que possuem cauda de poliA, como os RNAs mensageiros, ou para a eliminação de moléculas de RNAs ribossomais, que são a grande maioria das moléculas presentes no *pool* de RNA total extraído. À vista disso, antes da fragmentação dos RNAs para construção das bibliotecas para sequenciamento, técnicas de enriquecimento são usadas para selecionar os RNAs de interesse, cada uma com suas próprias vantagens e vieses (LOWE, ROHAN; SHIRLEY, NEIL; BLEACKLEY, MARK; DOLAN, STEPHEN *et al.*, 2017; SHANKER; PAULSON; EDENBERG; PEAK *et al.*, 2015). Como geralmente a tecnologia de sequenciamento irá ler fragmentos menores que a sequência de RNA original, ela é fragmentada em tamanhos menores, geralmente entre 200 a ~300nt, sendo este um aspecto importante na construção da biblioteca. (Figura 7). A preparação para o sequenciamento envolve amplificações por PCR, com a adição de sequências *barcode* e adaptadores que serão usados na fixação na lâmina do DNA molde, e também durante a amplificação em ponte, no caso do sequenciamento por síntese da Illumina (BENNETT, 2004; BENTLEY; BALASUBRAMANIAN; SWERDLOW; SMITH *et al.*, 2008). Após o sequenciamento dos produtos de amplificação, a análise dos dados gerados depende de um grande poder computacional para o alinhamento das sequências no genoma ou transcriptoma de referência e, posteriormente, a quantificação dos transcritos, permitindo obter informação precisa sobre a expressão de cada gene detectado (GRIFFITH; WALKER; SPIES; AINSCOUGH *et al.*, 2015).

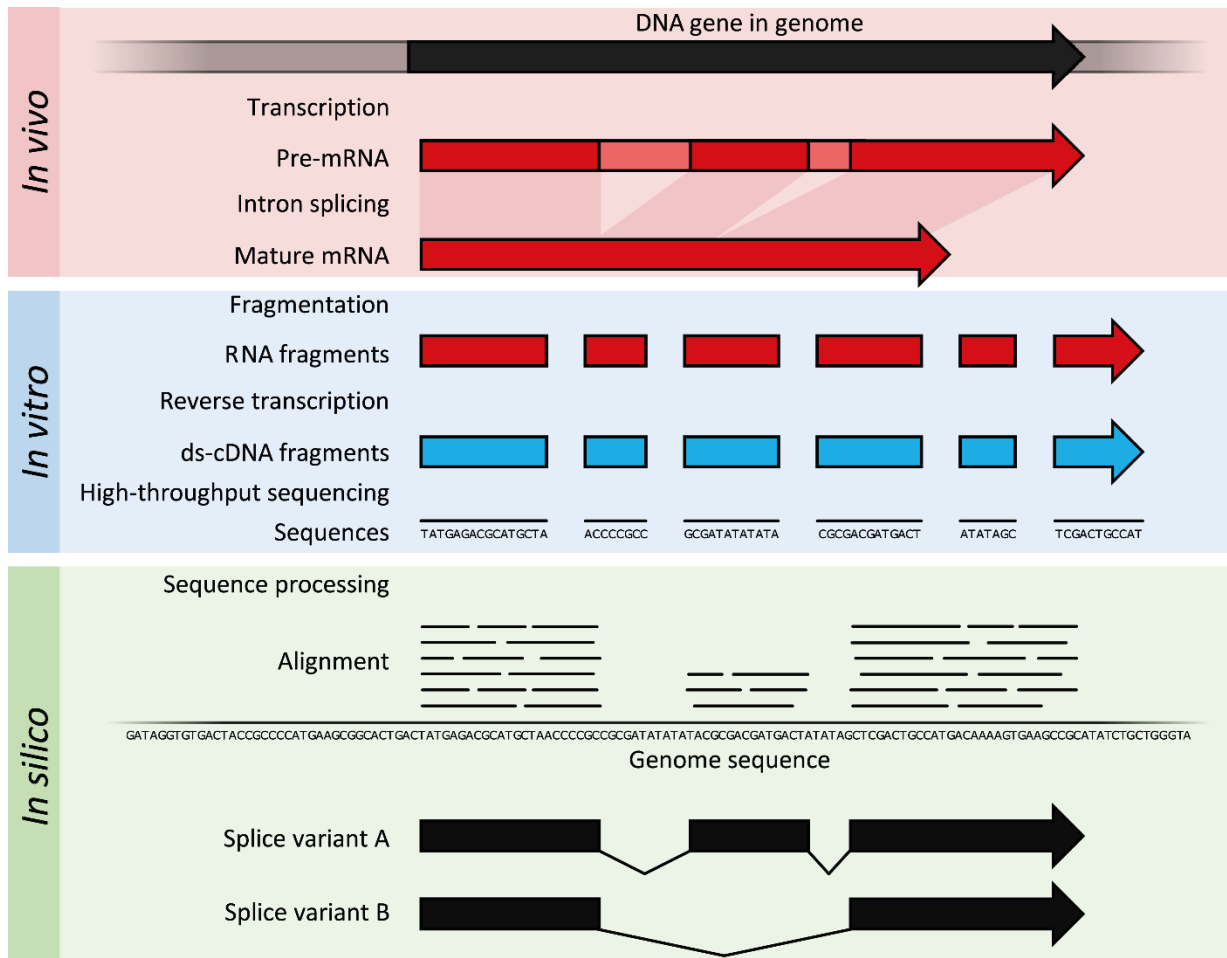


Figura 7: Resumo da construção de bibliotecas e do sequenciamento pelo método do RNA-Seq. Adaptado de: (LOWE, R.; SHIRLEY, N.; BLEACKLEY, M.; DOLAN, S. *et al.*, 2017).

1.5 O estudo do transcriptoma do *T. cruzi* e do seu hospedeiro mamífero infectado.

O microarranjo tem sido, até o momento, a principal ferramenta para o estudo do transcriptoma do hospedeiro infectado pelo *T. cruzi*. Foi proposto que o parasito é capaz de modular a maquinaria celular de cardiomiócitos para se replicar e persistir no hospedeiro (TANOWITZ; MUKHOPADHYAY; ASHTON; LISANTI *et al.*, 2011). Estudos com microarranjos mostraram a superexpressão de genes envolvidos com a ativação do sistema imune em cardiomiócitos (Manque *et al.*, 2011) e também a subexpressão de genes da via de fosforilação oxidativa mitocondrial em corações de camundongos (GARG; POPOV; PAPACONSTANTINOU, 2003; MANQUE; PROBST; PROBST; PEREIRA *et al.*, 2011). Vale ressaltar que o efeito da infecção por misturas de cepas não tem sido contemplado em nenhum estudo de expressão gênica do

hospedeiro, sendo que a maioria dos trabalhos que tratam do efeito no hospedeiro da infecção pelo parasito utilizou células em cultura ou camundongos não isogênicos como modelo da doença de Chagas experimental.

Além dos estudos do transcriptoma do hospedeiro infectado, o *T. cruzi* já teve sua expressão gênica avaliada por microarranjos em situações de estresse e durante as mudanças na progressão do seu ciclo de vida (GRYNBERG; PASSOS-SILVA; DE MOURÃO; JR *et al.*, 2012). Apenas recentemente o transcriptoma do *T. cruzi* tem começado a ser intensamente estudado utilizando-se do RNA-Seq (BERNA; CHIRIBAO; GREIF; RODRIGUEZ *et al.*, 2017; HOUSTON-LUDLAM; BELEW; EL-SAYED, 2016; LI; SHAH-SIMPSON; OKRAH; BELEW *et al.*, 2016; NAVARRO; FERREIRA; NAKAYA; BARON *et al.*, 2015). Contudo, estes estudos não são comparativos entre diferentes cepas de *T. cruzi*, que apresentam grande variabilidade em sua capacidade de infecção, replicação intracelular, performance de evasão do sistema imune e variabilidade genética (BELEW; JUNQUEIRA; RODRIGUES-LUIZ; VALENTE *et al.*, 2017; CARDOSO; REIS-CUNHA; BARTHOLOMEU, 2015; NARDY; FREIRE-DE-LIMA; MORROT, 2015; PABLOS; OSUNA, 2012; WAGNER; SO, 1990).

2. JUSTIFICATIVA

A notável capacidade do *T. cruzi* de causar manifestações clínicas diferentes em pacientes durante o curso da doença de Chagas tem sido reconhecida por pesquisadores desde os primórdios das pesquisas com o *T. cruzi*. Essa variabilidade tem sido relacionada com características genéticas do hospedeiro, do parasito, ou até mesmo com regiões cromossômicas específicas, como o complexo H-2 dos loci de MHC do camundongo. Além disso, tem sido hipotetizado que a ocorrência de infecções mistas poderia potencialmente contribuir para aumentar ainda mais a variabilidade patológica da doença, caso o comportamento das cepas em grupo fosse distinto daquele em infecções simples. Até o presente momento, entretanto, o efeito na expressão gênica de camundongos, causado por diferentes cepas de *T. cruzi* e por infecções mistas, não foi estudado com grande profundidade.

As novas tecnologias de sequenciamento do RNA permitem que se adquira um grande volume de informação de sequências derivadas do conjunto de moléculas de um tecido. Seu barateamento e eficiência, somados ao avanço computacional atual, possibilitam que o processamento de uma enorme quantidade de dados seja realizado rapidamente e com suporte estatístico sofisticado. A utilização dessas ferramentas apresenta uma ótima oportunidade para elucidar as diferenças de resposta do hospedeiro murino a duas cepas de *T. cruzi* e a uma mistura artificial de ambas, contribuindo, assim, para o avanço do entendimento da patologia da doença de Chagas.

3. OBJETIVOS

3.1 OBJETIVO GERAL

Investigar as alterações causadas por duas cepas de *T. cruzi* (Col1.7G2/TCI e JG/TCII) no modelo da doença de Chagas experimental em camundongos BALB/c durante a fase aguda da doença, utilizando-se de uma perspectiva transcriptômica.

3.2 OBJETIVOS ESPECÍFICOS

- Identificação, avaliação do perfil e classificação funcional dos genes diferencialmente expressos do coração e do reto de camundongos BALB/c durante a fase aguda da infecção com diferentes populações com *T. cruzi* (JG, Col1.7G2 e a mistura de ambas).
- Identificação de possíveis mecanismos e/ou vias bioquímicas que possam ajudar a explicar as diferentes “preferências teciduais” do *T. cruzi* no coração de camundongos BALB/c infectados com as diferentes cepas.

4. MATERIAIS E MÉTODOS

4.1 Animais Experimentais

Foram utilizados camundongos BALB/c machos, de 6-8 semanas, fornecidos pelo Centro de Bioterismo (CEBIO) da UFMG e mantidos em biotério aberto, onde receberam ração comercial e água *ad libitum* e ciclo de luz de 12 horas. A manipulação fez-se de acordo com as recomendações do Comitê de Ética em Experimentação Animal da UFMG (Protocolo Nº 64/2012).

Os grupos experimentais foram designados como animais não infectados (controle) e animais infectados (experimental), compreendendo cada grupo 5 animais. Três grupos foram infectados, por via intraperitoneal (ip), respectivamente com JG, Col 1.7G2 e uma mistura de ambas as cepas, recebendo 50 parasitos de cada. O grupo mistura recebeu 100 parasitos.

Após serem infectados, os animais foram mantidos em microisoladores, e a evolução da parasitemia foi acompanhada por um período de até 15 dias após a infecção, distinguindo-se o início da fase aguda em amostras de sangue periférico. Após o período de 15 dias, os animais foram anestesiados com uma mistura de xilazina (2,9mg/kg) e ketamina (14mg/kg) e sacrificados. O coração e o reto foram retirados, lavados com PBS estéril e imediatamente congelados em nitrogênio líquido, para impedir a degradação de quaisquer moléculas de mRNA. Por fim, foram armazenados em temperatura de -80°C para serem processados posteriormente (Figura 8).

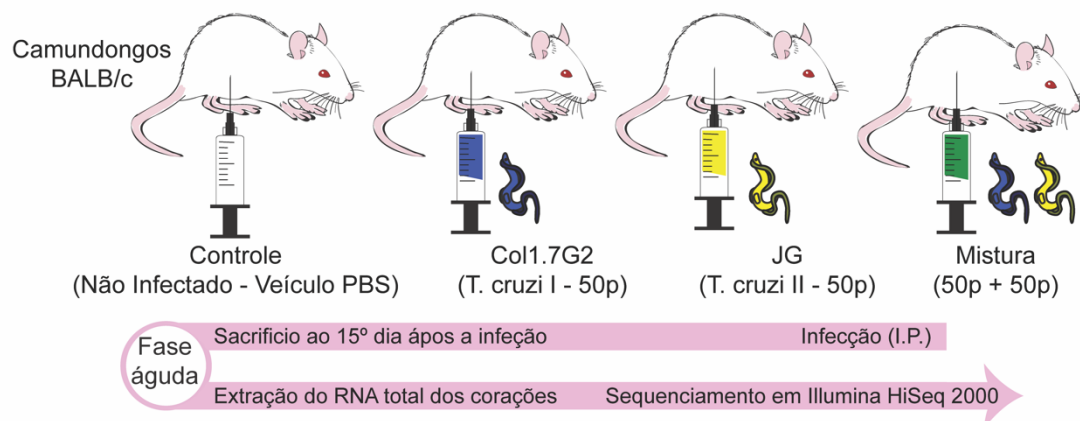


Figura 8: Design experimental. Camundongos BALB/c foram divididos em um grupo controle e três grupos experimentais, infectados com Col1.7G2, JG, ou uma mistura de quantidades equivalentes das duas cepas. Após 15 dias, durante o início da fase

aguda, os camundongos foram eutanasiados e os corações e retos extraídos para a extração do RNA total e posterior sequenciamento na Illumina HiSeq2000.

4.2 Aquisição e genotipagem dos parasitos

Populações de parasitos congeladas cedidas gentilmente pelo Laboratório de Biologia do *Trypanossoma cruzi*, liderado pelo Professor Egler Chiari, foram recuperadas e inoculadas em camundongos SWISS para a sua expansão e posterior utilização na infecção de camundongos BALB/c. O sangue desses animais foi coletado na fase aguda da infecção e teve o DNA extraído com o Kit DNeasy – Blood & Tissue Kit (QIAGEN). Para a genotipagem dos parasitos foi utilizada a amplificação do gene mitocondrial de *T. cruzi* COII, realizada segundo metodologia proposta por Freitas et al. (2006), utilizando dois iniciadores (Tcmit-10 e Tcmit-21). Após a amplificação, os amplicons foram digeridos com a enzima de restrição Alu I (Promega), de acordo com as indicações do fabricante. Como padrão de comparação dos RFLPs, foi utilizado DNA de cepas e clones característicos das linhagens de *T. cruzi* (clone Col.17G2, fragmentos: 30, 81, 264pb; haplótipo A) e *T. cruzi* II (cepa JG, fragmentos: 81 e 212pb; haplótipo C).

4.3 Obtenção do RNA total

O procedimento para realização da extração de RNA total dos corações e retos dos camundongos seguiu as recomendações e protocolo do TRIZOL (Invitrogen, Carlsbad, CA, USA). As amostras foram retiradas do freezer -80°C e imediatamente acondicionadas em nitrogênio líquido e moídas com a utilização de um cadinho de porcelana estéril, autoclavado e acondicionado em forno à 400°C por 4 horas, para a destruição de quaisquer enzimas que pudessem degradar os ácidos nucleicos. Durante todo o procedimento, as amostras foram mantidas congeladas, tendo sido a moagem feita com os tecidos imersos em nitrogênio líquido. O pó resultante foi colocado em tubo (Falcon) de 55 mL, ao qual foram adicionados 2 mL de TRIZOL, e incubado de 5 a 10 minutos à temperatura ambiente, para permitir a completa dissociação de complexos nucleoproteicos. Em seguida, foram adicionados 200 µL de clorofórmio (MERCK), para cada 1 mL de TRIZOL. Foi realizada a agitação em vortex, seguida da incubação por 3 minutos à temperatura ambiente. A amostra homogeneizada foi centrifugada a 12.000g, por 15 minutos, a 4°C. Após a

centrifugação, o homogenato foi separado em uma fase inferior (rosa), contendo material proteico e lipídico, uma fase intermediária, contendo DNA, e uma fase aquosa transparente superior, contendo o RNA. O RNA foi então transferido para um microtubo novo e foi precipitado com isopropanol (MERCK), na proporção de 500 µL de isopropanol para cada 1 mL de TRIZOL. As amostras foram incubadas por 10 minutos à temperatura ambiente e então centrifugadas a 12.000g, por 10 minutos, a 4°C. Após descarte do sobrenadante, foi adicionado 1 mL de etanol 75% (diluído em água tratada com DEPC) para cada mL de TRIZOL. O tubo foi agitado no vortex, para o desprendimento do precipitado. Então foi feita nova centrifugação a 7.500g, por 5 minutos, a 4°C. Após descarte do sobrenadante, o tubo foi deixado aberto por 5 minutos para secagem do sedimento que foi, em seguida, diluído em água milli-Q estéril, em 50 µL.

4.3.1 Controle de qualidade

A quantificação das amostras de RNA foi feita no espectrofotômetro NanoDrop ND1000 (Nano Drop Technologies, Wilmington, DE, EUA). Primeiramente, foi feita a calibração do aparelho com água ultrapurificada (MilliQ, Millipore) e, em seguida, foi selecionada a opção de leitura de RNA já definida no aparelho. A leitura do branco também foi feita com água Milli-Q. O RNA foi quantificado pela absorção a 260 nm, utilizando-se 1 µL da amostra de RNA obtido. A concentração foi fornecida em ng/µL e a pureza do RNA foi avaliada pela correlação 260/280 (RNA/proteínas). O RNA foi estocado a -80°C.

Apenas amostras consideradas de alta ou boa qualidade foram encaminhadas para o sequenciamento, sendo enviadas três amostras por grupo. Do total de vinte amostras de RNA, as doze melhores foram escolhidas para serem testadas pelo equipamento Bioanalyzer (Agilent Technologies), que verifica a quantidade e integridade do RNA, qualificando-o por meio de um score RIN (*RNA Integrity Number*). Apenas amostras com RIN maior que 7 foram qualificadas para serem sequenciadas.

4.4 Sequenciamento das amostras

4.4.1 Sequenciamento de corações

Para o envio, três alíquotas de RNA total de coração de camundongo (correspondentes a três réplicas biológicas) de cada grupo, experimental ou controle, foram precipitadas, utilizando-se o kit de estabilização e recuperação em temperatura ambiente RNA STABLE (Biomatrica). O procedimento realizado segundo o protocolo do fabricante consistiu em adicionar de 10 a 15 µg de RNA contidos em até 50 µl de volume de água ao tubo, misturando ao dissecante e desidratando por 1 hora em SpeedVac. Os 12 tubos, contendo amostras de 4 grupos, foram enviados para a empresa do BGI (China, [https:// http://www.genomics.cn/](https://http://www.genomics.cn/)), pela qual a qualidade do RNA total foi novamente avaliada com o equipamento Agilent 2100 Bioanalyzer (Agilent Technologies).

