

**UNIVERSIDADE FEDERAL DO PARANÁ**  
**PROGRAMA DE PÓS-GRADUAÇÃO EM BIOLÓGICA CELULAR E MOLECULAR**

**RAFAEL LUIS KESSLER**

**TRANSCRIPTÔMICA DE *TRYPANOSOMA CRUZI* EM RESPOSTA  
A INIBIDORES DA SÍNTESE DE ESTERÓIS**

**CURITIBA**

**2010**

**RAFAEL LUIS KESSLER**

**TRANSCRIPTÔMICA DE *TRYPANOSOMA CRUZI* EM RESPOSTA  
A INIBIDORES DA SÍNTESE DE ESTERÓIS**

Dissertação apresentada ao Programa de Pós-Graduação em Biologia Celular e Molecular da Universidade Federal do Paraná, como pré-requisito para obtenção do título de Mestre em Biologia Celular e Molecular

Orientadores: Dr. Marco Aurélio Krieger  
Dr. Christian Macagnan Probst

CURITIBA

2010

# PROGRAMA DE PÓS-GRADUAÇÃO EM BIOLOGIA CELULAR E MOLECULAR

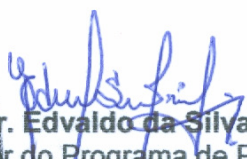
Departamento de Biologia Celular e Departamento de Fisiologia  
Setor de Ciências Biológicas  
Universidade Federal do Paraná  
Instituto Carlos Chagas (ICC)

## DECLARAÇÃO

Declaramos para os devidos fins que **Rafael Luis Kessler**, no dia 30 de setembro de 2010, no Instituto Carlos Chagas-ICC, Curitiba, defendeu sua Dissertação de Mestrado em Biologia Celular e Molecular, intitulada: ***"Transcriptômica de Trypanosoma cruzi em resposta a inibidores da síntese de esteróis"***, com a banca examinadora constituída, pelos pesquisadores: Dr. Augusto Savio Peixoto Ramos, do Instituto Carlos Chagas – ICC, Curitiba, Dra. Daniela Parada Pavoni, do Instituto Carlos Chagas - ICC, Curitiba, e Dr. Christian Macagnan Probst, como co-orientador e presidente da Banca Examinadora, do Instituto Carlos Chagas- ICC, Curitiba, e tendo sido Aprovado, recebendo o título de Mestre em Biologia Celular e Molecular, desde que cumpra integralmente todas as exigências estabelecidas nas Normas Internas deste Programa.

Curitiba, 30 de setembro de 2010.



  
Prof. Dr. Edvaldo da Silva Trindade  
Coordenador do Programa de Pós-Graduação  
Biologia Celular e Molecular

Aos meus pais, Sergio e Sirlei.

## **AGRADECIMENTOS**

Ao Dr. Marco Aurélio Krieger e Dr. Christian Macagnan Probst pela confiança e orientação.

A todos meus amigos, colegas e profissionais que auxiliaram para que este projeto se tornasse realidade.

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO</b>	<b>1</b>
1.1	CLASSIFICAÇÃO TAXONÔMICA DE <i>TRYPANOSOMA CRUZI</i>	1
1.2	CICLO DE VIDA DE <i>T. CRUZI</i>	2
1.3	CARACTERÍSTICAS CELULARES DE <i>T. CRUZI</i>	4
1.4	DOENÇA DE CHAGAS	9
1.5	A BIOSÍNTESE DE ESTERÓIS COMO ALVO QUIMIOTERÁPICO DA DOENÇA DE CHAGAS	13
1.5.1	Esteróis: estrutura, funções e biossíntese em células eucarióticas	13
1.5.2	Esteróis em <i>Trypanosoma cruzi</i> : síntese, funções e inibidores	20
1.5.2.1	Inibidores de 14- $\alpha$ demetilase: azóis	26
1.5.2.2	Inibidores de HMGR: estatinas	33
1.6	TRANSCRIÇÃO POLICISTRÔNICA EM TRIPANOSSOMATÍDEOS	37
1.7	REGULAÇÃO DA EXPRESSÃO GÊNICA EM TRIPANOSSOMATÍDEOS	41
1.8	GENÔMICA	45
1.8.1	Novas tecnologias de sequenciamento de DNA	46
1.8.2	Transcriptômica: microarranjo versus RNA-Seq	49
1.8.3	Organização genômica do <i>T. cruzi</i>	54
1.8.4	O projeto Reguloma de <i>T. cruzi</i>	56
<b>2</b>	<b>OBJETIVOS</b>	<b>59</b>
2.1	OBJETIVOS GERAIS	59
2.2	OBJETIVOS ESPECÍFICOS	59
<b>3</b>	<b>JUSTIFICATIVA</b>	<b>60</b>
<b>4</b>	<b>MATERIAIS E MÉTODOS</b>	<b>61</b>
4.1	ORIGEM DOS REAGENTES	61
4.2	MEIO DE CULTURA	61
4.3	SOLUÇÕES E TAMPÕES	61
4.4	CULTIVO DO PARASITA	62
4.5	INIBIDORES DA BIOSÍNTESE DE ERGOSTEROL (ibeS)	62
4.6	ATIVIDADE ANTIPROLIFERATIVA <i>IN VITRO</i> DOS IBES	62
4.6.1	Experimentos de recuperação	63
4.7	MICROSCOPIA	64

4.7.1	Microscopia óptica .....	64
4.7.2	Microscopia eletrônica de transmissão.....	64
4.7.3	Microscopia de fluorescência .....	65
4.8	CITOMETRIA DE FLUXO .....	65
4.9	AMOSTRAS DE RNA.....	66
4.9.1	RNA total.....	66
4.9.2	RNA polissomal .....	66
4.9.3	Determinação da quantidade e qualidade do RNA extraído.....	67
4.10	RNA-SEQ.....	68
4.10.1	Preparo das bibliotecas .....	68
4.10.2	Preparo da beads para seqüenciamento .....	70
4.10.2.1	PCR em emulsão .....	70
4.10.2.2	Quebra da emulsão e lavagem das beads .....	71
4.10.2.3	Enriquecimento das beads contendo DNA.....	71
4.10.2.4	Modificação da extremidade 3' e quantificação das beads .....	72
4.10.3	Deposição das beads e sequenciamento .....	73
4.10.4	Análise dos dados .....	74
<b>5</b>	<b>RESULTADOS .....</b>	<b>77</b>
5.1	EFEITO TRIPANOCIDA DOS FÁRMACOS .....	77
5.1.1	Lovastatina .....	77
5.1.2	Cetoconazol .....	80
5.2	ALTERAÇÕES MORFO-FISIOLÓGICAS EM RESPOSTA ÀS DROGAS.....	82
5.2.1	IC <sub>50</sub> .....	83
5.2.2	DL100.....	91
5.3	RNA-SEQ.....	94
5.3.1	Desenho experimental .....	94
5.3.2	Sequências geradas e alinhamento no genoma de <i>T. cruzi</i> .....	96
5.3.3	Réplicação técnica e biológica .....	101
5.3.4	Genes Diferencialmente expressos.....	104
5.3.4.1	IC <sub>50</sub> Cetoconazol .....	104
5.3.4.2	IC <sub>50</sub> Lovastatina.....	120
5.3.4.3	DL100 Cetoconazol.....	131
5.3.4.4	DL100 Lovastatina .....	139

5.3.4.5	IC <sub>50</sub> : respostas discrepantes entre as drogas .....	146
5.3.4.6	DL100: respostas distintas entre as drogas.....	149
<b>6</b>	<b>DISCUSSÕES.....</b>	<b>153</b>
6.1	REGULAÇÃO DA EXPRESSÃO GÊNICA EM <i>t. CRUZI</i> .....	153
6.2	INIBIÇÃO DA SÍNTESE DE ESTERÓIS: DE GENÔMICA A BIOLOGIA CELULAR .....	156
6.3	REGULAÇÃO DA SÍNTESE DE ESTERÓIS EM <i>T. CRUZI</i> .....	163
6.4	DL100: ANÁLISE GENÔMICA DA MORTE CELULAR EM <i>T. CRUZI</i> .....	167
6.5	REGULOMA .....	169
<b>7</b>	<b>CONCLUSÕES.....</b>	<b>174</b>
<b>8</b>	<b>PERSPECTIVAS .....</b>	<b>176</b>
<b>9</b>	<b>REFERÊNCIAS .....</b>	<b>177</b>



## LISTA DE FIGURAS

Figura 1.1. CICLO DE VIDA DO <i>TRYPANOSOMA CRUZI</i> .....	4
Figura 1.2 – ESTRUTURA CELULAR DE <i>T. CRUZI</i> (EPIMASTIGOTA).....	5
Figura 1.3 – ESTRUTURA QUÍMICA DOS ESTERÓIS .....	14
Figura 1.4 – AS QUATRO ETAPAS DA BIOSÍNTESE DE ESTERÓIS .....	15
Figura 1.5 – VIA DO MEVALONATO.....	16
Figura 1.6 – CONVERSÃO DO MEVALONATO A UNIDADES ATIVADAS DE ISOPRENO .....	17
Figura 1.7 – FORMAÇÃO DO ESQUALENO .....	18
Figura 1.8 – FORMAÇÃO DO NÚCLEO ESTERÓIDE.....	19
Figura 1.9 – VIA DE SÍNTESE DE ESTERÓIS EM <i>TRYPANOSOMA CRUZI</i> .....	23
Figura 1.10 – ESTRUTURA QUÍMICA DOS AZÓIS .....	28
Figura 1.11 – ESTRUTURA QUÍMICA DAS ESTATINAS .....	35
Figura 1.12 – RNA-SEQ .....	51
Figura 4.1 – PREPARO DAS BIBLIOTECAS DE RNA-SEQ (SOLID) .....	68
Figura 4.2 – PCR EM EMULSÃO.....	71
Figura 4.3 – ENRIQUECIMENTO DE BEADS MOLDES .....	72
Figura 4.4- SEQUENCIAMENTO DO BC E FRAGMENTO DE 50 BASES PELO SOLID .....	74
Figura 5.1 – EFEITO ANTIPROLIFERATIVO DE LOVASTATINA .....	77
Figura 5.2 – MORTE CELULAR COM ALTAS DOSES DE LOVASTATINA.....	79
Figura 5.3 – RECUPERAÇÃO: LOVASTATINA .....	80
Figura 5.4 – EFEITO ANTIPROLIFERATIVO DE CETOCONAZOL.....	81
Figura 5.5 – DOSE/RESPOSTA CETOCONAZOL: DM28C <i>VERSUS</i> Y .....	81
Figura 5.6 – MORTE CELULAR E RECUPERAÇÃO A CETOCONAZOL.....	82
Figura 5.7 – MICROSCOPIA ÓPTICA IC <sub>50</sub> .....	83
Figura 5.8 - IC <sub>50</sub> : CITOMETRIA DE FLUXO COM LARANJA DE ACRIDINA.....	85
Figura 5.9 – MICROSCOPIA CONFOCAL: LYSO TRACKER RED .....	86
Figura 5.10 – M.E.T. DE <i>T. CRUZI</i> : CÉLULAS CONTROLE .....	87
Figura 5.11 – ALTERAÇÕES ULTRAESTRUTURAIS À IC <sub>50</sub> DE CETOCONAZOL.....	88
Figura 5.12 – ALTERAÇÕES ULTRAESTRUTURAIS À IC <sub>50</sub> DE LOVASTATINA.....	89
Figura 5.13 – IC <sub>50</sub> : CITOMETRIA DE FLUXO COM RHODAMINA 123 .....	90
Figura 5.14 – MICROSCOPIA ÓPTICA DL100.....	91
Figura 5.15 – DL100: CITOMETRIA DE FLUXO .....	92
Figura 5.16 – ALTERAÇÕES ULTRAESTRUTURAIS EM RESPOSTA A DL100.....	93
Figura 5.17 – DESENHO EXPERIMENTAL RNA-SEQ .....	95
Figura 5.18 – SEQUÊNCIAS PRODUZIDAS NOS EXPERIMENTOS DE RNA-SEQ .....	97
Figura 5.19 – PRESENÇA DE RNAs NÃO CODIFICADORES NOS DADOS .....	100
Figura 5.20 – RÉPLICAS TÉCNICAS SOLID.....	102
Figura 5.21 - RÉPLICAS BIOLÓGICAS SOLID.....	103

Figura 5.22 – IC <sub>50</sub> CETOCONAZOL: N DE GENES SIGNIFICATIVOS EM CADA PONTO .....	106
Figura 5.23 – EXEMPLOS DE GENES SIGNIFICATIVOS PARA IC <sub>50</sub> DE CETOCONAZOL .....	107
Figura 5.24 – IC <sub>50</sub> CETOCONAZOL: GRÁFICO DE VENN PARA GENES SIGNIFICATIVOS EM DIFERENTES MÉTODOS ESTATÍSTICOS .....	110
Figura 5.25 – IC <sub>50</sub> CETOCONAZOL: CLUSTERIZAÇÃO DE GENES SIGNIFICATIVOS PARA DIFERENTES MÉTODOS ESTATÍSTICOS .....	112
Figura 5.26 – IC <sub>50</sub> CETOCONAZOL: CLUSTERIZAÇÃO DOS SUPERGENES SIGNIFICATIVOS E ENRIQUECIMENTO DE G.O. ....	114
Figura 5.27 – VIA DE PROCESSAMENTO PROTÉICO NO r.e. ENRIQUECIDA COM GENES DE EXPRESSÃO DIMINUÍDA PARA IC <sub>50</sub> CETOCONAZOL.....	118
Figura 5.28 – IC <sub>50</sub> LOVASTATINA: N DE GENES SIGNIFICATIVOS EM CADA PONTO .....	121
Figura 5.29 – EXEMPLOS DE GENES SIGNIFICATIVOS PARA IC <sub>50</sub> LOVASTATINA .....	122
Figura 5.30 – IC <sub>50</sub> LOVASTATINA: GRÁFICO DE VENN PARA GENES SIGNIFICATIVOS EM DIFERENTES MÉTODOS ESTATÍSTICOS .....	124
Figura 5.31 – IC <sub>50</sub> LOVASTATINA: CLUSTERIZAÇÃO DE GENES SIGNIFICATIVOS PARA DIFERENTES MÉTODOS ESTATÍSTICOS .....	126
Figura 5.32 – IC <sub>50</sub> LOVASTATINA: CLUSTERIZAÇÃO HIERÁRQUICA DOS SUPERGENES SIGNIFICATIVOS E ENRIQUECIMENTO DE G.O.....	127
Figura 5.33 – IC <sub>50</sub> LOVASTATINA: VIAS ENRIQUECIDAS DO KEGG.....	129
Figura 5.34 – DL100 CETOCONAZOL: DIAGRAMA DE VENN.....	132
Figura 5.35 – DL100 CETOCONAZOL: AGRUPAMENTO HIERÁRQUICO .....	134
Figura 5.36 – DL100 CETOCONAZOL: VIA DE DEGRADAÇÃO DE AMINOÁCIDOS (KEGG).....	137
Figura 5.37 – DL100 CETOCONAZOL: INIBIÇÃO DA SÍNTESE E PROCESSAMENTO DE RNA .....	138
Figura 5.38 – DL100 LOVASTATINA: DIAGRAMA DE VENN .....	140
Figura 5.39 – DL100 LOVASTATINA: AGRUPAMENTO HIERÁRQUICO .....	141
Figura 5.40 – DL100 LOVASTATINA: FOSFORILAÇÃO OXIDATIVA (KEGG) .....	144
Figura 5.41 – DL100 LOVASTATINA: ATIVAÇÃO DA VIA DA GLUTATIONA (KEGG).....	145
Figura 5.42 – IC <sub>50</sub> : COMPARAÇÃO ENTRE AS DROGAS.....	147
FIGURA 5.43 – IC <sub>50</sub> : GENES DIFERENCIALMENTE REGULADOS ENTRE AS DROGAS .....	148
Figura 5.44 – DL100: COMPARAÇÃO ENTRE AS DROGAS .....	151
Figura 5.45 – DL100: GENES DIFERENCIALMENTE REGULADOS ENTRE AS DROGAS.....	152
Figura 6.1 – ALINHAMENTO DOS <i>READS</i> DE RNA-SEQ NO GENOMA .....	154
Figura 6.2 – IC <sub>50</sub> CETOCONAZOL: PROTEÍNAS RESERVOSSOMAS DIFERENCIALMENTE EXPRESSAS.....	160
Figura 6.3 – IC <sub>50</sub> LOVASTATINA: PROTEÍNAS RESERVOSSOMAS DIFERENCIALMENTE EXPRESSAS .....	161
Figura 6.4 – IC <sub>50</sub> CETOCONAZOL: REGULAÇÃO DA SÍNTESE DE ESTERÓIS.....	164
Figura 6.5 - REGULOMA.....	171
Figura 6.6 – POSSÍVEL <i>REGULON</i> CONTENDO PROTEÍNA LIGADORA DE RNA.....	172



## LISTA DE TABELAS

Tabela 1.1 – ATIVIDADE DE AZÓIS TESTADOS CONTRA EPIMASTIGOTAS DE <i>T. CRUZI</i> EM CULTURA .....	29
Tabela 1.2 – ATIVIDADE DE AZÓIS TESTADA CONTRA AMASTIGOTAS DE <i>T. CRUZI</i> EM CULTURAS DE CÉLULAS DE MAMÍFEROS.....	29
Tabela 1.3 – ATIVIDADE DE AZÓIS (E BENZONIDAZOL) CONTRA A FORMA AMASTIGOTA DA CEPA TULAHUEN DE <i>T. CRUZI</i> CULTIVADA EM FIBROBLASTOS 3T3 DE CAMUNDONGOS .....	30
Tabela 1.4 – CARACTERÍSTICAS DOS SEQUENCIADORES DE DNA.....	47
Tabela 5.1 – RESULTADOS DOS DADOS DE ALINHAMENTO .....	99
Tabela 5.2 – IC <sub>50</sub> CETOCONAZOL: GENES SIGNIFICATIVOS .....	105
Tabela 5.3 – IC <sub>50</sub> CETOCONAZOL: GENES SIGNIFICATIVOS PARA 0+24 <i>VERSUS</i> 96+120.....	108
Tabela 5.4 – IC <sub>50</sub> CETOCONAZOL: GENES SIGNIFICATIVOS TESTE SAM 3 CLASSES.....	109
Tabela 5.5 – IC <sub>50</sub> CETOCONAZOL: N DE GENES SIGNIFICATIVOS PELO MÉTODO EDGER GLOBAL.....	113
Tabela 5.6 – IC <sub>50</sub> LOVASTATINA: GENES SIGNIFICATIVOS .....	121
Tabela 5.7 – IC <sub>50</sub> LOVASTATINA: GENES SIGNIFICATIVOS PARA 0+24 <i>VERSUS</i> 96+120 .....	123
Tabela 5.8 – IC <sub>50</sub> LOVASTATINA: GENES SIGNIFICATIVOS TESTE SAM 3 CLASSES .....	123
Tabela 5.9 – IC <sub>50</sub> LOVASTATINA: N DE GENES SIGNIFICATIVOS PELO MÉTODO EDGER GLOBAL.....	124
Tabela 5.10 – DL100 CETOCONAZOL: N DE GENES SIGNIFICATIVOS .....	131
Tabela 5.11 – ENRIQUECIMENTO DE G.O. PARA DL100 CETOCONAZOL .....	135
Tabela 5.12 – DL100 LOVASTATINA: N DE GENES SIGNIFICATIVOS.....	139
Tabela 5.13 – ENRIQUECIMENTO DE G.O. PARA DL100 LOVASTATINA.....	142
Tabela 6.1 – PROTEÍNAS RESERVOSSOMAS AUMENTADAS PARA IC <sub>50</sub> DE AMBAS AS DROGAS .	162

## RESUMO

O *Trypanosoma cruzi* é o protozoário causador da doença de Chagas, enfermidade que afeta cerca de 15 milhões de pessoas no continente americano. Diferentemente de mamíferos que sintetizam colesterol, *T. cruzi* produz principalmente ergosterol, esterol característico de fungos. Por este motivo, a via de síntese deste lipídio representa um potencial alvo quimioterápico da Doença de Chagas. Dentro de uma etapa inicial do projeto Reguloma de *T. cruzi*, o presente trabalho teve como objetivo analisar o transcriptoma deste parasita em resposta a duas drogas inibidoras que agem em pontos distintos desta via: Cetoconazol e Lovastatina. Foram calculadas as doses das drogas capazes de inibir o crescimento das culturas celulares em 50% após 3 dias de exposição (IC<sub>50</sub>: 32µM Cetoconazol e 50µM Lovastatina); ou matar todas as células em 24 horas (DL100: 120µM Cetoconazol e 100µM Lovastatina). Foram extraídas amostras de RNA polissomal em tempos precoces de exposição (30 e 60 minutos) à DL100 (triplicata) e em tempos tardios (1 à 5 dias) à IC<sub>50</sub> (duplicata). Os cDNAs gerados a partir das frações poli-A<sup>+</sup> foram analisados através de sequenciamento paralelo massivo (RNA-Seq) com a plataforma SOLiD 3. Foram gerados cerca de 320 milhões de reads de 50 bases para os experimentos DL100, sendo que a análise dos dados evidenciou quase 3.000 genes diferencialmente expressos em resposta a Lovastatina e cerca de 1.500 frente a Cetoconazol, com FDR menor de 0,1% e *foldchange* (FC) de 1,5 vez. Análises de enriquecimento de ontologia gênica (GO) mostraram que a resposta a DL100 envolve mecanismos estereotipados de expressão gênica relacionados à ativação de vias de catabolismo, principalmente de aminoácidos, e enriquecimento de genes envolvidos na síntese de ATP por fosforilação oxidativa, além de uma forte inibição da síntese e processamento de RNA. Foi evidenciado também o aumento da expressão de cisteína-peptidases do tipo Calpaína e Catepsina na resposta a DL100, as quais podem estar envolvidas em morte celular necrótica. Esses resultados são compatíveis com experimentos de microscopia eletrônica de transmissão e citometria de fluxo. Para os experimentos IC<sub>50</sub>, foram gerados cerca de 207 milhões de reads, selecionando-se 863 genes diferencialmente expressos em resposta a Cetoconazol e 462 em resposta a Lovastatina, com FDR 5% e FC 1,5 vez. Experimentos de microscopia eletrônica e de fluorescência e citometria de fluxo evidenciaram um aumento do número de reservossomos, enquanto os dados de RNA-Seq indicam o aumento de expressão de possíveis proteínas localizadas nessa organela, apontando para uma possível função dessa estrutura celular na síntese de esteróis. Análises de GO demonstraram um enriquecimento de vias envolvidas no transporte de aminoácidos e atividade peptidásica (para genes aumentados) e proteínas nucleossomais (genes diminuídos), além de respostas droga-específicas. Os dados gerados constituem a primeira avaliação global da resposta do parasita a essas drogas, e sua análise em conjunto com os demais ensaios representa uma abordagem riquíssima para o entendimento mais aprofundado das modificações moleculares que ocorrem com a inibição da biossíntese do ergosterol. Embora a quantidade de dados gerada no presente trabalho não seja suficiente para análises robustas de reguloma, resultados preliminares demonstram que tal metodologia permite uma compreensão integrada e muito poderosa da biologia básica do *Trypanosoma cruzi*, principalmente com relação a mecanismos pós-transcricionais da expressão gênica. A inclusão dos dados gerados no presente trabalho na base maior do projeto reguloma representa um importante avanço na identificação das redes de regulação da expressão gênica.

## ABSTRACT

*Trypanosoma cruzi* is the protozoan that causes Chagas' disease, illness that affects about 15 million people in the Americas. Unlike mammals that synthesize cholesterol, *T. cruzi* produces mainly ergosterol, sterol characteristic of fungi. For this reason, the route of synthesis of this lipid represents a potential chemotherapeutic target for Chagas disease. Within an initial phase of *T. cruzi* Regulome project, the present study aim to analyze the transcriptome of this parasite in response to two inhibitory drugs that act at different points of the ergosterol biosynthesis pathway: ketoconazole and lovastatin. We calculated the drug doses capable of inhibiting the growth of cell cultures by 50% after 3 days of exposure (IC<sub>50</sub>: 32μM ketoconazole and 50μM lovastatin), or kill all cells at 24 hours (DL100: 120μM ketoconazole and 100μM lovastatin). Polysomal mRNA samples were extracted at very early exposure times (30 and 60 minutes) to the LD100 (triplicate) and at later times (1 to 5 days) to IC<sub>50</sub> (duplicate). The cDNAs generated from poly A+ fractions were analyzed by massive parallel sequencing (RNA-Seq) with the SOLiD 3 platform. We generated about 320 million reads of 50 bases for the DL100 experiments, and the data analysis revealed nearly 3,000 genes differentially expressed in response to lovastatin and about 1,500 compared to ketoconazole, with FDR less than 0.1% and foldchange (FC) of 1.5. Analysis of gene ontology enrichment (GO) showed that the response to LD100 involves a stereotypical expression pattern that include mechanisms related to activation of catabolic pathways, especially amino acids, and enrichment of genes involved in ATP synthesis by oxidative phosphorylation, in addition to a strong inhibition of RNA synthesis and processing. It was also shown increased expression of cysteine peptidases of type calpain and cathepsin in response to LD100, which may be involved in necrotic cell death. These results are consistent with experiments of transmission electron microscopy and flow cytometry. For the IC<sub>50</sub> experiments, we generated about 207 million reads, selecting 863 genes differentially expressed in response to ketoconazole and 462 in response to Lovastatin, with FDR 5% and FC 1.5. Experiments of electron and fluorescence microscopy and flow cytometry showed an increased number of reservosomes, while the RNA-Seq data indicate the increased expression of possible proteins located in this organelle, suggesting a possible role of these cell structure in the synthesis of sterols. GO analysis showed an enrichment of pathways involved in amino acid transport and peptidase activity (for increased genes) and proteins nucleosomes (genes decreased), as well as drug-specific responses. The data generated constitute the first comprehensive assessment of the response of parasites to these drugs, and their analysis in conjunction with other tests represents a rich approach to the deeper understanding of the molecular changes that occur through the inhibition of the ergosterol biosynthesis. Although the amount of data generated in this work is not sufficiently robust reguloma analysis, preliminary results show that this methodology allows a very powerful and integrated understanding of the basic biology of *Trypanosoma cruzi*, mainly in relation to post-transcriptional mechanisms of gene expression. The inclusion of data generated in this study on the larger reguloma base project represents an important advance in the identification of regulatory networks of gene expression.

# 1 INTRODUÇÃO

## 1.1 CLASSIFICAÇÃO TAXONÔMICA DE *TRYPANOSOMA CRUZI*

*Trypanosoma cruzi* (CHAGAS, 1909) é um protozoário flagelado pertencente ao reino Protista, subreino Protozoa, filo Sarcomastigophora, subfilo Mastigophora, classe Zoomastigophora e ordem Kinetoplastida (LEVINE *et al.*, 1980).

A ordem Kinetoplastida abrange protozoários flagelados que contêm uma região especializada, conhecida como cinetoplasto, onde localiza-se o DNA mitocondrial, também denominado de kDNA (SHAPIRO & ENGLUND, 1995; TELLERIA *et al.*, 2006). Entre os cinetoplastídeos incluem-se duas famílias: **Trypanosomatidae**, compreendendo parasitas causadores de doenças em humanos (gêneros *Trypanosoma* e *Leishmania*, entre outros); e **Bodonidae**, cinetoplastídeos de vida livre.

O gênero *Trypanosoma* é um dos gêneros mais importantes, no ponto de vista de saúde pública, da família Trypanosomatidae, pois inclui, além do *Trypanosoma cruzi*, agente etiológico da Doença de Chagas (CHAGAS, 1909), outras duas subespécies do *Trypanosoma brucei* que ocasionam doenças em humanos, como o *T. b. rhodesiense* e *T. b. gambiense*, ambos agentes da doença do sono. Adicionalmente, ainda há o *T. b. brucei*, causador da enfermidade conhecida como nagana em animais (HIDE, 1999).

De acordo com a forma de desenvolvimento no hospedeiro invertebrado, o gênero *Trypanosoma* foi dividido em dois grupos: Stercoraria e Salivaria. O primeiro inclui parasitas que se desenvolvem no tubo digestivo do vetor e progridem em direção a sua porção final, culminando na liberação das suas formas infectivas através das excretas, como ocorre com *T. cruzi*. Já no grupo Salivaria estão incluídos os parasitas que se desenvolvem no trato digestivo do vetor, mas posteriormente migram para as glândulas salivares, para deste modo serem inoculados mecanicamente durante a picada do inseto vetor, como no caso do *T. brucei* (DIAS, 2000).

Embora *T. cruzi* seja descrito como um único táxon, este parasita mostra heterogeneidade genotípica e fenotípica considerável (STURM *et al.*, 2003; revisto por MACEDO *et al.*, 2004). Através da análise conjunta de diversos marcadores moleculares, *T. cruzi* foi subdividido em duas linhagens principais, designadas como *T. cruzi* I e II

(Recommendations, 1999), ou DTUs (*discrete typing unit*) I e II (TIBAYRENC, 2003). *T. cruzi* I é associado com o ciclo selvagem de transmissão e mamíferos arbóreos, enquanto *T. cruzi* II parece relacionar-se com o ciclo doméstico de transmissão e nichos terrestres (YEO *et al.*, 2005; revisto por STURM & CAMPBELL, 2009). Mais tarde, *T. cruzi* II foi subdividido em cinco grupos, denominados IIa, IIb, IIc, IId e IIe (BRISSE *et al.*, 2000), culminando com a presença de seis grupos distintos dentro de *T. cruzi*. Embora a correlação evolutiva desses seis grupos não seja completamente compreendida, aparentemente os DTUs I e IIb representam as linhagens parentais, a partir das quais divergem os demais grupos atuais (WESTENBERGER *et al.*, 2005). Uma hibridização entre I e IIb, seguida por evolução clonal independente, teria sido responsável pela gênese dos grupos homocigotos mosaico IIa e IIc. Um segundo evento mais recente de hibridização, entre as linhagens IIb e IIc, parece ter formado os atuais grupos híbridos heterocigotos de *T. cruzi* (IId e IIe). Dentro deste possível quadro evolutivo, Sturm e Campbell (2009) destacam a presença de quatro grupos biologicamente relevantes em *T. cruzi*: as duas linhagens parentais originais I e IIb; um terceiro grupo formado pelas linhagens híbridas homocigotas IIa e IIc; e um quarto grupo composto pelos híbridos heterocigotos IId e IIe.

Um novo consenso acerca da nomenclatura intraespecífica de *T. cruzi* foi proposto em 2009, no XXXVI Annual Meeting of Basic Research in Chagas Disease (ZINGALES *et al.*, 2009). As cepas de *T. cruzi* devem ser classificadas em seis DTUs, nomeados *T. cruzi* I a VI. A correspondência com a nomenclatura anterior é a seguinte: I-I, II-IIb; III-IIc; IV-IIa; V-IId, VI-IIe.

## 1.2 CICLO DE VIDA DE *T. CRUZI*

O parasita *Trypanosoma cruzi* é uma espécie de protozoário digenético (heteroxeno), possuindo, portanto, dois hospedeiros distintos: o barbeiro, inseto da subfamília Triatominae, e um hospedeiro vertebrado mamífero. Diversas alterações estruturais e morfológicas são observadas durante seu ciclo de vida, sendo que diferentes formas evolutivas podem ser associadas a determinados estágios do desenvolvimento do parasita (TYLER & ENGMAN, 2001). Nos hospedeiros vertebrados, no citoplasma de diversos tipos celulares (células do músculo liso, estriado e cardíaco, células do Sistema Nervoso Central, entre outras), são encontradas as formas replicativas amastigotas, que possuem característica esférica ou oval

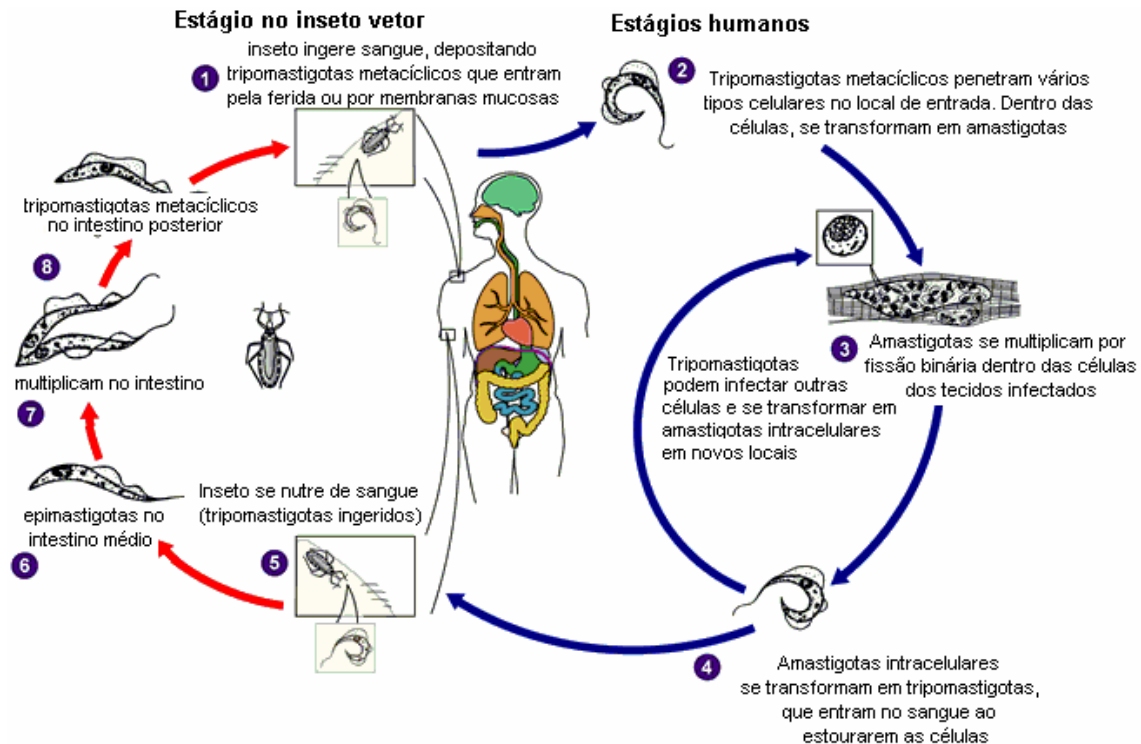


sem flagelo livre; e extracelularmente (no sangue circulante) as formas não replicativas tripomastigotas sanguíneas com morfologia alongada, cinetoplasto na extremidade posterior, membrana ondulante e flagelo livre. Os tripomastigotas sanguíneos apresentam variações morfológicas, como formas delgadas ou largas, denominadas “polimorfismos” que guardam correlações importantes com características fisiológicas do parasita. As formas delgadas do parasita seriam relativamente mais frequentes no início da infecção do hospedeiro vertebrado, quando ainda não existe uma imunidade humoral específica contra o parasita. Gradualmente estas formas seriam substituídas pelas formas largas, menos sensíveis à ação de anticorpos e que, portanto, passariam a predominar na fase mais tardia da infecção, quando a imunidade já se estabeleceu (TYLER & ENGMAN, 2001).

No hospedeiro invertebrado, são encontradas inicialmente formas arredondadas, denominadas esferomastigotas, presentes no estômago e intestino do triatomíneo; também são encontradas formas replicativas epimastigotas (forma oval alongada, com cinetoplasto localizado anteriormente com relação ao núcleo, pequena membrana ondulante e flagelo livre) presentes em todo intestino, e tripomastigotas metacíclicos presentes no reto. O tripomastigota metacíclico constitui a forma mais natural de infecção para o hospedeiro vertebrado (TYLER & ENGMAN, 2001).

O ciclo de vida do *T. cruzi* inicia-se quando um hospedeiro invertebrado ingere sangue de um mamífero infectado. No tubo digestivo do inseto, as formas tripomastigotas sanguíneas ingeridas iniciam sua diferenciação em epimastigotas. No intestino, os epimastigotas se dividem repetidamente por um processo de fissão binária e podem aderir às células do epitélio retal por hemidesmossomas (ZELEDON *et al.*, 1997). Estas formas aderidas sofrem alongamento e torção e transformam-se em tripomastigotas metacíclicos, processo este conhecido como metaciclogênese (revisto por DE SOUZA, 2000).

Durante o repasto sanguíneo, o triatomíneo libera junto com as fezes e urina as formas metacíclicas na pele ou mucosa do vertebrado. Através de descontinuidades dos tecidos, os parasitas podem atingir a corrente sanguínea e infectar outros tipos celulares.



**FIGURA 1.1. CICLO DE VIDA DO *TRYPANOSOMA CRUZI***

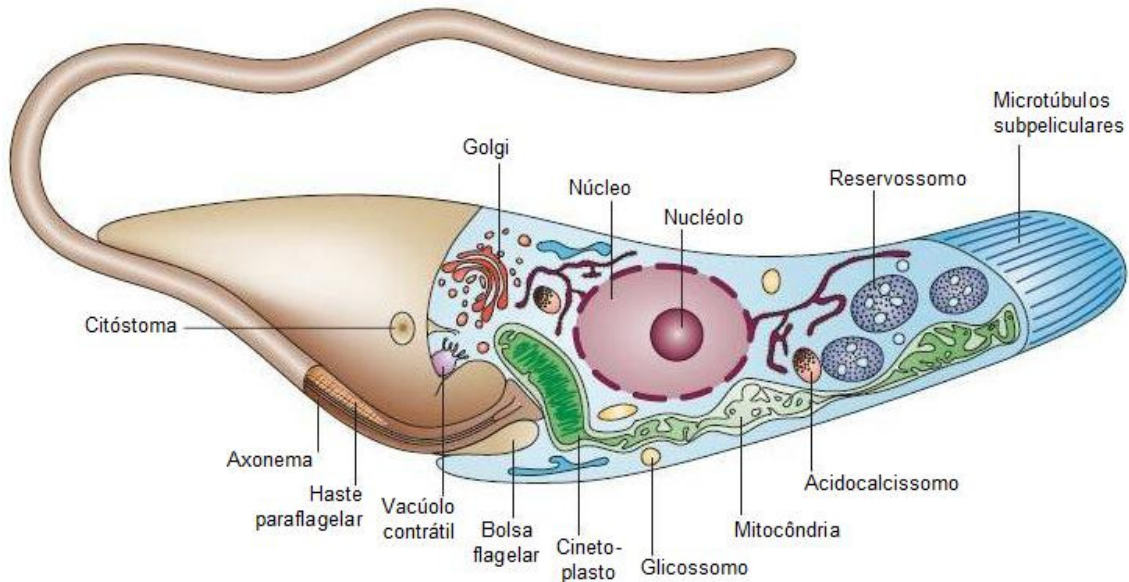
FONTE: modificados de [www.dpd.cdc.gov/dpdx](http://www.dpd.cdc.gov/dpdx).

Os tripomastigotas metacíclicos começam a invasão na célula hospedeira ao aderirem à superfície celular. Após a penetração, ocorre a formação do vacúolo parasitóforo, o qual se funde com lisossomos propiciando a acidificação do vacúolo e diferenciação dos metacíclicos em amastigotas, sendo que simultaneamente a membrana do vacúolo vai sendo destruída por enzimas secretadas pela forma tripomastigota (BURLEIGH & WOOLSEY, 2002). A forma amastigota é liberada no citoplasma da célula hospedeira, iniciando o processo de multiplicação. Posteriormente, as formas amastigotas sofrem diferenciação em tripomastigotas sanguíneas. Com a intensa movimentação do flagelo e devido ao grande número de parasitas, a célula se rompe, liberando os parasitas na corrente sanguínea, os quais podem infectar outras células ou serem ingeridos pelo inseto vetor, dando reinício ao ciclo (revisto por DE SOUZA, 2000), ilustrado na Figura 1.1.

### 1.3 CARACTERÍSTICAS CELULARES DE *T. CRUZI*

A Figura 1.2 mostra a estrutura geral de um epimastigota de *T. cruzi*, indicando diversas estruturas e organelas importantes para a função biológica do parasita. Algumas

dessas estruturas são peculiares ao subgrupo de organismos ao qual *T. cruzi* pertence. Devido à abrangência de um estudo de genômica funcional, iremos revisar algumas dessas estruturas que apresentam um papel importante para esse parasita.



**FIGURA 1.2 – ESTRUTURA CELULAR DE *T. CRUZI* (EPI-MASTIGOTA)**

FONTE: modificado de DOCAMPO *et al.*, 2005.

NOTA: as cores não tem significado biológico, visando apenas maior clareza das estruturas.

Peculiar da ordem Kinetoplastida, faz-se presente no *T. cruzi* uma mitocôndria única e diferenciada, a qual se entende por todo corpo do parasita (PAULIN, 1975). Numa certa porção da matriz mitocôndrial, localizada próximo ao corpúsculo basal, existe um arranjo complexo de fibrilas de DNA, formando uma estrutura conhecida como cinetoplasto, a qual é característica dos cinetoplastídeos. O DNA do cinetoplasto, também conhecido como kDNA, forma uma estrutura em forma de bastão perpendicular ao eixo do flagelo. A posição do cinetoplasto, em relação ao núcleo, varia nas diferentes fases do parasita (anterior em epimastigotas; posterior em tripomastigotas), mas está sempre localizada próximo ao corpo basal. Recentemente, foi demonstrada a presença de filamentos conectando o cinetoplasto ao corpo basal, provendo uma explicação para o posicionamento espacial e segregação do genoma mitocondrial (OGBADOIYI *et al.*, 2003). O kDNA representa cerca de 30% do DNA celular total, sendo composto por moléculas circulares denominadas maxicírculos e minicírculos, as quais são topologicamente relaxadas e interconectadas para formar uma única rede. Existem algumas dezenas de maxicírculos, com tamanho variando entre 20 e 40 kb, os quais contêm genes codificadores de proteínas mitocondriais e RNA ribossômico,

basicamente envolvidos na via de fosforilação oxidativa (SIMPSON *et al.*, 1987; SHLOMAI, 2004), enquanto os minicírculos são encontrados aos milhares, com tamanho variando entre 0,5 e 2,5 kb, e codificam moléculas de RNA guia que participam no processo de edição, através da inserção e, menos frequentemente, deleção de uridinas dos mRNA transcritos a partir dos maxicírculos (MADISON-ANTENUCCI *et al.*, 2002). Apenas nos tripanossomatídeos, o kDNA encontra-se como uma rede complexa de DNA circular condensado, formando uma massa única. Nos bodonídeos, o kDNA pode permanecer difuso pela mitocôndria ou apresentar-se em nódulos distintos (SIMPSON *et al.*, 2002).

O núcleo apresenta-se esférico, com um nucléolo grande e pequenas quantidades de heterocromatina, nas formas replicativas. Ao contrário, nas formas infectivas, o nucléolo desaparece e o núcleo, agora alongado, fica repleto de heterocromatina. Os níveis de compactação da cromatina, observados ao longo do ciclo celular do *T. cruzi*, estão associados aos padrões de transcrição. Assim, em tripomastigotas são observados os níveis mais elevados de compactação da cromatina e mais reduzidos de transcrição (ELIAS *et al.*, 2001). Intrigantemente, o envelope nuclear permanece intacto durante toda a divisão celular do parasita, com aparecimento de microtúbulos intranucleares, dispersão da cromatina e aparecimento de placas densas (SOLARI, 1995).

O flagelo único dos tripanossomatídeos é uma organela essencial, com funções críticas na mobilidade do parasita, morfogênese e divisão celular (revisto por RALSTON & HILL, 2008). Como na maioria dos eucariotos, o flagelo de *T. cruzi* emerge do corpo basal e é formado pelo axonema, composto de microtúbulos no padrão típico 9 + 2, que conduz o movimento flagelar. O flagelo de tripanossomatídeos possuiu uma estrutura interessante designada como haste paraflagelar ou haste paraxial, PFR (*paraflagellar rod*), a qual é essencial para a sobrevivência do parasita. A PFR é composta de um arranjo complexo de filamentos que correm paralelamente ao eixo do flagelo, estando ligados tanto ao axonema quanto à FAZ (descrita abaixo). Um grande número de proteínas formam o PFR, mas a grande maioria ainda não foi caracterizada. O flagelo é envolto por membrana própria, sendo que grande parte do seu comprimento permanece ligada ao corpo celular, principalmente na forma tripomastigota, através da zona de ligação do flagelo, FAZ (*flagellum attachment zone*), uma rede de filamentos e conexões de membrana que unem o flagelo ao longo do corpo celular do parasita. A FAZ é bem caracterizada estruturalmente, mas os componentes são em sua maioria desconhecidos (revisto por RALSTON & HILL, 2008).

A membrana plasmática deste parasita é coberta por um glicocálice delgado composto por diversos glicoconjugados (glicoproteínas, glicolípídeos e lipofosfoglicanos). Uma característica importante é que a maioria desses glicoconjugados são ligados à membrana por âncoras glicosilfosfatidilinositol, sendo que diversas dessas moléculas de superfície estão envolvidas no processo de interação do protozoário com a célula hospedeira (HEIZE *et al.*, 1996). Regiões especializadas da membrana, formando microdomínios distintos, podem ser observadas no parasita (revisto por De SOUZA, 2007), destacando-se a zona de ligação do flagelo, FAZ (revisto por De SOUZA, 2008).

Na face interna da membrana plasmática existe uma camada de microtúbulos, designados microtúbulos subpeliculares, dispostos paralelamente entre si e conectados uns aos outros através de filamentos especializados. Conexões com a membrana plasmática, cisternas periféricas do retículo endoplasmático e componentes da via endocítica também foram observados (SOUTO-PADRON *et al.*, 1984). Os microtúbulos subpeliculares conferem resistência mecânica e rigidez ao parasita (Kohl & Gull, 1998).

Tripanossomatídeos possuem uma organela esférica especializada, com uma matriz homogênea, na qual ocorre a maioria das reações da via glicolítica (OPPERDOES & BORST, 1977), motivo pelo qual tal estrutura foi denominada glicossomo. Outras vias metabólicas também parecem ocorrer no glicossomo, como fixação de dióxido de carbono, salvamento de purinas e biossíntese *de novo* de pirimidinas, alongamento de ácido graxos e síntese de isoprenos (revisto por De SOUZA, 2008), vias estas localizadas no citosol na maioria das células eucarióticas. Todas proteínas glicossomais são codificadas por genes nucleares, sendo sintetizadas no citoplasma e posteriormente importadas devido à presença de um peptídeo sinal (SHIH *et al.*, 1998).

Tripanossomatídeos também possuem uma estrutura vacuolar contendo depósitos elétron-densos, o acidocalcissomo, organela assim denominada devido à sua capacidade de transportar prótons e cálcio. Em epimastigotas de *T. cruzi* tal organela apresenta-se esférica com diâmetro médio de 0,2  $\mu\text{m}$ , podendo variar entre 50 e 600 nm (revisto por DOCAMPO *et al.*, 2005; revisto por De SOUZA, 2008). Os acidocalcissomos realizam importantes funções celulares como: (i) estoque de cálcio, magnésio, sódio, potássio, zinco, ferro, compostos contendo fósforo e especialmente pirofosfato e polifosfato inorgânicos; (ii) homeostase do pH através de uma H<sup>+</sup>-ATPase e (iii) osmorregulação em associação com o vacúolo contrátil (revisto por DOCAMPO *et al.*, 2005). O vacúolo contrátil de *T. cruzi* é formado por diversos túbulos conectados a um vacúolo central próximo à bolsa flagelar. Recentemente, foi

identificada a presença de aquaporina, uma proteína envolvida no transporte de água, em epimastigotas de *T. cruzi*, com localização dupla, acidocalcissomo e vacúolo contrátil (MONTALVETTI *et al.*, 2004). Essas estruturas parecem estar envolvidas na osmorregulação, sendo que a fusão de acidocalcissomos com o vacúolo contrátil ocorre em um processo mediado por cAMP (revisto por ROHLOFF & DOCAMPO, 2008).

Estudos em *T. cruzi* têm demonstrado que este protozoário exibe certas peculiaridades em sua via endocítica que o diferenciam de outros organismos, incluindo tripanossomatídeos (revisto por De SOUZA *et al.*, 2009). A endocitose ocorre de forma significativa somente em epimastigotas, a partir de dois sítios de ingestão de macromoléculas: a bolsa flagelar e o citóstoma. Além disso, o material endocitado é entregue a estruturas não usuais, designadas reservossomos, localizadas na região posterior da célula. A bolsa flagelar, característica da ordem Kinetoplastida, é uma depressão na região anterior da célula, de onde emerge o flagelo. Ela é formada por uma invaginação da membrana plasmática que estabelece continuidade direta com a membrana flagelar, formando um compartimento extracelular especial, de certa forma isolado do meio externo. A membrana da bolsa flagelar é altamente especializada, sendo a única área do parasita que não possui microtúbulos subpeliculares, permitindo intensa atividade endocítica e exocítica (revisto por De SOUZA *et al.*, 2009).

O citóstoma e reservossomo são estruturas encontradas somente em parasitas do subgênero *Schizotrypanum*. O citóstoma aparece como uma invaginação da membrana plasmática e microtúbulos adjacentes que penetram profundamente na célula até a região do núcleo. A abertura dessa estrutura tem um diâmetro de cerca de 0,3 µm, enquanto porções mais profundas (citofaringe) possuem um diâmetro significativamente menor, lembrando um funil. Cerca de 85% da atividade endocítica de epimastigotas de *T. cruzi* ocorre no citóstoma (PORTO-CARREIRO *et al.*, 2000). Epimastigotas incubados com laranja de acridina, um corante fluorescente acidófilo, o acumulam na citofaringe, sugerindo que esse compartimento tem um caráter ácido (revisto por De SOUZA *et al.*, 2009).

Em *T. cruzi*, macromoléculas endocitadas concentram-se em estruturas vesiculares conhecidas como reservossomos, assim denominadas devido à sua grande capacidade de acumular material ingerido. Cada epimastigota possui diversos reservossomos, principalmente na região posterior da célula, envoltos por membrana única e com diâmetro médio de 0,6 µm. Esta estrutura também concentra hidrolases ácidas, como a cruzipaina (CAZZULO *et al.*, 1997) e serina carboxipeptidase, sendo o principal sítio de degradação e regulação protéica neste parasita (revisto por De SOUZA *et al.*, 2009), o que, juntamente com o pH ácido dessa

estrutura, sugere que reservossomos sejam organelas relacionadas aos lisossomos. O pH 6 (SOARES *et al.*, 1992) é mantido por uma H<sup>+</sup>-ATPase do tipo P (VIEIRA *et al.*, 2005), diferentemente de demais organelas ácidas de eucariotos que possuem uma bomba do tipo V (vacuolar). Reservossomos acumulam lipídeos, sendo que éster de colesterol e ergosterol são massivamente concentrados nessa organela (revisto por De SOUZA *et al.*, 2009). Os reservossomos desaparecem gradualmente quando epimastigotas são incubados em meio pobre, condição que dispara a metaciclogênese (revisto por De SOUZA, 2008).

#### 1.4 DOENÇA DE CHAGAS

A doença de Chagas é assim denominada em homenagem ao seu descobridor, o médico brasileiro Dr. Carlos Justiniano Ribeiro das Chagas. Ele fez a descrição do agente etiológico, do transmissor e do modo de transmissão da doença em 1909, quando realizava uma campanha contra a malária que atingia operários que trabalhavam na construção de um trecho da Estrada de Ferro Central do Brasil, na região norte do Estado de Minas Gerais (CHAGAS, 1909). Tal doença é característica do continente americano, motivo pelo qual é chamada tripanossomíase americana, e está atualmente distribuída em 18 países, atingindo México, parte da América Central e toda América do Sul. Há aproximadamente 15 milhões de infectados, e 28 milhões correm risco de contaminação (DIAS, PRATA e CORREIA, 2008).

É uma doença transmissível, principalmente propagada por insetos hematófagos, popularmente conhecidos como barbeiros. Estes insetos, também chamados de triatomíneos, pertencem à ordem Hemiptera, família Reduviidae, subfamília Triatominae. Tal forma de transmissão vetorial é responsável pela grande maioria dos casos registrados (revisto por DIAS, 2000), mas a transmissão através de transfusão de sangue adquiriu uma importância considerável na epidemiologia da doença. Em 1950, Nussensweig e colaboradores demonstraram em São Paulo os primeiros casos de transmissão transfusional da doença. Desde então esse problema vem sendo estudado com interesse e importância crescentes (DIVISÃO DE DOENÇA DE CHAGAS SUCAM/MS, 1989). Recentemente, após o controle da transmissão por *T. infestans* e por transfusão sanguínea no Brasil, a infecção por via oral tem sido o foco de atenção, sendo a forma mais importante e frequente de transmissão (COURA, 2006). Surto de doença de Chagas aguda, ou microepidemias, através de ingestão de comida contaminada com formas metacíclicas têm ocorrido em distintas localidades do

país nos últimos anos, principalmente na região Amazônica, onde outros triatomíneos diferentes de *T. infestans* predominam. Como revisto por Coura e colaboradores (2002), mais de 50% de casos agudos de doença de Chagas registrados na Amazônia entre 1968 e 2000 foram atribuídas a microepidemias de infecção por transmissão oral. Outras formas de transmissão menos comuns são: transmissão materno-fetal, acidentes laboratoriais, transplantes de órgãos e transmissão sexual (COURA, 2007).

A doença de Chagas caracteriza-se por apresentar fases aguda e crônica distintas. A fase aguda pode ser assintomática e desenvolve-se logo após a infecção e breve período de incubação. Quando sintomática, o indivíduo infectado apresenta alta parasitemia, febre, mal-estar, cefaléia, edema subcutâneo, disfunção cardíaca, hepatomegalia e esplenomegalia. Tais sintomas persistem por cerca de dois meses, podendo ocorrer morte por insuficiência cardíaca ou meningoencefalite, principalmente em crianças (2 a 8% dos casos) (COURA, 2007). Neste período inicial, o parasita invade e se multiplica em vários tipos celulares, como macrófagos, células musculares lisas e estriadas, fibroblastos e células nervosas. Especificamente no coração, pode-se observar inflamação severa acompanhada de necrose de células parasitadas e não-parasitadas (revisto por COURA & DE CASTRO, 2002). Degranulação de eosinófilos, patologias microvasculares, edema, trombose, coagulação sanguínea e isquemia também foram demonstrados (ANDRADE, 1999). Diferentemente da infecção pelas fezes do triatomíneo, a infecção por via oral normalmente gera casos agudos fortemente sintomáticos, visto que o número de formas infectivas do parasita que entram em contato com a mucosa gástrica é muito maior (revisto por YOSHIDA, 2009).

Após a infecção aguda, os pacientes normalmente continuam infectados, geralmente apresentando o desenvolvimento de imunidade contra *T. cruzi*. Alguns parasitas escapam do sistema imune do hospedeiro e lesões inflamatórias focais podem ser vistas em vários órgãos. Neste momento, a combinação de aumento de imunidade do paciente contra o parasita com a supressão imunológica específica da hipersensibilidade e redução de reações inflamatórias parecem ser as principais vias que conduzem para a fase indeterminada da doença de Chagas (ANDRADE, 1999).

Depois de um período de anos ou décadas, se não tratados, os casos indeterminados então evoluem para quadros crônicos, geralmente assintomáticos, mas que entre 20 e 50% dos casos, dependendo da área endêmica, são acompanhados por debilitação severa ou morte (revisto por COURA & DE CASTRO, 2003). Durante a fase crônica, uma variedade de estados clínicos pode ser vista, incluindo distúrbios cardíacos (megacárdio), digestivos



(megaesôfago e megacólon) ou distúrbios neurológicos (revisto por BRENER *et al.*, 2000). Esta variedade de sintomas pode estar correlacionada com a diversidade genética, tanto do parasita quanto do hospedeiro, o que provavelmente influencia a história natural da infecção (revisto por STURM & CAMPBELL, 2009).

Os distúrbios cardíacos que se desenvolvem na fase crônica da doença, normalmente denominados como miocardite fibrosa crônica, eram atribuídos a reações auto-imunes devido à hipersensibilidade do hospedeiro vertebrado a antígenos do parasita. Embora tenha sido demonstrada a presença de antígenos de reação cruzada entre células musculares cardíacas e o *T. cruzi*, a auto-imunidade não explica totalmente a forma cardíaca da doença de Chagas. A alta positividade do xenodiagnóstico e hemocultura e a reativação de doença crônica por imunossupressão demonstram a presença do parasita em casos crônicos (revisto por COURA & DE CASTRO, 2002), o que chama atenção para a necessidade de combate contra o parasita no tratamento da doença.

A patogenicidade da doença de Chagas não é completamente definida e compreendida, mas dois mecanismos parecem atuar em conjunto no desenvolvimento da Doença: a persistência do parasita resulta em reações inflamatórias crônicas, o que induz respostas auto-imunes contra tecidos do hospedeiro (revisto por TARLETON, 2001). Ou seja, a presença do parasita e suas interações particulares com o sistema imune do hospedeiro vertebrado induzem a complexa patogenicidade da doença de Chagas.

Desde o final da década de 60 e início da década de 70, duas drogas têm sido usadas no tratamento da doença de Chagas: nifurtimox, um derivado de nitrofurano, e benzonidazol, um derivado de nitroimidazol. Os resultados obtidos com ambas as drogas variam de acordo com a fase da doença, o período de tratamento e a idade e origem geográfica do paciente (revisto por COURA & DE CASTRO, 2002). Em geral, os resultados são bons na fase aguda da doença, principalmente em crianças, com até 80% de eficácia, enquanto na fase crônica os resultados clínicos encontrados não são satisfatórios (10 a 20% de pacientes curados de acordo com diferentes análises). A dose recomendada na fase aguda da doença é 8 a 10 mg/kg/dia para nifurtimox e 5-7,5 mg/kg/dia para benzonidazol, durante 30 a 60 dias consecutivos. Desde a década de 80, nifurtimox teve sua comercialização descontinuada, primeiro no Brasil, depois na Argentina, Chile e Uruguai. Assim sendo, a droga padrão encontrada atualmente no mercado brasileiro para o tratamento da doença de Chagas é o benzonidazol (revisto por COURA & DE CASTRO, 2002).

Nifurtimox e benzonidazol atuam através da formação de radicais livres e metabólitos eletrofilicos. Como os mecanismos celulares de defesa contra o estresse oxidativo são defectivos em *T. cruzi*, tais radicais livres são responsáveis pelo efeito tripanocida destas drogas (revisto por MAYA *et al.*, 2007).

Embora nifurtimox e benzonidazol sejam tripanocidas para todas as formas evolutivas do parasita (revisto por COURA & DE CASTRO, 2002), tais drogas podem causar toxicidade sistêmica e diversos efeitos colaterais: anorexia, náusea, vômitos, cefaléia, depressão do sistema nervoso central ou sintomas psicóticos, dermatites e parestesia (revisto por MAYA *et al.*, 2007). Outro detalhe que complica o tratamento farmacológico da doença de Chagas com nifurtimox e benzonidazol, contribuindo para a baixa efetividade clínica das mesmas, é a diferente suscetibilidade de distintas cepas do parasita a tais drogas (FILARDI & BRENER, 1987; BRENER *et al.*, 1993). Além disso, como já citado, estas drogas não são capazes de eliminar o parasita durante a fase crônica da doença.

Seguindo os requerimentos da organização Mundial da Saúde, a droga ideal para o tratamento da doença de Chagas deve completar os seguintes critérios: (i) promover cura parasitológica na fase aguda e crônica da doença, (ii) ser efetiva em uma ou poucas doses, (iii) ser acessível financeiramente, (iv) não possuir efeitos colaterais ou teratogênicos, (v) não necessitar hospitalização para o tratamento e (vi) não induzir resistência (WHO, 1997). Essa droga ideal não existe, e provavelmente levará um longo período para ser desenvolvida, tendo em vista a falta de interesse das indústrias farmacêuticas em relação a doenças parasitárias. Dessa forma, o controle da doença vem sendo realizado principalmente através do controle vetorial (eliminação de *Triatoma infestans* por pesticidas), o que possibilitou um decréscimo do número de pessoas infectadas nos últimos anos. Entretanto, o fato de ser uma zoonose elimina a possibilidade de erradicação da doença, visto que várias espécies silvestres atuam como reservatório natural e outras espécies de triatomíneos podem transmitir a doença. Portanto, estratégias visando o controle da doença de Chagas devem focar no bloqueio da infecção ou no tratamento dos pacientes cronicamente afetados (COURA, 2007). Além da falta de medicamentos, não existem vacinas eficazes disponíveis para a doença de Chagas (revisto por DUMONTEIL, 2009).

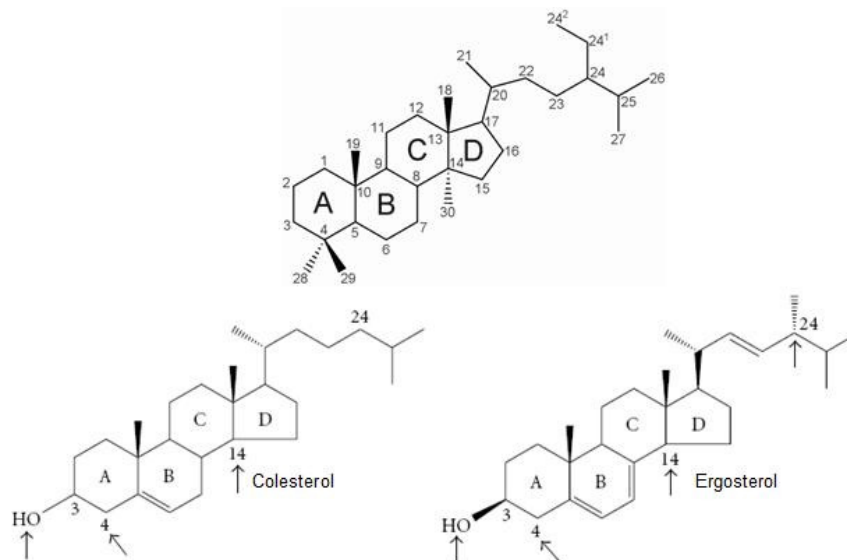
Recentes desenvolvimentos no entendimento da bioquímica básica do *T. cruzi* têm permitido a identificação de alvos potenciais para o estudo de novas drogas tripanocidas. Dentro desse contexto, destacam-se as drogas inibidoras da biossíntese do ergosterol, as quais encontram-se sob desenvolvimento experimental relativamente avançado (revisto por

URBINA, 2009a). Outras agentes tripanocidas atualmente em desenvolvimento são: inibidores de cisteína-protease (cruzipaína); compostos que interferem com o salvamento de purinas; inibidores do metabolismo de pirofosfato; inibidores do metabolismo de inositol e inibidores do metabolismo e síntese de tripanotiona (revisto por URBINA, 2009b).

## 1.5 A BIOSÍNTESE DE ESTERÓIS COMO ALVO QUIMIOTERÁPICO DA DOENÇA DE CHAGAS

### 1.5.1 Esteróis: estrutura, funções e biossíntese em células eucarióticas

Esteróis são lipídeos estruturais presentes na membrana da maioria das células eucarióticas. São moléculas derivadas de isoprenos com a característica comum de apresentarem um núcleo esteróide composto de quatro anéis carbônicos fusionados, três deles com cinco átomos de carbono e um com quatro (Figura 1.3). Dependendo dos organismos analisados, duplas ligações distintas podem estar presentes no núcleo esteróide, ou grupos químicos laterais distintos podem ser encontrados ligados ao núcleo esteróide (na posição C-17), o que diferencia os esteróis: colesterol em animais, ergosterol em fungos e protistas e sitosterol, entre outros, em plantas (SONG & NES, 2007). Estruturalmente, colesterol e ergosterol são muito similares, sendo que algumas características de ambos são essenciais para suas atividades em membranas celulares: o grupamento  $3\beta$ -OH e a ausência de grupamentos metil nas posições C-4 e C-14 são obrigatórios para o crescimento celular. Além disso, as características únicas do ergosterol (veja Figura 1.3) são essenciais para a viabilidade celular de fungos e tripanossomatídeos (revisto por DE SOUZA & RODRIGUES, 2009).

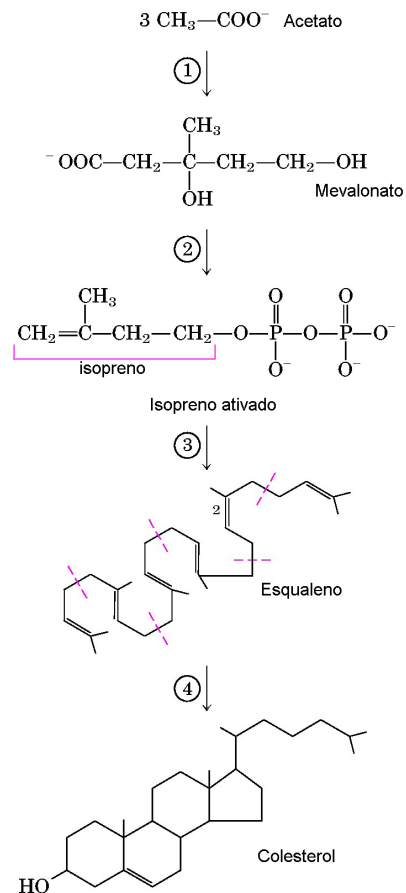


**FIGURA 1.3 – ESTRUTURA QUÍMICA DOS ESTERÓIS**

LEGENDA: na parte superior da figura está representado a estrutura geral dos esteróis, com a nomenclatura química dos átomos de carbono; os quatro anéis carbônicos do núcleo esteróide são nomeados de A – D. Abaixo, é evidenciado a estrutura do colesterol (principal esterol de mamíferos) e ergosterol (principal esterol de fungos e protozoários); as setas indicam partes das moléculas essenciais para a funcionalidade destas moléculas.

FONTE: modificado de De Souza & Rodrigues, 2009.

Esteróis possuem funções essenciais em eucariotos em geral. São componentes integrais da bicamada lipídica estando diretamente relacionados com a regulação da fluidez e permeabilidade da membrana plasmática (DEMEL & DE KRUYFF, 1976). Funções regulatórias também têm sido atribuídas aos esteróis em vários sistemas; como exemplo, foi demonstrado que a ligação covalente do colesterol a Hedgehog (uma molécula sinalizadora) está envolvida no desenvolvimento embrionário de vertebrados (GUY, 2000). Além disso, esteróis são precursores de compostos com alta atividade fisiológica, como os brassinosteróides, uma importante classe de hormônios envolvidos no crescimento e desenvolvimento de plantas superiores (LI *et al.*, 1996), e hormônios esteróides em vertebrados (NELSON & COX, 2003). Em *T. cruzi*, os esteróis endógenos parecem ter função direta na viabilidade celular e na regulação da atividade de enzimas de membrana (CONTRERAS, VIVAS e URBINA, 1997).



#### FIGURA 1.4 – AS QUATRO ETAPAS DA BIOSÍNTESE DE ESTERÓIS

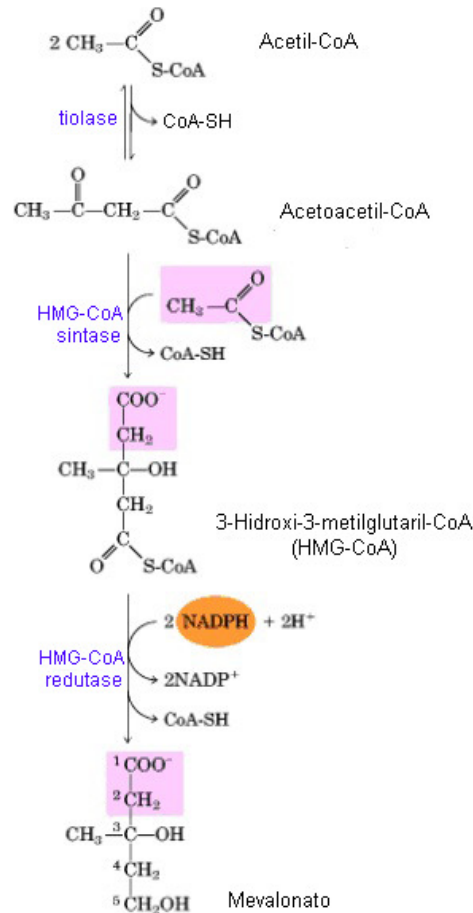
As unidades de isoprenos na molécula de esqualeno são separadas por linhas pontilhadas vermelhas. É colocado como exemplo a síntese de colesterol;

FONTE: modificado NELSON & COX, 2005.

O esqueleto das moléculas de esteróis são derivados de acetil-coenzima A, com exceção do grupo metil em C-24 da cadeia lateral do ergosterol. A via de síntese pode ser didaticamente dividida em quatro etapas, como ilustrado na Figura 1.4 acima (BLOCH, 1965; revisto por NELSON & COX, 2005). A seguir, cada uma das etapas apresentadas na Figura 1.4 é descrita detalhadamente:

**Via do mevalonato:** condensação de três moléculas de acetato para formar um intermediário de seis carbonos denominado mevalonato (detalhes na Figura 1.5; passo 1 da Figura 1.4). Primeiramente duas moléculas de AcetilCoA são condensadas através da ação da enzima tiolase para formar Acetoacetil-CoA. Esta, por sua vez, é condensada novamente com AcetilCoA, pela enzima HMG-CoA sintase, para formar o composto de seis carbonos conhecido como 3-hidroxi-3-metilglutaril-CoA (HMG-CoA). Finalmente, HMG-CoA é

reduzido a mevalonato pela enzima HMG-CoA redutase. Esse conjunto de reações é conhecido como via do mevalonato.



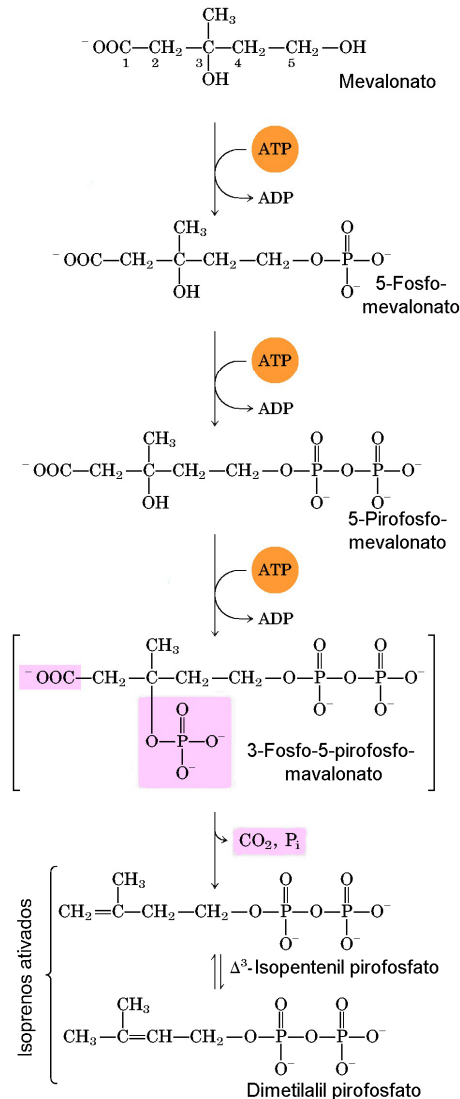
**FIGURA 1.5 – VIA DO MEVALONATO**

A origem dos carbonos C-1 e C-2 do mevalonato, a partir de acetil-CoA, são mostrados em rosa. As enzimas da via são nomeadas em azul.

FONTE: NELSON & COX, 2005.

**Conversão do mevalonato para unidades ativadas de isoprenos (detalhes na Figura 1.6, passo 2 Figura 1.4):** três grupos fosfato são transferidos sequencialmente ao mevalonato pela ação das enzimas mevalonato 5-fosfotransferase, fosfomevalonato quinase e fosfomevalonato descarboxilase. No próximo passo, o grupamento fosfato ligado ao C-3 do mevalonato é retirado da molécula, juntamente com o grupo carboxil C-1, pela ação da enzima pirofosfomevalonato descarboxilase, produzindo uma dupla ligação na molécula de cinco carbonos restante, denominada  $\Delta^3$ -isopentenil pirofosfato. A isomerização deste composto

produz o segundo isopreno ativado utilizado na formação de esteróis, dimetilalil pirofosfato (ver Figura 1.6).



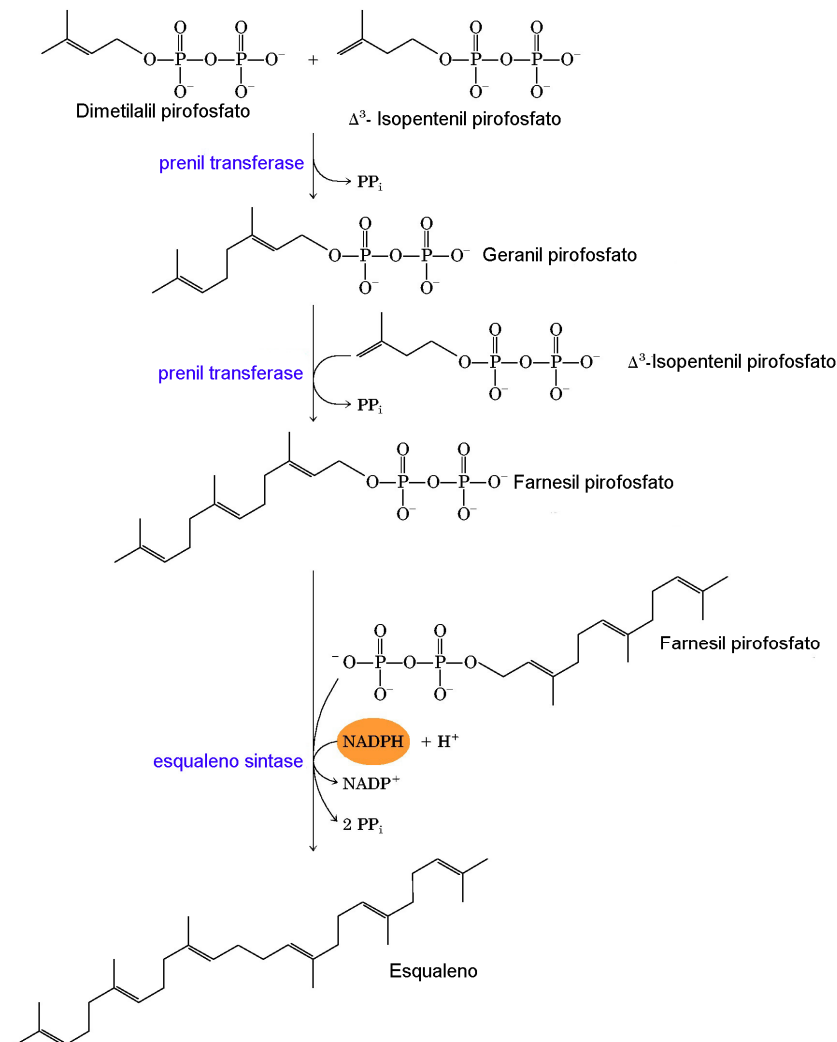
**FIGURA 1.6 – CONVERSÃO DO MEVALONATO A UNIDADES ATIVADAS DE ISOPRENO**

Os grupos retirados da molécula de 3-fosfo-5-pirofosfomevalonato, para formar o isopreno ativado, são sombreados em rosa. O intermediário em colchetes é hipotético.

FONTE: NELSON & COX, 2005.

**Polimerização de 6 subunidades isoprenos ativadas, cada uma com cinco carbonos, para formar a molécula linear de esqualeno com 30 carbonos:** as duas moléculas de isopreno produzidas na etapa anterior passam por uma condensação tipo cabeça-cauda, catalisada pela enzima prenil transferase, sendo que um grupo pirofosfato é deslocado, formando, portanto, uma cadeia de 10 carbonos denominada geranyl pirofosfato. A mesma enzima realiza uma nova condensação do tipo cabeça-cauda, unindo geranyl pirofosfato com mais uma molécula

de isopreno ativado produzindo o intermediário de 15 carbonos farnesil pirofosfato. Finalmente, duas moléculas de farnesil pirofosfato são unidas cabeça-cabeça, pela ação da enzima esqualeno sintase, para formar a molécula de esqualeno (veja Figura 1.7). Todas as etapas envolvidas na formação de farnesil pirofosfato, a partir do mevalonato (Figura 1.6 e Figura 1.7), são conhecidas como via dos isoprenos, a qual está presente em todos organismos investigados. Além de base para formação de esteróis, o isopreno farnesilo pirofosfato é o substrato para a síntese de diversos compostos essenciais para as células, incluindo ubiquinonas, dolicol, grupamento heme A e prenilação de proteínas (revisto por DE SOUZA & RODRIGUES, 2009).



**FIGURA 1.7 – FORMAÇÃO DO ESQUALENO**

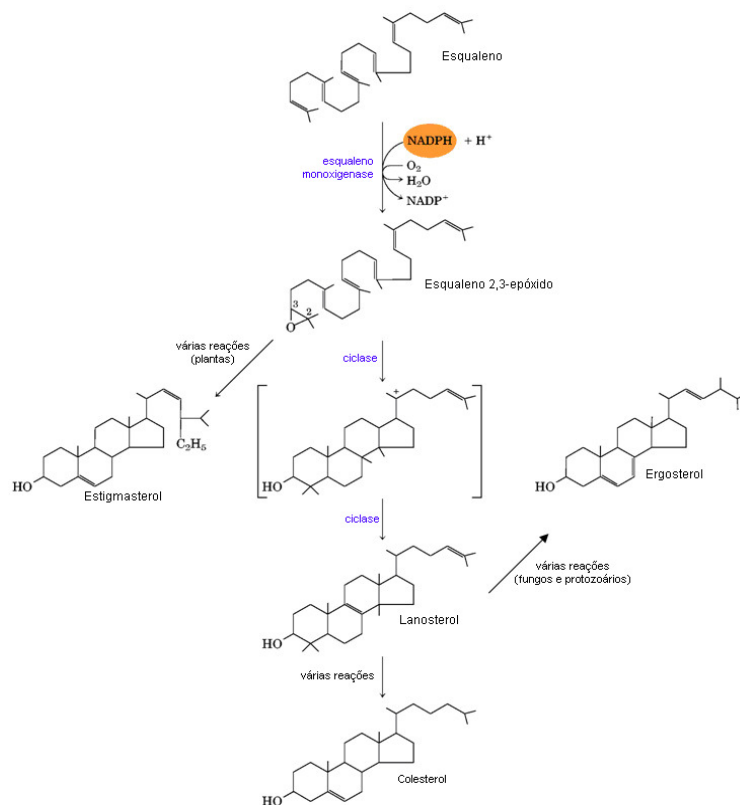
As enzimas da via são nomeadas em azul.

NOTA: a estrutura de 30 carbonos do esqualeno é formada através de sucessivas condensações de unidades de isoprenos ativadas de 5 carbonos.

FONTE: NELSON & COX, 2005.



**Ciclização do esqualeno para formar os quatro anéis do núcleo esteróide, seguido de uma série de modificações (oxidações, remoção ou migração de grupos metil) para produzir o esterol maduro:** quando a molécula de esqualeno é representada como na Figura 1.8, a relação de sua estrutura linear com a estrutura cíclica do núcleo esteróide fica clara. Todos esteróis possuem os quatro anéis fusionados que formam o núcleo esteróide, e todos são alcoóis com uma hidroxila no C-3, por isso o nome “esterol”. A enzima esqualeno monooxigenase adiciona um átomo de oxigênio proveniente de  $O_2$  à extremidade da cadeia do esqualeno, formando esqualeno 2,3 epóxido, enquanto NADPH reduz o outro átomo de oxigênio do  $O_2$  à  $H_2O$  (veja Figura 1.8). As duplas ligações do epóxido formado posicionam-se de maneira que a molécula linear de esqualeno 2,3 epóxido pode ser facilmente reestruturada para formar os anéis do núcleo esteróide. Em animais, fungos e protozoários tal reação de ciclização produz o esterol lanosterol; enquanto em plantas é formado cicloartenol. Lanosterol é então transformado em colesterol (em animais) ou ergosterol (em fungos e protozoários) numa sequência de pelo menos 20 reações.



