

PRISCILA GRYNBERG

ANÁLISE DA EXPRESSÃO GÊNICA EM *Trypanosoma cruzi*  
EM RESPOSTA À RADIAÇÃO IONIZANTE PELA TÉCNICA DE  
MICROARRANJO DE DNA

---

Tese de Doutorado apresentada ao  
Programa de Doutorado em Bioinformática,  
Instituto de Ciências Biológicas, Universidade  
Federal de Minas Gerais como requisito parcial  
à obtenção do grau de Doutor em  
Bioinformática.

Área de Concentração: Bioinformática  
Funcional

Orientadora: Dr<sup>a</sup> Glória Regina Franco

Co-Orientadora: Dr<sup>a</sup> Daniella Castanheira Bartholomeu

Colaboradores: Dr. Carlos Renato Machado (ICB/UFMG)

10 de Março de 2011



## AGRADECIMENTOS

---

O desenvolvimento e conclusão desta Tese de Doutorado só foram possíveis com ajuda essencial de muitas pessoas, além das agências de fomento CAPES pelo fornecimento da bolsa de estudo, CNPq pelo financiamento deste projeto, e FAPEMIG pelos auxílios fornecidos para participações em congressos. Deixo aqui os meus mais sinceros agradecimentos.

À minha querida orientadora e amiga Prof. Dra. Glória Regina Franco, pela incrível oportunidade de desenvolver a minha tese sob a sua orientação, que resultou em uma relação muito além de orientador-orientado. Agradeço pelas conversas, conselhos e discussões científicas, pelo conhecimento enciclopédico de bioquímica e biologia molecular e pelo imensurável apoio no andamento deste projeto.

À minha querida co-orientadora Prof. Dra. Daniella Castanheira Bartholomeu, pela intenção inicial em me orientar, e por depois aceitar a co-orientação desta tese. O seu conhecimento sobre a biologia e estrutura genômica de *Trypanosoma cruzi* foram muito importantes.

Ao Prof. Dr. Carlos Renato Machado, que deu a idéia central deste trabalho, e profundo conhecedor da biologia do reparo de DNA em *T. cruzi*. Os momentos de conversa sobre os resultados foram sempre iluminadores.

À Profa. Dra. Andrea Mara Macedo e ao Dr. Sérgio Danilo Pena, pelas sugestões dadas durante conversas e nas reuniões de laboratório, e pelas inúmeras companhias para os almoços.

Ao Centro de Desenvolvimento de Tecnologia Nuclear (CDTN), e em especial à equipe do Laboratório de Radiação Gama (LIG) pela prontidão em nos receber para realizar

as radiações, e ao grupo de Aplicações das Radiações Ionizantes na Saúde e Biologia, pelo empréstimo das instalações laboratoriais para manipulação das amostras.

À Neuza Antunes Rodrigues, pela enorme competência em gerenciar o LGB.

À equipe do Dr. Marco Aurélio Krieger (principalmente à Dra. Daniela P. Pavoni) por terem me recebido no ICC/Fiocruz (Curitiba/PR). O conhecimento que adquiri foi fundamental para o sucesso das hibridações das lâminas de microarranjo.

Ao Prof. Dr. Roberto Hirata Jr. por ter me dedicado o seu tempo em analisar a qualidade dos meus dados, e ter me dado dicas valiosas sobre o R.

Aos colegas atuais e do passado do LGB pelas conversas, festas, reuniões e pela convivência. Foram quatro anos de muito prazer em fazer parte dessa grande família. Agradeço em especial aos Dr. Michael Waisberg por ter me treinado em experimentos de microarranjos (bancada e análise) e à Dra. Danielle G. Passos Silva, pelo treinamento em cultivo e irradiação de células epimastigostas de *T. cruzi*, passos fundamentais para o desenvolvimento deste trabalho. Aos grandes amigos e queridas amigas que ganhei durante esse tempo, muito obrigada por tornarem o LGB a minha segunda (quando não primeira) casa.

À minha família e ao Felipe, alicerces da minha vida. Agradeço aos meus pais por me darem o suporte para eu pudesse dedicar todo o meu tempo à minha formação acadêmica, desde a graduação até o doutorado. Ao Felipe, companheiro de quase uma década e companheiro para a vida, agradeço por ter permanecido ao meu lado em todos os momentos.

# SUMÁRIO

---

Agradecimentos .....	iii
Lista de Figuras .....	ix
Lista de Tabelas .....	xi
Lista de abreviaturas .....	xii
Resumo .....	xiv
Abstract.....	xv
1 - Introdução.....	1
1.1. O <i>Trypanosoma cruzi</i> e a Doença de Chagas: aspectos históricos, epidemiológicos e biológicos .....	1
1.2 – O projeto genoma de <i>T. cruzi</i> : características da cepa CL Brener .....	3
1.3 – O efeito da radiação ionizante e o processo de reparo de quebra de fita dupla nas moléculas de DNA.....	4
1.4 – As vias de reparo de DNA .....	7
1.5 – O efeito da radiação ionizante em <i>T. cruzi</i> e em outros organismos.....	8
1.6 - A técnica de Microarranjos.....	12
1.7 - Aplicação da técnica de microarranjo de DNA em estudos com o <i>T. cruzi</i> .....	16
2 - Objetivos .....	19
2.1 – Objetivo Geral.....	19
2.2 – Objetivos específicos .....	19
3 – Material e métodos .....	20
3.1 – Cultura de <i>T. cruzi</i> .....	20
3.2 – Irradiação das células com radiação gama .....	20