4.4.1.1 Preparação da biblioteca de cDNA e sequenciamento no equipamento Illumina Hiseq 2000

A preparação da biblioteca de cDNA foi realizada com o Truseq kit (Illumina), que é iniciado a partir do RNA total extraído dos tecidos. Sucintamente, os mRNAs são capturados com oligo-dTs ligados a *beads* magnéticas que se complementam à cauda de poli-A do mRNA em duas rodadas de purificação. Na segunda rodada, também é feita a fragmentação dos mRNAs e a inserção de primers aleatórios de 6 nucleotídeos para a síntese do cDNA. Uma transcrição reversa com a enzima reverse transcriptase é então usada para sintetizar cDNA complementar ao RNA, e um processo é feito para remover o molde de RNA e substituir por um cDNA de fita dupla. Após essa etapa, é realizado um procedimento para transformar as extremidades rombudas em cegas e, em seguida, são ligados os adaptadores. Na sequência, os fragmentos com adaptadores são amplificados com uma PCR em ponte na *flow cell* para enriquecimento. Cada uma das réplicas biológicas foi sequenciada com a configuração *paired-end* de 90 pares de base e *non-stranded*.

4.4.2 Sequenciamento de retos

As amostras de RNA total foram extraídas dos retos dos mesmos camundongos cujos corações foram estudados. Cerca de 4 µg de RNA total tiveram

sua qualidade verificada pela forma descrita no tópico 4.3.1. As amostras foram armazenadas em gelo seco para o seu transporte até ao Instituto Nacional do Câncer (INCA-Rio de Janeiro/RJ). Em seguida, o RNA foi processado para sequenciamento em Illumina 2500 com a configuração *paired-end*, 101 pares de bases de tamanho e *stranded*.

4.4.2.1 Preparação da biblioteca de cDNA e sequenciamento no equipamento Illumina HiSeq 2500

Amostras de RNA total do reto foram processadas utilizando-se o kit *SureSelect Strand-Specific* (Agilent), que possui algumas diferenças em relação ao Truseq kit da Illumina para manter a *strandness* (conhecimento da fita de origem do mRNA). Após o isolamento de moléculas de mRNA, utilizando-se oligo(dt) *beads*, e a sua posterior fragmentação, a primeira fita de cDNA é feita com a presença de actinomicina D, que bloqueia a síntese de DNA dependente de DNA, mas não a dependente de RNA. Desse modo, impede-se a formação espúria de DNA complementar a partir de cópias de cDNA já sintetizadas. A especificidade da fita é realizada por meio da incorporação de dUTPs no cDNA da segunda fita, ao invés de dTTPs. A digestão de dUTPs pela uracil-DNA glycosylase (UGDase) impede que a fita seja amplificada pela PCR, conferindo-lhe a especificidade (Figura 9).

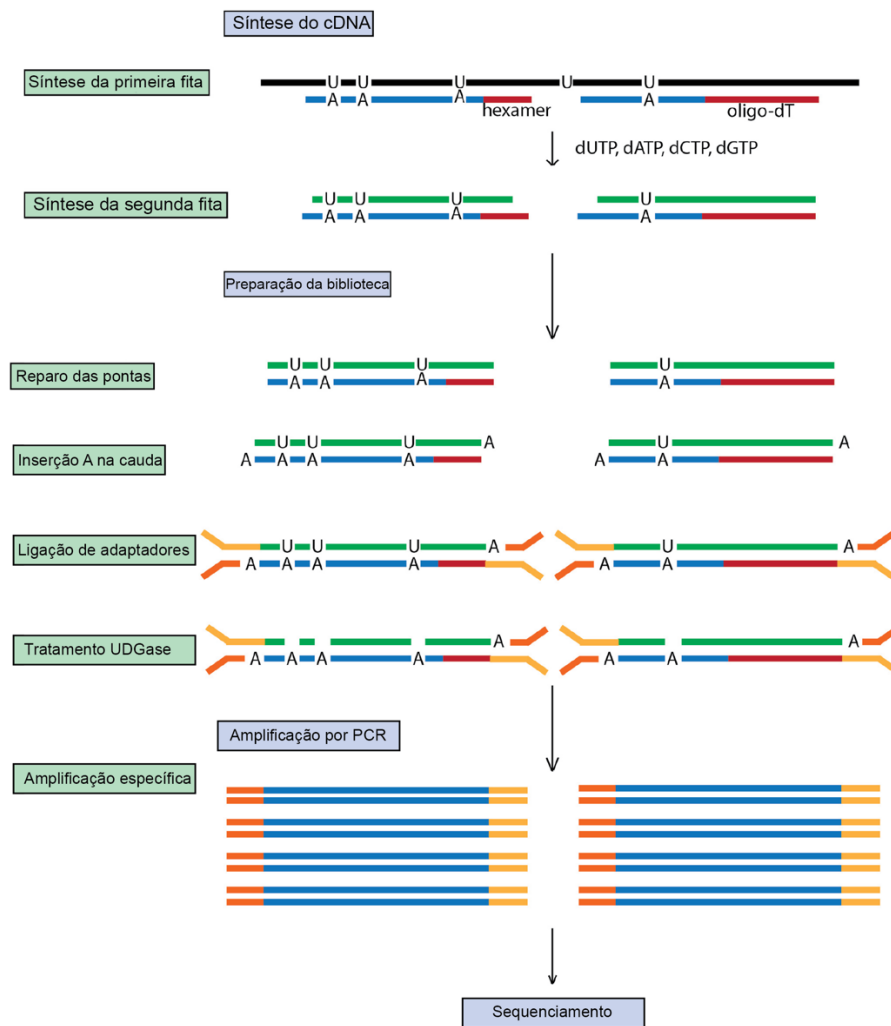


Figura 9: Criação de bibliotecas *stranded-specific*. Adaptado de (MARTIN; FEI; GIOVANNONI; ROSE, 2013).

4.5 Pipeline para análise dos dados de RNA-Seq gerados na plataforma Illumina Hiseq 2000/2500

4.5.1 Análise de qualidade

A qualidade total das leituras de sequências foi avaliada utilizando-se o programa FastQC (<https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>), que possui diferentes módulos de análises para o controle de qualidade das sequências brutas produzidas pelos equipamentos de sequenciamento de nova geração (NGS). Os gráficos gerados pelo programa permitem avaliar a qualidade geral das bibliotecas.

4.5.2 Quantificação da abundância de transcritos

A quantificação de transcritos foi feita através da técnica de pseudoalinhamento introduzida por Bray et al (2016). Para quantificar a abundância de leituras, foi utilizado o transcriptoma do camundongo mais recente até o momento (Gencode M21), que foi baseado no genoma GRCm38.p6, publicado pelo *Genome Reference Consortium Mouse* em 15/09/2017. O arquivo fasta, com todas as sequências de transcritos anotadas, foi usado de *input* para o programa Kallisto (v0.45) criar um índice contendo grafos De Bruijn de todos os k-mers encontrados no transcriptoma (BRAY; PIMENTEL; MELSTED; PACHTER, 2016). Com o intuito de realizar a quantificação de todas as amostras de maneira eficiente, foi empregado o programa GNU-Parallel para realizar múltiplos comandos em paralelo (TANGE, 2018). Primeiramente, uma lista contendo todos os nomes de arquivos de saída e arquivos fastq (*forward* e *reverse*) de replicatas técnicas de uma replicata biológica foi criada na mesma ordem (Figura 10).

```

1. #kallisto index -i [nome do arquivo de saída] [transcriptoma.fasta]
2. kallisto index -i kallisto_mouse.m21.index gencode.vM21.transcripts.fa
3. #comando cat imprime as linhas de texto do arquivo de entrada
4. cat lista1.txt #lista de nomes para output
5. output_amostra1
6. output_amostra2
7. cat lista2.txt #lista de nomes de input
8. amostra_1replicata_tecnica_1_forward.fastq
   amostra_1replicata_tecnica_1_reverse.fastq
   amostra_1replicata_tecnica_2_forward.fastq
   amostra_1replicata_tecnica_2_reverse.fastq
9. amostra_2replicata_tecnica_1_forward.fastq
   amostra_2replicata_tecnica_1_reverse.fastq
   amostra_2replicata_tecnica_2_forward.fastq
   amostra_2replicata_tecnica_2_reverse.fastq

```

Figura 10: Código para criação de um índice para o kallisto e as listas de *input* para kallisto quant.

Cada linha dessa lista foi usada como um *input* para o programa Parallel e substituída em uma chave de substituição. Em seguida, foi realizada a quantificação por meio do comando kallisto quant, embebido dentro de um comando GNU-Parallel, que disparou, em sequência, todos os comandos gerados usando a lista de arquivos fastq. A quantificação pelo kallisto foi executada por meio da opção *-b*, que determina

o número de amostragens de *bootstrap* realizadas pelo kallisto e fornece uma medida da acurácia de quantificação por reamostragem aleatória com reposição (Figura 11).

```

1. parallel kallisto quant \ #parallel e kallisto
2. -i kallisto mouse.m21.index \ #arquivo de indice
3. -b 20 \ # número de amostragens de bootstrap
4. --rf-stranded #opção apenas para amostras do intestino,
5. -o output/(ADAMS; KELLEY; GOCAYNE; DUBNICK et al.) {2} :::: lista1.txt
   :::: lista2.txt
6.
7. #Explicação dos Parâmetros:
8. #-o local e nome para os arquivos de saída
9. #{1} chave de substituição da lista 1
10. #{2} chave de substituição da lista 2
11. #::: 4 dois pontos, separa diferentes listas para gnuParallel

```

Figura 11: Código para execução de kallisto e quantificação da abundância de todas as amostras.

Após a quantificação, foram gerados três arquivos de saída por comando/amostra: Uma tabela em formato TSV (campos separados por tabulações), contendo, para cada isoforma gênica, o tamanho da isoforma, a estimativa do número de *reads* detectada para a isoforma e o valor de transcritos por quilobase milhão (*transcripts per kilobase million or TPM*). Esse valor de TPM é muito similar ao RPKM/FPKM, anteriormente muito utilizado como medida da abundância de leituras de DNA. Entretanto, para se calcular valores de TPM, é primeiro necessária a normalização pelo tamanho do gene e, em seguida, pela profundidade de sequenciamento, ao contrário do cálculo de RPKM, cuja ordem dos fatores é invertida. Como consequência, torna-se muito mais fácil, com valores de TPM, a comparação da proporção de *reads* que mapearam algum gene na amostra.

4.5.3 Normalização e análise

Como próximo passo, foi realizada a normalização entre as diferentes amostras, a fim de que os valores de abundância encontrados pudessem ser comparáveis. Para a execução dessa etapa, utilizou-se o pacote de análise baseado na linguagem R, Sleuth (v0.30), desenvolvido para facilitar o processamento estatístico de arquivos originados da quantificação pelo programa kallisto (PIMENTEL; BRAY; PUENTE; MELSTED *et al.*, 2017). Uma tabela, da qual cada linha se refere a uma única amostra, foi criada com as informações de metadados referentes às

variáveis do modelo experimental para a utilização no Sleuth. A análise foi efetuada ao nível de gene, ou seja, as abundâncias de diferentes isoformas foram agregadas a um único gene. A detecção de genes diferencialmente expressos entre as amostras foi feita então de duas formas, por meio do teste de razão de verossimilhança (TRV) e por meio do teste de Wald (Figura 12).

```

1. head(s2c)
2. ##      sample condition path
3. ## Sample_1 Controle ./results/Sample_1
4. ## Sample_2 Controle ./results/Sample_2
5. ## Sample_3 Controle ./results/Sample_3
6. ## Sample_4 Tratado ./results/Sample_4
7. ## Sample_5 Tratado ./results/Sample_5
8. ## Sample_6 Tratado ./results/Sample_6
9.
10. head(t2g)
11.
12.      ensembl_transcript_id_version      ensembl_gene_id external_gene_name
13.      ENSMUST00000082387.1 ENSMUSG00000064336          mt-Tf
14.      ENSMUST00000082388.1 ENSMUSG00000064337          mt-Rnr1
15.      ENSMUST00000082389.1 ENSMUSG00000064338          mt-Tv
16.      ENSMUST00000082390.1 ENSMUSG00000064339          mt-Rnr2
17.      ENSMUST00000082391.1 ENSMUSG00000064340          mt-Tl1
18.      ENSMUST00000082392.1 ENSMUSG00000064341          mt-Ndl
19.
20. so <- sleuth_prep(s2c, #tabela de metadados
21.                   target_mapping = t2g, #anotação (gene_names/ensembl IDs)
22.                   aggregation_column = 'ext_gene', #indica a fusão de isoformas
23.                   extra_bootstrap_summary = TRUE, #inclusão dos dados de bootstrap
24.                   read_bootstrap_tpm = TRUE, #leitura de bootstrap de tpm
25.                   max_bootstrap = 20, #número máximo de bootstrap
26.                   gene_mode=TRUE, #execução em nível de gene
27.                   transform_fun_counts= function(x) log2(x + 0.5)) #log2
28.
29. #Likelihood Ratio Test
30.
31. so <- sleuth_fit(so, ~Strain, 'full') #Modelo experimental
32. so <- sleuth_fit(so, ~1, 'reduced') #Modelo reduzido
33. so <- sleuth_fit(so, 'reduced', 'full') #likelihood ratio test
34. models(so)
35.
36. #Modelos disponiveis para teste usando o Wald Test. Intercept é o grupo não
   infectado.
37.
38. #[ full ]
39. #formula: ~Strain
40. #data modeled: obs_counts
41. #transform sync'ed: TRUE
42. #coefficients:
43. #      (Intercept)
44. #      StrainColl.7G2
45. #      StrainJG
46. #      StrainMixture
47.
48. #[ reduced ]
49. #formula: ~1
50. #data modeled: obs_counts
51. #transform sync'ed: TRUE
52. #coefficients:
53. #      (Intercept)
54.
55. #Wald Test

```

```

56. #Controle versus Col1.7G2
57. so <- sleuth_wt(so, StrainCol1.7G2)
58.
59. #Controle versus JG
60.
61. so <- sleuth_wt(so, StrainJG)
62.
63. #Controle versus Mistura
64.
65. so <- sleuth_wt(so, StrainMixture)
66.

```

Figura 12: Código em R para execução da análise de detecção de genes diferencialmente expressos usando o pacote Sleuth.

4.5.3.1 Teste de razão de verossimilhança

Em uma análise com o Sleuth, o TRV é feito usando-se dois modelos. O modelo completo ou *full*, que é o modelo de interesse dos testes desta pesquisa, e o modelo reduzido, que é a hipótese nula. Neste trabalho foram utilizados como modelo completo os grupos experimentais, a saber, Controle, Col1.7G2, JG e Mistura. O Sleuth estima para cada gene a razão das duas probabilidades e calcula um valor-q (*qvalue*) (valor-p ajustado para múltiplas comparações). Basicamente, se o modelo completo melhor predisser como os dados se comportam, o TRV gerará um pequeno valor-p. Vale ressaltar que o TRV não produz nenhum tipo de medida das diferenças entre as expressões dos genes, como faz, por exemplo, o *fold-change* (FC). Para ser obtida também uma medida de diferença de expressão entre pares de grupos experimentais realiza-se o teste de Wald.

4.5.3.2 Teste de Wald

O teste de Wald é derivado do TRV e permite que sejam feitas comparações par a par entre grupos. Hauck e Donner (1977) examinaram o desempenho do teste de Wald e alertam que o teste de wald pode se comportar de maneira errônea em alguns casos, não rejeitando a hipótese nula quando o coeficiente é significativo (HAUCK; DONNER, 1977). O teste de Wald foi introduzido no pacote Sleuth posteriormente por uma grande demanda dos usuários, para que os mesmos adquirissem valores de *fold change* entre pares de grupos experimentais. É recomendado que o TRV seja o teste padrão para avaliar a significância dos resultados.

4.5.4 Análises de categorias funcionais

A análise de enriquecimento de processos biológicos (PB) foi conduzida utilizando-se o pacote em R (`org.Mm.eg.db`), contendo as anotações completas do genoma do camundongo (CARLSON, 2017), e o pacote `topGO` (ALEXA; RAHNENFUHRER, 2016). Primeiramente, os genes detectados como significantes pelo TRV foram agrupados utilizando *k-means* na sua matriz de expressão. O valor de *k-means* foi escolhido pelo *Elbow Method*, o que levou à definição de quatro grupos ou *clusters* distintos. Os genes de cada *cluster* foram utilizados como entrada para o pacote `topGO`, que utiliza do teste exato de Fisher para a detecção de PBs enriquecidos em um determinado termo de ontologia gênica (GO). O teste exato de Fisher compara, ao acaso, o número esperado de genes ao número de genes significantes, para chegar a uma probabilidade de um termo de GO estar realmente enriquecido. Categorias de GO com valor-p menor que 0.01 foram considerados como significantes nas nossas análises subsequentes.

```

1.  if(!requireNamespace("BiocManager", quietly = TRUE))
2.    install.packages("BiocManager")
3.
4.  BiocManager::install("topGO")#instala topGO
5.
6.  BiocManager::install("org.Mm.eg.db")#instala a anotação
7.  library(topGO)
8.
9.  topGOterms <- function( fg.genes = NULL, #foreground genes
10.    bg.genes = NULL, #background genes
11.    organism = "Mouse",
12.    ontology.use = "BP", #Biological process
13.    stats.use = "fisher.elim", #Fisher
14.    algorithm.use = "weight01", #algoritmo de
    eliminação (elim) + weight
15.    topnodes.print=20, #top GO para visualizar
16.    num.char=100){ #número de caracteres máximo do
    termo de GO
17.
18.    if(is.null(fg.genes) | is.null(bg.genes)){
19.      stop("Error : Both gene lists are empty")
20.    }
21.
22.    require(topGO)
23.    if(organism == "Mouse"){
24.      mapping.use = "org.Mm.eg.db" #banco de dados para mouse
25.      library(org.Mm.eg.db)
26.    } else if (organism == "Human"){
27.      mapping.use = "org.Hs.eg.db" #banco de dados para human
28.      library(org.Hs.eg.db)
29.    } else {
30.      stop("Error : Organisms other than mouse and human are not
    supported")

```

```

31. }
32.
33. n = length(bg.genes)
34. geneList = integer(n)
35. names(geneList) = bg.genes
36. geneList[intersect(names(geneList), fg.genes)]=1
37. print(paste0("Total ", length(geneList), " genes. ", sum(geneList), "
genes in the foreground"))
38. geneList = factor(geneList)
39.
40. if(ontology.use %in% c("BP", "CC", "MF")){
41.   print(paste0("Using Ontology : ", ontology.use))
42. } else {
43.   stop("Error: Ontology not available. Should be one of BP, CC or
MF")
44. }
45.
46. # Fazer o objeto GO
47. GOdata <- new("topGOdata",
48.             description = "GOanalysis",
49.             ontology = ontology.use,
50.             allGenes = geneList,
51.             annot = annFUN.org,
52.             mapping = mapping.use,
53.             ID = "SYMBOL",
54.             nodeSize = 10)
55.
56. print(paste0("Using the ", stats.use, " statistic with the ",
algorithm.use, " algorithm"))
57. res.result <- runTest(GOdata, statistic = stats.use, algorithm =
algorithm.use)
58. to.return = list()
59. to.return$GOdata = GOdata
60. to.return$res.table <- GenTable(GOdata, pval = res.result, topNodes =
topnodes.print, numChar = num.char)
61.
62. #Para pegar os genes significantes em cada termo de GO
63. allGO = genesInTerm(GOdata)
64.
65. #Genes significantes
66. ANOTATION_DEx = lapply(allGO,function(x) x[x %in% fg.genes])
67. #Genes significantes ou não
68. ANOTATION_Ex = lapply(allGO,function(x) x[x %in% bg.genes])
69.
70. for(i in seq(1,length(ANOTATION_DEx))){
71.   ANOTATION_DEx[[i] <- bg.genes[match(ANOTATION_DEx[[i]], bg.genes)]
72. }
73. for(i in seq(1,length(ANOTATION_Ex))){
74.   ANOTATION_Ex[[i] <- bg.genes[match(ANOTATION_Ex[[i]], bg.genes)]
75. }
76. tableGO.genesDEx <- ANOTATION_DEx[to.return$res.table$GO.ID]
77. tableGO.genesEx <- ANOTATION_Ex[to.return$res.table$GO.ID]
78.
79. n.obs <- sapply(tableGO.genesDEx, length)
80. seq.max <- seq_len(max(n.obs))
81. data.frame1 <- t(sapply(tableGO.genesDEx, "[", i = seq.max))
82. data.frame1 <- t(data.frame1)
83.
84. #Salva a tabela com os genes significantes pra cada termo de GO
85. write.xlsx(data.frame1, "GenesperGO.xlsx", row.names = F)
86.