**FIGURA 1.8 – FORMAÇÃO DO NÚCLEO ESTERÓIDE**

NOTA: até a formação de esqualeno 2,3-epóxido, a via de síntese e a mesma em todos eucariotos. A ciclização desta molécula e seguintes modificações químicas para formação do esterol maduro são diferenciadas em distintos organismos.

FONTE: modificado de NELSON & COX, 2005.

A via de síntese de esteróis de todos eucariotos compartilha os mesmos passos até a formação de esqualeno 2,3-epóxido, ponto no qual a via diverge de forma a produzir diferentes esteróis em distintos organismos (veja Figura 1.8). A via pode divergir em rotas que são filogeneticamente distintas, baseadas na produção do intermediário lanosterol (animais e fungos) ou cicloartenol (plantas) (SONG & NES, 2007).

Após a formação do lanosterol, diversas transformações sequenciais ocorrem para formar colesterol em mamíferos e ergosterol em fungos e protozoários. Algumas dessas reações são comuns a todos eucariotos, enquanto outras são exclusivas a certos organismos, com diferenças mesmo entre tripanossomatídeos e fungos (revisto por DE SOUZA & RODRIGUES, 2009). A seguir serão discutidas as peculiaridades da via de síntese de esteróis em *T. cruzi*.

### 1.5.2 Esteróis em *Trypanosoma cruzi*: síntese, funções e inibidores

Tripanossomatídeos possuem muitas similaridades com fungos em relação à composição celular de esteróis e sua biossíntese. Os principais esteróis destes protozoários são compostos  $\Delta^{5,7}$  de 28 carbonos, C28 ergostano (grupo 24-metil), ou de 29 carbonos, C29-estigmasta (grupo 24-etil) (revisto por ROBERTS *et al.*, 2003). A forma epimastigota de *T. cruzi* contém principalmente ergosterol (ergosta-5,7,22-trien-3 $\beta$ -ol) e ergosta-5,7-dien-3 $\beta$ -ol, os quais juntos perfazem cerca de 40% do total de esteróis, e uma apreciável quantidade (~30%) de estigmasta-5,7-dien-3 $\beta$ -ol e estigmasta-5,7,22-trien-3 $\beta$ -ol (URBINA *et al.*, 1995) (veja Figura 1.9). Quando cultivado em meio contendo soro, colesterol representa cerca de 25% dos esteróis de membrana de epimastigotas de *T. cruzi*. Se adicionado acetato ou mevalonato marcado radioativamente ao meio, os esteróis maduros (4-desmetilados) endógenos passam a apresentar radioatividade, demonstrando que a síntese *de novo* de esteróis ocorre independentemente da presença de colesterol no meio (URBINA *et al.*, 1991; revisto por BUCKNER, 2008).

Os esteróis da fase replicativa amastigota de *T. cruzi* também foram caracterizados (LIENDO *et al.*, 1999). O esterol exógeno originário do hospedeiro (colesterol) representa aproximadamente 80% do peso total de esteróis desta forma do parasita. Os principais esteróis endógenos, em ordem de concentração, são: ergosta-7-en-3 $\beta$ -ol, 24-etil-7-en-colesta-3 $\beta$ -ol e ergosta-7,24(241)-dien-3 $\beta$ -ol (ou episterol), o que indica a ausência de atividade da enzima

$\Delta 5$ -desaturase, a qual adiciona uma dupla ligação no C5 do núcleo esteróide nos esteróis da forma epimastigota (veja Figura 1.9). Embora o parasita incorpore esteróis do hospedeiro, principalmente na forma amastigota, ele se mantém altamente suscetível a inibidores da via de síntese de esteróis. De fato, os estágios intracelulares do parasita são muito mais sensíveis aos inibidores da biossíntese dos esteróis (IBEs) em relação aos epimastigotas (veja Tabela 1.1 e Tabela 1.2) (revisado por URBINA, 2001). Tal suscetibilidade elevada de formas amastigotas pode estar relacionada a um menor *pool* pré-formado de esteróis e/ou a uma maior taxa de renovação de esteróis nessa forma do parasita (LIENDO *et al.*, 1999).

Os detalhes da via biossintética para produção de esteróis C28 e C29 em *T. cruzi* têm sido deduzidos por (i) análises bioquímicas dos esteróis constituintes do parasita, (ii) estudos de marcação com [14C]-acetato e [14C]-mevalonato e (iii) identificação de esteróis intermediários que são acumulados após tratamento das células com distintos IBEs (revisado por ROBERTS *et al.*, 2003). Com base nestes estudos, construiu-se o esquema apresentado na Figura 1.9, o qual apresenta os principais esteróis da via, as enzimas que catalisam os diversos passos biossintéticos e os respectivos IBEs já utilizados em *T. cruzi*. O principal substrato da enzima 14 $\alpha$ -desmetilase é lanosterol (LEPESHEVA *et al.*, 2006), de forma que o passo inicial preferencial da via lanosterol-esterol maduro (4-desmetilado) é a remoção do radical metil do C14 (veja Figura 1.9) (URBINA *et al.*, 1995; LIENDO *et al.*, 1999). Subsequentemente são removidos os dois grupos metil de C4, formando o intermediário zimosterol. Essa ordem preferencial de reações difere da maioria dos fungos e muitos protozoários nos quais lanosterol é primeiramente metilado em C-24 e desmetilado em C-4 para então ser desmetilado em C-14 (SONG & NESS, 2007). Em *T. cruzi*, zimosterol é então metilado em C24, formando fecosterol, o qual então sofre a ação da enzima C-8 esterol isomerase para formar episterol. A partir deste composto diferentes modificações são realizadas em epimastigotas e amastigotas para formar o conjunto distinto de esteróis maduros destas fases do parasita (veja Figura 1.9). A ausência completa de esteróis endógenos  $\Delta 5,7,22$  (como ergosterol) em amastigotas, os principais lipídeos neutros encontrados em epimastigotas (URBINA *et al.*, 1995), indica que a via de biossíntese de esteróis é significativamente mais simples nas formas intracelulares do parasita, com ausência de atividade de  $\Delta 5$  e  $\Delta 22$  dessaturase (LIENDO *et al.*, 1999), o que pode ser evidenciado na Figura 1.9. Além disso, a presença de desmosterol (precursor de colesterol em mamíferos) e esteróis  $\Delta 5$  em amastigotas, como ergosta-5-en-3 $\beta$ -ol e seu análogo 24-etilado, indica a ocorrência de

metilação direta de desmosterol, em C24, na forma amastigota do parasita, visto que esteróis  $\Delta 5$  não são normalmente sintetizados em tripanossomatídeos (LIENDO *et al.*, 1999).

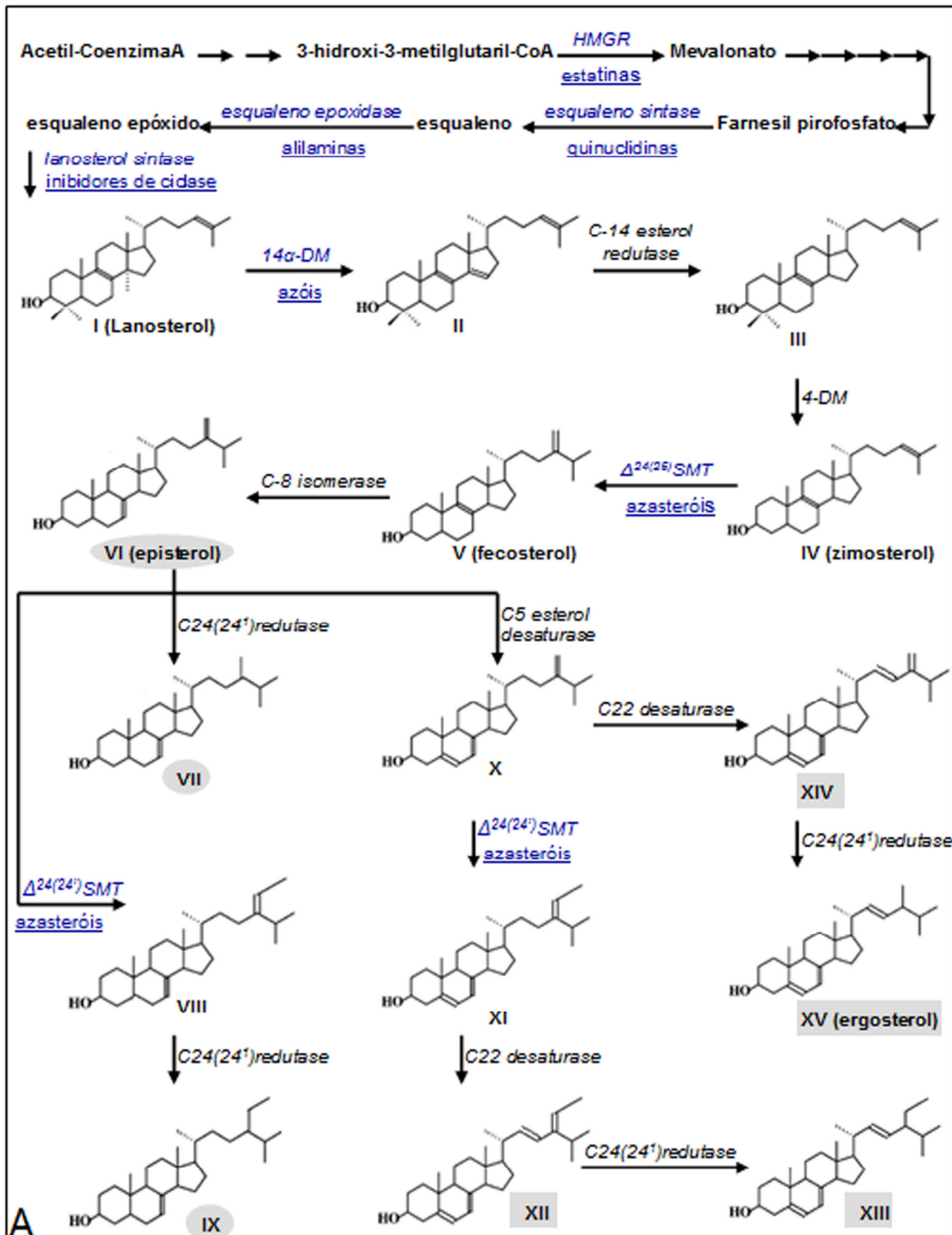
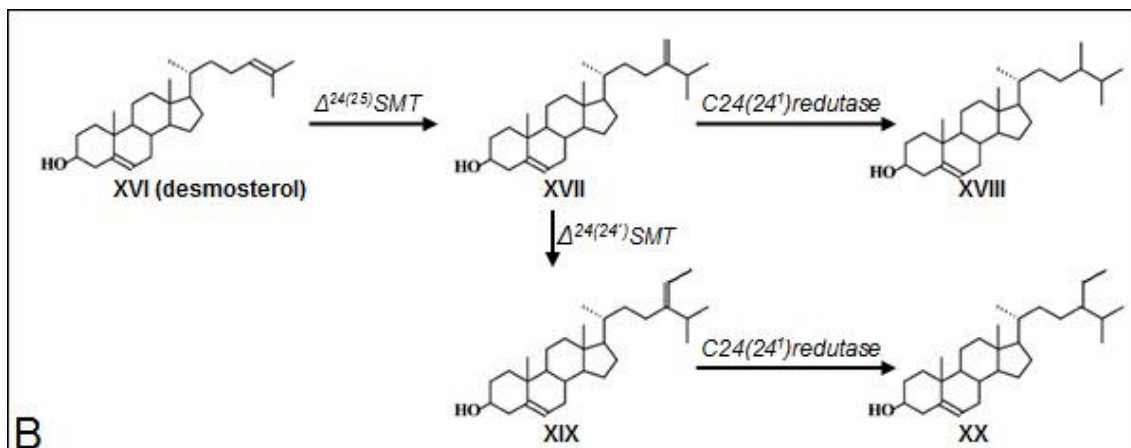


FIGURA 1.9 – VIA DE SÍNTESE DE ESTERÓIS EM *TRYPANOSOMA CRUZI*

CONTINUA



**FIGURA 1.9 – VIA DE SÍNTESE DE ESTERÓIS EM *TRYPANOSOMA CRUZI***

**FIM**

A – principais esteróis sintetizados em amastigotas (sobre elipse cinza) e epimastigotas (sobre quadrado cinza) de *T. cruzi*. B – metilação de esteróis do hospedeiro vertebrado em amastigotas (detalhes no texto). Cada seta representa um passo enzimático, com respectivas enzimas denominadas em itálico. Enzimas testadas como alvo quimioterápico estão destacadas em azul, com respectivas drogas inibitórias sublinhadas (em azul); o substrato e produto das principais reações estão evidenciados. I: 4,4,14 $\alpha$ -trimetilcolesta-8,24-dien-3 $\beta$ -ol (lanosterol); II: 4,4-dimetilcolesta-8,24-dien-3 $\beta$ -ol; III: 4 $\beta$ -metilcolesta-8,24-dien-3 $\beta$ -ol; IV: colesta-8,24-dien-3 $\beta$ -ol (zimosterol); V: Ergosta-8,24(241)-dien-3 $\beta$ -ol (fecosterol); VI: Ergosta-7,24(241)-dien-3 $\beta$ -ol (episterol); VII: Ergosta-7-en-3 $\beta$ -ol; VIII: Stigmasta-7,24(241)-dien-3 $\beta$ -ol; IX: Stigmasta-7-en-3 $\beta$ -ol; X: Stigmasta-5,7,24(241)-trien-3 $\beta$ -ol; XI: Stigmasta-5,7-dien-3 $\beta$ -ol; XII: Stigmasta-5,7,22-trien-3 $\beta$ -ol; XIII: Ergosta-5,7,24(241)-trien-3 $\beta$ -ol; XIV: Ergosta-5,7-dien-3 $\beta$ -ol; XV: Ergosta-5,7,22-trien-3 $\beta$ -ol (ergosterol); XVI: colesta-5,24-dien-3 $\beta$ -ol (desmosterol); XVII: Ergosta-5,24(241)-dien-3 $\beta$ -ol; XVIII: Ergosta-5-en-3 $\beta$ -ol; XIX: Stigmasta-5,24(241)-dien-3 $\beta$ -ol; XX: Stigmasta-5-en-3 $\beta$ -ol. HMGR: 3-hidroxi-3-metilglutarilCoA redutase; 14 $\alpha$ -DM: 14 $\alpha$ -desmetilase; 4-DM: 4-desmetilase; SMT: esterol metiltransferase; FPPS: farnesil pirofosfato sintase; GPP: geranyl pirofosfato; IPP: isopentenil pirofosfato.

Um dos mais importantes estágios da síntese de ergosterol, que não existe na síntese de colesterol, é a adição de um grupo metil na posição C-24 da cadeia lateral do esterol, o qual é transferido de S-Adenosil-L-metionina para esteróis insaturados na posição 24 ( $\Delta$ 24) (veja Figura 1.9). Essa metilação é catalisada pela enzima S-Adenosil-L-metionina: $\Delta$ 24-esterol metiltransferase (24-SMT), enzima esta presente em plantas, fungos e protozoários, mas ausente na via de síntese de esteróis de mamíferos. Portanto, 24-SMT deve representar um interessante alvo para o desenvolvimento de agentes anti-tripanosomais (revisto por DE SOUZA & RODRIGUES, 2009).

Baseado principalmente em estudos com leveduras, tem sido proposto que esteróis 24-alquilados devem desempenhar uma função metabólica na regulação do ciclo celular, de maneira que uma quantidade mínima de esteróis endógenos, com características químicas específicas, deve ser necessária para que a célula mantenha suas características estruturais normais e mantenha a normalidade do ciclo celular (revisto por ROBERTS *et al.*, 2003). De fato, um estudo das características alteradas da membrana plasmática de *T. cruzi* na presença de diferentes IBEs evidenciou um papel regulatório dos esteróis endógenos na atividade de

enzimas da membrana do parasita (CONTRERAS, VIVAS e URBINA, 1997). Demonstrou-se que um dos efeitos bioquímicos da depleção dos esteróis endógenos de epimastigotas de *T. cruzi* é uma marcada alteração no conteúdo da cabeça polar de fosfolipídeos e ácidos graxos de membrana devido a alterações na atividade de enzimas de membrana. Tais características, juntamente com o acúmulo de precursores de esteróis, devem modificar as características físico-químicas da matriz lipídica da membrana, o que, por sua vez, deve afetar a atividade de outras proteínas ligadas a membrana levando à inviabilidade celular (CONTRERAS, VIVAS e URBINA, 1997). Nas células tratadas com IBEs demonstrou-se uma diminuição de cerca de 50% na atividade da  $Mg^{++}$ -ATPase de membrana, a qual parece ser responsável pela manutenção de um potencial através da membrana de cinetoplastídeos, potencial este crucial para o transporte de íons e outras substâncias através da membrana. Assim sendo, a atividade reduzida desta ATPase em parasitas tratados com IBEs pode levar a drásticas alterações no conteúdo iônico e no balanço osmótico das células (CONTRERAS, VIVAS e URBINA, 1997). De fato, estudos das alterações ultra-estruturais induzidas por diferentes IBEs mostram que uma característica comum é o arredondamento celular e inchaço mitocondrial (LAZARDI, URBINA e DE SOUZA, 1990; revisto por ROBERTS *et al.*, 2003), o que pode estar relacionado com alterações osmóticas no parasita. Outras alterações estruturais comuns são padrões anormais do sistema de endomembranas, inclusões lipídicas, bolhas na membrana plasmática, aumento do vacúolo contrátil, núcleos ou cinetoplastos múltiplos (ciclo celular aberrante), condensação da cromatina e alterações no citoesqueleto (revisto por DE SOUZA & RODRIGUES, 2009). Portanto, os esteróis endógenos 24-alquilados de *T. cruzi* parecem ter função essencial em um número significativo de processos celulares, incluindo ciclo celular, morte celular e manutenção da estrutura, estabilidade e função da membrana plasmática (CONTRERAS, VIVAS e URBINA, 1997; revisto por DE SOUZA & RODRIGUES, 2009).

Em mamíferos, os passos iniciais da síntese de esteróis ocorrem principalmente no citosol, enquanto os passos finais devem ocorrer no retículo endoplasmático. Em fungos e tripanossomatídeos, duas das enzimas da via do mevalonato (HMG-CoA sintase e HMG-CoA redutase) encontram-se na mitocôndria. A clonagem, seqüenciamento e expressão heteróloga de 3-hidroxi-3-metilglutarilCoA redutase (HMGR) em *T. cruzi* mostraram que esta enzima não possuiu o domínio N-terminal de membrana sendo, portanto, uma proteína solúvel (PEÑA-DIAZ *et al.*, 1997). Estudos da localização celular de HMGR de *T. cruzi* são um pouco controversos: através de técnicas de fracionamento celular evidenciou-se que tal

enzima está principalmente associada à matriz de glicossomos (~80% da atividade) (CONCEPCION *et al.*, 1998); entretanto, estudos mais recentes utilizando as técnicas de microscopia imunoeletrônica, com anticorpo monoclonal para HMGR, e permeabilização celular com concentrações crescentes de digitonina (PEÑA-DIAZ *et al.*, 2004), evidenciaram que HMGR está principalmente associado à mitocôndria em *T. cruzi* e *Leishmania major*, sendo que a sequência amino-terminal é responsável pelo direcionamento para a matriz mitocondrial. Além disso, a clonagem e expressão de HMGR de *T. cruzi* em levedura duplo mutante para HMGR (*hmg1* e *hmg2*) possibilitou o crescimento normal das células na ausência de mevalonato, evidenciando que HMGR recombinante de *T. cruzi* é funcional em *Saccharomyces cerevisiae* sendo eficientemente encaminhada para a mitocôndria, demonstrando um mecanismo de importação conservado entre estes organismos (PEÑA-DIAZ *et al.*, 2004).

Em vertebrados, leveduras e plantas, a enzima esqualeno sintase (SQS) (primeira reação comprometida com a síntese de esteróis; veja Figura 1.7) é uma proteína conservada de membrana exclusivamente associada com o retículo endoplasmático, juntamente com demais enzimas das etapas finais da via de síntese de esteróis (THOMPSON, 1998). Já em *T. cruzi*, SQS é uma proteína de membrana principalmente associada a glicossomos e mitocôndria, com alguma presença no retículo endoplasmático (URBINA *et al.*, 2002), a mesma localização dupla (glicossomo/mitocôndria) encontrada para HMGR (CONCEPCION *et al.*, 1998). Entretanto, tal distribuição parece ser fortemente dependente da fase de crescimento das culturas analisadas, estando provavelmente relacionada com alterações no metabolismo energético do parasita (URBINA *et al.*, 2002). Foi detectado um enriquecimento de esteróis endógenos maduros na membrana mitocondrial interna (RODRIGUES *et al.*, 2001) e na membrana de glicossomos (QUINONES *et al.*, 2004), o que condiz com a possível localização distinta da via de síntese de ergosterol em *T. cruzi*. Estudos em *L. major* (Jiménez-Jiménez, 2008) e *T. cruzi* (PRETI, 2007) mostram que a enzima esterol metiltransferase, a qual catalisa etapas finais na via de síntese do ergosterol (Figura 1.9), está associada principalmente ao retículo endoplasmático, embora dados recentes também tenham detectado tal proteína nos reservossomos de *T. cruzi* (SANT'ANNA *et al.*, 2009) localização esta também observada para 14 $\alpha$ -desmetilase (LEPESHEVA *et al.*, 2010). Assim sendo, intrigantemente, parte da síntese de esteróis em *T. cruzi* parece ocorrer na mitocôndria e/ou glicossomo, diferente da maioria dos eucariotos, nos quais ocorre no citosol/retículo endoplasmático.

Devido ao fato da necessidade intrínseca de *T. cruzi* por esteróis endógenos para crescimento e sobrevivência celular e diferenças entre as enzimas do parasita e de humanos, a via de síntese de esteróis é considerada um dos alvos mais atrativos para o tratamento específico da doença de Chagas (revisto por URBINA, 2009). Dessa forma, paralelamente a estudos com derivados de azoles, a atividade tripanocida e o mecanismo de ação de diversos compostos estão sob intensa investigação. As enzimas já testadas como alvo para inibição da biossíntese de esteróis em *T. cruzi* são (veja Figura 1.9): HMGC<sub>o</sub>A redutase (URBINA *et al.*, 1993), farnesil pirofosfato sintase (FPPS) (MARTIN *et al.*, 2001), esqualeno sintase (URBINA *et al.*, 2002), esqualeno epoxidase (URBINA *et al.*, 1988), lanosterol sintase (também conhecida como oxidoesqualeno ciclase) (BUCKNER *et al.*, 2001), 14 $\alpha$ -desmetilase (revisto por BUCKNER, 2008) e esterol 24-metiltransferase (MAGARACI *et al.*, 2003). Tais enzimas têm sido estudadas individualmente ou em combinação, com diferentes graus de sucesso, sendo que diferentes passos da via podem ser inibidos sinergicamente (MALDONADO *et al.*, 1993; URBINA *et al.*, 1995).

Devido ao uso de cetoconazol e lovastatina no presente trabalho, a utilização desses dois tipos de IBEs contra *T. cruzi*, tanto em experimentos *in vitro* e *in vivo* como clinicamente no tratamento do doença de Chagas, será revisado em mais detalhes nos tópicos seguintes.

#### 1.5.2.1 Inibidores de 14- $\alpha$ demetilase: azóis

Azol é o nome genérico dado a compostos cíclicos de cinco moléculas que contenham pelo menos um átomo diferente de carbono (nitrogênio, enxofre ou oxigênio) (Eicher & Hauptmann, 2003). Os imidazóis possuem um anel heterocíclico com dois átomos de nitrogênio nas posições 1 e 3 e diferentes grupos laterais N-ligados distinguindo os distintos imidazóis (veja Figura 1.10); enquanto os triazóis possuem 3 átomos de nitrogênio no anel, nas posições 1, 2 e 4, e grupos laterais N-ligados particulares (veja Figura 1.10).

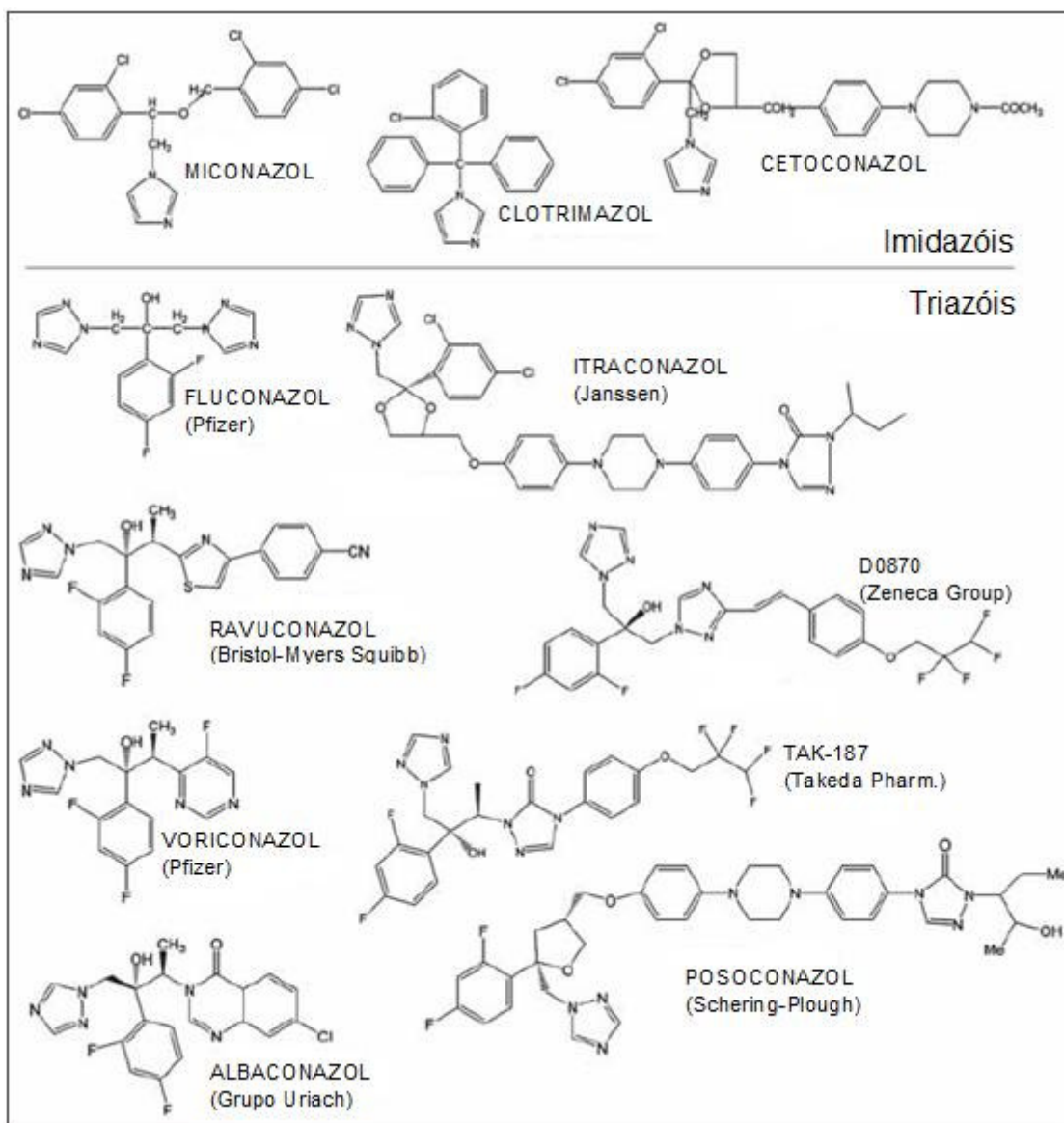
Após a descoberta destes compostos, em meados da década de 80, como interferindo especificamente no metabolismo fúngico, principalmente em um ou mais passos da via lanosterol-ergosterol, o estudo da biossíntese de ergosterol em fungos e protozoários se tornou de particular interesse (SONG & NES, 2007). A via de síntese de ergosterol passou então a ser vista como um alvo potencial de intervenção no tratamento de micoses e de parasitoses causadas por qualquer microorganismo que sintetizasse tal esterol. Além disso, IBEs se



tornaram uma importante ferramenta no estudo dos detalhes bioquímicos das vias de síntese de esteróis em diferentes organismos (revisto por BUCKNER, 2008).

Os azóis interagem com a enzima 14 $\alpha$ -desmetilase, também conhecida como CYP51, inibindo a desmetilação do carbono 14 do anel esteróide (veja Figura 1.9) (revisto por BUCKNER, 2008). Esta enzima é um membro da superfamília do citocromo P450, grupo de proteínas assim nomeado devido à presença de um grupamento heme que possui espectro característico com absorvância máxima em 450 nm na forma reduzida e ligada a CO (monóxido de carbono) (DANIELSON, 2002). Estas proteínas normalmente catalisam reações de óxido-redução, com o átomo de ferro do grupamento heme estando diretamente envolvido nos eventos de transferência de elétrons. O anel heterocíclico dos azóis interage diretamente com o átomo de ferro do citocromo P450, com o ligante lipofílico N-substituído do azol ocupando o sítio ativo do substrato da enzima (lanosterol, no caso de CYP51 de *T. cruzi*), bloqueando, portanto, tanto a ligação e a ativação de oxigênio molecular quanto a acomodação do substrato na enzima (LEPESHEVA *et al.*, 2007). O efeito inibitório depende diretamente do tamanho e configuração da região N-substituída da molécula do azol, a qual forma interações múltiplas adicionais com a parte protéica da enzima CYP51 (LEPESHEVA *et al.*, 2007). Assim sendo, moléculas mais volumosas, como cetoconazol (veja Figura 1.10), são mais potentes inibidores de CYP51 de fungos, embora possuam menor seletividade e, portanto, maiores problemas com relação à toxicidade em humanos devido à inibição de outros citocromos P450 além da CYP51 (revisto por BUCKNER, 2008).

A enzima esterol-14 $\alpha$ -desmetilase também está presente em humanos, entretanto, devido a diferenças estruturais na enzima de fungos, mamíferos e protozoários, o que reflete diferenças nos substratos preferências das mesmas, os azóis antifúngicos possuem IC<sub>50</sub> pelo menos 30 vezes maior para células humanas, o que possibilita que tais drogas possam ser usadas seguramente como antifúngicos em humanos (LAMB *et al.*, 1999; revisto por BUCKNER, 2008).



**FIGURA 1.10 – ESTRUTURA QUÍMICA DOS AZÓIS**

Na parte superior do quadro são evidenciados os imidazóis, enquanto na parte inferior são evidenciados os triazóis. Os nomes dentro de parênteses são as empresas farmacêuticas fabricantes do respectivo azól.

NOTA: note a presença do grupamento heterocíclico em todas as moléculas, contendo dois átomos de nitrogênio nos imidazóis e três no triazóis.

FONTE: modificado de BUCKNER, 2008.

Devido à similaridade da via de síntese de esteróis em *T. cruzi* e fungos, com ergosterol e ergosterol-like sendo o principal esterol de membrana deste parasita, azóis antifúngicos vêm sendo testados contra *T. cruzi* há mais de 25 anos. Docampo e colaboradores (1981) realizaram o primeiro trabalho *in vitro* e demonstraram potente atividade inibitória dos imidazóis miconazol e econazol em culturas de *T. cruzi*. Subsequentemente, uma variedade de azóis foi testada contra *T. cruzi*, com alguns compostos mais recentes apresentando IC<sub>50</sub> (concentração que inibe o crescimento em cultura em 50%)

na faixa nanomolar, especialmente Posaconazole (URBINA *et al.*, 1998), ravuconazole (URBINA *et al.*, 2003a) e TAK-187 (URBINA *et al.*, 2003b). Assim sendo, os azóis são hoje considerados como o grupo de compostos tripanocidas mais potente já testado *in vitro* contra *T. cruzi*. A Tabela 1.1 e Tabela 1.2 sumarizam os azóis já testados *in vitro* contra *T. cruzi* e suas respectivas IC<sub>50</sub> e MIC (*minimum inhibitory concentration*, concentração mínima capaz de inibir totalmente o crescimento de parasitas em cultura) contra as fases epimastigotas e amastigotas do parasita. Estudos recentes com cetoconazol e Posaconazol mostram que tais IBEs podem erradicar amastigotas de cardiomiócitos em cultura, permitindo a remontagem do citoesqueleto e do aparato contrátil da célula hospedeira (SILVA *et al.*, 2006).

**TABELA 1.1 – ATIVIDADE DE AZÓIS TESTADOS CONTRA EPIMASTIGOTAS DE *T. CRUZI* EM CULTURA**

Azól	Cepa	MIC (uM)	IC <sub>50</sub> (uM)	Referências
Miconazol	Tulahuen	x	20	Docampo, 1981
Econazol	Tulahuen	x	20	Docampo, 1981
Cetoconazol	Peru	x	0.2	Beach, 1986
Cetoconazol	Y	x	0.1	Urbina, 1993
Cetoconazol	EP ou Y	30	x	Urbina, 1991
Cetoconazol	Tulahuen	x	25	Hankins, 2005
D0870	EP	1-3,0	0.1	Liendo, 1998
Posoconazol	EP	0.003	0.00025	Benaim, 2006
Posoconazol	EP	0.03	x	Urbina, 1998
Albaconazol	EP ou Y	0.03	x	Urbina, 2000
Ravuconazol	EP ou Y	0.3	0.1	Urbina, 2003a
TAK-187	EP ou Y	0,3-1	x	Urbina, 2003b

NOTA: x-fatores não determinados nos trabalhos.

**TABELA 1.2 – ATIVIDADE DE AZÓIS TESTADA CONTRA AMASTIGOTAS DE *T. CRUZI* EM CULTURAS DE CÉLULAS DE MAMÍFEROS**

Azól	Cepa	MIC (nM)	IC <sub>50</sub> (nM)	Célula hospedeira	Referências
Fluconazol	Tulahuen	x	8000	Fibroblasto 3T3	Buckner, 1996
Itracanazol	x	20	2	Macrófagos	McCabe, 1986
Cetoconazol	Y	x	1	Macrófagos	McCabe, 1984
Cetoconazol	Peru	x	10	VA-13	Goad, 1989
Cetoconazol	Y	x	~2	Vero	Urbina, 1993
Cetoconazol	EP ou Y	10	x	Vero	Urbina, 1991
D0870	x	10	x	Vero	Liendo, 1998
Posoconazol	EP	30	0.25	Vero	Benaim, 2006
Albaconazol	EP ou Y	10	1	Vero	Urbina, 2000
Ravuconazol	EP ou Y	1	0.1	Vero	Urbina, 2003a
TAK-187	EP ou Y	1	0.3	Vero	Urbina, 2003b

NOTA: x-fatores não determinados nos trabalhos.

Embora a Tabela 1.1 e Tabela 1.2 infiram quais compostos são mais ativos contra culturas de *T. cruzi*, comparações diretas das concentrações inibitórias de crescimento são

complicadas devido às diferentes cepas utilizadas nos trabalhos independentes, visto que *T. cruzi* possui substancial diversidade genotípica e fenotípica intra-específica (revisito por STURM & CAMPBELL, 2009). De forma a retirar tal variável das análises, Buckner (2008) testou diferentes azóis contra a forma amastigota em cultura da cepa Tulahuen, demonstrando que posaconazol é o composto mais potente contra tal cepa do parasita, com efeitos inibitórios na célula hospedeira somente em concentrações cerca de 10.000 vezes maiores que a IC<sub>50</sub> para o parasita. Além disso, chama atenção a potência extremamente superior dos azóis em relação à droga usada clinicamente no tratamento da doença de chagas, benzonidazol (veja Tabela 1.3).

**TABELA 1.3 – ATIVIDADE DE AZÓIS (E BENZONIDAZOL) CONTRA A FORMA AMASTIGOTA DA CEPA TULAHUEN DE *T. CRUZI* CULTIVADA EM FIBROBLASTOS 3T3 DE CAMUNDONGOS**

Azól	IC <sub>50</sub> em Amastigotas (uM)	IC <sub>50</sub> em Fibroblastos 3T3 (uM)	IC <sub>50</sub> ama/IC <sub>50</sub> fib
Posaconazol	0.0005	5	10000
Cetoconazol	0.001	30	30000
Itraconazol	0.001	1	1000
Voriconazol	0.004	30	7500
Miconazol	0.02	10	500
Benzonidazol	1.5	>20	<13
Fluconazol	8	100	12.5

NOTA: os compostos estão ordenados de acordo com sua potência em *T. cruzi*. Na última coluna está representada a potencia das drogas no parasita em relação à célula hospedeira.

FONTE: modificado de BUCKNER, 2008.

Estudos *in vivo* com derivados de azóis no tratamento da doença de Chagas também vêm sendo realizados há décadas. O modelo animal mais utilizado em infecções experimentais por *T. cruzi* sem dúvida é o modelo murino (revisito por URBINA, 2001). Tais estudos demonstram que o antifúngico mais difundido atualmente, cetoconazol, possuiu atividade cepa-dependente na fase aguda da doença (BRENER *et al.*, 1993) e não é capaz de curar animais na fase crônica da doença de Chagas (MCCABE, 1988). Entretanto, derivados de azóis de última geração têm demonstrado resultados promissores: D0870 foi o primeiro composto com atividade curativa tanto na fase aguda quanto crônica da doença de Chagas em modelo murino (URBINA *et al.*, 1996); embora o desenvolvimento deste composto tenha sido interrompido devido a efeitos cardiotoxicos em estudos clínicos como droga antifúngica. Outros azóis têm demonstrado resultados similares: Posaconazol (SCH 56592) (URBINA *et al.*, 1998; MOLINA *et al.*, 2000) e TAK-187 (URBINA *et al.*, 2003b) parecem ser capazes de induzir cura parasitológica em ambas as fases da doença de Chagas, tanto em cepas suscetíveis quanto resistentes à droga padrão benzonidazol, além de serem efetivos em

camundongos imunossuprimidos. Tal potente ação *in vivo* destes compostos parece estar relacionada tanto à alta atividade anti-*T. cruzi* intrínseca (veja Tabela 1.1 e Tabela 1.2), quanto a melhores características farmacológicas (longa meia-vida e amplos volumes de distribuição) destas drogas em animais (URBINA *et al.*, 1998). Tais compostos encontram-se em testes clínicos como antifúngicos sistêmicos e são candidatos lógicos para futuros testes clínicos no tratamento da doença de Chagas. Embora albaconazol (UR-9825) seja um dos azóis mais potente *in vitro* contra *T. cruzi*, sua meia-vida breve em camundongos desencorajou estudos *in vivo* (URBINA *et al.*, 2000), situação similar observada para ravuconazol (URBINA *et al.*, 2003a; revisto por URBINA, 2009a).

Até o presente momento, somente dois azóis foram testados em humanos no tratamento da doença de Chagas: cetoconazol e itraconazol. Cetoconazol foi administrado na dose padrão para tratamento de micoses (4,5-8,8mg/kg/dia) durante 10-14 semanas em oito pacientes na fase crônica da doença, sendo que não foram encontradas evidências de cura por hemocultura e sorologia (BRENER *et al.*, 1993); estes pacientes eram brasileiros originários do estado de Minas Gerais, sendo que a cepa de *T. cruzi* responsável pela infecção dos mesmos não foi caracterizada.

Dois estudos com itraconazol foram reportados. Em um deles, realizado por pesquisadores da Universidade de São Paulo, 18 pacientes na fase crônica da Doença de Chagas foram tratados com doses de 100 ou 200mg/kg/dia, durante 12 semanas, sendo que não foram encontradas evidências de cura parasitológica (MOREIRA *et al.*, 1992). O outro estudo foi realizado por pesquisadores do Chile com pacientes de regiões urbanas e rurais do Chile e Argentina (APT *et al.*, 1998): 404 indivíduos na fase crônica da Doença de Chagas foram divididos em três grupos de tratamento recebendo itraconazol (6mg/kg/dia por 120 dias; 135 pacientes), alopurinol (8,5mg/kg/dia, durante 60 dias; 104 pacientes) ou placebo (165 pacientes). Não foram observados efeitos colaterais significativos, e, após um período de monitoramento de quatro anos, não foram encontradas alterações nos títulos de anticorpos anti-*T. cruzi* usando métodos sorológicos convencionais. O xenodiagnóstico foi utilizado para avaliar a cura parasitológica, entretanto, tal análise foi prejudicada pelo fato de que um grande número de pacientes já apresentava xenodiagnóstico negativo antes do tratamento, característica comum da fase crônica da doença. Entretanto, no grupo de tratamento com itraconazol, dos pacientes com xenodiagnóstico positivo antes do tratamento (43), um número significativo converteu para xenodiagnóstico negativo após o tratamento (38 dos 43, ou 88%) (APT *et al.*, 1998). Os pacientes deste último estudo foram acompanhados por onze anos pós-

tratamento (APT *et al.*, 2005), sendo que a maioria manteve positividade nos testes sorológicos. Portanto, parece que o tratamento com itraconazol não acarreta em cura parasitológica significativa.

Assim sendo, os estudos clínicos disponíveis não estabelecem uma função clara na utilização de cetoconazol ou itraconazol no tratamento da doença de Chagas (revisto por BUCKNER, 2008). Entretanto, baseado nos dados encontrados com novos azóis em modelos animais (URBINA *et al.*, 1998; MOLINA *et al.*, 2000; URBINA *et al.*, 2003b), os quais possuem atividade superior a cetoconazol ou itraconazol, novos estudos clínicos devem ser realizados num futuro próximo com inibidores de 14 $\alpha$ -desmetilase mais ativos contra *T. cruzi*. Segundo Coura (2009), Posaconazol é atualmente a maior esperança para o tratamento da Doença de Chagas, sendo que a fase inicial de experimentação em seres humanos foi iniciada em junho de 2010.

Todos os compostos mostrados na Figura 1.10 foram desenvolvidos especificamente como drogas antifúngicas. Conseqüentemente, o uso dos mesmos no tratamento de infecções por *T. cruzi* é uma aplicação secundária de tais drogas, a qual se torna válida levando em conta a falta de interesse de indústrias farmacêuticas em doenças parasitárias. Considerando a baixa similaridade das enzimas CYP51 de *T. cruzi* e *Candida albicans*, 28%, parece fortuito que drogas antifúngicas apresentem atividade contra CYP51 de *T. cruzi*, fato este que chama atenção para a importância da via de síntese de esteróis em *T. cruzi* (revisto por BUCKNER, 2008). Tais dados sugerem que há uma oportunidade de desenvolver inibidores de CYP51 específicos para *T. cruzi* que tenham, portanto, maior potência contra este parasita do que os azóis antifúngicos.

Com isso em mente, trabalhos vêm sendo desenvolvidos em busca de inibidores específicos de CYP51 de *T. cruzi*. A interação de diversos compostos com a enzima 14 $\alpha$ -desmetilase recombinante foi monitorada por alterações no espectro de absorvância do citocromo P450, demonstrando-se uma série de bis-imidazóis sintéticos capazes de interagir com TcCYP51 (CYP51 de *T. cruzi*) (BUCKNER *et al.*, 2003); dois desses compostos foram testados em modelo murino da Doença de Chagas e mostraram inibição total da mortalidade dos camundongos na fase aguda da doença. Em um trabalho mais recente, 30 derivados de imidazóis foram testados contra TcCYP51 recombinante através da análise de respostas espectroscópicas seguidas da quantificação do potencial inibitório da reação enzimática reconstituída *in vitro* (LEPESHEVA *et al.*, 2007); os compostos com parâmetros de ligação mais potentes eram então testados com relação à inibição da capacidade infectiva de

tripomastigotas em cardiomiócitos e com relação ao potencial inibitório na multiplicação de amastigotas em cardiomiócitos em cultura. Recentemente, demonstrou-se que a droga anti-câncer tipifarnib (um inibidor da enzima farnesil transferase de humanos) inibe a via do ergosterol em *T. cruzi* devido ligação a TcCYP51 (HUCKE *et al.*, 2005), um resultado fortuito que abriu caminho para a busca de formas modificadas da droga, mais específicas para *T. cruzi*, mas sem alterar as características farmacológicas de tipifarnib (KRAUS *et al.*, 2009). Mais recentemente, uma nova classe de inibidores de TcCYP51, derivada de um rastreamento para inibidores de CYP51 de *Mycobacterium tuberculosis* e baseados na fração molecular N-[4-piridil]-formamida foi descrita (CHEN *et al.*, 2009); um desses novos compostos foi capaz de erradicar amastigotas de *T. cruzi* de macrófagos em cultura. Tais metodologias de rastreamento de drogas demonstrou-se promissora na busca por novas drogas tripanocidas, ou até mesmo novos antifúngico, sendo que o estabelecimento de testes em larga escala na busca por compostos com afinidade de ligação à 14 $\alpha$ -desmetilase de *T. cruzi* poderá propiciar o desenvolvimento de compostos mais potentes no tratamento da doença de Chagas (revisto por BUCKNER, 2008).

#### 1.5.2.2 Inibidores de HMGR: estatinas

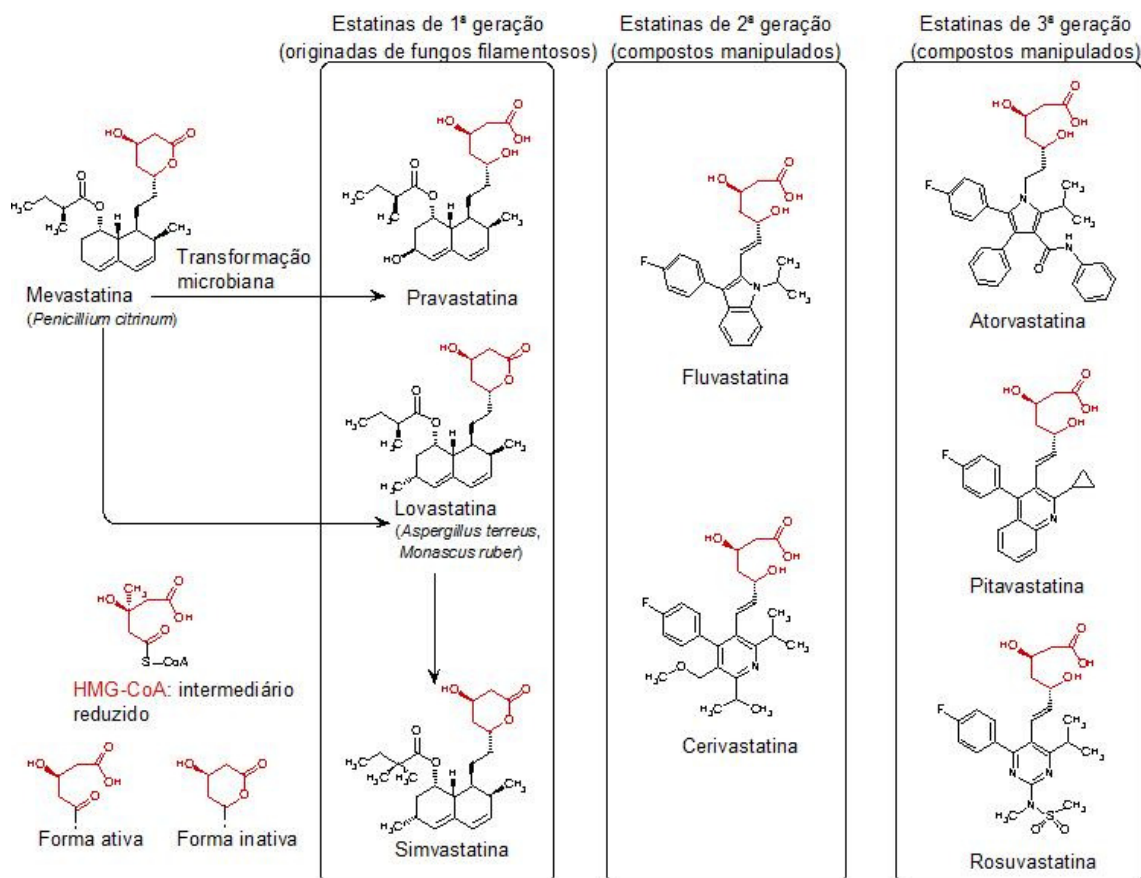
Os isoprenos (dolicol, coenzima Q, proteínas heme e proteínas preniladas, entre outros) e esteróis de eucariotos são sintetizados a partir de um mesmo composto celular (mevalonato). Este, por sua vez, é sintetizado pela redução de quatro elétrons do intermediário 3-hidroxi-3-metilglutaril CoA (HMG-CoA), catalisada pela enzima HMGR (veja Figura 1.5). Dessa forma, tal reação é o passo enzimático comprometido com a via biossintética do mevalonato, da qual irão se originar diversos compostos essenciais da célula (veja Figura 1.5, Figura 1.6 e Figura 1.7) (ISTVAN, 2002).

Há décadas, a inibição da via do mevalonato vem sendo estudada como uma forma de diminuir os níveis de colesterol plasmático em humanos. Cerca de 80% do colesterol presente em mamíferos é sintetizado de novo no fígado, sendo o restante absorvido da dieta. O principal ponto de regulação desta síntese endógena de colesterol em mamíferos é a modulação da atividade da enzima HMGR, através de um mecanismo de *feedback* negativo pelo colesterol exógeno absorvido no intestino. Dessa forma, a inibição da enzima HMGR deve ser uma forma efetiva de diminuir os níveis de LDL (*low-density-lipoprotein*) plasmático de forma a prevenir doenças coronarianas (ENDO, 1992). Assim sendo, no início

da década de 70 iniciou-se um projeto em busca de inibidores de HMGR de origem microbiana, o que levou a descoberta de um potente inibidor, isolado de *Penicillium citrinum*, nomeado mevastatina. Subsequentemente os mecanismos bioquímicos de ação desse composto foram elucidados e, no início da década de 80, foi demonstrado que mevastatina era capaz de diminuir os níveis de colesterol total e de LDL plasmático tanto em animais quanto em humanos. Estes resultados estimularam o desenvolvimento de análogos de mevastatina, e, no início da década de 90, três drogas – lovastatina (formalmente chamada mevenolin, isolada de *Aspergillus terreus*), simvastatina e pravastatina – haviam sido aprovadas e comercializadas em muitos países. Tais drogas vêm sendo utilizadas como efetivos agentes anti-hipercolesterolemia, demonstrando efeitos benéficos em pacientes com doenças coronarianas (ENDO, 1992). Tais resultados satisfatórios com compostos isolados de fungos filamentosos levaram ao desenvolvimento de drogas similares sintetizadas quimicamente, conhecidas genericamente como estatinas (veja Figura 1.11).

Comparando a estrutura química das estatinas com o substrato da enzima HMGR, HMG-CoA (veja Figura 1.11), evidencia-se que todos esses inibidores compartilham uma região HMG-CoA-like. Dessa forma, as estatinas são inibidores competitivos de HMGR com respeito à ligação de HMG. Estudos de cristalografia de raio-X resolveram o mecanismo pelo qual as estatinas interagem com a enzima HMGR (ISTVAN & DEISENHOFER, 2001): a região HMG-CoA-like do inibidor (região em vermelho na Figura 1.11) ocupa a bolsa de ligação para o substrato (HMG-CoA) da enzima HMGR, enquanto os grupos hidrofóbicos posicionam-se em um sulco raso formado como resultado do rearranjo da região C-terminal da enzima. Ou seja, as estatinas inibem HMGR ligando-se ao sítio ativo da enzima impedindo a ligação estérica do substrato (ISTVAN & DEISENHOFER, 2001; revisto por ISTVAN, 2002).





**FIGURA 1.11 – ESTRUTURA QUÍMICA DAS ESTATINAS**

Em vermelho está representada a região HMG-like das estatinas. A forma inativa das drogas apresenta região HMG-like na forma lactona (éster cíclico); in vivo, essas pró-drogas são hidrolisadas para a forma ácido-carboxílico ativa (veja detalhe no canto inferior esquerdo).

FONTE: modificado de [http://www.genome.ad.jp/dbget-bin/show\\_pathway?map07024+C00356](http://www.genome.ad.jp/dbget-bin/show_pathway?map07024+C00356)

Como os passos iniciais de síntese de esteróis são similares em todos eucariotos, a capacidade das estatinas de inibirem a enzima HMGR também vem sendo analisada em *T. cruzi*. O primeiro trabalho foi realizado por Florin-Christensen e colaboradores (1990), demonstrando-se que lovastatina era capaz de inibir o crescimento de epimastigotas em cultura de forma dose-dependente em concentrações de 10-30  $\mu\text{g/ml}$  (~25-75  $\mu\text{M}$ ), sendo que tais efeitos eram convertidos pela adição de 100  $\mu\text{M}$  de esqualeno no meio. Além disso, lovastatina mostrou-se capaz de inibir a incorporação de [14C]acetato nos esteróis endógenos do parasita (Florin-Christensen *et al.*, 1990; URBINA *et al.*, 1993). A 50  $\mu\text{g/ml}$  (~124  $\mu\text{M}$ ), lovastatina foi capaz de matar a maioria dos parasitas, concentração essa abaixo dos níveis tóxicos para mamíferos. Portanto, tal droga passou a ser analisada como um possível agente quimioterápico na doença de Chagas (Florin-Christensen *et al.*, 1990).

Em um trabalho subsequente, Urbina e colaboradores (1993) demonstraram que lovastatina em concentrações de 50  $\mu\text{M}$  e 75  $\mu\text{M}$  era capaz de induzir inibição total do

crescimento e lise celular de epimastigotas em cultura após 144 e 96 horas, respectivamente. Além disso, o tratamento levou a uma redução significativa na presença de esteróis endógenos maduros (4-desmetilados), com um concomitante aumento de colesterol exógeno. Análises da sinergia *in vitro* de diferentes IBEs mostraram que 7,5  $\mu\text{M}$  de lovastatina juntamente com 0,1  $\mu\text{M}$  de cetoconazol (concentrações estas capazes de inibir individualmente somente 20 a 30% do crescimento do parasita) eram capazes de induzir inibição total do crescimento seguida de lise celular em 144 horas (URBINA *et al.*, 1993); ou seja, a presença de 0,1  $\mu\text{M}$  de cetoconazol no meio diminui a concentração de lovastatina capaz de produzir um efeito tripanocida em 10 vezes, demonstrando o efeito sinérgico das drogas. Em amastigotas cultivadas em células VERO, o efeito de lovastatina, por si só, é moderado, diminuindo em cerca de 30% o número de células infectadas a 1  $\mu\text{M}$ ; acima desta concentração foram observados efeitos deletérios na célula hospedeira. Portanto, lovastatina tem uma atividade terapêutica *in vitro* muito pobre contra amastigotas de *T. cruzi* (URBINA *et al.*, 1993). Entretanto, a presença de 0,75  $\mu\text{M}$  de lovastatina diminuiu a MIC de cetoconazol em amastigotas de 10 nM para 1 nM. Devido ao forte efeito sinérgico destas drogas *in vitro*, particularmente em amastigotas, combinações de lovastatina e cetoconazol foram testadas em modelo murino da Doença de Chagas, demonstrando efeito sinérgico em relação à diminuição da parasitemia e sobrevivência dos camundongos. Estes resultados apóiam a idéia de que lovastatina pode ser usada para potencializar o efeito terapêutico de azóis no tratamento da Doença de Chagas, de forma a diminuir as doses efetivas dos mesmos evitando possíveis efeitos colaterais (URBINA *et al.*, 1993).

De forma a obter uma fonte apropriada da enzima HMGR para estudos funcionais e análises de inibição por drogas, Hurtado-Guerrero e colaboradores (2002) expressaram HMGR de *T. cruzi* em *E. coli*. Através de ensaios de inibição enzimática demonstraram que a estatina fluvastatina é um inibidor competitivo de HMGR, com  $K_i=27$  nM, enquanto a ligação de cerivastatina se dá por um processo bimodal, característico de um inibidor do tipo lento e com ligação forte (Hurtado-Guerrero *et al.*, 2002).

Embora os inibidores de HMGR, como lovastatina, aumentam os efeitos anti-proliferativos de azóis contra *T. cruzi*, tanto *in vitro* quanto *in vivo* (URBINA *et al.*, 1993), a baixa especificidade dessas drogas contra este parasita desencorajou estudos subsequentes como quimioterápicos no tratamento da doença de Chagas. Por outro lado, a localização celular distinta de HMGR em *T. cruzi* (glicossomo/mitocôndria), chama atenção para tal

enzima e para possíveis mecanismos distintos de controle da via de síntese de esteróis nesse parasita (PEÑA-DIAZ *et al.*, 2004).

## 1.6 TRANSCRIÇÃO POLICISTRÔNICA EM TRIPANOSSOMATÍDEOS

Em eucariotos, o processo de transcrição gênica é na maioria das vezes regulado por uma sequência promotora a montante da região codificadora do gene (revisto por Wray *et al.*, 2003). O processamento do pré-mRNA ocorre co-transcricionalmente e envolve a adição de um resíduo metilguanossina-trifosfato (m<sup>7</sup>Gppp, ou cap) na extremidade 5', *cis-splicing* (retirada dos íntrons para ligação dos éxons) e poliadenilação da extremidade 3' (revisto por PROUDFOOT *et al.*, 2002).

Tripanossomatídeos apresentam diferenças com relação a esse padrão. Embora a maquinaria de *cis-splicing* esteja presente (PALFI *et al.*, 2005), ela apresenta diferenças estruturais grandes com relação a de demais eucariotos, sendo que somente um gene, poli-A polimerase de *T. brucei* e *T. cruzi* (MAIR *et al.*, 2000), foi identificado como possuindo íntron em tripanossomatídeos. A transcrição gênica nesses organismos é muito peculiar e está diretamente relacionada à forma como os genes codificadores de proteínas estão dispostos no genoma destes parasitas. Dezenas a centenas de genes são codificados em ilhas em uma única fita do DNA, com pequenas regiões intergênicas, sendo que tais ilhas são precedidas por uma região não transcrita de 1 a 13 kb, referida como região de troca de fita (SSR, *strand switch region*). Na outra extremidade da SSR começa um segundo conjunto de genes, codificados na fita oposta do DNA, de forma que os arranjos gênicos em tandem irradiam-se bidirecionalmente de uma região SSR (revisto por PALENCHAR & BELLOFATTO, 2006). Usando informações de sequências de segmentos de cromossomos homólogos, tem sido verificado que, a despeito do alto nível de divergência nas sequências, *T. brucei*, *T. cruzi* e *L. major* exibem uma alta conservação do ordenamento gênico (GHEDIN *et al.*, 2004; EL-SAYED *et al.*, 2005b). Especula-se que tal conservação de grupos sintênicos deve estar relacionada com a transcrição policistrônica observada nesses organismos (VARGAS *et al.*, 2004), sendo que as SSRs seriam os sítos de início de transcrição.

Todos os genes presentes nas ilhas em torno das SSRs são transcritos em conjunto por uma RNA polimerase (RNAP) tipo II, produzindo unidades policistrônicas (BEN AMAR *et al.*, 1988; revisto por VANHAMME & PAYS, 1995) que são representadas por uma única

molécula de RNA geralmente contendo diversos genes que não são relacionados a uma determinada via metabólica, como ocorre nos operons de procariotos. Além disso, os genes de uma mesma unidade de transcrição policistrônica podem apresentar diferenças grandes nos níveis de expressão, demonstrando a importância de mecanismos pós-transcricionais de controle da expressão gênica (revisto por VANHAMME & PAYS, 1995). A molécula de RNA policistrônica deve ser processada, pelos mecanismos acoplados de *trans-splicing* (SUTTON & BOOTHROYD, 1986) e poliadenilação para gerar unidades monocistrônicas traduzíveis.

Durante o *trans-splicing*, a região codificadora de cada gene é separada do policistron e é adicionada um éxon especial na extremidade 5' da molécula, o qual é uma sequência espécie-específica, altamente conservada, contendo 39 nucleotídeos, conhecida como sequência-líder (SL) ou mini-éxon. A SL também é encontrada em nematódeos (KRAUSE & HIRSH, 1987), trematódeos (RAJKOVIC *et al.*, 1990) e em *Euglena* sp. (TESSIER *et al.*, 1991).

Nesta reação, participam duas moléculas de RNA, uma doadora e outra acceptora. A doadora é um precursor do mini-éxon, a qual, em *T. cruzi*, possui cerca de 110 nucleotídeos (ZWIERZYNSKI & BUCK, 1991), é transcrita por RNAP II e possuiu em sua extremidade 5' um cap modificado, conhecido como cap 4 devido a modificações químicas dos primeiros 4 nucleotídeos após a 7-metilguanosina (PERRY *et al.*, 1987). A molécula acceptora é o transcrito derivado da unidade policistrônica ao qual o mini-éxon é transferido. O *trans-splicing* ocorre através de duas reações de transesterificação, análogo ao *cis-splicing*, mas com a formação de uma estrutura intermediária em forma de “Y”, em vez de um laço, visto que os éxons provêm de moléculas distintas (LAIRD, 1989; revisto por LIANG *et al.*, 2003). Dessa forma, durante o processo de *trans-splicing*, dois genes, antes contíguos no pré-mRNA policistrônico, são separados, havendo adição da cauda poli-A na extremidade 3' do primeiro e ligação simultânea do SL na extremidade 5' do segundo (revisto por PALENCHAR & BELLOFATTO, 2006). Devido ao *trans-splicing*, todos mRNAs de tripanossomatídeos possuem uma SL na extremidade 5' (DE LANGE *et al.*, 1984).

Em tripanossomatídeos, não há um sinal específico dedicado à poliadenilação, sendo que tal sítio é determinado pelo *trans-splicing* do mRNA à jusante no RNA policistrônico. Esse evento sempre ocorre a uma distância relativamente fixa (~100-400 nt) a montante do sítio de *splicing*, sendo que deleções nestas regiões impedem o correto processamento destes transcritos (MATTHEWS *et al.*, 1994; revisto por LIANG *et al.*, 2003). Regiões ricas em

resíduos de pirimidinas, a distâncias determinadas dos sítios de poliadenilação e trans-splicing, regulam a adição da cauda poli-A e do mini-éxon, respectivamente (CAMPOS *et al.*, 2008). A sequência consenso AG no sítio 3' de *trans-splicing* também é essencial. As regiões não traduzidas 5'-UTR e 3'-UTR dos mRNAs maduros de *T. cruzi* possuem tamanho médio de 35 e 264 nucleotídeos, respectivamente (CAMPOS *et al.*, 2008), tamanho este significativamente menor que em *T. brucei* (BENZ *et al.*, 2005), o que condiz com o genoma mais compacto de *T. cruzi* (CAMPOS *et al.*, 2008).

Análises *in silico* dos genomas de *T. brucei*, *T. cruzi* e *L. major* mostram que estes parasitas não possuem homólogos para três das doze subunidades de RNA polimerase (RNAP) tipo II de eucariotos (IVENS *et al.*, 2005). Embora poucas diferenças estruturais tenham sido observadas, a cauda C-terminal (domínio CTD) não apresenta as repetições de aminoácidos típicas de eucariotos, mas também é rica em serina, sendo capaz de ser fosforilada. Tal CTD não usual deve estar relacionada com o mecanismo de transcrição fundamentalmente diferente de outros eucariotos (revisto por PALENCHAR & BELLOFATTO, 2006). Além disso, muitos dos fatores de transcrição típicos de RNAP II parecem estar ausentes (IVENS *et al.*, 2005). A aparente ausência desses fatores pode ser explicada de diferentes formas: (1) formas altamente divergentes destes fatores basais estão presentes, impedindo identificação bioinformática; (2) proteínas específicas destes parasitas desempenhariam funções similares ou (3) a regulação relaxada da transcrição policistrônica eliminaria a necessidade para tais fatores (revisto por PALENCHAR & BELLOFATTO, 2006).

Inspeção das SSRs revelam uma ausência de elementos promotores típicos de eucariotos para interação com fatores de transcrição, como TATA *box*, e nenhuma sobre-representação de elementos repetitivos que poderiam funcionar como promotores para RNAP II. Isso deve indicar que a transcrição inicia-se de forma aleatória nas SSRs ou que os promotores não são usuais (revisto por PALENCHAR & BELLOFATTO, 2006). O único promotor de RNAP II identificado até o momento é para os genes de SL, o qual se assemelha com o promotor do gene U1 snRNA de humanos (GUNZL *et al.*, 1997; revisto por PALENCHAR & BELLOFATTO, 2006).

Dados recentes indicam que estruturas específicas da cromatina nas SSRs devem compensar a aparente ausência de promotores definidos, sendo que os prováveis sítios de início de transcrição por RNAP II são marcados por uma cromatina mais frouxa, com nucleossomos distintos e menos estáveis devido a presença de formas variantes das histonas

H2A e H2B, H2AZ e H2BV, respectivamente (SIEGEL *et al.*, 2009). Além disso, foram observados elevados níveis de acetilação das histonas H3 e H4 na extremidade 5' das unidades policistrônicas de *T. cruzi* (RESPUELA *et al.*, 2008), *L. major* (THOMAS *et al.*, 2009) e *T. brucei* (SIEGEL *et al.*, 2009), juntamente com metilação de H3 (RESPUELA *et al.*, 2008). A acetilação de H4 pode funcionar como plataforma para ligação de proteínas auxiliaadoras na iniciação da transcrição, como BDF3 (*bromodomain-containing factor*), o qual recrutaria complexos de remodelamento da cromatina para os sítios de início de transcrição. Nesse contexto, a presença de motivos de poliguanina, nos sítios ricos em histonas H4 acetiladas, também parecem ser importante na determinação dos sítios de início da transcrição (SIEGEL *et al.*, 2009), uma característica comum de promotores de humanos (HUPPERT & BALASUBRAMANIAN, 2007). Já nos sítios de terminação da transcrição das unidades policistrônicas, observou-se uma elevada presença de formas variantes das histonas H3 e H4 (H3V e H4V, respectivamente) (SIEGEL *et al.*, 2009). Portanto, a transcrição policistrônica por RNAP II em tripanossomatídeos parece ser regulada por fatores epigenéticos, o que apóia a hipótese de que os sítios de início da transcrição não contêm motivos de sequência típicos para interação com fatores de transcrição (RESPUELA *et al.*, 2008; SIEGEL *et al.*, 2009).

Além dos mRNAs, a RNAP II de tripanossomatídeos também é responsável pela transcrição dos pequenos RNAs nucleolares, snoRNAs, os quais são transcritos como parte de unidades policistrônicas (DUNBAR *et al.*, 2000).

A função da RNAP I em *T. cruzi* é similar a de outros eucariotos: transcrição de genes de rRNA. Diferentemente dos promotores para RNAP II, a estrutura do promotor de genes rRNA é bem caracterizada e lembra a de outros eucariotos (revisto por CAMPBELL *et al.*, 2003). A noção de que promotores de rRNA são espécie-específicos também é verdade para tripanossomatídeos, sendo notável que a atividade destes promotores correlacione-se com as principais divisões filogenéticas de *T. cruzi* (ZINGALES *et al.*, 1999).

A RNAP III de cinetoplastídeos é responsável pela transcrição de genes de rRNA 5S, tRNAs e snRNAs. A sequência promotora consenso de eucariotos para genes rRNA 5S (promotor de RNAP III classe 1) e para tRNAs (promotor de RNAP III classe 2) foram identificadas. Entretanto, o promotor clássico para U6 snRNAs (promotor de RNAP III classe 3) não foi encontrado em tripanossomatídeos, sendo que todas as moléculas de snRNA são transcritas por RNAP III dirigidas por promotores de genes de tRNA dispostos a montante, e na fita oposta, dos genes de snRNA (revisto por CAMPBELL *et al.*, 2003). Diferentemente

destes parasitas, em mamíferos, os genes dos snRNAs U1-U5 são transcritos por RNAP II (WILLIS, 1993).

## 1.7 REGULAÇÃO DA EXPRESSÃO GÊNICA EM TRIPANOSSOMATÍDEOS

Todos os organismos adaptam-se a alterações no seu ambiente ajustando sua expressão gênica, sendo que o principal ponto de controle, desde bactérias até humanos, é a iniciação da transcrição. Entretanto, tripanossomatídeos parecem ter perdido quase que completamente a capacidade de regular a iniciação da transcrição por RNAP II, visto que os genes codificadores de proteínas são organizados em grandes grupamentos gênicos direcionais, ou unidades policistrônicas, havendo evidências de que a transcrição da maioria destes genes é ubíqua (revisto por CLAYTON, 2002; revisto por VANHAMME & PAYS, 1995). Por outro lado, estes parasitas apresentam diferentes estágios durante seu ciclo de vida, com características morfológicas e metabólicas únicas, consistentes com um alto nível de regulação gênica, o que é apoiado por dados recentes de microarranjo que indicam um grande número de genes regulados durante o ciclo de vida de *T. cruzi* (MINNING *et al.*, 2009) e *T. brucei* (JENSEN *et al.*, 2009). Assim sendo, os mecanismos de regulação da expressão gênica nesses organismos dependem quase que exclusivamente de vias regulatórias que agem em níveis posteriores à transcrição, num processo denominado regulação pós-transcricional. Nesse contexto, destacam-se os seguintes pontos de controle em tripanossomatídeos:

**Processamento do mRNA e exportação do núcleo:** o tamanho dos tratos de polipirimidina, e sua distância em relação aos sítios de poliadenilação e *trans-splicing*, são essenciais no processo acoplado de *trans-splicing*-poliadenilação (CAMPOS *et al.*, 2008). Diferenças na localização e tamanho destes tratos podem ser responsáveis por eficiências distintas no *trans-splicing* de vários genes (MATTHEWS *et al.*, 1994; revisto por LIANG *et al.*, 2003). Além disso, fatores protéicos estágio-específicos, que interagem com a maquinaria de *splicing* e poliadenilação podem ser responsáveis pelo *trans-splicing* alternativo de genes específicos em distintas fases do ciclo de vida de *T. cruzi* (MANNING-CELA, 2002), ou regularem o sucesso do amadurecimento do mRNA (revisto por LIANG *et al.*, 2003). A exportação de mRNAs maduros do núcleo ainda é pouco entendida em tripanossomatídeos, mas o controle seletivo

da passagem destas moléculas pelo poro nuclear é um potencial ponto de controle pós-transcricional da expressão gênica (revisto por CLAYTON & SHAPIRA, 2007).

**Degradação citoplasmática dos mRNAs:** o mecanismo de degradação de mRNAs no citoplasma é bem estabelecido em eucariotos. O apuramento da cauda poli(A) desencadeia o decapamento do mRNA, o qual é, em seguida, degradado por éxonucleases citoplasmáticas no sentido 5'-3', ou no sentido 3'-5' por um grande complexo de éxonucleases, o exosomo (BUTLER, 2002). Este é um mecanismo não regulado, ou constitutivo, de degradação de mRNA sendo dependente de desadenilação e operando em uma cinética mais lenta (revisto por HAILE & PAPADOPOULOU, 2007). Além da via constitutiva, há uma via regulada de degradação de mRNAs, a qual parece ser desadenilação independente; proteínas regulatórias podem ligar motivos específicos na região 3'-UTR e estabilizar ou desestabilizar mRNAs inibindo ou promovendo a desadenilação, respectivamente. A degradação regulada também pode envolver clivagem endonucleotídica na região 3'-UTR ou enzimas que degradam o cap (revisto por CLAYTON, 2002). Tripanossomatídeos claramente possuem esta maquinaria geral, sendo que estes mecanismos de controle da degradação de mRNAs parecem ser essenciais no controle da expressão gênica destes parasitas (IVENS *et al.*, 2005; revisto por CALYTON & SHAPIRA, 2007; revisto por HAILE & PAPADOPOULOU, 2007). Um exemplo são os genes de *T. cruzi* que codificam para proteínas semelhantes a mucinas (mucin-like), os quais são expressos predominantemente nas formas do parasita presentes no hospedeiro vertebrado. Tal regulação é realizada por um elemento rico nas bases AU (AU-rich element, ARE) na 3'-UTR do mRNA (DI NOIA *et al.*, 2000), no qual se liga uma proteína ligadora de RNA (TcUBP-1) que causa desestabilização do mRNA da mucina (D'ORSO & FRASCH, 2001). Como TcUBP-1 tem expressão regulada, ela é capaz de regular negativamente a expressão de diversos mRNAs que possuem a sequência específica para sua ligação na região 3'-UTR.

**Associação à maquinaria de tradução:** caso o mRNA seja processado e esteja disponível no citoplasma, sem estar sendo degradado, ainda é possível que tal elemento não tenha sua informação transformada em proteína, a qual seria a forma mais correta de considerarmos a expressão gênica. Fatores controlando o acesso dos mRNAs citoplasmáticos à maquinaria de tradução podem atuar de forma seletiva, criando um mecanismo gene-específico de controle pós-transcricional da expressão. Repressores da síntese protéica podem impedir a



circularização do mRNA e a formação do complexo necessário para iniciação da tradução (RICHTER & SONENBERG, 2005). Em tripanossomatídeos, a regulação da tradução normalmente deve-se a sinais específicos na região 3'-UTR dos mRNAs, como ocorre para os genes de choque térmico HSP83 e HSP70 (revisto por CLAYTON & SHAPIRA, 2007). Em *T. cruzi*, foi demonstrado que moléculas de RNA mensageiro podem ser mantidas estáveis e não associadas a polissomos, desmobilizadas no citoplasma, sendo passíveis de tradução *in vitro* (GOLDENBERG *et al.*, 1985). Há evidências de que o controle da associação à maquinaria de tradução seja o mecanismo principal na expressão diferencial de alguns genes diferencialmente expressos durante a metaciclogênese de *T. cruzi* (ÁVILA *et al.*, 2001; DALLAGIOVANNA *et al.*, 2001; FRAGOSO *et al.*, 2003). Análises de microarranjo mostram que durante o ciclo de vida (PAVONI, 2005) e metaciclogênese (PROBST, 2005) de *T. cruzi*, o número de genes diferencialmente expressos chega a ser 10 vezes maior avaliando-se mRNA polissomal em comparação com mRNA total, sendo muito provável que esse mecanismo de mobilização polissomal também esteja atuando em outros processos de diferenciação e resposta ao ambiente. Análises do ribossomo de *T. cruzi* demonstram a presença de uma estrutura não usual próximo ao canal de saída do mRNA, na subunidade 40S, devido expansões em domínios específicos do rRNA 18S (GAO *et al.*, 2005). Estas alterações são conservadas em *T. brucei* e *L. major*, o que, juntamente com as características únicas e conservadas do cap 4 e SL, sugerem a existência de interações únicas entre o ribossomo e mRNAs desses parasitas em um processo particular de iniciação da tradução (GAO *et al.*, 2005), o que apóia a hipótese da mobilização polissomal como essencial no controle da expressão gênica em *T. cruzi*. Além disso, o número expandido de potenciais fatores de tradução em *Trityps* (*T. cruzi*, *T. brucei* e *L. major*) sugere um alto nível de especialização e controle da tradução nesses organismos (IVENS *et al.*, 2005).

**Estabilidade protéica:** esse é um fenômeno pós-traducional, mas sem dúvida é uma forma de um organismo sem um controle restrito sobre a transcrição de seus genes poder criar diferenças no conjunto de proteínas disponíveis em um determinado momento. Entretanto, muito pouco é conhecido acerca da regulação gênica pós-traducional em tripanossomatídeos (revisto por HAILE & PAPADOPOULOU, 2007). Os genomas de *T. cruzi*, *T. brucei* e *L. major* contêm um grande número de possíveis proteínas quinases, enfatizando mecanismos de controle da atividade protéica por fosforilação/desfosforilação (PARSONS *et al.*, 2005). Estudos combinados de proteômica e transcriptômica em *L. infantum* mostram que

modificações pós-traducionais devem ter uma importante função durante a diferenciação desse parasita (MCNICOLL *et al.*, 2006), devido à presença de variantes moleculares de diversas proteínas em géis bidimensionais. A degradação protéica é parcialmente responsável pela diminuição estágio-específica de proteínas ligadoras de RNA em *T. brucei* (revisto por HAILE & PAPADOPOULOU, 2007). Além de fosforilação, outros processos típicos de eucariotos para estabilização e/ou ativação protéica estão presentes em tripanossomatídeos, como glicosilação e lipidificação (IVENS *et al.*, 2005).

Recentemente, maior atenção vem sendo dada a mecanismos pós-transcricionais de controle da expressão gênica em mamíferos. Como hipotetizado (2002) e confirmado (2007) por Jack D. Keene, múltiplos mRNAs podem ser co-regulados por uma ou mais proteínas ligadoras de RNA sequência-específica, as quais são capazes de orquestrar o *splicing*, exportação, estabilidade, localização e tradução de seus mRNAs. Tais complexos ribonucleoprotéicos podem ser denominados como operons pós-transcricionais, ou regulons, visto que coordenam a co-expressão de proteínas funcionalmente relacionadas (revisto por KEENE, 2007). Devido à aparente ausência de regulação da transcrição em tripanossomatídeos, tal modelo de operon pós-transcricional toma ainda maior importância nesses organismos. Os genomas de *Trypanosoma* contêm cerca de uma centena de genes codificando proteínas com domínios RRM (RNA recognition motifs) (IVENS *et al.*, 2005; DE GAUDENZI *et al.*, 2005), em torno de 40 proteínas contendo o domínio de interação com RNA dedo de zinco CCCH (contra somente 7 em *Saccharomyces cerevisiae* e 12 em *S. pombe*) (IVENS *et al.*, 2005) e 13 proteínas contendo domínio Pumilio (PUF), enfatizando a importância dos mecanismos de regulação da expressão gênica por interação dos mRNAs com proteínas ligadoras de RNA (revisto por CLAYTON & SHAPIRA, 2007). Portanto, é provável que a regulação pós-transcricional de RNAs e proteínas nesses parasitas baseie-se no conjunto de motivos presentes no mRNA, os quais especificam para um determinado padrão de expressão durante o ciclo de vida ou em condições de stress. Tais motivos devem ser ligados por diferentes proteínas regulatórias, sendo que a estabilidade e tradução dos mRNAs dependem da combinação de fatores e suas modificações pós-traducionais.

## 1.8 GENÔMICA

Na biologia molecular moderna, genoma é definido como a totalidade de informação genética presente num organismo, ou seja, a sequência nucleotídica completa de todos os cromossomos de um ser vivo. Dessa forma, genômica é a área da ciência dedicada ao estudo dos genomas dos organismos. Esse campo inclui esforços intensivos para determinar a sequência de DNA completa de diversas espécies, seguido da análise em larga escala dos genes e regiões não codificadoras presentes no genoma.

A genômica pode ser didaticamente subdividida em diferentes campos de estudo com escopos distintos. A Genômica funcional utiliza a grande quantidade de dados gerados pelos projetos genomas para descrever funções gênicas e suas interações. Dessa forma, a genômica funcional foca em aspectos dinâmicos do genoma, como transcrição, tradução e interações protéicas, diferentemente do aspecto relativamente estático da informação genômica (sequências de DNA e sua estrutura) (JOOS *et al.*, 2003). Por apoiar-se em técnicas de análise da expressão gênica em larga escala (microarranjos de DNA, RNA-seq, proteômica), a genômica funcional também é simplesmente definida como a análise em larga escala da expressão gênica de uma célula (ou conjunto de células) em uma determinada condição (JOOS *et al.*, 2003).

Com base nos termos “genoma” e “genômica”, neologismos para descrever conjuntos de dados biológicos completos, e respectivas áreas de estudo, foram criados com os sufixos “oma” e “ômica”. Transcriptoma é o conjunto completo de moléculas de RNA presentes numa célula, tecido ou organismo, sendo que a análise em larga escala destas moléculas é conhecida como transcriptômica. De forma similar, proteoma é o conjunto completo de proteínas presentes numa célula, tecido ou organismo, e seu estudo é denominado proteômica (JOOS *et al.*, 2003). A genômica funcional engloba diversas “ômicas”, incluindo transcriptômica e proteômica, sendo que a análise conjunta das mesmas permite quantificar os processos biológicos dinâmicos de forma a melhorar nosso entendimento das funções gênicas (e protéicas) e suas interações.