3.3 – Extração, purificação ( <i>Clean-up</i> ) e amplificação de RNA .....	21
3.4 – Design experimental.....	22
3.5 – Marcação e hibridização das amostras .....	25
3.6 – Análises das imagens .....	27
3.7 – Correção do <i>background</i> , normalização dos dados e aplicação de modelo linear para geração da tabela de genes diferentemente expressos.....	29
3.8 – Avaliação dos perfis de expressão gênica ao longo do tempo .....	33
3.9 – Análises das sequências obsoletas e identificação de sequências mitocondriais diferentemente expressas .....	33
3.10 – Atribuições de categorias de ontologia gênica ( <i>Gene Ontology</i> – GO) e análise funcional dos dados.....	34
3.11 – PCR quantitativa em tempo real (qRT-PCR).....	35
3.12 – Análise dos genes RHS .....	37
4 - Resultados .....	38
4.1 – Curva de crescimento após a irradiação.....	38
4.2 – Avaliação da integridade do RNA de <i>T. cruzi</i> após a irradiação .....	40
4.3 – Amplificação, marcação e hibridização do RNA.....	43
4.4 – Análises de microarranjos: correção do <i>background</i> e normalização dos dados..	44
4.4.1 – Correção do <i>background</i> : teste de dois diferentes algoritmos.....	44
4.4.2 – Normalização dentro e entre as lâminas .....	50
4.5 – Avaliação das alterações na expressão gênica induzida por radiação.....	58
4.5.1 – Genes diferentemente expressos e sua distribuição ao longo do tempo...	58
4.5.2 – Categorias de genes diferentemente expressos.....	62
4.6 – O efeito da irradiação na expressão das sequências obsoletas.....	66
4.7 – O efeito da irradiação na expressão dos genes RHS.....	70

4.8 – O efeito da irradiação na expressão de genes hipotéticos.....	73
4.9 – Agrupamento de genes baseado no perfil de expressão gênica ao longo do tempo.....	76
4.10 – Validação dos experimentos de microarranjos por PCR quantitativa em tempo real (qRT-PCR).....	78
4.11 - Análise funcional dos genes diferentemente expressos.....	84
4.11.1 – Efeitos da irradiação na expressão dos genes envolvidos na atividade motora e ciclo celular.....	94
4.11.2 – Efeito da irradiação na expressão dos genes envolvidos com a síntese e processamento de proteínas .....	95
4.11.3 – Efeito da irradiação na expressão dos genes envolvidos com a síntese de ATP.....	96
4.11.4 – Efeito da irradiação na expressão dos genes envolvidos com a resposta ao estresse oxidativo.....	97
4.11.5 – Efeito da irradiação na expressão dos genes envolvidos com o reparo de DNA.....	98
5 - Discussão.....	100
6 - Conclusões .....	114
7 – Referências bibliográficas .....	116
8 - Anexos.....	133
Anexo 1: Script em linguagem estatística R desenvolvido para as análises das lâminas de microarranjo usadas neste trabalho.....	133
Anexo 2: Escolha da melhor associação de métodos para a correção do <i>background</i> e normalização dos dados.....	136

Anexo 3 – Lista dos genes formadores dos clusters definidos pelo <i>heatmap</i> .....	140
Anexo 4 – Curvas de dissociação dos iniciadores dos genes usados nos experimentos de qRT-PCR. ....	143
Anexo 5 – Genes associados com processos de reparo de DNA que não apresentam valores de intensidade .....	145
Anexo 6 – Manuscrito referente a este trabalho submetido para a revista internacional indexada “PLoS Neglected Tropical Diseases” .....	147



## LISTA DE FIGURAS

---

Figura 1: O efeito da radiação ionizante sobre a molécula de DNA. ....	6
Figura 2: Resposta de <i>T. cruzi</i> à radiação ionizante. ....	11
Figura 3: Esquematização de um experimento de microarranjo. ....	13
Figura 4: Desenho experimental.....	23
Figura 5: Esquema de hibridização das lâminas.....	24
Figura 6: Imagens do programa <i>ScanArray Express</i> , usado para a leitura e análise das imagens das lâminas .....	28
Figura 7: <i>Pipeline</i> de análise dos dados de microarranjo após a leitura das lâminas .....	32
Figura 8: Curvas de crescimento de células epimastigotas de <i>T. cruzi</i> .....	39
Figura 9: Avaliação da integridade do RNA de <i>T. cruzi</i> após a irradiação. ....	41
Figura 10: Boxplots de cada lâmina ilustrando os valores de <i>background</i> e sinal brutos, e os valores do sinal corrigidos pelo método <i>Subtract</i> ou <i>NormExp</i> para o canal verde (Cy3).....	47
Figura 11: Boxplots de cada lâmina ilustrando os valores de <i>background</i> e sinal brutos, e os valores do sinal corrigidos pelo método <i>Subtract</i> ou <i>NormExp</i> para o canal vermelho (Cy5).....	48
Figura 12: MA <i>plots</i> para a lâmina 1 com diferentes métodos de normalização .....	53
Figura 13: <i>Boxplots</i> representando os valores de M ( $\log_2(R/G)$ ) para cada agulha de impressão.....	54
Figura 14: Distribuição empírica suavizada dos sinais dos canais verde e vermelho para todas as lâminas. ....	56
Figura 15: <i>Boxplots</i> representado os valores de M ( $\log_2(R/G)$ ) antes (primeira figura) e após a normalização (segunda figura) .....	57