```

```

87. n.obs <- sapply(tableGO.genesEx, length)
88. seq.max <- seq_len(max(n.obs))
89. data.frame2 <- t(sapply(tableGO.genesEx, "[", i = seq.max))
90. data.frame2 <- t(data.frame2)
91.
92. #Salva a tabela com todos os genes detectados expressos para cada termo
    de GO
93. write.xlsx(data.frame2, "GenesperGO_all.xlsx", row.names = F)
94.
95. return(to.return)
96. }
97. library(openxlsx)
98.
99.
100.
101.
102. #Codigo para usar a função criada acima
103. GO.Results <- topGOterms(
104.   fg.genes = c(Cluster_1.genes), #lista de genes detectada no
    cluster_1
105.   bg.genes = sleuth_table$target_id, #lista com todos os genes
    detectados significantes ou não
106.   organism = "Mouse",
107.   ontology.use = "BP",
108.   stats.use = "fisher",
109.   algorithm.use = "elim",
110.   topnodes.print= 100,
111.   num.char = 100)
112.
113. write.xlsx(file = "GO.Results_Cluster_4.xlsx", x =
    GO.Results$res.table, row.names = FALSE)

```

Figura 13: Código em R para execução da análise de detecção de enriquecimento de termos de GO usando o pacote topGO.

4.5.5 Análises de redes de processos biológicos

As redes de processos biológicos enriquecidos foram construídas usando o *plug-in* ClueGO do programa Cytoscape (BINDEA; MLECNİK; HACKL; CHAROENTONG *et al.*, 2009; SHANNON; MARKIEL; OZIER; BALIGA *et al.*, 2003). Cada uma das três listas de genes diferencialmente expressos, derivadas dos testes de Wald (Controle x Col1.7G2, Controle x JG, Controle x mistura), foram independentemente usadas como entrada, juntamente com anotações de GO do banco de dados EBI-QuickGO. Valores de valor-p menores que 0.01 foram considerados significantes após a correção por Bonferroni.

4.5.6 Construção do diagrama de Venn

Para a criação de diagramas de Venn, utilizou-se o pacote Vennerable (Figura 14).

```

1. install_github("devtools")
2. library(devtools)
3. install_github("js229/Vennerable")
4. library(Vennerable)
5.
6. #Listas de genes diferencialmente expressos encontrados em cada amostra
   pelo teste de Wald
7. venn.cruzi <- list(Col = "Control_x_Col1.7G2", #Lista 1
8.                  JG= "Control_x_JG", #Lista 2
9.                  Mix= "Control_x_Mixture") #Lista 3
10.
11. venn.cruzi
12. venn.cruzi <- Venn(venn.cruzi)
13.
14. plot(venn.cruzi, doWeights = FALSE)

```

Figura 14 Código em R para criação de diagramas de Venn de três conjuntos, usando o pacote Vennerable.

4.5.7 Construção de gráficos de vulcão

Para a criação de gráficos de vulcão, foi criada uma função em R, usando-se o pacote ggplot2 para a automatização da customização e a criação das tabelas de genes (Figura 15).

```

1. vol <- function(DF=DF, User.Sig = User.Sig, met = "qval",
   Log2FC=Log2FC, Label.Sig = 10e-4, Label.f=2, Comp.1= "Group1", Comp.2=
   "Group2", save.excel=FALSE,
2. limit.x=c(10,10), limit.y=c(10,10),
3. breaks.x=c(-10,-5,-2,0,2,5,10),
4. ann1.x=5,
5. ann1.y=11,
6. ann2.x=-7,
7. ann2.y=11){
8.
9.   DF.sig <- DF
10.  DF.sig$super <- DF$b > Log2FC & DF$q < User.Sig
11.  DF.sig$sub <- DF$b < -Log2FC & DF$q < User.Sig
12.  DF.sig$threshold <- as.factor(abs(DF$b) > Log2FC & DF[[met]] <
   User.Sig)
13.  DF.sig <- DF.sig[DF.sig$threshold == TRUE,]
14.  out <- DF.sig[-log10(DF.sig[[met]]) > max(limit.y),]
15.
16.  if(nrow(out) > 0){
17.    out[[met]] <- as.numeric(sprintf("1e-%s",max(limit.y) ))

```

```

18.   } else()
19.
20.   volcano_plot <- ggplot(data=DF, aes_string(x= "b", y=sprintf("-
log10(%s)",met))) +
21.   geom_point(data=DF, size = 1.5, colour= "#63384b", alpha = 0.3) +
22.
23.   #Detecta outliers e marca todos no teto do plot
24.   geom_point(data=out[out$super == TRUE,], shape=17, size=2, colour=
"#cc0000", alpha = 1) +
25.   geom_text_repel(data = out[out$super == TRUE,], aes(label=out[out$super
== TRUE,]$target_id) ) +
26.   geom_point(data=out[out$sub == TRUE,], shape=17, size=2,
colour="#3399ff", alpha = 1) +
27.   geom_text_repel(data = out[out$sub ==T,], aes(label=out[out$sub ==
TRUE,]$target_id) ) +
28.
29.   #Plota todos os pontos no gráfico
30.   geom_point(data=DF[DF$b > Log2FC & DF[[met]] < User.Sig,], size=2,
colour="#cc0000") +
31.   geom_text_repel(data=DF[DF$b > Label.f & DF[[met]] < Label.Sig,],
aes(label=DF[DF$b > Label.f & DF[[met]] < Label.Sig,]$target_id),
hjust=-0.2, vjust=0.1) +
32.   geom_point(data=DF[DF$b < -Log2FC & DF[[met]] < User.Sig,], size=2,
colour="#3399ff") +
33.   geom_text_repel(data=DF[DF$b < -Label.f & DF[[met]] < Label.Sig,],
aes(label=DF[DF$b < -Label.f & DF[[met]] < Label.Sig,]$target_id),
hjust=-0.2, vjust=0.1) +
34.   xlab("Log2FC") +
35.   ylab("-log10 q-value") +
36.   ggtitle(sprintf("%s versus %s",Comp.1,Comp.2) )+
37.   scale_x_continuous(limit=limit.x, breaks=breaks.x) +
38.   scale_y_continuous(limit=limit.y) +
39.   theme_classic() +
40.   theme(axis.title.y = element_text(face="bold", size=16),
41.   axis.title.x = element_text(face="bold", size=16, colour="black"),
42.   axis.text = element_text(size=12),
43.   legend.title = element_blank() ,
44.   legend.text = element_text(size = 1) ) +
45.   theme(legend.position= "none") +
46.   annotate("text", label = sprintf("High in %s: %s", Comp.1,
sum(DF.sig$sub == TRUE) ), x = ann1.x, y = ann1.y, size = 4, colour
= "#3399ff") + #cor vermelha para upregulated
47.   annotate("text", label = sprintf("High in %s: %s", Comp.2,
sum(DF.sig$super == TRUE) ), x = ann2.x, y = ann2.y, size = 4, colour
= "#cc0000") #cor azul para downregulated
48.
49.   print(volcano_plot)
50.
51.
52.   assign(sprintf("%s_x_%s_sig", Comp.1,Comp.2),DF.sig,.GlobalEnv)
53.
54.   if(save.excel == TRUE){
55.   write.xlsx(DF.sig, file =
sprintf("%s_versus_%s_sig.xlsx",Comp.1,Comp.2) )
56.   write.xlsx(DF, file = sprintf("%s_versus_%s.xlsx",Comp.1,Comp.2) )
57.   print("Excel Table Saved")
58.   print("Finished")
59.   }else{
60.   print("Finished")
61.   }
62.

```

Figura 15: Código em R para criação de gráficos de dispersão em formato vulcão (*Volcano plot*) usando o pacote ggplot2.

4.5.8 Detecção de sequências de parasito *T. cruzi*

Para se conseguir a detecção de sequências de *T. cruzi* nas nossas amostras, primeiramente se realizou o alinhamento das sequências no genoma do esmeraldo-like (versão TriTrypDB-43), usando-se o alinhador BWA. Após o alinhamento, foi feita a conversão dos arquivos de alinhamento BAM novamente para arquivos fastq. Essas leituras filtradas foram então utilizadas no programa de montagem *denovo* Trinity (HAAS; PAPANICOLAOU; YASSOUR; GRABHERR *et al.*, 2013). Os transcritos montados foram então usados como entrada para o programa de alinhamento BLAST (ALTSCHUL; GISH; MILLER; MYERS *et al.*, 1990). Finalmente, apenas transcritos que tiveram seu melhor *score* de alinhamento em um gene de *T. cruzi* foram mantidos. A partir daí, o programa BWA foi usado novamente para alinhar as sequências, desta vez na referência criada pela montagem *denovo*.

4.5.7 Pipeline resumido

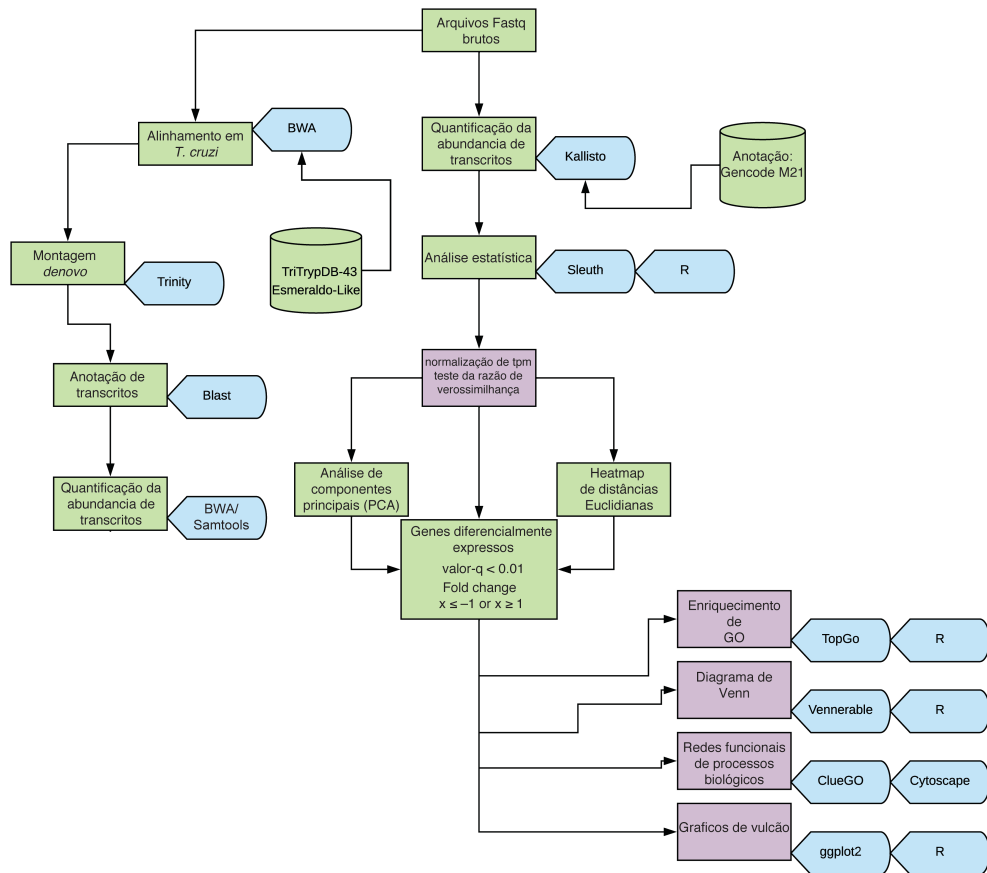


Figura 16: Design experimental completo

5. RESULTADOS

5.1 Infecções de camundongos BALB/c com JG e Col1.7G2

Para a execução deste estudo foram realizadas duas infecções em camundongos BALB/c, que foram acompanhadas até o 15º dia, quando se inicia a fase aguda, e até o 60º dia, marca do início da fase crônica, conforme explicado em detalhes na metodologia. Nas infecções que duraram até o 15º dia, não foram observadas alterações visíveis no comportamento ou aparência física dos animais. Por outro lado, nas infecções acompanhadas até o 60º dia, foi observado que, a partir do 20º dia, animais infectados com Col1.7G2 ou a com a mistura de JG e Col1.7G2 se apresentaram mais doentes, exibindo diarreia e prolapso retal, o que não foi observado em animais infectados apenas com JG.

5.1.1 Tipagem das cepas utilizadas nas infecções

Para verificar se as linhagens das populações de *T. cruzi* utilizadas nas infecções estavam corretas, foi utilizado o gene COII, como descrito por Freitas et al 2006 (Figura 16). Após amplificação e digestão com a enzima Alu I e um gel de acrilamida, confirmou-se que as populações descongeladas de *T. cruzi* oriundas do banco de cepas do laboratório de Biologia coordenado pelo professor Égler Chiari, denominadas JG e Col1.7G2, eram respectivamente, de acordo com o esperado, *T. cruzi* II e *T. cruzi* I (DE FREITAS; AUGUSTO-PINTO; PIMENTA; BASTOS-RODRIGUES *et al.*, 2006).

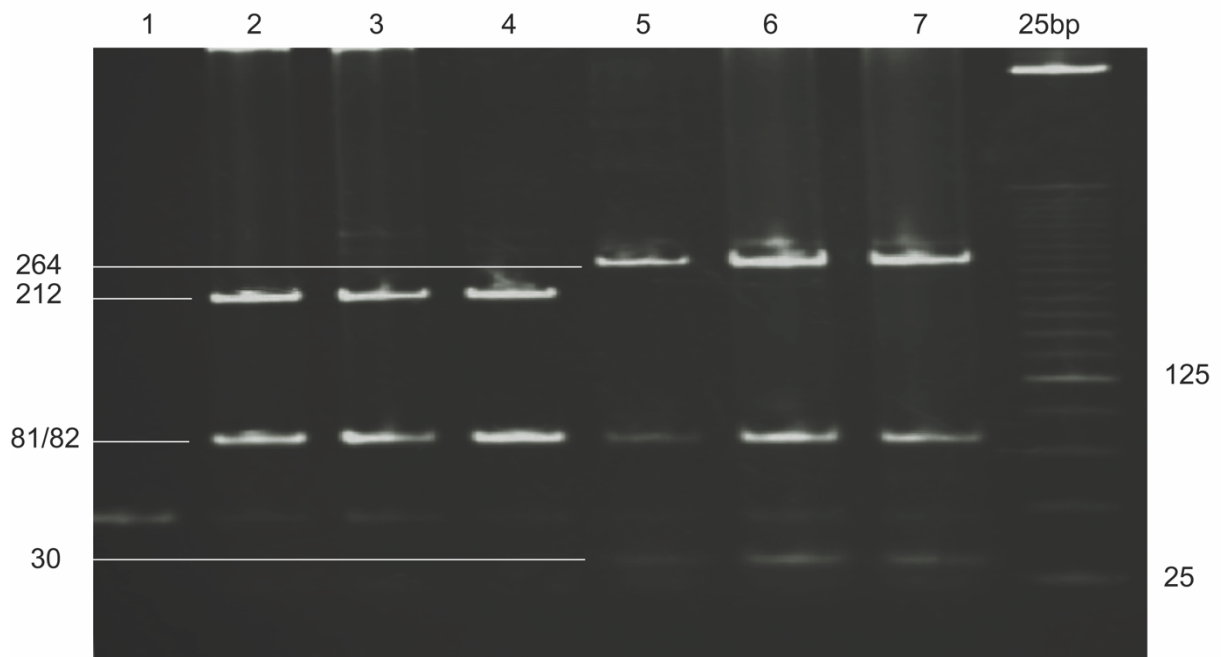


Figura 17: Genotipagem das cepas utilizadas nas infecções. Controle negativo da PCR (1). Parasitos tripomastigotas de JG utilizados nas infecções (2 e 3). Controle positivo contendo DNA da cepa JG (4). Controle positivo contendo DNA da cepa Col1.7G2 (5). Parasitos tripomastigotas de Col1.7G2 utilizados nas infecções (6 e 7).

5.1.2 Mortalidade e parasitemia dos camundongos infectados até 15 e 60 dias

Durante a avaliação da sobrevivência até o 60º dia dos animais infectados, pôde-se observar uma alta mortalidade entre os inoculados com Col1.7G2 e com a mistura de JG e Col1.7G2 (Figura 17A). Por outro lado, os animais infectados apenas com JG não morreram, nem aparentaram estar doentes, de forma semelhante aos animais controle não infectados. Surpreendentemente, foi constatado que animais infectados apenas com JG ou com a mistura apresentaram, na média, elevada parasitemia (10×10^4 /ml de sangue) em comparação aos animais infectados apenas com Col1.7G2 (Figura 17B), que apresentaram a metade do valor.