### 1.8.1 Novas tecnologias de seqüenciamento de DNA

O desenvolvimento das técnicas de seqüenciamento enzimático (Sanger *et al.*, 1977) e químico (Maxam & Gilbert, 1977) do DNA, na década de 70, revolucionaram a prática científica. Devido principalmente à sua adaptabilidade para automação, o método de Sanger, utilizando dideoxynucleotídeos terminadores de cadeia, se tornou o método padrão para determinação de seqüências de DNA na comunidade científica. Durante a década de 90, a introdução de sondas coloridas para detecção de cada uma das quatro bases do DNA em uma única reação de seqüenciamento, o desenvolvimento da eletroforese em capilar e a detecção computadorizada de fluorescência dos dideoxynucleotídeos marcados permitiram o desenvolvimento de máquinas de seqüenciamento de DNA totalmente automatizadas, capazes de analisar 96 ou 384 amostras em paralelo (revisto por Strausberg *et al.*, 2008). Esses seqüenciadores de Sanger foram utilizados com sucesso para seqüenciar o primeiro genoma de um organismo de vida-livre, a bactéria *Haemophilus influenzae* (FLEISCHMAN *et al.*, 1995), o genoma da levedura *Saccharomyces cerevisiae* (GOFFEAU *et al.*, 1997), o genoma de vários mamíferos, como o genoma humano (LANDER *et al.*, 2001), camundongo (WATERSTON *et al.*, 2002) e chimpanzé (MIKKELSEN *et al.*, 2005), genomas de vários protozoários (revisto por COPPEL & BLACK, 2005), como o *T. cruzi* (EL-SAYED *et al.*, 2005a), entre outros.

Atualmente, os seqüenciadores baseados em capilar que empregam a química de Sanger tem a habilidade de gerar pelo menos 1 a 2 milhões de bases de informação num período de 24 horas, com *reads* longos de até 900 bases e alta acurácia (veja Tabela 1.4). Entretanto, o processo laborioso de preparação das amostras, envolvendo longas etapas de clonagem *in vivo*, a limitada capacidade para análise de múltiplas amostras em paralelo, e o alto custo desta tecnologia impedem a análise de genomas em larga escala (revisto por Strausberg *et al.*, 2008).

Em 2004, o Instituto Nacional de Pesquisa Genômica Humana (National Human Genome Research Institute, NHGRI), parte do Instituto Nacional de Pesquisa dos E.U.A. (National Institute of Health, NIH), lançou dois novos programas com o objetivo de baixar o custo de seqüenciamento completo de um genoma humano (ou genoma de tamanho equivalente ao humano, 3 bilhões de nucleotídeos) para 100 mil dólares e, eventualmente, para mil dólares (SERVICE, 2006). O primeiro grupo que fosse capaz de cumprir tal meta receberia prêmios em dinheiro, bem como financiamentos para pesquisas futuras (SERVICE,

2006). Tais programas tiveram um efeito impressionante neste campo, estimulando o desenvolvimento de diversas metodologias inovadoras de seqüenciamento de DNA, sendo que algumas já vêm sendo comercializados nos últimos anos. A segunda geração de seqüenciadores de DNA vem abrindo oportunidades fascinantes na pesquisa científica, motivo pelo qual foi o método do ano de 2007 escolhido pela renomada revista científica *Nature Methods* (SCHUSTER *et al.*, 2008).

Estas novas metodologias possuem uma característica em comum: enorme paralelismo de reações de seqüenciamento. Isso permite uma produtividade de dados por corrida diversas ordens de grandeza acima do seqüenciamento clássico de Sanger. Por esse motivo, os métodos de seqüenciamento de DNA de segunda geração também são conhecidos como seqüenciamento paralelo massivo (*massively parallel sequencing*, MPS). Juntamente com preparações de amostras mais simples, tais métodos baixaram enormemente o custo por megabase de dados gerados (veja Tabela 1.4), de forma que o seqüenciamento genômico para análises clínicas (genômica pessoal), até hoje utópico, poderá estar disponível nos próximos anos (ANSORGE, 2009).

Atualmente, três novas tecnologias de seqüenciamento de DNA já estão disponíveis e consolidadas no mercado, com respectivas vantagens e desvantagens, discutidas em recentes publicações científicas (revisto por METZKER, 2010): 454 GenomeSequencer FLX (Roche Applied Science), Illumina Genome Analyser e SOLiD (Applied Biosystems) (para um entendimento detalhado do funcionamento destas tecnologias, veja ANEXO I).

**TABELA 1.4 – CARACTERÍSTICAS DOS SEQUENCIADORES DE DNA**

Equipamento	Química de sequenciamento	<i>reads/run</i>	tamanho médio dos <i>reads</i>	produtividade por corrida	custo por Mb (US\$)
3730xl	Sanger	96	900	86 Kb	500
454 FLX	pirosequenciamento	400 mil	250	100 Mb	60
Illumina	terminadores reversíveis	140 milhões	50	7 Gb	2
SOLiD	seqüenciamento com ligase	400 milhões	50	20 Gb	2

LEGENDA – Kb: kilo bases, mil bases; Mb: mega bases, um milhão de bases; Gb: giga bases, um bilhão de bases;

NOTA – custo por Mb segundo SHENDURE & JI, 2008.

Embora essas plataformas sejam bem distintas na química de seqüenciamento, bem como na forma como o arranjo para sequenciamento é gerado, o fluxo de trabalho é conceitualmente similar. O preparo da biblioteca é realizado por fragmentação aleatória do DNA, seguida por ligação *in vitro* de adaptadores em ambas as extremidades das moléculas. A formação de moléculas molde para o sequenciamento é realizada por amplificação clonal de cada um dos fragmentos originais, através de um dos métodos conhecidos como colônias de

polimerase: PCR em emulsão, emPCR (454 e SOLiD), e PCR de ponte em fase sólida no Illumina (revisto por SHENDURE & JI, 2008). Como cada molécula original é amplificada independentemente (em microrreatores na emPCR e em clusters independentes na PCR de ponte em fase sólida), a individualidade de cada molécula original se mantém. O que é comum em qualquer uma dessas tecnologias é que os amplicons gerados por PCR são depositados sobre a célula de fluxo, onde ocorrem as reações de seqüenciamento, sendo que todos os grupos de moléculas são seqüenciados simultaneamente pelos mesmos reagentes que banham a totalidade da célula de fluxo, sendo que cada leitura é individualizada por sua posição na célula de fluxo. Finalmente, o processo de seqüenciamento consiste de ciclos alternados de química dirigida por enzimas e aquisição de dados por imagem, sendo que as leituras são geradas por interpretação computacional das imagens geradas (revisto por SHENDURE & JI, 2008).

As principais vantagens das estratégias de seqüenciamento de segunda geração, em relação ao seqüenciamento clássico de Sanger, são: (i) ausência do laborioso processo de clonagem *in vivo*; (ii) alto grau de paralelismo devido a seqüenciamento em células de fluxo; (iii) utilização de volumes reduzidos de reagentes e (iv) produtividade de dados enormemente maior (revisto por SHENDURE & JI, 2008).

O tamanho das leituras das novas tecnologias de seqüenciamento de DNA está muito abaixo do método tradicional de Sanger (Tabela 1.4), o que dificulta a montagem de genomas *de novo*. Entretanto, o grande número de pequenas leituras geradas provê oportunidades para o desenvolvimento de novas aplicações que se beneficiam deste formato particular de dados. Dessa forma, estas tecnologias vêm sendo amplamente utilizadas no contexto de genomas referência, ou seja, para organismos que já possuem um genoma seqüenciado (revisto por Morozova & Marra, 2008).

O método “censo de sequencia” (*sequence census*) usa leituras pequenas para determinar o sítio de origem desta seqüência no genoma de referência (WOLD & MYERS, 2008). O princípio por trás deste método é muito simples: para entender o conteúdo de uma mistura complexa de ácidos nucléicos, sequencie a mesma. Não é necessário seqüenciar a totalidade de cada molécula presente na mistura inicial, uma leitura de cerca de 30 nucleotídeos (“tag”) já possibilita identificar o sítio de origem de cada fragmento através de alinhamento da mesma no genoma de referência. Desta forma, o número de leituras que mapeiam em uma determinada região do genoma correlaciona-se com a abundância da mesma na amostra original. A chave para este método é o grande número de leituras

individuais seqüenciadas, cada uma correspondente a uma única molécula na amostra original. Conceitualmente, o método censo de sequencia é similar ao método SAGE (*serial analysis of gene expression*) (VELCULESCO *et al.*, 1995) e MPSS (*massively parallel signature sequencing*) (BRENNER *et al.*, 2000), mas com a vantagem de ser mais barato, mais geral (diferentes aplicabilidades) e capaz de gerar muito mais informação (WOLD & MYERS, 2008). Este método propiciou uma nova maneira de analisar transcriptomas (RNA-Seq), sítios genômicos de ligação de proteínas (CHIP-Seq) e padrões de metilação do genoma (Metil-Seq).

Embora melhoramentos nas técnicas de sequenciamento de DNA de segunda geração continuem a impressionar, a grande novidade são as tecnologias conhecidas como sequenciamento de DNA de terceira geração (MUNROE & HARRIS, 2010). Estas plataformas possuem algumas características cruciais: sequenciamento de moléculas únicas (não é necessário amplificação), baixo custo por base, simples preparação do material, tempos de corrida bem menores e leituras longas (centenas a milhares de pares de bases). Com essas novas tecnologias, a utilização da genômica personalizada na medicina está muito próxima, sendo que o sequenciamento de um genoma humano por U\$1.000 deve ser possível em 2011 (MUNROE & HARRIS, 2010).

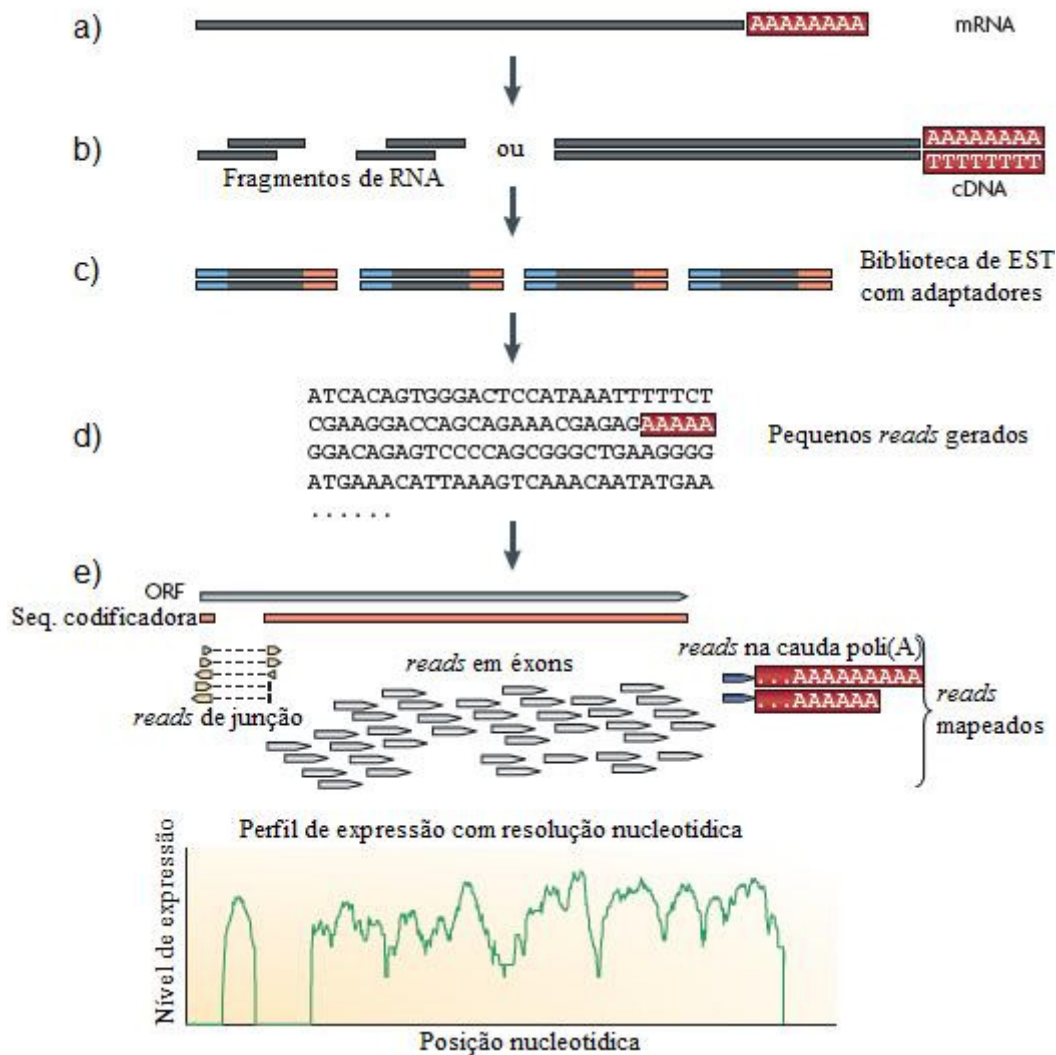
### 1.8.2 Transcriptômica: microarranjo versus RNA-Seq

Os Microarranjos de DNA foram inicialmente desenvolvidos na Universidade Stanford por Mark Schena e colaboradores em meados da década de 90. O primeiro trabalho descrevendo a técnica foi publicado na revista Science em 1995, no qual foi demonstrado a capacidade da técnica na análise de expressão diferencial de 45 genes de *Arabidopsis thaliana* em diferentes tecidos da planta (SCHENA *et al.*, 1995). Os anos subsequentes presenciaram uma explosão na aplicabilidade da técnica, destacando-se a produção de um microarranjo representando o transcriptoma completo de *S. cerevisiae* (LASHKARI *et al.*, 1997), o qual foi utilizado na análise da expressão gênica deste organismo modelo durante a esporulação (CHU *et al.*, 1998), ciclo celular (SPELLMAN *et al.*, 1998) e alteração metabólica – fermentação para respiração – (DERISI *et al.*, 1997), o desenvolvimento de um microarranjo contendo sondas para 8600 genes humanos, o qual foi utilizado no monitoramento da resposta de fibroblastos ao soro (IYER *et al.*, 1999), e estudos da biologia linfocitária (ALIZADEH *et al.*,

1998). Em conjunto com trabalhos pioneiros acerca da análise da grande quantidade de dados gerados e aplicação de algoritmos de clusterização para determinação de genes co-regulados (EISEN *et al.*, 1998; TORONEN *et al.*, 1999; TAMAYO *et al.*, 1999; TAVAZOIE *et al.*, 1999), o grande potencial da técnica de microarranjos de DNA foi afirmado, abrindo caminho para a era pós-genoma.

Mais recentemente, a metodologia “censo de sequencia” proporcionou uma nova maneira de mapear e quantificar transcriptomas (revisto por WANG *et al.*, 2009). O método conhecido como RNA-Seq baseia-se no isolamento de RNA celular (total ou poli(A)+), conversão do mesmo em uma biblioteca de fragmentos de cDNA ligados a adaptadores em ambas as extremidades e seqüenciamento de cada uma das moléculas usando uma das técnicas de seqüenciamento em larga escala. Os *reads* resultantes são alinhados a um genoma referência, resultando em um mapa transcricional em escala genômica que contém informação tanto da estrutura física dos transcritos quanto do nível de expressão para cada gene (veja Figura 1.12). Embora o RNA-Seq ainda seja uma tecnologia em desenvolvimento, ela já oferece diversos avanços sobre técnicas existentes para análise de transcriptoma. Diferentemente de métodos baseados em hibridização (microarranjo e northern-blot), RNA-Seq não necessita de conhecimento *a priori* do transcriptoma e não se limita a avaliação de genes para os quais existam sondas; desta forma, novos transcritos e novas variantes de *splicing* podem ser identificados. Além disso, é possível determinar polimorfismos em regiões transcritas com resolução de um nucleotídeo (revisto por WANG *et al.*, 2009). Outra vantagem em relação aos microarranjos de DNA é que RNA-Seq é muito pouco influenciado por sinal de fundo (*background*), visto que as sequências geradas normalmente podem ser mapeadas inequivocadamente a uma única região no genoma.





**FIGURA 1.12 – RNA-SEQ**

a) a amostra de RNA utilizada normalmente é RNA poli(A)+; b) a biblioteca a ser seqüenciada é formada por fragmentação do RNA seguida de síntese de cDNA usando *random primer*, ou síntese inicial do cDNA com oligo-dT seguida da fragmentação do cDNA; c) biblioteca de pequenos fragmentos a ser seqüenciada, EST: *expressed sequence tags*; d) *reads* gerados por MPS; d) exemplo de alinhamento dos *reads* ao genoma ou transcriptoma referência. Os *reads* são classificados de 3 formas: *reads* em éxons (dentro de éxons conhecidos), *reads* de junção (passam por dois éxons distintos) e *reads* na cauda poli(A)+. Estes três tipos de *reads* são usados para gerar um perfil de expressão com resolução nucleotídica para cada gene. É mostrado uma ORF (*open reading frame*) com apenas um íntron.

FONTE: modificado de WANG *et al.*, 2009.

RNA-Seq não possui um limite máximo de detecção, o qual se correlaciona com o número de seqüências obtidas, de forma que os limites do alcance dinâmico (*dynamic range*) são determinados apenas pela quantidade de leituras obtidas. Isso significa que através do seqüenciamento contínuo de uma determinada biblioteca, seria possível medir a expressão de qualquer transcrito presente e, portanto, o alcance dinâmico só representaria a diversidade biológica do transcriptoma analisado (revisto por WILHELM & LANDRY, 2009). Um

alcance dinâmico de mais de 5 ordens de grandeza foi estimado para 40 milhões de *reads* de camundongo (MORTAZAVI *et al.*, 2008) e na análise de mais de 35 milhões de *reads* de uma linhagem celular cancerígena (HASHIMOTO *et al.*, 2009). Em contraste, microarranjos de DNA perdem sensibilidade para genes expressos em baixos níveis (sinal da sonda é sobrepujada pela fluorescência de fundo do microarranjo) ou em altos níveis (saturação do sinal da sonda), tendo um alcance dinâmico somente de cerca de 3 ordens de grandeza (revisto por WANG *et al.*, 2009). A capacidade de detecção de transcritos raros (menos que 5 cópias por célula) por RNA-Seq, os quais podem representar cerca de 60% dos genes expressos numa célula, acarreta na detecção de expressão de 25% mais genes em comparação com microarranjo (HASHIMOTO *et al.*, 2009).

Como RNA-Seq é quantitativo, ele pode ser usado para determinar o nível de expressão gênica de forma mais acurada que em microarranjos. Em princípio, é possível determinar a quantidade absoluta de cada molécula de mRNA numa população de células (quantificação absoluta) e comparar os resultados entre experimentos independentes, diferentemente de microarranjos nos quais normalmente é feito uma quantificação relativa entre duas amostras (revisto por WANG *et al.*, 2009). De forma a determinar o nível de expressão para cada gene, o número de leituras de RNA-Seq deve ser convertido num valor quantitativo, o que é normalmente feito dividindo o número total de leituras que mapeiam em um gene sobre o tamanho do gene. Para possibilitar a subsequente comparação entre bibliotecas, também é feita uma normalização em relação ao número total de seqüências obtidas por biblioteca (total de leituras por condição de análise) (WILHELM & LANDRY, 2009). De fato, a quantificação absoluta por RNA-Seq tem se mostrado muito acurada na quantificação de níveis de expressão, como determinado por PCR quantitativa (PCR em tempo real) (NAGALAKSHMI *et al.*, 2008), além de ser muito reprodutível (MORTAZAVI *et al.*, 2008; NAGALAKSHMI *et al.*, 2008; revisto por WANG *et al.*, 2009). Finalmente, como não há passos de clonagem, RNA-Seq requer pequenas quantidades de amostras, sendo que a análise do transcriptoma de células únicas foi recentemente publicada (TANG *et al.*, 2009).

Considerando as vantagens descritas acima, RNA-Seq vem se consolidando como o método padrão para análise em larga escala de transcriptomas (revisto por WILHELM & LANDRY, 2009). Entretanto, inúmeros desafios ainda permeiam o desenvolvimento da técnica. Ainda não há um consenso no método para produção da biblioteca de fragmentos para seqüenciamento a partir do transcriptoma (Figura 1.12). Alguns trabalhos usam

fragmentação do cDNA, após sua síntese com oligo-dT, o que resulta em uma maior representação da extremidade 3' dos mRNAs originais nos *reads* gerados (TANG *et al.*, 2009; SULTAM *et al.*, 2008), enquanto outros trabalhos fragmentam o RNA a ser analisado para posteriormente sintetizar cDNA com *random primer*, o que ocasiona uma cobertura de sequenciamento mais homogênea ao longo de um gene mas com sub-representação de ambas as extremidades do transcrito (MARIONI *et al.*, 2008; CLOONAN *et al.*, 2008; revisto por WANG *et al.*, 2009).

RNA-Seq também possui diversos desafios bioinformáticos, incluindo métodos eficientes para armazenar, acessar e processar a imensa quantidade de dados gerados pelas técnicas de seqüenciamento em paralelo massivo. Além disso, novos programas para alinhamento simultâneo de milhões de pequenas leituras contra um genoma (ou transcriptoma) referência são necessários, sendo que diversos programas já vêm demonstrando ótimos resultados (PALMIERI & SCLOTTERER, 2009). Para genomas contendo muitas famílias gênicas, com parálogos similares, o alinhamento de pequenas leituras de cerca de 35pb pode ser muito complexo, gerando o problema de alinhamento em regiões múltiplas do genoma (revisto por WANG *et al.*, 2009). Este problema pode ser resolvido aumentando-se o tamanho das leituras ou com a utilização do método de seqüenciamento de ambas as extremidades de um fragmento de DNA de tamanho conhecido (*paired-end sequencing*) (Fullwood *et al.*, 2009). Além disso, para transcriptomas complexos, o alinhamento de leituras pequenas pode ser muito dificultado devido a variantes alternativas de *splicing* e leituras de junção (veja Figura 1.12) (TANG *et al.*, 2009; SULTAM *et al.*, 2008).

Além da técnica de RNA-Seq, a qual foca em RNAs codificadores de proteínas, o transcriptoma pode ser avaliado de diferentes formas e com diferentes intuitos (revisto por MOROZOVA *et al.*, 2009). Pequenos RNAs não codificadores (18 a 30 nucleotídeos) foram recentemente reportados como reguladores pós-transcricionais cruciais da expressão gênica, sendo que a análise dos mesmos por seqüenciamento em larga escala permite a descoberta de novos transcritos bem como a análise quantitativa desta importante classe de moléculas. O seqüenciamento de ambas as extremidades de um cDNA de tamanho conhecido (*paired-end sequencing*) (FULLWOOD *et al.*, 2009) permite a descoberta de transcritos aberrantes gerados por rearranjos genômicos, os quais são comuns em células cancerígenas.

### 1.8.3 Organização genômica do *T. cruzi*

O Projeto Genoma do *Trypanosoma cruzi* foi iniciado em 1994 por iniciativa do WHO *Special Programme for Research and Training in Tropical Diseases* com o objetivo de estabelecer um detalhado mapa físico e essencialmente toda sequência do genoma do agente etiológico da Doença de Chagas. O clone CL Brener (DTU IIe), derivado em 1987 da cepa CL isolada no sul do Brasil e há longo tempo estudada, foi selecionado como cepa referência do Projeto Genoma por ser bem caracterizado (BRISSE *et al.*, 1998).

A primeira versão da montagem do seqüenciamento do genoma de *T. cruzi* foi publicada recentemente (EL-SAYED *et al.*, 2005a). A montagem do genoma foi difícil, pois o Clone CL Brener possui origem híbrida e, conseqüentemente, um genoma muito complexo. Tal seqüenciamento foi realizado pela técnica de “whole genome shotgun” (WGS), sendo que o alto conteúdo de repetições do genoma, aliado à sua natureza híbrida, limitaram a técnica de seqüenciamento progressivo de clones em cromossomos artificiais de bactérias. Inicialmente, estimava-se o conteúdo repetitivo do genoma em 35%, mas com o seqüenciamento, mostrou-se ser acima de 50% (EL-SAYED *et al.*, 2005a).

No total, foi produzida uma cobertura genômica de 19 vezes. Os parâmetros de montagem tiveram de ser adaptados para lidar com a variabilidade dos alelos divergentes entre as cepas parentais, e também foi necessário gerar posteriormente uma cobertura genômica de 2,5 vezes da cepa Esmeraldo, pertencente ao subgrupo IIb (uma das linhagens parentais da cepa CL Brener), para que fosse possível distinguir os dois haplótipos e auxiliar na montagem.

Com base em análise densitométrica da separação por PFGE (CANO *et al.*, 1995), estimava-se que a cepa CL Brener continha aproximadamente 64 cromossomos, com um genoma de 87 Mb. Entretanto, os resultados do projeto genoma de *T. cruzi* evidenciaram que o genoma diplóide possui um tamanho entre 106,4 e 110,7 Mb. Desse total, 67 Mb foram selecionados para anotação, já que o restante contém uma grande quantidade de genes codificadores de moléculas de RNA, como rRNA e SLRNA, geralmente organizados em tandem.

A montagem destes 67Mb consiste de 5489 scaffolds, contendo 8740 contigs, os quais demonstram muito bem o estado muito fragmentado da atual versão do genoma. A análise da fração anotada, que perfaz 60,7 Mb, demonstra que 30,5 Mb são compostos de sequências encontradas pelo menos duas vezes, o que sugere que elas são oriundas dos diferentes

haplótipos (Iib e Iic) que compõe o genoma da cepa CL Brener. Os dois haplótipos demonstram alto nível de sintenia, com a maioria das diferenças causadas por inserções ou deleções em regiões intergênicas e subteloméricas, ou por ampliações de sequências repetitivas. A divergência média entre os dois haplótipos é de 5,4%, sendo que as regiões codificadoras de proteínas são consideravelmente mais conservadas (2,2% de diferenças) (EL-SAYED *et al.*, 2005a).

Com base na análise dos haplótipos, estimou-se que o genoma haplóide contenha cerca de 12.000 genes. Atualmente, 22.570 modelos gênicos foram preditos, sendo que 6.159 representam alelos presentes no haplótipo Iib, 6.043 representam alelos do outro haplótipo e 10.368 representam sequências que não puderam ser atribuídas a um haplótipo em particular (EL-SAYED *et al.*, 2005a).

Pelo menos 50% do genoma de *T. cruzi* contém sequências repetitivas, as quais são geralmente compostas por famílias gênicas de proteínas de superfície, retrotransposons e sequências teloméricas. As maiores famílias gênicas codificam proteínas de superfície associadas à mucina (MASP), membros da superfamília de trans-sialidases (TS), mucinas e a glicoproteína de superfície gp63, as quais são frequentemente *T. cruzi*-específicas e somam cerca de 18% dos genes codificadores de proteínas. Esses genes ocorrem em agrupamentos dispersos em tandem com repetições intercaladas, frequentemente em regiões subteloméricas (EL-SAYED *et al.*, 2005a).

Um dos resultados mais impressionantes, com relação à análise das vias biológicas presentes em *T. cruzi*, foi a ausência de diversas classes de moléculas de sinalização, incluindo receptores serpentina, proteínas G heterotriméricas, muitas classes de receptores catalíticos, domínios de interação SH2 e SH3 e fatores de transcrição regulatórios. Entretanto, *T. cruzi* possui um grande e complexo grupo de proteína-quinases (PKs), assim como uma diversidade de proteínas fosfatases. O genoma de *T. cruzi* codifica 167 proteína-quinases distintas potencialmente ativas, sendo esse número duas vezes maior que o quinoma de *P. falciparum* (EL-SAYED *et al.*, 2005a) e 30% maior que o de *S. cerevisiae*. A maioria das proteína-quinases de tripanossomatídeos não contém domínios acessórios discerníveis, em contraste com as quinases de mamíferos, nos quais esse padrão é muito comum. No entanto, a maioria das proteína-quinases de tripanossomatídeos são muito maiores do que um domínio catalítico simples, sugerindo a possível existência de domínios acessórios ainda não identificados.

O genoma de outros dois tripanossomatídeos, *T. brucei* (BERRIMAN *et al.*, 2005) e *L. major* (IVENS *et al.*, 2005), também foi recentemente publicado, permitindo fazer comparações entre os três patógenos. Mesmo tendo divergido há mais de 200 milhões de anos (OVERATH *et al.*, 2001; EL-SAYED *et al.*, 2005b), os genomas dos tripanossomatídeos apresentam um alto grau de sintenia. Assim sendo, a elucidação de vias de sinalização críticas para o reparo de DNA, replicação e meiose e a identificação de numerosas proteínas quinases e fosfatases, proporcionadas pelo estudo dos genomas de tripanossomatídeos, indica novos alvos para o desenvolvimento de drogas contra estes parasitas, sendo que o trabalho já está bem avançado nessa área (EL-SAYED *et al.*, 2005a).

Em um trabalho recente (WEATHERLY *et al.*, 2009), os fragmentos genômicos (contigs e scaffolds) gerados pelo projeto genoma foram montados em cromossomos baseados em mapas de sintenia de TriTryps e no sequenciamento de bibliotecas de cromossomos artificiais de bactérias (BAC) de *T. cruzi*, trazendo o genoma deste parasita para um nível similar de montagem e anotação em relação aos demais TriTryps. O resultado é um genoma com 41 pares cromossômicos, variando de 78 kb a 2,4 Mb (WEATHERLY *et al.*, 2009).

#### 1.8.4 O projeto Reguloma de *T. cruzi*

O grupo do Instituto Carlos Chagas (ICC) vem analisando a expressão gênica de *Trypanosoma cruzi* em larga escala. Neste contexto, etapas de análise pós-genômica começaram a ser feitas pela equipe do IBMP desde 2001, tendo culminado com o desenvolvimento de um microarranjo contendo milhares de sequências gênicas diferentes de *T. cruzi*, resultando numa cobertura de boa parte dos genes transcritos do parasita. Desde 2003 possuímos a capacidade plena do processo de produção e análise dos dados do microarranjo de *T. cruzi*. Além de todo o processo de produção e análise estar sendo feita no ICC, o gerenciamento também é facilitado pelo uso do software *CruziGeneDB*, o qual foi inteiramente desenvolvido no laboratório de Bioinformática do instituto, servindo para cobrir uma lacuna presente na maioria dos softwares voltados para a área de microarranjos (PROBST, 2005).

Esta técnica de microarranjo se mostrou muito poderosa para avaliar de forma sistêmica as modificações nos níveis de mRNA de uma célula, permitindo observar

características gerais que não são vistas em trabalhos mais focados. Estudos preliminares do transcriptoma do parasita em diversos estágios do seu ciclo de vida e durante o processo de diferenciação de epimastigotas a tripomastigotas metacíclicos (metaciclogênese), permitiram estabelecer o perfil de mRNAs expressos nos diferentes momentos estudados. Os dados de microarranjo indicam diversos genes diferencialmente expressos durante a metaciclogênese e ciclo de vida do *T. cruzi*, e que se tornam, portanto, interessantes para estudos por codificarem proteínas estágio específicas, as quais poderiam ser investigadas como alvos potenciais para a busca de alternativas quimioterápicas no tratamento da doença de Chagas (revisto por ÁVILA *et al.*, 2003). Esses dados de expressão diferencial, em nível de mRNA, foram associados com análises do proteoma, permitindo a identificação de uma grande quantidade de proteínas diferencialmente expressas em *T. cruzi* (PROBST *et al.*, em preparação).

Entretanto, a maior parte desses genes com padrão de expressão estágio específico (variando de 50 a 75% dependendo do estágio) encontra-se na categoria de genes que codificam proteínas hipotéticas ou proteínas hipotéticas conservadas, de forma que não se conhecem suas funções (PROBST, 2005). Assim sendo, embora as abordagens de microarranjo e proteômica permitam inferências acerca da função de grupos de genes co-regulados e abra diversas perspectivas para estudos da biologia deste parasita, necessita-se ampliar a quantidade de informações para predição mais confiável da função gênica. Nesse contexto, a possibilidade de se modular a expressão gênica de *T. cruzi in vitro*, através do uso de drogas, é uma importante ferramenta que pode gerar dados cruciais para a determinação de redes de genes co-regulados em *T. cruzi*, e que, portanto, devem ter funções associadas.

Com isso em mente, atualmente o ICC está desenvolvendo o Projeto Reguloma de *Trypanosoma cruzi*. O termo reguloma diz respeito ao conjunto de elementos que regulam a expressão de diferentes grupos de genes numa célula, ou seja, são os elementos (genes, proteínas reguladoras, sequências regulatórias, entre outros) que controlam as funções biológicas e celulares de um organismo. Dessa forma, tal projeto tem como objetivo geral a identificação de redes gênicas regulatórias (grupo de genes co-regulados e seus elementos regulatórios) do parasita através da modulação do transcriptoma do mesmo com o uso de modificadores moleculares (drogas tripanocidas). A integração dos dados de transcriptômica relativos a diversos fármacos, através de ferramentas computacionais, propiciará análises das redes gênicas regulatórias do protozoário *T. cruzi*, de forma a elucidar fenômenos biológicos básicos particulares ao parasita bem como identificar elementos essenciais para a determinação de alvos terapêuticos.

O projeto como um todo pode ser dividido em duas etapas principais, com distintos objetivos e graus de abrangência do ponto de vista da análise da regulação da expressão gênica de *T. cruzi*:

1. Etapa inicial focada na determinação do transcriptoma do parasita frente a diferentes moduladores moleculares, de forma a estudar as características biológicas individuais da resposta do *T. cruzi* a cada uma das substâncias tripanocidas utilizadas, o que podemos chamar de farmacogenômica.
2. Com os dados individuais de cada fármaco obtidos, a etapa final do projeto terá como objetivo a identificação de módulos de co-regulação, visando à caracterização de elementos e mecanismos reguladores da expressão gênica do *Trypanosoma cruzi*, bem como identificar módulos de genes com função comum possibilitando a atribuição indireta de função a proteínas hipotéticas.

A primeira etapa está diretamente relacionada aos resultados da modulação de cada fármaco, enquanto que a segunda etapa se apóia na utilização de métodos computacionais na base de dados produzidos para a identificação de padrões de regulação não perceptíveis nos trabalhos individuais.

Os moduladores moleculares que serão utilizados são drogas tripanocidas *in vitro* por alterarem (inibirem) mecanismos moleculares essenciais do parasita, como a biossíntese de esteróis, funcionalidade de proteíno-quinases, entre outros.



## 2 OBJETIVOS

O presente trabalho se propõe a estudar a resposta biológica de *Trypanosoma cruzi* frente à inibição da síntese de esteróis endógenos através da análise do padrão de expressão gênica do parasita em larga escala (RNA-Seq). Tal trabalho representa os primeiros esforços na caracterização de módulos regulatórios do parasita, enquadrando-se, portanto, como um sub-projeto dentro do Projeto Reguloma de *Trypanosoma cruzi*, sendo que os objetivos principais são de origem biológica e bioinformática, e podem ser sumarizados da seguinte forma:

### 2.1 OBJETIVOS GERAIS

- Caracterizar o transcriptoma de *T. cruzi* em resposta a duas drogas inibidoras da biossíntese de esteróis: cetoconazol e lovastatina;
- Estudar as características biológicas da resposta do *T. cruzi* frente à inibição da biossíntese de esteróis pelos fármacos.

### 2.2 OBJETIVOS ESPECÍFICOS

- Determinar a suscetibilidade de diferentes cepas de *T. cruzi* (Dm28c e Y) aos IBEs (inibidores da biossíntese de esteróis) utilizados;
- Analisar alterações morfológicas de epimastigotas em cultura, através de microscopia óptica e eletrônica, em resposta aos IBEs utilizados;
- Identificar a modulação da expressão gênica do parasita frente às drogas, por RNA-Seq;
- Comparar os resultados de expressão diferencial com a base de dados de Genômica Funcional de *T. cruzi*, visando identificar módulos de co-regulação na resposta do parasita ao tratamento inibitório.

### 3 JUSTIFICATIVA

A doença de Chagas é um importante problema de saúde, afetando cerca de 15 milhões de pessoas em diferentes regiões das Américas. Por se tratar de uma doença de grandes proporções, o estudo de seu agente etiológico, o *Trypanosoma cruzi*, tornou-se de grande interesse, de forma a entender as características biológicas básicas deste parasita, bem como para desenvolver novos quimioterápicos, já que os existentes não são satisfatórios devido aos seus efeitos colaterais.

Além disso, pouco se sabe a respeito da regulação da expressão gênica e sobre a função da maioria dos genes de *T. cruzi*, portanto, a metodologia de análise da expressão gênica em larga escala utilizada no presente projeto (RNA-Seq) pode fornecer importantes dados para um maior entendimento dos mecanismos de regulação gênica pós-transcricional deste organismo, bem como um maior entendimento da biologia básica do parasita. A experiência de nosso grupo com microarranjos, associado ao uso de uma população de mRNA (polissomal) que está se mostrando muito informativa, torna a realização desses objetivos viável.

Além disso, tendo em vista a importância da via de biossíntese de esteróis como um dos principais alvos no desenvolvimento de novos agentes quimioterápicos na Doença de Chagas, o presente projeto propiciará um enfoque em larga escala nunca antes desenvolvido acerca da resposta de *T. cruzi* a inibidores desta via. Tal metodologia pode fornecer um entendimento sistêmico acerca da importância e interconexões da síntese de esteróis no parasita, propiciando novos alvos e caminhos no desenvolvimento terapêutico. Além disso, os mecanismos distintos de controle da via de síntese de esteróis em *T. cruzi* oferecem uma ótima oportunidade para estudar respostas homeostáticas nesse parasita.

## 4 MATERIAIS E MÉTODOS

### 4.1 ORIGEM DOS REAGENTES

**Amresco:** detergente não iônico Nonidet<sup>®</sup> NP-40;

**Invitrogen:** sacarose;

**Sigma:**  $\beta$ -mercaptoetanol, cicloheximida, DEPC (dietil pirocarbonato), DMSO (dimetil sulfóxido), heparina, hepes, Cetoconazol, Lovastatina, Rhodamina 123, Laranja de Acridina.

### 4.2 MEIO DE CULTURA

**Meio LIT (*Liver Infusion Tryptose*):** extrato de levedura 15 g/l, fosfato dibásico de sódio 11,56 g/l, glicose 2,2 g/l, hemina 0,02 g/l, infuso de fígado 5 g/l, KCl 0,4 g/l, NaCl 4,4 g/l, soro fetal bovino 10 % (v/v), triptose 5 g/l. O pH é ajustado para 7,2 com HCl.

### 4.3 SOLUÇÕES E TAMPÕES

**Água DEPC (livre de RNase):** água ultrapura tratada com DEPC (dietil pirocarbonato), na proporção 1:1000, por 30 minutos à temperatura ambiente, e autoclavada por 15 minutos.

**Colchão de sacarose:** sacarose 2 M, em tampão hipotônico de lise celular.

**Tampão hipotônico de lise celular:**  $\beta$ -mercaptoetanol 5 mM, cicloheximida 10  $\mu$ g/ml, heparina 20 U/ml, MgCl<sub>2</sub> 5 mM, NaCl 10 mM, tris-HCl pH 7,6 10 mM. Cicloheximida, heparina e  $\beta$ -mercaptoetanol são adicionados somente antes do uso.

**Tampão NKM:** cicloheximida 10  $\mu$ g/ml, hepes pH 7,4 10 mM, KCl 5 mM, MgCl<sub>2</sub> 1,5 mM, NaCl 140 mM, em água DEPC.

**Tampão PBS (10x):** KCl 2,7 mM, KH<sub>2</sub>PO<sub>4</sub> 1,5 mM, Na<sub>2</sub>HPO<sub>4</sub>.7H<sub>2</sub>O 4,3 mM, NaCl 137mM.

#### 4.4 CULTIVO DO PARASITA

As culturas axênicas de formas epimastigotas de *T. cruzi* eram mantidas a 28°C através de passagens semanais em meio LIT. Para obtenção de parasitas em fase logarítmica de crescimento, eram realizadas requiques a cada três dias, mediante inóculo de  $1 \times 10^6$  células por mililitro de solução para a cepa Dm28c (CONTRERAS *et al.*, 1988) e  $2 \times 10^6$  células por mililitro de solução para a cepa Y (CAMARGO, 1964). Culturas de três dias (fase mid-log) eram usadas para iniciar todos os experimentos. Tais condições são estabelecidas segundo a curva de crescimento do parasita em meio LIT a 28°C (CAMARGO, 1964).

#### 4.5 INIBIDORES DA BIOSÍNTESE DE ERGOSTEROL (IBES)

O inibidor da enzima 14-desmetilase, cetoconazol, e da enzima 3-hidroxi-3-metilglutaril-CoA, lovastatina (mevenolina), foram obtidos da SIGMA-ALDRICH® com os IDs K1003 e M2147, respectivamente. Os compostos foram diluídos em 100% DMSO (dimetil sulfóxido), de forma a obter soluções estoque à 50 mM.

#### 4.6 ATIVIDADE ANTIPROLIFERATIVA *IN VITRO* DOS IBES

A análise antiproliferativa das drogas foi realizada através de curvas de crescimento de epimastigotas de *T. cruzi* em cultura em diferentes concentrações das drogas. As culturas eram iniciadas com uma densidade celular de  $2 \times 10^6$  cel/ml e as drogas eram adicionadas no dia seguinte, quando a densidade celular era  $4-6 \times 10^6$  cel/ml. A densidade celular era analisada por contagem direta em hemocitômetro (câmara de Neubauer), a cada 24 horas, durante sete dias. A viabilidade celular era analisada pela morfologia e motilidade celular. Os experimentos foram realizados em triplicata, e foram utilizados como controles as curvas de crescimento do parasita sem droga e somente com DMSO. A concentração final de DMSO na cultura nunca passou de 0,1% e não teve efeito algum no crescimento dos parasitas.

A inibição do crescimento foi quantificada definindo-se um percentual de fator de crescimento (FC) (Sanz-Rodríguez *et al.*, 2007):  $\%FC = (FC_{droga}/FC_{controle}) \times 100$ , onde  $FC = (DC_{xh} - DC_{0h})/DC_{0h}$ , sendo que  $FC_{droga}$  é o fator de crescimento numa certa concentração de droga e  $FC_{controle}$  é o fator de crescimento para células não tratadas, e

DC<sub>xh</sub> é a densidade celular da cultura em x horas após a adição da droga. Portanto, o percentual de inibição do crescimento (%IC) é:  $\%IC = 100 - \%FC$ .

A concentração experimental necessária para reduzir a proliferação do parasita em 50%, IC<sub>50</sub>, foi calculada com base nos dados de %IC versus concentração da droga, por regressão não linear com o programa Graphad prism. Para tal, foram utilizados os dados do quarto dia de cultura, 3 dias de tratamento inibitório (72 horas).

Além da determinação da faixa de concentração de droga capaz de inibir o crescimento de epimastigotas em 72 horas de exposição, e respectiva IC<sub>50</sub>, foram determinadas as concentrações mínimas das drogas capazes de matar todas as células em cultura dentro de 24 horas de exposição. Para tal, considerou-se as células mortas quando esféricas e imóveis (AMEISEN *et al.*, 1995; KOSEC *et al.*, 2006), além de incapazes de recuperar o crescimento após retirada da droga (experimentos de recuperação, veja tópico 4.6.1). Denominamos tal dose como DL100 (dose letal para 100% das células em cultura). Diferentemente da IC<sub>50</sub>, esta dose foi obtida de forma empírica, através do teste de algumas altas concentrações das drogas.

As concentrações efetivas das drogas (IC<sub>50</sub> e DL100) foram usadas para obtenção das amostras de RNA polissomal para os experimentos de RNA-Seq.

#### 4.6.1 Experimentos de recuperação

Para analisar os tempos de exposição às drogas na qual os parasitas estariam comprometidos com a morte celular, foram realizados experimentos de retirada de droga do meio de cultura após determinados tempos de exposição a altas concentrações das mesmas. Cerca de  $5 \times 10^6$  células, em seus respectivos meios de cultura contendo droga, foram recuperados por centrifugação a  $7.000 \times g$  por 5 minutos e lavados três vezes com PBS. As células foram passadas para novo meio de cultura sem droga na densidade de  $2 \times 10^6$  cel/ml e o comprometimento com a morte celular foi acompanhado por contagens diárias da densidade celular.

## 4.7 MICROSCOPIA

As alterações morfológicas e estruturais de *T. cruzi* frente a inibição da biossíntese de esteróis foram analisadas por três tipos distintos de microscopia:

### 4.7.1 Microscopia óptica

Os parasitas eram recuperadas por centrifugação a 7000 x g e lavados duas vezes com PBS. As células eram fixadas por 30 minutos com paraformaldeído 4% e novamente lavadas com PBS. Cerca de  $5 \times 10^5$  células eram depositadas sobre lâminas de microscopia e, após secagem ao ar, os esfregaços eram cobertos com metanol 100% e deixados secar. Após lavagem em água e secagem ao ar, os esfregaços eram clarificados através da imersão das lâminas em HCl 5 N por 10 minutos. As lâminas eram lavadas com água e, após secagem ao ar, coradas pelo procedimento Panótico rápido LB (Laborclin). Este kit consiste de três soluções de coloração, nas quais os esfregaços eram imergidos por 1,5 minutos. A solução 1 consiste de triarilmetano 0,1%; a solução 2 possui corantes xantenos à 0,1% e a solução 3 possui corantes tiazinas à 0,1%. Após imersão seqüencial nas três soluções, as lâminas eram lavadas com água e, depois da secagem dos esfregaços, as lâminas eram finalizadas através da montagem das lamínulas sobre uma gota de Permount® (Fisher Scientific).

Os esfregaços foram observadas em microscópio de Luz (Nikon), e as imagens capturadas usando a câmara CoolSnap (Media Cybernetics) com auxílio do programa Image Pro-Plus v. 4.5.1.22 (Media Cybernetics). Finalmente, os quadros de imagens foram montados usando o programa Adobe Photoshop 7.0.1.

### 4.7.2 Microscopia eletrônica de transmissão

Os parasitas foram recuperados por centrifugação a 7000 x g, lavados duas vezes com PBS e fixados por uma hora em glutaraldeído a 2,5% em tampão fosfato 0,1 M pH 7,2. Após duas lavagens em tampão fosfato 0,1 M, as células foram pós-fixadas por uma hora em solução contendo 1% tetróxido de Ósmio/ 0,8% ferricianeto de potássio/ 5 mM cloreto de cálcio / tampão cacodilato 0,1 M pH 7,2. Em seguida os parasitas foram desidratados em concentrações crescentes de acetona (50%, 70%, 90% e 100%) e incluídos *overnight* em

mistura 1:1 de acetona 100% e resina PolyBed 812. As amostras foram então incluídas em resina PolyBed 812 pura por 4 a 6 horas e emblocadas para polimerizar por 48 horas a 60°C.

Cortes ultrafinos foram obtidos no ultramicrotomo Leica EM UC6 e contrastados por 40 minutos em acetado de uranila a 5% e por 3 minutos em citrato de chumbo. Finalmente, as amostras foram analisadas em um microscópio eletrônico de transmissão Jeol 1200 EXII. Os quadros de imagens foram montados usando o programa Adobe Photoshop 7.0.1.

#### 4.7.3 Microscopia de fluorescência

O corante fluorescente acidotrófico LysoTracker® Red DND-99 (Invitrogen) foi utilizado para monitorar a presença de compartimentos ácidos no parasita em resposta as drogas. Cerca de  $1 \times 10^7$  células foram recuperadas por centrifugação a 7000 x g por 5 minutos e lavadas com PBS. As células foram incubadas por 30 minutos a 28°C em PBS contendo 0,5  $\mu$ M de LysoTracker® Red DND-99. A marcação foi interrompida através da fixação com 4% de paraformaldeído por uma hora. Após lavagem com PBS, os parasitas foram depositados, em poços limitados por Teflon, em lâminas de imunofluorescência cobertas com 0,1% de Poli-L-lisina e deixados aderir por 30 minutos. Em seguida as lâminas eram lavadas por imersão em PBS para retirada do excesso de parasitas não aderidos, e o DNA era marcado por 5 minutos pela adição de 20  $\mu$ l do corante Hoechst 33342 (Invitrogen) a 2  $\mu$ g/ $\mu$ l em PBS. Após cinco lavagens por imersão em PBS, as lâminas eram finalizadas pela adição de 10  $\mu$ l de n-propil-galato (solução anti-*fade*) e lacradas utilizando lamínulas de microscopia e selador (esmalte). As amostras foram abservadas em um microscópio confocal Leica SP5, sendo que as imagens obtidas foram posteriormente processadas com auxílio do programa Adobe Photoshop 7.0.1.

#### 4.8 CITOMETRIA DE FLUXO

Experimentos de citometria de fluxo foram realizados para monitorar a resposta biológica de *T. cruzi* frente a inibição da biossíntese de esteróis. O corante Rhodamina 123 foi utilizado para análise do potencial transmembrana mitocondrial e o fluoróforo acidotrófico laranja de acridina para monitoramento de vesículas ácidas no parasita. Para ambos os corantes o procedimento experimental foi similar. Pelo menos  $5 \times 10^6$  células eram recuperadas

por centrifugação a 7.000 x g e lavadas em PBS. A seguir, os parasitas eram incubados com PBS contendo 5 µg/ml de laranja de acridina ou 10µg/ml de Rhodamina 123. Após 15 minutos a 28°C, os parasitas eram lavados três vezes em PBS e passados para tubos de citometria. As amostras foram quantificadas em um citometro FACSCalibur, e as análises posteriores realizadas no programa WinMDI versão 2.9. Para Rhodamina 123 foram usados os dados do detector FL1 (luz verde) e para Laranja de Acridina, FL3 (luz vermelha).

Além disso, analisou-se a possível exposição do fosfolípido fosfatidilserina (característica de células apoptóticas) através da utilização do material BD Pharmingen Annexin V-FITC Apoptosis Detection Kit, segundo instruções do fabricante. Basicamente, é realizada uma marcação dupla das células com o corante vital Iodeto de Propídeo (0,5 µg/ml) e Anexina V-FITC. Devido a exposição de fosfatidilserina, a qual liga-se a Anexina V, células apoptóticas apresentam marcação somente com FITC, células integras não possuem marcação para nenhum dos corantes e células rompidas (necrose) apresentam marcação dupla. A ligação de Anexina V foi acompanhada pela fluorescência de FITC com o detector FL1, enquanto a ligação de Iodeto de Propídeo foi acompanhada pelo detector FL2. Além disso, foram realizadas marcações independentes com cada um dos corantes para compensação dos sinais dos detectores. Os dados foram analisados como descrito acima.

## 4.9 AMOSTRAS DE RNA

### 4.9.1 RNA total

Amostras de RNA total de epimastigotas de *T. cruzi* foram isoladas a partir de  $5 \times 10^7$  células através do kit RNeasy (Qiagen). A purificação foi realizada segundo instruções do fabricante utilizando o protocolo "Animal Cells I" com o passo adicional de tratamento com DNase na própria coluna (RNase-free DNase set – Qiagen).

### 4.9.2 RNA polissomal

A extração das moléculas de RNA associadas a polissomos foi realizada segundo GOLDENBERG *et al.*, 1985, com pequenas modificações. Inicialmente  $5 \times 10^8$  epimastigotas foram mantidos em seus respectivos meios de cultivo acrescidos de cicloheximida 10 µg/ml,



durante 10 minutos, a temperatura ambiente. As células foram recuperadas por centrifugação a 7000 x g, por 5 minutos, a 4°C e lavadas duas vezes em tampão NKM acrescido de cicloheximida 10 µg/ml e Hepes 10 mM pH 7,4. Os *pellets* celulares foram estocados a -70°C.

Para posterior extração do RNA polissomal, 7,9 ml de tampão de lise hipotônico, acrescido de cicloheximida 10 µg/ml, heparina 20 µg/ml e β-mercaptoetanol 5 mM, foi adicionado diretamente aos *pellets* celulares, seguido de breve ressuspensão por vortex. Em seguida foi adicionado 875 µl de detergente Nonidet P-40 e, após breve homogeneização por vortex, a lise celular foi rapidamente interrompida pela adição de sacarose 0,25 M, acrescida de cicloheximida 10 µg/ml, heparina 20 µg/ml e β-mercaptoetanol 5 mM. O volume nesse passo era 10 ml.

A seguir, foram sedimentados os núcleos por centrifugação a 10.000 x g, por 10 minutos a 4°C; e as mitocôndrias, a 10.000g, por 30 minutos a 4°C. A fração polissomal era obtida após ultra-centrifugação do sobrenadante pós-mitocondrial em um colchão de sacarose 2M (acrescido de cicloheximida 10 µg/ml, heparina 20 µg/ml e β-mercaptoetanol 5 mM), a 365.000 x g (39.000 rpm) por 2 horas a 4°C. Imediatamente após o término da centrifugação, a fração polissomal sedimentada era estocada a -70°C, para futura extração de RNA.

O RNA polissomal é finalmente purificado através do kit RNeasy (Qiagen), da mesma forma que para o RNA total (tópico acima).

#### 4.9.3 Determinação da quantidade e qualidade do RNA extraído

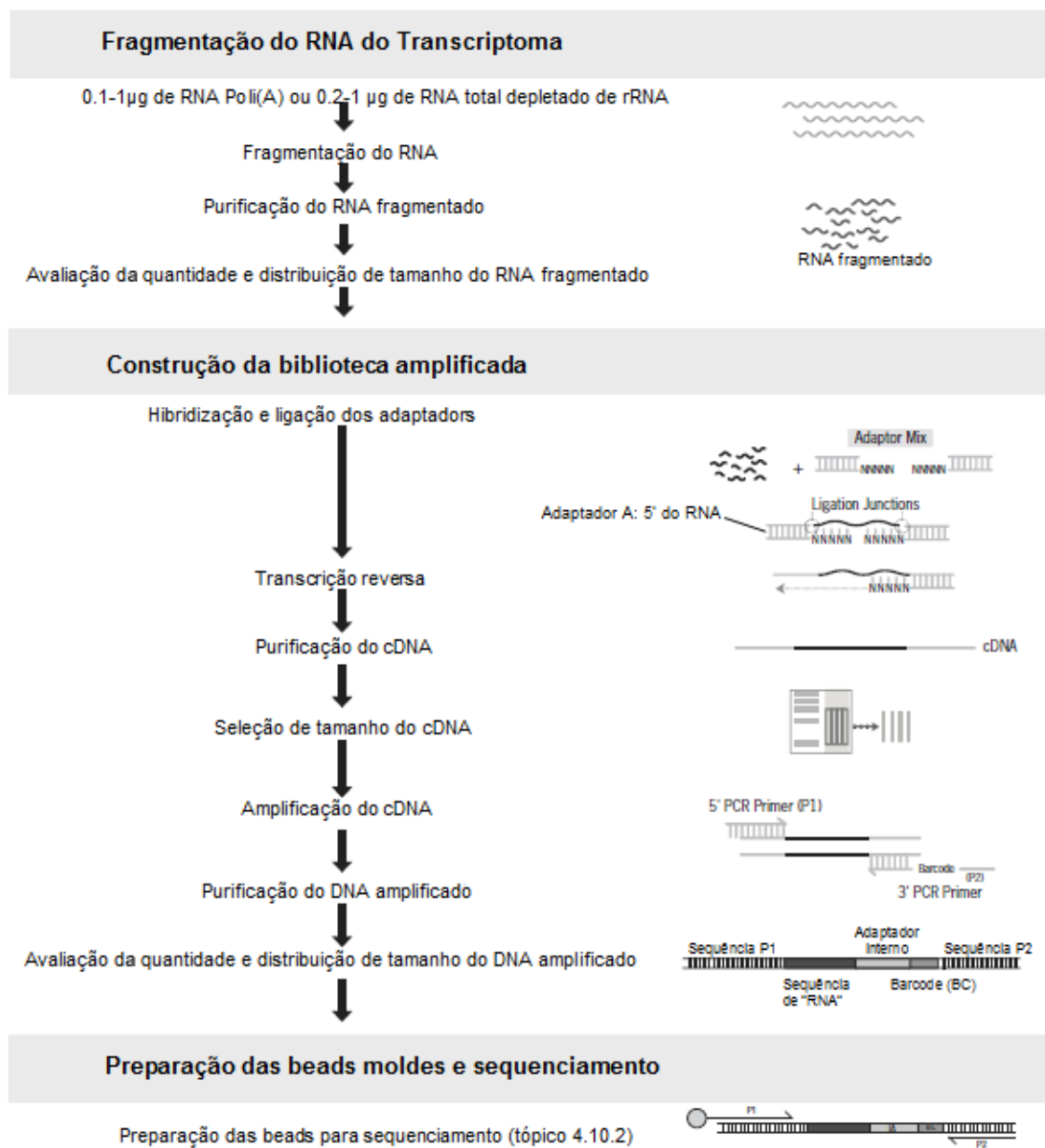
A quantidade de RNA foi determinada através da medição da absorbância a 260 nm, em um espectrofotômetro NanoDrop, considerando que o valor da absorbância igual a 1 equivale a 40 µg/ml de RNA. A qualidade dos ácidos nucleicos foi determinada pela razão entre a absorbância a 260nm e a 280nm: amostras puras devem ter tal relação maior ou igual a 2.

Além disso, as amostras de RNA foram analisadas pelo equipamento Agilent 2100 Bioanalyser, com o kit RNA 6000 Nano LabChip®, segundo especificações do fabricante. A qualidade e integridade das amostras de RNA foram investigadas pela inspeção visual dos eletroferogramas e das imagens gel-like gerados (VITALE, 2001).

## 4.10 RNA-SEQ

### 4.10.1 Preparo das bibliotecas

O preparo das amostras foi realizado segundo o protocolo *Whole Transcriptome Library Preparation for SOLiD™ Sequencing*. Na Figura 4.1 é apresentado o fluxograma de trabalho para preparação das bibliotecas contendo SOLiD™ System Barcode.



**FIGURA 4.1 – PREPARO DAS BIBLIOTECAS DE RNA-SEQ (SOLID)**

Os passos presentes na figura resumem o protocolo apresentado no tópico 4.10.1.

FONTE – modificado do manual *Whole Transcriptome Library Preparation for SOLiD™ Sequencing*.

A partir de 0,1-1 µg de mRNA poli-A, obtido através da purificação das amostras de RNA com kit Poly(A) Purist™ (Applied Biosystems), procedeu-se a fragmentação por RNase III, utilizando-se 1 µl de RNase III e 1 µl de tampão em um volume final de 10 µl. A reação foi incubada em um termociclador a 37°C por 10 minutos e interrompida pela adição de 90 µl de água sem nuclease. O RNA foi purificado através do produto *RiboMinus Concentration Module* (Invitrogen), e teve sua qualidade verificada pelo equipamento Bioanalyzer.

Com o RNA fragmentado, foi feita a hibridização e ligação dos adaptadores em ambas as extremidades dos fragmentos de RNA usando componentes do kit SOLiD Small RNA Expression. Para tal, eram utilizados 2 µl de solução contendo adaptadores, 3 µl de solução de hibridização e 50-100 ng de RNA fragmentado em 3 µl. O material era colocado em um termociclador, a 65°C por 10 minutos e a 16°C por 5 minutos; concomitantemente, era adicionado 10 µl de tampão de ligação 2X e 2 µl da mistura da enzima ligase e deixado incubar a 16°C por 16 horas.

Após a ligação dos adaptadores, foi realizada a transcrição reversa, com os componentes do kit SOLiD Small RNA Expression. Aos 20 µl da reação de ligação, adicionou-se 13 µl de água sem nuclease, 4 µl de tampão RT 10X, 2 µl de dNTP 2,5 mM, e 1 µl de transcriptase reversa, vortexou-se brevemente e incubou-se em um termociclador a 42°C por 30 minutos. Após o término da reação de transcrição reversa, o material era purificado com o kit MinElute PCR Purification (Qiagen) de acordo com as instruções do fabricante.

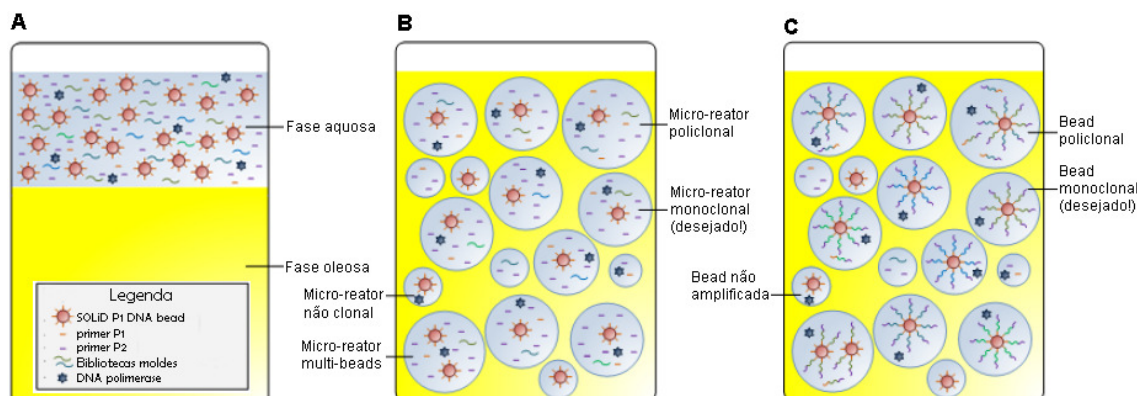
O material resultante da transcrição reversa foi separado em gel de poliacrilamida 6% desnaturante usando o sistema Novex de gel com uréia e TBE (Invitrogen). Após a corrida das amostras em uma cuba XCell SureLock™ Mini-Cell (Invitrogen), o gel foi corado com SYBR Gold (Invitrogen) e seguiu-se a seleção dos fragmentos de 100-200 bases através da excisão da banda do gel com o tamanho adequado. O fragmento de gel excisado era subdividido em 4 fatias longitudinais, a partir das quais foi realizada a amplificação do cDNA. Duas fatias de cada amostras eram prontamente usadas para a PCR e outras duas eram estocadas a -20°C. Cada fatia era colocada em um tubo de 0,2 ml, ao qual era adicionado 2 µl da solução de iniciadores (contendo SOLiD™ System Barcode) e 98 µl do mix para PCR utilizando-se reagentes do kit SOLiD Small RNA Expression. A amplificação foi realizada em um termociclador através de 15 ciclos de 95°C por 30 segundos, 62°C por 30 segundos e 72°C por 30 segundos, sucedidos por incubação a 72°C por 7 minutos. Finalmente, o material amplificado era purificado com o kit PureLink™ PCR Micro (Invitrogen) e sua

concentração e qualidade final eram analisada com os equipamentos NanoDrop e Bioanalyzer.

#### 4.10.2 Preparo da beads para seqüenciamento

##### 4.10.2.1 PCR em emulsão

As PCRs em emulsão (ePCR) foram realizadas segundo protocolo SOLiD™ 3 System Template Bead Preparation Guide, mais especificamente para ePCR em mini-escala (produção de 75-150 milhões de beads). A fase oleosa era preparada através da mistura de 1,8 ml de estabilizador de emulsão 1, 400 µl de estabilizador de emulsão 2 e 37,8 ml de óleo de emulsão para um volume final de 40 ml (veja Figura 4.2A). Na fase aquosa da emulsão eram combinadas quantidades iguais das amostras de cDNA amplificado contendo SOLiD™ System Barcodes, para uma concentração final de 0,5 e 1 pM de DNA molde, os reagentes para ePCR (14 mM dNTP mix, 25 mM Cloreto de Magnésio, 40 nM ePCR primer 1, 3 µM ePCR primer 2, 0,54U/µl AmpliTaq Gold DNA Polimerase e água para volume final de 2720 µl) e 800 milhões de beads revestidas com oligo P1 (P1 DNA Bead) para amplificação nos microrreatores (veja Figura 4.2A). As fases eram emulsificadas por 5 minutos usando o sistema ULTRA-TURRAX™ Tube Drive da IKA®, num volume final de 11,8 ml, de forma a gerar microgotículas de fase aquosa no interior da emulsão (microrreatores, veja Figura 4.2B) onde ocorre a amplificação clonal do DNA em torno da P1 DNA Bead durante a ePCR. A emulsão é otimizada para gerar microrreatores monoclonais (somente uma bead e uma molécula de DNA molde), os quais podem gerar beads úteis para o seqüenciamento. 100 µl da emulsão eram transferidos para cada um dos poços de uma placa de PCR, onde seguia-se a ePCR: incubação inicial à 95°C por 5 minutos, 40 ciclos de 93°C por 15 segundos, 62°C por 30 segundos e 72°C por 75 segundos e incubação final a 72°C por 7 minutos. Durante a ePCR, cerca de 30 mil cópias clonais do DNA molde são geradas em cada P1 DNA Bead, com o adaptador P1 ligado a bead (veja Figura 4.2C).



**FIGURA 4.2 – PCR EM EMULSÃO**

A: fase aquosa e oleosa antes da emulsificação; B: emulsão antes da amplificação; C: emulsão após amplificação (ePCR).

FONTE: modificado do manual SOLiD™ 3 System Template Bead Preparation Guide.

#### 4.10.2.2 Quebra da emulsão e lavagem das beads

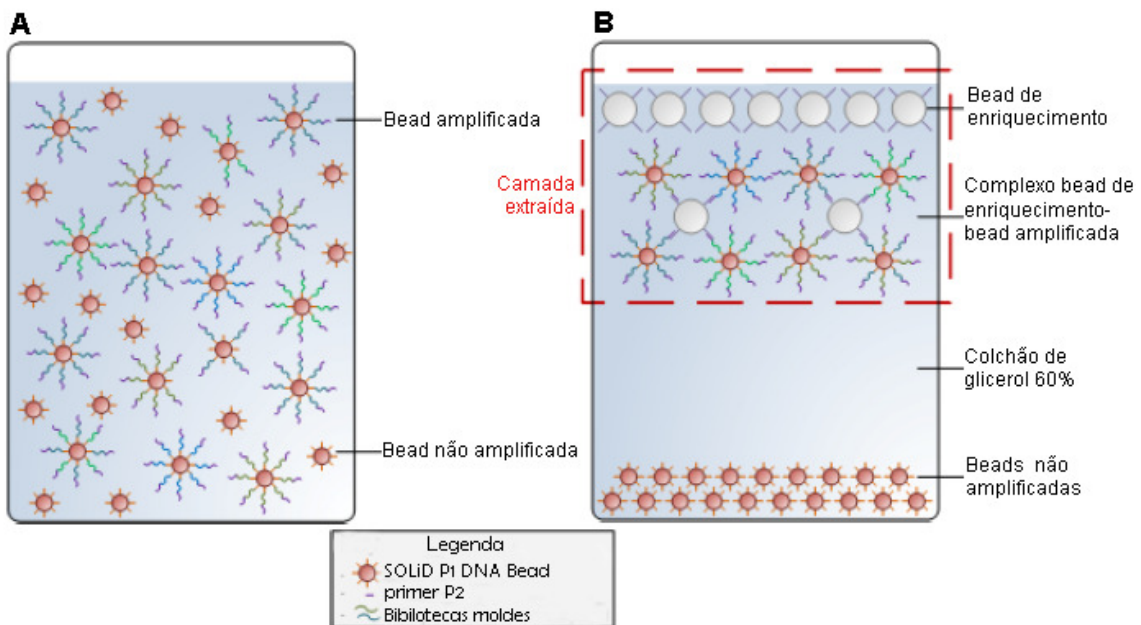
Após a ePCR, os microrreatores na emulsão são quebrados com 2-butanol e as beads são lavadas para eliminar o óleo e emulsificadores. O volume total da emulsão (~11ml) era passado para um sistema SOLiD™ Emulsion Collection Tray onde era adicionado 10 ml de 2-Butanol. Após homogeneização da solução, a mistura era transferida para um tubo de 50 ml, completando-se o volume para 30 ml com 2-Butanol e centrifugada a 2200 x g por 5 minutos. A fase superior era descartada e o *pellet* contendo as beads era recuperado com 1X Bead Wash Buffer, passado para tubo de 1,5 ml LoBind e lavado com 1X Bead Wash Buffer e 1X TEX Buffer num volume final de 200 µl (veja Figura 4.3A).

#### 4.10.2.3 Enriquecimento das beads contendo DNA

O procedimento de enriquecimento de beads contendo DNA molde para sequenciamento (beads molde) tem por objetivo isolar as beads contendo produtos de PCR completos após a ePCR. Beads de poliestireno com adaptadores P2 simples-fita (complementares a extremidade livre do DNA amplificado presente nas beads) são usados para capturas as beads moldes. Tanto beads policlonais como monoclonais são enriquecidas, produzindo cerca de 75-150 milhões de beads molde por ePCR em mini escala.

Após lavagens em tampão TEX 1X, as beads de enriquecimento eram combinadas com as beads de sequenciamento em um volume final de 150 µl e deixadas à 61°C por 15 minutos para permitir o anelamento das extremidades complementares dos adaptadores. A mistura contendo beads de enriquecimento, complexos bead de enriquecimento-bead molde e

beads não amplificadas era passada para o topo de um colchão de glicerol 60% em um tubo 1,5 ml LoBind. Após centrifugação à 21.000 g por 3 minutos, as beads de enriquecimento e os complexos bead de enriquecimento-bead molde concentram-se em uma camada superior (Figura 4.3B) e são recuperados com ajuda de uma pipeta e passados para um novo tubo 1,5 ml LoBind contendo 1 ml de tampão TEX. Após lavagens com tampão TEX, os complexos bead de enriquecimento-bead molde eram mantidos em 400 µl de tampão TEX. A seguir seguiu-se a desnaturação dos complexos através de 3 lavagens sucessivas com 400 µl de tampão de desnaturação, sendo que as beads moldes eram seletivamente recuperadas com auxílio de um rack magnético e mantidas em 400 µl de tampão TEX.



**FIGURA 4.3 – ENRIQUECIMENTO DE BEADS MOLDES**

A: beads amplificadas e não amplificadas após a quebra da emulsão e lavagem das beads; B: beads de enriquecimento, beads amplificadas e beads não amplificadas após a centrifugação em colchão de glicerol 60%.

FONTE: modificado do manual SOLiD™ 3 System Template Bead Preparation Guide.

#### 4.10.2.4 Modificação da extremidade 3' e quantificação das beads

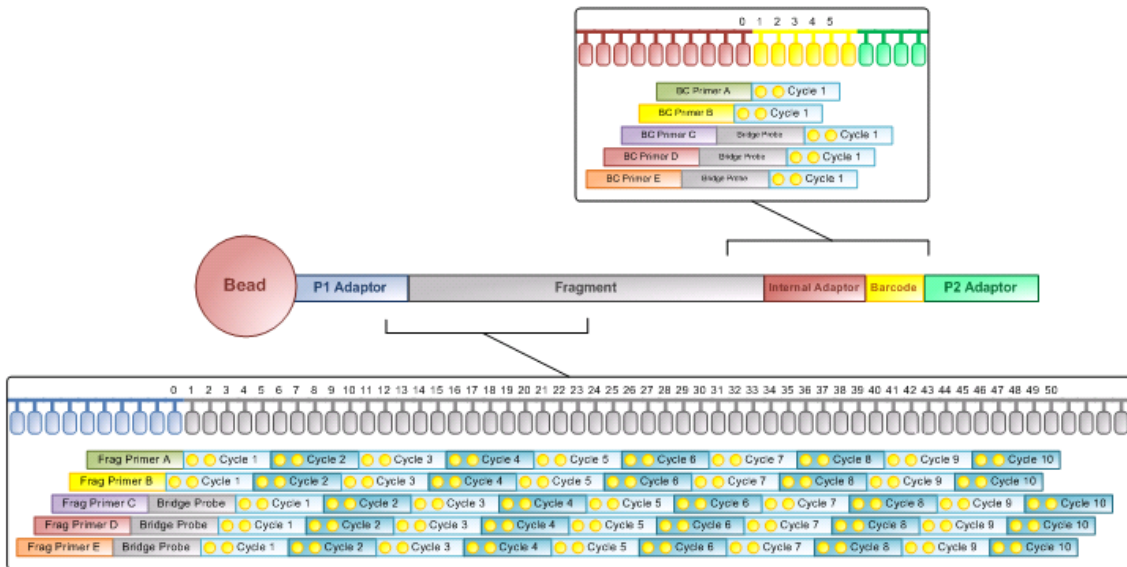
Para preparar as beads enriquecidas para deposição na lâmina, um nucleotídeo dUTP modificado é adicionado a extremidade do DNA molde usando uma reação com terminal transferase. Este nucleotídeo possui modificações que permitem a ligação na sílica da lâmina de seqüenciamento de maneira a formar arranjos estáveis de beads na célula de fluxo. As beads eram passadas para 178 µl de Tampão de Terminal Transferase 1X, onde era adicionado 20 µl de solução Bead Linker (dUTP modificado) e 2 µl da enzima Terminal Transferase

(20U/ $\mu$ l). Após incubação a 37°C por duas horas as beads eram lavadas e recuperadas em 200  $\mu$ l de tampão TEX.

A concentração final de beads molde obtidas era estimada usando contagem em câmara de Neubauer. A solução de beads era diluída cerca de 15X e, após deposição na câmara, eram contadas as beads na região quadrangular central da câmara, em um total de 5 subquadrantes. A concentração de beads na solução original era obtida usando a fórmula: concentração = (média de beads por subquadrante) x 250 x (fator de diluição).

#### 4.10.3 Deposição das beads e sequenciamento

O preparo das lâminas e sequenciamento foi realizado segundo o protocolo *SOLiD™ 3 System Instrument Operation Guide*. Antes do sequenciamento das amostras, a qualidade das beads era analisada através do procedimento *workflow analysis* (WFA) para um total de 15 milhões de beads por amostra múltipla. A partir dos resultados do WFA, seguiu-se o sequenciamento propriamente dito. Cerca de 72 milhões de beads (ou 36 milhões de beads para lâmina de 8 quadrantes) eram lavadas com tampão de deposição e depositadas em cada um dos poços de uma câmara de deposição de 4 quadrantes. As beads eram deixadas aderir na lâmina por 1,5 horas à 37°C e finalmente a lâmina era instalada no aparelho SOLiD™ 3 System. Os sequenciamentos foram realizados segundo protocolo SOLiD Multiplex; primeiramente é realizado o sequenciamento dos Barcodes através de um ciclo de ligação para cada um dos cinco primers de sequenciamento, resultado em um tag de cinco bases para cada barcode. A seguir é realizado o sequenciamento dos fragmentos através de dez ciclos de ligação para cada um dos cinco primers de sequenciamento, resultando em *reads* de 50 bases para cada bead molde (para detalhes da química de sequenciamento veja ANEXO I).



**FIGURA 4.4- SEQUENCIAMENTO DO BC E FRAGMENTO DE 50 BASES PELO SOLiD**

Na parte superior da figura é detalhado o sequenciamento do *barcode* (BC) utilizando primers complementares ao adaptador interno (em vermelho). A parte inferior explica o sequenciamento do fragmento proveniente do transcriptoma através de 10 ciclos de ligação para cada um dos 5 primers de sequenciamento. Note que cada posição é interrogada duas vezes (bolas amarelas).

O lançamento das corridas, bem como o processamento dos sequenciamentos para geração, obtenção e controle dos dados foi realizado através do programa SOLiD™ Instrument Control Software (ICS).

#### 4.10.4 Análise dos dados

As leituras geradas pelos sequenciamento na plataforma SOLiD foram alinhados ao genoma de *T. cruzi* usando a ferramenta *SOLiD™ WT Analysis Pipeline*, permitindo até 2 erros de pareamento de bases. As sequências foram alinhadas nos dois haplótipos do parasita e foram consideradas leituras de alinhamento múltiplos em até 10 posições distintas no genoma. O número de leituras alinhadas em cada gene anotado de *T. cruzi* foi somado, e este valor foi utilizado para as análises de expressão diferencial. Os dados brutos foram pré-processados para retirada de genes referentes a RNA não codificador (ncRNA).

Os dados de contagens de *reads* foram normalizados e analisados para expressão diferencial no pacote edgeR (ROBINSON *et al.*, 2010) do programa R, considerando diferentes níveis de confiabilidade estatística (*false Discovery rate*, FDR). Além disso, foi usado o valor de alteração absoluta (*fold-change*) de 1,5 vezes como critério biológico para seleção de genes diferencialmente expressos.



O nível de expressão dos genes também foi analisado usando o método RPKM (MORTAZAVI *et al.*, 2008), o qual leva em conta o tamanho total de cada biblioteca (número total de leituras alinhados no genoma) e o tamanho do gene, o que permite uma estimativa do nível absoluto de expressão de cada gene. Para tal, o número de leituras,  $R$  (*reads*), alinhados a cada gene foi normalizado pela seguinte fórmula:  $R = 10^9 \times C / N \times L$ ; onde  $C$  é o número de *reads* alinhados em um gene,  $N$  é o total de leituras alinhadas no genoma e  $L$  é o tamanho do gene em pares de bases.

Os genes identificados como diferencialmente expressos foram analisados por agrupamento hierárquico, utilizando distância Euclidiana no programa Cluster (EISEN *et al.*, 1998), sendo que as imagens geradas foram visualizadas no programa JavaTreeView.

Para análise do enriquecimento de termos de ontologia gênica (*gene ontology*, GO), foi retirada a redundância dos dados de expressão gênica (devido a presença dos dois haplótipos de *T. cruzi*), considerando apenas os supergenes do parasita (agrupamento dos alelos das duas cepas). Os supergenes diferencialmente expressos foram analisados na ferramenta Blast2GO (CONESA *et al.*, 2005). Para determinação dos termos de GO enriquecidos foi usado o teste exato de Fisher, considerando como limiar estatístico FDR de 5%, sendo considerado apenas os termos sobre-representados.

Para as análises de regulação da via de síntese de esteróis, os genes relacionados a tal via em *T. cruzi* foram selecionados através de análise de similaridade de sequência usando a ferramenta BLASTp (*Basic Local Alignment Search Tool*) (LADUNGA, 2009) e, como referência, as proteínas da via de síntese de ergosterol de *Saccharomyces cerevisiae*. Os genes selecionados foram utilizados na construção desta via de *T. cruzi* no programa GenMapp (DAHLQUIST *et al.*, 2002); finalmente, os dados de RNA-Seq foram associados ao mapa de síntese de esteróis construído, permitindo a visualização dos genes regulados em resposta aos inibidores da via.

As análises de vias enriquecidas em resposta as drogas foram realizadas por busca contra os bancos de dados do KEGG (*Kyoto Encyclopedia of Genes and Genomes*) (OGATA *et al.*, 1999), utilizando as listas de genes significativos como os dados de entrada para as buscas. Para visualização dos genes diferencialmente expressos dentro das vias foi utilizada a ferramenta *online* KEGG Mapper, a qual permite a visualização de padrões de cores associadas aos genes de expressão aumentada ou diminuída.

Para as análises de reguloma, todos os experimentos de seqüenciamento foram normalizados em conjunto no programa edgeR (ROBINSON *et al.*, 2010). Os dados

normalizados foram então processados de forma a se obter o valor Z (leituras alinhadas - média de leituras alinhadas/ desvio padrão) para cada gene em cada amostra. Esses dados foram analisados por clusterização hierárquica no programa Cluster (EISEN *et al.*, 1998). Nesse momento, foram usados filtros de variação (número de desvios padrão acima da média), de forma que somente os genes com maior variação fossem selecionados. Estes genes foram então clusterizados, seguido da inspeção visual das imagens geradas em busca de grupos de genes co-regulados (*regulons*).

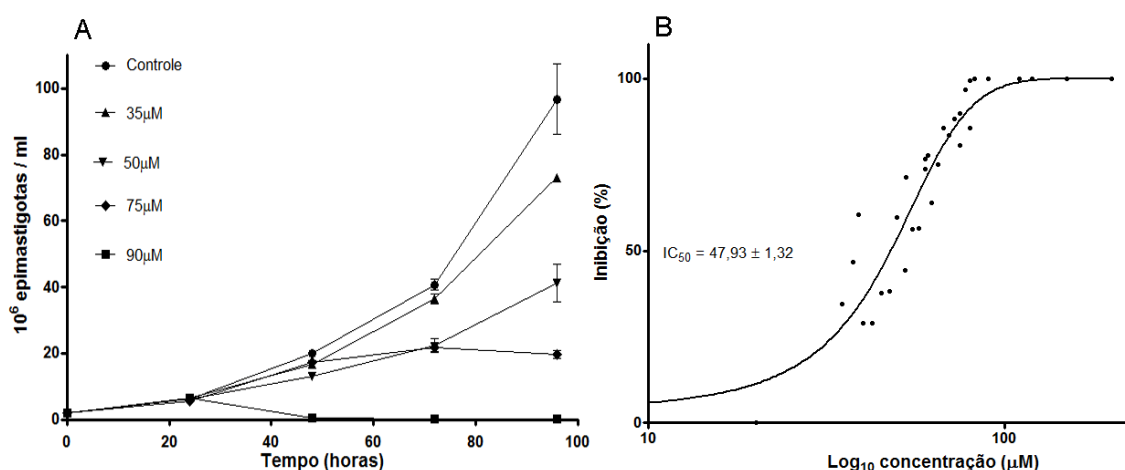
## 5 RESULTADOS

### 5.1 EFEITO TRIPANOCIDA DOS FÁRMACOS

A primeira etapa do trabalho consistiu na determinação das concentrações inibitórias dos IBEs na cepa de *Trypanosoma cruzi* utilizada em nosso laboratório, Dm28c. Embora já existam relatos na literatura acerca das doses inibitórias destas drogas (detalhes no tópico 1.5.2), não é adequado utilizar tais concentrações diretamente visto que diferentes cepas de *Trypanosoma cruzi* foram usadas.

#### 5.1.1 Lovastatina

Seguindo os métodos descritos (tópico 4.3), demonstrou-se que as doses de lovastatina capazes de inibir o crescimento de epimastigotas de *Trypanosoma cruzi* (Dm28c) em cultura estavam dentro da faixa de concentração previamente demonstrada (FLORIN-CHRISTENSEN *et al.*, 1990; URBINA *et al.*, 1993). Observou-se uma inibição do crescimento dos parasitas de forma dose dependente entre as concentrações de cerca de 20  $\mu\text{M}$  a 100  $\mu\text{M}$  (Figura 5.1). A partir das porcentagens de inibição do crescimento obtidos para



**FIGURA 5.1 – EFEITO ANTIPROLIFERATIVO DE LOVASTATINA**

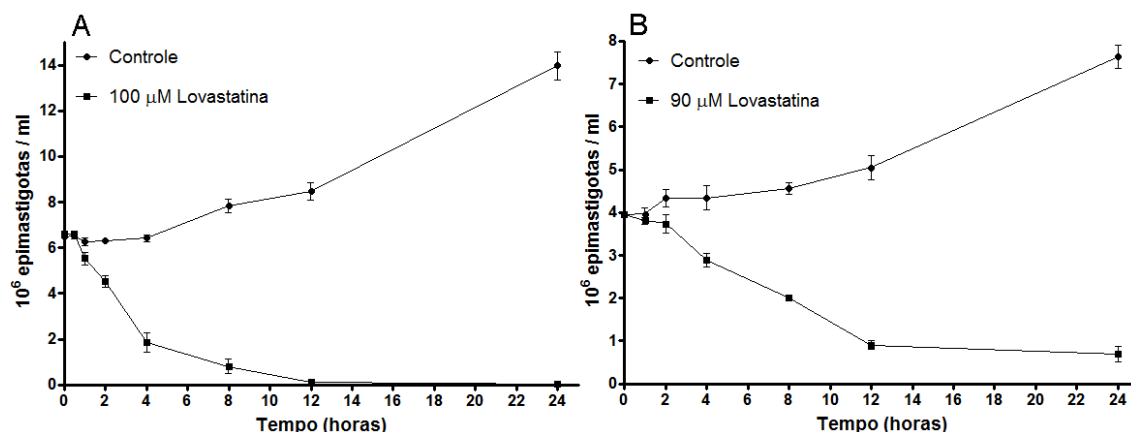
A- após 96 horas de cultura os parasitas entram em fase estacionária de crescimento. Parasitas na presença somente de DMSO não demonstraram inibição do crescimento (dados não mostrados). B- curva dose/resposta para Lovastatina.

diversas concentrações de lovastatina, após 72 horas de exposição à mesma, foi construída uma curva dose-resposta da qual foi extraído, por regressão não-linear, o valor da  $IC_{50}$  para lovastatina de  $47,93 \pm 1,36 \mu M$  (Figura 5.1B). Tais resultados corroboram dados presentes na literatura. O primeiro trabalho utilizando lovastatina como inibidor da biossíntese de esteróis foi realizado por Florin-Christensen e colaboradores (1990), demonstrando-se que tal droga era capaz de inibir o crescimento de epimastigotas em cultura em concentrações de 10 a 30  $\mu g/ml$  ( $\sim 25-75 \mu M$ ), concentrações similares às aqui descritas (Figura 5.1). 50  $\mu g/ml$  ( $\sim 124 \mu M$ ) de lovastatina foi capaz de matar a maioria dos parasitas, o que também foi observado no presente trabalho.

Urbina e colaboradores (1993) demonstraram que lovastatina em concentrações de 50  $\mu M$  e 75  $\mu M$  era capaz de induzir inibição total do crescimento e lise celular de epimastigotas em cultura após 144 e 96 horas (veja a lise celular no tópico 5.2.1). Como pode ser observado na Figura 5.1, 75  $\mu M$  de Lovastatina gerou o mesmo efeito de inibição total do crescimento.

Portanto, os dados de inibição do crescimento de epimastigotas em cultura corroboram dados já descritos na literatura, o que valida os métodos usados no presente trabalho.

Além da determinação da faixa de concentração de droga capaz de inibir o crescimento de epimastigotas em 72 horas de exposição, e respectiva  $IC_{50}$ , foram determinadas as concentrações mínimas das drogas capazes de matar todas as células em cultura dentro de 24 horas de exposição (detalhes no tópico 4.3). 100  $\mu M$  de lovastatina induz a morte gradual dos parasitas a partir de cerca de 1 hora de exposição, culminando com a ausência quase total de parasitas detectados após 12 e 24 horas de exposição (Figura 5.2A). A 90  $\mu M$  (Figura 5.2B), a morte celular não é tão pronunciada em uma ou duas horas de exposição, mas após duas horas também ocorre morte gradual dos parasitas até 12 horas de exposição. Noventa  $\mu M$  de lovastatina permite a sobrevivência de cerca de 10 vezes mais células após 12 e 24 horas de exposição em relação à 100  $\mu M$  (compare Figura 5.2A e B).

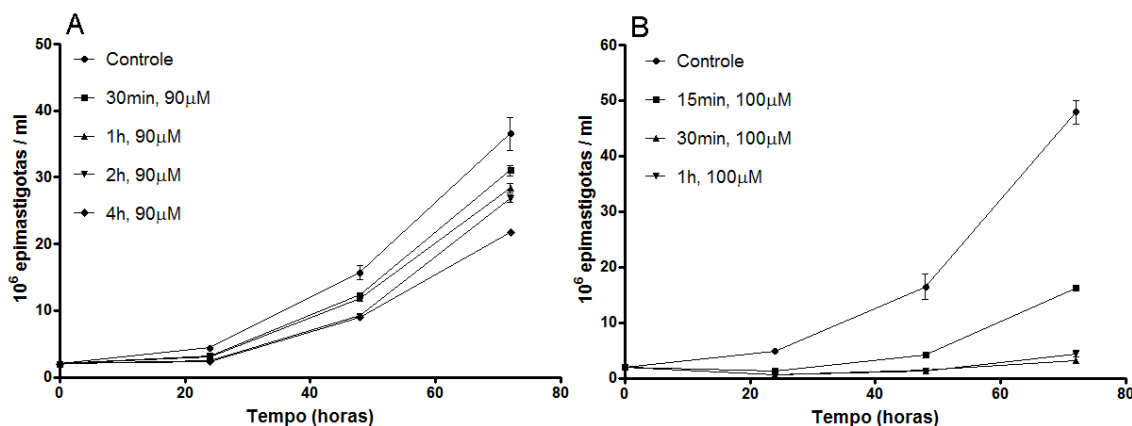


**FIGURA 5.2 – MORTE CELULAR COM ALTAS DOSES DE LOVASTATINA**

A - a densidade celular em 24 horas de exposição a 100µM de Lovastatina era de  $3,3 \times 10^4$  cel/ml; B - a densidade celular em 24 horas de exposição a 90µM de Lovastatina era de  $6,8 \times 10^5$  cel/ml.

De forma a entender mais profundamente a cinética de ativação da morte celular em epimastigotas de *T. cruzi* expostos a altas concentrações de droga e, conseqüentemente, determinar mais criteriosamente os tempos de exposição a serem usados para extração de RNA em experimentos de RNA-Seq, foram realizados experimentos de recuperação. Após tempos determinados de exposição a doses letais das drogas os parasitas eram recuperados por centrifugação e lavados com tampão fosfato estéril três vezes de forma a retirar a droga previamente presente na cultura. As células eram então passadas para um novo meio de cultura sem droga na densidade celular de  $2 \times 10^6$  cel/ml e a ativação da morte celular, devido ao prévio contato com a droga, era analisado por acompanhamento do crescimento da cultura até 72 horas pós-retirada da droga.

Como pode ser observado na Figura 5.3A, à 90 µM de lovastatina as células são capazes de retomar o crescimento em cultura mesmo após 4 horas de exposição a droga, embora haja uma diminuição do número de células proporcionalmente ao tempo de exposição à droga, provavelmente relacionado ao número de células nas quais é ativada a morte celular. À 100 µM o padrão já se torna distinto, sendo que em apenas 15 minutos de exposição é ativada morte celular em cerca de 50% dos parasitas, e em 30 minutos ou 1 hora de exposição a cultura não é capaz de retomar o crescimento (Figura 5.3B). Portanto, a dose de 100 µM é uma concentração mais apropriada para a etapa de análise genômica funcional da morte celular de *T. cruzi* em tempos curtos de exposição a droga. O objetivo dessa etapa do projeto é avaliar quais são os módulos transcricionais comuns e, mais importantemente, quais são os possíveis módulos, provavelmente mais raros, distintos, que se alteram no transcriptoma em resposta a uma agressão externa forte, que invariavelmente leva à morte celular.

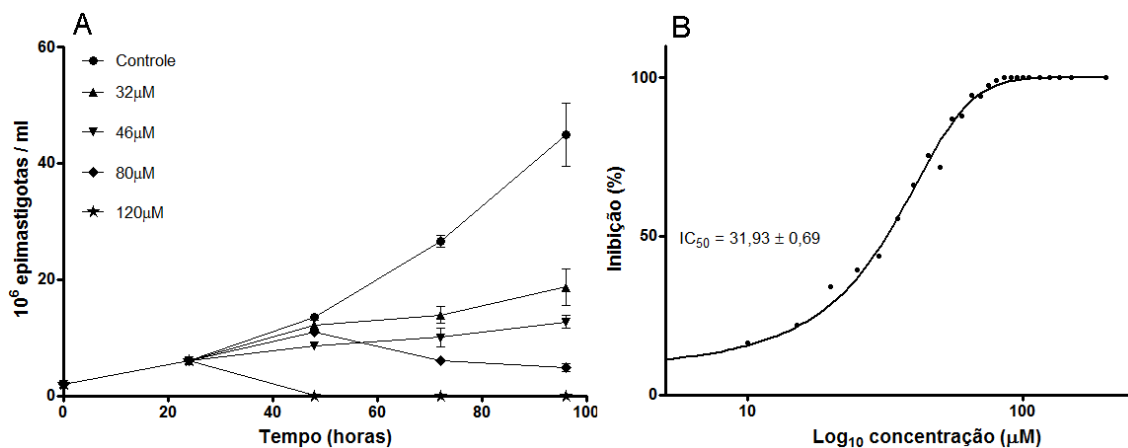


**FIGURA 5.3 – RECUPERAÇÃO: LOVASTATINA**

Após o tempo indicado na legenda a droga foi retirada do meio de cultura por seguidas lavagens com tampão fosfato (detalhes em 4.3.1) e as células foram passadas para um novo meio de cultura na densidade celular de  $2 \times 10^6$  cel/ml. A: exposição prévia a  $90 \mu\text{M}$  de lovastatina; B:  $100 \mu\text{M}$  de lovastatina.

### 5.1.2 Cetoconazol

De forma similar aos experimentos realizados com lovastatina, inicialmente foram determinadas as concentrações inibitórias de cetoconazol na cepa Dm28c. Como observado na Figura 5.4, cetoconazol é capaz de inibir o crescimento de epimastigotas em cultura de forma dose-dependente entre as concentrações de cerca de 1 a  $90 \mu\text{M}$ . A partir da curva dose-resposta foi obtida a  $\text{IC}_{50}$  de  $31,93 \pm 0,69 \mu\text{M}$ . Diferentemente dos dados obtidos para lovastatina, a  $\text{IC}_{50}$  de cetoconazol difere de dados presentes na literatura para outras cepas de *T. cruzi*. Segundo Beach e colaboradores (1986) e Urbina e colaboradores (1993), a  $\text{IC}_{50}$  de cetoconazol nas cepas Peru e Y, respectivamente, é de 0,2 e  $0,1 \mu\text{M}$ , enquanto segundo Hankins e colaboradores (2005), a  $\text{IC}_{50}$  de Cetoconazol na cepa Tulahuen é  $25 \mu\text{M}$ . Assim sendo, o valor de  $\text{IC}_{50}$  calculado no presente trabalho está duas ordens de grandeza acima dos valores descritos para as cepas Peru e Y, mas muito semelhante à da cepa Tulahuen. Esse resultado poderia ser explicado de duas formas: (i) suscetibilidade diferente entre as cepas de *T. cruzi* a cetoconazol e/ou (ii) uso de diferentes métodos para cálculo de  $\text{IC}_{50}$ . Para responder tal questão, calculou-se a  $\text{IC}_{50}$  de cetoconazol na cepa Y seguindo o mesmo procedimento realizado para Dm28c.

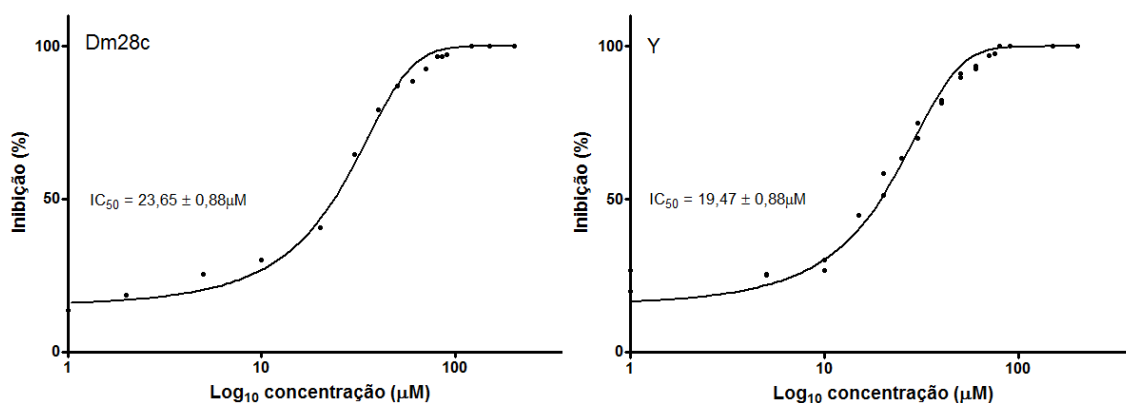


**FIGURA 5.4 – EFEITO ANTIPROLIFERATIVO DE CETOCONAZOL.**

A: curva de crescimento; B: curva dose/resposta para cálculo de IC<sub>50</sub>.

NOTA: após 96h de cultura os parasitas entram em fase estacionária de crescimento. Parasitas na presença somente de DMSO não demonstraram inibição do crescimento (dados não mostrados).

Como culturas da cepa Y formam grumos, outros grupos crescem estas células sob forte agitação (120 rpm) (URBINA *et al.*, 1993; URBINA *et al.*, 2003; URBINA *et al.*, 2004). Como em nosso laboratório não havia a possibilidade de crescer as células com agitação, optou-se por homogeneizar fortemente as culturas (vortex) antes das contagens em câmara de Neubauer de forma a desfazer os grumos de parasitas e evitar a subestimação da densidade celular. Como esse método altera o valor da IC<sub>50</sub> obtida, também foi recalculado a IC<sub>50</sub> para a cepa Dm28c. Os resultados (Figura 5.5) mostram que as cepas usadas possuem IC<sub>50</sub> para cetoconazol muito similar (Dm28c: 23,65±0,88 μM, Y: 19,47±0,88 μM), e muito próximo ao valor descrito por Hankins e colaboradores (2005) para a cepa Tulahuen (25 μM), o que diminui a importância de uma possível resistência intrínseca de Dm28c a tal azol e aponta para o impacto de diferentes métodos no cálculo da IC<sub>50</sub>.

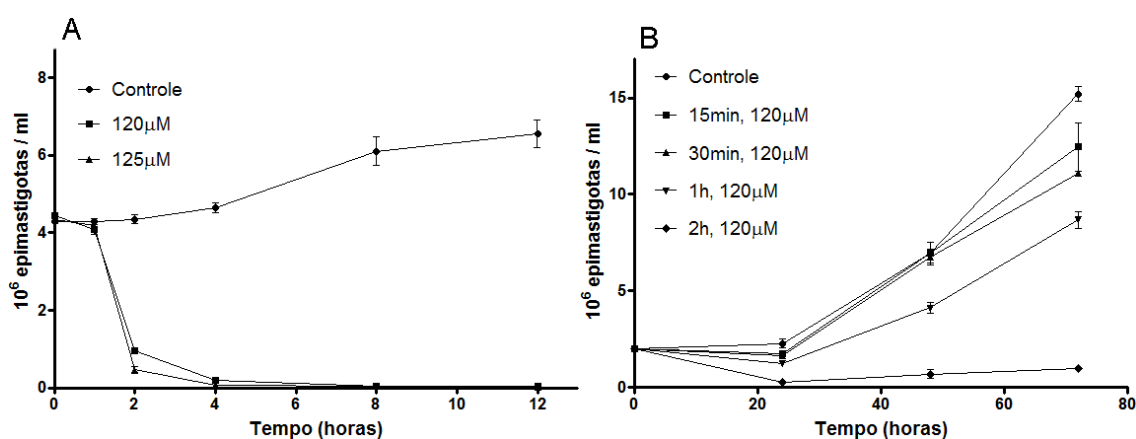


**FIGURA 5.5 – DOSE/RESPOSTA CETOCONAZOL: DM28C VERSUS Y**

Curva dose/resposta das cepas Dm28c (à esquerda) e Y (à direita) a Cetoconazol, usando homogeneização antes das contagens da densidade celular.

De forma similar aos experimentos feitos com lovastatina, buscou-se concentrações de cetoconazol capazes de matar a grande maioria dos parasitas em curtos períodos de tempo. Após análise de algumas altas concentrações da droga (100-130  $\mu\text{M}$ ), observou-se que a 120  $\mu\text{M}$  e 125  $\mu\text{M}$ , a maioria das células começam a morrer a partir de 2 horas de exposição à droga. Diferentemente do padrão de morte celular com 100  $\mu\text{M}$  de Lovastatina (queda gradual da densidade celular), a partir de 2 horas de exposição observa-se uma queda brusca na densidade celular, culminando com ausência total de parasitas vivos após 12 horas de exposição (veja Figura 5.6A).

A seguir investigou-se a capacidade de recuperação de *T. cruzi* após curtos períodos de exposição a doses altas de cetoconazol. A 120  $\mu\text{M}$ , após 15 minutos de exposição os parasitas parecem recuperar o crescimento quase normal em relação ao controle (Figura 5.6B). Entretanto, com uma hora de exposição, cerca da metade da população de parasitas morre e após duas horas de exposição à 120  $\mu\text{M}$ , praticamente não se observa proliferação posterior das células. Portanto, determinou-se a DL100 de cetoconazol como 120  $\mu\text{M}$ .



**FIGURA 5.6 – MORTE CELULAR E RECUPERAÇÃO A CETOCONAZOL.**  
A – Cinética de morte celular através da contagem da densidade celular; B – Recuperação das células após curtos períodos de exposição.

## 5.2 ALTERAÇÕES MORFO-FISIOLÓGICAS EM RESPOSTA ÀS DROGAS

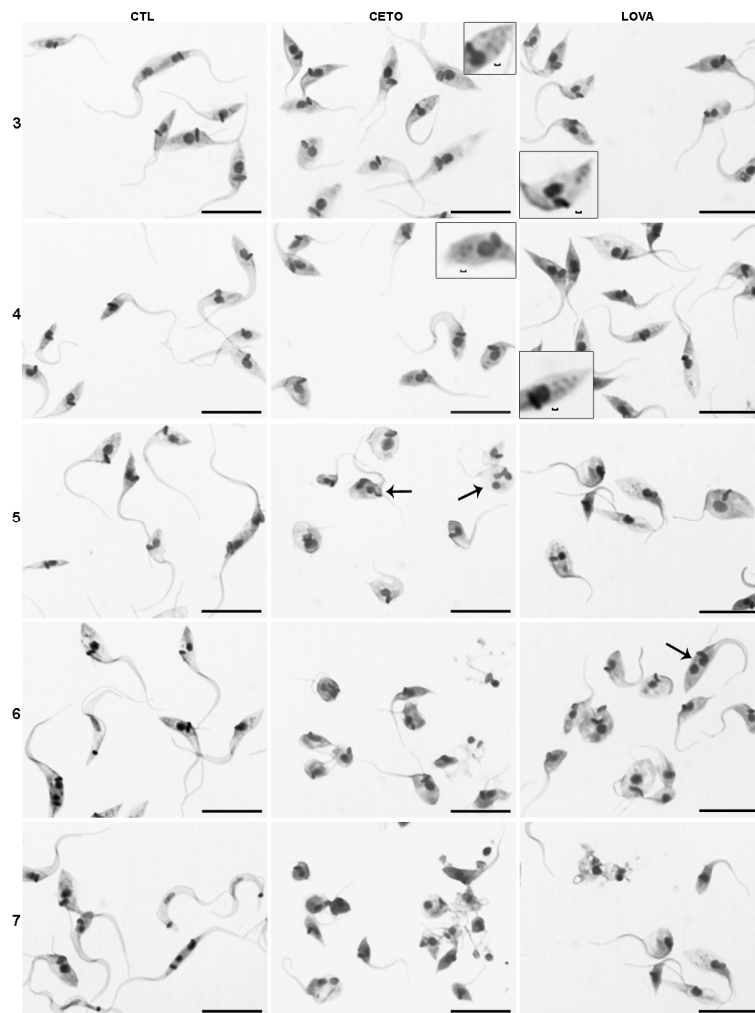
As doses  $\text{IC}_{50}$  e DL100 das drogas foram utilizadas em experimentos de microscopia e citometria de fluxo de forma a acompanhar as alterações morfológicas e fisiológicas de epimastigotas em resposta às drogas.



### 5.2.1 IC<sub>50</sub>

Culturas de epimastigotas tratadas com lovastatina e cetoconazol foram analisadas por microscopia óptica segundo protocolo descrito no tópico 4.4.1. Embora este método propicie pouca resolução de estruturas celulares, permite a análise de um grande número de células e, conseqüentemente, a identificação de padrões robustos na resposta dos parasitas às drogas.

Na dose IC<sub>50</sub>, os parasitas apresentam alterações morfológicas tardias, devido à presença de um *pool* pré-formado de esteróis nas células (URBINA *et al.*, 1993). Dessa forma, para ambas as drogas analisadas, algumas das alterações morfológicas clássicas na resposta a IBEs (revisito por DE SOUZA & RODRIGUES, 2008) devem coincidir com a depleção dos esteróis endógenos do parasita (URBINA *et al.*, 1993).



**FIGURA 5.7 – MICROSCOPIA ÓPTICA IC<sub>50</sub>**

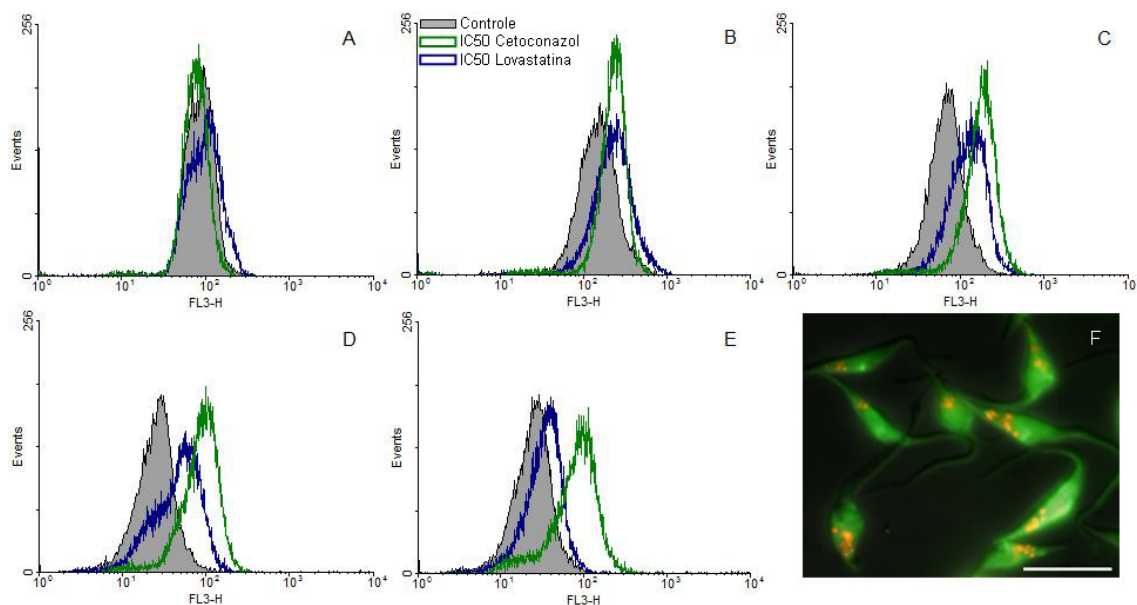
Os números do lado esquerdo da figura referem-se aos dias de exposição das células as drogas. CTL: controle; CETO: Cetoconazol; LOVA: Lovastatina. As caixas em destaque mostram células com grande quantidade de vesículas na porção posterior da célula. As barras de aumento possuem 10µm, a não ser nas células em destaque (0,5µm). As setas indicam células com dois núcleos e/ou cinetoplastos.

Como pode ser observado na Figura 5.7, a exposição dos epimastigotas de *T. cruzi* a  $IC_{50}$  de ambas as drogas induz um inchaço celular gradual a partir de quatro dias de exposição, culminando com a lise celular com seis ou sete dias, o que corrobora dados da literatura com diversos IBEs (LAZARDI *et al.*, 1991; URBINA *et al.*, 1993; VIVAS, URBINA & DE SOUZA, 1997; BRAGA *et al.*, 2005). Tal lise celular foi mais evidente nos parasitas tratados com cetoconazol (Figura 5.7). Após seis dias de exposição, os parasitas também apresentaram retração celular (diminuição do tamanho da célula), o que novamente foi mais evidente no tratamento com cetoconazol.

Outra alteração observada foi a presença de dois núcleos e/ou cinetoplastos (setas na Figura 5.7), indicando descontrole do ciclo celular e parada na divisão celular, como já observado para inibidores das enzimas 14 $\alpha$ -desmetilase (LAZARDI *et al.*, 1991; SANTA-RITA *et al.*, 2005), esqualeno sintase (BRAGA, URBINA & DE SOUZA, 2004) e 24-SMT (BRAGA *et al.*, 2005). Além disso, as drogas parecem inibir a metaciclogênese, visto que culturas controle apresentam quantidade crescente de formas metacíclicas a partir de cinco ou seis dias de cultivo (Figura 5.7), o que não foi observado nas culturas expostas as drogas, mesmo embora esses dados não tenham sido avaliados mais a fundo por contagens diferenciais.

Outra alteração muito clara é o aumento de vesículas fortemente coradas na região posterior das células (destacado na Figura 5.7). A forte coloração destas vesículas pelo procedimento panótico sugere que estas possuam um baixo pH, ou características ácidas. Para corroborar essa hipótese, realizaram-se experimentos de citometria de fluxo com o corante acidotrófico laranja de acridina e de microscopia confocal com o corante LysoTracker® Red.

Quando ligado ao DNA ou livre no citoplasma, laranja de acridina fluoresce na faixa de luz verde. Entretanto, ao acumular-se em vesículas ácidas, altera suas características espectrais e passa a fluorescer no comprimento de onda de luz vermelha (CLERC & BARENHOLZ, 1998; veja Figura 5.8F). Experimentos de citometria de fluxo com parasitas tratados com a  $IC_{50}$  de ambas as drogas evidenciaram um aumento na intensidade de fluorescência na faixa do vermelho (detector FL3) (Figura 5.8) a partir de 48 horas de exposição, o que é compatível com a hipótese de um aumento de vesículas ácidas nos parasitas em resposta a  $IC_{50}$  dos IBEs aqui utilizados.



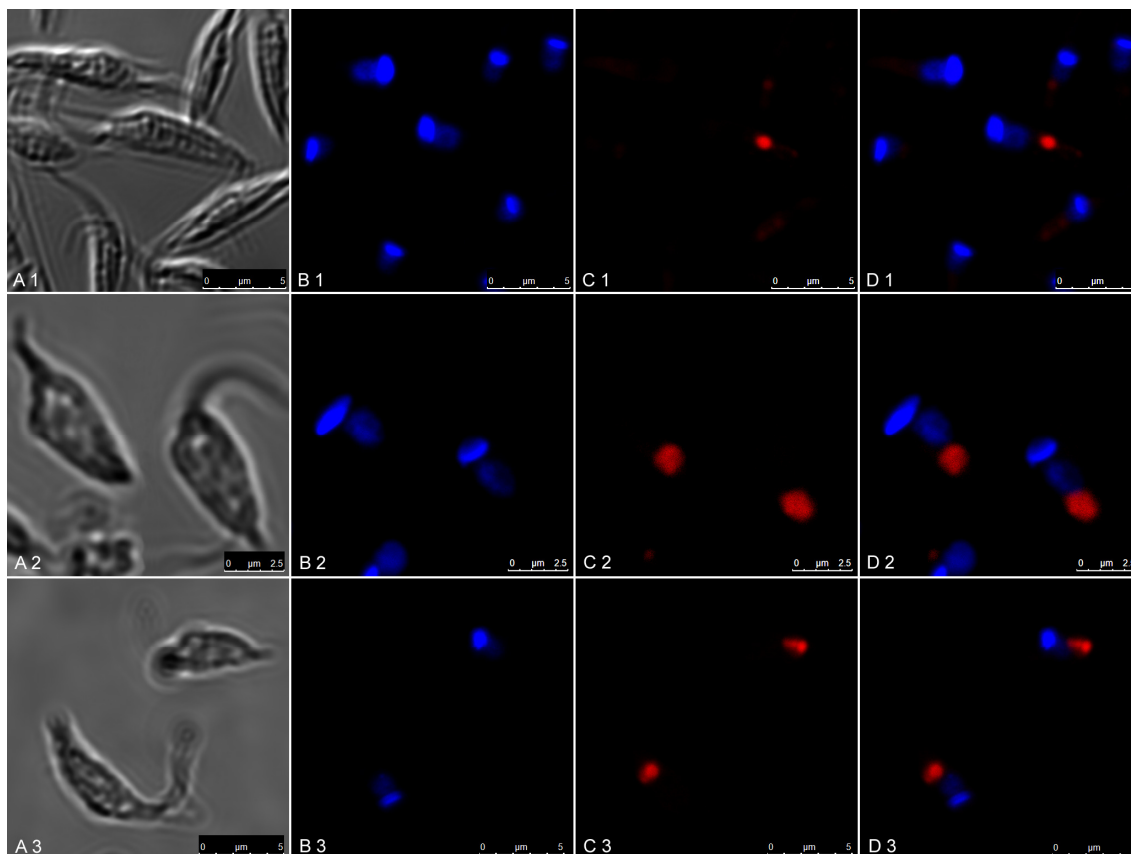
**FIGURA 5.8 - IC<sub>50</sub>: CITOMETRIA DE FLUXO COM LARANJA DE ACRIDINA**

Os parasitas foram tratados com Cetoconazol ou Lovastatina na dose IC<sub>50</sub> e analisados por citometria de fluxo com o corante Laranja de Acridina. A: 24 horas de exposição a droga; B: 48 horas; C: 72 horas; D: 96 horas; E: 120 horas; F: microscopia de fluorescência de uma cultura controle (não tratada com droga) para ilustração das vesículas ácidas (Reservossomos) corados em vermelho; a barra de aumento possui 10µm.

Nota: devido alta sensibilidade do corante Laranja de Acridina e a utilização de células vivas, não foi possível fotografar os parasitas tratados com os IBEs por microscopia de fluorescência com tal corante. Por este motivo foi usado o corante LysoTracker® Red em células fixadas.

A posição e morfologia das vesículas observadas através de microscopia confocal é apresentada na Figura 5.9. Nesse experimento, foram utilizadas células tratadas por 72 horas, ponto no qual há aumento significativo no conteúdo de vesículas ácidas pelos dados de citometria (Figura 5.8C), mas ainda não ocorreram alterações drásticas na morfologia celular (Figura 5.7). Pode ser observado um claro aumento na intensidade de fluorescência na região posterior das células tratadas com a IC<sub>50</sub> das drogas (Figura 5.9).

Devido ao tamanho aproximado de 0,5 µm (Figura 5.7) e localização na região posterior das células (Figura 5.7 e Figura 5.9), tais vesículas ácidas podem corresponder a reservossomos, o que chama atenção para a importância dessa organela na síntese e estocagem de esteróis em *T. cruzi* (revisto por De SOUZA *et al.*, 2009). De fato, um estudo recente com a metodologia de microscopia eletrônica de transmissão demonstra um aumento no tamanho e número de reservossomos em *T. cruzi* a partir de 48 horas de exposição a cetoconazol (SANTA-RITA *et al.*, 2005).



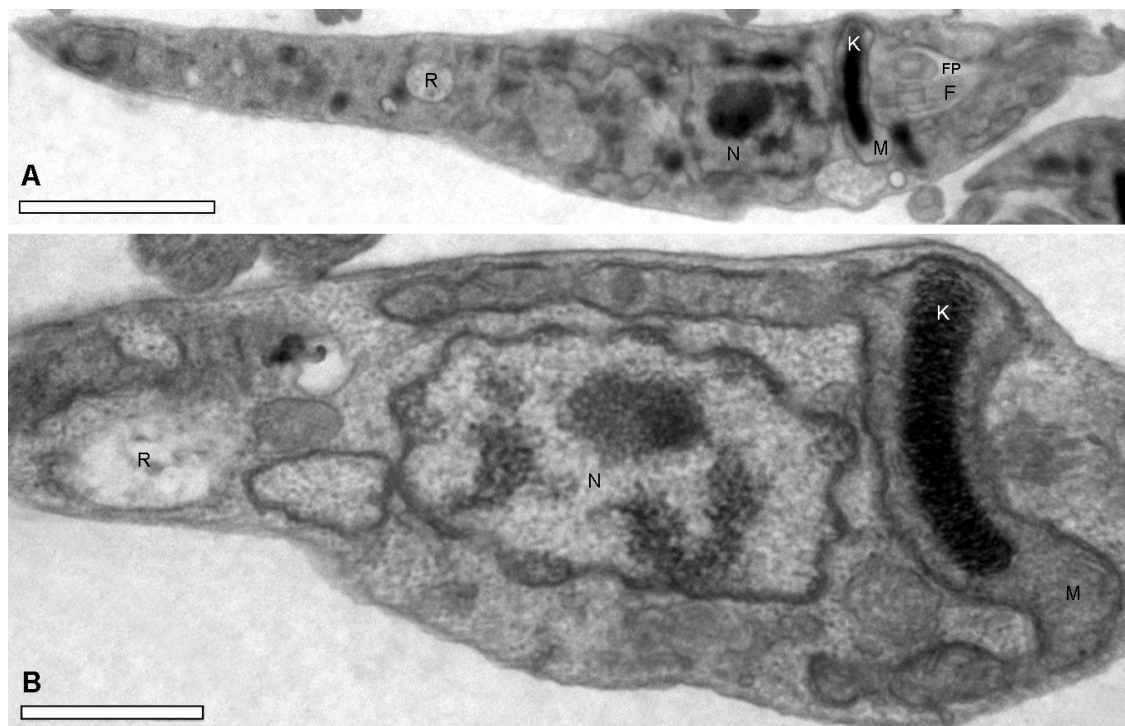
**FIGURA 5.9 – MICROSCOPIA CONFOCAL: LYSO TRACKER RED**

Os parasitas foram tratados por 72 horas com 32 $\mu$ M de Cetoconazol (linha 2), 50 $\mu$ M de Lovastatina (linha 3) ou não tratados (linha 1). Coluna A: DIC; coluna B: DNA, corante Hoechst 33342; coluna C: vesículas ácidas, corante LysoTracker® Red; coluna D: sobreposição das colunas B e C.

De forma a confirmar as alterações observadas por microscopia confocal e analisar mais profundamente as alterações morfológicas de *T. cruzi* em resposta aos inibidores da síntese de esteróis, foram realizadas análises de microscopia eletrônica de transmissão (M.E.T.). Para a dose IC<sub>50</sub>, foram utilizados os tempos de exposição de 72 horas e 120 horas.

A Figura 5.10 mostra a morfologia normal de células de *T. cruzi*; pode-se observar o formato alongado do corpo celular, com o cinetoplasto posicionado anteriormente ao núcleo e associado ao corpo basal. Na região posterior das células é possível observar vesículas de cerca de 0,5  $\mu$ m de diâmetro, os reservossomos.

Parasitas tratados com cetoconazol (Figura 5.11) ou lovastatina (Figura 5.12) apresentaram drásticas alterações ultraestruturais. Algumas das alterações aqui observadas já estão bem caracterizadas na resposta de *T. cruzi* a inibidores da síntese de esteróis (revisto por DE SOUZA & RODRIGUES, 2009), o que dá suporte ao presente trabalho como uma análise das respostas frente à inibição desta via.



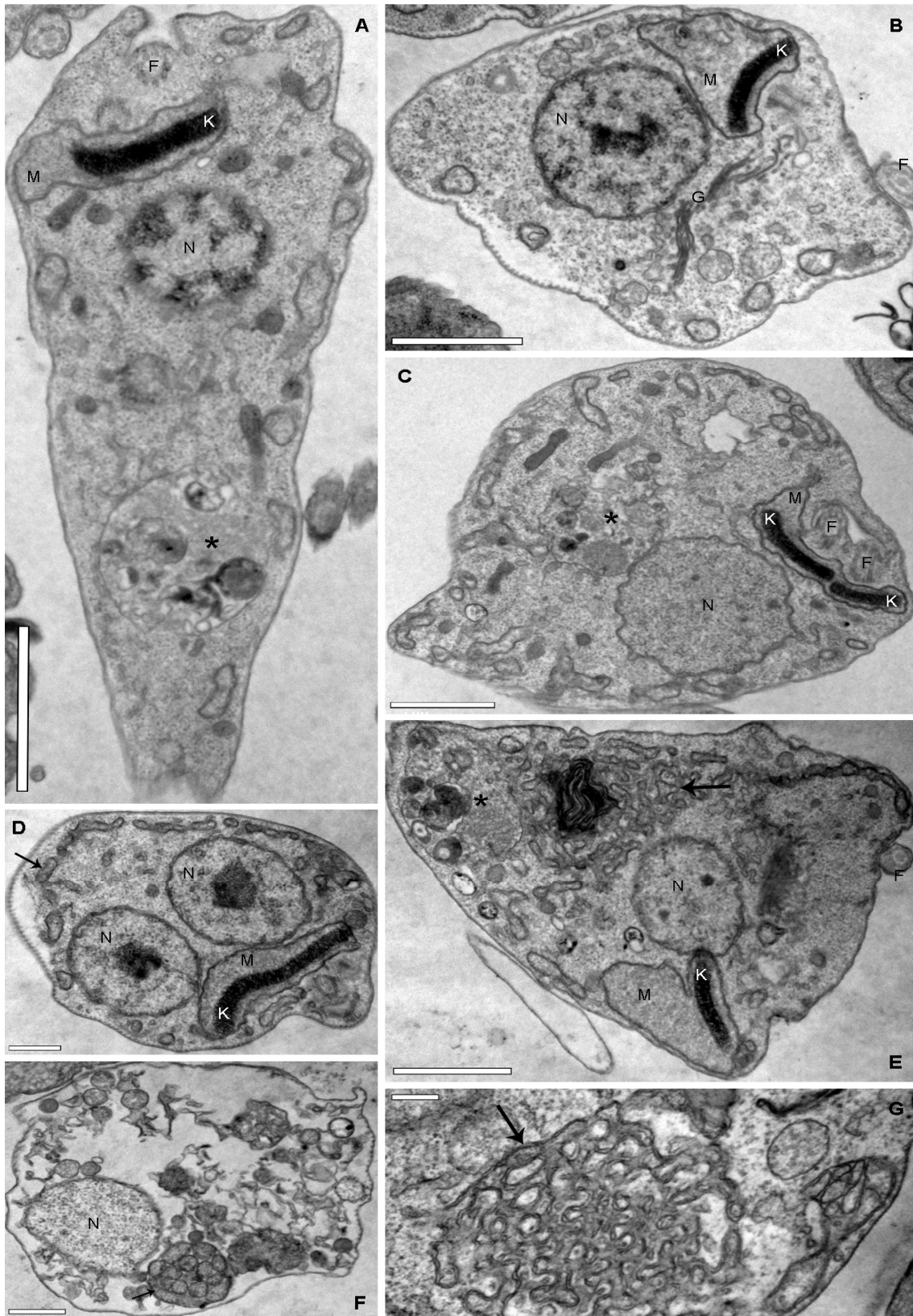
**FIGURA 5.10 – M.E.T. DE *T. CRUZI*: CÉLULAS CONTROLE**

As imagens mostram a morfologia normal de *T. cruzi* em culturas em fase exponencial (A) e estacionária (B) do crescimento. Algumas estruturas importantes são indicadas com letras; N: núcleo, K: Cinetoplasto, M: mitocôndria, F: flagelo, FP: bolsa flagelar (*flagellar pocket*), R: reservossomo. As barras de aumento possuem 2 $\mu$ m (A) e 0,5 $\mu$ m (B).

Uma das alterações mais presentes na literatura frente à inibição da síntese de esteróis em *T. cruzi* é a desorganização das membranas mitocondriais, com a presença de padrões aberrantes de membranas na matriz, seguido de intenso inchaço e perda do conteúdo da matriz. Essas alterações podem ser observadas na Figura 5.11B e E e Figura 5.12A e F.

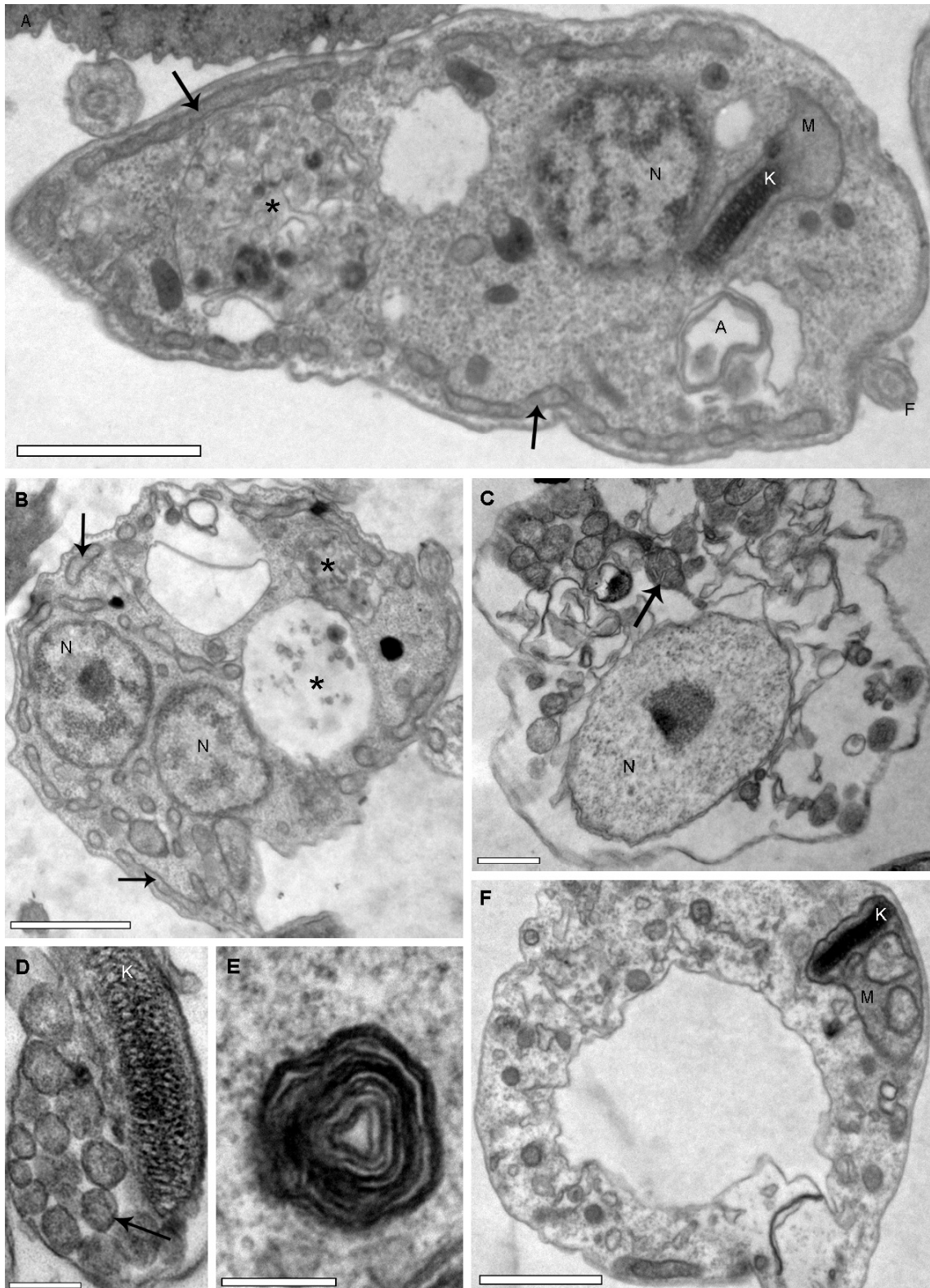
Como já observado nos experimentos de microscopia óptica (Figura 5.7), os parasitas tratados apresentam problemas no ciclo celular e divisão celular, sendo comum a presença de parasitas com dois núcleos e/ou cinetoplastos (Figura 5.11C e D e Figura 5.12B), principalmente em tempos tardios de exposição.

A presença de vacúolos lembrando autofagossomos (Figura 5.12A) e de figuras mielínicas (Figura 5.11E, em destaque na Figura 5.12E) podem indicar uma possível morte celular semelhante à autofagia em resposta à inibição da síntese de ergosterol.



**FIGURA 5.11 – ALTERAÇÕES ULTRAESTRUTURAIS À IC<sub>50</sub> DE CETOCONAZOL**

Na imagens A a C, os parasitas foram expostos por 72 horas à 32uM de Cetoconazol, enquanto em D-G por 120 horas. Os asteriscos pretos (\*) indicam os reservossomos extremamente aumentados em resposta as drogas. Setas pretas apontam para a aberrante ramificação mitocondrial. N: núcleo, K: cinetoplasto, M: mitocôndria, F: flagelo, G: Golgi. As barras de aumento possuem 1µm (A, B, C e E), 0,5µm (D e F) e 0,2µm (G).

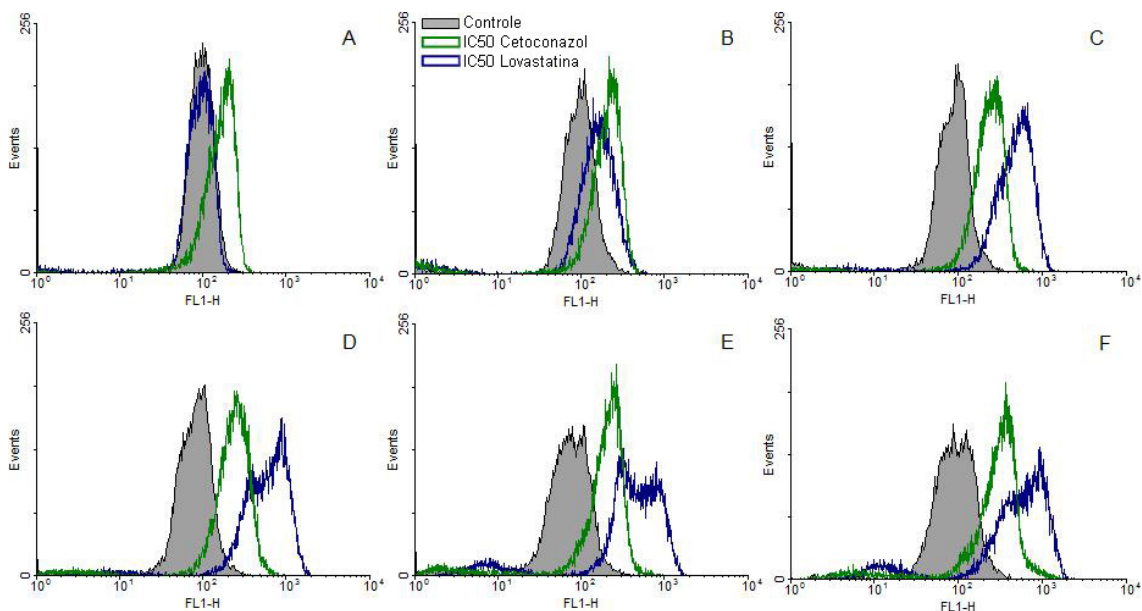


**FIGURA 5.12 – ALTERAÇÕES ULTRAESTRUTURAIS À IC<sub>50</sub> DE LOVASTATINA**

Na imagens A e E, os parasitas foram expostos por 72 horas à 50 $\mu$ M de Lovastatina, enquanto em B, C, D e F por 120 horas. Os asteriscos pretos (\*) indicam os reservossomos extremamente aumentados em resposta as drogas. Setas pretas apontam para a aberrante ramificação mitocondrial. Em E é destacado uma figura mielínica típica de células em autofagia. N: núcleo, K: quinetoplasto, M: mitocôndria, F: flagelo, A: autofagossomo. As barras de aumento possuem 1 $\mu$ m (A, B e F), 0,5 $\mu$ m (C) e 0,2 $\mu$ m (D e E).

Uma alteração que chamou muito atenção foi o inchaço muito significativo dos reservossomos, compatível com os dados de citometria de fluxo (Figura 5.8) e microscopia confocal (Figura 5.9). Tais organelas apresentaram-se extremamente hipertrofiadas, com diâmetro de pelo menos 1  $\mu\text{m}$  (asterisco na Figura 5.11A, C e E, e Figura 5.12A e B), pelo menos duas vezes maior do que em células controle (Figura 5.10). Embora existam diversos trabalhos na literatura acerca das alterações estruturais de *T. cruzi* em resposta a inibidores da síntese de esteróis, somente um deles (SANTA-RITA *et al.*, 2005), mais especificamente com cetoconazol, também demonstrou tal alteração nos reservossomos.

Uma alteração não destacada na literatura, mas facilmente observada no presente trabalho, foi a hipertrofia das membranas mitocondriais. Esta organela apresenta-se amplamente ramificada em resposta às drogas (setas na Figura 5.11D, E e F, e na Figura 5.12A, B, C e D). Tal alteração é diferente do inchaço mitocondrial já comentado, visto que as cristas mitocondriais apresentam-se compactadas e amplamente ramificadas (setas na Figura 5.11F e G, Figura 5.12C e D). Em maior aumento, é possível visualizar a dupla membrana (em destaque na Figura 5.11G) e as cristas mitocondriais (seta na Figura 5.12C). Nos cortes ultrafinos, as membranas mitocondriais podem ser visualizadas como diversos círculos vizinhos (seta na Figura 5.11F), algumas vezes em contato com o cinetoplasto (Figura 5.12D).



**FIGURA 5.13 – IC<sub>50</sub>: CITOMETRIA DE FLUXO COM RHODAMINA 123**

Os parasitas foram tratados com Cetoconazol ou Lovastatina na dose IC<sub>50</sub> e analisados por citometria de fluxo com o corante Laranja de Acridina. A: 24 horas de exposição a droga; B: 48 horas; C: 72 horas; D: 96 horas; E: 120 horas; F: 144 horas. Note o aumento do sinal de fluorescência em resposta as drogas.



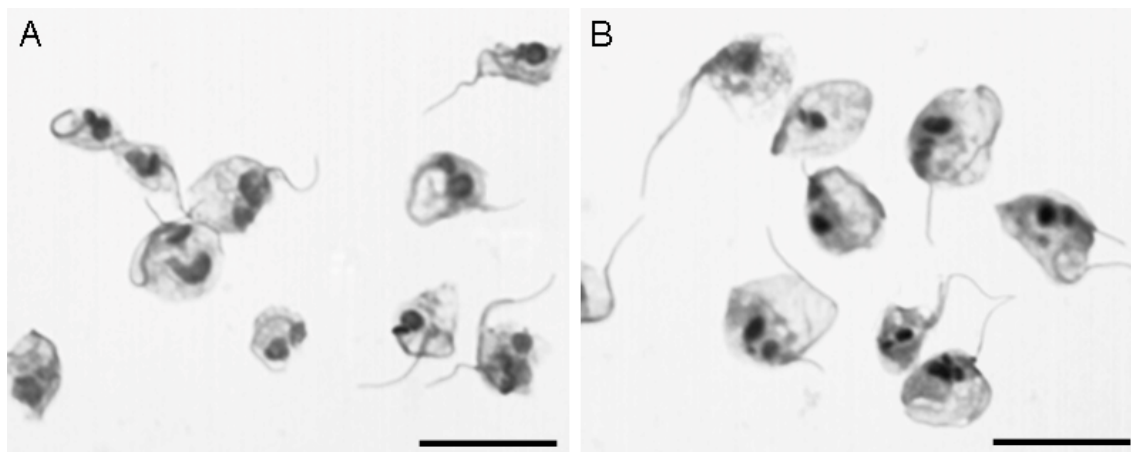
Tal proliferação das membranas mitocondriais foi analisada por citometria de fluxo com o corante rhodamina 123, o qual intercala-se na membrana mitocondrial interna devido à presença do potencial transmembrana, marcando, portanto, somente mitocôndrias funcionais. A Figura 5.13 mostra que células tratadas com as droga apresentam maior fluorescência, o que corrobora os dados de M.E.T.

A exposição dos epimastigotas de *T. cruzi* à  $IC_{50}$  de ambas as drogas culmina com a lise celular, o que acarreta na perda da elétron-densidade típica do citoplasma (Figura 5.11F e Figura 5.12C). Tais células lisadas apresentam as ramificações de membranas mitocondriais já comentadas.

Uma alteração só observada para o tratamento com  $IC_{50}$  de cetoconazol foi a desorganização do Complexo de Golgi (Figura 5.11B).

### 5.2.2 DL100

Doses letais (DL100) dos IBEs utilizados induzem drásticas alterações morfológicas às células. Com somente 12 horas de exposição os parasitas apresentam formato esférico e intensa vacuolização (Figura 5.14). Tais características são comuns em células comprometidas com a morte (revisto por KROEMER *et al.*, 2009). Portanto, buscou-se entender o tipo de morte celular programada induzida por altas doses dos IBEs.

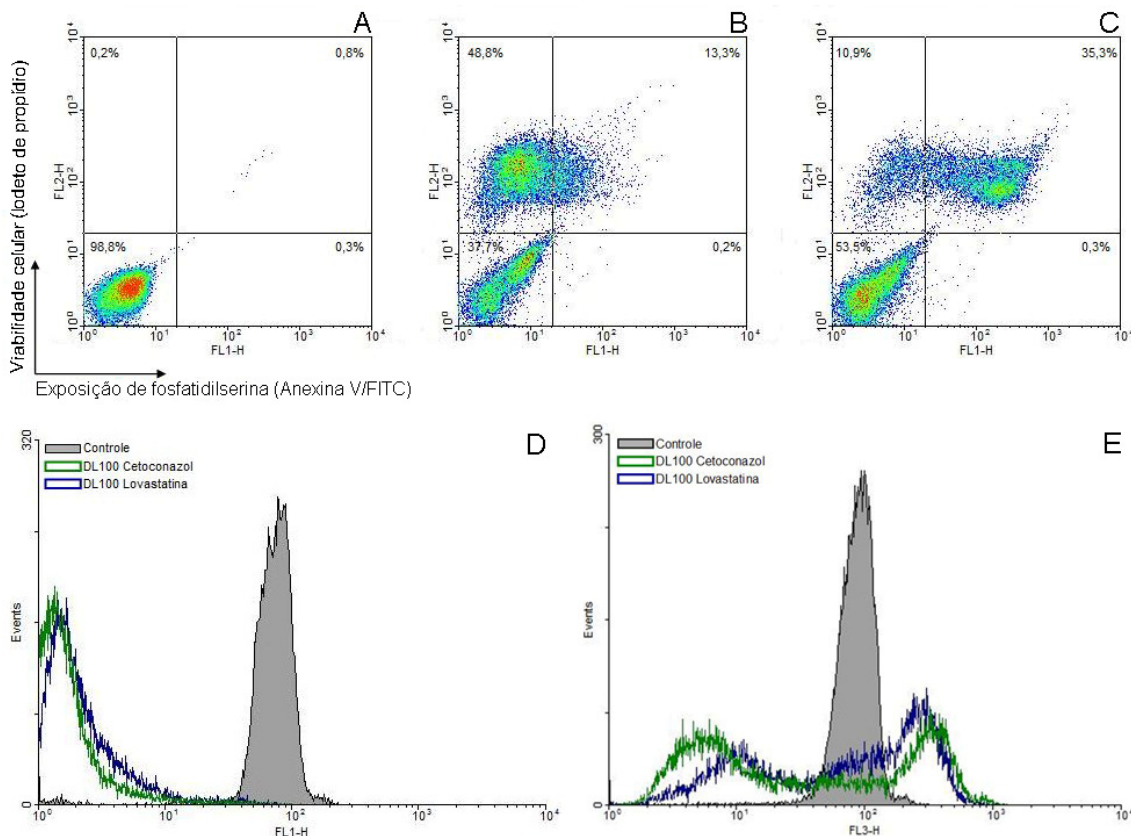


**FIGURA 5.14 – MICROSCOPIA ÓPTICA DL100.**

Os parasitas foram expostos por 12 horas a 120 $\mu$ M de Cetoconazol (A) ou 100 $\mu$ M de Lovastatina (B). A barra de tamanho possui 10 $\mu$ m.

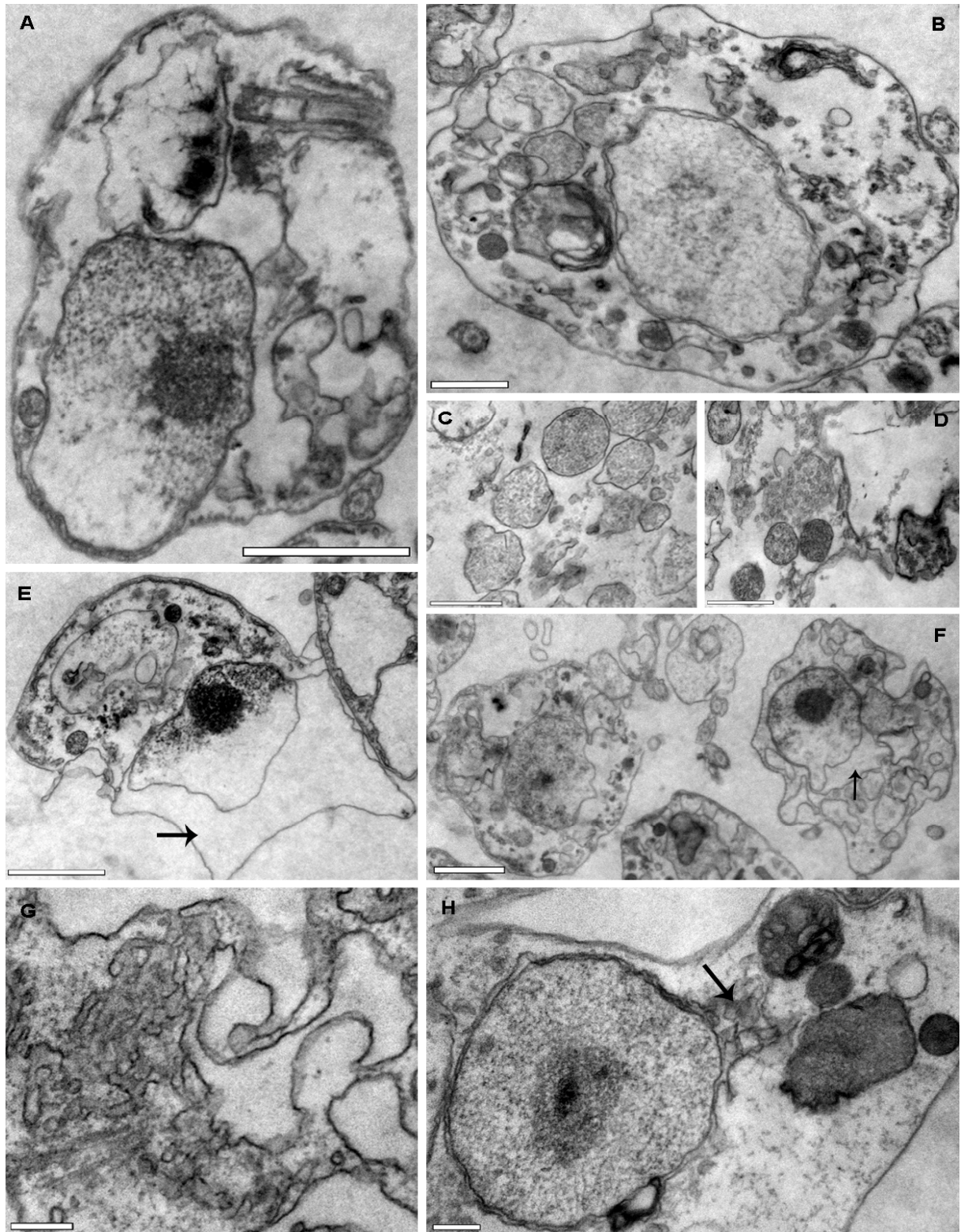
Duas características clássicas de células apoptóticas de mamíferos são a perda do potencial transmembrana mitocondrial e a exposição do fosfolípido fosfatidilserina (KROEMER *et al.*, 2009), as quais podem ser facilmente quantificadas por citometria de fluxo com o corante Rhodamina 123 e com kits para detecção de fosfatidilserina (detalhes no tópico 4.5), respectivamente. Como pode ser observado na Figura 5.15D, após 24 horas de exposição às drogas, os parasitas apresentam forte despolarização mitocondrial.

Entretanto, os percentuais de células positivas para marcação com FITC se mantêm baixos e comparáveis às células controle (~0,3%), indicando ausência de exposição de fosfatidilserina (Figura 5.15A, B e C) na morte celular induzida por altas doses dos IBEs. Cerca da metade das células em cultura (62% para lovastatina e 46% para cetoconazol, veja Figura 5.15B e C, respectivamente) perdem a integridade da membrana plasmática após 24 horas de exposição à DL100 de ambas as drogas, o que é evidenciado pela marcação positiva para o corante vital iodeto de propídio.



**FIGURA 5.15 – DL100: CITOMETRIA DE FLUXO**

Na parte superior da figura são apresentados density-plots da exposição de fosfatidilserina nas células controle (A), após 24 horas de exposição a lovastatina (B) ou cetoconazol (C). D: sobreposição dos histogramas das culturas marcadas com rhodamina 123; E: sobreposição dos histogramas das culturas marcadas laranja de acridina.



**FIGURA 5.16 – ALTERAÇÕES ULTRAESTRUTURAIS EM RESPOSTA A DL100**

Os parasitas foram expostos por 18 horas a 120 $\mu$ M de Cetoconazol (A-D) ou 100 $\mu$ M de Lovastatina (E-H). Note a intensa degradação dos componentes celulares (A, B, E e F), incluindo a desestruturação do cinetoplasto (A), inchaço de organelas citoplasmáticas (A, B, F) e intenso inchaço delular. Uma resposta comum a Cetoconazol foi a lise dos reservosossomos (B, em detalhe em C e D). Para Lovastatina foi observado um aberrante inchaço do espaço perinuclear (seta em E e F), com padrões anormais de membrana (seta em H). Os padrões morfológicos assemelham-se a morte celular por Necrose, culminando com a lise celular e presença de debirs na cultura (F). As barras de tamanho possuem 1 $\mu$ m (A, E e F), 0,5 $\mu$ m (B, C e D) ou 0,2 $\mu$ m (G e H).

Além disso, a marcação com o corante acidotrófico laranja de acridina evidenciou a presença de duas populações em resposta à DL100 das drogas: uma com maior sinal de fluorescência e outra com menor sinal (Figura 5.15E). É possível que as células com maior marcação representem parasitas com número e/ou tamanho de reservossomos aumentados, os quais possuem um pH baixo, enquanto células com menor marcação representem células com número significativo de reservossomos já lisados (Figura 5.16B, C e D). Assim sendo, é provável que altas doses das drogas desencadeiem um processo similar à necrose, o qual envolve um ganho no volume celular (Figura 5.14 e Figura 5.16), inchaço de organelas (Figura 5.16), despolarização mitocondrial (Figura 5.15D) e permeabilização de lisossomos/reservossomos (Figura 5.16), culminando na perda da integridade da membrana, ou lise celular (revisto por KROEMER *et al.*, 2009).

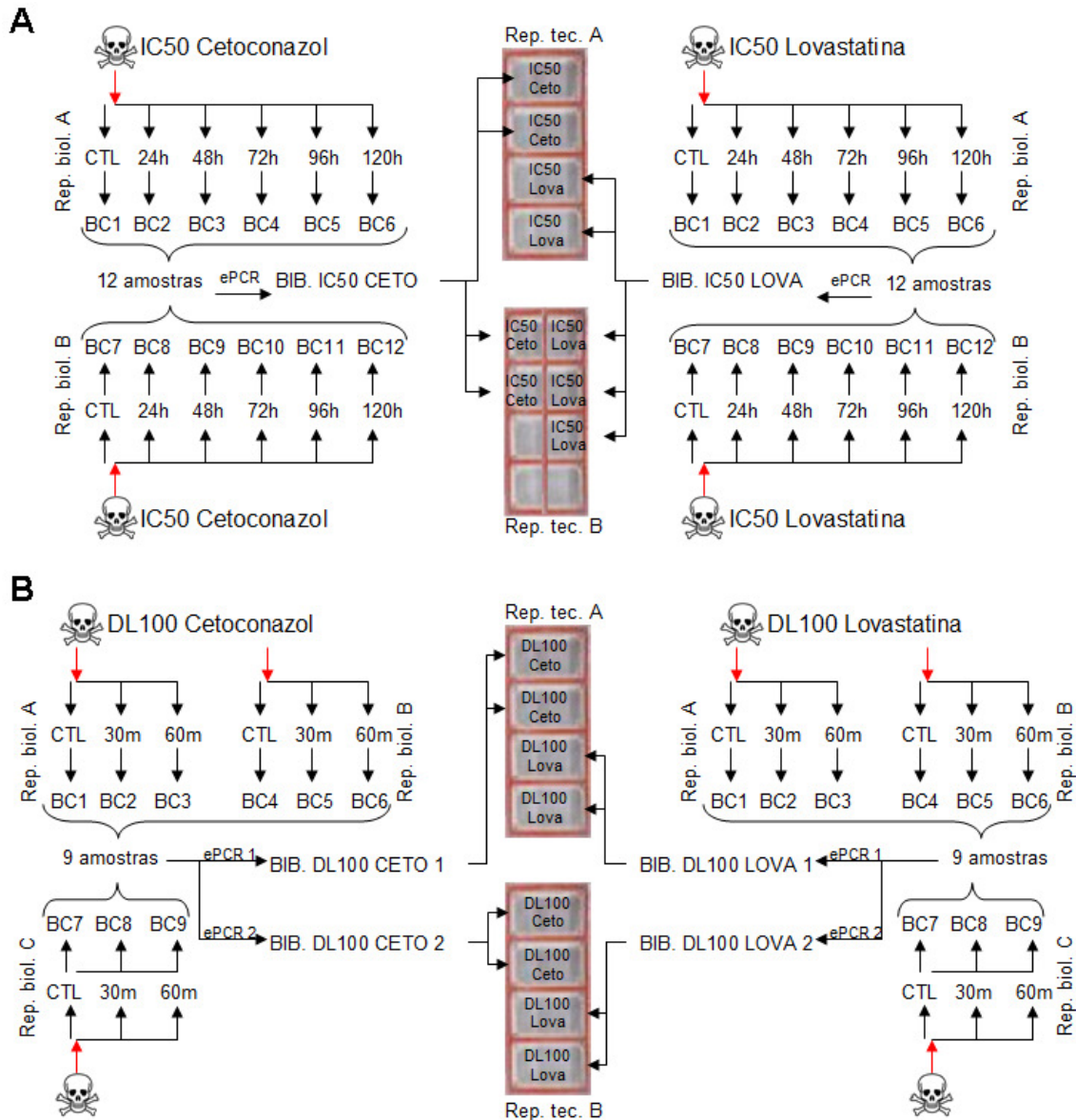
### 5.3 RNA-SEQ

#### 5.3.1 Desenho experimental

O presente trabalho se propõe a investigar a mudança do padrão de expressão gênica de *T. cruzi* em larga escala em resposta aos IBEs utilizados de duas formas distintas:

1. DL100 por períodos curtos de exposição (minutos);
2. IC<sub>50</sub> por períodos maiores (dias);

Devido à alta concentração de droga (DL100), o primeiro tipo de experimento tem como objetivo a análise de padrões de genes co-regulados no parasita de forma mais genérica, drástica e muito provavelmente não necessariamente relacionado à inibição da via de síntese de esteróis diretamente, uma vez que os efeitos de tais inibidores são tardios em *T. cruzi* (cerca de três dias) devido à presença de um *pool* de esteróis pré-formado. Estes experimentos buscam perturbar brutalmente a homeostase celular, para que padrões de expressão possam ser investigados, de forma a elucidar fenômenos biológicos básicos particulares ao parasita, analisar funções de proteínas hipotéticas, identificar elementos essenciais para a determinação de alvos terapêuticos, bem como analisar a morte celular de *T. cruzi* sobre forte pressão das drogas com um enfoque diferente (genômica funcional). Assim sendo, este tipo de experimento está relacionado diretamente, e praticamente de forma única, ao projeto reguloma.



**FIGURA 5.17 – DESENHO EXPERIMENTAL RNA-SEQ**

O esquema na figura resume a preparação das amostras e seqüenciamentos realizados no presente trabalho. As linhas horizontais representam a escala temporal e respectivos tempos de extração das amostras de RNA polissomal. O *barcode* usado para marcar cada amostra é especificado com o código BC. A: para as doses IC<sub>50</sub> foram realizadas duplicatas biológicas, compondo 12 *barcodes* (BC) em cada biblioteca (BIB. IC<sub>50</sub> CETO e BIB. IC<sub>50</sub> LOVA); estas bibliotecas foram seqüenciadas em duplicata técnica como demonstrado na parte central da figura. B: para as doses DL100 foram realizadas triplicatas biológicas, somando 9 BC para cada biblioteca (BIB. DL100 CETO e BIB. DL100 LOVA); novamente, tais bibliotecas foram seqüenciadas em duplicata técnica, mas desta vez foram incluídas as ePCRs no processo de replicação devido a falta de beads para reseqüenciamento.

Por outro lado, o segundo tipo de experimento tem como objetivo estudar as características biológicas da resposta do *T. cruzi* frente à inibição da biossíntese de esteróis pelos fármacos, sendo que a utilização da IC<sub>50</sub> permite o acompanhamento da regulação da expressão gênica por períodos de dias, de forma a caracterizar os pontos de controle desta via e suas interações com demais redes gênicas celulares, como possíveis mecanismos de

resistência às drogas. Portanto, no segundo tipo de experimento foi proposto uma análise farmacogenômica: análise dos produtos do genoma relacionados com o processo inibitório frente ao tratamento com os IBEs. Esses dados também podem ser integrados ao banco do reguloma, provendo uma perspectiva distinta dos módulos de co-regulação da célula.

Dentro deste contexto e a partir dos dados obtidos nos experimentos de inibição do crescimento de *T. cruzi* com lovastatina e cetoconazol, determinou-se o desenho experimental para as extrações de RNA polissomal a serem usadas nos experimentos de RNA-Seq para tais drogas (Figura 5.17). No experimento de IC<sub>50</sub> foram usados tempos de exposição múltiplos de 24 horas, até o ponto final de 120 horas (cinco dias) de exposição, com duplicata biológica. Já para o experimento DL100, foram realizadas triplicatas biológicas, usados os tempos de exposição de 30 e 60 minutos, tempos estes que demonstraram alta porcentagem de genes regulados em experimentos prévios de microarranjo (dados não mostrados).

As extrações de RNA polissomal foram realizadas conforme descrito no tópico 4.9.2 e foram recuperados, ao final do processo, cerca de 10 µg de RNA polissomal para cada amostra. Tais amostras foram usadas no preparo das bibliotecas de RNA-Seq (veja Figura 5.17) para seqüenciamento no sistema SOLiD (para entendimento de todo processo de preparação das amostras com inclusão do SOLiD System Barcodes, veja Figura 4.1).

### 5.3.2 Sequências geradas e alinhamento no genoma de *T. cruzi*

Ao final de cada um dos quatro seqüenciamentos realizados, as leituras geradas para cada um dos *barcodes*, em cada uma das quatro bibliotecas produzidas (veja Figura 5.17), foram somados e os dados são apresentados na Figura 5.18. As duas réplicas técnicas de seqüenciamento das bibliotecas DL100 geraram um total de cerca de 200 milhões e 118 milhões de leituras. Somando os dados das duas corridas foram obtidos quase 320 milhões de leituras para os dados DL100. Combinando as réplicas técnicas, cada uma das amostras do experimento com cetoconazol tiveram um número médio de quase 19 milhões de seqüências, enquanto para lovastatina foi produzido aproximadamente 16 milhões de leituras por amostra (em média) (detalhes na Figura 5.18, parte superior).

Já para as duas corridas das amostras IC<sub>50</sub>, foi produzido um total de 141 milhões e 66 milhões de leituras, totalizando cerca de 208 milhões de seqüências. Cada amostra da biblioteca IC<sub>50</sub> cetoconazol (BIB IC<sub>50</sub> CETO) teve uma média de 8 milhões de leituras,

enquanto para as amostra da biblioteca IC<sub>50</sub> Lovastatina (BIB IC<sub>50</sub> LOVA) foi gerado mais de 9 milhões de leituras (em média). Portanto, produziu-se cerca da metade de leituras para cada uma das amostras de IC<sub>50</sub> em comparação com os dados DL100, devido ao próprio desenho experimental (Figura 5.17) e por variações na quantificação e deposição das beads nas células de fluxo para seqüenciamento (detalhes no tópico 4.10.2.4).

Nos seqüenciamentos das amostras IC<sub>50</sub>, evidencia-se bibliotecas com número de leituras produzidas bem acima da média (C120a, C96a, L0b e L120a), o que pode ser explicado por erro na quantificação da concentração de cDNA amplificado e conseqüente sobre-representação dessas amostras no momento de combinar os cDNAs amplificados com diferentes *barcodes* para a ePCR (detalhes no tópico 4.10.2.1).

BIB DL100 CETO				BIB DL100 LOVA			
Amostra	Rep. tec. A	Rep. tec. B	total	Amostra	Rep. tec. A	Rep. tec. B	total
C0a	11,516,475	6,922,814	18,439,289	L0a	12,675,768	6,013,906	18,689,674
C0b	10,780,290	6,928,290	17,708,580	L0b	12,073,484	6,538,804	18,612,288
C0c	12,301,121	7,744,908	20,046,029	L0c	8,435,502	4,726,474	13,161,976
C30a	11,321,842	7,575,712	18,897,554	L30a	13,487,165	6,830,031	20,317,196
C30b	10,752,045	6,748,810	17,500,855	L30b	9,830,589	5,208,816	15,039,405
C30c	13,296,893	8,088,588	21,385,481	L30c	8,318,466	4,644,660	12,963,126
C60a	12,431,332	8,017,272	20,448,604	L60a	9,989,183	5,615,545	15,604,728
C60b	9,752,137	5,845,473	15,597,610	L60b	14,102,407	7,222,935	21,325,342
C60c	12,281,966	8,104,075	20,386,041	L60c	8,146,108	5,156,196	13,302,304
média	11,603,789	7,330,660	18,934,449	média	10,784,297	5,773,041	16,557,338
soma	104,434,101	65,975,942	170,410,043	soma	97,058,672	51,957,367	149,016,039
				DL100	Rep. Tec. A	Rep. Tec. B.	
				reads totais	201,492,773	117,933,309	

BIB IC50 CETO				BIB IC50 LOVA			
Amostra	Rep. tec. A	Rep. tec. B	total	Amostra	Rep. tec. A	Rep. tec. B	total
C0a	3,589,399	1,450,941	5,040,340	L0a	4,446,916	2,248,286	6,695,202
C0b	3,864,021	1,554,427	5,418,448	L0b	8,014,820	3,933,824	11,948,644
C120a	14,207,537	6,749,024	20,956,561	L120a	17,507,393	8,816,645	26,324,038
C120b	4,661,859	2,041,772	6,703,631	L120b	4,963,288	2,569,697	7,532,985
C24a	3,529,056	1,496,259	5,025,315	L24a	4,684,634	2,425,481	7,110,115
C24b	4,594,021	1,966,119	6,560,140	L24b	4,572,813	2,304,453	6,877,266
C48a	5,228,365	2,033,120	7,261,485	L48a	4,789,868	2,382,967	7,172,835
C48b	4,193,094	1,679,220	5,872,314	L48b	4,283,413	2,131,403	6,414,816
C72a	4,292,815	1,825,437	6,118,252	L72a	5,820,471	2,882,546	8,703,017
C72b	3,836,543	1,610,168	5,446,711	L72b	4,450,282	2,163,072	6,613,354
C96a	11,035,682	4,675,082	15,710,764	L96a	5,658,778	2,827,389	8,486,167
C96b	4,295,832	1,739,693	6,035,525	L96b	4,975,444	2,402,828	7,378,272
média	5,610,685	2,401,772	8,012,457	média	6,180,677	3,090,716	9,271,393
soma	67,328,224	28,821,262	96,149,486	soma	74,168,120	37,088,591	111,256,711
				IC50	Rep. Tec. A	Rep. Tec. B.	
				reads totais	141,496,344	65,909,853	

#### FIGURA 5.18 – SEQUÊNCIAS PRODUZIDAS NOS EXPERIMENTOS DE RNA-SEQ

As quatro tabelas demonstram o número de *reads* produzidos para cada uma das 18 amostras do experimento DL100 (9 amostras na biblioteca de tratamento cetoconazol, BIB DL100 CETO), e 9 amostras para Lovastatina, BIB DL100 LOVA), e 24 amostras dos experimentos IC<sub>50</sub> (12 amostras para cada biblioteca, BIB IC<sub>50</sub> CETO e BIB IC<sub>50</sub> LOVA). As amostras estão nomeadas de acordo com a droga utilizada, C (Cetoconazol) e L (Lovastatina), o tempo de exposição (em minutos para os experimento DL100 e em horas para os experimentos IC<sub>50</sub>) e a réplica biológica (a, b e c). Abaixo das tabelas DL100 e IC<sub>50</sub> encontram-se os valores do total de *reads* produzidos em cada seqüenciamento.