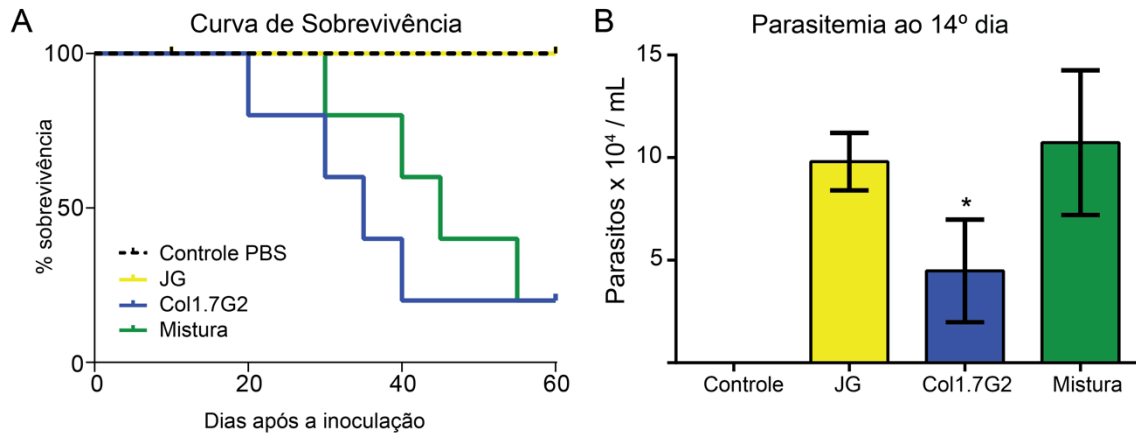


Figura 18: JG e Col1.7G2 diferem quanto a parasitemia e mortalidade de camundongos BALB/c na fase aguda. (A) Curva de sobrevivência de camundongos infectados pelas cepas JG e Col1.7G2 até o 60º dia. (B) Parasitemia dos camundongos no 14º dia após a infecção. Dados representam a média \pm SEM, $n=5$ camundongos/grupo * $p<0.05$ comparando o grupo Col1.7G2 com JG e mistura.

5.2 Estudo do transcriptoma de corações infectados por diferentes cepas de *T. cruzi*

5.2.1 Análise quantitativa e integridade do RNA total

As amostras de corações e retos congelados foram submetidas à extração de RNA total e foram avaliadas através de um gel de agarose desnaturante. As doze melhores amostras identificadas nesse momento preliminar foram selecionadas e submetidas a um controle de qualidade mais refinado por meio do equipamento Bioanalyser (Agilent Technologies), cujo resultado se encontra nas tabelas 1 e 2. Os valores de concentração, massa total, e razão da subunidade 28S e 18S foram convertidos para um valor de RIN (*RNA Integrity Number*), que representa um valor da integridade e qualidade do RNA. O RIN mínimo requerido para o sequenciamento (7) foi obtido em todas as amostras.

Tabela 1: Integridade do RNA total do coração enviado ao sequenciamento.

Número da Amostra	Grupo Experimental	RNA Concentração ng/μl	RNA Massa Total (μg)	Relação 28S/18S	RIN
3	Controle	675	15.5	0.7	7
4	Controle	480	13	1	7
5	Controle	380	10.5	1	7.3
6	Col1.7G2	590	16	1.1	7.3
8	Col1.7G2	560	16	1.1	7.7
9	Col17.G2	612	17	0.8	7.1
11	JG	460	13	1.4	7
13	JG	900	25.5	1.4	8.1
14	JG	1140	32	1.4	8.6
16	Mistura	550	15	1.5	8.2
17	Mistura	330	9.3	1.4	8
18	Mistura	870	23.5	1.7	8.5

Tabela 2: Integridade do RNA total do reto enviado ao sequenciamento.

Número da Amostra	Grupo Experimental	RNA Concentração ng/ μ l	RNA Massa Total (μ g)	RIN
3	Controle	300	4	8.2
4	Controle	800	4	8
5	Controle	600	4	8.1
6	Col1.7G2	1150	4	8
8	Col1.7G2	1200	4	7
9	Col17.G2	260	4	7.7
11	JG	450	4	7
13	JG	634	4	7.6
14	JG	700	4	7.6
16	Mistura	800	4	8.2
17	Mistura	950	4	8
18	Mistura	350	4	8.5

5.2.2 Controle de qualidade das bibliotecas

As bibliotecas foram avaliadas no programa FASTQC para determinar possíveis problemas e/ ou artefatos relacionados ao sequenciamento (Tabela 3). Pela tabela podemos observar um problema ocorrido com as amostras de intestino. Além da biblioteca estar com um número de leituras bastante variada e apresentar diversos *warnings* dados pro programa FASTQC, uma amostra do grupo infectado com JG foi perdida.

Tabela 3: Tamanho e qualidade das bibliotecas 1) *Per base sequence quality*, 2) *Per tile sequence quality*, 3) *Per sequence quality scores*, 4) *Per base sequence content*, 5) *Per sequence GC content*, 6) *Per base N content*, 7) *Sequence Length Distribution*, 8) *Sequence Duplication Levels*, 9) *Overrepresented sequences*, 10) *Adapter Content*

Nome da amostra	Grupos	Número de reads totais (Coração)	Problemas detectados pelo FASTQC	Nome da amostra	Número de reads totais (Reto)	Problemas detectados pelo FASTQC
3	Controle	12.282.839	4,5,8	3	12.253.584	1,2,4,5,7,8,9,10
4	Controle	12.471.492	4,5,8	4	11.456.657	1,2,4,5,7,8,9,10
5	Controle	12.499.734	4,5,8	5	15.773.099	1,2,4,5,7,8,9,10
6	Col1.7G2	12.512.968	4,5,8	6	11.459.357	1,2,4,5,7,8,9,10
8	Col1.7G2	12.395.128	4,5,8	7	9.817.450	1,2,4,5,7,8,9,10

9	Col1.7G2	12.512.255	4,5,8	9	6.141.103	1,2,4,5,7,8,9,10
11	JG	12.512.375	4,5,8	12	291	1,2,4,5,7,8,9,10
13	JG	12.052.513	4,5,8	13	9.233.810	1,2,4,5,7,8,9,10
14	JG	12.448.530	4,5,8	14	6.145.492	1,2,4,5,7,8,9,10
16	Mistura	12.469.466	4,5,8	16	10.937.186	1,2,4,5,7,8,9,10
17	Mistura	12.382.484	4,5,8	17	19.527.827	1,2,4,5,7,8,9,10
18	Mistura	12.448.580	4,5,8	18	12.678.836	1,2,4,5,7,8,9,10
Total:		148.988.364			125.424.692	

5.2.3 Perfil diferencial de infecção no coração por JG, Col1.7G2 e por mistura

Para acessar as mudanças de expressão gênica nos corações dos camundongos em resposta à infecção por Col1.7G2, JG e pela mistura dessas cepas, executou-se uma análise de componentes principais (PCA) na matriz de expressão das amostras. Foi observado que uma das amostras de Col1.7G2 apresentou alto nível de agrupamento com o grupo controle não infectado (Figura 18A), sugerindo um problema do *design* experimental, por exemplo uma falha ou atraso da infecção. Devido à alta correlação de expressão com os corações de camundongos não infectados, optamos pela remoção dessa amostra das análises posteriores. O perfil de expressão gênica geral capturado pela PCA mostrou que as amostras de JG, Col1.7G2 e a mistura apresentam significativa alteração com 15 dias de infecção, e essa alteração está representada pelo primeiro componente principal (PC1) (Figura 18B,C). Vale ressaltar que uma amostra de JG (amostra 11), apresentou uma menor expressão de genes do PC1, o que poderia indicar um atraso da infecção em relação às outras amostras de JG. Interessantemente, o segundo componente principal (PC2) capturou alterações únicas de cada uma das cepas (Figura 18B,D). Além disso, verificou-se que, como um grupo, a mistura apresentou-se em um estado intermediário entre Col1.7G2 e JG. Entretanto, certa variabilidade foi verificada entre as amostras, com algumas sendo mais parecidas com JG (amostra 17), outra possuindo um estado mais intermediário (amostra 16), e outra sendo mais parecida com Col1.7G2 (amostra 18). Não obstante, todas as amostras de mistura apresentam em menor ou maior nível a assinatura de expressão de ambas as cepas. As distâncias entre cada amostra foram capturadas pelo *heatmap* de distância euclidiana, que reforçou o perfil distinto de JG e Col1.7G2, ao mesmo tempo que demonstrou a semelhança da mistura entre ambas as cepas (Figura 18E).

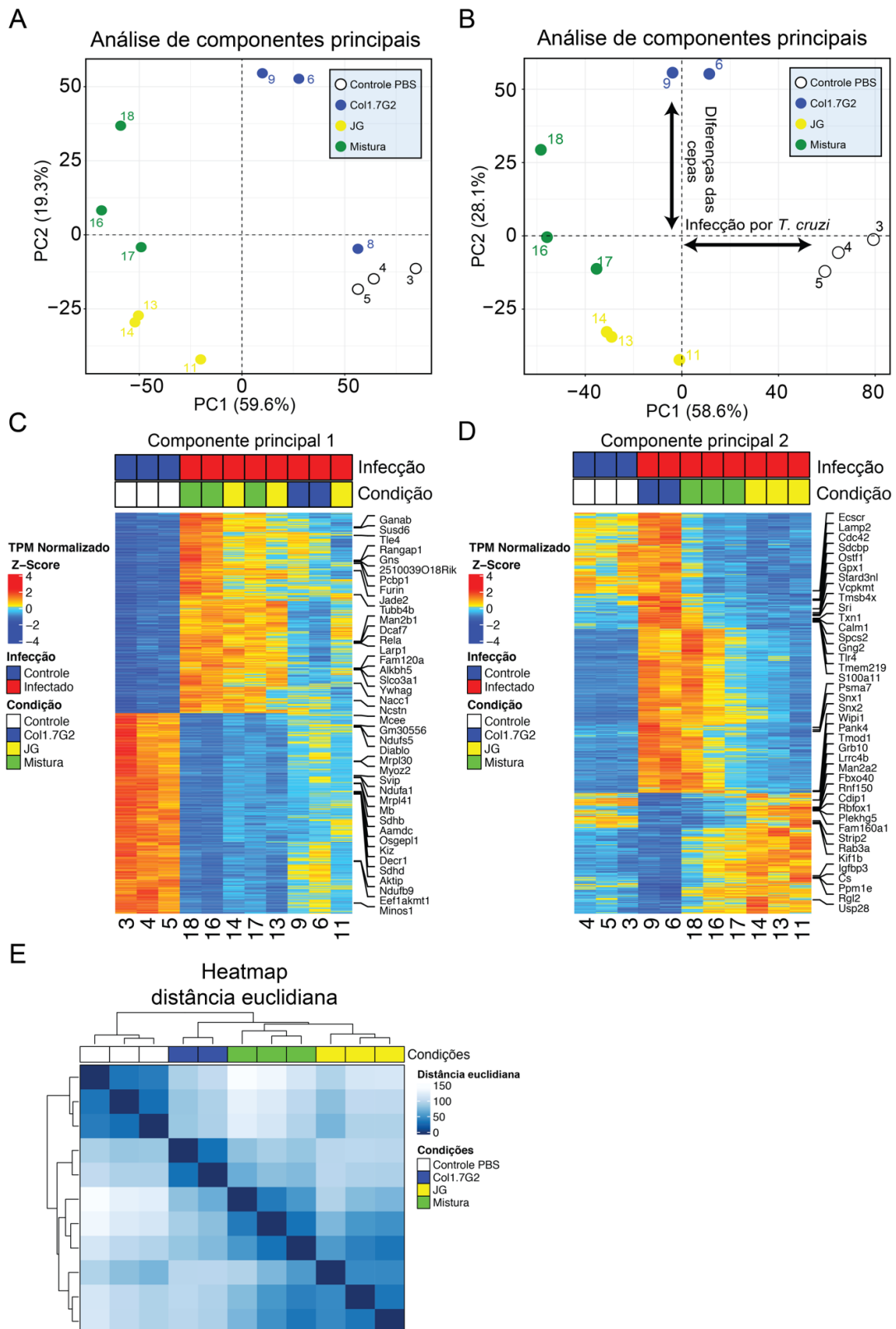


Figura 19: JG e Col1.7G2 apresentam perfil de expressão gênica distintos no coração (A) PCA com todas as bibliotecas, (B) PCA com as bibliotecas mantidas para análises posteriores. (C-D) *Heatmap* com os top 500 genes com autovalores positivos

e top 500 genes com autovalores negativos. Os top 20 genes mais significantes de cada direção do eixo são mostrados. (E) Distância euclidiana das amostras. Escalas de cor para o *heatmap* foram determinadas para cada gene como o z-score do TPM normalizado, sendo que o valor 0 representa a média.

5.2.5 Genes diferencialmente expressos pela infecção em relação ao controle não infectado

Com o objetivo de avaliar a mudança geral da expressão gênica de amostras infectadas por *T. cruzi*, realizaram-se comparações par a par por meio do teste de Wald entre grupos infectados e o grupo controle. Foram encontrados aproximadamente 16.400 genes únicos, expressos pelo coração dos camundongos em todos os grupos experimentais. A distribuição de genes diferencialmente expressos (GDEs) em cada grupo está representada por gráficos de vulcão (Figura 19A-C). É interessante notar as diferenças marcantes de GDEs entre os grupos. Col1.7G2 mostrou mais genes superexpressos (960 em Col1.7G2 e 216 em JG) e menos genes subexpressos (160 em Col1.7G2 e 508 em JG) em relação ao controle. Por sua vez, corações das infecções mistas mostraram tanto um alto número de genes subexpressos (849) quanto de genes superexpressos (1332), tendo em consideração o grupo controle. Em seguida, para capturar os genes que são compartilhados ou exclusivos de cada grupo, construiu-se um diagrama de Venn, com todos os 2.628 genes detectados das comparações par a par (Figura 19D). GDEs compartilhados por todos os grupos constituíram apenas 5% do total, compartilhados pelos infectados por Col1.7G2 e a mistura, 27,5%, e por JG e a mistura, 16,5%. JG e Col1.7G2 praticamente não apresentam GDEs compartilhados que também não são presentes com a mistura. Cabe destacar que 10,5% desses GDEs são exclusivos da infecção por Col1.7G2 e 6,3% por JG. Além disso, a mistura possui 34,5% de genes exclusivos, o que sugere uma singularidade da infecção mista. Para verificar esse fenômeno, decidiu-se avaliar a distribuição de valor-q (valor-p ajustado) dos genes exclusivos dos grupos infectados com Col1.7G2, JG e a mistura em grupos cujos mesmos genes não tivessem sido considerados GDEs. Como resultado, observou-se que o *qvalue* de GDEs exclusivos de Col.7G2 no grupo JG (CeJ-Col1.7G2 em JG) ou vice-versa (JeC-JG em Col1.7G2), apresentavam uma distribuição uniforme em todos os possíveis valores de q. Ao contrário, os GDEs exclusivos da mistura, tanto no grupo infectado por Col1.7G2 (MeC-Mistura em Col1.7G2) quanto por JG (MeJ-Mistura em JG) apresentavam grande concentração de GDEs no valor de corte para serem considerados significantes (Figura 19E).

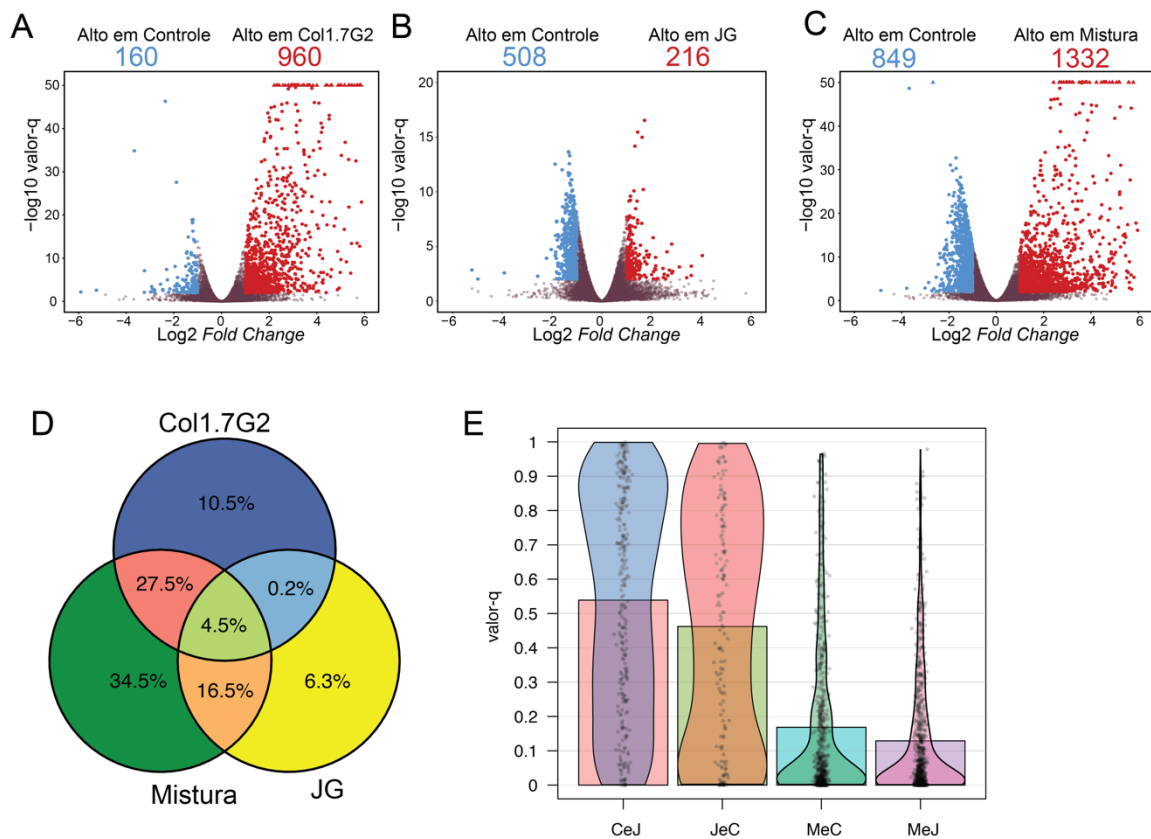


Figura 20: JG e Col1.7G2 modulam a expressão gênica no coração (A-C) Gráficos de dispersão do tipo vulcão. Genes em vermelho e azul representam genes super e subexpressos respectivamente. Genes com valores de *Fold Change* maiores que 1 ou menores que -1 e valores de *q* menores que 0.01 foram considerados como GDEs. (B) Diagrama de Venn. (E) Gráfico de dispersão do tipo violino mostrando a distribuição do valor *q* de genes GDEs de Col1.7G2 no grupo JG (CeJ), genes GDEs de JG no grupo Col1.7G2 (JeC) e genes GDEs de mistura no grupo Col1.7G2 (MeC) ou mistura no grupo JG (MeJ).

5.2.5 Diferenças globais de expressão gênica por Col1.7G2, JG e a mistura

Tencionando compreender as diferenças de expressão entre os grupos a partir de uma abordagem global, realizou-se a detecção de genes alterados significativamente por meio do teste da razão de verossimilhança. Em seguida, foi construída uma matriz com todos os genes que obtiveram um valor de q (valor-p ajustado) menor que 0,01, a fim de que os genes fossem agrupados pelo método de *k-means*. O valor de km utilizado foi obtido a partir do método do *Elbow* (Figura 20A), obtendo-se um valor ótimo de 4 *clusters*. Os *clusters* de genes obtidos foram representados por um *heatmap* (Figura 20B). O *cluster* 1 exibe genes que possuem expressão reduzida em todos os grupos experimentais, embora mais acentuada nos grupos JG e mistura. O *cluster* 2 é formado por genes que são fortemente subexpressos nos grupos infectados com JG e mistura, enquanto Col1.7G2 possui o mesmo nível de expressão desses genes em relação aos controles não infectados. Finalmente, os *clusters* 3 e 4 apresentam genes que o grupo infectado com a mistura compartilha com JG e Col1.7G2.