NOTA: o tempo zero (0) corresponde as amostras controle.

Combinando os dados dos quatro seqüenciamentos realizados, produziu-se mais de 500 milhões de leituras, o que destaca a alta produtividade dos seqüenciadores de DNA de segunda geração.

A seguir, as leituras geradas para cada amostra de RNA foram alinhados no genoma de *T. cruzi* usando a ferramenta SOLiD<sup>TM</sup> WT Analysis Pipeline (detalhes no tópico 4.10.4). Nos trabalhos de RNA-Seq normalmente os *reads* gerados são alinhados contra o genoma haplóide representativo do organismo em estudo. Entretanto, devido à complexidade e grande quantidade de seqüências repetitivas em *T. cruzi*, não há uma montagem haplóide que defina o genoma desse parasita (EL-SAYED *et al.*, 2005). Em vez disso, o genoma é apresentado como dois haplótipos distintos: Esmeraldo e não-Esmeraldo, nomes estes denominativos das linhagens parentais da cepa usada para o seqüenciamento do genoma (CL-Brener). Isso levanta o seguinte problema: qual genoma referência deve ser usado para alinhamento das leituras geradas no presente trabalho? Uma solução seria agrupar genes com alta similaridade e representá-los por uma única seqüência referência, como foi feito num recente trabalho de RNA-Seq de *Trypanosoma brucei* (SIEGEL *et al.*, 2010). Embora esse método seja útil para alinhar seqüências normalmente alocadas em múltiplos locais no genoma, ele não tem capacidade de detectar expressão diferencial entre membros de certas famílias multigênicas (visto que todos os membros da família serão representados pela mesma seqüência referência) ou expressão de alelos específicos; além disso, este método enfrenta a dificuldade de se produzir seqüências referência para todos os grupos de genes similares, o que pode introduzir desvios nos dados dependendo da linha de corte de similaridade usada para representar dois genes por uma mesma seqüência referência. Por estes motivos, não utilizamos tal método. Outra hipótese seria trabalhar apenas com um dos haplótipos parentais de CL-Brener; embora essa seja uma solução simples, seria uma representação pobre da expressão gênica de *T. cruzi*, visto que diversos genes desse parasita só estão presentes em um dos haplótipos.

Por estes motivos, as leituras foram alinhadas contra os dois haplótipos e foram levados em conta seqüências com alinhamento múltiplo em até 10 posições diferentes do genoma. Como o presente trabalho pretende identificar genes diferencialmente expressos entre diferentes situações, os desvios na quantificação dos níveis de expressão introduzidos pela consideração de leituras de alinhamento múltiplo devem afetar as diferentes amostras de forma similar. Esse método propicia uma maior amostragem nos dados de alinhamento, proporcionando maior poder nas análises estatísticas posteriores. Tais vantagens discutidas acima justificam a utilização de leituras com alinhamento múltiplo no presente trabalho,



mesmo embora a maioria dos trabalhos de RNA-Seq até hoje utilizem apenas leituras de alinhamento em locais únicos do genoma, o que é inviável em *T. cruzi*. Além disso, se não considerados os alinhamentos múltiplos, observamos que usando os dois haplótipos do genoma de *T. cruzi* apenas cerca de 13% das leituras alinham em um único local no genoma, de forma que mais de 80% das sequências produzidas são desperdiçadas.

Como possíveis problemas dessa técnica, é criada uma situação semelhante à hibridação cruzada presente em microarranjos, nos quais genes muito semelhantes tem a quantificação de seu sinal misturada. No entanto, foi possível avaliar que, em média, somente 10% das leituras mapeiam em mais de 10 posições (uma situação muito comum para a grande quantidade de famílias gênicas de *T. cruzi*, que podem perfazer até 30% do transcriptoma sequencial), e que a grande maioria mapeia somente em uma, duas (situação mais comum, pela existência de dois alelos no genoma referência) ou três posições (situação também comum, pela existência de parálogos idênticos), o que demonstra que o problema de hibridação cruzada (sinal oriundo de dois genes similares, mas não idênticos) não constitui uma distorção grande nos dados.

Como resultado dos alinhamentos, o número de leituras alocadas em cada gene é transformado em uma contagem digital da expressão gênica para todos os genes em cada condição de estudo (especificadas por um *barcode*, veja Tabela 5.1).

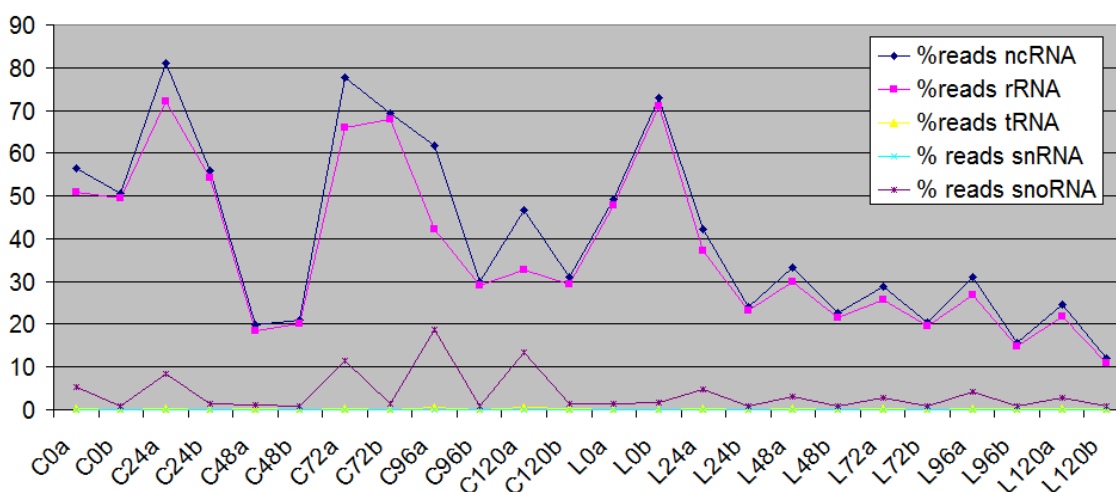
**TABELA 5.1 – RESULTADOS DOS DADOS DE ALINHAMENTO**

geneID	C0a	C0b	C0c	C30a	C30b	C30c	C60a	C60b	C60c
TcChr32-S 255	71	81	75	43	32	66	50	49	66
TcChr29-P 132	16	8	14	4	12	11	14	12	6
TcChr16-S 219	171	189	205	185	174	208	209	153	181
TcChr28-P 195	23	12	15	14	13	11	19	16	14
TcChr35-P 236	364	353	337	361	294	426	463	348	421
TcChr15-S 102	796	685	722	557	531	838	637	511	641
TcChr30-S 58	294	291	367	357	284	384	368	334	349
TcChr17-P 71	365	404	465	287	333	405	373	321	312
TcChr11-P 139	6122	8066	8984	9329	10513	14465	10574	10273	11296
TcChr29-S 147	2644	2156	2456	2352	2253	2604	2752	2431	2341

A tabela mostra como exemplo o resultado do alinhamento para 10 genes nas 9 amostras presentes na biblioteca BIB. DL100 CETO 1. Na primeira coluna é especificado o código de cada gene e nas colunas vizinhas o dado bruto de contagem para cada amostra. O nome das amostras é o mesmo que na Figura 5.17.

Normalmente, mais de 90% do RNA presente numa célula eucariótica representa rRNA. Para evitar a quantificação exacerbada deste tipo de molécula nos dados de RNA-Seq, e consequente perda de cobertura em RNAs codificadores, faz-se necessário a retirada do rRNA do transcriptoma. No presente trabalho foi utilizada a purificação de RNA-poli(A) com matriz oligodT-celulose (detalhes na Figura 4.1). Após o alinhamento das leituras,

observamos que para os dados DL100, menos de 3% (média de 2,6%) do total de leituras alinhadas em cada amostra representam rRNA. No entanto, a purificação de RNApoli(A) nas amostras dos experimentos IC<sub>50</sub> apresentou uma proporção muito maior de rRNA, sendo que algumas amostras apresentaram cerca de 80% de ncRNA (*non-coding* RNA: rRNA, tRNA, snRNA e snoRNA) em sua composição (veja Figura 5.19). Por este motivo, após a retirada das sequências de ncRNA, o universo de leituras úteis para análises de expressão diferencial caiu drasticamente; entretanto, a biblioteca IC<sub>50</sub> com o menor número de leituras ainda possuía quase um milhão de leituras alinhadas (C24a, ~887 mil leituras alinhadas), com média de cerca de 4,7 milhões de leituras alinhadas.



**FIGURA 5.19 – PRESENÇA DE RNAs NÃO CODIFICADORES NOS DADOS**

Para cada uma das 24 amostras está plotado a porcentagem das sequências da mesma que alinham em cada um dos tipos de RNAs não codificadores. ncRNA: *non-coding* RNA (soma de todos RNAs não codificadores); rRNA: RNA ribossomal; tRNA: RNA transportador; snRNA: *small nuclear* RNA; snoRNA: *small nucleolar* RNA.

Em um trabalho recente, Li e colaboradores (2008) analisaram a quantidade necessária de leituras alinhadas por amostra para análises digitais do transcriptoma por RNA-Seq. Para o genoma humano, é necessário cerca de 1,2 milhões de leituras mapeáveis para detectar quais genes são expressos e em torno de 5,5 milhões de leituras mapeáveis para analisar a expressão diferencial entre condições distintas, em um nível no qual a adição de tags não produzirá mais de 5% de novas descobertas (LI *et al.*, 2008). Assim sendo, mesmo embora os dados de IC<sub>50</sub> estejam contaminados com ncRNA, ainda foi produzido uma quantidade de *reads* alinhados por amostra (leituras de alinhamento múltiplo) suficiente para análises da expressão diferencial, mas com um poder de detecção bem menor em comparação com os dados DL100

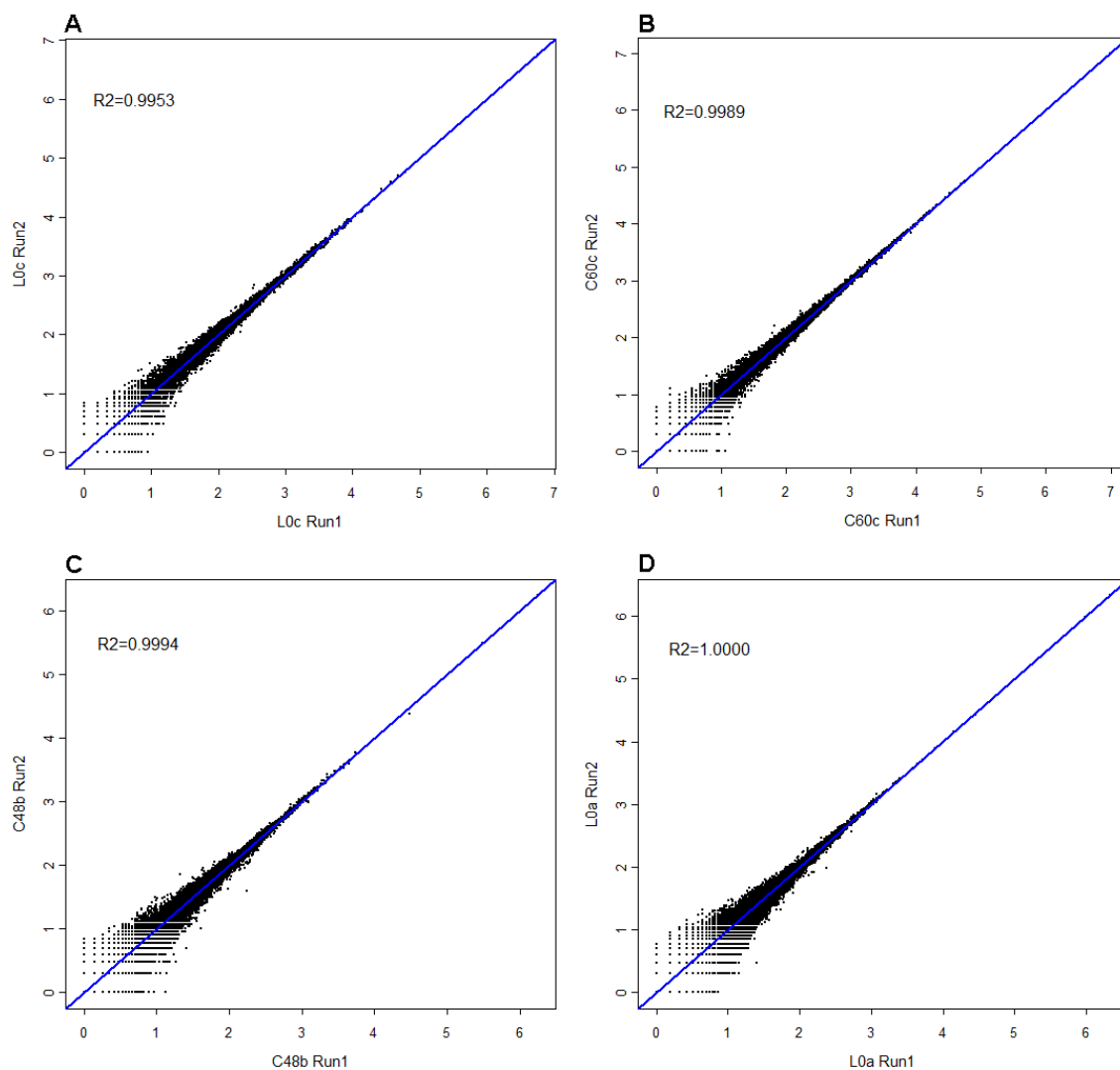
nos quais foram obtidos muito mais do que o limiar teórico de 5,5 milhões de leituras alinhadas.

A correção dessa distorção será feita futuramente, ao produzir uma nova corrida incorporando as amostras que apresentaram um número de leituras mapeáveis muito baixo.

### 5.3.3 Réplicação técnica e biológica

Como demonstrado na Figura 5.17, os experimentos realizados no presente trabalho foram desenhados de forma a controlar tanto a variabilidade técnica do protocolo de RNA-Seq quanto a variabilidade biológica inerente à resposta das cultura de *T. cruzi* às drogas.

A partir dos dados de alinhamento de cada uma das réplicas técnicas comparou-se a replicação do seqüenciamento através de regressões lineares entre os dados de contagem da mesma amostra nas duas corridas distintas. Para tal foram usados os dados brutos de contagem, incluindo ncRNAs. Para os dados DL100 (veja Figura 5.20A e B), biblioteca cetoconazol, obtivemos um valor de correlação linear ( $R^2$ ) médio de 0,9969, valor similar à média para a biblioteca lovastatina: 0,9957. Para os dados de  $IC_{50}$  (veja Figura 5.20C e D) a média foi de 0,9999 para ambas as bibliotecas (cetoconazol e lovastatina); chamou atenção o valor  $R^2$  médio das bibliotecas de  $IC_{50}$  ser maior do que para as bibliotecas DL100, visto que no experimento DL100 foi obtida uma cobertura bem maior. Quando desconsiderado ncRNA, as amostras  $IC_{50}$  da biblioteca cetoconazol tem um  $R^2$  médio de 0,9861, e as de lovastatina 0,9960, valores similares aos das bibliotecas DL100.



**FIGURA 5.20 – RÉPLICAS TÉCNICAS SOLiD**

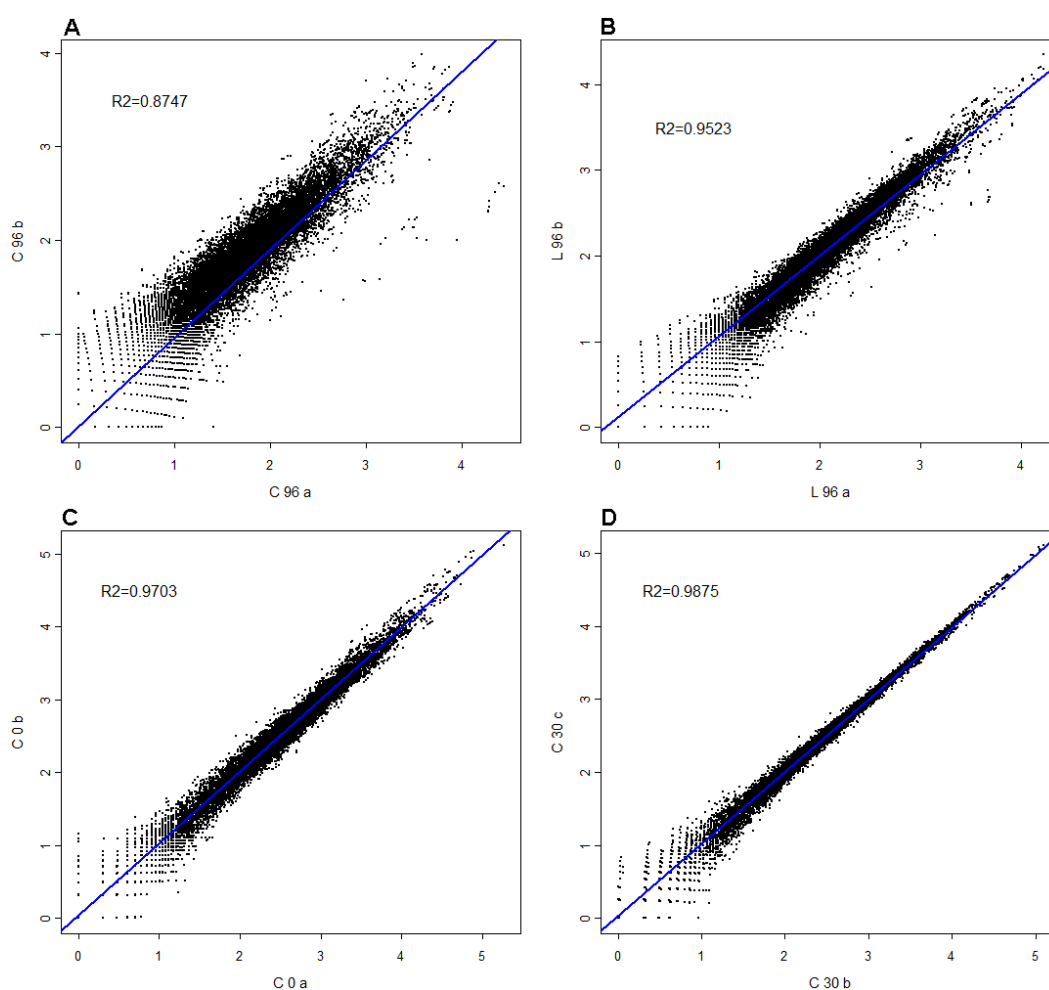
Na figura estão plotados quatro exemplos de correlação linear entre as réplicas técnicas do SOLiD. De forma a homogeneizar as escalas nos eixos x e y, as bibliotecas de dados brutos foram normalizadas por número total de *reads* alinhados, sendo plotado o valor de  $\log_{10}(x+1)$ , onde x representa o número de *reads* alinhados em cada gene. A reta de regressão linear é evidenciada em azul, com o respectivo valor de  $R^2$  especificado dentro de cada gráfico. Na parte de cima da figura, a amostra com menor (A) e maior (B) valor de  $R^2$  nos dados DL100. De forma similar, na parte inferior está plotado a amostra com menor (C) e maior (D) valor de correlação linear para os dados IC<sub>50</sub>. As amostras são nomeadas da mesma forma que na Figura 5.18. Run1: réplica técnica A; Run2: réplica técnica B.

Como os dados de seqüenciamento na plataforma SOLiD se mostraram muito reprodutíveis, é possível desconsiderar as réplicas técnicas, e as duas contagens foram somadas de forma a ter uma única contagem final por amostra biológica.

A seguir, foram realizadas comparações das réplicas biológicas através de regressões lineares dos dados de alinhamento, de forma similar as comparações para réplicas técnicas. Para a biblioteca cetoconazol, dados DL100 (Figura 5.21C e D), obtivemos um valor de correlação linear ( $R^2$ ) médio de 0,9805 (variando de 0,9703 a 0,9875), valor similar à média

para a biblioteca lovastatina: 0,9840 (variando de 0,9811 a 0,9857), o que demonstra uma alta correlação entre as réplicas biológicas do experimento DL100.

Para os dados IC<sub>50</sub> (Figura 5.21A e B), as réplicas biológicas mostraram maior variabilidade biológica, com valores de correlação linear menores, o que pode ser explicado pelo menor número de leituras nesse experimento e pela própria complexidade biológica desse tipo de tratamento com longos tempos de exposição, além de flutuações estocásticas inerentes à manipulação. Para as réplicas biológicas do tratamento com cetoconazol, obteve-se uma média de regressão de 0,8996, variando entre 0,8747 e 0,9346; enquanto para lovastatina, a média de correlação linear foi 0,9326 (variando de 0,8997 a 0,9523).



**FIGURA 5.21 - RÉPLICAS BIOLÓGICAS SOLiD**

Na figura estão plotados quatro exemplos de correlação linear entre as réplicas biológicas sequenciadas no SOLiD. De forma a homogeneizar as escalas nos eixos x e y, as bibliotecas de dados brutos foram normalizadas pelo programa edgeR (detalhes no tópico seguinte), sendo plotado o valor de  $\log_{10}(x+1)$ , onde x representa o número de *reads* alinhados em cada gene. A reta de regressão linear é evidenciada em azul, com o respectivo valor de  $R^2$  especificado dentro de cada gráfico. Na parte de cima da figura, a amostra com menor (A) e maior (B) valor de  $R^2$  nos dados IC<sub>50</sub>. De forma similar, na parte inferior está plotado a amostra com menor (C) e maior (D) valor de correlação linear para os dados DL100. As amostras estão nomeadas da mesma forma que na Figura 5.18. Note a maior dispersão, e conseqüente menor valor de  $R^2$ , nas réplicas biológicas de IC<sub>50</sub>.

### 5.3.4 Genes Diferencialmente expressos

Após a análise técnica dos experimentos de RNA-Seq, foi feita a análise biológica dos genes diferencialmente expressos em resposta às drogas. Para tal, foi utilizado o pacote edgeR (*empirical digital gene expression analysis using R*) (ROBINSON *et al.*, 2010), o qual funciona dentro do *software R*, programa de acesso livre para análises computacionais estatísticas e gráficas.

Como recentemente demonstrado, a normalização dos dados é um passo essencial nas análises de RNA-Seq (BULLARD *et al.*, 2010; ROBINSON & OSHLACK, 2010). Os dados brutos de alinhamentos foram inicialmente normalizados pelo método TMM do pacote edgeR (*trimmed mean of MxA values*) (ROBINSON & OSHLACK, 2010), o qual calcula fatores de normalização entre as bibliotecas excluindo valores extremos de M ( $\log_2[x/y]$ : variação na expressão) e A ( $\log_2[x*y]$ : intensidade) e constrói bibliotecas normalizadas a partir destes fatores. Segundo os autores (ROBINSON & OSHLACK, 2010), este método corrige possíveis alterações no número de *reads* gerados devido diferentes composições dos transcriptomas sendo comparados. Dentro do programa edgeR, ainda foi feita uma segunda normalização, pelo método quantil, antes de realizar as comparações de genes diferencialmente expressos.

Os diferentes tempos de exposição às drogas foram comparados com o tempo zero (controle), sendo que foram usados diferentes níveis de significância estatística (FDR) e biológica (*fold change*).

#### 5.3.4.1 IC<sub>50</sub> Cetoconazol

A Tabela 5.2 mostra o número de genes diferencialmente expressos em resposta à IC<sub>50</sub> de Cetoconazol. Devido à utilização de apenas duas réplicas biológicas para cada ponto de análise e grande variação biológica intrínseca das amostras (veja Figura 5.21A e B), o número de genes detectados como diferencialmente expressos não foi muito alto, principalmente nos pontos de menor tempo de exposição as drogas (24 e 48 horas) e quando considerados valores mais exstringentes de FDR ou limiares de *fold change* (FC). Isso se deve ao baixo poder estatístico inerente a um experimento com duas réplicas.

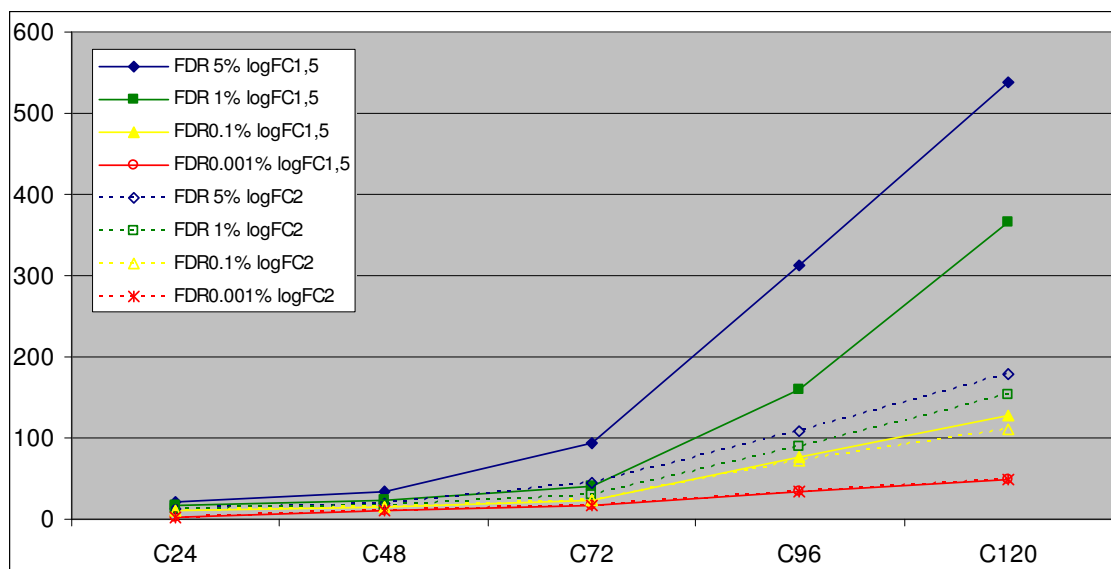
TABELA 5.2 – IC<sub>50</sub> CETOCONAZOL: GENES SIGNIFICATIVOS

sem critério FC				
Comparação	FDR 5%	FDR 1%	FDR 0.1%	FDR 0.01%
C24xC0	22 (21, 1)	16 (1, 15)	11 (11, 0)	2 (2, 0)
C48xC0	35 (10, 25)	23 (3, 20)	15 (0, 15)	11 (0, 11)
C72xC0	94 (82, 12)	40 (37, 3)	23 (23, 0)	17 (17, 0)
C96xC0	346 (268, 78)	159 (135, 24)	77 (65, 12)	35 (31, 4)
C120xC0	661 (491, 170)	366 (274, 62)	127 (109, 18)	49 (47, 2)
Total	752	405	163	74
log <sub>2</sub> (FC) 1,5				
C24xC0	22 (21,1)	16 (1, 15)	11 (11, 0)	2 (2, 0)
C48xC0	35 (10, 25)	23 (3, 20)	15 (0, 15)	11 (0, 11)
C72xC0	94 (82, 12)	40 (37, 3)	23 (23, 0)	17 (17, 0)
C96xC0	313 (249, 64)	159 (135, 24)	77 (65, 12)	35 (31, 4)
C120xC0	538 (403, 135)	366 (291, 75)	127 (109, 18)	49 (47, 2)

Para cada uma das comparações foram usados diferentes limiares de FDR; o total de genes detectados como diferencialmente expressos em cada comparação é apresentado, sendo que os números entre parênteses (x, y) dizem respeito a quantidade de genes com expressão aumentada, x, e diminuída, y, em relação ao controle. A linha “Total” representa a soma do número de genes significativos em cada ponto de comparação, sem redundância. Na parte superior da tabela são apresentados os resultados sem critério de *fold change*, enquanto na parte inferior só foram levados em conta genes com  $\log_2(\text{fold change})$  maior ou igual a 1,5 (alteração absoluta mínima na expressão de 2,83 vezes).

Para o experimento com cetoconazol, quando usado FDR menor ou igual a 1%, todos os genes selecionados possuem um valor absoluto de FC maior ou igual a 2,8 vezes ( $\log_2\text{FC}$  maior que 1,5; compare a parte superior e inferior da Tabela 5.2), o que aponta para uma significância biológica dos dados quando usado maior estringência estatística, mas um baixo poder estatístico no teste usado devido à presença de apenas duas réplicas. Ou seja, alterações mais sutis na expressão (*fold change* de 1,5 vezes, ou  $\log_2\text{FC}$  de 0,58) não são passíveis de serem identificadas, de forma que diversos genes biologicamente significativos na resposta à droga passam despercebidos.

Outra forma de visualizar o número de genes significativos em cada ponto temporal é apresentado na Figura 5.22. Fica claro que o número de genes significativos aumenta gradualmente com o tempo de exposição à dose IC<sub>50</sub> da droga, independentemente da significância estatística (FDR) e biológica ( $\log_2\text{FC}$ ) utilizada, o que condiz com a ação tardia de inibidores da síntese de esteróis em *T. cruzi* e uma resposta tempo-dependente. Isso é uma representação indireta de que, de maneira geral, os resultados são oriundos da modulação da expressão gênica causada pela droga.



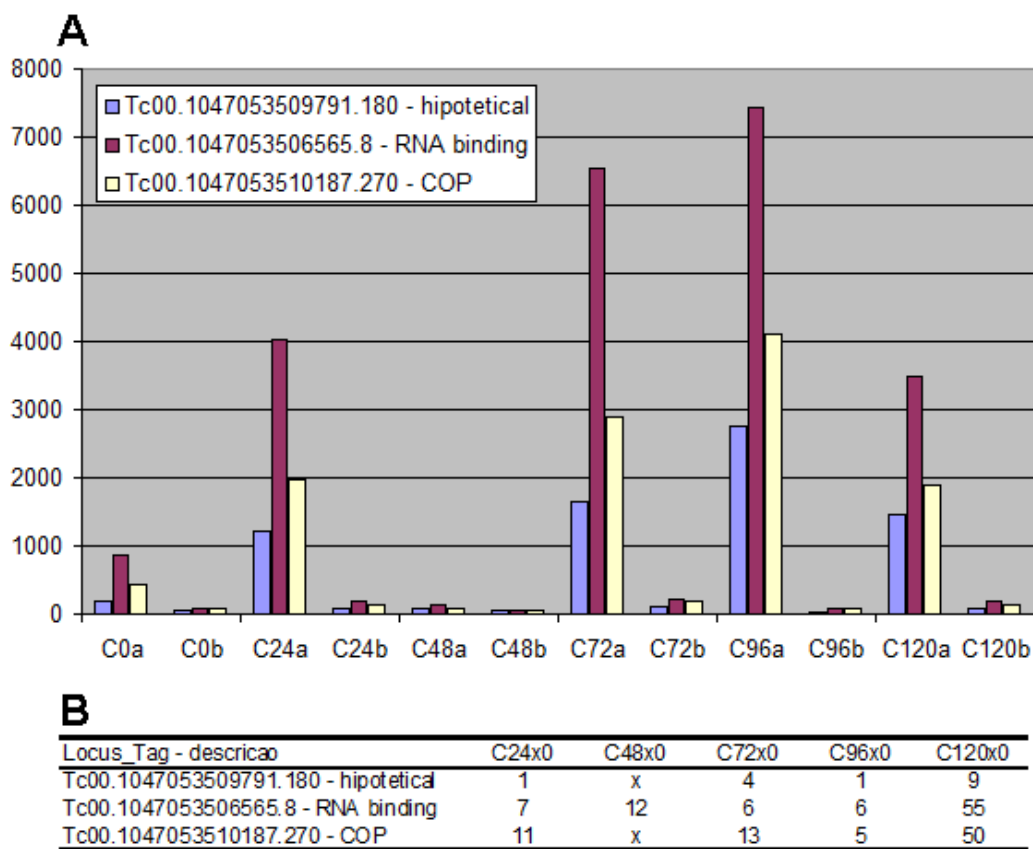
**FIGURA 5.22 – IC<sub>50</sub> CETOCONAZOL: N DE GENES SIGNIFICATIVOS EM CADA PONTO**

Cada linha representa o número de genes detectados como diferencialmente expressos, em relação ao controle, ao longo dos tempos de exposição a droga para um determinado FDR e logFC limite. No eixo x está plotado cada um dos pontos de comparação com o controle (zero horas), enquanto o eixo y representa o número de genes detectado como diferencialmente expressos. As linhas cheias representam os mesmos dados da parte inferior da Tabela 5.2, enquanto as linhas pontilhadas são os dados para logFC maior ou igual a 2 (4 vezes de alteração mínima na expressão).

Inicialmente, os genes diferencialmente expressos em resposta à droga foram analisado através de comparações de cada ponto de exposição (1 a 5 dias), contra o controle. Entretanto, análises detalhadas dos dados demonstraram que esta não era a melhor forma de analisar tal resposta sutil ao longo do tempo, principalmente com apenas duas réplicas biológicas. Na resposta à cetoconazol, considerando todos os pontos de análise, foram encontrados 163 genes diferencialmente expressos com FDR0,1%.

Todos os genes detectados na comparação 48 horas x Controle possuíam expressão diminuída (veja Tabela 5.2), mesmo embora estes mesmos genes estejam diferencialmente aumentados nas demais comparações (24 x 0, 72 x 0, 96 x 0 e 120 x 0; Tabela 1 do ANEXO II). De forma a entender esse problema, analisou-se os dados das leituras alinhadas e normalizadas mais profundamente. É possível observar na Figura 5.23 que a causa dessa mudança é técnica, sem significado biológico, pois para diversos genes analisados, a réplica biológica B não responde da mesma forma que a réplica A, diluindo a expressão diferencial presente na primeira réplica; além disso, o ponto 48 horas apresenta sinal bem mais baixo para as duas réplicas, de forma que os genes são detectados como diminuídos em relação ao controle; isso pode ser explicado por problemas na extração de RNA polissomal ou questões técnicas no seqüenciamento dessas amostras, o que será explorado futuramente.





**FIGURA 5.23 – EXEMPLOS DE GENES SIGNIFICATIVOS PARA IC<sub>50</sub> DE CETOCONAZOL**

Em A está plotado um gráfico de barras com 3 genes (cada um representado por uma cor) detectados como diferencialmente expressos em resposta a IC<sub>50</sub> de cetoconazol. No eixo x são nomeados cada um dos pontos de análise; C (Cetoconazol), seguido do tempo de exposição em horas (0-120 horas) e a réplica biológica (a e b). Para cada gene, está plotado o nível de expressão normalizado pelo tamanho da biblioteca e tamanho do gene segundo método RPKM. A tabela em B mostra a posição de cada gene, em cada um das comparações contra o controle (Tabela I do ANEXO II).

Levando em conta essa variabilidade entre as réplicas, agregado ao fato da existência de somente duas réplicas biológicas, os testes estatísticos utilizados no programa edgeR só detectam genes com alto grau de mudança (veja Tabela 5.2), quando comparados cada ponto de exposição da droga com o controle. Como a resposta à dose IC<sub>50</sub> é tempo-dependente, com aumento de genes diferencialmente expressos a cada dia de exposição a droga (Figura 5.22), analisamos os dados de uma forma diferente, de maneira a compensar o número de réplicas biológicas nos dados. Para tal, comparou-se os pontos C0+C24 contra C96+C120; obtendo-se quatro réplicas para cada ponto de análise, ao mesmo tempo que se aumenta a variabilidade biológica das réplicas, pois pontos temporais distintos estão sendo agregados. Em resumo, perde-se informação por aumentar a variabilidade biológica, mas essa perda é amplamente compensada pelo ganho em poder estatístico ao se aumentar o número de réplicas de dois para

quatro; obviamente, diferenças pontuais que sejam específicas de um dos dias agrupados não serão observadas, mas essas muito provavelmente não seriam identificadas pelo baixo número de réplicas.

**TABELA 5.3 – IC<sub>50</sub> CETOCONAZOL: GENES SIGNIFICATIVOS PARA 0+24 VERSUS 96+120**

FC	Comparação	FDR 5%	FDR 1%	FDR 0.1%	FDR 0.01%
nenhum	0+24x96+120	1792 (1156, 636)	915 (648, 267)	449 (339, 110)	255 (205, 50)
nenhum	Total ponto a ponto	752	405	163	74
1,5	0+24x96+120	1792 (1156, 636)	915 (648, 267)	449 (339, 110)	255 (205, 50)
2	0+24x96+120	1276 (859, 417)	911 (645, 266)	449 (339, 110)	255 (205, 50)
2,83	0+24x96+120	340 (264, 76)	309 (243, 66)	255 (208, 47)	219 (182, 37)

A primeira coluna especifica o tipo de limiar de *fold change* utilizado. Os testes foram feitos no programa edgeR utilizando as amostra controle (zero hora) e 24 horas de exposição como réplicas do início do processo de tratamento e as amostras 96 e 120 horas como réplicas do final do tratamento. A linha “Total ponto a ponto” diz respeito ao número total de genes detectados como significativos (sem redundância) para as comparações individuais contra o controle, como mostra a Tabela 5.2.

A Tabela 5.3 mostra o número de genes detectados como diferencialmente expressos com diferentes limiares de FDR e FC. Fica claro que a comparação dos pontos extremos no processo de resposta a IC<sub>50</sub> de cetoconazol possibilita a identificação de um número bem maior de genes diferencialmente expressos em relação às comparações individuais (ponto a ponto), independentemente do limiar de FDR usado. Considerando um FDR máximo de 0,1%, observou-se que 137 genes são identificados como diferencialmente expressos em ambos os métodos de análise (ponto a ponto ou comparando os extremos), 26 genes são únicos das comparações ponto a ponto e 312 genes são exclusivamente detectados na comparação dos extremos do tratamento. Esses resultados indicam que a comparação individual de cada ponto de extração de RNA polissomal contra o controle é pouco informativa, para este tipo de experimento de resposta lenta com apenas duas réplicas biológicas.

Embora o programa edgeR tenha sido desenvolvido especificamente para análises de RNA-Seq, a possibilidade de fazer somente testes pareados impede uma análise global da resposta temporal dos genes a IC<sub>50</sub> das drogas, o que propiciaria a detecção de alterações mais sutis ao longo dos dias de exposição a droga. Pelos diversos motivos aqui apresentados, também utilizamos a ferramenta estatística desenvolvida para análises de microarranjo denominada SAM (*significant analysis of Microarray*) (TUSHER *et al.*, 2001). A grande vantagem desse programa é a possibilidade de se analisar os dados por testes multiclasse que comparam todos os pontos experimentais e buscam padrões de alteração ao longo do processo

analisado; ou seja, repostas individualmente sutis nas comparações ponto a ponto (e com baixa significância devido a presença de apenas duas réplicas por ponto) se tornam alterações mais robusta quando analisado o processo como um todo.

Para as análises no SAM, foram usadas as tabelas de contagens de leituras normalizadas pelo programa edgeR. Quando utilizado a análise *multiclass* com 6 classes (cada um dos pontos experimentais), cada uma com duas réplicas biológicas, o número de genes detectados como diferencialmente expressos foi muito baixo (18 genes com FDR de 10%). Essa fato novamente mostra o baixo poder dos métodos estatísticos quando usado somente duas réplicas biológicas para cada ponto de análise, principalmente quando tais réplicas possuem alta variação intrínseca (veja Figura 5.21). De forma a aumentar o grau de confiabilidade biológica das análises, e sobrepujar a grande variação entre as réplicas A e B, consideramos os dados como apenas 3 classes no teste SAM multiclasse: 0+24, 48+72 e 96+120 horas. Tal análise aumentou significativamente a quantidade de padrões de expressão diferencial biologicamente e estatisticamente significantes identificados: 649 genes identificados com FDR de 10%, 338 destes com FC mínimo de 2 vezes. Este resultado novamente mostra que a resposta a IC<sub>50</sub> dos inibidores da síntese de esteróis é gradual, com aumento da significância estatística e biológica dos genes diferencialmente expressos em tempos mais tardios do processo, e reforça a importância de mais réplicas biológicas em experimentos com alta variação biológica, o que geralmente é complexo de ser obtido pelos altos custos dos experimentos.

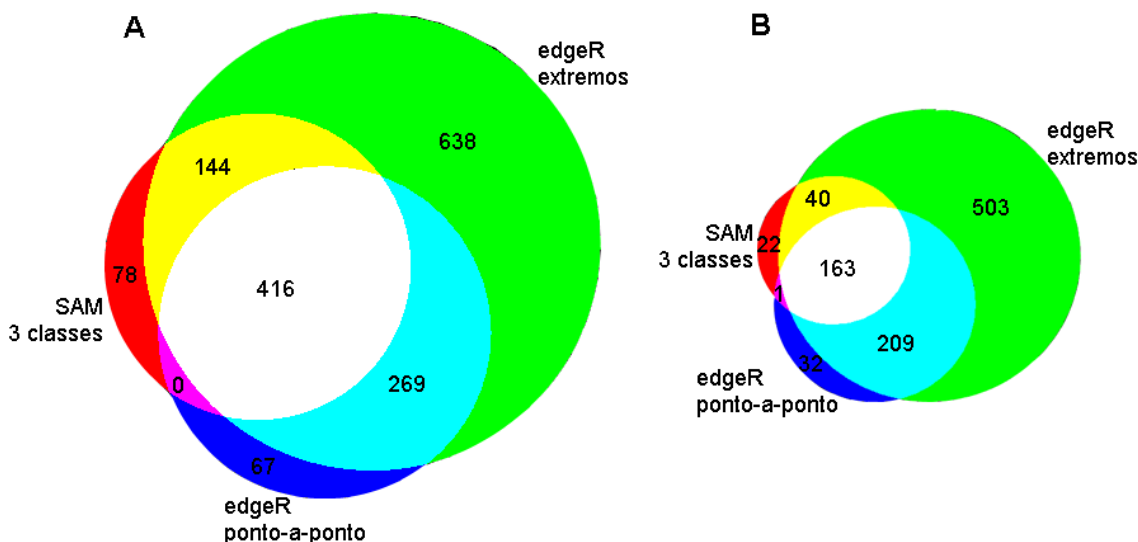
**TABELA 5.4 – IC<sub>50</sub> CETOCONAZOL: GENES SIGNIFICATIVOS TESTE SAM 3 CLASSES**

FC	Comparação	FDR10%	FDR5%	FDR1%
sem FC	SAM: 3 classes	649	226	19
FC1,5	SAM: 3 classes	649	226	19
FC2	SAM: 3 classes	638	224	19
FC3	SAM: 3 classes	411	199	19
FC4	SAM: 3 classes	201	113	18

As colunas trazem o total de genes detectados como significativos com diferentes níveis de significância estatística (FDR). A coluna FC (*foldchange*) refere-se a diferentes níveis de significância biológica na alteração da expressão gênica (pelo menos um ponto em relação ao controle); FC2 corresponde a um mínimo de 100% de alteração nos níveis de expressão, e assim sucessivamente. As comparações foram feitas no programa SAM, usando 3 classes (detalhes no texto).

Como demonstrado, foram utilizados três métodos estatísticos distintos para identificar os genes diferencialmente expressos em resposta a IC<sub>50</sub> de cetozonazol: edgeR ponto-a-ponto (comparação de cada ponto contra a controle), edgeR extremos (0+24 x 96+120) e SAM com 3 classes (0+24, 48+72, 96+120). A seguir, foram comparados os genes identificados como

significativos em cada um dos métodos e na sobreposição deles, de forma a selecionar o método mais informativo para as análises posteriores.



**FIGURA 5.24 – IC<sub>50</sub> CETOCONAZOL: GRÁFICO DE VENN PARA GENES SIGNIFICATIVOS EM DIFERENTES MÉTODOS ESTATÍSTICOS**

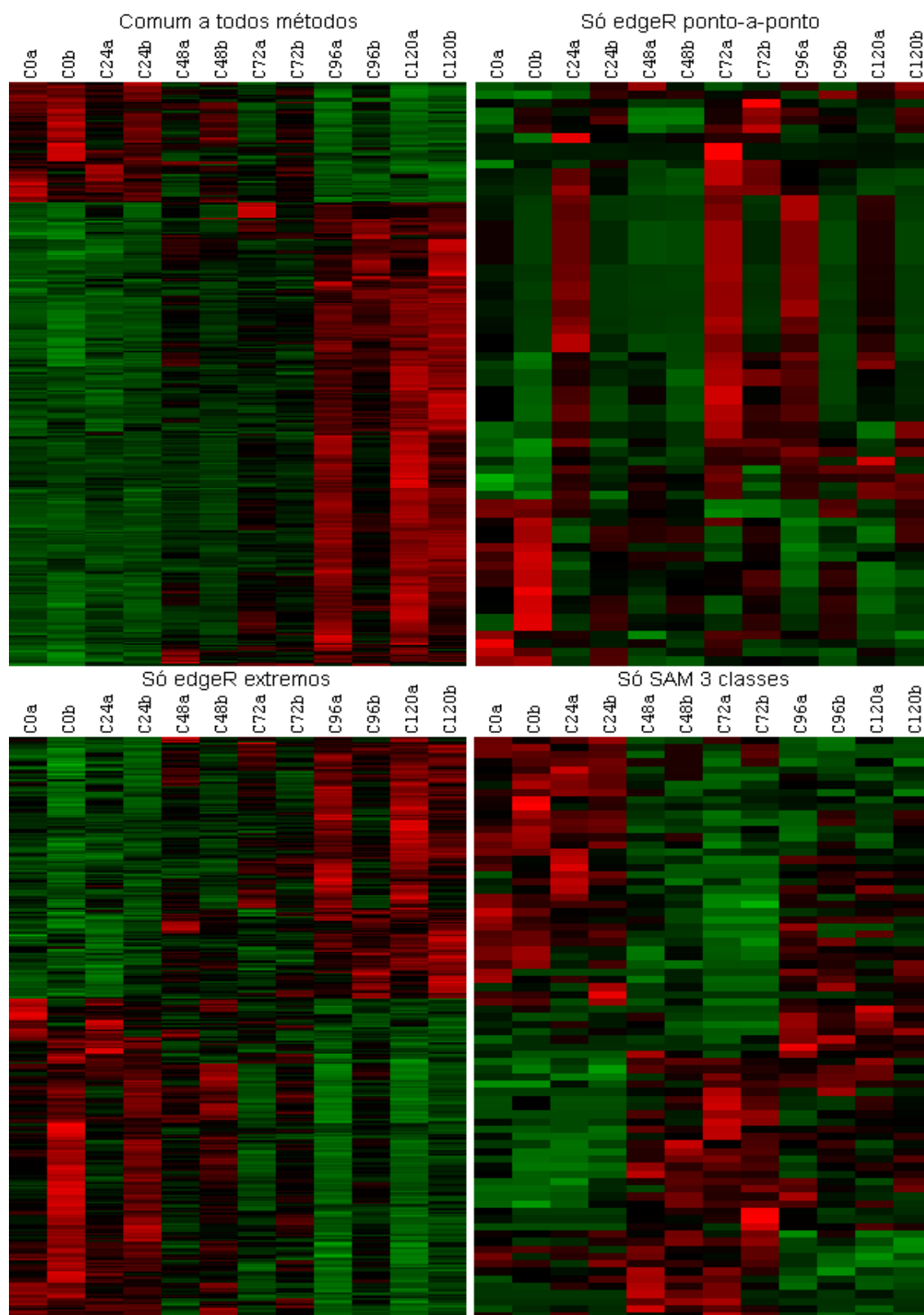
A: só estão plotados os genes com FDR menor ou igual a 5% (a não ser para o teste SAM: 10%), e alteração mínima na expressão de 100% (FC2); B: FDR menor ou igual a 1% (para o teste SAM 5%), e FC1,5. Os números dentro dos gráficos representam o número de genes diferencialmente expressos em cada método e/ou sobreposição dos mesmos.

A visualização da Figura 5.24 deixa claro que o método “edgeR extremos” é o mais abrangente. Cerca de 90% dos genes identificados como diferencialmente expressos pelos métodos “edgeR ponto-a-ponto” e “SAM 3 classes” estão contidos na lista de genes significativos para o método “edgeR extremos”, independentemente do nível estatístico (FDR) e biológico (FC) usado (Figura 5.24). Além disso, o método “edgeR extremos” identifica pelo menos 95% mais genes diferencialmente expressos que os outros dois métodos.

Para visualizar o padrão de expressão dos genes identificados como diferencialmente expressos, foram realizadas clusterizações hierárquicas dos genes identificados pelos 3 métodos em comum, ou somente por um dos métodos (Figura 5.25). Para realizar estas clusterizações, foi utilizado o valor Z de cada gene significativo para todas as réplicas biológicas, com alta estrigência biológica (FC limiar de 2, os mesmos dados da Figura 5.24A). Esse tipo de análise permite visualizar os padrões de expressão dos genes diferencialmente expressos do ponto de vista de significância biológica (replicação dos dados e níveis de alteração na expressão), de forma a contribuir na seleção do melhor método estatístico para as análises posteriores dos dados IC<sub>50</sub>.

Quando clusterizados os genes significativos para os três métodos em comum (416 genes), obteve-se os padrões visualmente mais robustos, com maior correlação entre as réplicas biológicas (Figura 5.25, canto superior esquerdo). Por outro lado, quando considerados somente os genes significativos para o método “edgeR ponto-a-ponto” (67 genes), as réplicas biológicas não são tão consistentes (Figura 5.25, canto superior direito; compare as réplicas biológicas de cada ponto experimental). Tal incoerência também pode ser observada nos genes exclusivos para os métodos “edgeR extremos” (638 genes; Figura 5.25, canto inferior esquerdo) e “SAM 3 classes” (78 genes; Figura 5.25, canto inferior direito), mas em menor intensidade. Embora o método “SAM 3 classes” identifique poucos genes exclusivos, tal método foi o único capaz de identificar padrões robustos de genes diferencialmente expressos nos pontos intermediários do processo (veja Figura 5.25, canto inferior direito).

Com base nos resultados apresentados, parece claro que o método mais abrangente para análise dos genes diferencialmente expressos em resposta a  $IC_{50}$  das drogas seja o “edgeR extremos”, já que engloba quase a totalidade dos genes identificados nos outros dois métodos e acrescenta genes não identificados pelos outros. Como só o método “SAM 3 classes” identifica genes regulados em tempos intermediários do processo, utilizamos o programa edgeR para também analisar os pontos intermediários (48+72 horas) contra os tempos iniciais (0+24 horas) e tardios (96+120) e os pontos intermediários (48+72) contra os pontos extremos (0+24+96+120): chamamos esse método de “edgeR global”. Embora essa análise permita a identificação de alguns genes antes só identificados no método SAM (6 genes), a maioria deles (72 genes) continuam sendo únicos do método SAM (veja Figura 5.24).



**FIGURA 5.25 – IC<sub>50</sub> CETOCONAZOL: CLUSTERIZAÇÃO DE GENES SIGNIFICATIVOS PARA DIFERENTES MÉTODOS ESTATÍSTICOS**

A imagem no canto superior esquerdo representa a clusterização hierárquica dos genes identificados como significativos nos 3 diferentes métodos usados (note a maior replicação biológica desses dados), enquanto as demais imagens referem-se aos genes só identificados como diferencialmente expressos por um dos métodos (especificado acima da figura).

A Tabela 5.5 mostra o número total de genes (IDs) identificados pelo método edgeR global fazendo todas as comparações citadas acima, para diferentes níveis de FDR e FC.

**TABELA 5.5 – IC<sub>50</sub> CETOCONAZOL: N DE GENES SIGNIFICATIVOS PELO MÉTODO EDGER GLOBAL**

FC	FDR 5%	FDR 1%	FDR 0.1%	FDR 0.01%
sem	1834	936	458	259
1,5	1834	936	458	259
2	1255	889	458	259

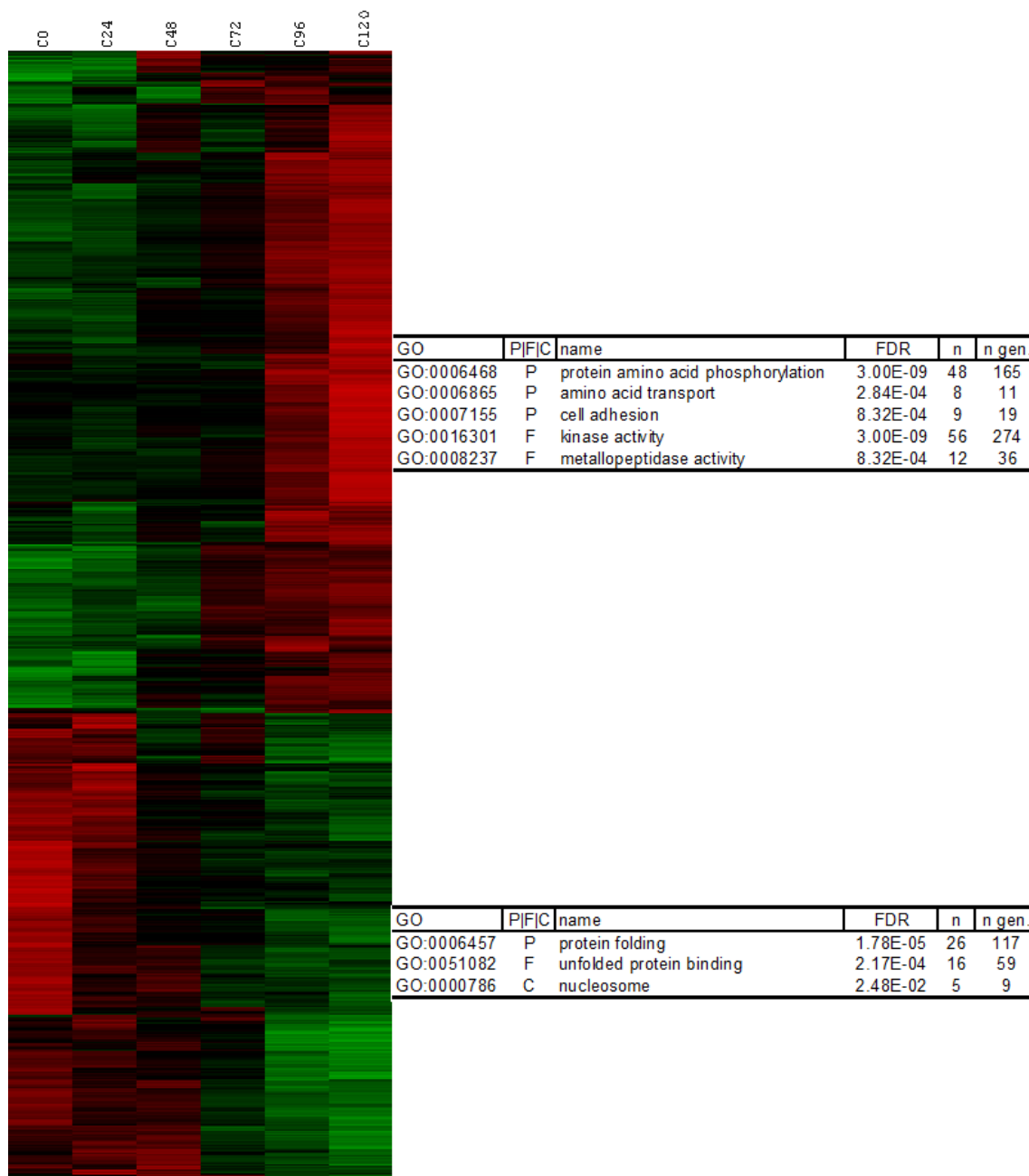
Os números representam o total de genes (IDs) detectados como significativos para diferentes níveis de FDR e FC (*fold change*), considerando todos os pontos de comparação pelo método “edgeR global”.

Como os dados foram trabalhados usando simultaneamente os dois haplótipos de *T. cruzi*, grande parte dos genes selecionados como diferencialmente expressos estão repetidos nos dados, ou seja, há mais de um ID (geralmente 2, um de cada haplótipo), para o mesmo gene. Para retirar tal redundância dos dados de expressão diferencial, foram considerados os “supergenes” de *T. cruzi*, sendo que só foi permitido um ID por supergene nos dados finais (o ID de cada supergene foi escolhido ao acaso).

Além disso, devido a problemas na montagem e anotação do genoma de *T. cruzi*, diversos genes repetitivos, truncados e/ou mal anotados não estão associados a nenhum supergene, sendo, em sua grande maioria, representantes de famílias multigênicas. Como estes genes não estão bem resolvidos no genoma deste parasita, não foram analisados dentro do contexto de expressão diferencial no presente trabalho.

Após exclusão das famílias repetitivas e retirada da redundância dos dados, obteve-se 863 genes significativos (~59% com expressão aumentada), com FDR5% e FC1,5, valor bem menor (cerca da metade) do que quando considerado todos IDs de *T. cruzi*. Se considerada alta estrinência estatística (FDR 1%) e biológica (FC 2), foram identificados 409 genes significativos, sem redundância.

Na Figura 5.26, os 863 supergenes selecionados estão agrupados por clusterização hierárquica. Tal imagem deixa claro que a resposta a IC<sub>50</sub> é lenta, com alteração mais significativa dos níveis de expressão em tempos mais tardios de exposição a droga. Dos 863 supergenes, 507 (59%) possuem um aumento de expressão em resposta ao inibidor, enquanto o restante (356, 41%) tem expressão diminuída.



**FIGURA 5.26 – IC<sub>50</sub> CETOCONAZOL: CLUSTERIZAÇÃO DOS SUPERGENES SIGNIFICATIVOS E ENRIQUECIMENTO DE G.O.**

Estão agrupados os 863 supergenes detectados como diferencialmente expressos na resposta a IC<sub>50</sub> de cetoconazol pelo método edgeR global. O padrão de cores é o mesmo da Figura 5.35. Ao lado da imagem de clusterização são listadas as anotações de G.O. enriquecidas nos genes aumentados (tabela superior) e diminuídos (tabela inferior) em resposta a IC<sub>50</sub> de Cetoconazol, considerando o nível de ontologia (PIFIC: processo celular, função molecular e componente celular, respectivamente) e respectivo nome do termo G.O. (name), além da significância estatística (FDR), o número de supergenes identificados dentro de cada termo G.O. (n) e a quantidade de supergenes anotados com tal termo no genoma como um todo (n gen.).

Na tabela em anexo (Tabela 2, ANEXO II), estão listados os 863 supergenes selecionados pelo método “edgeR global” com limiares FDR5% e FC 2. Chama atenção a grande quantidade de genes de função desconhecida (*hypothetical protein*) detectados como



diferencialmente expressos em resposta a cetoconazol. Dos 863 supergenes selecionados, 490 (~57%) codificam proteínas hipotéticas. Embora a presença de diversos desses genes seja pouco informativa, fornece uma ótima oportunidade para a inferência de função a partir de dados de co-regulação.

De forma a buscar mecanismos celulares significativamente alterados em resposta às drogas, foram utilizadas análises de enriquecimento de termos de ontologia gênica (GO). O Consórcio de Ontologia Gênica, *Gene Ontology Consortium* (ASNURNER *et al.*, 2000), tem como objetivo produzir um vocabulário dinâmico, controlado e precisamente definido para descrever as funções dos genes em qualquer organismo. Nesse contexto, cada gene é nomeado em três níveis de ontologia; (i) **processo biológico**: refere-se a um objetivo biológico ao qual o gene, ou produto gênico, contribui. Um processo é realizado por um ou mais conjuntos ordenados de função moleculares. Muitas vezes envolvem processos químicos ou físicos de transformação, no sentido que algo entra em um processo e algo diferente é produzido. Tais processos biológicos podem ser bem amplos, como “crescimento celular e manutenção”, ou mais específicos, como “tradução” ou “metabolismo de pirimidinas”. (ii) **função molecular**: atividade bioquímica do produto gênico (incluindo ligação específica a estruturas ou ligantes). Essa definição só descreve o que é feito sem levar em conta onde ou quando o evento ocorre. Também pode ter definições amplas, como “enzima” ou “ligante”, até definições mais específicas, como “adenilato ciclase” ou “ligante de receptor Toll”; (iii) **componente celular**: refere-se ao local na célula onde o produto gênico é ativo. Inclui termos como “ribossomo”, “proteossomo”, onde múltiplos produtos gênicos serão encontrados; ou termos mais gerais como “membrana nuclear” ou “aparato de Golgi”.

As anotações de GO são trabalhosas e focadas em organismos modelo; assim sendo, *T. cruzi* não as possui. Entretanto, uma das vantagens das ontologias gênicas é a possibilidade de transferência automática de anotação de GO através de similaridade gênica ou protéica. Neste trabalho, utilizou-se a ferramenta Blast2GO (CONESA *et al.*, 2005) para realizar uma anotação de GO de *T. cruzi* baseado em similaridade gênica ou protéica. Uma vez obtidas as anotações de GO, foram realizadas análises estatísticas de enriquecimento de termos de GO nos grupos de genes de interesse (genes diferencialmente expressos) em comparação com o grupo referência (genoma completo), dentro da plataforma Blast2GO.

As análises de enriquecimento de GO foram realizadas tanto na lista geral de genes diferencialmente expressos, bem como considerando somente os genes com sentido de alteração na expressão, aumentada ou diminuição, similar. Obtivemos resultados mais

significativos e específicos quando considerado separadamente os genes de expressão aumentada ou diminuída. Além disso, destacamos os processos biológicos enriquecidos, já que este nível mais geral de ontologia costuma demonstrar os resultados mais conclusivos para análises de enriquecimento em genes agrupados por padrões similares de expressão gênica (ASHBURNER *et al.*, 2000).

Após as análises de enriquecimento, diversos termos de GO redundantes eram listados nos resultados (Ex.: *quinase activity*, *protein quinase activity*, *post-translational protein modification*, *phosphotransferase activity*, *phosphorylation*), sendo que os supergenes identificados para um determinado termos de GO normalmente estavam contidos dentro do grupo de supergenes associados a um GO similar. Desta forma, a listagem de GOs enriquecidos foi filtrada manualmente de forma a se obter os termos mais específicos e que englobassem o maior número de genes relacionados a certo processo redundante.

Considerando apenas os genes de expressão aumentada em resposta a cetoconazol, foram identificados 3 processos celulares (sem redundância) enriquecidos (veja Figura 5.26): fosforilação protéica de aminoácidos, transporte de aminoácidos e adesão celular.

Ao termo “fosforilação protéica de aminoácidos” foram associados 48 supergenes; destes, 46 possuem o termo “quinase” na descrição do gene, enquanto os outros dois são proteínas hipotéticas. A presença de proteínas hipotéticas associadas a termos definidos de GO chama atenção para a importância da inferência cruzada de anotação de GO através da análise de similaridade em banco de dados de proteína ou domínios protéicos, o que favorece as próprias análises de enriquecimento e permite a inferência funcional para proteínas de papel desconhecido. A validade do enriquecimento do processo “fosforilação protéica de aminoácidos” foi apoiada pelo enriquecimento, em nível de ontologia função molecular, do termo “atividade quinase” (veja Figura 5.26), ao qual foram associadas os mesmos 48 supergenes do processo “fosforilação protéica de aminoácidos”, além de outras oito quinases.

O segundo processo celular enriquecido foi “transporte de aminoácidos” (veja Figura 5.26). Os oito supergenes associados a tal processo estão descritos no genoma de *T. cruzi* como “*aminoacid transporter*” (4 deles), “*aminoacid permease*” (3 deles) ou proteína hipotética (um).

O terceiro e último processo celular identificado como enriquecido na resposta a cetoconazol foi “adesão celular”, contendo nove supergenes associados a tal termo. Um dado interessante é que oito destes estão descritos como proteases de superfície da família GP63, e um como proteína hipotética. Este enriquecimento é apoiado, em nível de função molecular,

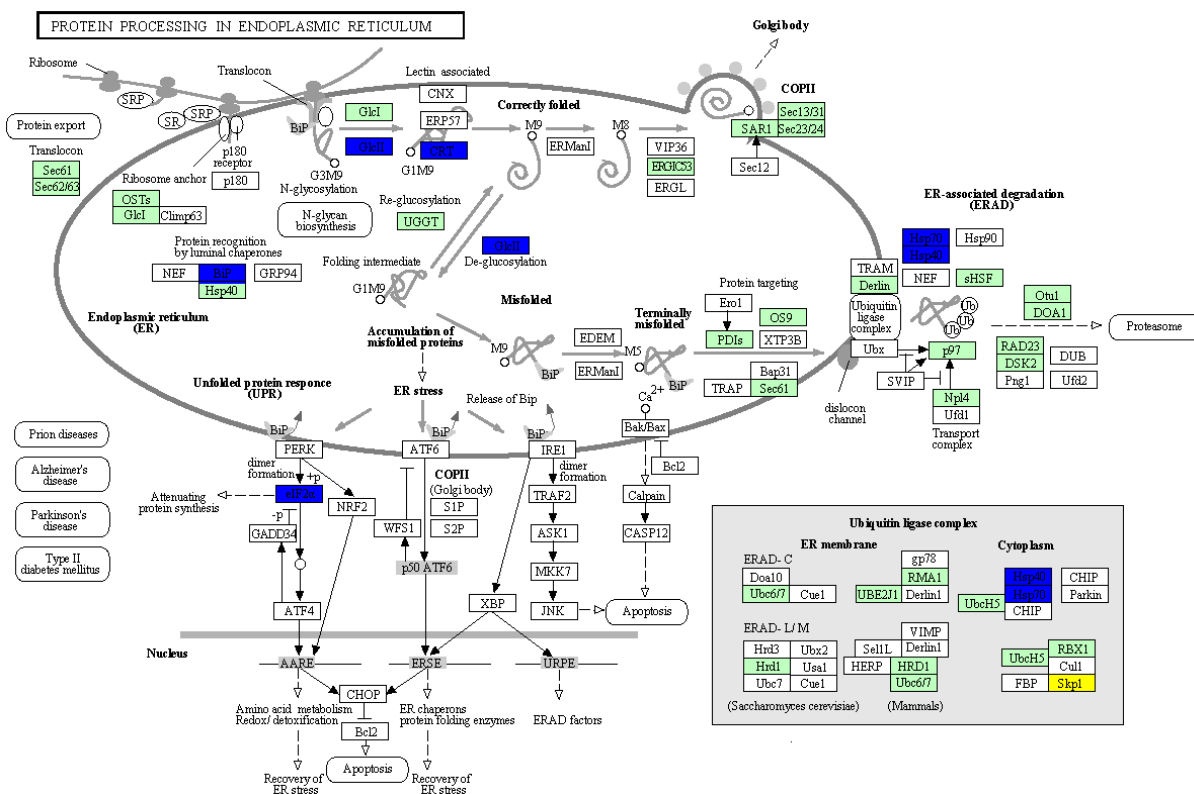
pelo termo “atividade metalopeptidase”. As glicoproteínas de superfície GP63 formam a sexta maior superfamília gênica em *T. cruzi* (425 membros) (EL-SAYED *et al.*, 2005), com importância na infecção à célula hospedeira (KULKARNI *et al.*, 2009), o que parece estar associado a presença de atividade metaloprotease (CUEVAS *et al.*, 2003). Portanto, a identificação do enriquecimento dos termos “adesão celular” e “atividade metalopeptidase” seria explicado pela presença das GP63. A presença de membros da família gênica GP63 dentro das anotações de supergene e nas listagem de genes diferencialmente expressos, chama atenção para a presença de genes mais bem resolvidos para esta família gênica de *T. cruzi*.

Para os genes de expressão diminuída em relação ao controle, os diversos processos celulares redundantes identificados como enriquecidos podem ser representados, de forma geral, pelo processo “dobramento protéico”, ao qual 29 supergenes foram associados (Figura 5.26). Destes genes, 9 estão anotados como chaperonas, 6 como proteínas de choque térmico, e as 15 demais como proteínas acessórias ou relacionadas aos mecanismos de dobramento protéico (como 4 peptilprolil isomerases). O enriquecimento do processo celular “dobramento protéico”, dentre os supergenes com expressão diminuída em relação com controle, é apoiado pelo enriquecimento da função molecular “ligação a proteína mal-dobrada”, a qual estão associados 16 genes, todos também presentes no processo “dobramento protéico”. O processo celular “resposta a estímulo” (GO:0050896) também está enriquecido dentre os genes diminuídos em resposta a cetoconazol; além de diversos genes redundantes ao processo “dobramento protéico”, destaca-se a presença de uma proteína conjugadora de ubiquitina e dois genes de histona (H3 e H3 variante). A presença de diversos genes de histonas com expressão diminuída fica mais evidente no enriquecimento do componente celular “nucleossomo”, associado a 5 supergenes: duas histonas H2A, H3, H3 variante e H2B.

Quando os genes diferencialmente expressos em resposta a IC<sub>50</sub> de Cetoconazol são analisados dentro do banco de dados de vias do KEGG, observa-se um enriquecimento de genes de expressão diminuída na via de síntese de “Processamento protéico no retículo endoplasmático” (veja Figura 5.27), compatível com o resultado da análise dos dados de GO (“dobramento protéico”).

Embora as análises de enriquecimento de termos de GO indiquem grupos de genes enriquecidos com alta significância estatística, diversas respostas comuns podem passar despercebidas devido a problemas nas anotações de GO, o que é ainda mais evidente em organismos menos estudados para os quais as anotações são realizadas por similaridade

protéica, o que pode resultar em perda de significância para respostas específicas de determinados organismos. Isso fica evidente pelo grande número de genes detectados como



**FIGURA 5.27 – VIA DE PROCESSAMENTO PROTÉICO NO R.E. ENRIQUECIDA COM GENES DE EXPRESSÃO DIMINUÍDA PARA IC<sub>50</sub> CETOCONAZOL**

É demonstrado um exemplo de via do KEGG enriquecidas com genes diferencialmente expressos em resposta a IC<sub>50</sub> de Cetoconazol e corroborada por dados de enriquecimento de G.O. Cada retângulo representa uma enzima que catalisa o passo especificado na imagem. Os retângulos coloridos em verde indicam genes identificados no genoma de *T. cruzi* mas não detectados como diferencialmente expressos nos dados de RNA-Seq, enquanto retângulos amarelos e azuis indicam genes de expressão aumentada e diminuída, respectivamente, em relação ao controle.

diferencialmente expressos mas não associados a nenhum dos termos enriquecidos de GO (~90%). Portanto, além das análises de GO, buscou-se por genes com alta representatividade dentre as listagens de supergenes diferencialmente expressos, simplesmente pela busca de anotações específicas (Ex.: *RNA-binding protein*, *phosphatase*, *helicase*, etc).

Foi identificada a presença de diversos supergenes anotados como proteínas ligadoras de RNA, 10 de expressão aumentada e 5 de expressão diminuída em relação ao controle. Contando com outras duas proteínas anotadas como “dedos de zinco” (*zinc finger*), ambas aumentadas, foram identificados 17 supergenes de proteínas ligadoras de RNA, o que chama atenção para a importância dessas proteínas na regulação pós-transcricional da expressão gênica em *T. cruzi* e possível papel das mesmas na resposta a inibidores da síntese de esteróis.

Considerando a anotação “RNA-binding protein”, 1,5% dos supergenes da lista de diferencialmente expressos para IC<sub>50</sub> de Cetoconazol são identificados, valor este que cai para 0,52% quando considerado todo genoma (todos supergenes): ou seja, a proporção de proteínas ligadoras de RNA entre os genes diferencialmente expressos é 3 vezes maior que a proporção desse tipo de gene no genoma. Foram identificados 4 supergenes diferencialmente expressos anotados como helicases de RNA (3 aumentados), sendo 3 deles da família DEAD.FH. Estas proteínas possuem importante papel no metabolismo de RNA, como splicing, transporte, biogênese do ribossomo, tradução e decaimento de RNA (GUSTAFSON & WESSEL, 2010). É possível que algumas dessas proteínas atuem em conjunto com proteínas ligadoras de RNA em complexos ribonucleoprotéicos relacionados ao controle da expressão gênica.

Além dos transportadores de aminoácidos identificados nos dados de enriquecimento de GO, destaca-se a presença de 4 supergenes anotados como transportadores ABC e diferencialmente aumentados em resposta a IC<sub>50</sub> de cetoconazol. Comparando com o genoma, a lista de supergenes diferencialmente expressos possui 2 vezes mais genes ABC que o genoma. Estas proteínas transmembrana transportam uma grande variedade de substratos através da membrana extra ou intra-celular, incluindo produtos metabólicos, lipídios (inclusive esteróis) e drogas. A ativação de bombas que auxiliam no efluxo de azóis de células de fungos foi demonstrado para cetoconazol, voriconazol e fluconazol (revisto por CANNON *et al.*, 2009). Além disso, a indução de resistência em laboratório a fluconazol foi demonstrada para a cepa Tulahuen de *T. cruzi* (BUCKNER *et al.*, 1998): quando crescido em baixas concentrações de fluconazol, os parasitas desenvolvem resistência cruzada para vários azóis, enquanto se mantêm suscetíveis ao tratamento com benzonidazol.

A sequência protéica de genes conhecido relacionados a extrusão de azóis em fungos (revisto por CANNON *et al.*, 2009) foram alinhados ao genoma de *T. cruzi* através da ferramenta BLASTp. Dentre os genes identificados, encontramos um dos transportadores ABC de expressão aumentada a IC<sub>50</sub> de cetoconazol (Tc00.1047053508231.190, *ABC transporter*, mais de duas vezes de aumento;  $\text{value} = 6e^{-48}$  para gene Cdr1 de *Candida albicans*). Embora não possamos afirmar ao certo se estes genes representam ortólogos, esse resultado aponta para o possível papel desse transportador em mecanismos de resistência a azóis em *T. cruzi*, de forma que tal proteína pode representar uma alvo quimioterápico interessante no tratamento da doença de Chagas. É provável que os demais transportadores ABC de expressão aumentada em resposta a cetoconazol também estejam relacionados a extrusão desse azol do parasita.

A relação destes transportadores com a extrusão de cetoconazol da célula parece provável, visto que tais supergenes não são aumentados em resposta a lovastatina.

Diversos outros tipos de transportadores foram identificados como aumentados em resposta a Cetoconazol: 12 transportadores para aminoácidos (*amino acid transporter* ou *amino acid permase*), um para metal (metal-ion transporte), 2 para açúcar (*sugar transporte* e *hexose transporter*) e 2 para nucleosídeos.

Também destaca-se a presença de oito supergenes de expressão aumentada anotados como cisteína-peptidase. Juntamente com uma proteína lisossomal e uma próton-ATPase vacuolar, estes dados indicam uma possível correlação da expressão gênica com o aumento dos reservossomos observado por microscopia eletrônica (Figura 5.11) e de fluorescência (Figura 5.9), visto que tal organela é rica em proteases do tipo cisteína-peptidase.

#### 5.3.4.2 IC<sub>50</sub> Lovastatina

A resposta de *T. cruzi* a lovastatina foi analisada da mesma forma que para cetoconazol. A Tabela 5.6 mostra o número de genes significativos em cada ponto de comparação com o controle. Novamente o número de genes detectados não foi muito alto, principalmente quando considerado valores mais estridentes de FDR (205 genes com FDR 0,1%; número similar a cetoconazol).

Novamente, o número de genes diferencialmente expressos aumenta ao decorrer do tratamento com a droga (veja Tabela 5.6 e Figura 5.28), o que condiz com uma resposta lenta aos inibidores da síntese de esteróis quando usada a dose IC<sub>50</sub>.

A grande maioria dos genes significativos tem expressão diminuída em relação ao controle (no máximo 20% dos genes significativos tem expressão aumentada), a não ser no ponto 120 horas, o qual apresenta cerca da metade dos genes significativos como diferencialmente aumentados (veja Tabela 3 do ANEXO II).

TABELA 5.6 – IC<sub>50</sub> LOVASTATINA: GENES SIGNIFICATIVOS

sem critério FC				
Comparação	FDR 5%	FDR 1%	FDR 0.1%	FDR 0.01%
L24xL0	56 (7, 49)	23 (3, 20)	5 (0, 5)	3 (0, 3)
L48xL0	164 (31, 133)	69 (6, 63)	28 (1, 27)	16 (0, 16)
L72xL0	403 (143, 260)	153 (40, 113)	68 (12, 56)	32 (3, 29)
L96xL0	812 (263, 549)	339 (85, 254)	117 (23, 94)	53 (9, 44)
L120xL0	681 (334, 347)	347 (184, 163)	131 (69, 62)	58 (28, 30)
Total	1185	556	205	99
log <sub>2</sub> (FC) 1,5				
L24xL0	45 (4, 41)	23 (3, 20)	5 (0, 5)	3 (0, 3)
L48xL0	102 (16, 86)	60 (6, 54)	28 (1, 27)	16 (0, 16)
L72xL0	168 (48, 120)	111 (31, 80)	68 (12, 56)	32 (3, 29)
L96xL0	307 (75, 232)	209 (41, 168)	117 (23, 94)	53 (9, 44)
L120xL0	276 (145, 131)	219 (115, 104)	131 (69, 62)	58 (28, 30)

Os dados são apresentado da mesma forma que na Tabela 5.2.

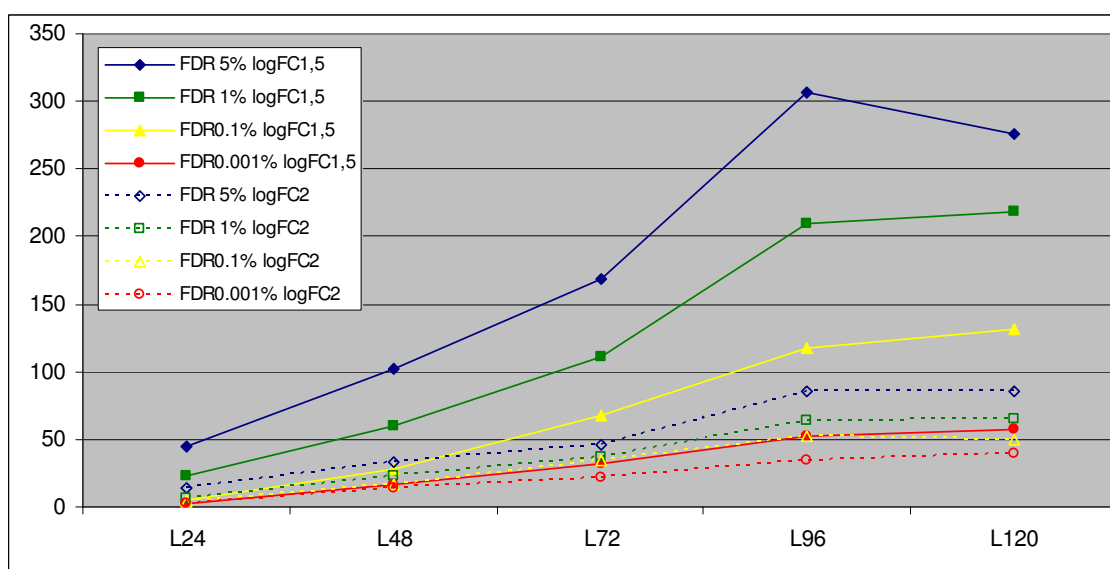
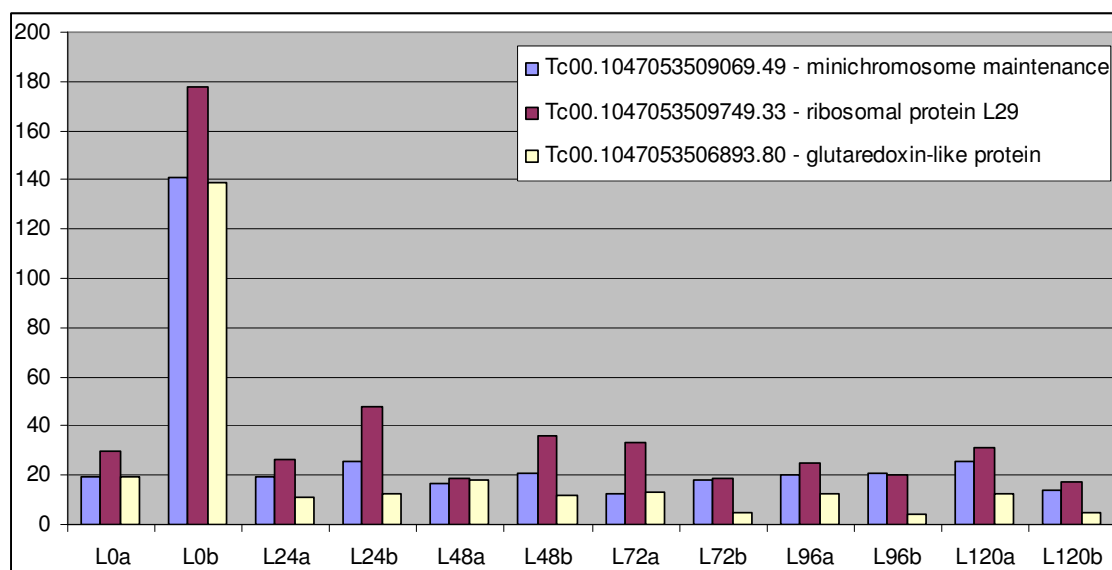


FIGURA 5.28 – IC<sub>50</sub> LOVASTATINA: N DE GENES SIGNIFICATIVOS EM CADA PONTO

Cada linha representa o número de genes detectados como diferencialmente expressos nos diversos tempos de comparação com o controle para um determinado FDR e logFC limite. No eixo x está plotado cada um dos pontos de comparação com o controle (zero horas), enquanto o eixo y representa o número de genes detectado como diferencialmente expressos. As linhas cheias representam os mesmos dados da parte inferior da Tabela 5.2, enquanto as linhas pontilhadas são os dados para logFC maior ou igual a 2 (4 vezes de alteração mínima na expressão).

A Figura 5.29 mostra claramente que esse grande número de genes com expressão significativamente diminuída nos testes estatísticos deve-se a problemas na réplica L0b (ponto zero da réplica biológica b). Para diversos genes, esta amostra possui contagens muito superiores em relação a todos os outros pontos de análise, de forma que os pontos tratados são detectados como tendo expressão diminuída. Podemos observar na Figura 5.29 que os genes plotados têm expressão constante ao longo do processo, mas devido ao problema já citado com a amostra L0b são marcados como diminuídos em relação ao controle.



**FIGURA 5.29 – EXEMPLOS DE GENES SIGNIFICATIVOS PARA IC<sub>50</sub> LOVASTATINA**

Está plotado um gráfico de barras com 3 genes (cada um representado por uma cor) detectados como diferencialmente expressos em resposta a IC<sub>50</sub> de lovastatina. No eixo x são nomeados cada um dos pontos de análise; L (lovastatina), seguido do tempo de exposição em horas (0-120 horas) e a réplica biológica (a e b). Para cada gene, está plotado o nível de expressão normalizado pelo tamanho da biblioteca e tamanho do gene segundo método RPKM.

Por estes problemas com os dados IC<sub>50</sub> lovastatina, a utilização de outros métodos estatísticos para identificação de genes diferencialmente expressos é ainda mais evidente. De forma similar às análises realizadas com os dados do tratamento com cetoconazol, realizamos as análises estatísticas com os outros dois métodos já discutidos (“edgeR extremos”, veja Tabela 5.7, e “SAM 3 classes”, veja Tabela 5.8) e comparamos os genes significativos identificados em cada um dos métodos. Diferentemente dos resultados para IC<sub>50</sub> de cetoconazol (Figura 5.24), o método que identificou maior número de genes diferencialmente expressos, independentemente do limiar de FDR e FC usado, foi “edgeR ponto-a-ponto” (veja figura Figura 5.30). Entretanto, análises de clusterização hierárquica dos genes detectados somente neste método (Figura 5.31, canto superior direito) evidenciam que tais genes têm um comportamento similar ao já demonstrado na Figura 5.29, ou seja: a grande maioria tem expressão “diminuída” em relação ao controle devido à alta contagem na amostra controle da réplica B (veja alta intensidade em vermelho da amostra L0b na Figura 5.31 para dados edgeR ponto-a-ponto). Assim sendo, parece claro que a grande maioria dos genes detectados como diminuídos somente pelo método “edgeR ponto-a-ponto” devem ser artefatos devido ao baixo número de réplicas biológicas utilizadas, aliado à alta variação intrínseca das amostras, e à comparação de cada ponto com o controle.



**TABELA 5.7 – IC<sub>50</sub> LOVASTATINA: GENES SIGNIFICATIVOS PARA 0+24 VERSUS 96+120**

FC	Comparação	FDR 5%	FDR 1%	FDR 0.1%	FDR 0.01%
nenhum	0+24x96+120	872 (421, 451)	384 (199, 185)	124 (56, 68)	58 (22, 36)
nenhum	Total ponto a ponto	1185	556	205	99
1,5	0+24x96+120	872 (421, 451)	384 (199, 185)	124 (56, 68)	58 (22, 36)
2	0+24x96+120	462 (243, 219)	339 (174, 165)	124 (56, 68)	58 (22, 36)
2.83	0+24x96+120	61 (40, 21)	42 (24, 18)	32 (15, 17)	30 (14, 16)

A primeira coluna especifica o tipo de limiar de *fold change* utilizado. Os testes foram feitos no programa edgeR utilizando as amostra controle (zero hora) e 24 horas de exposição como réplicas do início do processo de tratamento e as amostras 96 e 120 horas como réplicas do final do tratamento. A linha “Total ponto a ponto” diz respeito ao número total de genes detectados como significativos (sem redundância) para as comparações individuais contra o controle, como mostra a Tabela 5.6.

Quando considerado somente os genes detectados pelos três métodos (247 genes, mesmos dados da Figura 5.30A), o padrão de expressão é bem mais homogêneo (Figura 5.31, canto superior esquerdo), de forma similar ao observado para os dados de cetoconazol.

Embora os métodos “edgeR extremos” e “SAM 3 classes” identifiquem um menor número de genes significativos exclusivos (71 e 97, respectivamente; veja Figura 5.30) em comparação com “edgeR ponto-a-ponto”, os padrões de expressão destes genes são mais robustos, com menor influência do problema com a amostra L0b (veja parte inferior da Figura 5.31) e, portanto, aparentemente mais confiáveis. Os dados “edgeR extremos” contêm cerca de 70% dos genes detectados pelo método SAM, mas, por outro lado, possuem centenas de genes não detectados pelo último método.

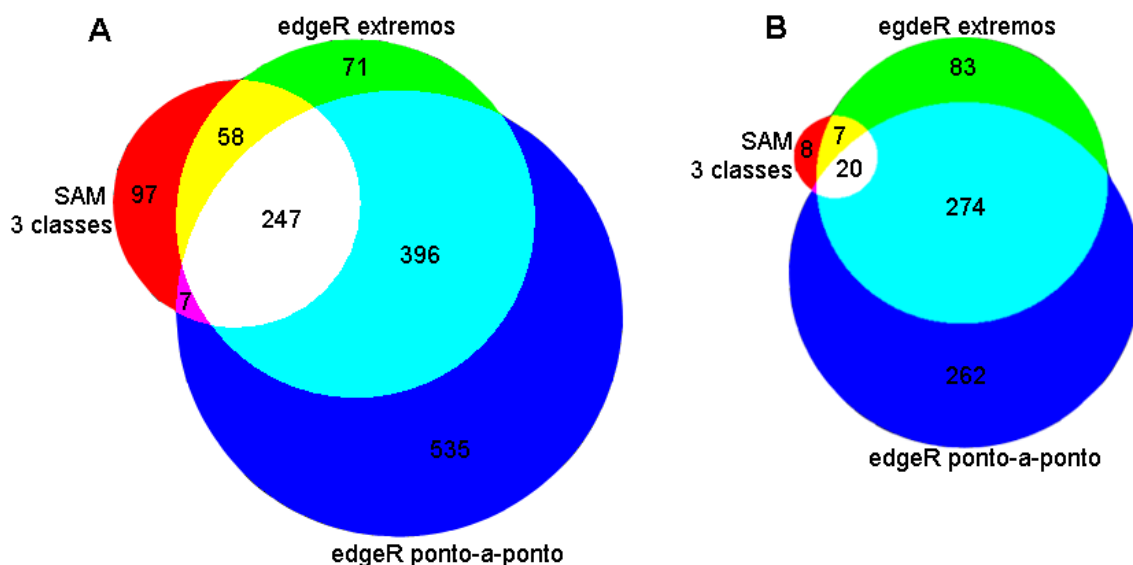
**TABELA 5.8 – IC<sub>50</sub> LOVASTATINA: GENES SIGNIFICATIVOS TESTE SAM 3 CLASSES**

FC	Comparação	FDR10%	FDR5%	FDR1%
sem FC	SAM: 3 classes	477	35	13
FC1,5	SAM: 3 classes	477	35	13
FC2	SAM: 3 classes	409	35	13
FC3	SAM: 3 classes	176	21	11
FC4	SAM: 3 classes	92	13	10

As colunas trazem o total de genes detectados como significativos com diferentes níveis de significância estatística (FDR). A coluna FC (*foldchange*) refere-se a diferentes níveis de significância biológica na alteração da expressão gênica (pelo menos um ponto em relação ao controle); FC2 corresponde a um mínimo de 100% de alteração nos níveis de expressão, e assim sucessivamente. As comparações foram feitas no programa SAM, usando 3 classes (detalhes no texto).

Da mesma forma que para os dados IC<sub>50</sub> cetoconazol, o método mais abrangente e robusto para análise dos dados IC<sub>50</sub> lovastatina parece ser o “edgeR extremos”. De forma a captar uma gama maior de alterações, como aquelas detectáveis somente pelo método “SAM 3 classes”, também foram realizadas análises estatísticas comparando os pontos intermediários (48+72) contra os pontos iniciais (0+24) e finais (96+120), além da

comparação dos pontos intermediários (48+72) contra os extremos (0+24+96+120), método “edgeR global”. De forma similar ao ocorrido para os dados IC<sub>50</sub> cetoconazol, essas análises permitem a identificação de poucos genes antes únicos do método SAM (5 de 97 genes, veja Figura 5.30). Ou seja, alguns poucos padrões de expressão diferencial são apenas identificados pelo método SAM, mostrando a forma distinta como este método analisa os dados.



**FIGURA 5.30 – IC<sub>50</sub> LOVASTATINA: GRÁFICO DE VENN PARA GENES SIGNIFICATIVOS EM DIFERENTES MÉTODOS ESTATÍSTICOS**

A: só estão plotados os genes com FDR menor ou igual a 5% (a não ser para o teste SAM: 10%), e alteração mínima na expressão de 100% (FC2); B: FDR menor ou igual a 1% (para o teste SAM 5%), e FC1,5. Os números dentro dos gráficos representam o número de genes diferencialmente expressos em cada método e/ou sobreposição dos mesmos.

O número final de genes (IDs) considerados diferencialmente expressos em todos os pontos de comparação pelo método edgeR global é demonstrado na Tabela 5.9.

**TABELA 5.9 – IC<sub>50</sub> LOVASTATINA: N DE GENES SIGNIFICATIVOS PELO MÉTODO EDGER GLOBAL**

FC	FDR 5%	FDR 1%	FDR 0.1%	FDR 0.01%
sem	964	421	146	70
1,5	964	421	146	70
2	538	379	146	70

Os números representam o total de genes (IDs) detectados como significativos para diferentes níveis de FDR e FC (*fold change*), considerando todos os pontos de comparação pelo método “edgeR global”.

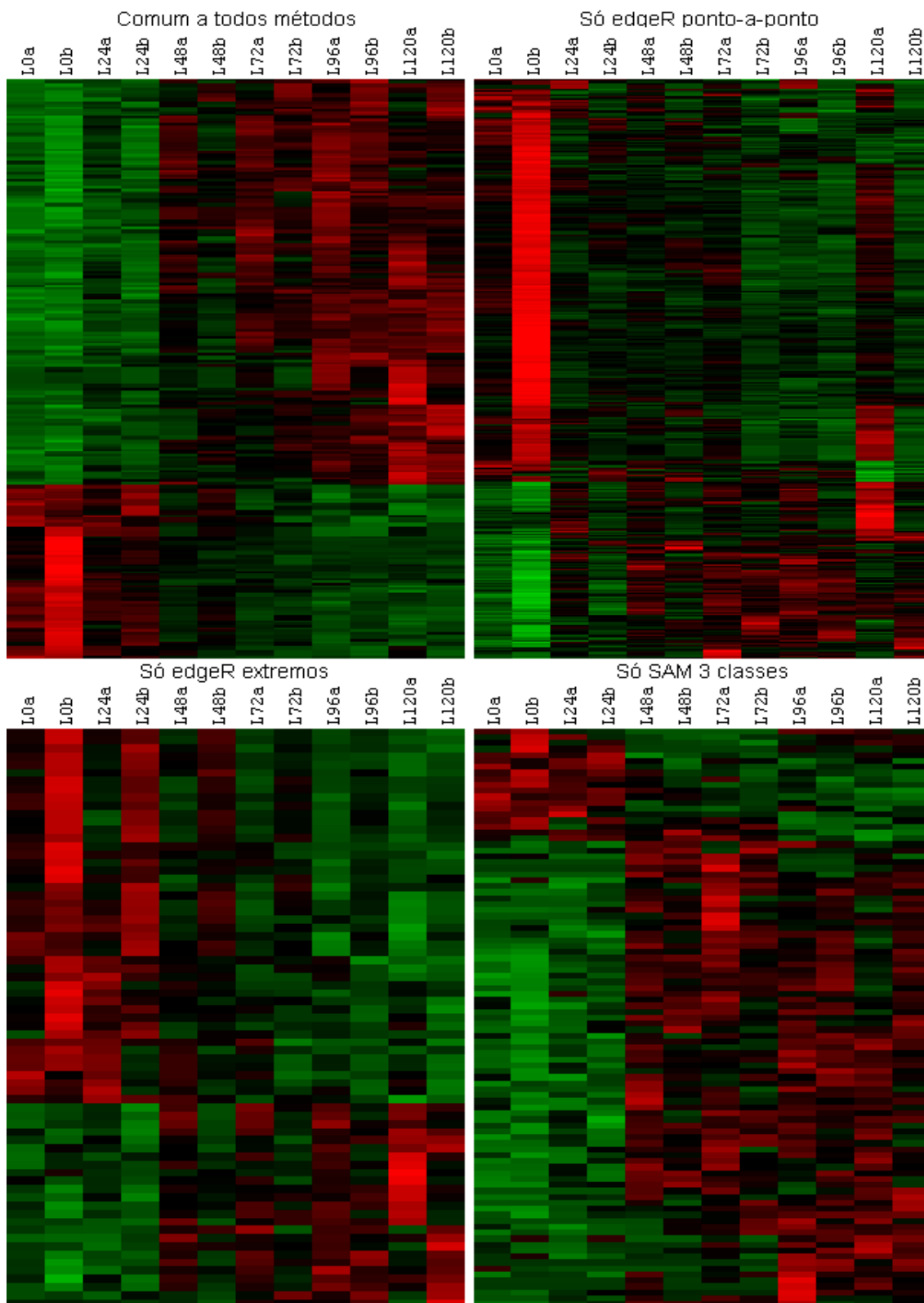
Quando considerada a redundância dos dados, como anteriormente discutido, os 964 IDs diferencialmente expressos com FDR5% e FC1,5 representam 462 supergenes.

Diferentemente da resposta para cetoconazol, a maioria dos genes detectados na resposta a lovastatina possuem expressão diminuída (281, 61%). Esse resultado parece estranho, visto que nesse tipo de resposta seria esperado uma indução de expressão de genes relacionados a resposta a droga. É possível que este baixo número de genes aumentados (181, 39%) deva-se ao problema de alta amostragem na amostra controle “L0b”, como demonstrado na figura Figura 5.29. Quando agrupados por clusterização hierárquica (Figura 5.321), fica evidente a influência deste problema, visto que muitos genes de expressão diminuída demonstram forte diminuição já no ponto 24 horas de exposição, e, a seguir, mantêm-se relativamente estáveis até 120 horas (veja região central da Figura 5.2932).

Novamente, as proteínas hipotéticas representam mais da metade dos supergenes diferencialmente expressos 239 (~52%), destacando o pouco conhecimento funcional acerca dos genes de *T. cruzi* e reforçando a necessidade da utilização de métodos em larga escala para inferir funções a tais proteínas, pois o desconhecimento de suas funções limita muito a aplicação de trabalhos como o presente projeto.

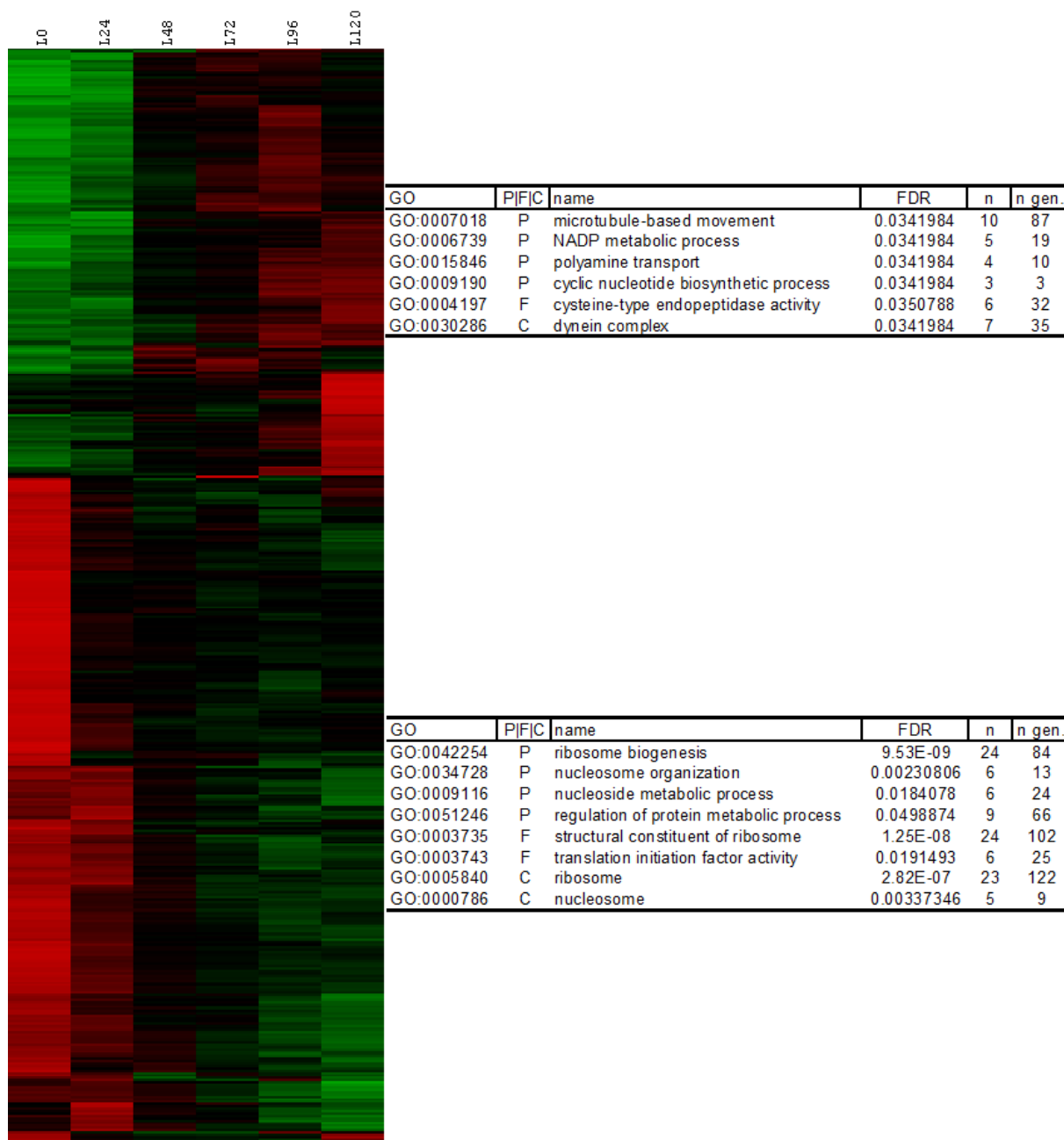
Como previamente descrito, seguiu-se a análise de enriquecimento de termos de GO. Para os 181 supergenes de expressão aumentada, foram identificados 4 processos celulares enriquecidos (veja figura Figura 5.32): “movimento baseado em microtúbulos” (10 genes), “processo metabólico de NADP” (5 genes), “transporte de poliaminas” (4 genes) e “processo biológico de nucleotídeos cíclicos” (3 genes).

Ao processo “movimento baseado em microtúbulos” foram associados 7 supergenes anotados como dineínas e 3 como cinesinas. Além destes, mais uma dineína foi identificada pela análise manual das anotações dos genes significativos. Os dados de genes de expressão aumentada em relação ao controle possuem quase 4 vezes (3,79) mais genes anotados como dineínas em relação ao controle (2,16% versus 0,57%). Considerando apenas as cinesinas, o percentual de genes diferencialmente expressos assim anotado é similar à quantidade relativa no genoma como um todo (0,65% versus 0,59%). O processo celular “movimento baseado em microtúbulo” é corroborado pela presença do enriquecimento do componente celular “complexo de dineína”, associado a 7 supergenes redundantes ao processo citado acima.



**FIGURA 5.31 – IC<sub>50</sub> LOVASTATINA: CLUSTERIZAÇÃO DE GENES SIGNIFICATIVOS PARA DIFERENTES MÉTODOS ESTATÍSTICOS**

Para esta análise foram usados os 1185 genes significativos (FDR5%, FC1,5) na soma dos cinco pontos de comparação com o controle. Em cada uma das amostras, está plotado o valor Z (contagem-contagem média/desvio-padrão) para cada gene. Note o alto valor de Z somente na amostra L0b para os genes detectados como diminuídos (parte superior da figura); para alguns desses genes o nível de expressão está acima da média (principalmente na amostra L120a), evidenciando que tais genes não devem ter significância biológica.



**FIGURA 5.32 – IC<sub>50</sub> LOVASTATINA: CLUSTERIZAÇÃO HIERÁRQUICA DOS SUPERGENES SIGNIFICATIVOS E ENRIQUECIMENTO DE G.O.**

Estão agrupados os 462 supergenes detectados como diferencialmente expressos na resposta a IC<sub>50</sub> de lovastatina pelo método edgeR global. Os dados de enriquecimento de G.O. são apresentados da mesma forma que na Figura 5.26.

Outro processo celular enriquecido nos genes de expressão aumentada foi “processo metabólico de NADP”, associado a 5 supergenes. Dois destes representam enzimas da via glicolítica (6-fosfo-1-frutoquinase e glucose-6-fosfato isomerase), outros dois (6-fosfogluconate desidrogenase, descarboxilase, e glucose-6-fosfato desidrogenase) pertencem à via pentose fosfato e um pertence à família de isomerasas, mais especificamente uma

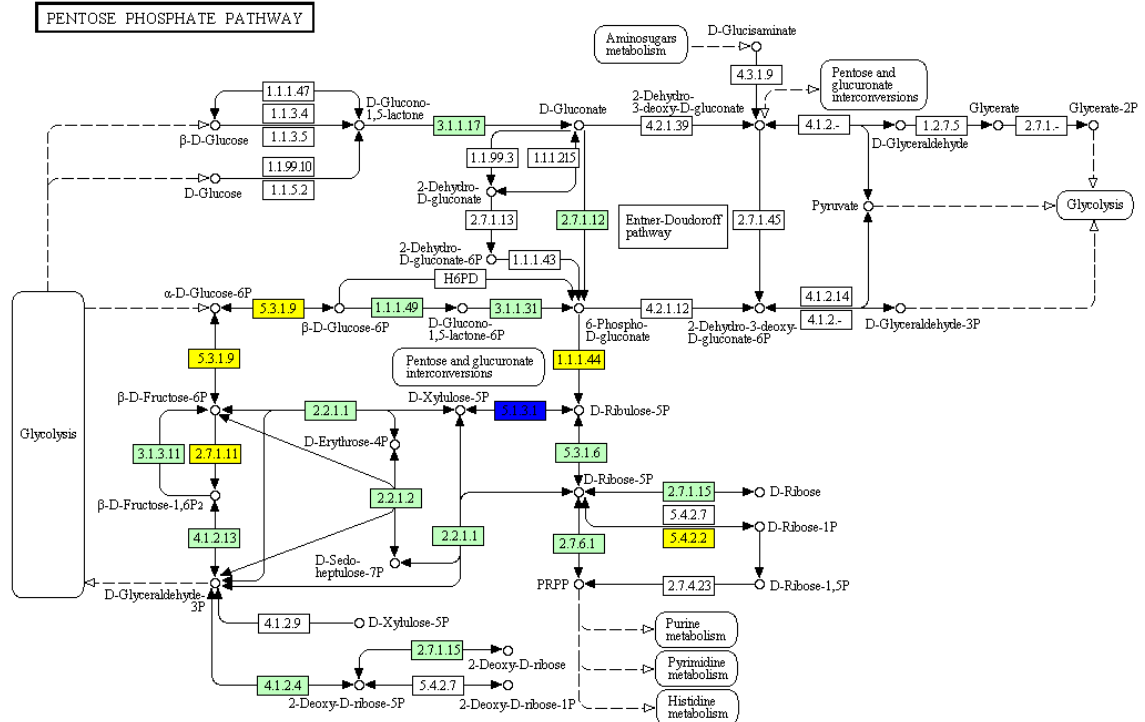
fosfomanomutase (phosphomannomutase-like protein, putative). O enriquecimento desse termo de GO é apoiado pela presença de 4 genes de expressão aumentada dentro da via “Pentose-fosfato” do KEGG (veja Figura 5.33A). Essa via, também conhecida como via do fosfogliconato, usa glicose como fonte energética para gerar NADPH e ribose-5-fosfato (NELSON & COX, 2002).

Para o processo “transporte de poliaminas”, foram associados 4 supergenes, todos anotados como transportador de aminoácidos (3 supergenes) ou permease de aminoácido (1 supergene). Além desses 4 supergenes identificados por GO, 3 outras permeases de aminoácido foram identificados como diferencialmente aumentados em resposta à lovastatina, mas não identificados por enriquecimento de GO. Esse fato aponta para a validade das análises de enriquecimento de GO para identificação de processos relacionados à resposta à droga, mas também pela relativa falta de poder para identificar genes associados a certos termos de G.O., evidenciando o estado inicial de anotações de GO de *T. cruzi*. A quantidade de genes anotados como permease de aminoácido dentro dos genes diferencialmente expressos está mais de 6 vezes aumentada em relação às anotações em todos os supergenes. Outros transportadores também foram identificados como aumentados em resposta à lovastatina: um transportador de hexose, um de nucleosídeo, dois transportadores de cátion e uma permease de fosfato. Diferentemente dos dados de cetoconazol, não foram identificados transportadores da família ABC como diferencialmente expressos.

O último termo de GO identificado como enriquecido dentre os genes de expressão aumentada foi: “processo biológico de nucleotídeos cíclicos”. Nesse grupo, estão presentes 3 supergenes anotados como possíveis receptores tipo adenilato ciclase, chamando atenção para ativação de vias de sinalização envolvidas na produção de cAMP em resposta à lovastatina.

Uma função molecular enriquecida dentre os supergenes de expressão aumentada, e de composição não redudante em comparação com os processos celulares já citados, foi a “atividade endopeptidase tipo cisteína”. Nesse grupo estão presentes 6 supergenes anotados como cisteína-peptidases, 3 deles são calpaínas (peptidases citoplasmáticas), uma é uma glutamil-carboxipeptidase e duas são catepsinas lisossomais/reservossomais (incluindo a cruzipaína). Juntamente com o aumento de expressão de um precursor de alfa-manosidase lisossomal (Tc00.1047053506195.120), a resposta genômica a IC<sub>50</sub> de lovastatina é compatível com o inchaço dos reservossomos observados por análises de biologia celular (veja Figura 5.12 e Figura 5.9).

A



B

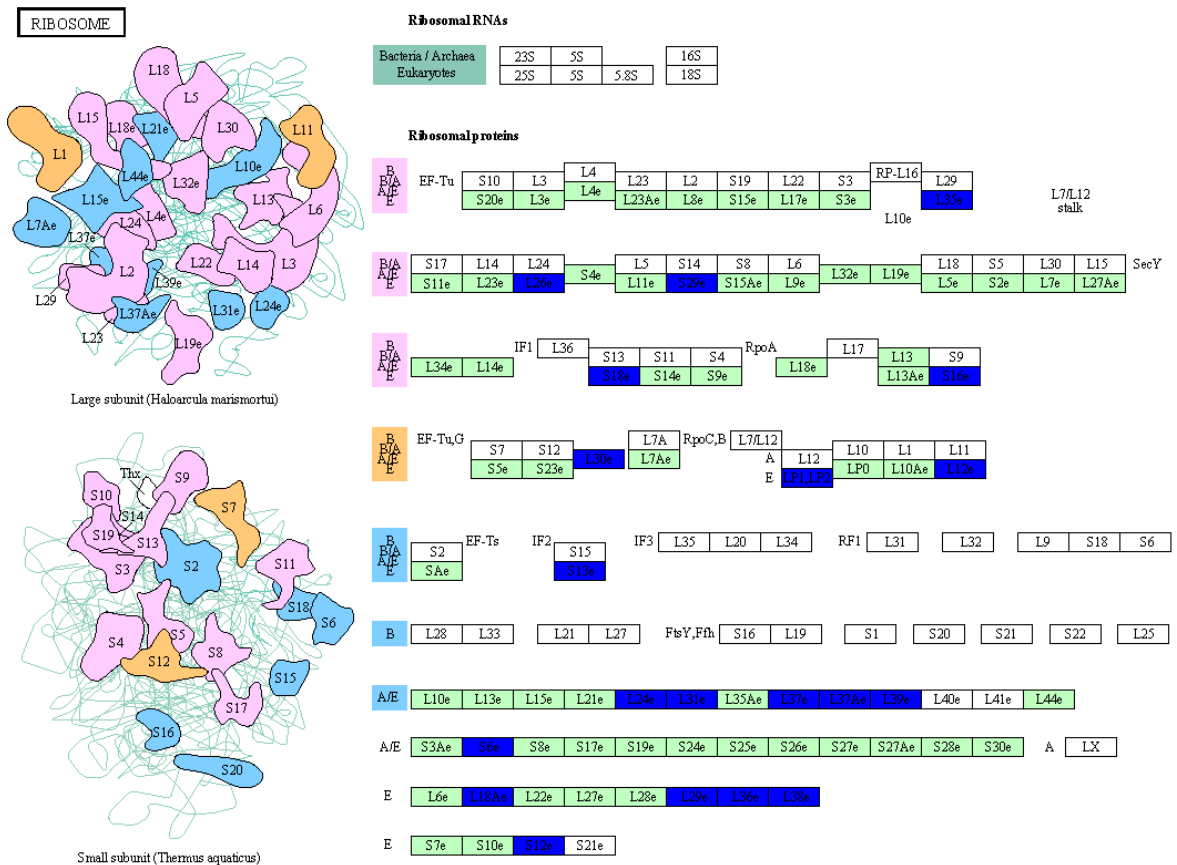


FIGURA 5.33 – IC<sub>50</sub> LOVASTATINA: VIAS ENRIQUECIDAS DO KEGG

Os dados são apresentados da mesma forma que na Figura 5.27. A: Via das Pentoses fosfato; B: ribossomo.

Para os genes de expressão diminuída em relação ao controle, foram identificados 4 processos celulares enriquecidos não-redundantes (veja Figura 5.32): “biogênese de ribossomos”, “organização do nucleossomo”, “processo metabólico de nucleosídeos” e “regulação de processo metabólico protéico”.

Para o termo “biogênese de ribossomos”, foram associados 24 supergenes. Destes, 20 estão anotados como proteínas ribossomais, uma como proteína reguladora da biogênese do ribossomo (RRS: *ribosome biogenesis regulatory protein*), uma fibrilarina (proteína associada a partículas snRNP e essencial nos primeiros passos de processamento do rRNA), uma proteína hipotética e uma proteína ligadora de GTP nuclear. Considerando apenas a anotação *ribosomal protein*, a tabela de genes diferencialmente expressos em resposta a lovastatina contém quase 4 vezes (3,86) mais genes em relação a porcentagem considerando todos os supergenes. O enriquecimento de proteínas ribossomais dentre os supergenes de expressão diminuída em resposta a lovastatina é apoiado pela detecção significativa dos termos redundantes função molecular “constituente estrutural de ribossomo” e componente celular “ribossomo” (veja Figura 5.32). Esta forte diminuição de genes codificadores de proteínas ribossomais também foi identificado pelas análises de vias do KEGG (veja Figura 5.33).

O segundo processo celular mais significativo dentre os genes de expressão diminuída foi “organização do nucleossomo”, ao qual estão associados 5 supergenes de histonas (H1, H2A, H2B, H3 e H4), além de uma proteína anotada como *nucleosome assembly protein-like*. A diminuição na expressão de genes de histonas também foi observado na resposta a IC<sub>50</sub> de cetoconazol (H2A, H3, H3 variante e H2B), indicando a alteração na estrutura da cromatina em resposta as drogas e/ou parada no ciclo celular e conseqüente diminuição na expressão de proteínas de interação com o DNA. Tal diminuição de histonas também foi identificada no nível componente celular, termo “nucleossomo”.

Outro processo celular significativamente enriquecido dentre os genes de expressão diminuída foi “processo metabólico de nucleosídeo”. Nesse grupo estão presentes 6 supergenes associados à síntese de nucleosídeos. A diminuição na síntese desse tipo de moléculas pode estar correlacionada à diminuição na biogênese de ribossomos, principal RNA (em quantidade) presente nas células.

Também foi identificado como significativamente enriquecido dentre os genes diminuídos o processo celular “regulação de processo metabólico protéico”. A este termo foram associados 9 supergenes: 5 estão anotados como fatores de iniciação da tradução, um como fator de alongação da tradução, uma proteína ativadora de ubiquitina E1, uma tirosina



fosfatase e uma proteína hipotética. Considerando a descrição *translation initiation factor*, os dados de genes significativos em resposta à lovastatina contem cerca de 3 vezes (3,12) mais genes em proporção à quantidade encontrada no genoma.

Diferentemente dos dados para cetoconazol, não foi observado um enriquecimento de quinases ou proteínas ligadoras de RNA. Foram detectadas como significativas apenas 11 quinases (9 com expressão aumentada) e 8 proteínas com descrição *RNA-binding protein*, todas elas de expressão diminuída.

### 5.3.4.3 DL100 Cetoconazol

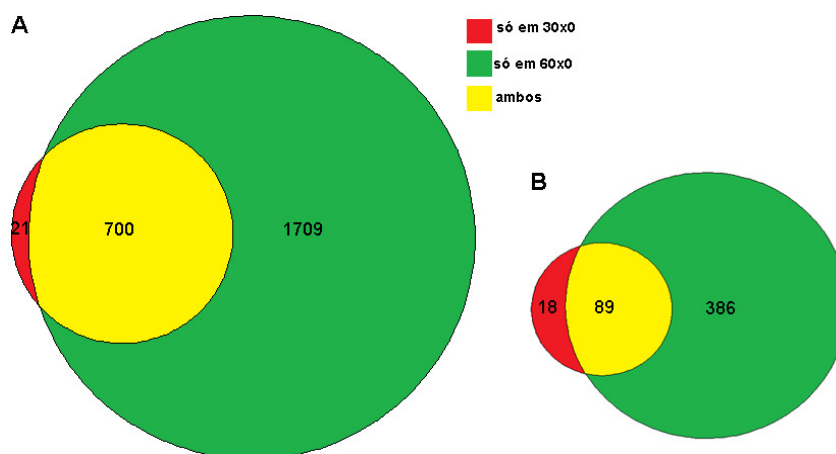
Na tabela Tabela 5.10 é mostrado o número de genes significativos para cada ponto de comparação nos testes estatísticos do programa edgeR. Diferentemente do experimento IC<sub>50</sub>, cada um dos pontos de tratamento (30 e 60 minutos) foram comparados diretamente com o controle (0 minutos), sendo que a quantidade de genes detectados como diferencialmente expressa foi bem grande, mesmo com limiar de FDR bem baixo (0,01%). Quando considerado limiar de FC de 1,5 vezes (mínimo de 50% de alteração no nível de expressão), o número de genes significativos cai bastante (compare parte superior e intermediária da Tabela 5.10). Isso demonstra que as réplicas biológicas possuem alta correlação nos níveis de expressão, e destaca o poder estatístico alcançado quando utilizado 3 réplicas biológicas por ponto de análise: pequenas variações nos níveis de expressão são detectadas como significativas. Quando considerado somente fortes alterações biológicas (FC de 2 vezes), o número de genes detectados é ainda menor.

**TABELA 5.10 – DL100 CETOCONAZOL: N DE GENES SIGNIFICATIVOS**

FC	Comparação	FDR5%	FDR1%	FDR0,1%	FDR0,01%
sem FC	C30xC0	3665 (57)	1914 (65)	880 (77)	482 (85)
sem FC	C60xC0	7693 (51)	5387 (53)	3358 (58)	2190 (64)
sem FC	C60xC30	83 (71)	22 (68)	4 (75)	0
FC1,5	C30xC0	960 (74)	853 (73)	721 (76)	482 (85)
FC1,5	C60xC0	2726 (62)	2585 (61)	2408 (60)	2107 (63)
FC1,5	C60xC30	49 (73)	22 (68)	4 (75)	0
FC2	C30xC0	107 (88)	93 (89)	78 (91)	73 (90)
FC2	C60xC0	475 (88)	447 (89)	415 (89)	404 (89)
FC2	C60xC30	3 (67)	1 (0)	0	0

A tabela mostra o número de genes significativos para cada uma das comparações especificadas na coluna “Comparação”, para diferentes níveis de significância estatística (cada uma das colunas “FDR”) e biológica (coluna “FC”: *fold change*). FC1,5 corresponde a um mínimo de 50% de alteração na expressão, enquanto FC2 a um mínimo de 100% de variação. O número entre (parênteses) é a porcentagem dos genes com expressão aumentada em cada comparação.

Para as análises subsequentes optou-se trabalhar com dois níveis de significância estatística e biológica: baixa significância estatística (FDR 5%) aliada a uma alta significância biológica (FC de 2 vezes), e alta significância estatística (FDR 0,1%) com significância biológica menor (FC de 1,5 vezes).



**FIGURA 5.34 – DL100 CETOCONAZOL: DIAGRAMA DE VENN**

Para cada um dos plots foi considerado o número de genes diferencialmente expressos segundo os testes estatísticos no edgeR. A: FDR0,1% e FC1,5; B: FDR5%, FC2.

A Figura 5.34A mostra que quase todos os genes (97%) diferencialmente expressos em 30 minutos de exposição a DL100 de Cetoconazol também estão selecionados em 60 minutos de exposição (sempre com o mesmo sentido de alteração na expressão em relação ao controle). Os poucos genes diferencialmente expressos somente em 30 minutos tem maior alteração na expressão em tempos curtos; mas tais genes não são significativos em 60 minutos pois a variação em relação ao controle é menor (embora o sentido de alteração seja o mesmo). Quando usado um limiar de significância biológica mais estridente (*fold change* de 2 vezes, 100% de alteração), a porcentagem de genes genes significativos em 30 minutos contidos em 60 minutos de exposição diminui (84%) (veja Figura 5.34B).

Como esperado, e mostrado na Figura 5.34, com 60 minutos de exposição diversos genes diferencialmente expressos são adicionados ao grupo de genes significativos em 30 minutos (independente do limiar de FDR e FC), demonstrando que as alterações são tempo-dependente. Considerando FDR máximo de 0,1% e FC de 1,5 vezes, 70% do total de genes diferencialmente expressos no experimento DL100 só são significativos em 60 minutos de exposição; este valor aumenta para 78% quando utilizado FDR5% e FC2.

Análises de agrupamento hierárquico dos genes diferencialmente expressos em resposta a DL100 de Cetoconazol (Figura 5.35) mostram que as réplicas biológicas são bem

mais homogêneas em comparação ao experimento IC<sub>50</sub>. Esse fato, aliado à presença de três réplicas biológicas, propiciou a detecção do grande número de genes diferencialmente expressos na comparação direta de cada ponto de exposição contra o controle, mesmo com níveis de FDR bem baixos.

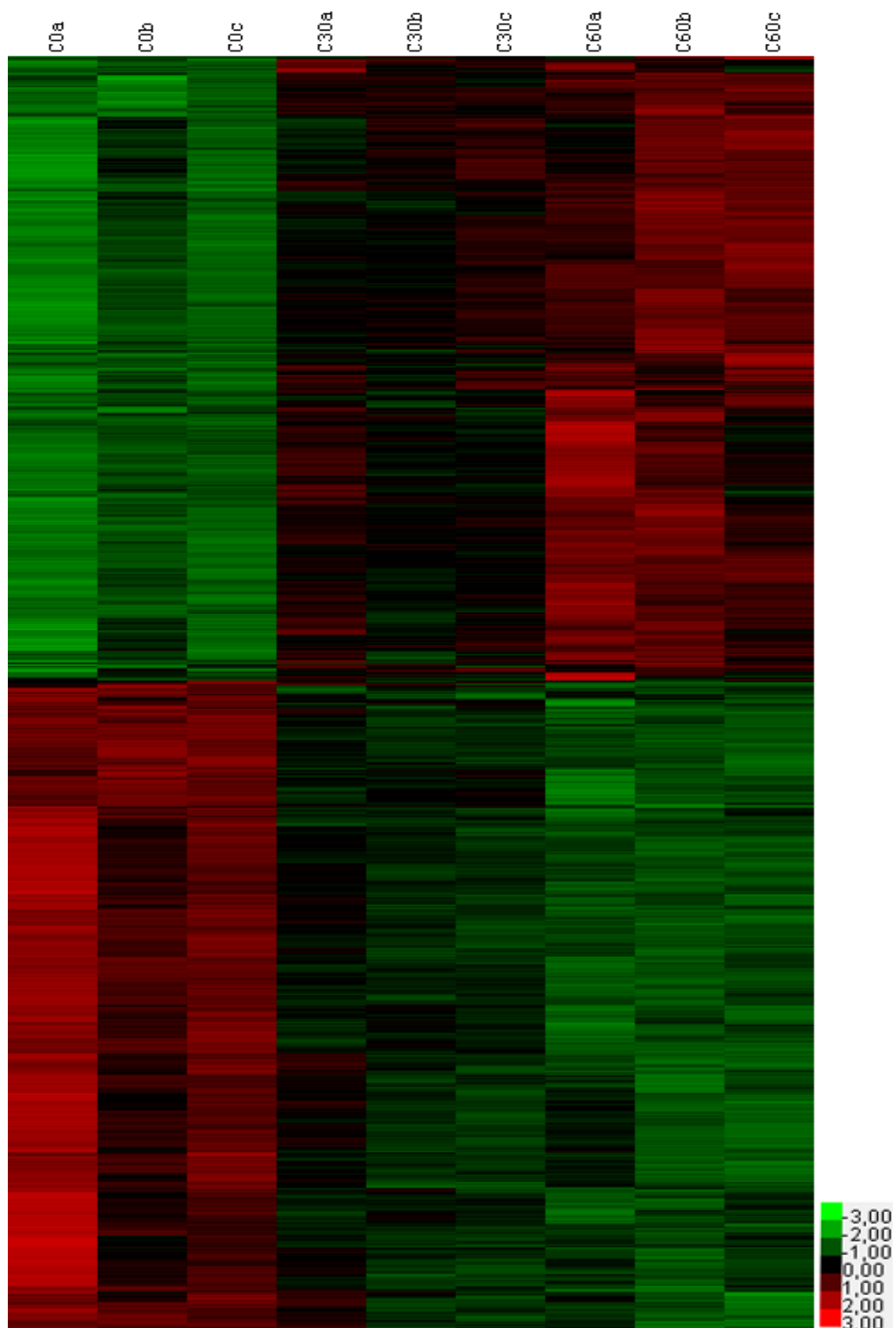
Acreditamos que esta alta correlação entre as réplicas biológicas deva-se ao próprio tipo de experimento sendo analisado. A dose letal da droga em menos de 24 horas de exposição (DL100) deve desencadear padrões de expressão fortemente estereotipados, os quais são naturalmente altamente reprodutíveis e, portanto, facilmente captados pelas análises de RNA-Seq.

Quando retirada a redundância dos dados, os 2565 genes (IDs) significativos com FDR0,1% e FC1,5 representam 1480 supergenes. Destes, 704 (~48%) possuem expressão aumentada em relação ao controle, enquanto 776 (~52%) têm expressão diminuída. Dentre os supergenes, 958 (~65%) estão anotados como proteínas hipotéticas. Quando considerado alta estringência de *foldchange* (FC 2 com FDR 5%), os 460 IDs significativos representam 272 supergenes, número muito baixo de genes para análises posteriores de ontologia gênica. Assim sendo, a seguir, trabalhamos apenas com os supergenes para FDR0,1% e FC1,5.

A listagem dos supergenes diferencialmente expressos em resposta a DL100 de Cetoconazol está presente na Tabela 5 do ANEXO II.

Quando agrupados por clusterização hierárquica, considerando todas as réplicas biológicas independentemente, os supergenes apresentam um padrão muito similar ao já apresentado na Figura 5.35: alta consistência entre as réplicas biológicas e alteração na expressão mais significativa em 60 minutos de exposição (para a grande maioria dos genes).

As análises de GO para os genes de expressão aumentada apontam para um sobre-representação de processos envolvidos na degradação de substratos e geração de energia: “movimento baseado em microtúbulo”, “fosforilação”, catabolismo de aminoácidos” e “proteólise” (veja Tabela 5.11).



**FIGURA 5.35 – DL100 CETOCONAZOL: AGRUPAMENTO HIERÁRQUICO**

Estão plotados os 2430 IDs (genes) detectados como diferencialmente expressos com FDR0,1% e FC1,5, na resposta a DL100 de Cetoconazol. Como já descrito, as cores representam o valor Z (legenda na parte inferior direita da figura). Repare a maior correlação visual das réplicas biológicas em comparação com os dados IC<sub>50</sub>.

TABELA 5.11 – ENRIQUECIMENTO DE G.O. PARA DL100 CETOCONAZOL

genes aumentados					
GO	PIFIC	name	FDR	n	n gen.
GO:0007018	P	microtubule-based movement	1.85E-07	32	65
GO:0016310	P	phosphorylation	2.16E-04	56	203
GO:0009063	P	cellular amino acid catabolic process	0.0316986	10	25
GO:0006508	P	proteolysis	0.037324	34	159
GO:0022857	F	transmembrane transporter activity	0.00747915	29	111
GO:0000104	F	succinate dehydrogenase activity	0.0089797	4	4
GO:0015645	F	fatty-acid ligase activity	0.0240583	4	4
GO:0004114	F	3',5'-cyclic-nucleotide phosphodiesterase activity	0.0281659	5	5
GO:0005875	C	microtubule associated complex	3.72E-06	31	73
GO:0045281	C	succinate dehydrogenase complex	0.0089797	4	4
genes diminuídos					
GO:0016070	P	RNA metabolic process	2.19E-05	47	198
GO:0034660	P	ncRNA metabolic process	0.00766996	20	69
GO:0034062	F	RNA polymerase activity	0.0477112	9	21
GO:0003723	F	RNA binding	0.0477112	28	145

As análises foram realizadas independentemente para os genes de expressão aumentada (parte superior da tabela) ou diminuída (parte inferior) em resposta a DL100 de cetoconazol. A coluna “name” indica o nome do termo anotado de G.O., com nível de ontologia indicado na coluna “PIFIC” (processo biológico, função molecular e componente celular). A coluna FDR representa a significância estatística do respectivo enriquecimento, com número de supergenes associados indicados na coluna “n”, enquanto o número total de supergenes no genoma como, para cada termo de G.O., está representado na coluna “n gen”.

Ao processo “movimento baseado em microtúbulo” foram associados 32 supergenes, 13 anotados como cinesinas, 15 como dineínas, uma tubulina-gama, uma quinases (*putative e phosphatidylinositol-4-phosphate 5-quinase-like protein*) e duas proteínas hipotéticas. O enriquecimento de proteínas motoras também é detectado no nível componente celular: “complexo associado a microtúbulo”.

Dentre os genes de expressão aumentada, o segundo processo biológico em significância estatística foi “fosforilação”. Nesse grupo estão presentes 56 supergenes; a grande maioria está anotada como quinase (38), nove são proteínas hipotéticas, uma chaperona, 3 genes envolvidos no metabolismo de NADP (*NADH-dependent fumarate reductase, putative*), duas subunidades da ATP sintase mitocondrial (*ATP synthase, alpha chain, mitochondrial precursor, ATPase alpha subunit (pseudogene), putative*), uma ATPase de membrana e uma flavoproteína (*succinate dehydrogenase flavoprotein, putative*). A possível ativação de vias de transdução de sinal é corroborada pela identificação de enriquecimento da função molecular “atividade fosfodiesterase de nucleosídeos cíclicos”, a qual foram associados 5 supergenes (duas fosfodiesterase de cAMP e 3 proteínas hipotéticas).

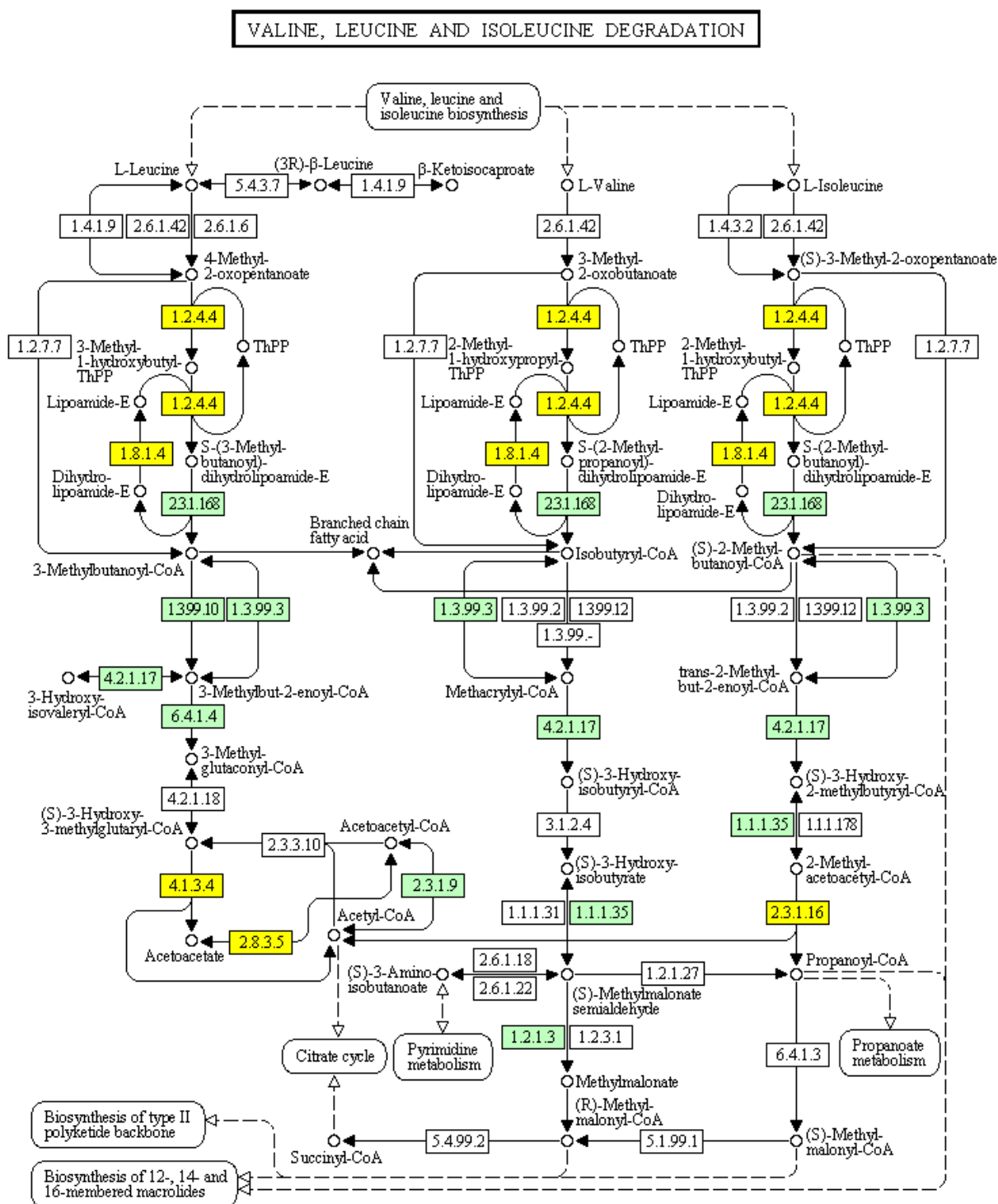
Para o termo “processo catabólico de aminoácidos” foram associados 10 supergenes. As anotações desses são bastante variáveis, mas de forma geral representam enzimas

envolvidas nas degradação de aminoácidos, muitas delas de localização mitocondrial, relacionaods com a geração de energia para o ciclo de Krebs. Tal aumento de vias de geração de energia é corroborado pelo enriquecimento das funções moleculares “atividade succinato desidrogenase” (4 supergenes envolvidos na síntese de NADH) e “atividade ligase de ácido graxo” (4 supergenes para *fatty acyl CoA synthetase, putative*). Em nível de componente celular, também foi identificado o enriquecimento do termo “complexo succinato desidrogenase”, o qual possuiu os mesmos 4 supergenes associados ao processo “atividade succinato desidrogenase”.

A ativação de degradação de aminoácidos em resposta a DL100 de Cetoconazol pode ser observada, pelo dados do KEGG, no aumento de expressão de genes associados a via de degradação dos aminoácidos Valina, Leucina e Isoleucina (veja Figura 5.36)

Outro processo enriquecido nos genes de expressão aumentada em resposta a DL100 de Cetoconazol foi “proteólise”. Nesse grupo estão presentes 34 supergenes, a maioria (21) anotado como peptidases (cisteína-peptidase, carboxi-peptidase, metalo-peptidase, amino-peptidase e oligo-peptidase), além de algumas proteínas hipotéticas (6), dineínas (4), e GP63 (2). Tal atividade proteolítica parece estar correlacionado com o aumento de vias de geração de energia na célula, como a degradação de aminoácidos associado a síntese de NADH e ATP. Além disso, a análise manual dos genes de expressão aumentada em resposta a DL100 de Cetoconazol revelou a presença de uma subunidade do complexo citocromo c oxidase da cadeia transportadora de elétrons mitocondrial.

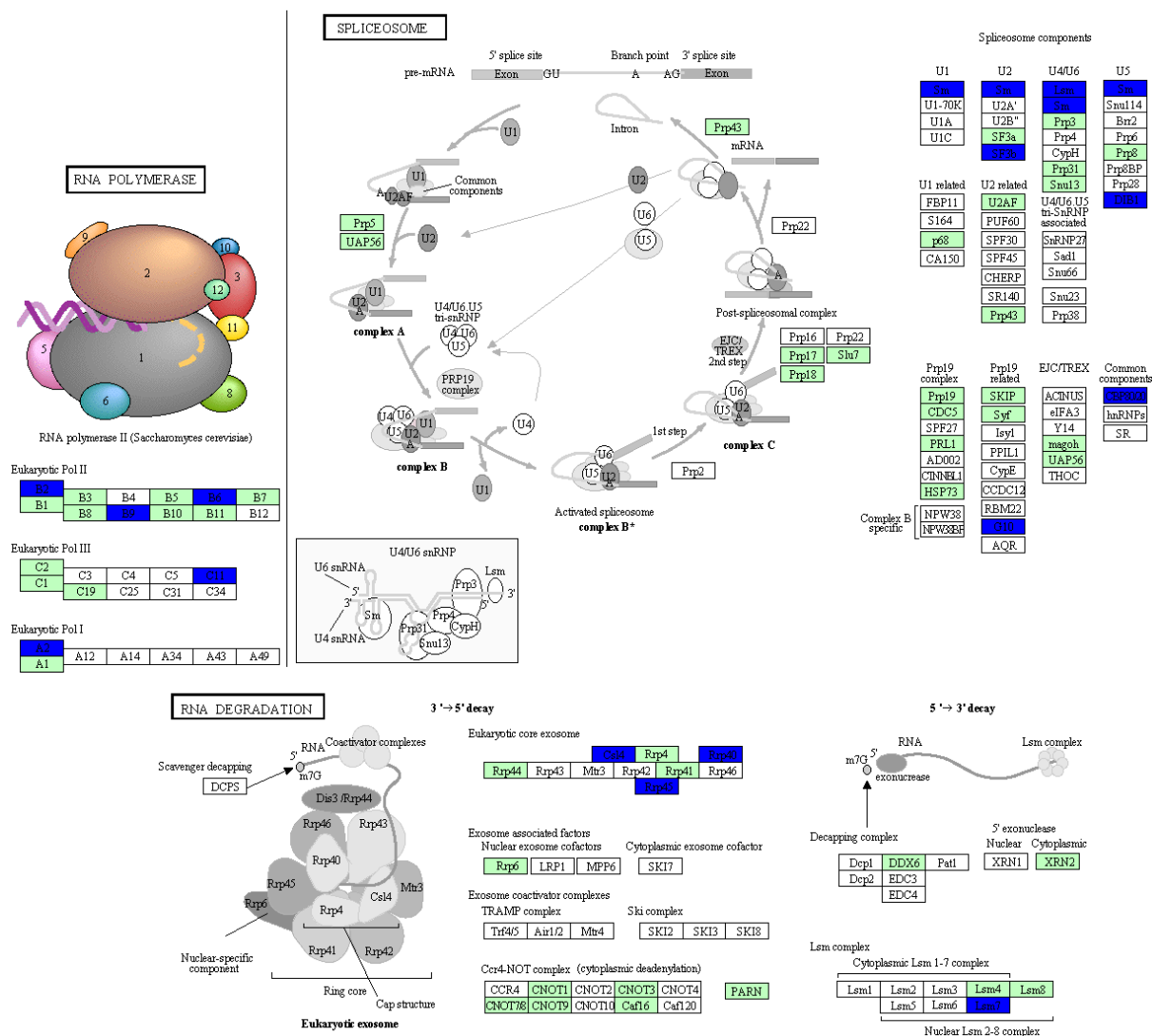
Uma função molecular não redundante identificada dentre os genes de expressão aumentada foi “atividade de transporte transmembrana”. Nesse grupo estão presentes 29 supergenes que codificam uma variedade de transportadores de membrana, incluindo: 6 transportadores ABC, 6 ATPases, uma ATP sintetase, transportadores de hexose, cátion e nucleobase, além de canal de cloreto. Chama atenção que quatro dos transportadores ABC aumentados em resposta a DL100 de Cetoconazol também são em resposta a IC<sub>50</sub> da mesma droga.



**FIGURA 5.36 – DL100 CETOCONAZOL: VIA DE DEGRADAÇÃO DE AMINOÁCIDOS (KEGG)**  
 O código de cores é o mesmo já apresentado nas figuras do KEGG anteriores.

Para os genes de expressão diminuída em relação ao controle, observou-se um claro enriquecimento de processos relacionados ao metabolismo de RNA: “processo metabólico de RNA” e “processo metabólico de ncRNA”, além das funções moleculares “atividade RNA

polimerase” e “ligação a RNA”. Assim sendo, altas doses de Cetoconazol parecem induzir uma estagnação geral da transcrição.



**FIGURA 5.37 – DL100 CETOCONAZOL: INIBIÇÃO DA SÍNTESE E PROCESSAMENTO DE RNA**

São apresentadas três vias relacionadas a síntese e processamento de RNA com diversos genes associados com expressão diminuída. No canto superior esquerdo é demonstrada a diminuição na expressão de subunidades das três RNA polimerases; Na canto superior direito está evidenciada a via de *splicing* e na parte inferior as vias de degradação de RNA.

Ao “processo metabólico de RNA” foram associados 47 supergenes, os quais codificam uma variedade de proteínas relacionadas à transcrição propriamente dita (como a própria RNA polimerase), processamento (snRNAs e snoRNAs) e regulação (proteínas de ligação de RNA) de RNAs. O “processo metabólico de ncRNA” tem composição parcialmente redundante em relação ao citado acima, mas possui diversos supergenes anotados como tRNA-sintetase antes não identificados. A diminuição da expressão de RNAs



polimerases fica evidente no enriquecimento da “atividade RNA polimerase” (veja Figura 5.37), ao qual foram associados 9 supergenes: 3 subunidades de RNAPol III, duas subunidades de RNAPol I e polIII, além de duas proteínas hipotéticas. A diminuição geral da transcrição também é identificada pela função molecular “ligação a RNA”, associada a uma variedade de proteínas de interação com o RNA, como: proteínas ribossomais, proteínas de complexos ribonucleoprotéicos snRNPs, proteínas ligadoras de RNA, proteínas associadas ao exossomo (veja Figura 5.37) e tRNA sintetases. Tal estagnação na transcrição e processamento de RNA é facilmente visualizada nas vias do KEGG (veja Figura 5.37).

#### 5.3.4.4 DL100 Lovastatina

A Tabela 5.12 apresenta o número de genes diferencialmente expressos para lovastatina, dose DL100, nos dois pontos de análise (30 e 60 minutos).

**TABELA 5.12 – DL100 LOVASTATINA: N DE GENES SIGNIFICATIVOS**

FC	Comparação	FDR5%	FDR1%	FDR0,1%	FDR0,01%
sem FC	L30xL0	11815 (48)	9797 (49)	7637 (49)	6080 (49)
sem FC	L60xL0	13558 (48)	11900 (48)	10008 (49)	8665 (49)
sem FC	L60xL30	2446 (65)	1288 (75)	607 (85)	322 (93)
FC1,5	L30xL0	3935 (41)	3815 (41)	3655 (42)	3552 (42)
FC1,5	L60xL0	6464 (45)	6322 (46)	6126 (46)	5972 (47)
FC1,5	L60xL30	242 (79)	210 (81)	172 (87)	147 (93)
FC2	L30xL0	512 (28)	482 (28)	446 (29)	437 (30)
FC2	L60xL0	1629 (44)	1599 (45)	1555 (46)	1533 (46)
FC2	L60xL30	30 (90)	25 (100)	22 (100)	21 (100)

A tabela mostra o número de genes significativos para cada uma das comparações especificadas na coluna “Comparação”, para diferentes níveis de significância estatística (cada uma das colunas “FDR”) e biológica (coluna “FC”: *fold change*). FC1,5 corresponde a um mínimo de 50% de alteração na expressão, enquanto FC2 a um mínimo de 100% de variação. O número entre (parênteses) é a porcentagem dos genes com expressão aumentada em cada comparação.

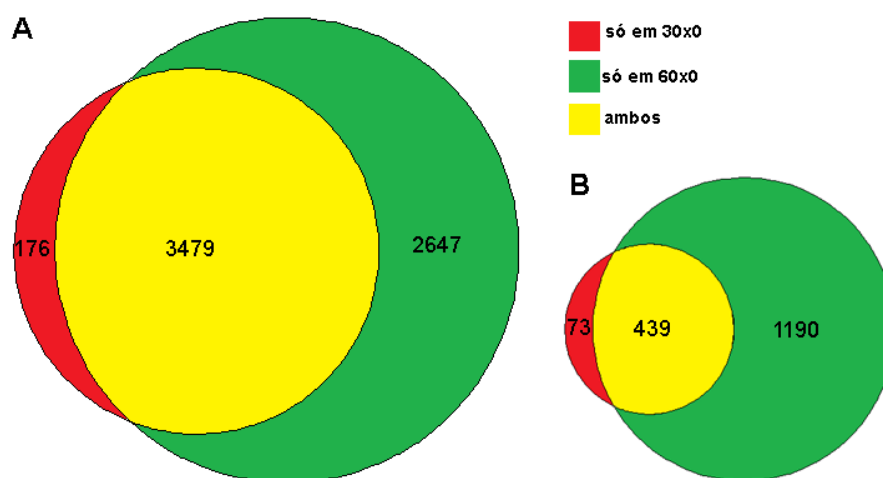
Há um grande número de genes (IDs) detectados como diferencialmente expressos, mesmo com limiares de significância estatística bem baixos (veja Tabela 5.12).

Da mesma forma que as análises feitas para cetoconazol, a sobreposição dos genes detectados como significativos em cada ponto de comparação foi analisada por diagrama de Venn. A Figura 5.38 mostra que as alterações na expressão são tempo dependentes, mas o tempo 60 minutos de exposição não incorpora tantos genes significativos como o tratamento com cetoconazol, principalmente quando considerado FDR0,1% e FC1,5 (42% do total de

genes significativos somente detectados em 60 minutos; 70% para cetoconazol). Entretanto, quando considerado FDR5% e FC2, o tempo 60 minutos acrescenta 70% do total de genes diferencialmente expressos na resposta à droga (valor similar para cetoconazol: 78%); ou seja: em tempos maiores de exposição a intensidade de variação de expressão é maior, motivo pelo qual estes genes não são detectados em 30 minutos com *fold change* de 2 vezes.

De forma similar aos resultados para DL100 de cetoconazol, a grande maioria dos genes detectados como diferencialmente expressos em 30 minutos de exposição (~97%) estão contidos no grupo de genes significativos em 60 minutos, todos estes com alteração na expressão no mesmo sentido em relação ao controle. Ou seja, as respostas significativas em 30 minutos de exposição a droga são exacerbadas em 60 minutos.

Quando os genes diferencialmente expressos são visualizados por um agrupamento hierárquico (veja Figura 5.39), fica evidente a alta correlação entre as réplicas biológicas do experimento DL100 Lovastatina, como já demonstrado no tópico 5.3.3 (R2 médio de 0.99, veja Figura 5.21).



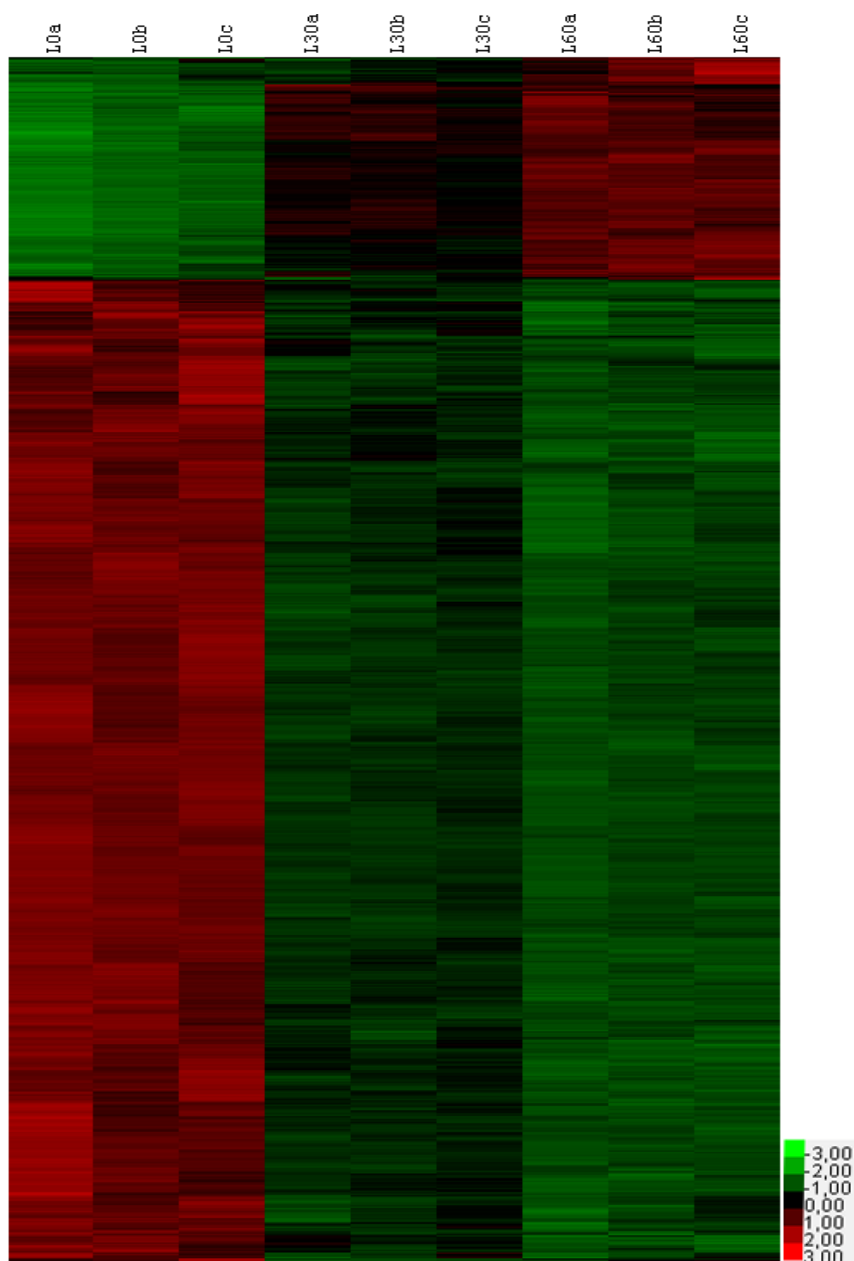
**FIGURA 5.38 – DL100 LOVASTATINA: DIAGRAMA DE VENN**

Para cada um dos plots foi considerado o número de genes diferencialmente expressos segundo os testes estatísticos no edgeR. A: FDR0,1% e FC1,5; B: FDR5%, FC2.

Como já postulado para a resposta com cetoconazol, esta alta reprodutibilidade biológica dos dados deve estar relacionada a respostas estereotipadas frente ao forte estresse celular desencadeado na presença de altas doses das drogas.

Um detalhe que chama atenção na Figura 5.39 é grande quantidade de genes com expressão diminuída em relação ao controle (~82% dos genes detectados). Essa resposta difere do padrão para cetoconazol para o qual cerca da metade dos supergenes detectados

como diferencialmente expressos (~48%) tem expressão diminuída (veja Figura 5.35). Portanto, a resposta a DL100 de lovastatina deve desencadear uma forte inibição de diversos processos celulares, em um mecanismo estereotipado de estresse abiótico, o que culmina com a diminuição dos níveis de mRNA de diversos genes.



**FIGURA 5.39 – DL100 LOVASTATINA: AGRUPAMENTO HIERÁRQUICO**

Estão plotados os 5603 IDs (genes) detectados como diferencialmente expressos com FDR0,01% e FC1,5, na resposta a DL100 de Lovastatina. Como já descrito, as cores representam o valor Z (legenda na parte inferior direita da figura). Da forma similar aos dados DL100 Cetoconazol, repare a maior correlação visual das réplicas biológicas em comparação com os dados IC<sub>50</sub>.

Para lovastatina, trabalhamos com níveis de significância estatística ainda mais baixos (FDR0,01%), aliado a FC de 1,5 vezes (5603 genes selecionados), ou FDR1% com maior significância biológica, FC2 (2884 genes selecionados).

Após retirada da redundância, obteve-se 2904 supergenes diferencialmente expressos à DL100 de Lovastatina, considerando FDR 0,01% e FC 1,5, e 1447 considerando FDR1% e FC2. Como já ocorrido para o total de genes, apenas 19% (542) dos supergenes tem expressão aumentada em relação ao controle (para FDR0,01% e FC1,5). Quando considerado maior estringência biológica, FC 2, aliada a um FDR de 1%, a porcentagem de genes aumentados é ainda menor (9%, 132 supergenes). Assim sendo, de forma a selecionar um maior número de genes aumentados em resposta a droga, optou-se trabalhar com menor significância biológica (FC1,5) aliada a alta significância estatística (FDR0,01%).

Dos 2904 supergenes selecionados, 2105 (72%) codificam proteínas anotadas como hipotéticas, porcentagem mais alta de genes de função desconhecida observada dentre todas as análises realizadas.

**TABELA 5.13 – ENRIQUECIMENTO DE G.O. PARA DL100 LOVASTATINA**

genes aumentados					
GO	PIFIC	name	FDR	n	n gen.
GO:0006084	P	acetyl-CoA metabolic process	8.68E-09	17	17
GO:0009063	P	cellular amino acid catabolic process	2.36E-06	15	20
GO:0006096	P	glycolysis	1.02E-05	15	23
GO:0006399	P	tRNA metabolic process	3.04E-04	17	44
GO:0015645	F	fatty-acid ligase activity	0.00612415	4	4
GO:0008233	F	peptidase activity	0.00649717	26	109
GO:0008135	F	translation factor activity, nucleic acid binding	0.0291798	12	45
GO:0045261	C	proton-transporting ATP synthase complex, catalytic core F(1)	0.0022647	4	4
GO:0030286	C	dynein complex	0.00252453	12	30
GO:0045239	C	tricarboxylic acid cycle enzyme complex	0.0143454	3	3
genes diminuídos					
GO:0016070	P	RNA metabolic process	0.00152569	89	156
GO:0034470	P	ncRNA processing	0.00500197	26	27
GO:0004518	F	nuclease activity	0.000632	33	33

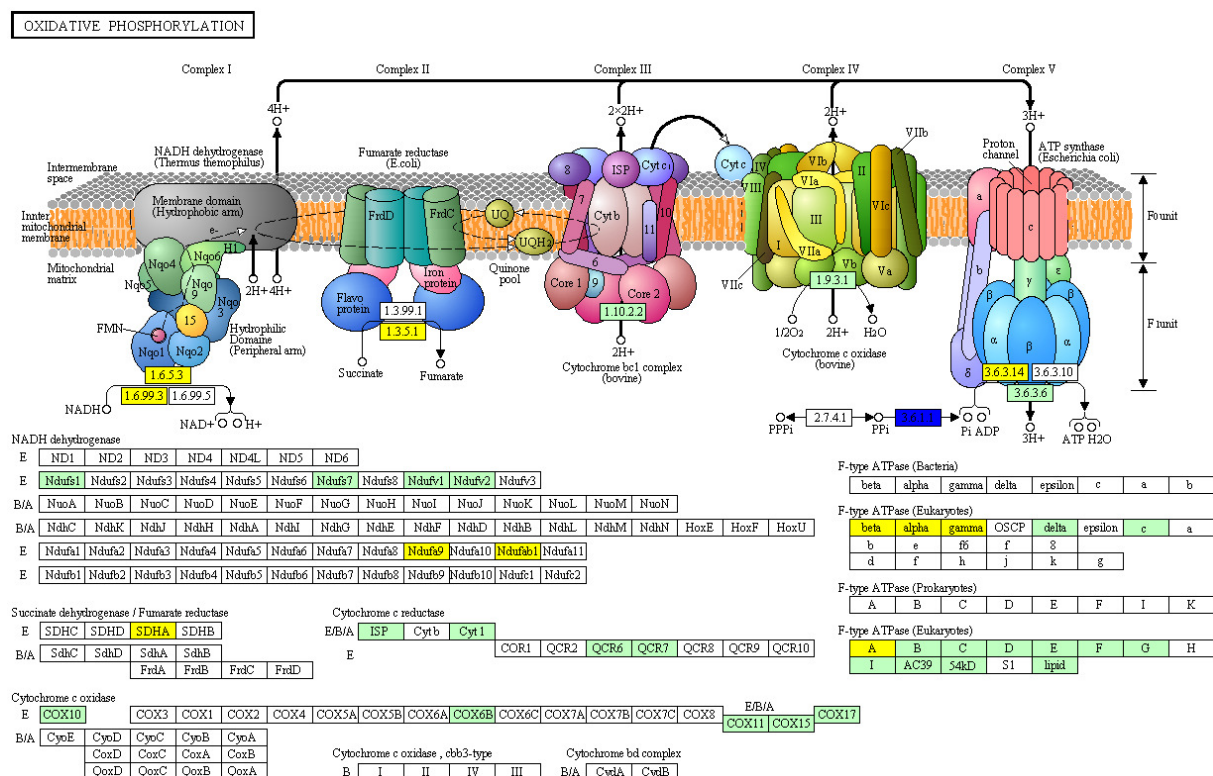
As análises foram realizadas independentemente para os genes de expressão aumentada (parte superior da tabela) ou diminuída (parte inferior) em resposta a DL100 de Lovastatina. A coluna “name” indica o nome do termo anotado de G.O., com nível de ontologia indicado na coluna “PIFIC” (processo biológico, função molecular e componente celular). A coluna FDR representa a significância estatística do respectivo enriquecimento, com número de supergenes associados indicados na coluna “n”; a quantidade de supergenes presentes no genoma associados a cada termo de G.O. é representada na coluna “n gen.”.

Após as análises de enriquecimento de termos de GO, foram identificados quatro processos biológicos enriquecidos dentre os genes de expressão aumentada em resposta à Lovastatina (veja Tabela 5.13): “processo metabólico de AcetilCoa”, “processo catabólico de aminoácidos”, “glicólise”, “processo metabólico de tRNA”.

Analisando em conjunto, parece claro que a resposta à DL100 da droga induz um aumento nos processos de degradação de metabólitos e produção de ATP. Para o catabolismo de aminoácidos foram associados 15 genes de funções e anotações variadas, o que chama atenção para a utilização de aminoácidos como fonte energética em *T. cruzi*. Ao termo “glicólise” foram relacionados 15 supergenes, os quais codificam enzimas acessórias ou da própria via glicolítica. Ao “processo metabólico de AcetilCoA” foram associados 17 supergenes relacionados a síntese de AcetilCoA e produção de NADH. Além disso, a função molecular “atividade ligase de ácido-graxo” possui 4 supergenes não redundantes, todos anotados como *fatty acyl CoA synthetase*. A ativação da cadeia transportadora de elétrons fica clara pelo enriquecimento do componente celular “*proton-transporting ATP synthase complex, catalytic core F(1)*”, ao qual foram associados 4 supergenes que codificam possíveis subunidades da ATP sintase. Além disso, também está enriquecido o componente celular “complexo enzimático ciclo do ácido cítrico”, ou ciclo de Krebs, relacionado a 3 supergenes do ciclo do ácido cítrico. A análise manual dos genes diferencialmente expressos revelou o aumento na expressão de diversas proteínas da cadeia transportadora de elétrons: 5 subunidades do complexo citocromo c oxidase, 2 supergenes anotados como citocromo B5 redutase e 1 citocromo B5-like. Análises no banco de dados de vias metabólicas do KEGG corroboram o aumento de expressão de genes relacionados ao catabolismo e síntese de ATP: via das pentoses fosfato, metabolismo de piruvato, ciclo do ácido cítrico, metabolismo do piruvato, fosforilação oxidativa (veja figura Figura 5.40), dentre outras.

O enriquecimento da função molecular “atividade peptidase”, dentre os genes de expressão aumentada, também deve estar relacionado a produção de substratos para ativação de vias de produção de ATP. A esse termo foram associados 26 supergenes, com atividade peptidásica variada: carboxipeptidase, metalopeptidase, cisteína-peptidase, subunidades do proteossomo, aminopeptidase e oligopeptidase.