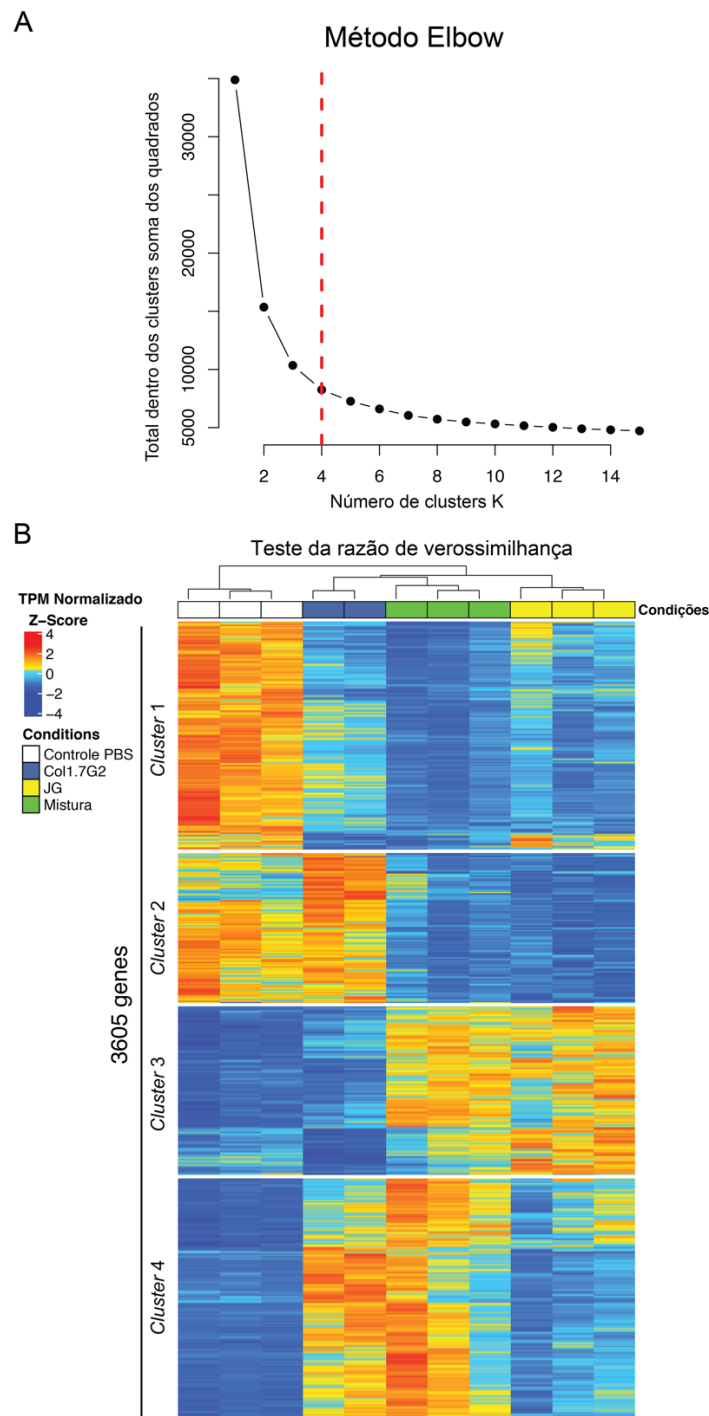


Figura 21: Teste da razão de verossimilhança e agrupamento de genes (A) Método de *Elbow* para determinar o número de km a ser usado no *k-means*. (B) *Heatmap* de genes detectados como significantes pelo teste da razão de verossimilhança e agrupados em 4 *clusters*.

5.2.6 Processos biológicos enriquecidos pela infecção por JG e Col1.7G2

Com o propósito de estudar os PBs presentes nos *clusters* de genes identificados, executou-se o teste de Fisher juntamente com o pacote topGO para a detecção de enriquecimento. Os dez PBs mais significantes para cada *cluster* foram representados em um gráfico de barras (Figura 21). É possível observar o alto número de processos enriquecidos relacionados a vias metabólicas que ocorrem na mitocôndria no *cluster 1*, tais como síntese de ATP acoplada ao transporte de prótons, com 14 genes significantes e 20 totais encontrados, (*ATP synthesis coupled proton transport*), complexo 1 da cadeia respiratória mitocondrial, 29/43, (*mitochondrial respiratory chain complex I*) e ciclo dos ácidos tricarboxílicos 12/26 (*tricarboxylic acid cycle*). Surpreendentemente, o *cluster 2* revelou grande número de genes, na sua maioria subunidades ribossomais e ubiquitinas que estão relacionadas à tradução e ao processamento proteico da célula. Os processos de tradução, 108/555, (*translation*), nedilação, 8/15, (*protein neddylation*) e processamento de RNA Ribossomal, 43/177, são exemplos dessas vias. É interessante observar no *cluster 3* a existência de PBs envolvidos com o funcionamento do músculo cardíaco, como por exemplo angiogênese, 44/424, (*angiogenesis*), diferenciação de células musculares, 43/315, (*muscle cell differentiation*) e regulação do ritmo do coração por condução cardíaca, 11/29, (*regulation of heart rate by cardiac conduction*). Finalmente, o *cluster 4* apresenta inúmeros PBs relacionados à resposta imunológica do hospedeiro à infecção pelo parasita *T. cruzi*. Genes de resposta de defesa a vírus, 75/182, (*defense response to virus*) provavelmente representam genes relacionados à resposta a patógenos intracelulares. Além disso, vias de processamento de antígeno e apresentação via MHC de classe I, 24/28, (*antigen processing and presentation via MHC class I*) e resposta imune inata, 200/537, (*innate immune response*) sugerem uma intensa resposta inflamatória em camundongos infectados por Col1.7G2.

Processos biológicos enriquecidos

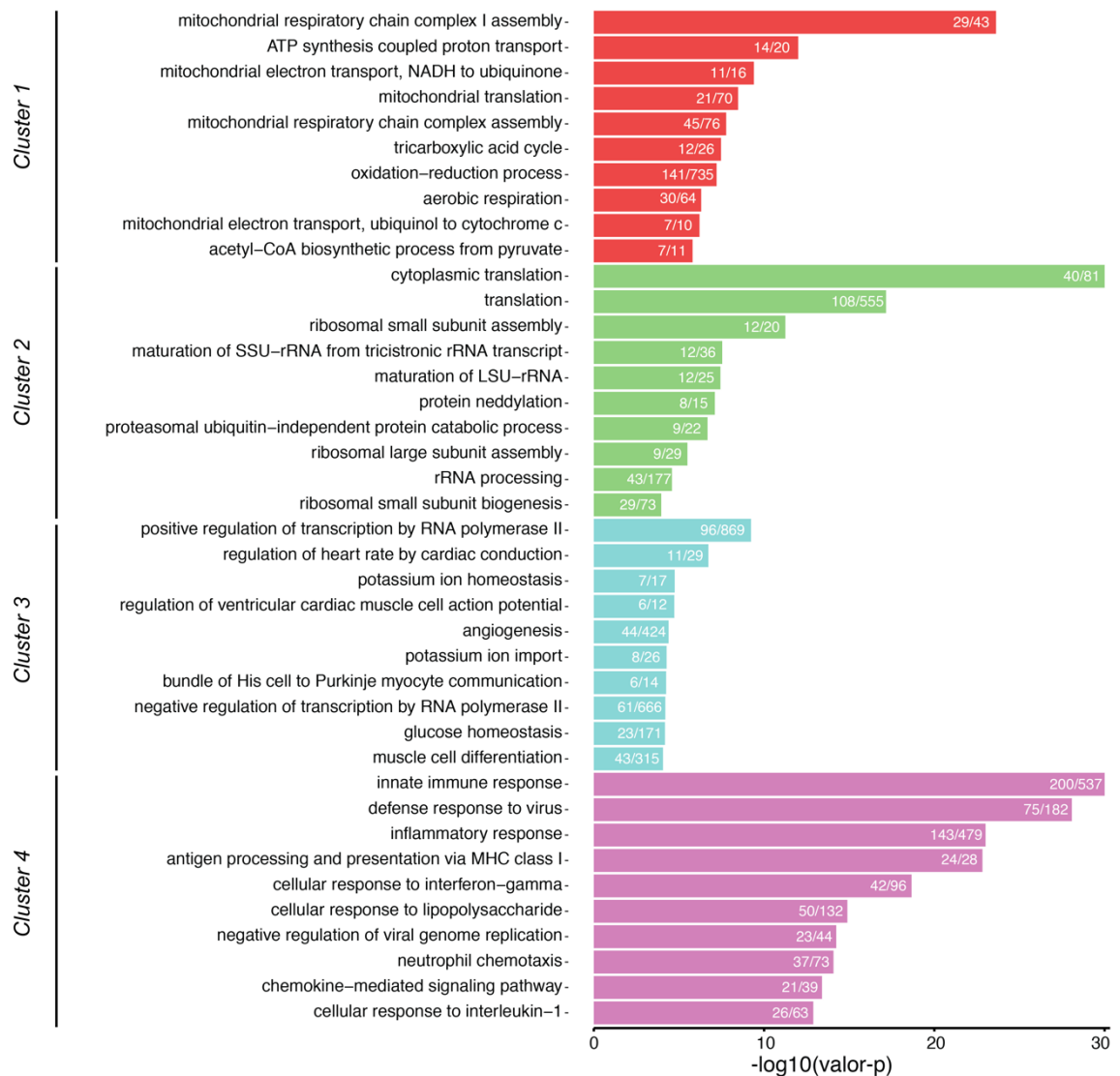


Figura 22: Análise de enriquecimento de processos biológicos. Os top 10 PBs mais significantes são representados neste gráfico de barras, separando-se os 4 *clusters* encontrados pelo *k-means*. Os valores das etiquetas representam o número de genes significantes no termo de GO sobre o número de genes totais detectados no termo.

5.2.7 Análise de redes de interação entre processos biológicos

Tencionando entender as interações entre os PBs encontrados nas listas de GDEs, utilizou-se o *plug-in* ClueGO do *software* Cytoscape para criação de redes. Pôde-se avaliar as relações existentes entre PBs em cada uma das listas de GDEs encontrados separadamente nas comparações par a par com o teste de Wald. É interessante observar a predominância de PBs relacionados com a resposta imune em animais infectados com Col1.7G2 (Figura 22A). A ativação de membros da família de proteíno-quinases por mitógenos (MAPK) é destacada na parte inferior da rede. As MAPK são reguladoras essenciais do sistema imune inato e adaptativo e participam de um importante papel na produção de citosinas inflamatórias e anti-inflamatórias. É possível observar a cadeia de eventos que vai desde migração de leucócitos (*Leukocyte migration*), resposta a protozoários (*Response to protozoan*), ativação de células B e T (*T/B cell activation*), endocitose (*Endocytosis*) até processamento e apresentação de antígenos (*Antigen processing and presentation*), demonstrando a intensa atividade inflamatória ativada por parasitos da cepa Col1.7G2. Por outro lado, esse perfil inflamatório não é observado para a cepa JG, visivelmente reduzido ou atrasado em relação a Col1.7G2 (Figura 22B). Contudo, os animais infectados com JG apresentaram genes subexpressos em vias de cadeia de transporte de elétrons (*Electron transport chain*) e de montagem do ribossomo (*Ribosome assembly*) devido ao alto número de subunidades ribossomais subexpressos nesse grupo. Surpreendentemente, a lista de GDEs do grupo infectado com a mistura apresenta ambos os fenótipos (Figura 22C), ou seja, alto número de PBs relacionados a ativação do sistema imune (perfil de Col1.7G2) e também PBs relacionados à cadeia transportadora de elétrons (perfil de JG).

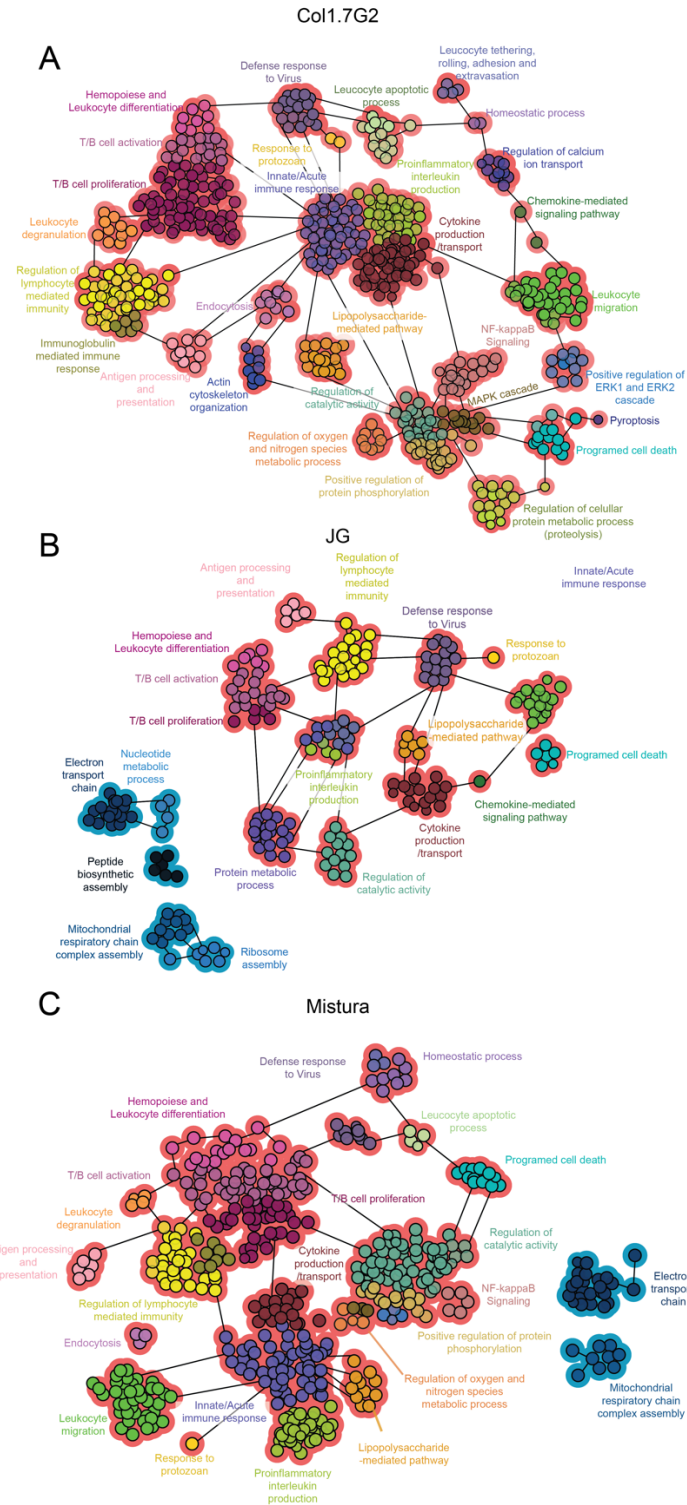


Figura 23: Rede de interação entre PBs significantes em Col1.7G2, JG e mistura
 Listas de GDEs detectadas pelo teste de Wald foram usadas no *software* Cytoscape/ClueGO para a detecção de conexões entre PBs enriquecidos em (A) Col1.7G2 (B) JG e (C) mistura.

5.2.8 Visualização de genes diferencialmente expressos em vias metabólicas do KEGG

A fim de que se visualizasse o perfil de expressão de genes dentro das vias e/ou PBs encontrados enriquecidos, foram criados *heatmaps* mesclando as listas de genes diferencialmente expressos pelo TRV e as listas de genes de vias do KEGG (Figura 23). Aqui é possível verificar a presença de uma modulação negativa generalizada em unidades ribossomais nos animais infectados com JG e a mistura (Figura 23A). Além disso, genes da via de fosforilação oxidativa estão subexpressos em todos os grupos, com uma maior proeminência nos animais infectados por JG e pela mistura (Figura 23B). É interessante notar que genes da via KEGG relacionados à doença de Chagas e à sinalização de quimosinas estão mais expressos nos grupos infectados por Col1.7G2 e pela mistura (Figura 23C-D).

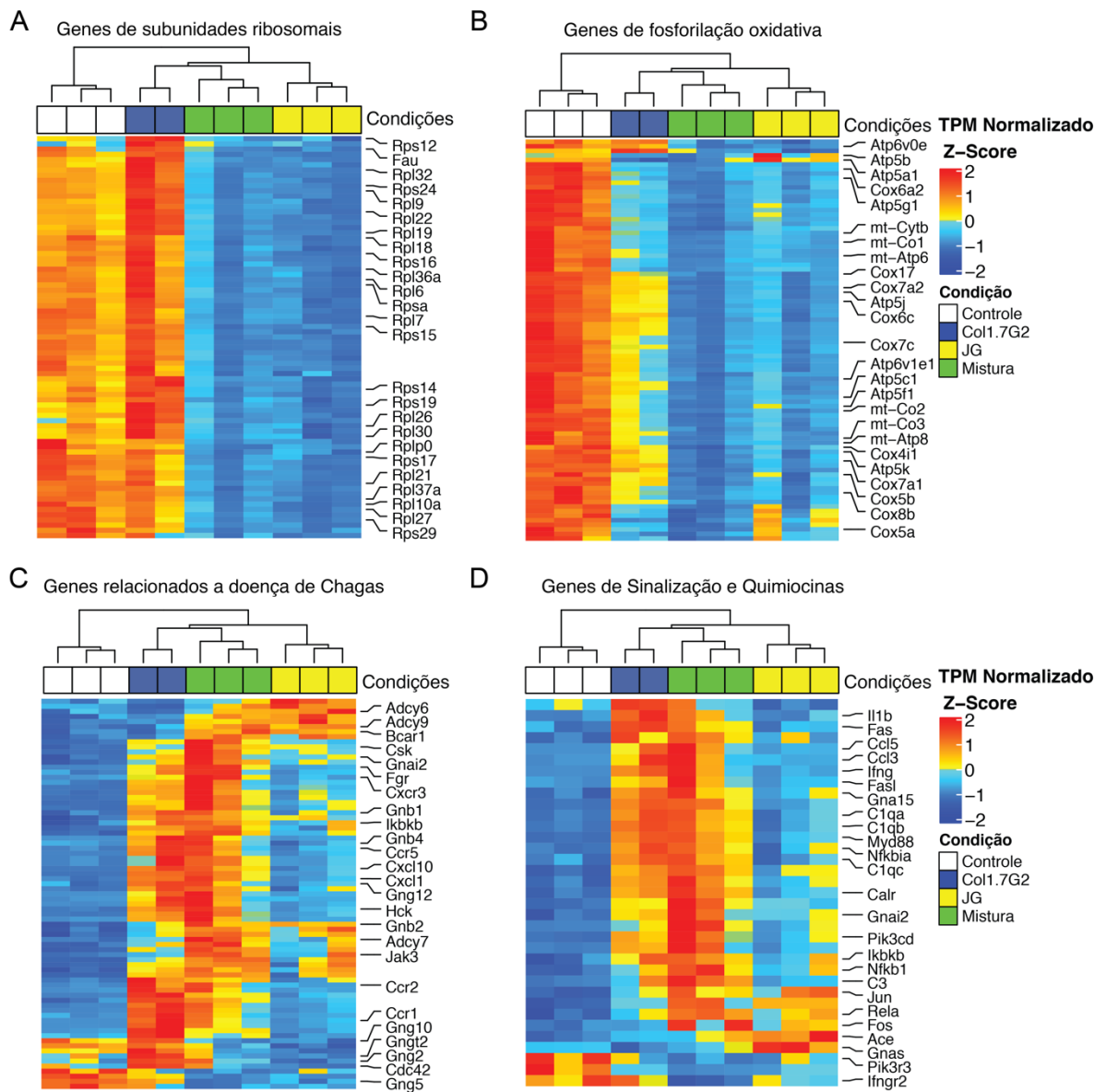


Figura 24: Heatmap de vias metabólicas e imunológicas do KEGG. (A) Genes de subunidades ribossomais (KEGG 03010), (B) da via de fosforilação oxidativa (KEGG 00190), (C) relacionados a infecções pelo *T. cruzi* (KEGG 05142), (D) via de sinalização de quimiocinas (KEGG 04062). Escalas de cor para o heatmap foram determinadas para cada gene como o z-score do TPM normalizado, sendo que o valor 0 representa a média.