Outro processo celular enriquecido dentre os genes de expressão aumentada em resposta a Lovastatina foi “processo metabólico de tRNA”. A esse termo foram associados 17 supergenes, dos quais 13 codificam para enzimas tRNA-sintase. Além disso, foi identificada a função molecular “atividade de fator de tradução, ligação a ácido nucléico”. Nesse grupo foram listados 5 fatores de iniciação da tradução, 4 fatores de alongação, um fator de liberação (terminação da tradução) e um fator de alongação mitocondrial.



**FIGURA 5.40 – DL100 LOVASTATINA: FOSFORILAÇÃO OXIDATIVA (KEGG)**

Note o aumento na expressão de dois genes codificadores de proteínas do complexo NADH desidrogenase, um gene do complexo Succinato desidrogenase e algumas subunidades da ATP sintase.

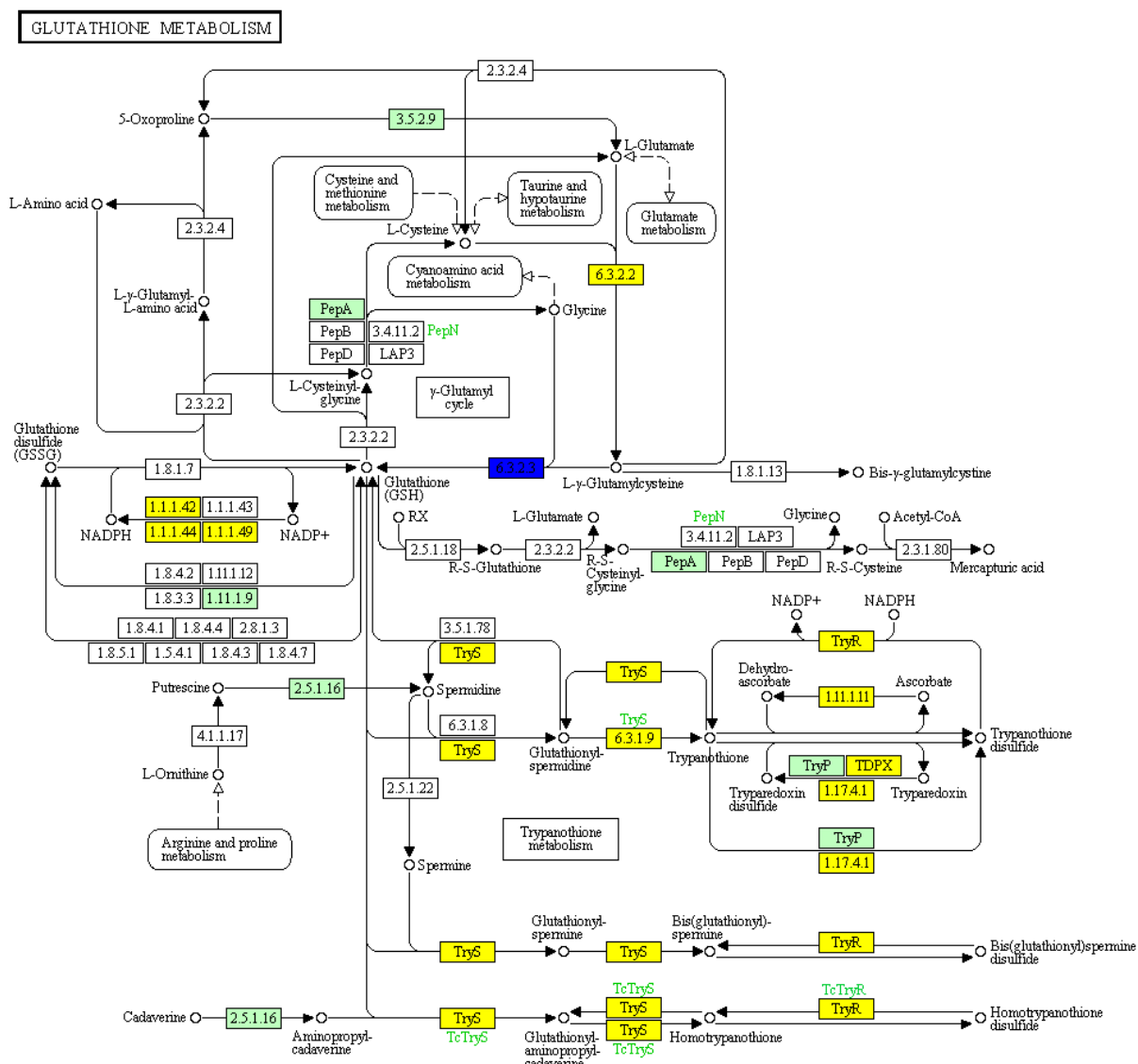
Por fim, foi identificado o componente celular “complexo dineína”, associado a 12 supergenes anotados como dineínas.

Analisando em conjunto, a resposta a DL100 de lovastatina parece estar dividida em duas categorias: degradação de metabólicos associado a produção de ATP por fosforilação oxidativa e aumento de síntese protéica.

Uma via interessante pelas análises de KEGG, enriquecida com genes de expressão aumentada em resposta a lovastatina, mas não detectada pelas análises de GO, foi a via de metabolismo de glutatona (veja Figura 5.41). Nessa via estão alguns genes essenciais à síntese de glutatona e tripanotona, principais moléculas relacionadas à defesa contra estresse oxidativo em *T. cruzi* e conhecidamente relacionadas à resposta a drogas (revisto por MAYA *et al.*, 2007).

Para os genes de expressão diminuída em resposta a DL100 de lovastatina, os resultados são similares a resposta a DL100 de cetoconazol, com forte impacto no processo de transcrição. Os processos biológicos gerais identificados que representam tal resposta são: “processo metabólico de RNA” e “processamento de ncRNA”, além da função molecular “atividade nucleásica”. Ao “processo metabólico de RNA” foram associados diversos tipos de

genes relacionados à síntese, processamento e degradação de RNA: diversas helicases, incluindo subunidades do exossomo, proteínas associadas a snRNPs, subunidades da RNA polimerase II (4 supergenes) e III (3 supergenes), algumas proteínas ligadoras de RNA, fatores de transcrição e proteínas associadas ao capeamento e poliadenilação de mRNAs maduros.



**FIGURA 5.41 – DL100 LOVASTATINA: ATIVAÇÃO DA VIA DA GLUTATIONA (KEGG)**

Note o aumento na expressão de diversos genes envolvidos na síntese de moléculas anti-oxidantes.

Ao “processamento de RNA” foram associados 26 supergenes, parcialmente redundantes ao processo biológico “processo metabólico de RNA”. Entretanto, o que chamou atenção nesse grupo foi a grande quantidade de proteínas hipotéticas (50%), associadas, de forma geral, a proteínas de processamento de RNA.

A função molecular “atividade nucleásica” reafirma a grande quantidade de supergenes com atividade nucleásica diminuídos em resposta a lovastatina, mas é o único grupo que inclui quatro supergenes anotadas como proteínas de reparo de DNA.

De forma geral, a dose DL100 das duas drogas induz forte inibição da transcrição.

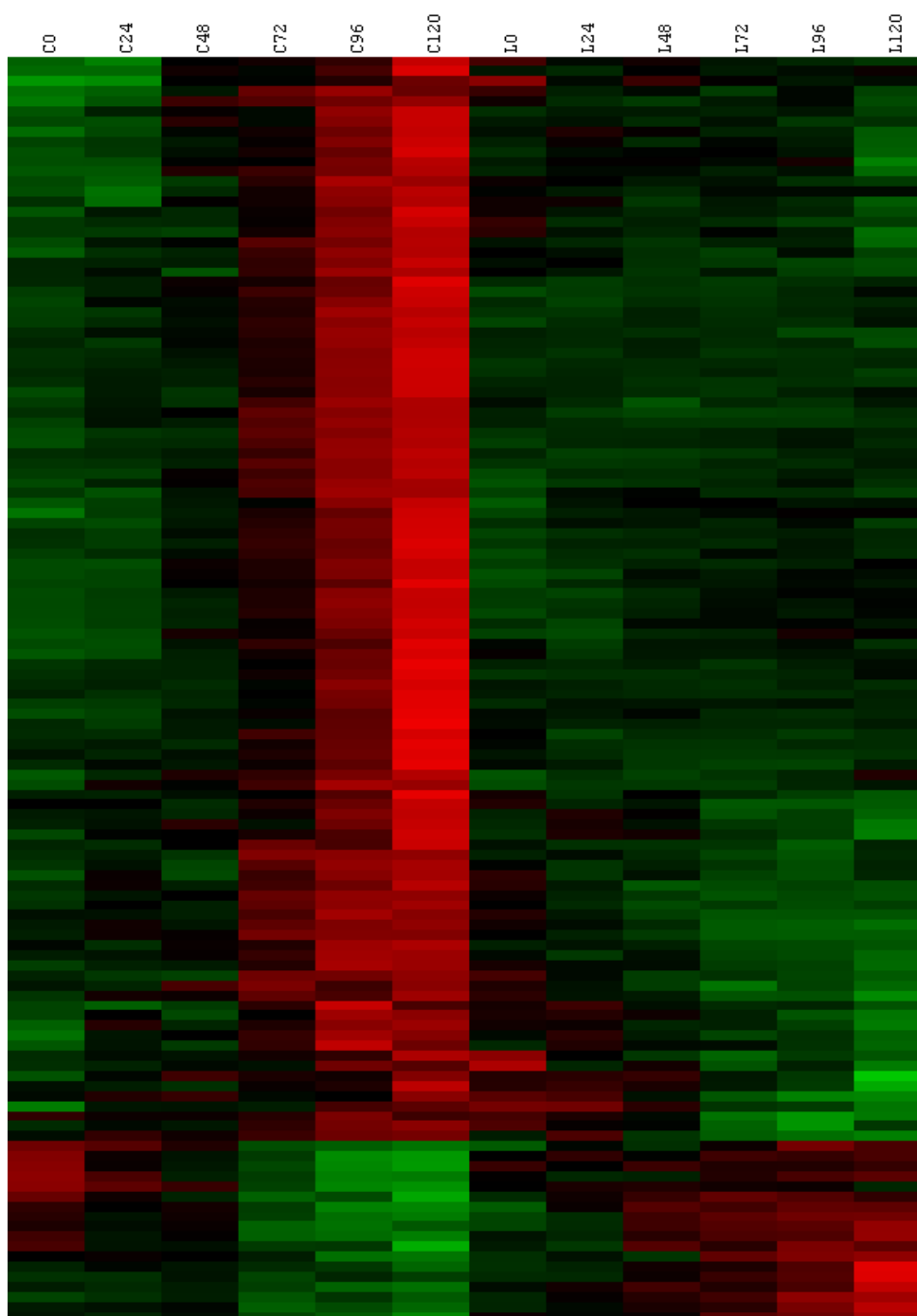
#### 5.3.4.5 IC<sub>50</sub>: respostas discrepantes entre as drogas

Além da análise individual de cada um dos inibidores da síntese de esteróis usados, comparou-se a resposta transcricional de *T. cruzi* a IC<sub>50</sub> de cetoconazol e lovastatina. Para tal, utilizamos o programa SAM, mais especificamente o método *two class unpaired timecourse*, o qual busca padrões distintos de resposta entre os experimentos sendo analisados. Nesse caso, foram utilizadas as tabelas de contagens de *reads* gerais (cetoconazol e lovastatina) normalizadas no programa edgeR.

O número de genes detectados como diferencialmente regulados entre as drogas foi pequeno, podendo ser explicado tanto pelo baixo poder estatístico de duas réplicas, mas também pela existência de poucas diferenças biológicas entre as respostas às drogas. Com FDR 10%, foram identificados 252 genes (IDs) significativos, número este que cai para 48 quando considerado FDR de 5%. De forma a selecionar somente genes fortemente diferentemente regulados, e portanto biologicamente significativos na comparação entre as drogas, utilizamos alta estringência de FC (mínimo de 2 vezes), com um critério mais frouxo de FDR (10%). Após seleção por FC 2, 206 IDs, dos 252 iniciais, passaram pelo critério de seleção e foram selecionados para as análises posteriores.

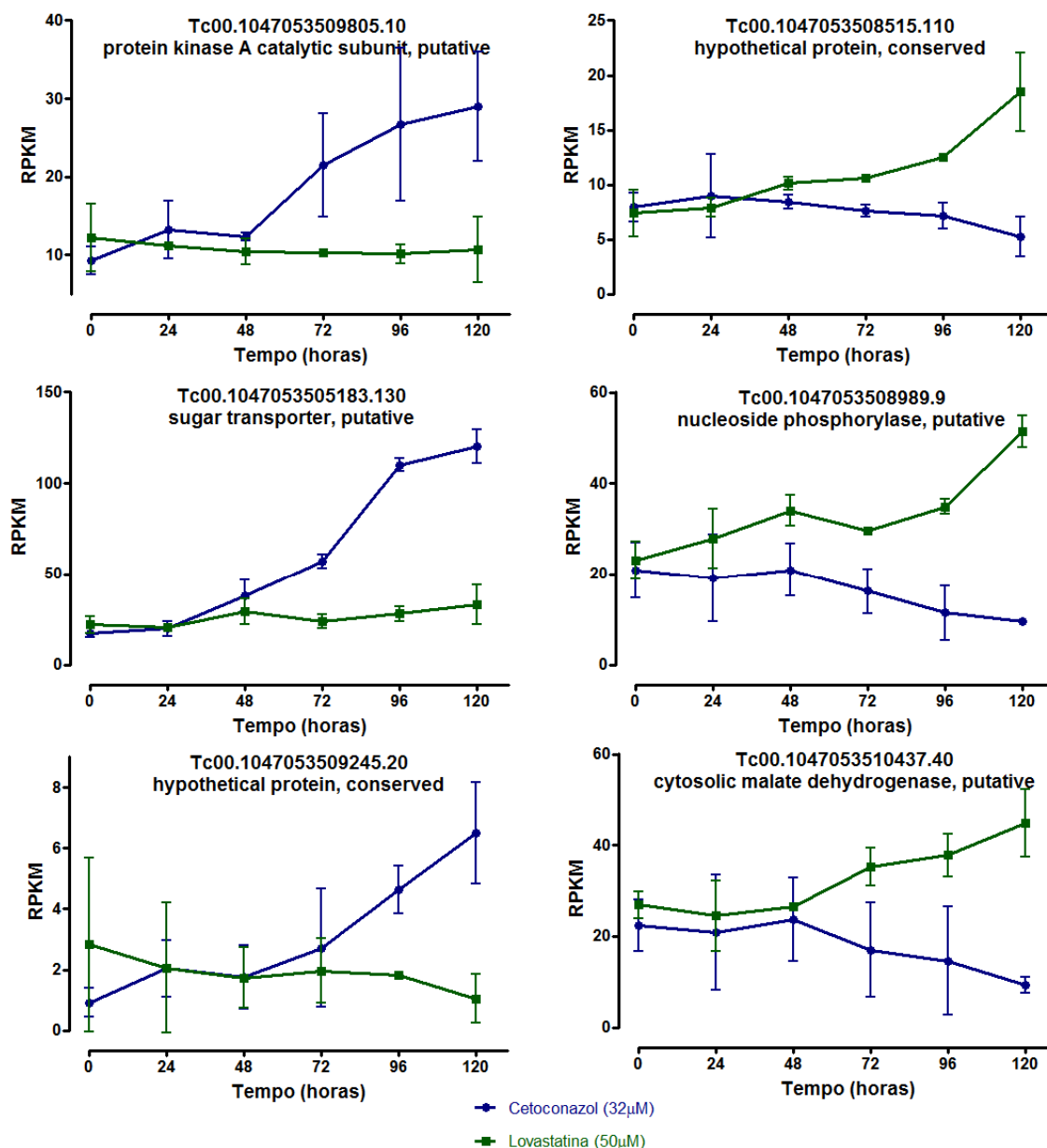
Após retirada da redundância e de famílias gênicas, obteve-se 126 genes diferencialmente regulados entre as drogas. A Figura 5.42 mostra uma clusterização hierárquica destes 126 genes. Fica fácil observar que a grande maioria deles (108, ~86%) são genes aumentados somente na resposta a cetoconazol, enquanto se mantêm estáveis, ou diminuem sua expressão, na resposta a lovastatina. Somente 18 genes apresentaram aumento de expressão unicamente em resposta a lovastatina (~14%). A Tabela 7 do ANEXO II lista todos os genes diferencialmente regulados na resposta a IC<sub>50</sub> entre as drogas.





**FIGURA 5.42 – IC<sub>50</sub>: COMPARAÇÃO ENTRE AS DROGAS**

A figura mostra uma clusterização hierárquica dos 126 genes selecionados como diferencialmente regulados na resposta a IC<sub>50</sub> das drogas. Foram usados os valores Z, considerando as tabelas de contagens de *reads* gerais para ambas as drogas (normalizadas no edgeR). A escala de cores é a mesma da figura Figura 5.39



**FIGURA 5.43 – IC<sub>50</sub>: GENES DIFERENCIALMENTE REGULADOS ENTRE AS DROGAS**

O lado esquerdo da figura contém 3 exemplos de genes de expressão aumentada somente na resposta a IC<sub>50</sub> de Cetoconazol, enquanto no lado direito estão plotados 3 genes com expressão aumentada somente para Lovastatina. Cada ponto (0-120 horas) representa a média das réplicas biológicas, com respectivas barras de desvio indicando o desvio-padrão da contagem de *reads* normalizadas pelo tamanho da biblioteca e tamanho do gene (RPKM).

Para alguns casos (veja os genes aumentados para lovastatina na Figura 5.42), a resposta é contrária, ou seja, aumento em resposta a uma das drogas e diminuição para a outra. Essas respostas droga-específica são muito informativas, tanto para a análise de reguloma (veja tópico 6.5), visto que adicionam variação biologicamente significativa nos dados de expressão gênica, quanto para uma melhor compreensão da biologia do parasita, em relação à sua capacidade distinta de modular genes específicos. Além disso, esses dados

chamam atenção para os diferentes alvos de ação das drogas usadas, mesmo embora atuem, de forma global, na mesma via biológica.

Análises de enriquecimento de GO para tais genes não geraram nenhum resultado significativo, o que deve estar associado à pequena quantidade de supergenes analisados (126 genes) e ao baixa determinação de termos GO para os genes de *T. cruzi*. Portanto, buscou-se genes enriquecidos pela análise manual da tabela de genes diferencialmente regulados entre a IC<sub>50</sub> das duas drogas. Grande parte dos supergenes estão anotados como proteínas hipotéticas (81,6%), o que dificulta tal análise. As proteínas mais presentes foram quinases (10 só aumentadas para cetoconazol), seguidas de proteínas ligadoras de RNA (5 só aumentadas para cetoconazol), transportadores de membrana (dois supergenes só aumentados para cetoconazol), peptidases (duas cisteína-peptidases só aumentadas para cetoconazol e uma só para lovastatina).

De forma geral, a comparação da resposta a IC<sub>50</sub> de cada droga reflete padrões bem determinados somente para uma das duas drogas, o que pode ser identificado pela comparação dos enriquecimentos de GO de cada um dos experimentos em separado. Entretanto, as respostas discrepantes entre as drogas acrescentam variação nos dados, o que pode ser útil nas análises de reguloma.

#### 5.3.4.6 DL100: respostas distintas entre as drogas

De forma similar às análises realizadas para a dose IC<sub>50</sub>, foram comparadas as respostas às doses DL100 das duas drogas utilizadas. Usando um FDR de 5% no teste *two class unpaired time course* do programa SAM, foram identificados 4261 genes diferencialmente expressos entre as drogas. Entretanto, somente 973 destes “genes” possuem um FC (*foldchange*) mínimo de 2 vezes, critério estabelecido para propiciar a seleção somente de genes fortemente significativos na sua diferença de resposta entre as drogas.

Esse resultado novamente aponta para a alta reprodutibilidade dos dados DL100, chamando atenção para a presença de respostas intensas e específicas para cada uma das drogas. Embora doses rapidamente letais das drogas desencadeiam resposta similares do ponto de vista morfológico e molecular (intensa vacuolização celular aliada a enriquecimento de processos envolvidos na produção de ATP e estagnação da síntese de RNA), diversos genes são individualmente regulados somente por uma das drogas, apontando para a presença tanto de uma resposta a estresse “químico” estereotipada, bem como para uma resposta

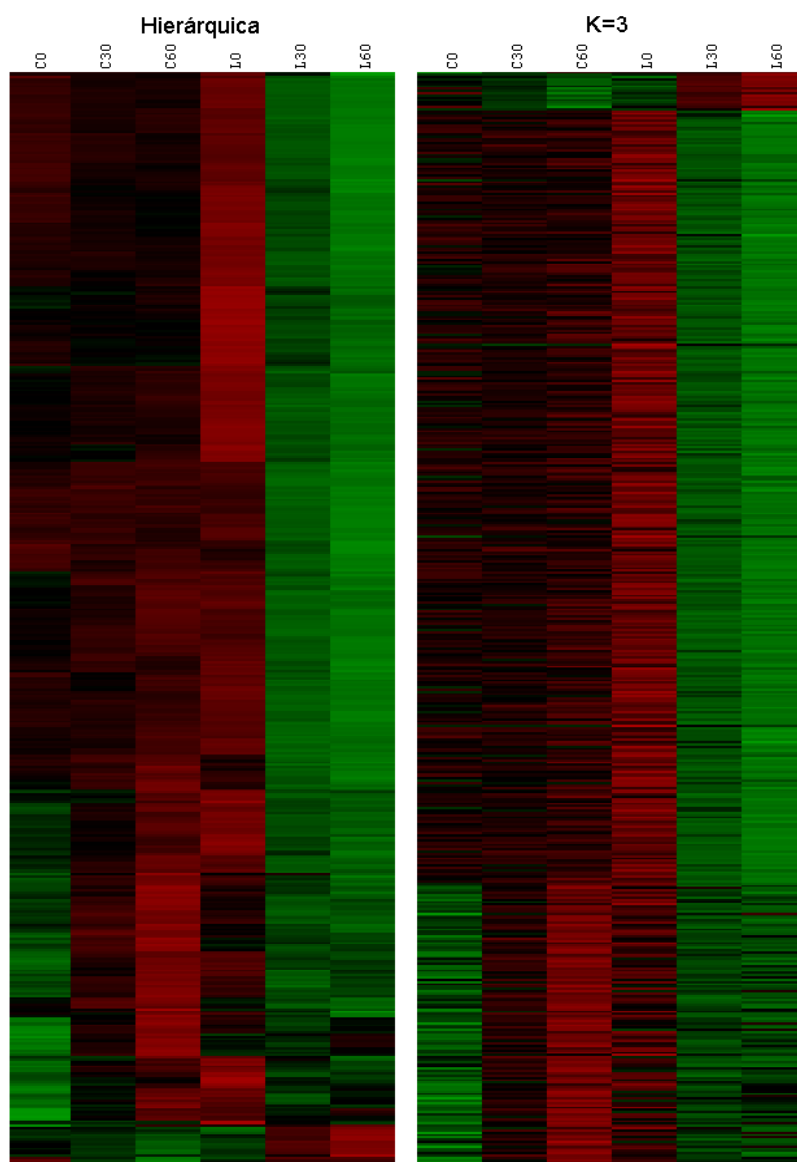
específica para cada tipo de droga. Como já comentado, tais variações biológicas nas respostas a diferentes drogas enriquecem os dados para análises de reguloma.

Retirando a redundância, os 973 genes representam 478 supergenes diferencialmente regulados na comparação das respostas a DL100 das drogas. A visualização destes genes por agrupamento hierárquico (Figura 5.44) permite a identificação de 3 grupos distintos de genes: genes que são diminuídos em resposta a DL100 de lovastatina, mas não alteram a expressão em resposta a DL100 de cetoconazol (340 genes, veja exemplo na Figura 5.45A e B); genes com resposta contrária: aumentam para cetoconazol e caem para lovastatina (122 genes, veja exemplo na Figura 5.45C); e genes aumentados para lovastatina e estáveis (ou diminuídos) para cetoconazol (16 genes, veja exemplo na Figura 5.45D). A visualização desses 3 grupos de genes fica bem mais clara quando usado método de clusterização Kmeans (Figura 5.44, lado direito), método este que exige a determinação, *a priori*, do número de clusters presentes nos dados.

O padrão mais comum de alteração distinta em resposta a DL100 das duas drogas é a diminuição na expressão apenas na resposta a lovastatina; esse resultado está de acordo com o padrão observado na resposta a DL100 em lovastatina, com grande quantidade de regulação negativa na expressão gênica, como já discutido no tópico 5.3.4.4.

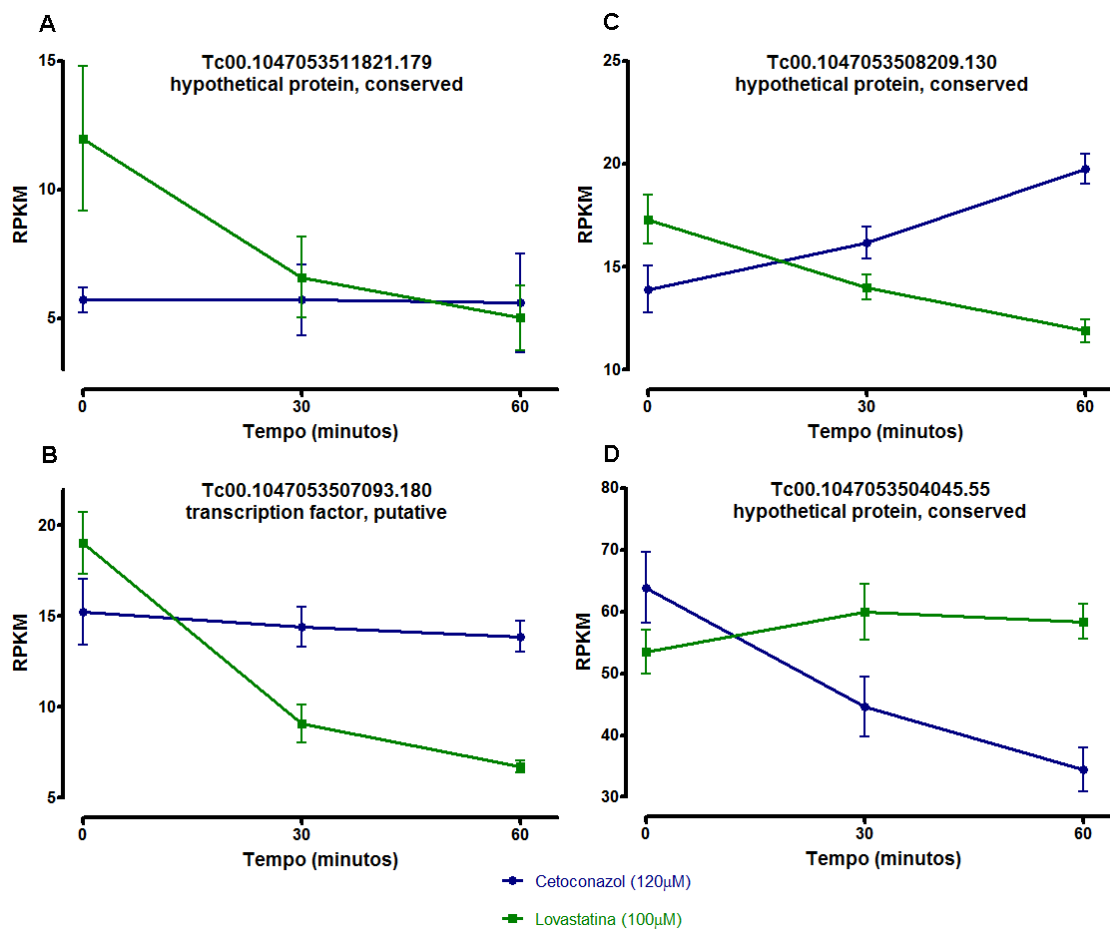
De forma similar ao ocorrido para a comparação entre as drogas na dose IC<sub>50</sub>, a análise de GO não identificou nenhum termo enriquecido nos dados. Assim sendo, destacamos aqui os genes com maior presença na lista de expressão diferencial. Mais de 70% dos supergenes estão anotados como proteínas hipotéticas (352 genes). Diversas quinases (24) alteram a expressão somente a uma das drogas, normalmente só aumentadas para cetoconazol, o que reflete o forte enriquecimento de termos de GO associados à fosforilação para a resposta à DL100 de cetoconazol. Algumas helicases (5) são expressas diferentemente entre as drogas, além de várias peptidases (7), destacando-se a cruzipafina (aumentada só em resposta a lovastatina), uma das peptidases mais abundantes no reservossomo.

Todos os 478 supergenes significativos na comparação das drogas podem ser visualizados na Tabela 8 do ANEXO II.



**FIGURA 5.44 – DL100: COMPARAÇÃO ENTRE AS DROGAS**

A figura mostra duas imagens de clusterização hierárquica dos 478 supergenes seleccionados como diferencialmente regulados na resposta a DL100 das drogas, por dois métodos distintos: clusterização hierárquica (esquerda) e Kmeans (direita) com 3 grupos. Foram usados os valores  $Z$ , considerando as tabelas de contagens de *reads* gerais para ambas as drogas (normalizadas no edgeR). A escala de cores é a mesma da figura Figura 5.39



**FIGURA 5.45 – DL100: GENES DIFERENCIALMENTE REGULADOS ENTRE AS DROGAS**

O lado esquerdo da figura (A e B) contém exemplos de genes de expressão diminuída somente na resposta a DL100 de lovastatina, enquanto no lado direito estão plotados genes com expressão contrária na resposta às duas drogas: aumento em resposta a cetoconazol e diminuição a lovastatina (C) e vice-versa (D). Cada ponto (0-30 minutos) representa a média das réplicas biológicas, com respectivas barras de desvio indicando o desvio-padrão da contagem de *reads* normalizadas pelo tamanho da biblioteca e tamanho do gene (RPKM).

## 6 DISCUSSÕES

### 6.1 REGULAÇÃO DA EXPRESSÃO GÊNICA EM *T. CRUZI*

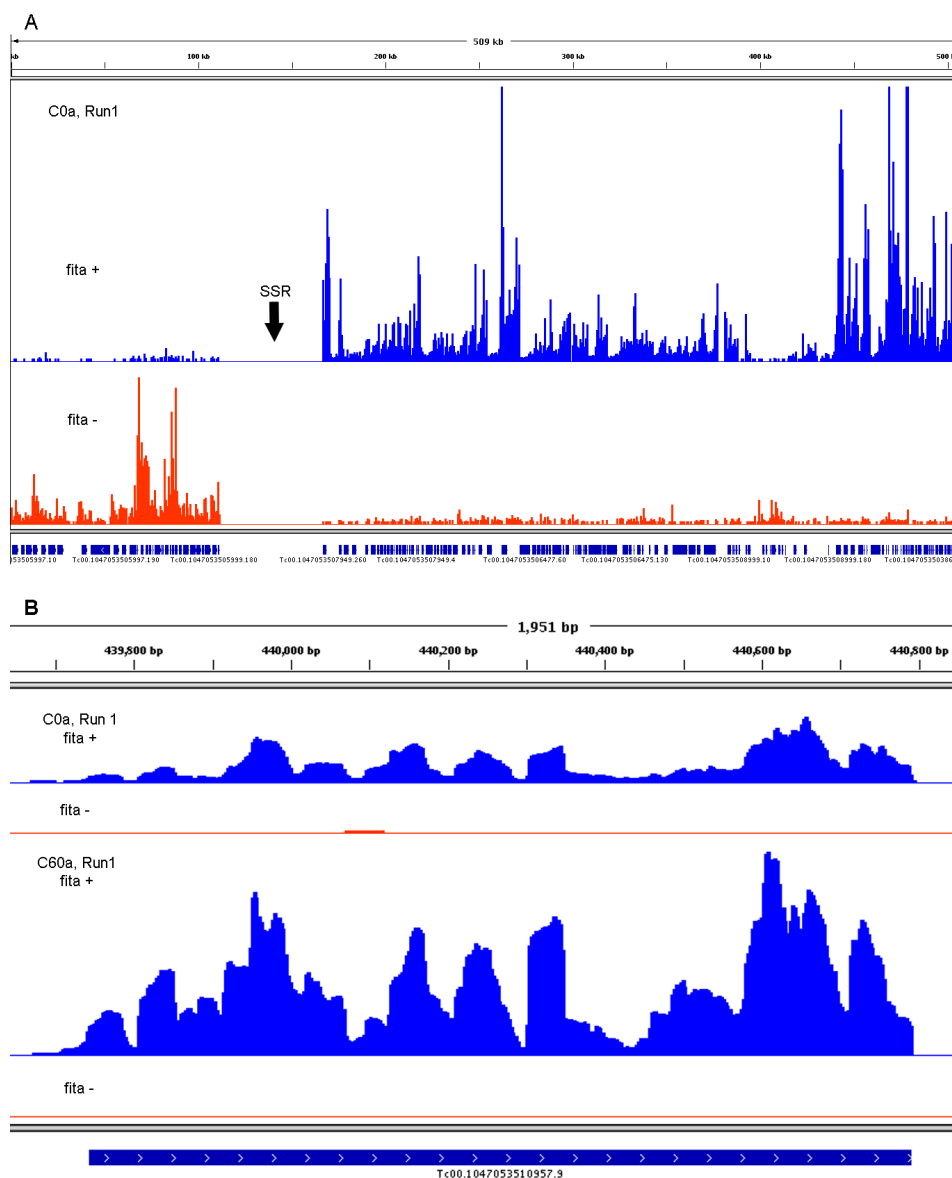
Como previamente exposto, tripanossomatídeos possuem uma organização genômica peculiar: os genes codificadores de proteínas são dispostos em unidades policistrônica contendo dezenas a centenas de genes na mesma fita de DNA (revisto por PALENCHAR & BELLOFATTO, 2006). Tais genes são transcritos em uma mesma molécula de RNA policistrônico, o qual deve ser processado pelo mecanismo de *trans-splicing* de forma a gerar mRNAs maduros passíveis de tradução. Assim sendo, a transcrição não representa um importante ponto de controle da expressão em *T. cruzi*, havendo evidências de que a transcrição da maioria dos genes é ubíqua (revisto por CLAYTON, 2002; revisto por VANHAMME & PAYS, 1995).

Análises de microarranjo realizadas no ICC mostraram que a metaciclógênese (PROBST, 2005) de *T. cruzi*, o número de genes diferencialmente expressos chega a ser 10 vezes maior avaliando-se mRNA polissomal em comparação com mRNA total, sendo muito provável que mecanismos de mobilização polissomal também estejam atuando em outros processos de diferenciação e resposta ao ambiente.

Visto que tal subpopulação do transcriptoma deve estar enriquecida para diferenças na expressão gênica, diferenças estas correlacionadas com os níveis protéicos (ÁVILA *et al.*, 2001; DALLAGIOVANNA *et al.*, 2001; FRAGOSO *et al.*, 2003), utilizou-se somente RNA polissomal no presente trabalho. De fato, observou-se grande parte do genoma de *T. cruzi* regulado em resposta aos inibidores da síntese de esteróis: 10,7% e 5,7% para IC<sub>50</sub> de cetoconazol e lovastatina, respectivamente. Para os experimentos DL100, essa porcentagem foi bem maior: 18% para cetoconazol e 36% para lovastatina (considerando os supergenes). Esse resultado reafirma a utilização de RNA polissomal como um grupo de moléculas muito informativas para análises genômicas em organismos sem controle transcricional da expressão gênica. No entanto, deve-se levar com cautela o grau de mudança observado em DL100, no qual potencialmente a população polissomal pode ser muito mais suscetível a modulações inespecíficas, porém reprodutíveis. Além disso, é importante avaliar a população de mRNA

total nessas mesmas amostras, para confirmar o maior poder de identificação de expressão diferencial. Esses experimentos com mRNA total serão processados futuramente.

Essa alta sensibilidade na detecção de expressão diferencial foi reforçada pelo método de RNA-Seq. A capacidade de detecção de transcritos raros por este método, os quais podem



### FIGURA 6.1 – ALINHAMENTO DOS READS DE RNA-SEQ NO GENOMA

As imagens mostram o alinhamento dos *reads* gerados por RNA-Seq no programa IGV (<http://www.broadinstitute.org/igv>) para a amostra C0a e C60a, réplica técnica 1 (Run1), contra o cromossomo 9 (haplótipo Esmeraldo) de *T. cruzi*. Os alinhamentos nas duas fitas do cromossomo são indicados pelas cores azul (fita +) e vermelha (fita -). A: está evidenciado todo o cromossomo, o que possibilita a identificação de uma região de fita (SSR; indicada com uma seta) e a transcrição das unidades polიცистрônicas. B: detalhe de uma região contendo um gene diferencialmente expresso (Tc00.1047053510957.9, *calpain-like cysteine peptidase*) em resposta a DL100 de Cetoconazol; na parte superior está indicado o alinhamento para a amostra controle (C0a, Run1), enquanto a parte inferior demonstra o alinhamento dos *reads* gerados para a amostra tratada (C60a, Run1); note a maior quantidade de *reads* alinhados para a amostra tratada.



representar grande parte do transcriptoma, aliada a um alcance dinâmico de mais de cinco ordens de magnitude (veja Figura 5.20), possibilita a comparação dos níveis de expressão para praticamente todos os genes do organismo em estudo (HASHIMOTO *et al.*, 2009), de forma que alterações no número de leituras alinhadas em cada gene devem refletir variações biológicas verdadeiras. Além disso, como RNA-Seq baseia-se no alinhamento das leituras geradas contra o genoma de referência (Figura 6.1), é possível visualizar as regiões dos genes sendo expressas e identificar regiões não traduzidas.

A Figura 6.1A mostra o alinhamento das leituras geradas para a amostra controle C0a (Run 1) no cromossomo 9 (haplótipo Esmeraldo) de *T. cruzi*. A utilização, neste trabalho, de um método que mantém a informação de direcionalidade do mRNA (*strand specific*) (veja Figura 4.1), permite a identificação das regiões de troca de fita (SSR, *strand switch region*) na transcrição das unidades policistrônicas, regiões interessantes pois devem representar os sítios de início da transcrição em *T. cruzi* (RESPUELA *et al.*, 2008). A visualização das sequências alinhadas no genoma do parasita (Figura 6.1A) também deixa claro que os genes de uma mesma unidade policistrônica podem apresentar níveis bem distintos de expressão. Além disso, a fragmentação do mRNA antes da síntese de cDNA com adaptadores aleatórios (veja Figura 4.1) possibilita uma análise global do transcriptoma, sem sobrerepresentação de uma das extremidades dos transcritos (veja Figura 6.1B). Dessa forma, a análise conjunta de dados de diversos experimentos de RNA-Seq sendo desenvolvidos em nosso laboratório propiciará a identificação de sítios de *trans-splicing* e poliadenilação, além de permitir a anotação mais acurada do genoma de *T. cruzi*, algo recentemente realizado para o tripanossomatídeo *T. brucei* (SIEGEL *et al.*, 2010; KOLEV *et al.*, 2010).

A análise dos genes diferencialmente expressos em resposta às drogas revelou uma grande quantidade de proteínas ligadoras de RNA. Para o experimento IC<sub>50</sub>, foram observadas 17 supergenes anotados como *RNA-binding protein* diferencialmente expressos para cetoconazol e 8 para lovastatina. Para o experimento DL100 foram 12 para cetoconazol e 9 para lovastatina. É provável que estas proteínas estejam regulando a estabilidade e associação a polissomos de seus parceiros, orquestrando a regulação pós-transcricional da expressão gênica em resposta às drogas utilizadas. Esta hipótese pode ser corroborada pela grande quantidade de helicases detectadas como diferencialmente expressas, diversas da família DEAD, proteínas importantes no metabolismo de RNA, como transporte, tradução e decaimento de RNA (GUSTAFSON & WESSEL, 2010). Além disso, a grande quantidade de

quinases observadas como reguladas, principalmente em resposta a cetoconazol, pode estar relacionada a tais mecanismos pós-transcricionais da expressão, funcionando como transdutores de sinal frente ao estresse desencadeado pelas drogas.

## 6.2 INIBIÇÃO DA SÍNTESE DE ESTERÓIS: DE GENÔMICA A BIOLOGIA CELULAR

A utilização de técnicas clássicas de biologia celular (microscopia e citometria) aliada à análise em larga escala da expressão gênica, permitiram um entendimento global das respostas de *T. cruzi* frente à inibição da síntese de esteróis endógenos.

Como previamente demonstrado, *T. cruzi* tratado com algumas drogas inibidoras da síntese de esteróis mostra desorganização das membranas mitocondriais seguido de intenso inchaço e perda do conteúdo da matriz (revisto por DE SOUZA & RODRIGUES, 2009). Tais alterações também foram observadas no presente trabalho (veja Figura 5.11 e Figura 5.12), o que, de forma secundária, confirma a inibição da via de síntese de esteróis. Entretanto, nossas análises de microscopia eletrônica em resposta a IC<sub>50</sub> das drogas revelaram que a principal alteração na mitocôndria é a proliferação e fragmentação da organela (veja setas na Figura 5.11 e Figura 5.12). Diferentemente de um simples inchaço mitocondrial (quando a matriz mitocondrial aumenta de volume expondo uma matriz mitocondrial pouco elétron-densa) parece haver proliferação da membrana mitocondrial interna, sendo possível visualizar a anormal quantidade de cristas mitocondriais (veja Figura 5.12C). Em alguns casos, é possível observar tal padrão ramificado de membranas em contato direto com o cinetoplasto (veja Figura 5.12D). O aumento da mitocôndria em resposta as drogas é corroborado pela maior marcação com o corante fluorescente Rhodamina 123 (veja Figura 5.13). Assim sendo, parece que estas membranas são funcionais do ponto de vista bioquímico, de forma que o estresse gerado pelas drogas deve induzir uma ativação da cadeia transportadora de elétrons.

De fato, observou-se genes da cadeia transportadora de elétrons aumentados em resposta a ambas as drogas. Para cetoconazol, foram detectados como diferencialmente aumentadas uma subunidade do complexo citocromo c oxidase (Tc00.1047053511389.110, cytochrome c oxidase VIII (COX VIII), aumento de ~2,5 vezes) e uma proteína envolvida na montagem de tal complexo (Tc00.1047053511211.70 cytochrome oxidase assembly protein, aumento de ~2,3 vezes). Para lovastatina, identificou-se outra subunidade do complexo

citocromo c oxidase como aumentada (Tc00.1047053508651.20 cytochrome c oxidase subunit 10, aumento de 1,62 vezes).

Também observou-se aumento de expressão de diversos transportadores de membrana para moléculas que podem estar relacionadas com, ou serem substrato, em vias de síntese de ATP: cátion, açúcar (hexose), aminoácidos, nucleosídeos, metais e fosfato. Além disso, em resposta a IC<sub>50</sub> de lovastatina foi observado o enriquecimento do processo biológico de GO “processo metabólico de NADP”. Dessa forma, parece que o aumento da membrana mitocondrial interna observada por microscopia eletrônica, e corroborada por citometria de fluxo (Figura 5.13), pode estar correlacionado com um aumento na síntese de ATP por fosforilação oxidativa. Além disso, a localização mitocondrial de enzimas relacionadas à síntese de esteróis, incluindo HMGR (PEÑA-DIAZ *et al.*, 2004; revisto por DESOUZA & RODRIGUES, 2009), chama atenção para o papel dessa organela na síntese de esteróis em *T. cruzi*. De fato, análises da composição mitocondrial de *T. cruzi* indicam a presença de esteróis endógenos na membrana interna (RODRIGUES *et al.*, 2001).

Outra alteração morfológica muito clara em resposta a IC<sub>50</sub> das duas drogas foi o aumento dos reservossomos (Figura 5.11 e Figura 5.12). Experimentos de microscopia confocal (Figura 5.9) e citometria de fluxo (Figura 5.8) demonstraram que tais vesículas possuem um baixo pH, corroborando a hipótese que as mesmas representem reservossomos. Esse é o primeiro trabalho que destaca tal impacto dos inibidores da síntese de esteróis nessas organelas únicas desse parasita, complementado com modulações da expressão gênica de proteínas potencialmente participantes dessa organela. Embora um trabalho recente com cetoconazol tenha mostrado um resultado similar (SANTA-RITA *et al.*, 2005) em relação ao aumento dos reservossomos, pouca ênfase é dada a essa observação e não há dados de expressão diferencial das proteínas dessa organela.

Portanto, analisamos os dados de RNA-Seq em busca de padrões de expressão possivelmente relacionados com a proliferação dessas organelas. Os reservossomos possuem alta concentração de serina carboxipeptidase e cisteína-peptidases pertencentes ao clã CA, família C1 (família de peptidases lisossomais conhecidas como catepsinas): as bem caracterizadas cruzipainas 1 e 2 (catepsina tipo L, CAZZULO *et al.*, 2002), além de outra catepsina L e uma catepsina B (KOSEC *et al.*, 2006b). Para lovastatina, foram detectadas 2 cisteína-peptidases da família C1 com expressão aumentada na resposta à dose IC<sub>50</sub> (uma cruzipaina e a outra catepsina L de *T. cruzi*, aumento de pelo menos 1,9 vezes), além de um precursor de alfa-manosidase (Tc00.1047053506195.120 *lysosomal alpha-mannosidase*

*precursor*, aumento de mais de 2 vezes). Para cetoconazol, detectamos um aumento de quase duas vezes do gene de serina carboxipeptidase (CBP1, Tc00.1047053508671.20) e um gene codificador de uma possível proteína de membrana do reservossomo (Tc00.1047053505555.10, *lysosomal membrane glycoprotein*, aumento de ~2,6 vezes). O aumento da atividade de cisteína-peptidase foi apoiado pelo enriquecimento da função molecular de GO “atividade endopeptidásica tipo cisteína” em resposta à IC<sub>50</sub> de lovastatina.

Considerando apenas as cisteínas peptidases do tipo catepsina, existem 4 supergenes (duas cruzipaínas, uma catepsina L e uma B) no genoma de *T. cruzi*, dos quais dois tem expressão aumentada em resposta à lovastatina. Assim sendo, os dados de supergenes de expressão aumentada para lovastatina possuem mais de 20 vezes mais catepsinas em relação a porcentagem presente no genoma do parasita (1,12% *versus* 0,05%).

Pelos resultados acima apresentados, parece que o aumento dos reservossomos em resposta a IC<sub>50</sub> dos IBEs (pelo menos para lovastatina) pode ser demonstrado tanto em nível celular quanto molecular.

Recentemente foi publicada uma análise proteômica dos reservossomos de *T. cruzi* (SANT’ANNA *et al.*, 2009), no qual foram identificadas cerca de 700 proteínas residentes nessa organela. Devido à hipótese aqui apresentada de um aumento dos reservossomos em resposta a IC<sub>50</sub> dos inibidores da síntese de esteróis, as proteínas identificadas por Sant’Anna e colaboradores (2009) foram correlacionadas com os dados de expressão diferencial de RNA-Seq. Embora os dados de Sant’Anna e colaboradores (2009) não representem uma análise definitiva da composição protéica dos reservossomos, por serem dados ruidosos, com um componente de contaminação potencialmente importante, a correlação com os dados presentes pode propiciar a identificação de genes interessantes. Além disso, como os reservossomos adquirem uma morfologia não usual, é possível que sua composição seja distinta da organela em epimastigotas normais, o que prejudicaria a correlação dos dados presentes de RNA-Seq com a proteômica de Sant’Anna (2009).

Para cetoconazol, 98 genes (IDs) associados ao reservossomo foram detectados como diferencialmente expressos, 38 deles com expressão aumentada. Para lovastatina, foram identificados 79 genes (31 com expressão aumentada).

A seguir, os genes diferencialmente expressos à IC<sub>50</sub> das drogas, e com localização protéica associada ao reservossomo (SANT’ANNA *et al.*, 2009), foram agrupados por clusterização hierárquica (Figura 6.2 e Figura 6.3). A visualização da descrição funcional dos

mesmos mostra que os genes aumentados em resposta as drogas e associados aos reservossomos (SANT'ANNA *et al.*, 2009), possuem diversas anotações funcionais.

Chamou atenção na resposta a cetoconazol o aumento de duas enzimas envolvidas na síntese de esteróis e localizadas no reservossomo: esterol-24C-metil-transferase e tiolase (detalhes no tópico a seguir). Além disso, todos os transportadores ABC detectados como aumentados em resposta a cetoconazol têm localização reservossomal segundo Sant'Anna e colaboradores (2009). O envolvimento dessa organela na via endocítica de *T. cruzi* é apoiado pelo aumento de uma proteína marcadora de vesículas de transporte (Tc00.1047053510187.270, COP-coated vesicle membrane protein erv25 precursor).

Para lovastatina, chama atenção o aumento de expressão do precursor de cruzipáina, uma das peptidases mais abundantes no reservossomo, além de uma cisteína-peptidase tipo calpaína. Como inferido anteriormente, o precursor de alfa-manosidade possivelmente está presente no reservossomo (SANT'ANNA *et al.*, 2009), de forma que seu aumento em resposta à lovastatina reafirma o aumento dessa organela na resposta às drogas. O aumento de uma proteína envolvida na síntese de tripanotona, triparedoxina peroxidase, chama atenção para o papel dessa organela na resposta ao estresse químico gerado por lovastatina.

Embora tenham sido detectados cerca de 35 genes (IDs) de localização protéica reservossomal aumentados em resposta a cada uma das drogas, a comparação dos mesmo mostra que somente 10 deles são aumentados na resposta a ambas as drogas (veja Tabela 6.1). Estes genes são interessantes pois podem ser marcadores reservossomais; destaca-se a presença de uma proteína de resistência a drogas, aumentada mais de 3 vezes em resposta a cetoconazol e 2,1 vezes em resposta a lovastatina. A presença de uma peptidase destaca o papel dessa organela na via endocítica e confirma a importância dessa classe de enzimas como componente reservossomal. Chama atenção o aumento de expressão em comum de

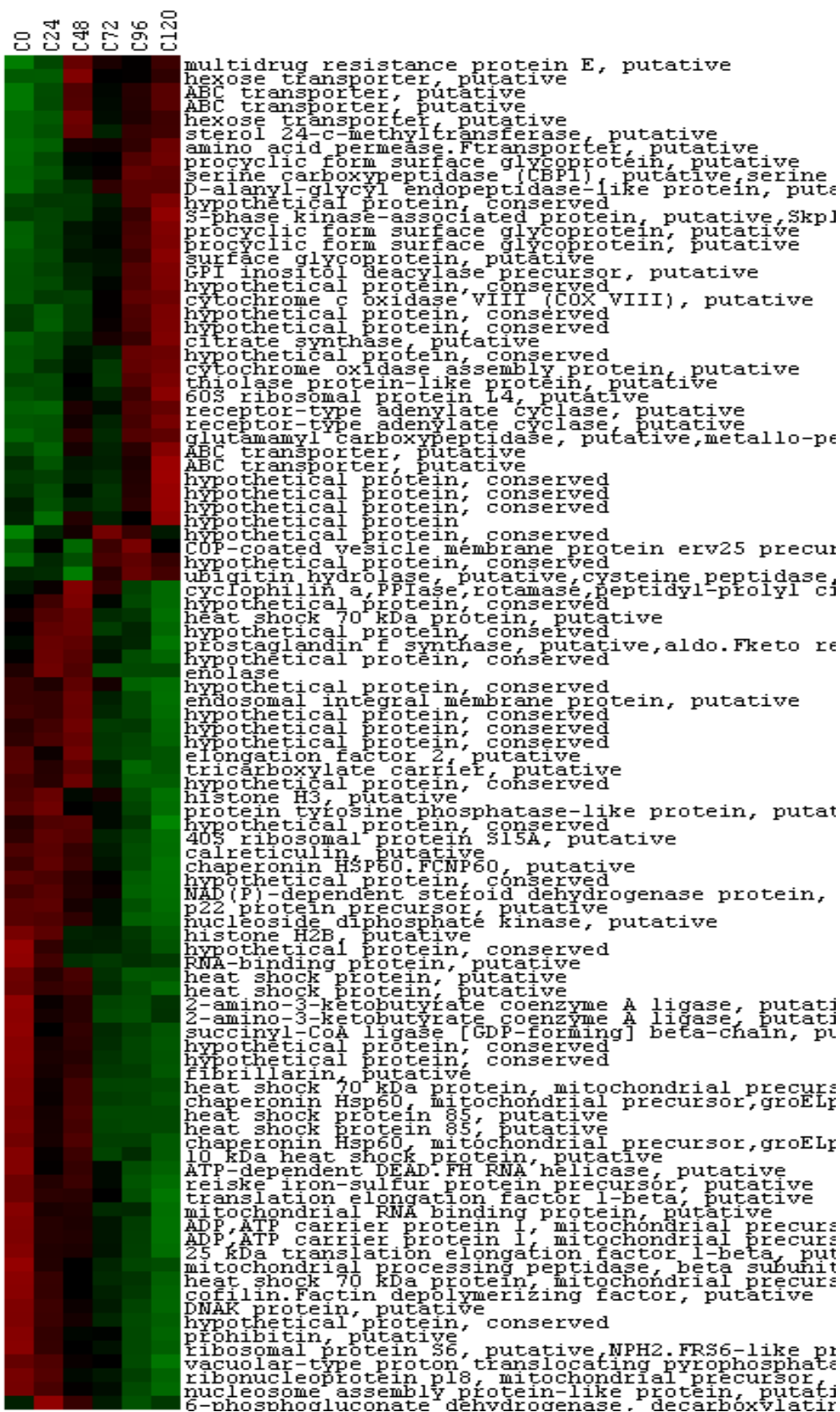
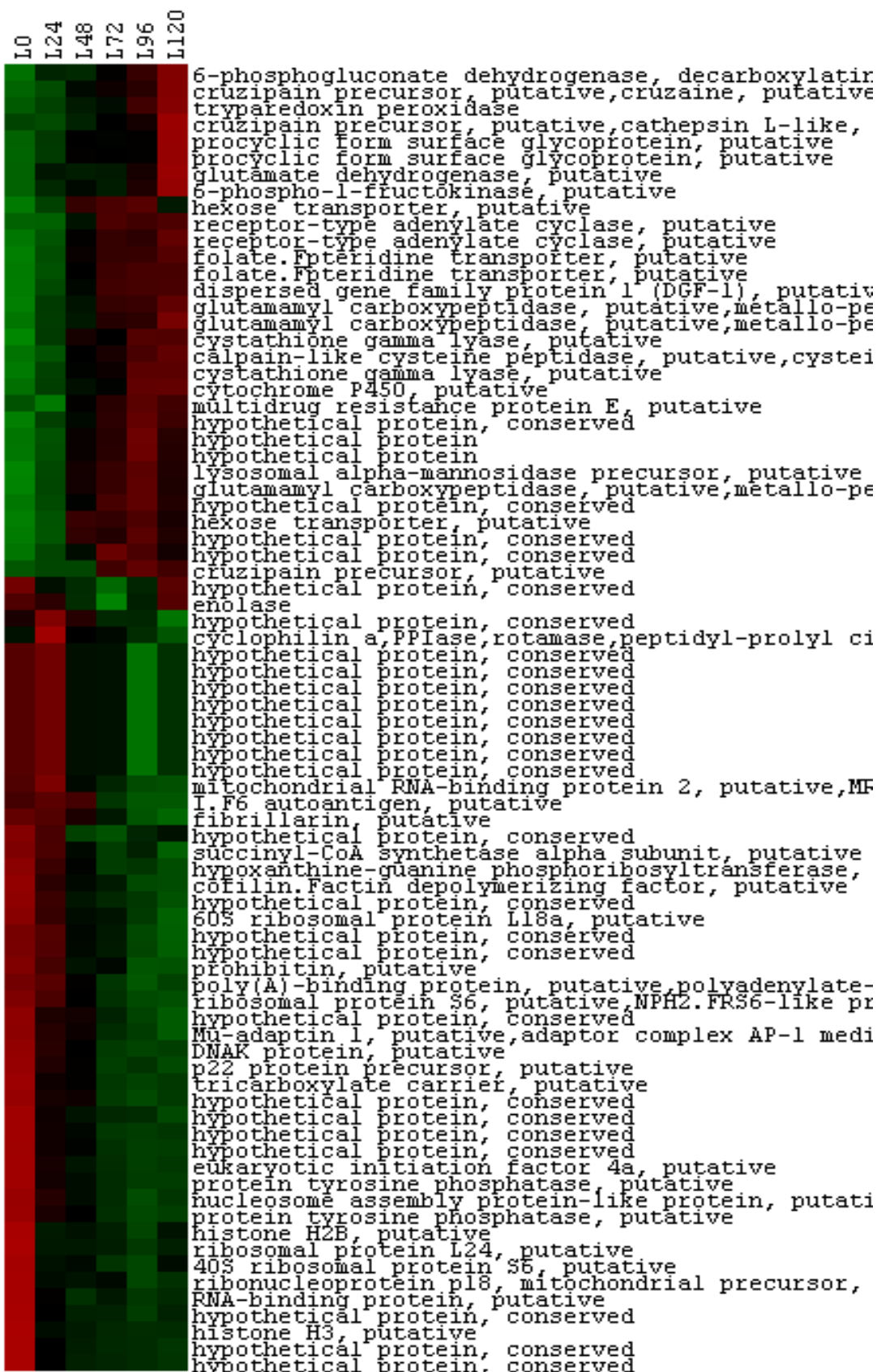


FIGURA 6.2 – IC<sub>50</sub> CETOCONAZOL: PROTEÍNAS RESERVOSSOMAS DIFERENCIALMENTE EXPRESSAS

Estão clusterizadas os 98 genes diferencialmente expressos em resposta a Cetoconazol e associados aos reservossomos segundo os dados de Sant'Anna e colaboradores (2009).



**FIGURA 6.3 – IC<sub>50</sub> LOVASTATINA: PROTEÍNAS RESERVOSSOMAS DIFERENCIALMENTE EXPRESSAS**

Estão clusterizadas os 79 genes diferencialmente expressos em resposta a Lovastatina e associados aos reservossomos segundo os dados de Sant'Anna e colaboradores (2009).

**TABELA 6.1 – PROTEÍNAS RESERVOSSOMAS AUMENTADAS PARA IC<sub>50</sub> DE AMBAS AS DROGAS**

Locus_Tag	Descrição
Tc00.1047053511391.120	glutamamyl carboxypeptidase, putative, metallo-peptidase, clan MH, family M18, putative
Tc00.1047053506355.10	hexose transporter, putative
Tc00.1047053508231.9	hexose transporter, putative
Tc00.1047053508137.20	hypothetical protein, conserved
Tc00.1047053510317.7	hypothetical protein, conserved
Tc00.1047053508965.14	multidrug resistance protein E, putative
Tc00.1047053507711.100	procyclic form surface glycoprotein, putative
Tc00.1047053509637.36	procyclic form surface glycoprotein, putative
Tc00.1047053507467.10	receptor-type adenylate cyclase, putative
Tc00.1047053510581.20	receptor-type adenylate cyclase, putative

transportadores de hexose e receptores de tipo adenilato ciclase. A presença de um possível proteína de membrana plasmática (*procyclic form surface glycoprotein*) reafirma o papel dessa organela como alvo final no tráfego de vesículas a partir da membrana plasmática.

A presença de contaminantes nos dados de proteômica fica mais evidente quando analisados os genes de expressão diminuída em resposta aos IBEs e de localização protéica reservossomal segundo Sant'Anna (2009). Nesse grupo encontramos duas histonas (H2B e H3), uma proteína de montagem do nucleossomo, uma proteína nucleolar (fibrilarina), uma enzima da via glicolítica (enolase), dentre outros. Assim sendo, este grupo de genes parece ser governado por contaminantes dos dados de proteômica, o que torna os dados acima não confiáveis.

O aumento dos reservossomos em resposta a IC<sub>50</sub> dos inibidores da síntese de esteróis deve estar relacionado à função dessa organela na síntese de esteróis. Como já descrito, esta organela é rica em lipídios, incluindo ergosterol e precursores, o que somado com a demonstração recente da presença de enzimas essenciais na síntese de ergosterol, 24-SMT e tiolase (SANT'ANNA *et al.*, 2009) e 14-alfa-demetilase (LEPESHEVA *et al.*, 2010), justifica o aumento significativo dessa organela frente à inibição da síntese de ergosterol em *T. cruzi*. Essa hipótese é corroborada pelo aumento na expressão de 24-SMT e tiolase em resposta a cetoconazol no presente trabalho. Por outro lado, o aumento dos reservossomos pode ser uma resposta indireta à depleção dos esteróis endógenos, sendo que o aumento de atividade endocítica da célula justificaria o aumento dos reservossomos; de fato, células tratadas com IBEs demonstram quantidades exacerbadas de esteróis exógenos, incluindo colesterol. Além disso, não podemos excluir a hipótese de que as enzimas da síntese de esteróis entram nos reservossomos no caminho para outros compartimentos celulares.



### 6.3 REGULAÇÃO DA SÍNTESE DE ESTERÓIS EM *T. CRUZI*

Em mamíferos e leveduras, o passo limitante na via de síntese de esteróis é a conversão de HMG-CoA a mevalonato, a reação catalisada por HMGR (GIMPL *et al.*, 2002). A regulação da atividade desta enzima é mediada por um ativador transcricional do gene que codifica HMGR, denominado SREBP (*sterol regulatory element-binding protein*). Quando os níveis de esteróis intracelulares estão baixos, o precursor de SREBP, localizado na face citosólica da membrana do retículo endoplasmático, migra para o complexo de Golgi e é clivado, liberando uma subunidade que é translocada para o núcleo para ativar a transcrição do gene de HMGR (GIMPL *et al.*, 2002; HUGHES *et al.*, 2005). Um estudo utilizando diversos inibidores da biossíntese de ergosterol em *T. cruzi* mostrou que HMGR não parece ser regulada nesse parasita, tanto em nível transcricional como protéico (HANKINS *et al.*, 2005). Por outro lado, a inibição da via induz um aumento de cerca de 10 vezes nos níveis de mRNA da enzima 14 $\alpha$ -desmetilase, em associação com um aumento de cerca de 2,5 vezes no nível protéico da mesma (HANKINS *et al.*, 2005), regulação esta que parece ser mediada pela região 3'-UTR do gene. Dados mais recentes indicam um aumento de quase 8 vezes dessa enzima em resposta a Posaconazol (LEPESHEVA *et al.*, 2010). Tal resultado prediz que a regulação da via de síntese de esteróis em *T. cruzi* é significativamente diferente de mecanismos bem caracterizados em outros organismos.

A análise da inibição da síntese de esteróis através da técnica de RNA-Seq possibilitou uma visualização inédita a respeito da regulação e interconexões desta via em *T. cruzi*. Como o efeito deste tipo de inibidores é tardio (URBINA *et al.*, 1993; URBINA *et al.*, 1988), sendo necessário que os estoques intracelulares de esteróis endógenos sejam depletados antes de uma resposta de alteração morfológica, os experimentos de IC<sub>50</sub> são os mais indicados para análise da regulação da via, visto que a morte induzida pela dose DL100 deve ser causada por algum mecanismo secundário desconhecido (que não a síntese de esteróis, FLORIN-CHRISTENSEN *et al.*, 1990).

Os genes de *T. cruzi* associados à via de síntese de esteróis não estão bem anotados no genoma (EL-SAYED *et al.*, 2005); assim sendo, de forma selecionar todos os genes relacionados com tal via, utilizamos a via bem descrita de *S. cerevisiae* como base para busca dos genes de *T. cruzi* por similaridade protéica com a ferramenta BLASTp. Os genes mais significativos foram selecionados e utilizados para reconstrução da via no programa GenMapp (DAHLQUIST *et al.*, 2002). Tal programa permite a visualização de dados de expressão

gênica em vias metabólicas, permitindo a identificação de genes relacionados ao controle do mecanismo em estudo.



**FIGURA 6.4 – IC<sub>50</sub> CETOCONAZOL: REGULAÇÃO DA SÍNTESE DE ESTERÓIS**

As etapas enzimáticas da via de síntese de esteróides em *T. cruzi* são representadas por flechas verticais, com respectivas enzimas nomeadas dentro do retângulo. Os substratos e produtos das reações estão especificados acima e abaixo de cada reação, respectivamente. Enzimas detectadas como diferencialmente expressas nos dados de RNA-Seq são coloridas com os padrões de cores especificados na legenda, onde vermelho representa aumento de expressão e verde diminuição.

Quando considerado somente os genes diferencialmente expressos com FDR5% e FC1,5, detectamos somente 3 enzimas com expressão aumentada em resposta a cetoconazol:

tiolase, poliprenil sintase e 24-SMT (veja Figura 6.4). A enzima HMGR, principal ponto de regulação da síntese de colesterol em mamíferos, não demonstrou alteração na expressão no presente trabalho.

Tiolase é a enzima responsável pela síntese de acetoacetil-CoA a partir de duas moléculas de acetil-CoA, representando a primeira de três reações necessária para a síntese de mevalonato (via do Mevalonato). A segunda enzima detectada como aumenta em resposta a cetoconazol foi poliprenil sintase, a qual condensa duas moléculas de 5 carbonos (isopentenil difosfato e dimetilalil difosfato) para formar um isopreno de 10 carbonos denominado geranyl difosfato. Estas enzimas estão muito a montante na via, e são compartilhadas por diferentes processos metabólicos na célula; o composto farnesil-PP (farnesil difosfato), o qual representa o produto final da via dos isoprenos, é o substrato de enzimas que catalisam a primeira reação comprometida com as vias de síntese de ubiquinonas, dolicol, hemeA, proteínas prenildadas e biosíntese de esteróis (revisto por DE SOUZA & RODRIGUES, 2009). Assim sendo, embora as enzimas tiolase e poliprenil sintase sejam essenciais na síntese de esteróis, o aumento na expressão das mesmas pode estar correlacionado a outros mecanismos além da síntese de ergosterol (via do mevalonato e via dos isoprenos), de forma que a determinação destas enzimas como pontos de controle da síntese de esteróis necessita uma análise mais criteriosa, mesmo embora a correlação seja interessante.

Por outro lado, a enzima 24-SMT está exclusivamente relacionada à síntese de esteróis, sendo um dos pontos finais desta via em tripanossomatídeos e fungos (ROBERTS *et al.*, 2003). Assim sendo, o aumento observado de expressão da mesma (veja Figura 6.4) indica que tal enzima deve representar um ponto de regulação da síntese de ergosterol em *T. cruzi*. A demonstração no presente trabalho de regulação desta enzima frente a inibição da síntese de esteróis, associada a ausência desta enzima em mamíferos, a torna um alvo quimioterápico interessante no tratamento da doença de Chagas (GIGANTE *et al.*, 2010).

Chama atenção que algumas enzimas tem expressão diminuída em resposta a IC<sub>50</sub> de cetoconazol. Duas delas representam enzimas de passos iniciais da via (veja Figura 6.4), de forma que a diminuição na expressão das mesmas não deve estar relacionado com um inibição da via de síntese de esteróis. A terceira enzima de expressão diminuída (esterol desidrogenase-NAD(P) dependente) representa uma etapa a jusante na via; sendo que sua diminuição foi uma resposta inesperada considerando a inibição da síntese de esteróis endógenos por cetoconazol. Entretanto, tal resultado pode ser uma erro na identificação dos

supergenes relacionados a via de síntese de esteróis através de similaridade protéica contra as enzimas de *S. cerevisiae*, ponto que será explorado mais profundamente em outro trabalho.

Como a enzima 24-SMT mostrou-se regulada em nível de RNA polissomal, atualmente estamos analisando a expressão da mesma em nível protéico. Experimentos preliminares de *western blot*, a partir de extratos protéicos de *T. cruzi* em resposta a IC<sub>50</sub> dos inibidores, indicam um aumento de pelo menos 50% na quantidade de 24-SMT após 48 horas de exposição à droga (dados não mostrados). Devido à presença de apenas uma réplica biológica até o momento, esses dados são considerados muito preliminares.

Como já discutido, a enzima 14-alfa-demetilase é fortemente regulada em *T. cruzi* em resposta à cetoconazol, com níveis de RNA (total) mais de cinco vezes aumentado após 24 horas de exposição. Além disso, Lepesheva e colaboradores (2010) demonstraram que esta enzima tem expressão fortemene aumentada em resposta aos azóis Posaconazol e fluconazol. Embora a dose de cetoconazol no trabalho de Hankins e colaboradores (2005), 25 µM, seja muito similar à aqui utilizada, não detectou-se alteração significativa da expressão dessa enzima no presente trabalho. É possível que estes resultados distintos devam-se aos diferentes métodos utilizados, visto que em Hankins e colaboradores (2005) foi utilizado PCR em tempo real. A utilização de apenas duas réplicas também pode estar relacionada à não detecção desta enzima como diferencialmente expressa no presente trabalho, embora o padrão visualizado diretamente não demonstre tendência de expressão diferencial.

Diferentemente de cetoconazol, a exposição à lovastatina não induz um aumento na expressão de 24-SMT em nível de RNA polissomal, mesmo embora os dados preliminares de *western blot* indiquem um leve aumento desta proteína em resposta a lovastatina. Além disso, quando os dados de RNA-Seq em resposta a IC<sub>50</sub> de lovastatina são plotados no mapa da via de ergosterol construído no presente trabalho (Figura 6.4), apenas uma enzima, farnesiltransferase, apresenta alteração significativa na expressão (diminuição).

O que chama atenção é que segundo os dados de Sant'Anna e colaboradores (2009), a enzima 24-SMT localiza-se nos reservossomos de *T. cruzi*. Juntamente com dados recentes que indicam que a enzima 14-alfa-desmetilase também está presente nesta organela (LEPESHEVA *et al.*, 2010), o presente trabalho aponta para uma possível correlação desta organela com a via de síntese de esteróis em *T. cruzi*.

#### 6.4 DL100: ANÁLISE GENÔMICA DA MORTE CELULAR EM *T. CRUZI*

A utilização de doses das drogas capazes de matar todos os parasitas em cultura em menos de 24 horas (DL100) é uma interessante oportunidade para estudar a morte celular em *T. cruzi* em escala genômica.

A morte celular por necrose é morfológicamente caracterizada por um aumento no volume celular (oncosis), inchaço de organelas, ruptura da membrana plasmática e subsequente perda de componentes intracelulares (revisado por KROEMER *et al.*, 2009). Além disso, o processo envolve alterações dramáticas na mitocôndria (incluindo despolarização). Devido à observação das características morfológicas citadas acima, aliada a completa despolarização mitocondrial frente ao tratamento com a DL100 das drogas (Figura 5.15D), o presente trabalho aponta para uma morte celular necrose-like em resposta a altas doses dos inibidores. A ocorrência de apoptose parece improvável pela ausência da exposição de fosfatidilserina na face externa da membrana plasmática (Figura 5.15A, B e C), processo já observado na morte celular em *T. cruzi* induzida por soro (KOSEC *et al.*, 2006) ou veneno de jararaca (DEOLINDO *et al.*, 2005).

A possível morte por necrose em resposta aos IBEs é corroborada pela observação da lise dos reservossomos de *T. cruzi* em resposta a DL100 de cetoconazol (Figura 5.16). Experimentos de citometria de fluxo com o corante acidotrófico laranja de acridina indicaram a presença de duas populações, uma com maior sinal e outra com menor sinal em relação ao controle (veja Figura 5.15E), o que pode ser explicado pelo aumento dos reservossomos seguido de sua lise (veja tópico 5.2.2).

Embora a longo tempo considerada acidental, evidências mostram que a morte por necrose é um mecanismo regulado por um conjunto de vias de transdução de sinal e mecanismos de catabolismo. Segundo as recomendações do comitê de nomenclatura de morte celular de 2009 (KROEMER *et al.*, 2009), tal processo é caracterizado, em nível molecular, pela ativação de cisteína-peptidases do tipo calpaína e catepsinas.

Calpaínas são peptidases dependentes de cálcio (BARRET & RAWLINGS, 2001). Assim sendo, um passo inicial e crucial para a ativação da necrose é um aumento repentino nos níveis desse cátion no citoplasma. Tal alteração parece ser o gatilho para a ativação massiva de calpaínas, as quais degradam proteínas estruturais e regulatórias, incluindo proteínas da membrana lisossomal. A consequente liberação de cisteína-peptidases lisossomais (catepsinas) para o citoplasma, devido à permeabilização do lisossomo, é um passo crítico e

irreversível no processo de necrose (GIUSTI *et al.*, 2009; revisto por TURK & TURK, 2009). A observação de lise dos reservossomos em resposta a DL100 de cetoconazol pode estar relacionado a liberação de cruzipaína e demais catepsinas reservossomais para a citoplasma, levando à completa degradação dos parasitas.

Análises do genoma de *T. cruzi* mostraram que este parasita contém 24 genes que codificam para possíveis calpaínas (ERSFELD *et al.*, 2005). Entretanto, somente 15 deles possuem o domínio II catalítico, além de nenhuma destas possuir os cinco resíduos de aminoácidos relacionados a ligação de cálcio para ativação da atividade proteolítica. Portanto, tais calpaínas podem ser não funcionais, ou a ativação por cálcio ocorre de forma desconhecida. Nenhuma destas calpaínas foi caracterizada, de forma que a possível função de protease não está demonstrada para nenhuma delas. Além das calpaínas, *T. cruzi* possuiu alguns genes para catepsinas: as bem caracterizadas cruzipaínas 1 e 2 (CAZZULO *et al.*, 2002), além de outra catepsina L e uma catepsina B (KOSEC *et al.*, 2006b).

Há uma grande quantidade de supergenes anotados como cisteína-peptidases de expressão aumentada em resposta a DL100 de ambas as drogas. Tal aumento de expressão de peptidases foi corroborado pelo enriquecimento do processo biológico “proteólise” em resposta a cetoconazol (Tabela 5.11), e da função molecular “cisteína-peptidase” para lovastatina (Tabela 5.13), comprovando a validade estatística do aumento de expressão de cisteína-peptidases. Quando considerado apenas os supergenes anotados como calpaínas, foram 10 genes aumentados para cetoconazol e 9 para lovastatina; 7 destes genes são aumentados para ambas as drogas. Se considerado apenas os genes de expressão aumentada em resposta as drogas, os genes de calpaínas estão 4,7 vezes mais presentes na resposta a cetoconazol (10 supergenes de 704 aumentados) em comparação ao genoma do parasita (24 dos 7946 supergenes analisados), enriquecimento este altamente significativo para teste  $X^2$  ( $p < 4 \times 10^{-8}$ ). Para lovastatina, quando considerados apenas os genes de expressão aumentada o enriquecimento é ainda maior, 5,5 vezes (9 supergenes dos 542 aumentados,  $p = 4 \times 10^{-6}$ ). Para cetoconazol, ainda foi observado o aumento na expressão das catepsinas B e L de *T. cruzi*, o que representa um aumento de cerca de 11 vezes na representatividade destas enzimas, dentre os genes de expressão aumentada, em relação ao genoma ( $p < 4 \times 10^{-8}$ ).

A ativação das calpaínas por influxo repentino de cálcio é um processo principalmente regulado em nível pós-traducional, visto que os níveis protéicos ou de mRNA de calpaínas correlacionam-se fracamente com a atividade das mesmas (FARR & BERGER, 2010). Entretanto, o aumento dos níveis protéicos destas enzimas podem favorecer a resposta

catastrófica envolvida na necrose, sendo que diversos estudos reportam a expressão diferencial de cisteína proteases no tecido nervoso afetado por necrose (ARTAL-SANZ & TAVERNARAKIS, 2005). Assim sendo, o observado aumento na associação a polissomos de mRNA de calpaínas e catepsinas de *T. cruzi*, em resposta ao tratamento, pode estar associado a um mecanismo de regulação da necrose que reflete as alterações pós-traducionais de ativação das peptidases em um aumento na tradução destas proteínas. Se tal via de ativação de necrose de fato ocorre em *T. cruzi*, o influxo de cálcio citoplasmático deve ser desencadeado de maneira a ativar as peptidases dependentes de cálcio. Outro dado interessante é o aumento de uma ATPase de transporte de cálcio em resposta a DL100 de cetoconazol (Tc00.1047053504141.9, *calcium-translocating P-type ATPase*, aumento de 2,7 vezes)

Embora a atividade peptidásica das calpaínas de *T. cruzi* seja duvidosa (ERSFELD *et al.*, 2005), o aumento de expressão de diversas delas durante a necrose induzida por IBEs chama atenção para um possível papel funcional das mesmas nesse processo, seja como peptidase ou como proteína reguladora. A extensa degradação celular observada por microscopia eletrônica, incluindo a lise dos reservossomos e possível liberação das catepsinas, chama atenção para a existência de atividade proteolítica generalizada em resposta às drogas. Entretanto, uma relação causal direta da ativação da necrose com o aumento da associação de mRNAs de calpaínas aos polissomos não pode ser estabelecida a partir dos dados aqui apresentados.

Em condições celulares normais, as calpaínas têm funções diversas, sendo que a clivagem precisa e limitada de proteínas estruturais e regulatórias é essencial para modular a função protéica. Entretanto, o influxo de cálcio ao citoplasma resulta numa ativação exacerbada e conseqüente degradação celular concomitante (ARTAL-SANZ & TAVERNARAKIS, 2005). Embora de atividade desconhecida, a expansão e conservação dos genes de calpaínas em *Trityps* chama atenção para um papel essencial destas proteínas nesses parasitas, sendo possível sua relação com a morte celular por necrose.

## 6.5 REGULOMA

Um passo inicial natural na análise de dados genômicos é examinar os extremos, ou seja, buscar os genes diferencialmente expressos na comparação de determinadas amostras ou em um experimento de série temporal. Como demonstrado nos tópicos anteriores dessa

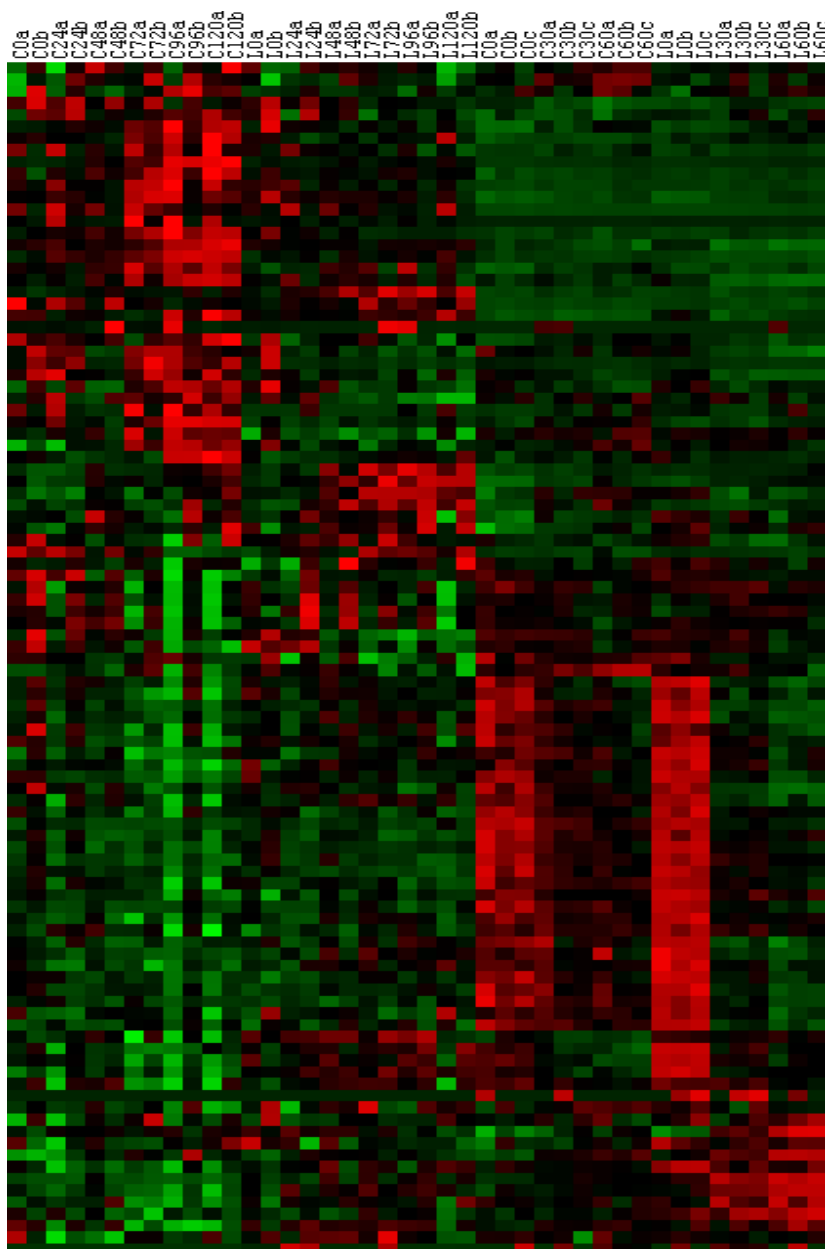
dissertação, essa análise inicial dos dados pode ser extremamente eficiente, por exemplo, na busca de alvos de drogas ou respostas globais a um certo tratamento. Entretanto, tal análise não aborda o imenso potencial de experimentos em escala. Para tal, faz-se necessário uma análise abrangente de todo o repertório de transcrições, numa análise do estado global de uma célula quando ela passa através de um processo biológico. Necessita-se uma abordagem holística para a análise de dados genômicos que incide em todo o conjunto de observações, permitindo biólogos desenvolverem uma compreensão integrada do processo que está sendo estudado (EISEN *et al.*, 1998).

Uma forma de organizar dados de expressão gênica é agrupar genes com padrões similares de expressão, o que, associado com representações gráficas dos dados, permite a identificação de grupos de genes co-expressos. Se tal análise for realizada em nível global, para todo o genoma de um organismo, incluindo um ampla gama de resposta biológicas, diversos grupos de genes co-regulados (*regulons*) podem ser extraídos dos dados, os quais naturalmente possuem genes funcionalmente relacionados (EISEN *et al.*, 1998; HU *et al.*, 2010).

A partir dos dados de RNA-Seq de todas as amostras analisadas, buscou-se grupos de genes co-regulados em diferentes situações (*regulons*). Para tal, os dados brutos dos seqüenciamentos foram normalizados em conjunto no programa edgeR e, em seguida, normalizados pelo fator Z. A tabela final, composta de 42 colunas de dados (cada uma das réplicas biológicas analisadas no presente trabalho) foi analisada no programa Cluster. Para agrupamento hierárquico dos supergenes, utilizamos apenas os genes com maior variação de expressão entre as amostras, ou seja, somente os genes com informações úteis para análises de co-regulação.

A Figura 6.5 mostra a clusterização dos 100 genes mais variantes dentre as 42 amostras sequenciadas no presente trabalho. A visualização da figura deixa claro que este tipo de análise possibilita a identificação de grupos de genes co-expressos





**FIGURA 6.5 - REGULOMA**

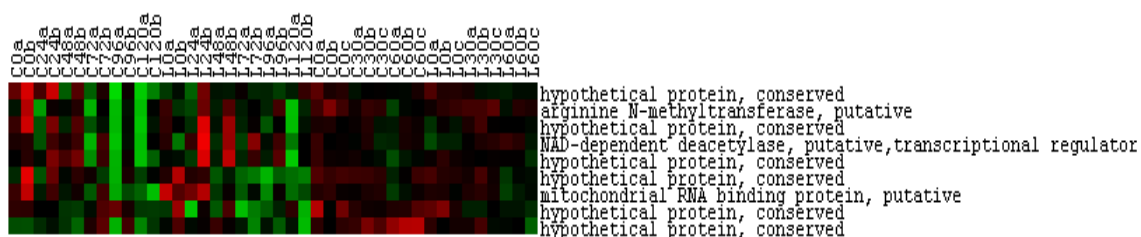
Clusterização hierárquica dos 100 genes mais variantes entre as 42 amostras analisadas no presente trabalho. Fica claro a presença de grupos de genes co-regulados em diferentes situações.

nos diferentes tratamentos. Entretanto, a pequena quantidade relativa de dados para análise de reguloma, juntamente com a utilização de drogas que atuam na mesma via, permite somente a identificação de grupos de genes de expressão diferencial em cada um dos tratamentos (observe Figura 6.5). Isso demonstra que análises robustas para identificação de *regulons* (genes co-regulados) necessitam dados biologicamente mais diversos que permeiem diferentes tipos de respostas celulares (ciclo de vida do parasita, tratamento com IBEs, diferentes tipos de stresse, infecção da célula hospedeira, decaimento de RNA, etc). Eisen e colaboradores (1998), num trabalho original de análises de reguloma em *S. cerevisiae*

utilizando a técnica de microarranjo, utilizaram 71 amostras do transcriptoma em uma diversidade de respostas biológicas (ciclo celular, esporulação, choque térmico em altas e baixas temperaturas, agentes redutores, dentre outras), o que permitiu a identificação de *regulons* corroborados em nível funcional. Em um trabalho recente com parasita *Plasmodium falciparum*, foram realizadas análises de microarranjo na resposta temporal a 20 compostos que inibem o crescimento deste parasita, totalizando 144 pontos experimentais (HU *et al.*, 2010). A predição de função de inúmeras proteínas desconhecidas foi realizada utilizando uma rede de interação construída a partir de co-expressão dos genes, homologia de sequências, e de dados de interação por experimentos de duplo-híbridos em levedura.

Embora os trabalhos citados acima usem dados muito amplos para identificação de *regulons*, mesmo com os dados das duas drogas utilizadas no presente trabalho foi possível observar grupos de genes co-regulados que representam possíveis *regulons*. Destaca-se a presença de diversas proteínas ligadoras de RNA nos dados de clusterização, o que indica que seus parceiros de co-expressão devam ser regulados por este tipo de proteína e chama atenção para os mecanismos pós-transcricionais de expressão gênica em *T. cruzi*. A presença desse tipo de proteína nos *regulons* é o caso ideal para estudo e identificação de motivos de ligação protéico nas regiões UTR dos mRNAs.

Dois exemplos interessantes de possíveis *regulons* são apresentados na Figura 6.6 e Figura 6.7. A presença de uma proteína ligadora de RNA mitocondrial no regulon da Figura 6.6 permite-nos inferir que esta proteína seja essencial na expressão de seus parceiros, destacando-se a presença de um regulador transcricional dependente de NAD.



**FIGURA 6.6 – POSSÍVEL REGULON CONTENDO PROTEÍNA LIGADORA DE RNA**

Outro grupo de genes co-expressos interessante está apresentado na Figura 6.7. A presença da proteína argonauta de *T. cruzi* nesse cluster chama atenção para o papel regulador dessa proteína em *T. cruzi*, destacando a possível funcionalidade desse grupo de proteínas, até então pouco conhecidas em tripanossomatídeos, principalmente em *T. cruzi* que não possuiu a via RNAi de degradação de RNAs.

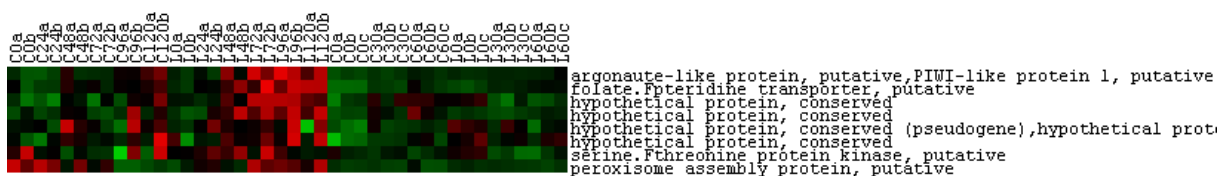


FIGURA 6.7 – POSSÍVEL *REGULON* CONTENDO PROTEÍNA ARGONAUTA

A confirmação de *regulons* pode ser principalmente demonstrada por análises de enriquecimento de termos de GO (ASHBURNER *et al.*, 2000). Além disso, a corroboração de *regulons* pode ser realizada cruzando dados de interação proteína-proteína (Interatoma ou dados de literatura) ou indícios de co-localização celular das proteínas. Ou seja, a confirmação de um *regulons* é uma análise multifatorial que engloba grande quantidade de dados, numa análise essencialmente em larga escala. Como os grupos de genes co-expressos nos dados acima são muito pequenos, e geralmente associados a expressão diferencial a um dos tratamentos, a busca por termos de GO ou motivos nas regiões UTR é desfavorecida.

Os dados apresentados acima demonstraram claramente que o projeto reguloma é uma forma muito útil de se avaliar a biologia básica de organismos com muitas proteínas de função desconhecida. A co-regulação de proteínas hipotéticas em *regulons* contendo proteínas já caracterizadas pode abrir novos caminhos na caracterização funcional das mesmas, além de gerar informações muito valiosas dos mecanismos de regulação pós-transcricional nesse parasita.

Como discutido anteriormente, a análise genômica correlacionada de uma ampla gama de experimentos de RNA-Seq será a base para as análises bem fundamentadas do projeto Reguloma de *T. cruzi*. Assim sendo, os dados acima demonstram o potencial dessa metodologia, visto que com um número relativamente pequeno de amostras já foi possível a identificação padrões interessantes. A correlação com os dados do Projeto Interatoma de *T. cruzi*, em desenvolvimento em nosso Instituto, deverá propiciar uma grande contribuição no entendimento da biologia básica deste parasita e facilitar a identificação de novas drogas.

## 7 CONCLUSÕES

A partir do exposto no presente trabalho concluímos que:

- RNA-Seq é um método muito rico para quantificação do transcriptômica de *T. cruzi*, o que, aliado a utilização de RNA polissomal, permitiu a identificação de um número muito grande de genes diferencialmente expressos em resposta aos inibidores da síntese de esteróis.
- As respostas transcricionais de *T. cruzi* frente a altas doses dos inibidores (DL100) desencadeiam respostas estereotipadas que envolvem uma sobrecarga mitocondrial, com aumento de expressão de genes da cadeia de transporte de elétrons e ATP sintase. Além disso, diversas vias catabólicas são ativadas, como degradação de aminoácidos e açúcar. Paralelamente, há forte inibição do metabolismo de RNA, com diminuição da expressão de subunidades das três RNA polimerases (Pol I, II e III) e diversas proteínas ligadoras de RNA.
- Considerando as alterações morfo-fisiológicas de *T. cruzi* em resposta a DL100, e aumento de expressão de inúmeras cisteína-peptidases do tipo calpaína e catepsina, acreditamos que a morte celular em resposta a altas doses de inibidores da síntese de esteróis ocorra por um processo similar à necrose, envolvendo a liberação de peptidases reservossomais para o citoplasma.
- Frente a inibição da síntese de esteróis com a dose  $IC_{50}$  das drogas, as principais organelas alteradas são mitocôndria e reservossomo. O aumento na expressão de subunidades do complexo citocromo c oxidase e de diversas proteínas reservossomais corroboram estes dados de microscopia eletrônica. Assim sendo, é possível que tais organelas tenham função essencial na síntese de esteróis nesse parasita.
- A enzima 24-C-metil-transferase é um possível ponto de controle da síntese de esteróis em *T. cruzi*. Devido a ausência desta enzima em mamíferos, estes dados reforçam a utilização desta enzima como alvo quimioterápico da doença de chagas;
- A metodologia de reguloma é uma forma extremamente informativa de se avaliar a biologia básica de *T. cruzi*, da qual podem ser extraídos dados esclarecedores acerca

da função de proteínas hipotéticas e de mecanismos de regulação pós-transcricional da expressão gênica.

## 8 PERSPECTIVAS

Este trabalho abre diversas perspectivas de atuação:

- Avaliar a redundância do RNA total nas modulações da expressão gênica em resposta aos inibidores da síntese de esteróis;
- Analisar a influência das drogas na metaciclogênese;
- Aprofundar a caracterização e entendimento funcional dos genes diferencialmente expressos em resposta as drogas;
- Caracterização do transcriptoma do parasita em resposta a outros inibidores da síntese de esteróis, que ajam em diferentes enzimas, visto que essa via demonstrou-se informativa do ponto de vista do projeto Reguloma;
- A análise de um grande conjunto de experimentos de RNA-Seq em *T. cruzi* possibilitará uma anotação criteriosa do genoma desse parasita, incluindo identificação de regiões não traduzidas nos mRNAs e sítios de *trans-splicing*.
- A demonstração da morte celular por necrose em resposta a DL100 das drogas é uma ótima oportunidade para estudos de morte celular programada em *T. cruzi*, envolvendo a caracterização funcional de calpaínas e catepsinas, enzimas aparentemente essenciais nesse processo.
- A demonstração da relação direta dos reservossomos na resposta a IC<sub>50</sub> das drogas encoraja a análise proteômica dessa organela frente a inibição da síntese de esteróis.
- Análises de regulação da expressão, localização celular e interação protéica da enzima 24-C-meti-transferase, um importante ponto de controle da via de esteróis em *T. cruzi*.
- Caracterização de transportadores do tipo ABC, aumentados em resposta as drogas e que, portanto, podem representar interessantes alvos de estudo para novos caminhos terapêuticos da Doença de Chagas.

## 9 REFERÊNCIAS

AMEISEN, J. C.; IZIDOREK, T.; BILLAUT-MULOT, O. *et al.* Apoptosis in a unicellular eukaryote (*Trypanosoma cruzi*): implications for the evolutionary origin and role of programmed cell death in the control of cell proliferation, differentiation and survival. **Cell Death Differ**, vol. 2, p. 285–300, 1995.

ANSORGE, W. J. Next-generation DNA sequencing techniques. **New Biotechnology**, vol. 25, n. 4, abril 2009.

ANDRADE, Z. A. Immunopathology of Chagas disease. **Mem Inst Oswaldo Cruz**, vol. 94 (Suppl I): p. 71-80, 1999.

APT, W.; AGUILERA, X.; ARRIBA, A. *et al.*. TREATMENT OF CHRONIC CHAGAS' DISEASE WITH ITRACONAZOLE AND ALLOPURINOL. **Am. J. Trop. Med. Hyg.**, vol. 59(1): 133–138, 1998.

APT, W.; ARRIBA, A.;ZULANTAY, I. *et al.*. Itraconazole or allopurinol in the treatment of chronic American trypanosomiasis: the results of clinical and parasitological examinations 11 years post-treatment. **Ann Trop Med Parasitol**, vol 99(8): 733-41, 2005.

ARTAL-SANZ, M.; TAVERNARAKIS, N. Proteolytic mechanisms in necrotic cell death and neurodegeneration. **FEBS Letters**, vol. 579, p. 3287–3296, 2005.

ASHBURNER, M.; BALL, C. A.; BLAKE, J. A.; BOTSTEIN, D.; BUTLER, H. *et al.* Gene Ontology: tool for the unification of biology. **Nature genetics**, vol. 25, 2000.

ÁVILA, A. R.; YAMADA-OGATTA, S. F.; MONTEIRO, V. S.; KRIEGER, M. A.; NAKAMURA, C. V.; DE SOUZA, W.; GOLDENBERG, S. Cloning and characterization of the metacyclogenin gene, which is specifically expressed during *Trypanosoma cruzi* metacyclogenesis. **Molecular and Biochemical Parasitology**, vol. 117, n. 2, p. 169-177, 2001.

BARBEE, K.D.; HUANG, X. Magnetic assembly of high-density DNA arrays for genomic analyses. **Anal. Chem.**, vol. 80, p. 2149–2154, 2008.

BEACH, D. H.; GOAD, L. J.; HOLZ, Jr. G. G. Effects of ketoconazole on sterol biosynthesis by *Trypanosoma cruzi* epimastigotes. **Biochem Biophys Res Commun**, vol 136, p. 851-856, 1986.

BENAIM, G.; SANDERS, J. M.;GARCIA-MARCHAM, Y. *et al.*. Amiodarone Has Intrinsic Anti-*Trypanosoma cruzi* Activity and Acts Synergistically with Posaconazole. **J Med Chem**, vol 49: 892-899, 2006.

BEN AMAR, M. F., A. PAYS, P. TEBABI, B. DERO, AND T. SEEBECK. Structure and transcription of the actin gene of *Trypanosoma brucei*. **Mol. Cell. Biol.** vol. 8, pg. 2166–2176, 1988.

BENTLEY, D. R. Whole-genome re-sequencing. **Current Opinion in Genetics & Development**, vol. 16, p. 545–552, 2006.

BENZ, C., NILSSON, D., ANDERSSON, B., CLAYTON, C., GUILBRIDE, D. L. Messenger RNA processing sites in *Trypanosoma brucei*. **Molecular and Biochemical Parasitology**, vol. 143, p. 125–134, 2005.

BERRIMAN, M.; GHEDIN, E.; HERTZ-FOWLER, C. *et al.* The genome of the African trypanosome *Trypanosoma brucei*. **Science**, v. 309, n. 5733, p. 416-422, 2005.

BLOCH, K. The biological synthesis of cholesterol. **Science**, vol 150: 19-28, 1965.

BRAGA, M. V.; URBINA, J. A.; DE SOUZA, W. Effects of squalene synthase inhibitors on the growth and ultrastructure of *Trypanosoma cruzi*. **International Journal of Antimicrobial Agents**, vol. 24, p. 72–78, 2004.

BRAGA, M. V.; MAGARACI, F.; LORENTE, S. O.; GILBERT, I.; DE SOUZA, W. Effects of inhibitors of Delta24(25)-sterol methyl transferase on the ultrastructure of epimastigotes of *Trypanosoma cruzi*. **Microscopy and Microanalysis**, vol. 11, p. 506–515, 2005.

BRENER, Z. Terapêutica experimental na doença de Chagas. In Brener, Z., Andrade, Z., Barral-Netto, M. **Trypanosoma cruzi e Doença de Chagas**, 2a ed., Guanabara Koogan, Rio de Janeiro, p. 379-388, 2000.

BRENER, Z.; CANÇADO, J. R.; GALVÃO, L. M.; DA LUZ, Z. M. P.; FILARDI, L. S.; PEREIRA, M. E. S.; SANTOS, L. M. T.; CANÇADO, C. B. An experimental and clinical assay with ketoconazole in the treatment of Chagas disease. **Mem Inst Oswaldo Cruz**, vol 88: 149–53, 1993.

BRENNER, S.; JOHNSON, M.; BRIDGHAM, J.; GOLDA, G.; LLOYD, D. H. *et al.* Gene expression analysis by massively parallel signature sequencing (MPSS) on microbead arrays. **Nature Biotechnology**, vol. 18, p. 630-634, 2000.

BRISSE, S. *et al.* A phylogenetic analysis of the *Trypanosoma cruzi* genome project CL Brener reference strain by multilocus enzyme electrophoresis and multiprimer random amplified polymorphic DNA fingerprinting. **Molecular and Biochemical Parasitology**, vol. 92, 1998.

BRISSE, S., BARNABE, C., TIBAYRENC, M. Identification of six *Trypanosoma cruzi* phylogenetic lineages by random amplified polymorphic DNA and multilocus enzyme electrophoresis. **Int. J. Parasitol.**, vol. 30, 2000.

BUCKNER, F. S.; VERLINDE, C. L. M. J.; LA FLAMME, A. C. *et al.* Efficient technique for screening drugs for activity against *Trypanosoma cruzi* using parasites expressing galactosidase. **Antimicrob Agents Chemother**, vol 40: 2592-2597, 1996.



BUCKNER,, F. S.; GRIFFIN, J.; WILSON, A. J.; VAN VOORHIS, W. C. Potent Anti-*Trypanosoma cruzi* Activities of Oxidosqualene Cyclase Inhibitors. **ANTIMICROBIAL AGENTS AND CHEMOTHERAPY**, vol. 45, n. 4, p. 1210-1215, 2001.

BUCKNER, F.; YOKOYAMA, K.; LOCKMAN, J. *et al.*. A class of sterol 14-demethylase inhibitors as anti-*Trypanosoma cruzi* agents. **PNAS**, Vol 100, n 25, dezembro/9, 2003.

BUCKNER, F. S. Sterol 14-Demethylase Inhibitors for *Trypanosoma cruzi* Infections. **Adv Exp Med Biol**, vol 625: 61-80, 2008.

BURLEIGH, B. A.; WOOLSEY, A. M. Cell signalling and *Trypanosoma cruzi* invasion. **Cell Microbiology**. Vol 4,n 11: 701-711, 2002.

BUTLER, J. S. The yin yang of the exosome. **Trends Cell Biol.**, vol 12, p. 90-96, 2002.

CAMARGO, E. P. GROWTH AND DIFFERENTIATION IN TRYPANOSOMA CRUZI. I. ORIGIN OF METACYCLIC TRYPANOSOMES IN LIQUID MEDIA. **Rev Inst Med Trop Sao Paulo**. Vol. 6, p. 93-100, 1964.

CAMPBELL, D. A.; THOMAS, S.; STURM, N. R. Transcription in kinetoplastid protozoa: why be normal? **Microbes and Infection**, vol. 5, pg. 1231–1240, 2003.

CANNON, R. D.; LAMPING, E.; HOLMES, A. R.; NIIMI, K.; BARET, P. V. *et al.* Efflux-Mediated Antifungal Drug Resistance. **CLINICAL MICROBIOLOGY REVIEWS**, vol. 22, n. 2, p. 291–321, 2009.

CANO, M.I.; GRUBER, A.; VAZQUEZ, M. *et al.* Molecular karyotype of clone CL Brener chosen for the *Trypanosoma cruzi* genome project. **Molecular and Biochemical Parasitology**, v. 71, n. 2, p. 273-278, 1995.

CAZZULO, J. J.; STOKA, V.; TURK, V. Cruzipain, the major cysteine proteinase from the protozoan parasite *Trypanosoma cruzi*. **Biol. Chem.**, vol. 378, n. 1, p. 1-10, 1997.

CHAGAS, C. Nova tripanozomíaze humana. Estudos sobre a morfologia e o ciclo evolutivo do *Schizotrypanum cruzi* n. gen., n. sp., agente etiolojico de nova entidade mórbida do homem. **Memórias do Instituto Oswaldo Cruz**, vol 1, n. 2, p. 159-218, 1909.

CHEN, C. K.; DOYLE, P. S.; YERMALITSKAYA, L.; MACKEY, Z. B.; ANG, K. K.; MCKERROW, J. H.; PODUST, L. M. *Trypanosoma cruzi* CYP51 inhibitor derived from a Mycobacterium tuberculosis screen hit. **PLoS Negl Trop Dis**, vol. 3, n. 2, e372, 2009.

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

CLAYTON, C.; SHAPIRA, M. Post-transcriptional regulation of gene expression in trypanosomes and leishmanias. **Molecular & Biochemical Parasitology**, vol. 156, p. 93–101, 2007.

CLOONAN, N.; FORREST, A. R. R.; KOLLE, G.; GARDINER, B. B. A.; FAULKNER, G. J. *et al.* Stem cell transcriptome profiling via massive-scale mRNA sequencing. **Nature Methods**, vol. 5, n.7, p. 613-619, 2008.

CONCEPCION, J. L.; GONZALEZ-PACANOWSKA, D.; URBINA, J. A. 3-hydroxy-3-methyl-glutaryl-CoA Reductase in *Trypanosoma* (*Schizotrypanum*) *cruzi*: subcellular localization and kinetic properties. **Arch Biochem Biophys**, vol 352:114-20, 1998.

CONESA, A.; GÖTZ, S.; GARCÍA-GÓMEZ, J. M.; TEROL, J.; TALÓN, M.; ROBLES, M. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. **Bioinformatics**, vol. 21, n. 18, p. 3674-3676, 2005.

CONTRERAS, L. M.; VIVAS, J.; URBINA, J. A. Altered Lipid Composition and Enzyme Activities of Plasma Membranes from *Trypanosoma* (*Schizotrypanum*) *Cruzi* Epimastigotes Grown in the Presence of Sterol Biosynthesis Inhibitors. **Biochemical Pharmacology**, Vol 53: 697-704, 1997.

COPPEL, R. L.; BLACK, C. G. Parasite genomes. **International Journal for Parasitology**, vol. 35, p 465–479, 2005.

COURA, J. R. Transmission of chagasic infection by oral route in the natural history of Chagas' disease. **Rev Soc Bras Med Trop**, vol. 39 (Suppl. 3): 113-117, 2006.

COURA, J. R. Chagas disease: what is known and what is needed – A background article. **Mem Inst Oswaldo Cruz**, Rio de Janeiro, Vol. 102(Suppl. I): 113-122, 2007.

COURA, J. R.; DE CASTRO, S. L. A Critical Review on Chagas Disease Chemoterapy. **Mem Inst Oswaldo Cruz**, Rio de Janeiro, Vol. 97(1): 3-24, janeiro, 2002.

COURA, J. R.; JUNQUEIRA, A. C. V.; FERNANDES, O.; VALENTE, S. A. S.; MILES, M. A. Emerging Chagas' disease in Amazonian Brazil. **Trends Parasitol**, vol. 18: 171-176, 2002.

CUEVAS, I. C.; CAZZULO, J. J.; SÁNCHEZ, D. O. gp63 Homologues in *Trypanosoma cruzi*: Surface Antigens with Metalloprotease Activity and a Possible Role in Host Cell Infection. **INFECTION AND IMMUNITY**, vol. 71, n. 10, p. 5739–5749, 2003.

DALLAGIOVANNA, B.; PLAZANET-MENUT, C.; OGATTA, S. F.; AVILA, A. R.; KRIEGER, M. A.; GOLDENBERG, S. *Trypanosoma cruzi*: a gene family encoding chitin. **Experimental Parasitology**, vol. 99, n. 1, p. 7-16, 2001.

DAHLQUIST, K. D.; SALOMONIS, N.; VRANIZAN, K.; LAWLOR, S. C.; CONKLIN, B. R. GenMAPP, a new tool for viewing and analysing microarray data on biological pathways. **Nature Genetics**. Vol. 31, n.11, p. 19-20, 2002.

DANIELSON, P. B. The Cytochrome P450 Superfamily: Biochemistry, Evolution and Drug Metabolism in Humans. **Current Drug Metabolism**, Vol 3, n 6: 561-597, 2002.

DE GAUDENZI, J.; FRASCH, A. C.; CLAYTON, C. RNA-binding domain proteins in Kinetoplastids: a comparative analysis. **Eukaryot Cell**, vol. 4, n. 12, p. 2106-2114, 2005.

DE LANGE, T.; MICHELS, P. A. M.; VEERMAN, H. G. J.; CORNELISSEN, A. W. C. A.; BORST, P. Many trypanosome messenger RNAs share a common 5' terminal sequence. **Nucleic Acids Research**, vol. 12, n 9, p 3777-3790, 1984.

DEMEL, R. A.; DE KRUYFF, B. The function of sterols in membranes. **Biochim. Biophys. Acta.**vol 457:109–32, 1976.

DE SOUZA, W. O Parasito e sua Interação com os Hospedeiros. In: Brenner, Z.; Andrade, A. Z.; Barral-Netto, M. **Trypanosoma cruzi e Doença de Chagas**. 2a Ed. Rio de Janeiro: Guanabara Koogan SA, 2000.

DE SOUZA, W. Macro, micro and nanodomains in the membrane of parasitic protozoa. **Parasitol Int**, vol. 56, p. 161-170, 2007.

DE SOUZA, W. Electron microscopy of trypanosomes - A historical view. **Mem Inst Oswaldo Cruz**, vol. 103, n. 4, p. 313-325, 2008.

DE SOUZA, W.; SANT'ANNA, C.; CUNHA-E-SILVA, N. L. Electron microscopy and cytochemistry analysis of the endocytic pathway of pathogenic protozoa. **Prog Histochem Cytochem.**, vol. 44, n. 2, p. 67-124, 2009.

DE SOUZA, W.; RODRIGUES, J. C. F. Sterol Biosynthesis Pathway as Target for Anti-trypanosomatid Drugs. **Interdisciplinary Perspectives on Infectious Diseases**, vol 2009, 2009.

DIAS, J. C. P. Epidemiologia. In: Brener Z, Andrade ZA, Barral-Netto M, editors. **Trypanosoma cruzi e a Doença de Chagas**. 2nd ed. Rio de Janeiro: Guanabara-Koogan; 2000. p.48-75.

DIAS, J. C. P.; PRATA, A.; CORREIA, D. Problems and perspectives for Chagas disease control: in search of a realistic analysis. **Revista da Sociedade Brasileira de Medicina Tropical**, Vol 41(2): 193-196, mar-abr, 2008.

DI NOIA, J. M.; D'ORSO, I.; SÁNCHEZ, D. O.; FRASCH, A. C. AU-rich elements in the 3'-untranslated region of a new mucin-type gene family of *Trypanosoma cruzi* confers mRNA instability and modulates translation efficiency. **J Biol Chem.**, vol. 275, p. 10218-10227, 2000.

DIVISÃO DE DOENÇA DE CHAGAS SUCAM/MS. Doença de Chagas e Transfusão de Sangue. In: **MINISTÉRIO DA SAÚDE. Superintendência de Campanhas de Saúde Pública. Doença de Chagas: Textos de apoio**. Brasília: Ministério da Saúde. Sucam. p. 52, 1989.

DOCAMPO, R.; MORENO, S. N.; TURRENS, J. F. Biochemical and ultrastructural alterations produced by miconazole and econazole in *Trypanosoma cruzi*. **Mol Biochem Parasitol**, vol 3: 169-180, 1981.

DOCAMPO, R.; DE SOUZA, W.; MIRANDA, K.; ROHLOFF, P.; MORENO, S. N. J. ACIDOCALCISOMES — CONSERVED FROM BACTERIA TO MAN. **Nat Rev Microbiol.**, vol. 3, n. 3, p. 251-261, 2005.

D'ORSO, I.; FRASCH, A. C. C. TcUBP-1, a developmentally regulated U-rich RNA-binding protein involved in selective mRNA destabilization in trypanosomes. **J. Biol. Chem.** vol. 276, p. 34801-34809, 2001.

DRESSMAN, D.; YAN, H.; TRAVERSO, G.; KINZLER, K. W.; VOGELSTEIN, B. Transforming single DNA molecules into fluorescent magnetic particles for detection and enumeration of genetic variations. **PNAS**, vol. 100, no. 15, p. 8817–8822, 2003.

DUMONTEIL, E. Vaccine development against *Trypanosoma cruzi* and Leishmania species in the post-genomic era. *Infection, Genetics and Evolution*, vol. 9, n. 6, p. 1075-82, 2009.

DUNBAR, D. A.; CHEN, A. A.; WORMSLEY, S.; BASERGA, S. J. The genes for small nucleolar RNAs in *Trypanosoma brucei* are organized in clusters and are transcribed as a polycistronic RNA. **Nucleic Acids Res**, vol 28, n 15, p. 2855-2861, 2000.

EICHER, T.; HAUPTMANN, S. **The Chemistry of Heterocycles: Structure, Reactions, Syntheses, and Applications**. Wiley-VCH, 2nd ed, 2003.

EISEN, M. B.; SPELLMAN, P. T.; BROWN, P. O.; BOTSTEIN, D. Cluster analysis and display of genome-wide expression patterns. **Proc. Natl. Acad. Sci. USA**, vol. 95, p. 14863–14868, 1998.

ELIAS, M. C.; MARQUES-PORTO, R.; FREYMULLER, E.; SCHENKMAN, S. Transcription rate modulation through the *Trypanosoma cruzi* life cycle occurs in parallel with changes in nuclear organisation. **Molecular and Biochemical Parasitology**, vol. 112, p. 79-90, 2001.

EL-SAYED, N. M. *et al.* The Genome Sequence of *Trypanosoma cruzi*, Etiologic Agent of Chagas Disease. **Science**. Vol 309, 15 julho 2005 (a).

EL-SAYED, N. M. *et al.* Comparative Genomics of Trypanosomatid Parasitic Protozoa. **Science**. Vol 309, 15 julho 2005 (b).

ENDO, A. The discovery and development of HMG-CoA reductase inhibitors. **Journal of Lipid Research**, vol 33, 1992

ERSFELD, K.; BARRACLOUGH, H.; GULL, K. Evolutionary Relationships and Protein Domain Architecture in an Expanded Calpain Superfamily in Kinetoplastid Parasites. **J. Mol. Evol.**, vol. 61, p. 742–757, 2005.

FILARDI, L. S.; BRENER, Z. Susceptibility and natural resistance of *Trypanosoma cruzi* strains to drugs used clinically in Chagas disease. **Trans. R. Soc. Trop. Med. Hyg.**, vol 81: 755–759, 1987.

FLEISCHMANN, R. D., ADAMS, M. D., WHITE, O., CLAYTON, R. A., KIRKNESS, E. F., KERLAVAGE, A. R., BULT, C. J., TOMB, J. F., DOUGHERTY, B. A., MERRICK, J. M. *et al.* Whole-genome random sequencing and assembly of *Haemophilus influenzae*. **Science**, vol 269, p 496-512, 2005.

FLORIN-CHRISTENSEN, M.; FLORIN-CHRISTENSEN, J.; GARIN, C.; ISOLA, E.; BRENNER, R. R.; RASMUSSEN, L. Inhibition of *Trypanosoma cruzi* growth and sterol biosynthesis by lovastatin. **Biochem Biophys Res Commun**, Vol 166, n. 3, p. 1441-1445, 1990.

FRAGOSO, S. P.; PLAZANET-MENUT, C.; CARREIRA, M. A.; MOTTA, M. C.; DALLAGIOVANA, B.; KRIEGER, M. A.; GOLDENBERG, S. Cloning and characterization of a gene encoding a putative protein associated with U3 small nucleolar ribonucleoprotein in *Trypanosoma cruzi*. **Molecular and Biochemical Parasitology**, v. 126, n. 1, p. 113-117, 2003.

FULLWOOD, M. J.; WEI, C-L.; LIU, E. T.; RUAN, Y. Next-generation DNA sequencing of paired-end tags (PET) for transcriptome and genome analyses. **Genome Research**, vol. 19, p. 521-532, 2009.

GHEDIN, E. *et al.* Gene synteny and evolution of genome architecture in trypanosomatids. **Mol Biochem Parasitol**, vol. 134, p. 183–91, 2004.

GIMPL, G.; BURGER, K.; FAHRENHOLZ, F. A closer look at the cholesterol sensor. **TRENDS in Biochemical Sciences**, vol. 27, n. 12, 2002.

GOAD, L. J.; BERENS, R. L.; MARR, J. J. The activity of ketoconazole and other azoles against *Trypanosoma cruzi*: Biochemistry and chemotherapeutic action in vitro. **Mol Biochem Parasitol**, vol 32: 179-190, 1989.

GOFFEAU, A., AERT, R., AGOSTINI-CARBONE, M. L., AHMED, A., AIGLE, M., ALBERGHINA, L., ALBERMANN, K., ALBERS, M., ALDEA, M., ALEXANDRAKI, D. *et al.* The yeast genome directory. **Nature**, vol. 387, Suppl., p 1-105, 1997.

GOLDENBERG, S.; SALLES, J.M.; CONTRERAS, V.T.; LIMA FRANCO, M.P.; KATZIN, A.M.; COLLI, W.; MOREL, C.M. Characterization of messenger RNA from epimastigotes and metacyclic trypomastigotes of *Trypanosoma cruzi*. **FEBS**, Vol. 180, p. 265-270, 1985.

GUNZL, A.; ULLU, E.; DORNER, M. *et al.* Transcription of the *Trypanosoma brucei* spliced leader RNA gene is dependent only on the presence of upstream regulatory elements. **Mol Biochem Parasitol**, vol. 85, pg. 67–76, 1997.

GUSTAFSON, E. A.; WESSEL, G. M. DEAD-box helicases: Posttranslational regulation and function. **Biochemical and Biophysical Research Communications**, vol. 395, p.1–6, 2010.

GUY, R. K. Inhibition of sonic hedgehog autoprocessing in cultured mammalian cells by sterol deprivation. **Proc. Nat. Acad. Sci USA**, vol 97: 7307–12, 2000.

- HAILE, S.; PAPADOPOULOU, B. Developmental regulation of gene expression in trypanosomatid parasitic protozoa. **Current Opinion in Microbiology**, vol. 10, p. 569–577, 2007.
- HALL, N. Advanced sequencing technologies and their wider impact in microbiology. **The Journal of Experimental Biology**, vol. 209, p. 1518-1525, 2007.
- HANKINS, E. G.; GILLESPIE, J. R.; AIKENHEAD, K.; BUCKNER, F. S. Upregulation of sterol C14-demethylase expression in *Trypanosoma cruzi* treated with sterol biosynthesis inhibitors. **Molecular & Biochemical Parasitology**, vol. 144, p. 68–75, 2005.
- HASHIMOTO, S. *et al.* High-Resolution Analysis of the 5'-End Transcriptome Using a Next Generation DNA Sequencer. **PLoS ONE**, vol. 4, 2009.
- HIDE, G. History of Sleeping Sickness in East Africa. **Clin. Microbiol. Rev.**, vol. 12, n. 1, p. 112-25, 1999.
- HEISE, N.; RAPER, J.; BUXBAUM, L. U.; PERANOVICH, T. M.; ALMEIDA, M. L. Identification of complete precursors for the glycosylphosphatidylinositol protein anchors of *Trypanosoma cruzi*. **J. Biol. Chem.**, vol. 271, n. 28, p. 16877-16887, 1996.
- HU, G.; CABRERA, A.; KONO, M.; MOK, S.; CHAAL, B. K.; HAASE, S.; ENGELBERG, K.; CHEEMADAN, S.; SPIELMANN, T.; PREISER, P. R.; GILBERGER, T. W.; BOZDECH, Z. Transcriptional profiling of growth perturbations of the human malaria parasite *Plasmodium falciparum*. **Nature Biotechnololy**, vol. 28, n. 1, p. 91-98, 2010.
- HUCKE, O.; GELB, M. H.; VERLINDE, C. L.; BUCKNER, F. S. The protein farnesyltransferase inhibitor Tipifarnib as a new lead for the development of drugs against Chagas disease. **J Med Chem**, vol. 48, p. 5415-5418, 2005.
- HUGHES, A. L.; TODD, B. L.; ESPENSHADE, P. J. SREBP Pathway Responds to Sterols and Functions as an Oxygen Sensor in Fission Yeast. **Cell**, vol. 120, p. 831–842, 2005.
- HUPPERT, J. L.; BALASUBRAMANIAN, S. G-quadruplexes in promoters throughout the human genome. **Nucleic Acids Res.**, vol. 35, p. 406–413, 2007.
- HURTADO-GUERRRERO, R.; PEÑA-DÍAZ, J.; MONTALVETTI, A. *et al.* Kinetic properties and inhibition of *Trypanosoma cruzi* 3-hydroxy-3-methylglutaryl CoA reductase. **FEBS Letters**, vol 510, p. 141-144, 2002.
- HUSE, S. M., HUBER, J. A.; MORRISON, H. G.; SOGIN, M. L.; WELCH, D .M. Accuracy and quality of massively parallel DNA pyrosequencing. **Genome Biol.**, vol. 8, 143, 2007.
- HYMAN, E.D. A new method of sequencing DNA. **Anal. Biochem.**, vol. 174, p. 423–436, 1988.
- ISTVAN, E. S.; DEISENHOFER, J. Structural mechanism for statin inhibition of HMG-CoA reductase. **Science**, Vol 292:1160-4, 2001.

ISTVAN, E. S. Structural mechanism for statin inhibition of 3-hydroxy-3-methylglutaryl coenzyme A reductase. **American Heart Journal**, December 2002.

IVENS, A.C.; PEACOCK, C.S.; WORTHEY, E.A. *et al.* The genome of the kinetoplastid parasite, *Leishmania major*. **Science**, v. 309, n. 5733, p. 436-442, 2005.

JENSEN, B. C.; SIVAM, D. S.; KIFER, C. T.; MYLER, P. J.; PARSONS, M. Widespread variation in transcript abundance within and across developmental stages of *Trypanosoma brucei*. **BMC Genomics**, vol. 10, 2009.

JIMÉNEZ-JIMÉNEZ, C.; CARRERO-LÉRIDA, J.; SEALEY-CARDONA, M.; PÉREZ, L. M. R.; URBINA, J. A.; PACANOWSKA, D. G.  $\Delta 24$  (25)-sterol methenyltransferase: Intracellular localization and azasterol sensitivity in *Leishmania major* promastigotes overexpressing the enzyme. **Molecular & Biochemical Parasitology**, vol. 160, p. 52–59, 2008.

KEENE, J. D. RNA regulons: coordination of post-transcriptional events. **Nat. Rev. Genet.**, vol 8, p. 533-543, 2007.

KEENE, J. D.; TENENBAUM, S. A. Eukaryotic mRNPs May Represent Posttranscriptional Operons. **Molecular Cell**, vol. 9, p. 1161–1167, 2002

KOHL, L.; GULL, K. Molecular architecture of the trypanosome cytoskeleton. **Mol. Biochem. Parasitol.**, vol. 93, n.1, p. 1-9, 1998.

KOLEV, N. G.; FRANKLIN, J. B.; CARMÍ, S.; SHI, H.; MICHAELI, S.; TSCHUDI, C. The Transcriptome of the Human Pathogen *Trypanosoma brucei* at Single-Nucleotide Resolution. **Plos Pathogens**, vol. 6, n. 9, 2010.

KOSEC, G.; ALVAREZ, V. E.; AGÜERO, F.; SÁNCHEZ, D.; DOLINAR, M.; TURK, B.; TURK, V.; CAZZULO, J. J. Metacaspases of *Trypanosoma cruzi*: possible candidates for programmed cell death mediators. **Mol. Biochem. Parasitol.**, vol. 145, n. 1, p. 18-28, 2006.

KOSEC, G.; ALVAREZ, V. E.; CAZZULO, J. J. Cysteine proteinases of *Trypanosoma cruzi*: from digestive enzymes to programmed cell death mediators. **BIOCELL**, vol. 30, n. 3, p. 479-490, 2006b.

KRAUS, J. M.; VERLINDE, C. L.; KARIMI, M.; LEPESHEVA, G. I.; GELB, M. H.; BUCKNER, F. S. Rational modification of a candidate cancer drug for use against Chagas disease. *J. Med. Chem.*, vol. 52, p. 1639-1647, 2009.

KRAUSE, M.; HIRSH, D. A trans-spliced leader sequence on actin mRNA in *C. elegans*. *Cell*, v. 49, n. 6, pg. 753-761, 1987.

KULKARNI, M. M.; OLSON, C. L.; ENGMAN, D. E.; MCGWIRE, B. S. *Trypanosoma cruzi* GP63 Proteins Undergo Stage-Specific Differential Posttranslational Modification and Are Important for Host Cell Infection. **Infection and Immunity**, vol. 77, n. 5, p. 2193-2200, 2009.

LAIRD, P. W. Trans splicing in trypanosomes--archaism or adaptation? **Trends in Genetics**, v. 5, n. 7, p. 204-208, 1989.

LAMB, D. C. *et al.*. Characteristics of the heterologously expressed human lanosterol 14 $\alpha$ -demethylase (other names: P45014DM, CYP51, P45051) and inhibition of the purified human and *Candida albicans* CYP51 with azole antifungal agents. **Yeast**, vol 15, n. 9, p. 755-63, 1999.

LANDER, E. S., LINTON, L. M., BIRREN, B., NUSBAUM, C., ZODY, M. C.; BALDWIN, J., DEVON, K., DEWAR, K., DOYLE, M., FITZHUGH, W. *et al.* Initial sequencing and analysis of the human genome. **Nature**, vol. 409, p. 860-921, 2001.

LAZARDI, K.; URBINA, J. A.; DE SOUZA, W. Ultrastructural Alterations Induced by Two Ergosterol Biosynthesis Inhibitors, Ketoconazole and Terbinafine, on Epimastigotes and Amastigotes of *Trypanosoma (Schizotrypanum) cruzi*. **ANTIMICROBIAL AGENTS AND CHEMOTHERAPY**, Vol 34 No. 11: 2097-2105, 1990.

LAZARDI, K.; URBINA, J. A.; DE SOUZA, W. Ultrastructural Alterations Induced by ICI 195,739, a Bis-Triazole Derivative with Strong Antiproliferative Action against *Trypanosoma (Schizotrypanum) cruzi*. **ANTIMICROBIAL AGENTS AND CHEMOTHERAPY**, vol. 35, n. 4, p. 736-740, 1991.

LEPESHEVA, G. I.; ZAITSEVA, N. G.; NES, W. D.; ZHOU, W.; ARASE, M.; LIU, J.; HILL, G. C.; WATERMAN, M. R. CYP51 from *Trypanosoma cruzi*: a phyla-specific residue in the B' helix defines substrate preferences of sterol 14 $\alpha$ -demethylase. **J. Biol. Chem.**, vol. 281, n. 6, p. 3577-3585, 2006.

LEPESHEVA, G. I.; OTT, R. D.; HARGROVE, T. Y.; KLESHCHENKO, Y. Y.; SCHUSTER, I. *et al.* Sterol 14 $\alpha$ -Demethylase as a Potential Target for Antitrypanosomal Therapy: Enzyme Inhibition and Parasite Cell Growth. **Chemistry & Biology**, vol 14, p. 283-1293, 2007.

LEPESHEVA, G. I.; HARGROVE, T. Y.; ANDERSON, S.; KLESHCHENKO, Y.; FURTAK, V.; WAWRZAK, Z.; VILLALTA, F.; WATERMAN, M. R. STRUCTURAL INSIGHTS INTO INHIBITION OF STEROL 14 $\alpha$ -DEMETHYLASE IN THE HUMAN PATHOGEN *TRYPANOSOMA CRUZI*. **J. Biol. Chem.**, vol. 285, n. 33, p. 25582-25590, 2010.

LEVINE, N. D.; CORLISS, J. O.; COX, F. E. G.; DEROUX, G.; BRAIN, J.; HONIGBERG, B. M.; *et al.* A newly revised classification of the Protozoa. **J. Protozool.**, vol. 27, n. 1, p. 37-58, 1980.

LI, J.; NAGPAL, P.; WITART, V.; MCMORRIS, T. C.; CHORY, J. A role for brassinosteroids in light-dependent development of *Arabidopsis*. **Science**, vol 272: 398-401, 1996.

LIANG, X.; HARITAN, A.; ULIEL, S.; MICHAELI, S. trans and cis Splicing in Trypanosomatids: Mechanism, Factors, and Regulation. **EUKARYOTIC CELL**, vol. 2, n. 5, p. 830-840, 2003.



LIENDO, A.; LAZARDI, K.; URBINA, J. A. In-vitro antiproliferative effects and mechanism of action of the bis-triazole D0870 and its S(-) enantiomer against *Trypanosoma cruzi*. **Journal of Antimicrobial Chemotherapy**, vol 41: 197–205, 1998.

LIENDO, A. *et al.*. Sterol composition and biosynthesis in *Trypanosoma cruzi* amastigotes. **Molecular and Biochemical Parasitology**, vol 104: 81-91, 1999.

LUDUNGA, L. Finding similar nucleotide sequences using network BLAST searches. **Curr Protoc Bioinformatics**. Chapter 3, Unit 3.3, 2009.

MACEDO, A. M.; MACHADO, C. R.; OLIVEIRA, R. P.; PENA, S. D. J. *Trypanosoma cruzi*: Genetic Structure of Populations and Relevance of Genetic Variability to the Pathogenesis of Chagas Disease. **Mem Inst Oswaldo Cruz**, vol. 99, n. 1, p. 1-12, 2004.

MADISON-ANTENUCCI, S.; GRAMS, J.; HAJDUK, S. L. Editing machines: the complexities of trypanosome RNA editing. **Cell.**, vol. 108, p. 435-438, 2002.

MAGARACI, F.; JIMENEZ, C. J.; RODRIGUES, C. *et al.* Azasterols as Inhibitors of Sterol 24-Methyltransferase in Leishmania Species and *Trypanosoma cruzi*. **J. Med. Chem.** Vol 46: 4714-4727, 2003.

MAIR, G.; SHI, H.; LI, H.; DJIKENG, A.; AVILES, H. O.; BISHOP, J. R.; FALCONE, F. H.; GAVRILESCU, C.; MONTGOMERY, J. L.; SANTORI, M. I.; STERN, L. S.; WANG, Z.; ULLU, E.; TSCHUDI, C. A new twist in trypanosome RNA metabolism: cis-splicing of premRNA. **RNA**, v. 6, n. 2, 163-169, 2000.

MALDONADO, R. A.; MOLINA, J.; PAYARES, G.; URBINA, J. A. Experimental Chemotherapy with Combinations of Ergosterol Biosynthesis Inhibitors in Murine Models of Chagas' Disease. **ANTIMICROBIAL AGENTS AND CHEMOTHERAPY**, Vol 37, no 6, p. 1353-1359, junho, 1993.

MANNING-CELA, R.; GONZÁLEZ, A.; SWINDLE, J. Alternative splicing of LYT1 transcripts in *Trypanosoma cruzi*. **Infect. Immun.**, vol. 70, p. 4726–4728, 2002.

MARDIS, E. R. Next-Generation DNA Sequencing Methods. **Annu. Rev. Genomics Hum. Genet.**, vol. 9, p. 387–402, 2008.

MARGULIES, M. *et al.*, Genome sequencing in microfabricated high-density picolitre reactors. **Nature**, vol. 437, p. 376–380., 2005.

MARIONI, J. C.; MASON, C. E.; MANE, S. M.; STEPHENS, M.; GILAD, Y. RNA-seq: An assessment of technical reproducibility and comparison with gene expression arrays. **Genome Research**, vol.18, p. 1509-1517, 2008.

MARTIN, M. B.; GRIMLEY, J. S.; LEWIS, J. C.; *et al.* Bisphosphonates inhibit the growth of *Trypanosoma brucei*, *Trypanosoma cruzi*, *Leishmania donovani*, *Toxoplasma gondii*, and *Plasmodium falciparum*: a potential route to chemotherapy. **Journal of Medicinal Chemistry**, vol. 44, n. 6, p. 909–916, 2001.

MATTHEWS, K.R.; TSCHUDI, C.; ULLU, E. A common pyrimidine-rich motif governs transsplicing and polyadenylation of tubulin polycistronic pre-mRNA in trypanosomes. **Genes and Development**, v. 8, n. 4, p. 491-501, 1994.

MAXAM, A.M.; GILBERT, W. A new method for sequencing DNA. **Proc. Natl. Acad. Sci. U. S. A.**, vol. 74, p 560–564, 1977.

MAYA, J. D.; CASSELS, B. K.; ITURRIAGA-VÁSQUEZ, P.; FERREIRA, J.; FAÚNDEZ, M. *et al.* Mode of action of natural and synthetic drugs against *Trypanosoma cruzi* and their interaction with the mammalian host. **Comparative Biochemistry and Physiology**, vol 146: 601-620, 2007.

MCCABE, R. E.; REMMINGTON, J. S.; ARAUJO, F. J. Ketoconazole inhibition of intracellular multiplication of trypanosome cruzi and protection of mice against lethal infection with the organism. **J Infect Dis**, vol 150:594-601, 1984.

MCCABE, R. E.; REMMINGTON, J. S.; ARAUJO, F. J. In vitro and in vivo effects of itraconazole against *Trypanosoma cruzi*. **Am J Trop Med Hyg**, vol 35:280-284, 1986.

MCCABE, R. E. Failure of Ketoconazole to cure chronic murine Chagas disease. **J Infect Dis**, 158: 1408-1409, 1988.

MCNICOLL, F.; DRUMMELSMITH, J.; MULLER, M.; MADORE, E.; BOILARD, N.; OUELLETTE, M.; PAPADOPOULOU, B. A combined proteomic and transcriptomic approach to the study of stage differentiation in *Leishmania infantum*. **Proteomics**, vol. 6, p. 3567-3581, 2006.

METZKER, M. L. Emerging technologies in DNA sequencing. **Genome Res.**, vol. 15, p. 1767-1776, 2005.

METZKER, M. L. Sequencing technologies - the next generation. **Nature Reviews in Genetics**. Vol. 11, p. 31-46, 2010.

MIKKELSEN, T., HILLER, L. W., EICHLER, E. E., ZODY, M. C., JAFFE, D. B., YANG, S. P., ENARD, W., HELLMANN, I., LINBALD-TOH, K. AND ALTHEIDE, T. K. Initial sequence of the chimpanzee genome and comparison with the human genome. **Nature**, vol 437, p. 69-87, 2005.

MINNING, T. A.; WEATHERLY, D. B.; ATWOOD, J.; ORLANDO, R.; TARLETON, R. L. The steady-state transcriptome of the four major life-cycle stages of *Trypanosoma cruzi*. **BMC Genomics**, vol. 10, 2009.

MOLINA, J.; MARTINS-FILHO, O.; BRENER, Z.; ROMANHA, A. J.; LOEBENBERG, D.; URBINA, J. A. Activities of the Triazole Derivative SCH 56592 (Posaconazole) against Drug-Resistant Strains of the Protozoan Parasite *Trypanosoma* (*Schizotrypanum*) *cruzi* in Immunocompetent and Immunosuppressed Murine Hosts. **ANTIMICROBIAL AGENTS AND CHEMOTHERAPY**, vol. 44, n. 1, p. 150-155, 2000.

MONTALVETTI, A.; ROHLOFF, P.; DOCAMPO, R. A functional aquaporin co-localizes with the vacuolar proton pyrophosphatase to acidocalcisomes and the contractile vacuole complex of *Trypanosoma cruzi*. **J Biol Chem.**, vol. 279, p. 3867-3882, 2004.

MOREIRA, A. A. *et al.*. Evaluation of the therapeutic activity of itraconazole in chronic infections, experimental and human, by *Trypanosoma cruzi*. **Rev Inst Med Trop Sao Paulo**, vol 34, n. 2, p. 177-80, 1992.

MOROZOVA, O.; MARRA, M. A. Applications of next-generation sequencing technologies in functional genomics. **Genomics**, vol. 92, p. 255-264, 2008.

MOROZOVA, O.; HIRST, M.; MARRA, M. A. Applications of New Sequencing Technologies for Transcriptome Analysis. *Annu. Rev. Genomics Hum. Genet.*, vol 10, p. 135–151, 2009.

MORTAZAVI, A., WILLIAMS, B. A., MCCUE, K., SCHAEFFER, L. & WOLD, B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. **Nature Methods**, vol. 5, p. 621–628, 2008.

MUNROE, D. J.; HARRIS, T. J. R. Third Generation sequencing fireworks at Marco Island. **NATURE BIOTECHNOLOGY**, vol. 28, n. 5, p. 426-428, 2010.

NAGALAKSHMI, U. *et al.* The transcriptional landscape of the yeast genome defined by RNA sequencing. **Science**, vol. 320, p. 1344–1349, 2008.

NELSON, D. L.; COX, M. M. **Lehninger Principles of Biochemistry**. Fourth Edition. New York: Freeman and Company, 2005.

NYREN, P.; LUNDIN, A. Enzymatic method for continuous monitoring of inorganic pyrophosphate synthesis. **Anal. Biochem.**, vol. 151, p. 504–509, 1985.

OGATA, H.; GOTO, S.; SATO, K.; FUJIBUCHI, W.; BONO, H.; KANEHISA, M. KEGG: Kyoto Encyclopedia of Genes and Genomes. **Nucleic Acids Res.**, vol. 27, n. 1, p. 29-34, 1999.

OGBADOIYI, E. O.; ROBINSON, D. R., GULL, K. A high-order trans-membrane structural linkage is responsible for mitochondrial genome positioning and segregation by flagellar basal bodies in trypanosomes. **Mol Biol Cell**, vol. 14, p. 1769-1779, 2003.

OPPERDOES, F. R.; BORST, P. Localization of nine glycolytic enzymes in a microbody-like organelle in *Trypanosoma brucei*. **FEBS Lett**, vol. 80, p. 360-364, 1977.

OVERATH, P.; HAAG, J.; LISCHKE, A.; O'HUIGIN, C. The surface structure of trypanosomes in relation to their molecular phylogeny. **Int J Parasitol.** May 1;31(5-6): 468-71, 2001.

PALENCHAR, J. B.; BELLOFATTO, V. Gene transcription in trypanosomes. **Molecular & Biochemical Parasitology**, vol. 146, pg. 135–141, 2006.

PALFI, Z.; SCHIMANSKI, B.; GUNZL, A.; LUCKE, S.; BINDEREIF, A. U1 small nuclear RNP from *Trypanosoma brucei*: a minimal U1 snRNA with unusual protein components. **Nucleic Acids Research**, v. 33, n. 8, p. 2493-2503, 2005.

PALMIERI, N.; SCHLOTTERER, C. Mapping Accuracy of Short Reads from Massively Parallel Sequencing and the Implications for Quantitative Expression Profiling. **PLOS ONE**, vol. 4, n. 7, 2009.

PARSONS, M.; WORTHEY, E. A.; WARD, P. N.; MOTTRAM, J. C. Comparative analysis of the kinomes of three pathogenic trypanosomatids: *Leishmania major*, *Trypanosoma brucei* and *Trypanosoma cruzi*. **BMC Genom.**, vol. 6, p. 127, 2005.

PAULIN, J. The chondriome of selected trypanosomatids. A threedimensional study based on serial thick sections and high voltage electron microscopy. **J Cell Biol.**, vol. 66, p. 404-413, 1975.

PAVONI, D. P. CARACTERIZAÇÃO GENÔMICA FUNCIONAL DAS FORMAS INFECTIVAS DO TRYPANOSOMA CRUZI. Curitiba, 2005. Tese (Doutorado em Biologia Celular e Molecular) - Instituto Oswaldo Cruz.

PEÑA-DÍAZ, J.; MONTALVETTI, A.; CAMACHO, A.; GALLEGO, C.; RUIZ-PEREZ, L. M.; GONZALEZ-PACANOWSKA, D. A soluble 3-hydroxy-3-methylglutaryl-CoA reductase in the protozoan *Trypanosoma cruzi*. **Biochem. J.**, Vol 324: 619-626, 1997.

PEÑA-DÍAZ, J.; MONTALVETTI, A.; FLORES, C. *et al.*. Mitochondrial Localization of the Mevalonate Pathway Enzyme 3-Hydroxy-3-methyl-glutaryl-CoA Reductase in the Trypanosomatidae. **Molecular Biology of the Cell**, Vol. 15: 1356–1363, 2004.

PERRY, K. L.; WATKINS, K. P.; AGABIAN, N. Trypanosome mRNAs have unusual "cap 4" structures acquired by addition of a spliced leader. **Proc Natl Acad Sci USA**. vol. 84, n. 23, p. 8190–8194, 1987.

PORTO-CARREIRO, I.; ATTÍAS, M.; MIRANDA, K.; DE SOUZA, W.; CUNHA-E-SILVA, N. *Trypanosoma cruzi* epimastigote endocytic pathway: cargo enters the cytostome and passes through an early endosomal network before storage in reservosomes. **Eur J Cell Biol**, vol. 79, p. 858–869, 2000.

PRETI, H. CARACTERIZAÇÃO FUNCIONAL DE GENES CODIFICADORES DE PROTEÍNAS HIPOTÉTICAS DIFERENCIALMENTE EXPRESSOS DURANTE A METACICLOGÊNESE DE *TRYPANOSOMA CRUZI*. Curitiba, 2007. Dissertação (Mestrado em Biologia Celular e Molecular) – Pós Graduação em Biologia Celular e Molecular/UFPR.

PRISCILA C. CAMPOS, P. C.; BARTHOLOMEU, D. C.; DAROCHA, W. D.; CERQUEIRA, G. C.; TEIXEIRA, S. M. R. Sequences involved in mRNA processing in *Trypanosoma cruzi*. **International Journal for Parasitology**, vol. 38, p. 1383–1389, 2008.

PROBST, C. M. Descrição da Metaciclogênese de *Trypanosoma cruzi* pelo uso de microarranjos de DNA. Curitiba, 2005. Tese (Doutorado em Biologia Celular e Molecular) - Instituto Oswaldo Cruz.

PROUDFOOT, N.; FURGER, A.; DYE, M. Integrating mRNA processing with transcription. **Cell**, v. 108, n. 4, p. 501-512, 2002.

QUINONES, W., URBINA, J. A., DUBOURDIEU, M., CONCEPCION, L. The glycosome membrane of *Trypanosoma cruzi* epimastigotes: protein and lipid composition. **J. Exp. Parasitol.**, vol. 106, p. 135-149, 2004.

RAJKOVIC, A.; DAVIS, R. E.; SIMONSEN, J. N.; ROTTMAN, F. M. A spliced leader is present on a subset of mRNAs from the human parasite *Schistosoma mansoni*. **Proceedings National Academy Sciences USA**, v. 87, n. 22, p. 8879-8883, 1990.

RALSTON, K. S.; HILL, K. L. The flagellum of *Trypanosoma brucei*: new tricks from an old dog. **Int J Parasitol.**, vol. 38, p. 869-884, 2008.

Recommendations from a Satellite Meeting. **Mem Inst Oswaldo Cruz**, Rio de Janeiro, Vol. 94, Suppl. I: 429-432, 1999.

RESPUELA, P.; FERELLA, M.; RADA-IGLESIAS, A.; ASLUND, L. Histone acetylation and methylation at sites initiating divergent polycistronic transcription in *Trypanosoma cruzi*. **J. Biol. Chem.**, vol. 283, p. 15884-15892, 2008.

RICHTER, J. D.; SONENBERG, N. Regulation of cap-dependent translation by eIF4E inhibitory proteins. **Nature**, vol. 433, p.477-480, 2005.

ROBERTS, C. W. Fatty acid and sterol metabolism: potential antimicrobial targets in apicomplexan and trypanosomatid parasitic protozoa. **Molecular & Biochemical Parasitology**, vol 126: 129-142, 2003.

ROBINSON, M. D.; MCCARTHY, D. J.; SMYTH, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. **Bioinformatics**, vol. 26, n. 1, p. 139-140, 2010.

RODRIGUES, C. O.; CATISTI, R.; UYEMURA, S. A.; VERCESI, A. E.; LIRA, R.; RODRIGUEZ, C.; URBINA, J. A.; DOCAMPO, R. The sterol composition of *Trypanosoma cruzi* changes after growth in different culture media and results in different sensitivity to digitonin-permeabilization. **J Eukaryot Microbiol.**, vol. 48, n. 5, p. 588-594, 2001.

ROHLOFF, P.; DOCAMPO, R. A contractile vacuole complex is involved in osmoregulation in *Trypanosoma cruzi*. **Exp Parasitol**, vol. 118, p. 17-24, 2008.

RONAGHI, M. *et al.* Real-time DNA sequencing using detection of pyrophosphate release. **Anal. Biochem.**, vol. 242, p. 84-89, 1996.

SANGER, F. *et al.* DNA sequencing with chain-terminating inhibitors. **Proc. Natl. Acad. Sci. U. S. A.**, vol. 74, p 5463-5467, 1977.

SANT'ANNA, C.; NAKAYASU, E. S.; PEREIRA, M. G.; LOURENÇO, D.; DE SOUZA, W.; ALMEIDA, I. C.; CUNHA-E-SILVA, N. L. Subcellular proteomics of *Trypanosoma cruzi* reservosomes. **Proteomics**, vol. 9, n. 7, p. 1782-94, 2009.

SANTA-RITA, R. M.; LIRA, R.; BARBOSA, H. S.; URBINA, J. A.; DE CASTRO, S. L. Anti-proliferative synergy of lysophospholipid analogues and ketoconazole against *Trypanosoma cruzi* (Kinetoplastida: Trypanosomatidae): cellular and ultrastructural analysis. **Journal of Antimicrobial Chemotherapy**, vol. 55, p. 780–784, 2005.

SCHUSTER, S. C. *et al.* Method of the year, next-generation DNA sequencing. Functional genomics and medical applications. **Nat. Methods**, vol. 5, p. 11–21, 2008.

SERVICE, R. F. The Race for the \$1000 Genome. **Science**, vol. 311, 2006.

SHAPIRO, T. A.; ENGLUND, P. T. The Structure and Replication of Kinetoplast DNA. **Annu. Rev. Microbiol.**, vol. 49, p. 117-43, 1995.

SHENDURE, J. Accurate Multiplex Polony Sequencing of an Evolved Bacterial Genome. **Science**, vol. 309, p.1728-1732, 2005.

SHENDURE, J.; JI, H. Next-generation DNA sequencing. **Nature Biotechnology**, vol. 26, n. 10, p. 1135-1145, 2008.

SHIH S.; HWANG, H-Y.; CARTER, D.; STENBERG, P.; ULLMAN, B. Localization and targeting of the *Leishmania donovani* hypoxanthineguanine phosphoribosyltransferase to the glycosome. **J Biol Chem**, vol. 273, p. 1534-1541, 1998.

SHLOMAI, J. The structure and replication of kinetoplast DNA. **Curr. Mol. Med.**, vol. 4, p. 623-647, 2004.

SIEGEL, T. N.; HEKSTRA, D. R.; KEMP, L. E.; FIGUEIREDO, L. M.; LOWELL, J. E.; FENYO, D.; WANG, X.; DEWELL, S.; CROSS, G. A. Four histone variants mark the boundaries of polycistronic transcription units in *Trypanosoma brucei*. **Genes Dev.**, vol. 23, n. 9, p. 1063-1076, 2009.

SIEGEL, T. M.; HEKSTRA, D. R.; WANG, X.; DEWELL, S.; CROSS, G. A. M. Genome-wide analysis of mRNA abundance in two life-cycle stages of *Trypanosoma brucei* and identification of splicing and polyadenylation sites. **Nucleic Acids Research**, vol. 38, n. 15, p.:4946-4957, 2010.

SILVA, D. T.; DE MEIRELLES, M. N.; ALMEIDA, D.; URBINA, J. A.; PEREIRA, M. C. Cytoskeleton reassembly in cardiomyocytes infected by *Trypanosoma cruzi* is triggered by treatment with ergosterol biosynthesis inhibitors. **Int J Antimicrob Agents**, vol. 27, p. 530-537, 2006.

SIMPSON, A. G. B.; LUKES, J.; ROGER, A. J. The Evolutionary History of Kinetoplastids and Their Kinetoplasts. **Mol. Biol. Evol.**, vol. 19, n. 12, p. 2071–83, 2002.

SIMPSON, L.; NECKELMANN, N.; DE LA CRUZ, V. F.; SIMPSON, A. M.; FEAGIN, J. E.; JASMER, D. P.; STUART, J. E. Comparison of the maxicircle (mitochondrial) genomes of *Leishmania tarentolae* and *Trypanosoma brucei* at the level of nucleotide sequence. **J Biol Chem.**, vol. 262, n.13, p. 6182-6196, 1987.

SOARES, M. J.; SOUTO-PADRON, T.; DE SOUZA, W. Identification of a large pre-lysosomal compartment in the pathogenic protozoon *Trypanosoma cruzi*. **J Cell Sci.**, vol. 102, p. 157–167, 1992.

SOLARI, A. J. Mitosis and genome partition in trypanosomes. **Biocell**, vol. 19, p. 65-84, 1995.

SONG, Z.; NES, W. D. Sterol Biosynthesis Inhibitors: Potential for Transition State Analogs and Mechanism-Based Inactivators Targeted at Sterol Methyltransferase. **Lipids**, vol 42: 15–33, 2007.

SOUTO-PADRON, T.; DE SOUZA, W; HEUSER, J. E. Quick-freeze, deepetch rotary replication of *Trypanosoma cruzi* and *Herpetomonas megaseliae*. **J Cell Sci**, vol. 69, p. 167-168, 1984.

STRAUSBERG, R. L.; LEVY, S.; ROGERS; Y. Emerging DNA sequencing technologies for human genomic medicine. **Drug Discovery Today**, vol. 13, n 13/14, julho 2008.

STURM, N. R.; CAMPBELL, D. A. Alternative lifestyles: The population structure of *Trypanosoma cruzi*. **Acta Tropica**, 2009.

STURM, N. R.; VARGAS, N. S.; WESTENBERGER, S. J.; ZINGALES, B.; CAMPBELL, D. A. Evidence for multiple hybrid groups in *Trypanosoma cruzi*. **International Journal for Parasitology**, vol. 33, p. 269-279, 2003.

SULTAN, M. *et al.* A global view of gene activity and alternative splicing by deep sequencing of the human transcriptome. **Science**, vol. 321, 956-960, 2008.

SUTTON, R. E.; BOOTHROYD, J. C. Evidence for trans splicing in trypanosomes. **Cell**, v. 47, n. 4, pg. 527-535, 1986.

TANG, F.; BARBACIORU, C.; WANG, Y.; NORDMAN, E.; LEE, C. *et al.* mRNA-Seq whole-transcriptome analysis of a single cell. **Nature Methods**, vol. 6, n. 5, p. 377-382, 2009.

TARLETON, R. L. Parasite persistence in the aetiology of Chagas disease. **Int J Parasitol**, vol 31: 550-554, 2001.

TELLERIA, J.; LAFAY, B.; VIRREIRA, M.; BARNABÉ, C.; TIBAYRENC, M.; SVOBODA, M. *Trypanosoma cruzi*: Sequence analysis of the variable region of kinetoplast minicircles. **Exp. Parasitol.**, vol. 114, p. 279–88, 2006.

TESSIER, L.H.; KELLER, M.; CHAN, R.L.; FOURNIER, R.; WEIL, J.H.; IMBAULT, P. Short leader sequences may be transferred from small RNAs to pre-mature mRNAs by transsplicing in *Euglena*. **Embo Journal**, v. 10, n. 9, p. 2621-2625, 1991.

THOMAS, S.; GREEN, A.; STURM, N. R.; CAMPBELL, D. A.; MYLER, P. J. Histone acetylations mark origins of polycistronic transcription in *Leishmania major*. **BMC Genomics**, vol. 10, 2009.

THOMPSON, J. F.; DANLEY, D. E.; MAZZALUPO, S.; MILOS, P. M.; LIRA, M. E.; HARWOOD, H. J. Jr. Truncation of human squalene synthase yields active, crystallizable protein. **Arch Biochem Biophys**, vol 350: 283-90, 1998.

TIBAYRENC, M. Genetic subdivisions within *Trypanosoma cruzi* (Discrete Typing Units) and their relevance for molecular epidemiology and experimental evaluation. **Kinetoplast. Biol. Dis.**, vol. 2: 12, 2003.

TURK, B.; TURK, V. Lysosomes as “Suicide Bags” in Cell Death: Myth or Reality? **THE JOURNAL OF BIOLOGICAL CHEMISTRY**, vol. 284, n. 33, p. 21783–21787, 2009.

TYLER, K. M.; ENGMAN, D. M. The life cycle of *Trypanosoma cruzi* revisited. **International Journal for Parasitology**, Vol 31: 472-481, 2001.

URBINA, J. A.; LAZARDI, K.; AGUIRRE, T.; PIRAS, M. M.; PIRAS, R. Antiproliferative Synergism of the Allylamine SF 86-327 and Ketoconazole on Epimastigotes and Amastigotes of *Trypanosoma* (*Schizotrypanum*) *cruzi*. **ANTIMICROBIAL AGENTS AND CHEMOTHERAPY**, Vol 32, no 8: 1237-1242, Agosto 1988.

URBINA, J. A.; LAZARDI, K.; AGUIRRE, T. *et al.* Antiproliferative Effects and Mechanism of Action of ICI 195,739, a Novel Bis-Triazole Derivative, on Epimastigotes and Amastigotes of *Trypanosoma* (*Schizotrypanum*) *cruzi*. **ANTIMICROBIAL AGENTS AND CHEMOTHERAPY**, Vol. 35, no. 4: 730-735, 1991.

URBINA, J. A.; LAZARDI, L.; MARCHAN, E.; VISBAL, G.; AGUIRRE, T. *et al.* Mevinolin (Lovastatin) Potentiates the Antiproliferative Effects of Ketoconazole and Terbinafine against *Trypanosoma* (*Schizotrypanum*) *cruzi*: In Vitro and In Vivo Studies. **ANTIMICROBIAL AGENTS AND CHEMOTHERAPY**, vol 37, no 3: 580-591, 1993.

URBINA, J. A.; VIVAS, J.; VISBAL, G.; CONTRERAS, L. M. Modification of the sterol composition of *Trypanosoma* (*Schizotrypanum*) *cruzi*, epimastigotes by 24(25)-sterol methyl transferase inhibitors and their combinations with ketoconazole. **Mol Biochem Parasitol**; vol 73: 199–210, 1995.

URBINA, J. A.; PAYARES, G.; MOLINA, J. *et al.* Cure of short- and long-term experimental Chagas disease using D0870. **Science**, vol 273: 969–71, 1996.

URBINA, J. A. Lipid biosynthesis pathways as chemotherapeutic targets in kinetoplastid parasites. **Parasitology**, vol. 114, suppl, p. S91-S99, 1997.

URBINA, J. A.; PAYARES, G.; CONTRERAS, L. M. *et al.* Antiproliferative effects and mechanism of action of SCH 56592 against *Trypanosoma* (*Schizotrypanum*) *cruzi*: in vitro and in vivo studies. **Antimicrob Agents Chemother**, vol 42: 1771–7, 1998.



URBINA, J. A. *et al.*. In Vitro Antiproliferative Effects and Mechanism of Action of the New Triazole Derivative UR-9825 against the Protozoan Parasite *Trypanosoma* (*Schizotrypanum*) *cruzi*. **ANTIMICROBIAL AGENTS AND CHEMOTHERAPY**, vol 44, n9: 2498–2502, 2000.

URBINA, J. A. Specific treatment of Chagas disease: current status and new developments. **Current Opinion in Infectious Diseases**, vol 14, p. 733-741, 2001.

URBINA, J. A.; CONCEPCION, J. L.; RANGEL, S.; VISBAL, B.; LIRA, R. Squalene synthase as a chemotherapeutic target in *Trypanosoma cruzi* and *Leishmania mexicana*. **Molecular & Biochemical Parasitology**, vol. 125, p. 35-45, 2002.

URBINA, J. A. *et al.*. In vitro and in vivo activities of ravuconazole on *Trypanosoma cruzi*, the causative agent of Chagas disease. **International Journal of Antimicrobial Agents**, vol 21, p. 27-38, 2003 (a).

URBINA, J. A. *et al.*. Parasitological cure of acute and chronic experimental Chagas disease using the long-acting experimental triazole TAK-187. Activity against drug-resistant *Trypanosoma cruzi* strains. **International Journal of Antimicrobial Agents**, vol 21: 39-48, 2003 (b).

URBINA, J. A. Ergosterol biosynthesis and drug development for Chagas disease. **Mem Inst Oswaldo Cruz**, vol. 104, suppl. I, p. 311-318, 2009 (a).

URBINA, J. A. Specific chemotherapy of Chagas disease: Relevance, current limitations and new approaches. **Acta Tropica**, 2009 (b).

URBINA, J. A.; DOCAMPO, R. Specific chemotherapy of Chagas disease: controversies and advances. **TRENDS in Parasitology**, vol. 19, n. 11, 2003.

VANHAMME, L.; PAYS, E. Control of gene expression in trypanosomes. **Microbiology Reviews**, vol. 59, p. 223-240, 1995.

VANNIER-SANTOS, M. A.; MARTINY, A.; LINS, U.; URBINA, J. A.; BORGES, V. M.; DE SOUZA, W. Impairment of sterol biosynthesis leads to phosphorus and calcium accumulation in *Leishmania acidocalcisomes*. **Microbiology**, vol. 145, p. 3213-3220, 1999.

VARGAS, N.; PEDROSO, A.; ZINGALES, B. Chromosomal polymorphism, gene synteny and genome size in *T. cruzi* I and *T. cruzi* II groups. **Molecular & Biochemical Parasitology**, vol. 138, p. 131–141, 2004.

VELCULESCU, V.E.; ZHANG, L. ; VOGELSTEIN, B.; KINZLER, K.W. Serial analysis of gene expression, **Science**, vol. 270, p. 484–487, 1995.

VIEIRA, M.; ROHLOFF, P.; LUO, S.; CUNHA-E-SILVA, N. L.; DE SOUZA, W.; DOCAMPO, R. Role for a P-type H<sup>+</sup>-ATPase in the acidification of the endocytic pathway of *Trypanosoma cruzi*. **Biochem J.**, vol 392, p. 467–474, 2005.

VITALE, D. Interpreting mRNA electropherograms. [www.agilent.com/chem/labonachip](http://www.agilent.com/chem/labonachip), número de publicação: 5988-3001EN, 2001.

VIVAS, J.; URBINA, J. A.; DE SOUZA, W. Ultrastructural alterations in *Trypanosoma* (*Schizotrypanum*) *cruzi* induced by A2 4(25)s terol methyl transferase inhibitors and their combinations with ketoconazole. **International Journal of Antimicrobial Agents**, vol. 8, p. 1-6, 1997.

WANG, Z.; GERSTEIN, M.; SNYDER, M. RNA-Seq: a revolutionary tool for Transcriptomics. **Nature Reviews**, vol. 10, p. 57-63, 2009.

WATERSTON, R. H., LINDBLAD-TOH, K., BIRNEY, E., ROGERS, J., ABRIL, J. F., AGARWAL, P., AGARWALA, R., AINSCOUGH, R., ALEXANDERSSON, M., AN, P. *et al.* Initial sequencing and comparative analysis of the mouse genome. **Nature**, vol. 420, p. 520-562, 2002.

WEATHERLY, D. B.; BOEHLKE, C.; TARLETON, R. L. Chromosome level assembly of the hybrid *Trypanosoma cruzi* Genome. **BMC Genomics**, vol. 10, 2009.

WESTENBERGER, S. J.; BARNABE, C.; CAMPBELL, D. A.; STURM, N. R. Two Hybridization Events Define the Population Structure of *Trypanosoma cruzi*. **Genetics**, vol. 171, p. 527-543, 2005.

WHO-World Health Organization. Chagas disease. **Thirteenth Programme Report UNDP/TDR**, Geneve, 1997.

WILHELM, B. T.; LANDRY, J-R. RNA-Seq: quantitative measurement of expression through massively parallel RNA-sequencing. **Methods**, vol. 48, p. 249-257, 2009.

WILLIS, I. M. RNA polymerase III. Genes, factors and transcriptional specificity, **Eur. J. Biochem.**, vol. 212, pg 1-11, 1993.

WOLD, B.; MYERS, R. M. Sequence census methods for functional genomics. **Nature Methods**, vol. 5, p.19-21, 2008.

WRAY, G. A.; HAHN, M. A.; ABOUHEIF, E.; BALHOFF, J. P.; PIZER, M.; ROCKMAN, M. V.; ROMANO, L. A. The Evolution of Transcriptional Regulation in Eukaryotes. **Mol. Biol. Evol.**, vol. 20 (9), pg. 1377-1419, 2003.

YEO, M.; ACOSTA, N.; LLEWELLYN, M.; SÁNCHEZ, ADAMSON, S. *et al.* Origins of Chagas disease: Didelphis species are natural hosts of *Trypanosoma cruzi* I and armadillos hosts of *Trypanosoma cruzi* II, including hybrids. **International Journal for Parasitology**, vol. 35, p. 225-233, 2005.

YOSHIDA, N. Molecular mechanism of *Trypanosoma cruzi* infection by oral route. **Mem Inst Oswaldo Cruz**, Rio de Janeiro, Vol. 104 (Suppl. I), p. 101-107, 2009.

ZELEDON, R.; ALVARENGA, N.J.; SCHOSINSKY, K. In "Chagas disease" pp.59. Pan American, 1997.

ZINGALES, B.; STOLF, B. S.; SOUTO, R. P.; FERNANDES, O.; BRIONES, M. R. S. Epidemiology, Biochemistry and Evolution of *Trypanosoma cruzi* Lineages Based on Ribosomal RNA Sequences. **Mem Inst Oswaldo Cruz**, Rio de Janeiro, Vol. 94, Suppl. I: 159-164, 1999.

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

ZWIERZYNSKI, T.A.; BUCK, G.A. RNA-protein complexes mediate in vitro capping of the spliced-leader primary transcript and U-RNAs in *Trypanosoma cruzi*. **Proceedings National Academy Sciences USA**, v. 88, n. 13, p. 5626-5630, 1991.