5.3 Estudo do transcriptoma de retos infectados por diferentes cepas de *T. cruzi*

Servindo-se do mesmo método de análises de RNA-Seq usado para as amostras de corações infectados, realizou-se o estudo das de reto sequenciadas (Figura 24). Aqui observou-se que o PCA não foi capaz de corretamente separar as amostras de acordo com o design experimental. Além disso, a variância contida nos dois primeiros componentes principais é baixa, e indica que existe pouca diferença entre as amostras (Figura 24A-B). De fato, ao se realizar as análises estatísticas, não foi possível detectar GDEs em nenhuma comparação. Ao se verificar os top 100 genes mais expressos, pôde-se verificar que vários deles são genes tipicamente encontrados em alta atividade no intestino, tais como, IGHA (Imunoglobulina alpha), JCHAIN (cadeia J) e MUC (proteínas de superfície de células intestinais) (Figura 24C). A falta de padrão de expressão entre as amostras, somada ao fato de que as bibliotecas de cDNA tiveram vários problemas detectados pelo controle de qualidade, pode ser um indicativo de que algum problema técnico possa ter inviabilizado as análises. Uma outra alternativa, que não pode ser descartada, é o fato de que a extração de RNA total do órgão completo, como feita com o coração e o reto, possa ter consequências diferentes devido à natureza de cada órgão. O reto é formado por um variado e complexo conjunto de tipos celulares e células inflamatórias residentes, diferentemente do coração, que tem uma estrutura simples com poucos tipos celulares. Assim sendo, é possível que a melhor estratégia de estudo de RNA-seq no reto usando este modelo de doença de Chagas experimental seria por meio de uma separação de células infectadas ou de tipos celulares específicos para o enriquecimento de determinados processos por meio de técnicas de separação celular (FACS). Os dados obtidos mostram que existe uma pequena separação entre animais controle e experimentais (Figura 24A), sugerindo que alterações foram capturadas, entretanto, devido a um forte ruído, essas mudanças não puderam ser avaliadas.

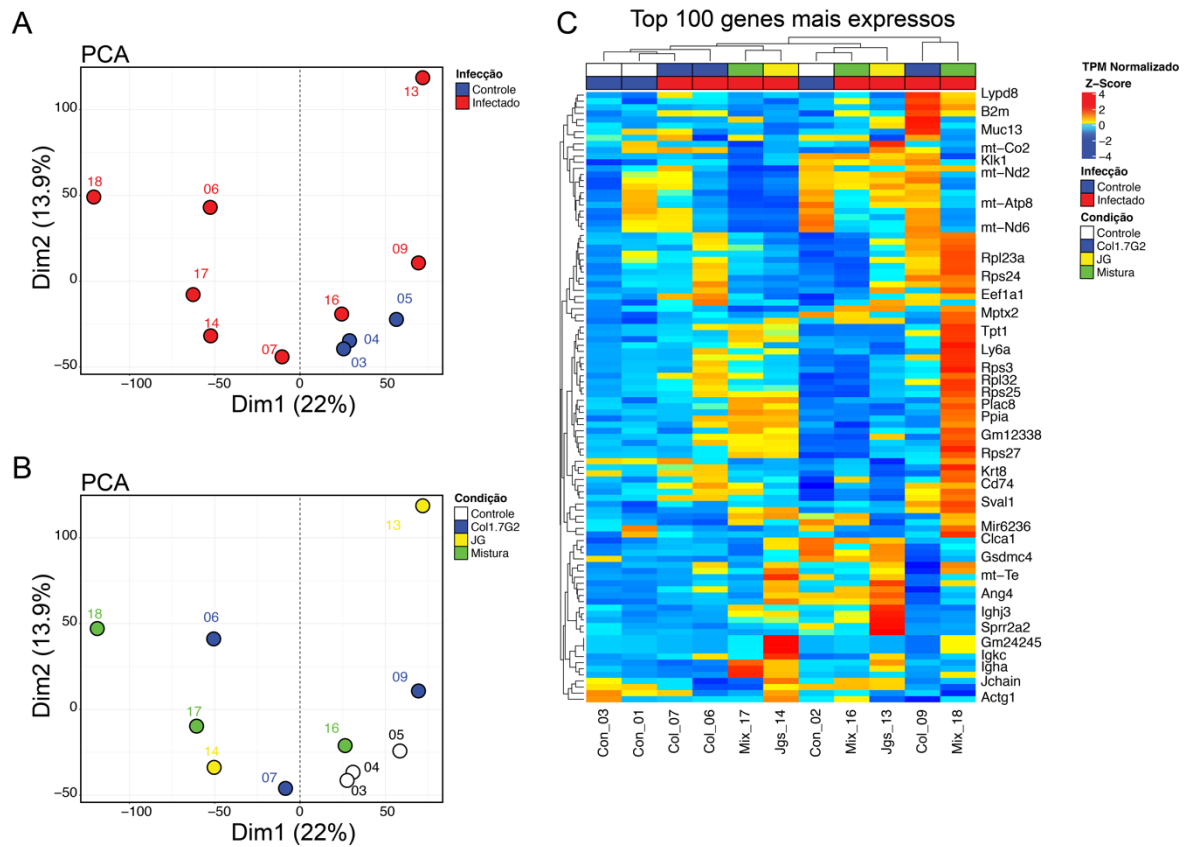


Figura 25: Análise do sequenciamento total do reto de camundongos infectados com JG, Col1.7G2 e a mistura. PCA mostrando a separação das amostras de acordo com a infecção (A) e o design experimental (B). Top 100 genes mais expressos nas amostras de acordo com a média (C).

6. DISCUSSÃO

Um dos aspectos mais críticos relacionados a cepas de *T. cruzi* distintas é a sua alta variabilidade genética e as diferenças da parasitemia ao infectar um hospedeiro mamífero. O nosso grupo de pesquisa, o Laboratório de Genética e Bioquímica, já havia estudado intensivamente o curso da infecção causada por Col1.7G2 (TCI) e JG (TCII) em diferentes linhagens de camundongo para a compreensão do tropismo tecidual (ANDRADE; GALVAO; MEIRELLES MDE; CHIARI *et al.*, 2010; ANDRADE; MACHADO; CHIARI; PENA *et al.*, 2002; ANDRADE; MACHADO; CHIARI; PENA *et al.*, 1999).

A cepa JG raramente causa morte em camundongos BALB/c, enquanto a cepa Col1.7G2 apresenta alta virulência, matando os animais nos períodos iniciais da infecção. Por outro lado, essa virulência observada não aparenta ser causada por uma alta parasitemia, já que os camundongos infectados por JG apresentam parasitemia mais elevada do que Col1.7G2 na fase aguda. Verificou-se também que os animais infectados com a mistura apresentaram tanto uma alta parasitemia quanto uma alta taxa de morte. Tal fenômeno foi também observado por Campos *et al.* em que animais BALB/c infectados por duas linhagens de *T. cruzi* I (AQ1-7 and MUTUM) apresentaram baixa parasitemia; ao serem, entretanto, co-infectadas com JG, um efeito sinérgico na parasitemia foi observado (SALES-CAMPOS; KAPPEL; ANDRADE; LIMA *et al.*, 2015). Sem embargo, Perez *et al.* mostraram que infecções mistas não aumentaram a mortalidade quando comparadas a infecções simples, apesar da alta parasitemia observada (PEREZ, C. J.; THOMPSON, R. C. A.; KEATLEY, S. K.; WALSH, A. L. *et al.*, 2018). Além disso, Rodrigues *et al.* observaram uma diminuição da mortalidade em animais infectados com a mistura de JG e CL Brenner quando comparada a infecções simples (RODRIGUES, C. M.; VALADARES, H. M.; FRANCISCO, A. F.; ARANTES, J. M. *et al.*, 2010). Esses estudos revelam a complexidade desses sistemas biológicos. Por exemplo, em infecções mistas, uma cepa pode se beneficiar de outra, potencializando sua capacidade de sobrevivência. Mas outros tipos de misturas podem levar a uma intensa competição entre as cepas e, em combinação do sistema imune do hospedeiro, levar a exclusão de uma cepa durante o curso da infecção. Não obstante, nossos dados sugerem que, durante as fases iniciais da infecção com o *T. cruzi*, ambas as cepas podem coexistir sem interferirem entre si.

Nos experimentos realizados, é improvável que muitas células, se alguma, estejam infectadas com ambas as cepas simultaneamente no momento da fase aguda, já que os camundongos são inoculados com uma carga parasitária baixa. Portanto, com base na parasitemia observada de JG e de Col1.7G2, sugere-se que JG prolifera a uma taxa mais alta que Col1.7G2. Esse aspecto da infecção pode explicar o motivo pelo qual JG predomina em infecções mistas com Col1.7G2 nos corações de camundongos BALB/c. Essa hipótese é corroborada por relatos anteriores mostrando que tanto JG como Col1.7G2 são detectados nos corações de camundongos BALB/c infectados com a mistura durante a fase aguda por PCR de baixa estrigência, entretanto, durante a fase crônica (3 a 6 meses após a infecção), encontra-se uma predominância de JG (ANDRADE; MACHADO; CHIARI; PENA *et al.*, 2002; ANDRADE; MACHADO; CHIARI; PENA *et al.*, 1999). Além disso, achados recentes demonstraram que Col1.7G2 replica menos eficientemente que JG em cardiomiócitos de camundongos BALB/c neonatos, o que reforça a hipótese aqui apresentada (DIAS; CAPILA; DO COUTO; ESTRADA *et al.*, 2017).

A razão pela qual ocorre preferência tecidual em diferentes cepas de *T. cruzi* ainda não é bem compreendida. Em humanos, tem sido sugerido que cepas distintas de *T. cruzi* podem prevalecer em diferentes órgãos e causar dano variável, culminando nos diferentes desfechos da fase crônica da doença (VAGO; ANDRADE; LEITE; D'AVILA REIS *et al.*, 2000; VAGO; MACEDO; OLIVEIRA; ANDRADE *et al.*, 1996). Portanto, diferentes formas da doença de Chagas seriam dependentes da cepa, do tropismo para determinados órgãos e da variabilidade genética do hospedeiro, um conceito que levou a criação do modelo histotrópico clonal para explicar a patogênese da doença de Chagas (MACEDO; PENA, 1998). As severas manifestações da fase crônica da doença de Chagas resultam da persistência do estágio intracelular do parasita (amastigotas) nos cardiomiócitos, células musculares, gânglios entéricos e tecido adiposo (BRENER, 1973; COMBS; MUKHERJEE; ALMEIDA; JELICKS *et al.*, 2005). Desse modo, ainda é controverso se os parasitos de *T. cruzi* exibem um tropismo intrínseco pelos cardiomiócitos devido a uma interação tecido-específica na superfície celular, ou se a proliferação diferencial e a seleção pelo sistema imune e/ou resistência a drogas irá afetar sua sobrevivência e persistência, levando à conclusão observada nas fases crônicas (PEREZ; LYMBERY; THOMPSON, 2014). Durante o estágio agudo da infecção, os parasitos de *T. cruzi*

podem infectar virtualmente qualquer tipo de célula nucleada, acarretando que o verdadeiro significado do conceito de tropismo seja questionado. Não se pode também descartar a possibilidade de que a preferência tecidual de cada cepa seja na verdade resultado da competição entre duas ou mais cepas de *T. cruzi*, como, por exemplo, uma determinação pela resistência de cada cepa e por sua adaptação ao sistema imune. Experimentos com ratos Holtzman verificaram a eficiência de JG em persistir no coração quando comparado com CL-Brenner (*T. cruzi* VI), tanto na fase aguda, quanto na crônica (FRANCO; VAGO; CHIARI; MEIRA *et al.*, 2003). Tem sido sugerido que o *T. cruzi* persiste preferencialmente em células com altas taxas de metabolismo de ácidos graxos, aquelas que podem fornecer o ambiente adequado e rico (CARADONNA; ENGEL; JACOBI; LEE *et al.*, 2013). Entretanto, também vale ressaltar que as células onde os parasitas de *T. cruzi* predominam na fase crônica, tais como adipócitos, cardiomiócitos, fibras esqueléticas e neurônios são as células com as menores taxas de renovação celular em mamíferos (RICHARDSON; ALLAN; LE, 2014), representando tecidos que podem abrigar parasitas latentes por longos períodos (SANCHEZ-VALDEZ; PADILLA; WANG; ORR *et al.*, 2018).

Pela razão de permanecer aberta a questão se infecções mistas levam à exacerbação dos efeitos patogênicos de cepas únicas (efeito sinérgico) ou se se trata apenas da soma dos efeitos de seus indivíduos únicos, (MAGALHÃES; VIANA; CHIARI; GALVÃO *et al.*, 2015; PEREZ, C. J.; THOMPSON, R. C.; KEATLEY, S. K.; WALSH, A. L. *et al.*, 2018; RODRIGUES, C.; VALADARES, H.; FRANCISCO, A.; ARANTES, J. *et al.*, 2010), avaliaram-se os GDEs compartilhados e exclusivos entre as amostras. Embora os dados tenham mostrado alta percentagem de genes exclusivos nas amostras infectadas com a mistura (próximo de 35%), uma avaliação mais a fundo da distribuição de valores-q conduziu à conclusão de que a mistura é muito mais similar aos grupos de infecções simples.

Para melhor compreensão das diferenças de vias metabólicas e processos biológicos que são modulados por JG e Col1.7G2, definiram-se quatro clusters de genes que foram encontrados por meio do TRV. Como já foi relatado previamente, observou-se uma modulação negativa de vias existentes na mitocôndria por todas as cepas. Entretanto, vale notar que em animais infectados por JG e a mistura essa modulação foi mais intensa (cluster 1). Esse fenômeno pode representar uma estratégia para aumentar a viabilidade da infecção, ocasionando uma adaptação do

parasito ao ambiente intracelular. As vias de tradução e processamento proteico foram drasticamente diminuídas tanto em JG como na mistura (cluster 2), mas nenhum efeito foi observado para Col1.7G2. Patógenos intracelulares são conhecidos por secretar proteínas efetoras capazes de modular e alterar vias metabólicas das células. Por exemplo a inibição de MAP Quinases e da via de NFkB são mecanismos que bactérias patogênicas podem explorar para aumentar suas chances de sobrevivência (CORNEJO; SCHLAERMANN; MUKHERJEE, 2017). Bactérias patogênicas possuem mecanismos capazes de inibir a tradução proteica para aumentar a disponibilidade de aminoácidos no citosol e aumentar a captação de nutrientes da célula (BEST; ABU KWAIK, 2019). Em vista disso, os dados obtidos sugerem que um mecanismo único de inibição da tradução proteica também pode estar presente em algumas cepas de *T. cruzi*, e pode ser uma importante estratégia capaz de aumentar a persistência do parasito. Ademais, a cepa JG e a mistura, mas não a cepa Col1.7G2, também foram capazes de induzir a expressão de PBs relacionados à angiogênese, diferenciação de células musculares, genes de canais iônicos e de regulação do ritmo cardíaco (cluster 3). Esse efeito revela a expressão de genes do tecido em resposta à infecção pela cepa JG e pode ser gerado pelo estresse de células cardíacas infectadas.

Os resultados sugerem que JG promove uma redução drástica no metabolismo oxidativo em cardiomiócitos infectados (cluster 1). Análises de microarranjos anteriores mostraram essa mesma redução de genes da via de fosforilação oxidativa em cardiomiócitos infectados por cepas de *T. cruzi* distintas (GARG; POPOV; PAPACONSTANTINOU, 2003; MANQUE; PROBST; PROBST; PEREIRA *et al.*, 2011). A modulação negativa de genes da cadeia transportadora de elétrons poderia aumentar ou diminuir as espécies reativas de oxigênio (*Reactive Oxygen Species* - ROS) no citosol, dependendo de diversos fatores, como a força próton-motriz e a razão de NADH/NAD⁺ e CoQH₂/CoQ, além da concentração de oxigênio (MURPHY, 2009). Todavia, estudos anteriores mostraram que ROS é uma faca de dois gumes para o *T. cruzi*, ou seja, agindo como uma molécula sinalizadora para a replicação celular em macrófagos se em baixa concentração (GOES; ROCHA; DINIZ; AGUIAR *et al.*, 2016; PAIVA; FEIJÓ; DUTRA; CARNEIRO *et al.*, 2012), mas causando danos letais em altas concentrações (AGUIAR; FURTADO; REPOLÊS; RIBEIRO *et al.*, 2013). Interessantemente, o tratamento em cultura com catalase, enzima que acelera a decomposição do peróxido de hidrogênio (H₂O₂), reduz a multiplicação de JG em

cardiomiócitos, mas não a de Col1.7G2, sugerindo que H₂O₂ age como molécula sinalizadora para levar JG a um crescimento acelerado, sendo que o mesmo não ocorre com Col1.7G2 (DIAS; CAPILA; DO COUTO; ESTRADA *et al.*, 2017). Notavelmente, o *T. cruzi* não possui o gene de catalase, o que pode ser relevante para o parasita perceber e superar o estresse oxidativo. Além do mais, parasitos transfectados com catalase mostraram maior resistência a H₂O₂ e sinais para diferenciação prejudicadas se comparados a animais (FREIRE; ALVES; GOES; RESENDE *et al.*, 2017). Por outro lado, a exposição excessiva a ROS leva a danos no DNA, como, por exemplo, acúmulo de 8-oxo-7,8-dihydroguanine (OG), que diminui o crescimento do parasito. Aliás, Campos *et al.* mostrou que as cepas JG e Esmeraldo (ambas *T. cruzi* II) são mais sensíveis ao dano oxidativo do DNA causada pelo tratamento com H₂O₂ quando comparadas a Col1.7G2 e Silvio (ambas *T. cruzi* I) (CAMPOS; SILVA; FURTADO; MACHADO-SILVA *et al.*, 2011). Eles também apontaram que, mesmo antes do tratamento com 200 µM de H₂O₂, JG apresentou mais lesões oxidativas no DNA que Col1.7G2, determinadas pela quantidade de OG presente no núcleo. O papel de OG na replicação do *T. cruzi* ainda é pouco compreendido, entretanto, estudos recentes mostraram que parasitas superexpressando MutT, que é responsável por remover OG do DNA, aumentaram a replicação dos parasitos da cepa CL Brener em fibroblastos significativamente quando comparados a parasitos WT (AGUIAR; FURTADO; REPOLÊS; RIBEIRO *et al.*, 2013). De fato, o produto de MutT, 8-oxo-dGMP, pode agir como um carreador de sinal de estresse, e lesões de OG no DNA podem agir como uma modificação epigenética, servindo como um sensor para o estresse oxidativo, consequentemente aumentando a expressão (FLEMING; DING; BURROWS, 2017). Desse modo, especula-se que cepas de *T. cruzi* que sejam mais sensíveis a ROS, como JG, possam ter sua proliferação aumentada dependendo do hospedeiro e do tecido envolvido durante a fase aguda, podendo, portanto, apresentar tropismo diferencial quando estudado em misturas de diferentes cepas.

O enriquecimento de PBs revelou também uma assinatura compatível com uma intensa resposta inflamatória contra os parasitos da cepa Col1.7G2 (cluster 4). Entretanto, camundongos infectados com JG também apresentaram assinatura similar, apesar de bem menos pronunciada. A superexpressão de genes relacionados à quimiotaxia de neutrófilos, apresentação de antígenos e resposta inflamatória em

um órgão como o coração, que normalmente não possui alto tráfico de células inflamatórias como o intestino possui, sugere que esses transcritos estão vindo de células recém-chegadas no coração para lidar com a infecção por *T. cruzi*. A expressão aumentada de citosinas como IFN- γ , IL-6, TNF e IL-12 em animais infectados por Col1.7 G2 está de acordo com resposta Th1 previamente estudada utilizando-se as cepas Y e Colombiana (ALIBERTI; CARDOSO; MARTINS; GAZZINELLI *et al.*, 1996). No entanto, animais infectados por JG exibiram uma expressão menor destes mesmos genes. Similar à expressão de citosinas CC e CXC, que são envolvidas no recrutamento de neutrófilos e macrófagos, os fatores de transcrição, como o transativador do complexo de histocompatibilidade principal do tipo II (*Class II Major Histocompatibility Complex Transactivator*-CIITA) e TBX21/TBET, clássicos marcadores de uma resposta Th1 potente, também apresentaram baixos níveis de expressão nos animais infectados apenas com JG em relação aos que foram infectados com Col1.7G2. É importante ressaltar ainda que camundongos que receberam a mistura destes parasitos exibiram os mesmos padrões de expressão que Col1.7G2. Esse padrão de expressão gênica sugere que Col1.7G2 é uma cepa de mais fácil reconhecimento pelo sistema imune nos momentos mais recentes da infecção, ativando uma resposta inflamatória mais exacerbada que a cepa JG.

A capacidade de evasão ao sistema imune de cada cepa pode explicar as diferenças de genes envolvidos com a resposta imune quando o hospedeiro é infectado com várias cepas diferentes de *T. cruzi*. Estudos anteriores demonstraram que células CD8⁺ citotóxicas específicas para a cepa Y de *T. cruzi* não são vistas na circulação sanguínea até o nono dia de infecção. Para confirmar que esse efeito era devido a imunossupressão ou evasão, os autores administraram agonistas de TLR-2 e TLR-9 para emular o reconhecimento prévio do sistema imune. Eles mostraram que camundongos desenvolveram células CD8⁺ específicas para o parasita precocemente em comparação ao grupo controle não tratado. (PADILLA; BUSTAMANTE; TARLETON, 2009; PADILLA; SIMPSON; TARLETON, 2009; TARLETON, 2007). Desse modo, a menor expressão de genes de resposta inflamatória em camundongos infectados com JG pode ser resultante de uma capacidade de evasão do sistema imune efetiva dessa cepa. Uma possível explicação para essa maior capacidade de está relacionada às proteínas de superfície do

parasita, tais como mucinas, proteínas associadas a mucinas (MASPs) e transialidases (TSs) que estão fortemente relacionadas à evasão do sistema imune, à adesão a membrana celular e à invasão celular (BARTHOLOMEU; CERQUEIRA; LEAO, 2009; EL-SAYED; MYLER; BARTHOLOMEU; NILSSON *et al.*, 2005; NARDY; FREIRE-DE-LIMA; MORROT, 2015; RUIZ DÍAZ; MUCCI; MEIRA; BOGLIOTTI *et al.*, 2015; YOSHIDA; DORTA; FERREIRA; OSHIRO *et al.*, 1997). Curiosamente, cepas do grupo *T. cruzi* I, como Col1.7G2, possuem menos genes codificando pra mucinas, MASPs e TSs se comparadas a *T. cruzi* II(PABLOS; OSUNA, 2012).

7. CONCLUSÕES

Em conjunto, este trabalho mostrou a habilidade de duas cepas de *T. cruzi* (JG e Col1.7G2) em modular a expressão gênica no tecido cardíaco de camundongos BALB/c durante as fases agudas da infecção. Mostrou-se, outrossim, como a mistura desses parasitos afeta a expressão gênica do hospedeiro. Foram descritos dois comportamentos distintos: Corações infectados com JG ou a mistura exibiram uma subexpressão de genes do metabolismo oxidativo e da tradução de proteínas quando comparados a controles não infectados. Por outro lado, corações de camundongos infectados com Col1.7G2 e a mistura mostraram forte ativação de genes do sistema imune inato e adaptativo quando comparados aos controles não infectados. Corroborando dados anteriores, esta pesquisa indica que a notável diferença na habilidade de diferentes tipos de cepas de *T. cruzi* em persistir em corações de camundongos BALB/c pode se dar por eventos que ocorram durante a fase aguda, a saber, as diferenças na velocidade de proliferação intracelular e na capacidade de ativar o sistema imune (Figura 25). O trabalho também sugere que JG pode ser beneficiado pela robusta redução do status energético das células, pelo aumento de ROS e pela evasão do sistema imune para se estabelecer de maneira mais eficiente nas células cardíacas. Por outro lado, Col1.7G2 pode ser mais capaz de infectar outros tecidos cuja competição com JG não é intensa. Os resultados obtidos podem ajudar a determinar quais parasitos sobreviverão durante a fase crônica da doença.

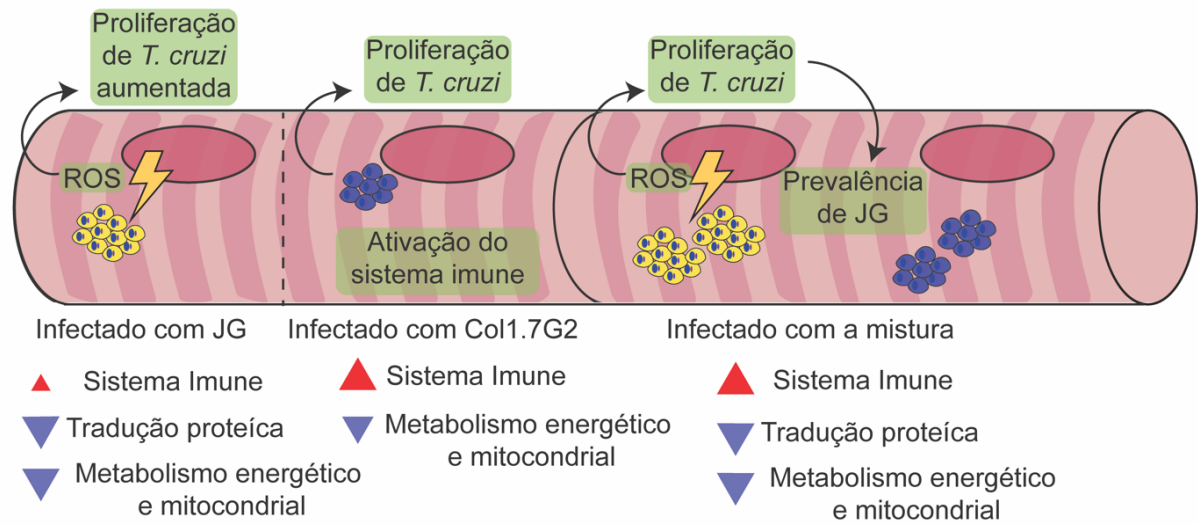


Figura 26: Modelo proposto para a distribuição diferencial de JG e Col1.7G2 em uma infecção mista.

8. REFERÊNCIAS BIBLIOGRÁFICAS

ADAMS, M. D. Serial analysis of gene expression: ESTs get smaller. **Bioessays**, 18, n. 4, p. 261–262, Apr 1996.

ADAMS, M. D.; KELLEY, J. M.; GOCAYNE, J. D.; DUBNICK, M. *et al.* Complementary DNA sequencing: expressed sequence tags and human genome project. **Science**, 252, n. 5013, p. 1651–1656, Jun 21 1991.

AGUIAR, P. H. N.; FURTADO, C.; REPOLÊS, B. M.; RIBEIRO, G. A. *et al.* Oxidative Stress and DNA Lesions: The Role of 8-Oxoguanine Lesions in Trypanosoma cruzi Cell Viability. **PLoS Neglected Tropical Diseases**, 7, n. 6, 2013.

ALEXA, A.; RAHNENFUHRER, J. topGO: Enrichment Analysis for Gene Ontology. **R package version 2.24.0.**, 2016.

ALIBERTI, J. C.; CARDOSO, M. A.; MARTINS, G. A.; GAZZINELLI, R. T. *et al.* Interleukin-12 mediates resistance to Trypanosoma cruzi in mice and is produced by murine macrophages in response to live trypomastigotes. **Infection and immunity**, 64, n. 6, p. 1961–1967, 1996.

ALTSCHUL, S. F.; GISH, W.; MILLER, W.; MYERS, E. W. *et al.* Basic local alignment search tool. **J Mol Biol**, 215, n. 3, p. 403–410, Oct 5 1990.

ANDRADE, L. O.; GALVAO, L. M.; MEIRELLES MDE, N.; CHIARI, E. *et al.* Differential tissue tropism of Trypanosoma cruzi strains: an in vitro study. **Mem Inst Oswaldo Cruz**, 105, n. 6, p. 834–837, Sep 2010.

ANDRADE, L. O.; MACHADO, C.; CHIARI, E.; PENA, S. *et al.* Trypanosoma cruzi: role of host genetic background in the differential tissue distribution of parasite clonal populations. **Experimental Parasitology**, 100, n. 4, p. 269–275, 2002.

ANDRADE, L. O.; MACHADO, C. R.; CHIARI, E.; PENA, S. D. *et al.* Differential tissue distribution of diverse clones of Trypanosoma cruzi in infected mice. **Molecular and biochemical parasitology**, 100, n. 2, p. 163–172, 1999.

ANDRADE, S. G.; CARVALHO, M. L.; FIGUEIRA, R. M. Caracterização Morfo-Biológica e Histopatológica de diferentes cepas do Trypanosoma cruzi. **Gazeta Médica Bahia**, Wed Apr 1 00:00:00 GMT–0300 1970 1970.

BAFICA, A.; SANTIAGO, H. C.; GOLDSZMID, R.; ROPERT, C. *et al.* Cutting edge: TLR9 and TLR2 signaling together account for MyD88-dependent control of parasitemia in Trypanosoma cruzi infection. **J Immunol**, 177, n. 6, p. 3515–3519, 2006.

BARBULOVIC-NAD, I.; LUCENTE, M.; SUN, Y.; ZHANG, M. *et al.* Bio-microarray fabrication techniques--a review. **Crit Rev Biotechnol**, 26, n. 4, p. 237–259, Oct–Dec 2006.

BARTHOLOMEU, D. C.; CERQUEIRA, G. C.; LEO, A. C. A. Genomic organization and expression profile of the mucin-associated surface protein (masp) family of the human pathogen Trypanosoma cruzi. **Nucleic acids ...**, 2009.

BELEW, A. T.; JUNQUEIRA, C.; RODRIGUES-LUIZ, G. F.; VALENTE, B. M. *et al.* Comparative transcriptome profiling of virulent and non-virulent *Trypanosoma cruzi* underlines the role of surface proteins during infection. **PLoS Pathog**, 13, n. 12, p. e1006767, Dec 2017.

BENNETT, S. Solexa Ltd. **Pharmacogenomics**, 5, n. 4, p. 433–438, Jun 2004.

BENTLEY, D. R.; BALASUBRAMANIAN, S.; SWERDLOW, H. P.; SMITH, G. P. *et al.* Accurate whole human genome sequencing using reversible terminator chemistry. **Nature**, 456, n. 7218, p. 53–59, 2008.

BERNA, L.; CHIRIBAO, M. L.; GREIF, G.; RODRIGUEZ, M. *et al.* Transcriptomic analysis reveals metabolic switches and surface remodeling as key processes for stage transition in *Trypanosoma cruzi*. **PeerJ**, 5, p. e3017, 2017.

BEST, A.; ABU KWAIK, Y. Nutrition and Bipartite Metabolism of Intracellular Pathogens. **Trends Microbiol**, 27, n. 6, p. 550–561, Jun 2019.

BINDEA, G.; MLECNIK, B.; HACKL, H.; CHAROENTONG, P. *et al.* ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. **Bioinformatics (Oxford, England)**, 25, n. 8, p. 1091–1093, 2009.

BRAY, N. L.; PIMENTEL, H.; MELSTED, P.; PACHTER, L. Near-optimal probabilistic RNA-seq quantification. **Nat Biotechnol**, 34, n. 5, p. 525–527, May 2016.

BRENER, Z. (COMPARATIVE STUDIES OF DIFFERENT STRAINS OF TRYPANOSOMA CRUZI.). **Annals of tropical medicine and parasitology**, 59, n. 1, p. 19–26, 1965.

BRENER, Z. Biology of *Trypanosoma Cruzi*. **Annual Review of Microbiology**, 27, n. 1, p. 347–382, 1973.

BUDDENBROCK, W. V. A criticism of the tropism theory of Jacques Loeb. **Journal of Animal Behavior**, 6, n. 5, p. 341, 1916.

BURGOS, J. M.; ALTCHER, J.; BISIO, M.; DUFFY, T. *et al.* Direct molecular profiling of minicircle signatures and lineages of *Trypanosoma cruzi* bloodstream populations causing congenital Chagas disease. **International journal for parasitology**, 37, n. 12, p. 1319–1327, 2007.

CAMPBELL, D. A.; WESTENBERGER, S. J.; STURM, N. R. The determinants of Chagas disease: connecting parasite and host genetics. **Curr Mol Med**, 4, n. 6, p. 549–562, Sep 2004.

CAMPOS, E. Estudos sobre uma raça neurotrópica de *Trypanosoma cruzi*. **Ann Fac Med São Paulo**, 2, p. 197–201, 1927.

- CAMPOS, M. A.; ALMEIDA, I. C.; TAKEUCHI, O.; AKIRA, S. *et al.* Activation of Toll-like receptor-2 by glycosylphosphatidylinositol anchors from a protozoan parasite. **J. Immunol**, 167, n. 1, p. 416–423, 2001.
- CAMPOS, P. C.; SILVA, V. G.; FURTADO, C.; MACHADO-SILVA, A. *et al.* Trypanosoma cruzi MSH2: Functional analyses on different parasite strains provide evidences for a role on the oxidative stress response. **Molecular and Biochemical Parasitology**, 176, n. 1, p. 8–16, 2011.
- CARADONNA, K. L.; ENGEL, J. C.; JACOBI, D.; LEE, C.-H. *et al.* Host Metabolism Regulates Intracellular Growth of Trypanosoma cruzi. **Cell Host & Microbe**, 13, n. 1, p. 108–117, 2013.
- CARDOSO, M. S.; REIS-CUNHA, J. L. L.; BARTHOLOMEU, D. C. Evasion of the Immune Response by Trypanosoma cruzi during Acute Infection. **Frontiers in immunology**, 6, p. 659, 2015.
- CARLSON, M. org.Mm.eg.db: Genome wide annotation for Mouse. 2017.
- CHAGAS, C. Nova tripanozomíaze humana: estudos sobre a morfologia e o ciclo evolutivo do Schizotrypanum cruzi n. gen., n. sp., agente etiológico de nova entidade morbida do homem. **Memórias do Instituto Oswaldo Cruz**, 1, p. 159–218, 1909.
- COMBS, T. P.; MUKHERJEE, S.; ALMEIDA, C. J. G.; JELICKS, L. A. *et al.* The adipocyte as an important target cell for Trypanosoma cruzi infection. **Journal of Biological Chemistry**, 280, n. 25, p. 24085–24094, 2005.
- CONNERS, E. E.; VINETZ, J. M.; WEEKS, J. R.; BROUWER, K. C. A global systematic review of Chagas disease prevalence among migrants. **Acta Tropica**, 156, p. 68–78, 2016.
- CORNEJO, E.; SCHLAERMANN, P.; MUKHERJEE, S. How to rewire the host cell: A home improvement guide for intracellular bacteria. **J Cell Biol**, 216, n. 12, p. 3931–3948, Dec 4 2017.
- COURA, J.; VIÑAS, P. Chagas disease: a new worldwide challenge. **Nature**, 465, n. n7301_suppl, 2010.
- COURA, J. R.; ANUNZIATO, N.; WILLCOX, H. P. F. Morbidade da doença de Chagas: I - Estudo de casos procedentes de vários estados do Brasil, observados no Rio de Janeiro. **Memórias do Instituto Oswaldo Cruz**, 78, p. 363–372, 1983.
- COURA, R. J.; ANUNZIATO, N.; WILLCOX, H. P. F. Morbidade da doença de Chagas: I - Estudo de casos procedentes de vários estados do Brasil, observados no Rio de Janeiro. **Memórias do Instituto Oswaldo Cruz**, 78, n. 3, p. 363–372, 1983.
- DE FREITAS, J. M.; AUGUSTO-PINTO, L.; PIMENTA, J. R.; BASTOS-RODRIGUES, L. *et al.* Ancestral genomes, sex, and the population structure of Trypanosoma cruzi. **PLoS pathogens**, 2, n. 3, 2006.

DE SOUZA, W. Basic cell biology of *Trypanosoma cruzi*. **Curr Pharm Des**, 8, n. 4, p. 269–285, 2002.

DESCHAMPS, P.; LARA, E.; MARANDE, W.; LOPEZ-GARCIA, P. *et al.* Phylogenomic analysis of kinetoplasts supports that trypanosomatids arose from within bodonids. **Mol Biol Evol**, 28, n. 1, p. 53–58, Jan 2011.

DIAS, E.; LARANJA, F. S.; MIRANDA, A.; NOBREGA, G. Chagas' disease; a clinical, epidemiologic, and pathologic study. **Circulation**, 14, n. 6, p. 1035–1060, Dec 1956.

DIAS, P. P.; CAPILA, R. F.; DO COUTO, N. F.; ESTRADA, D. *et al.* Cardiomyocyte oxidants production may signal to *T. cruzi* intracellular development. **PLoS Negl Trop Dis**, 11, n. 8, p. e0005852, Aug 23 2017.

DUTRA, W. O.; ROCHA, M. O.; TEIXEIRA, M. M. The clinical immunology of human Chagas disease. **Trends Parasitol**, 21, n. 12, p. 581–587, Dec 2005.

EL-SAYED, N. M.; MYLER, P. J.; BARTHOLOMEU, D. C.; NILSSON, D. *et al.* The genome sequence of *Trypanosoma cruzi*, etiologic agent of Chagas disease. **Science**, 309, n. 5733, p. 409–415, 2005.

FLEMING, A. M.; DING, Y.; BURROWS, C. J. Oxidative DNA damage is epigenetic by regulating gene transcription via base excision repair. **Proceedings of the National Academy of Sciences**, 114, n. 10, p. 2604–2609, 2017.

FRANCO, D. J.; VAGO, A. R.; CHIARI, E.; MEIRA, F. *et al.* *Trypanosoma cruzi*: mixture of two populations can modify virulence and tissue tropism in rat. **Experimental Parasitology**, 104, n. 1-2, p. 54–61, 2003.

FREIRE, A. C. G.; ALVES, C. L.; GOES, G. R.; RESENDE, B. C. *et al.* Catalase expression impairs oxidative stress-mediated signalling in *Trypanosoma cruzi*. **Parasitology**, p. 1–13, Jun 27 2017.

GABALDON, T.; GINGER, M. L.; MICHELS, P. A. Peroxisomes in parasitic protists. **Mol Biochem Parasitol**, 209, n. 1-2, p. 35–45, Sep – Oct 2016.

GARG, N.; POPOV, V. L.; PAPAConstantinou, J. Profiling gene transcription reveals a deficiency of mitochondrial oxidative phosphorylation in *Trypanosoma cruzi*-infected murine hearts: implications in chagasic myocarditis development. **Biochimica et biophysica acta**, 1638, n. 2, p. 106–120, 2003.

GILMORE, T. D. The Rel/NF-kappaB signal transduction pathway: introduction. **Oncogene**, 18, n. 49, p. 6842–6844, Nov 22 1999.

GOES, G. R.; ROCHA, P. S.; DINIZ, A. R. S.; AGUIAR, P. H. N. *et al.* *Trypanosoma cruzi* Needs a Signal Provided by Reactive Oxygen Species to Infect Macrophages. **PLOS Neglected Tropical Diseases**, 10, n. 4, 2016.

GRAVINA, H. D.; ANTONELLI, L.; GAZZINELLI, R. T.; ROPERT, C. Differential Use of TLR2 and TLR9 in the Regulation of Immune Responses during the Infection with *Trypanosoma cruzi*. **PLoS ONE**, 8, n. 5, 2013.

GRIFFITH, M.; WALKER, J. R.; SPIES, N. C.; AINSCOUGH, B. J. *et al.* Informatics for RNA Sequencing: A Web Resource for Analysis on the Cloud. **PLoS Comput Biol**, 11, n. 8, p. e1004393, Aug 2015.

GRYNBERG, P.; PASSOS-SILVA, D.; DE MOURÃO, M.; JR, R. *et al.* *Trypanosoma cruzi* Gene Expression in Response to Gamma Radiation. **PLoS ONE**, 7, n. 1, 2012.

GUSSIN, A. E. JACQUES LOEB: THE MAN AND HIS TROPISM THEORY OF ANIMAL CONDUCT. **Journal of the history of medicine and allied sciences**, 18, p. 321–336, 1963.

HAAS, B. J.; PAPANICOLAOU, A.; YASSOUR, M.; GRABHERR, M. *et al.* De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. **Nat Protoc**, 8, n. 8, p. 1494–1512, Aug 2013.

HAUCK, W. W.; DONNER, A. Wald's Test as Applied to Hypotheses in Logit Analysis. **Journal of the American Statistical Association**, 72, n. 360a, p. 851–853, 1977/12/01 1977.

HOUSTON-LUDLAM, G. A.; BELEW, T. A.; EL-SAYED, N. M. Comparative Transcriptome Profiling of Human Foreskin Fibroblasts Infected with the Sylvio and Y Strains of *Trypanosoma cruzi*. **PLOS ONE**, 11, n. 8, 2016.

HUSSON, R. [--Taxy, --tactism and --tropism]. **Bulletin of the World Health Organization**, 38, n. 5, p. 799–802, 1968.

KÖBERLE, F. Ciba Foundation Symposium 20 - Trypanosomiasis and Leishmaniasis (with Special Reference to Chagas' Disease). **wiley**, p. 137–158, 1970.

LI, Y.; SHAH-SIMPSON, S.; OKRAH, K.; BELEW, T. A. *et al.* Transcriptome Remodeling in *Trypanosoma cruzi* and Human Cells during Intracellular Infection. **PLOS Pathogens**, 12, n. 4, 2016.

LONGO, D. L.; BERN, C. Chagas' Disease. **The New England Journal of Medicine**, 373, n. 5, p. 456–466, 2015.

LOWE, R.; SHIRLEY, N.; BLEACKLEY, M.; DOLAN, S. *et al.* Transcriptomics technologies. **PLOS Computational Biology**, 13, n. 5, 2017.

LOWE, R.; SHIRLEY, N.; BLEACKLEY, M.; DOLAN, S. *et al.* Transcriptomics technologies. **PLoS Comput Biol**, 13, n. 5, p. e1005457, May 2017.

MACEDO, A. M.; OLIVEIRA, R. P.; PENA, S. D. Chagas disease: role of parasite genetic variation in pathogenesis. **Expert reviews in molecular medicine**, 4, n. 5, p. 1–16, 2002.

MACEDO, A. M.; PENA, S. D. J. Genetic Variability of *Trypanosoma cruzi*: Implications for the Pathogenesis of Chagas Disease. **Parasitology Today**, 14, n. 3, p. 119–124, 1998.

MAGALHÃES, L. M. M.; VIANA, A.; CHIARI, E.; GALVÃO, L. M. M. *et al.* Differential Activation of Human Monocytes and Lymphocytes by Distinct Strains of *Trypanosoma cruzi*. **PLoS neglected tropical diseases**, 9, n. 7, 2015.

MANQUE, P. A.; PROBST, C. M.; PROBST, C.; PEREIRA, M. C. S. *et al.* *Trypanosoma cruzi* Infection Induces a Global Host Cell Response in Cardiomyocytes. **Infection and Immunity**, 79, n. 5, p. 1855–1862, 2011.

MARCILI, A.; LIMA, L.; CAVAZZANA, M.; JUNQUEIRA, A. C. *et al.* A new genotype of *Trypanosoma cruzi* associated with bats evidenced by phylogenetic analyses using SSU rDNA, cytochrome b and Histone H2B genes and genotyping based on ITS1 rDNA. **Parasitology**, 136, n. 6, p. 641–655, 2009.

MARTIN, L. B.; FEI, Z.; GIOVANNONI, J. J.; ROSE, J. K. Catalyzing plant science research with RNA-seq. **Front Plant Sci**, 4, p. 66, 2013.

MARTINS, L.; CASTANHO, R.; DA ROSA, J.; DA SILVA, L. *et al.* Caracterização biológica, histopatológica e análise de ácido nucléico de uma cepa *Trypanosoma cruzi* da região de Marília, SP. **Revista da Sociedade Brasileira de Medicina Tropical**, 36, n. 1, p. 35–39, 2003.

MASOCHA, W.; KRISTENSSON, K. Passage of parasites across the blood-brain barrier. **Virulence**, 3, n. 2, p. 192–202, 2012.

MCCALL, L.-I.; SIQUEIRA-NETO, J. L.; MCKERROW, J. H. Location, Location, Location: Five Facts about Tissue Tropism and Pathogenesis. **PLOS Pathogens**, 12, n. 5, 2016.

MCFADDEN, G.; MOHAMED, M. R.; RAHMAN, M. M.; BARTEE, E. Cytokine determinants of viral tropism. **Nature Reviews Immunology**, 9, n. 9, p. 645–655, 2009.

MELO, R. C.; BRENER, Z. Tissue Tropism of Different *Trypanosoma cruzi* Strains. **The Journal of Parasitology**, 64, n. 3, p. 475, 1978.

MOGENSEN, T. H. Pathogen Recognition and Inflammatory Signaling in Innate Immune Defenses. **Clinical Microbiology Reviews**, 22, n. 2, p. 240–273, 2009.

MURPHY, M. P. How mitochondria produce reactive oxygen species. **The Biochemical journal**, 417, n. 1, p. 1–13, 2009.

NARDY, A.; FREIRE-DE-LIMA, C.; MORROT, A. Immune Evasion Strategies of *Trypanosoma cruzi*. **Journal of Immunology Research**, 2015, p. 1–7, 2015.

NAVARRO, I.; FERREIRA, F.; NAKAYA, H. I.; BARON, M. *et al.* MicroRNA Transcriptome Profiling in Heart of *Trypanosoma cruzi*-Infected Mice: Parasitological and Cardiological Outcomes. **PLOS Neglected Tropical Diseases**, 9, n. 6, 2015.

NETO, E.; GRUBER, A.; ZINGALES, B.; KALIL, J. Estudo da doença de Chagas: abordagem molecular. **Revista da Sociedade de Cardiologia**, Sat Apr 1 00:00:00 GMT-0300 1995.

OZSOLAK, F.; MILOS, P. M. RNA sequencing: advances, challenges and opportunities. **Nat Rev Genet**, 12, n. 2, p. 87–98, Feb 2011.

PABLOS, L.; OSUNA, A. Multigene Families in *Trypanosoma cruzi* and Their Role in Infectivity. **Infection and Immunity**, 80, n. 7, p. 2258–2264, 2012.

PADILLA, A. M.; BUSTAMANTE, J. M.; TARLETON, R. L. CD8+ T cells in *Trypanosoma cruzi* infection. **Current Opinion in Immunology**, 21, n. 4, p. 385–390, 2009.

PADILLA, A. M.; SIMPSON, L. J.; TARLETON, R. L. Insufficient TLR activation contributes to the slow development of CD8+ T cell responses in *Trypanosoma cruzi* infection. **Journal of immunology (Baltimore, Md. : 1950)**, 183, n. 2, p. 1245–1252, 2009.

PAIVA, C. N.; FEIJÓ, D. F.; DUTRA, F. F.; CARNEIRO, V. C. *et al.* Oxidative stress fuels *Trypanosoma cruzi* infection in mice. **Journal of Clinical Investigation**, 122, n. 7, p. 2531–2542, 2012.

PEREZ, C. J.; LYMBERY, A. J.; THOMPSON, R. C. Chagas disease: the challenge of polyparasitism? **Trends in Parasitology**, 30, n. 4, p. 176–182, 2014.

PEREZ, C. J.; THOMPSON, R. C.; KEATLEY, S. K.; WALSH, A. L. *et al.* The effect of reinfection and mixed *Trypanosoma cruzi* infections on disease progression in mice. **Acta Tropica**, 178, p. 107–114, 2018.

PEREZ, C. J.; THOMPSON, R. C. A.; KEATLEY, S. K.; WALSH, A. L. *et al.* The effect of reinfection and mixed *Trypanosoma cruzi* infections on disease progression in mice. **Acta Trop**, 178, p. 107–114, Feb 2018.

PIMENTEL, H.; BRAY, N. L.; PUENTE, S.; MELSTED, P. *et al.* Differential analysis of RNA-seq incorporating quantification uncertainty. **Nat Methods**, 14, n. 7, p. 687–690, Jul 2017.

RASSI, A.; RASSI, A.; DE REZENDE, J. Infectious Disease Clinics of North America. **Infectious Disease Clinics of North America**, 26, n. 2, p. 275–291, 2012.

RASSI, A.; RASSI, A.; MARIN-NETO, J. Chagas disease. **The Lancet**, 375, n. 9723, p. 1388–1402, 2010.

RICHARDSON, R. B.; ALLAN, D. S.; LE, Y. Greater organ involution in highly proliferative tissues associated with the early onset and acceleration of ageing in humans. **Experimental gerontology**, 55, p. 80–91, 2014.

RODRIGUES, C.; VALADARES, H.; FRANCISCO, A.; ARANTES, J. *et al.* Coinfection with Different *Trypanosoma cruzi* Strains Interferes with the Host Immune Response to Infection. **PLoS Neglected Tropical Diseases**, 4, n. 10, 2010.

RODRIGUES, C. M.; VALADARES, H. M.; FRANCISCO, A. F.; ARANTES, J. M. *et al.* Coinfection with different *Trypanosoma cruzi* strains interferes with the host immune response to infection. **PLoS Negl Trop Dis**, 4, n. 10, p. e846, Oct 12 2010.

ROGNER, U. C.; AVNER, P. Congenic mice: cutting tools for complex immune disorders. **Nature Reviews Immunology**, 3, n. 3, p. 243–252, 2003.

RUIZ DÍAZ, P.; MUCCI, J.; MEIRA, M. A. A.; BOGLIOTTI, Y. *et al.* *Trypanosoma cruzi* trans-sialidase prevents elicitation of Th1 cell response via interleukin 10 and downregulates Th1 effector cells. **Infection and immunity**, 83, n. 5, p. 2099–2108, 2015.

SALES-CAMPOS, H.; KAPPEL, H.; ANDRADE, C.; LIMA, T. *et al.* *Trypanosoma cruzi* DTU TcII presents higher blood parasitism than DTU TcI in an experimental model of mixed infection. **Acta Parasitologica**, 60, n. 3, p. 435–441, 2015.

SANCHEZ-VALDEZ, F. J.; PADILLA, A.; WANG, W.; ORR, D. *et al.* Spontaneous dormancy protects *Trypanosoma cruzi* during extended drug exposure. **Elife**, 7, Mar 26 2018.

SHANKER, S.; PAULSON, A.; EDENBERG, H. J.; PEAK, A. *et al.* Evaluation of Commercially Available RNA Amplification Kits for RNA Sequencing Using Very Low Input Amounts of Total RNA. **Journal of Biomolecular Techniques : JBT**, 2015.

SHANNON, P.; MARKIEL, A.; OZIER, O.; BALIGA, N. S. *et al.* Cytoscape: a software environment for integrated models of biomolecular interaction networks. **Genome research**, 13, n. 11, p. 2498–2504, 2003.

SOUTO, R. P.; FERNANDES, O.; MACEDO, A. M.; CAMPBELL, D. A. *et al.* DNA markers define two major phylogenetic lineages of *Trypanosoma cruzi*. **Molecular and biochemical parasitology**, 83, n. 2, p. 141–152, 1996.

SU, Z.; FANG, H.; HONG, H.; SHI, L. *et al.* An investigation of biomarkers derived from legacy microarray data for their utility in the RNA-seq era. **Genome Biol**, 15, n. 12, p. 523, Dec 03 2014.

TAN, H.; ANDREWS, N. W. Don't bother to knock – the cell invasion strategy of *Trypanosoma cruzi*. **Trends in Parasitology**, 18, n. 10, p. 427–428, 2002.

TANGE, O. GNU Parallel 2018. **Ole Tange**, 2018.

TANOWITZ, H. B.; MUKHOPADHYAY, A.; ASHTON, A. W.; LISANTI, M. P. *et al.* Microarray analysis of the mammalian thromboxane receptor-*Trypanosoma cruzi* interaction. **Cell cycle (Georgetown, Tex.)**, 10, n. 7, p. 1132–1143, 2011.

TARLETON, R. L. Immune system recognition of *Trypanosoma cruzi*. **Current Opinion in Immunology**, 19, n. 4, p. 430–434, 2007.

TIBAYRENC, M. Genetic epidemiology of parasitic protozoa and other infectious agents: the need for an integrated approach. **Int J Parasitol**, 28, n. 1, p. 85–104, Jan 1998.

VAGO, A. R.; ANDRADE, L. O.; LEITE, A. A.; D'AVILA REIS, D. *et al.* Genetic characterization of *Trypanosoma cruzi* directly from tissues of patients with chronic Chagas disease: differential distribution of genetic types into diverse organs. **The American journal of pathology**, 156, n. 5, p. 1805–1809, 2000.

VAGO, A. R.; MACEDO, A. M.; OLIVEIRA, R. P.; ANDRADE, L. O. *et al.* Kinetoplast DNA signatures of *Trypanosoma cruzi* strains obtained directly from infected tissues. **The American journal of pathology**, 149, n. 6, p. 2153–2159, 1996.

VELCULESCU, V. E.; ZHANG, L.; ZHOU, W.; VOGELSTEIN, J. *et al.* Characterization of the yeast transcriptome. **Cell**, 88, n. 2, p. 243–251, Jan 24 1997.

VIANNA, G. Contribuição para o estudo da anatomia patológica da "Molestia de Carlos Chagas" (Esquizotripanose humana ou tireoidite parasitaria). . **Memórias do Instituto Oswaldo Cruz**, 3, n. 2, p. 276–294, 1911.

WAGNER, W.; SO, M. Genomic variation of *Trypanosoma cruzi*: involvement of multicopy genes. **Infection and immunity**, 58, n. 10, p. 3217–3224, 1990.

WHO. Chagas disease (American trypanosomiasis). **Fact sheet**, 2017.

YOSHIDA, N.; DORTA, M. L.; FERREIRA, A. T.; OSHIRO, M. E. *et al.* Removal of sialic acid from mucin-like surface molecules of *Trypanosoma cruzi* metacyclic trypomastigotes enhances parasite-host cell interaction. **Mol Biochem Parasitol**, 84, n. 1, p. 57–67, Jan 1997.

ZINGALES, B.; ANDRADE, S. G.; BRIONES, M. R.; CAMPBELL, D. A. *et al.* A new consensus for *Trypanosoma cruzi* intraspecific nomenclature: second revision meeting recommends TcI to TcVI. **Mem Inst Oswaldo Cruz**, 104, n. 7, p. 1051–1054, Nov 2009.

ZINGALES, B.; MILES, M. A.; CAMPBELL, D. A.; TIBAYRENC, M. *et al.* The revised *Trypanosoma cruzi* subspecific nomenclature: Rationale, epidemiological relevance and research applications. **Infection, Genetics and Evolution**, 12, n. 2, p. 240–253, 2011.

ZINGALES, B.; MILES, M. A.; CAMPBELL, D. A.; TIBAYRENC, M. *et al.* The revised *Trypanosoma cruzi* subspecific nomenclature: rationale, epidemiological relevance and research applications. **Infect Genet Evol**, 12, n. 2, p. 240–253, Mar 2012.

9. ANEXO

ANEXO A – Artigo publicado

CASTRO, T. B. R.; CANESSO, M. C. C.; BORONI, M.; CHAME, D. F.; DE LAET SOUZA, D.; TOLEDO, N. E.; TAHARA, E. B.; PENA, S. D.; MACHADO, C. R.; CHIARI, E.; FRANCO, G. R. Differential modulation of mouse heart gene expression by infection with two *Trypanosoma cruzi* strains: a transcriptome analysis. *Frontiers in Genetics*, v. 11, p. 1031, 2020. DOI: 10.3389/fgene.2020.01031.