Figura 16: Número de genes diferentemente expressos nos tempos definidos após a irradiação .....	59
Figura 17: Gráficos do tipo vulcão para os tempos logo após a irradiação – I.a.i (A), 4 horas (B), 24 horas (C), 48 horas (D) e 96 horas (E) após a irradiação, comparados com a amostra não irradiada (WT).....	61
Figura 18: Categorias dos genes diferentemente expressos e suas distribuições ao longo do tempo .....	63
Figura 19: Perfil de expressão gênica de quatro genes do maxicírculo após a irradiação.....	68
Figura 20: Frequência dos genes RHSs e não-RHS superexpressos em resposta à radiação gama ao longo do tempo.....	72
Figura 21: A) <i>Heatmap</i> de genes nucleares de <i>T. cruzi</i> diferentemente expressos. B) Agrupamentos de genes baseados nos perfis similares de expressão ao longo do tempo.. .....	77
Figura 22: Comparação entre os valores de <i>fold-change</i> obtidos por microarranjo (colunas pretas) e qRT-PCR (colunas cinzas).....	80
Figura 23: Correlação de Spearman entre os resultados de microarranjo e qRT-PCR. ....	81
Figura 24: Categorias funcionais Slim de ontologia gênica dos genes diferentemente expressos.....	85
Figura A1: <i>MA plots</i> para a lâmina 1 com diferentes métodos de normalização.. .....	137
Figura A2: Comparação entre dados brutos e normalizados com o método Robust-spline. <i>MA plots</i> para as 20 lâminas usadas neste trabalho.....	138
Figura A3: Gráfico de densidades de canais verdes e vermelhos.....	139
Figura A4: Curvas de dissociação dos iniciadores utilizados nos experimentos de qRT-PCR .....	144

## LISTA DE TABELAS

---

Tabela 1: Quantificação do RNA total antes e após o <i>clean-up</i> de RNA das amostras....	42
Tabela 2: Quantificação do mRNA amplificado (aRNA) e análise do seu grau de contaminação com proteínas. ....	44
Tabela 3: Quantidade de genes diferentemente expressos em cada tempo analisado. ....	58
Tabela 4: Valores de <i>fold-change</i> dos genes diferentemente expressos que codificam proteínas de função conhecida.....	64
Tabela 5: Valores de <i>fold-change</i> das sequências obsoletas diferentemente expressas ....	69
Tabela 6: Valores de <i>fold-change</i> dos genes RHS diferentemente expressos.....	71
Tabela 7: Valores de <i>fold-change</i> dos genes diferentemente expressos que codificam proteínas hipotéticas .....	74
Tabela 8: Sequências dos pares de iniciadores usados para a PCR quantitativa em tempo real, e os valores respectivos de slope, $R^2$ e eficiência.....	82
Tabela 9: Comparação entre os resultados significativos entre as técnicas de microarranjo e qRT-PCR .....	83
Tabela 10: Termos GO-Slim para as categorias de processos biológicos e função molecular; e os respectivos genes.....	86
Tabela 11: Valores de <i>fold-change</i> de genes participantes de vias de reparo de DNA.....	99

## LISTA DE ABREVIATURAS

---

aRNA – RNA mensageiro amplificado

BSA – Albumina de soro bovino (*Bovine serum albumin*)

cDNA - DNA complementar

CDS – Região codificadora de sequência

Cy3/Cy5 – cianinas 3/cianina 5

DDT – Ditiotreitól

DGF-1 – *Dispersed Gene Family - 1*

DSB - *Double Strand Break* (Quebra de fita dupla de DNA ( )

EDTA - Ácido etilenodiamino tetra-acético

EST –*Expressed Sequence Tag* (Etiqueta de sequência expressa)

FDR – *False Discovery Rate* (Taxa de descoberta falsa)

GAPDH – Gliceraldeído 3-fosfato desidrogenase

GO – *Gene ontology* (Ontologia gênica)

Gy – Gray

H<sub>2</sub>O<sub>2</sub> - Peróxido de hidrogênio

HO• - Radical hidroxil

HR – *Homologous Recombination* (Recombinação homóloga)

KCl – Cloreto de potássio

kDNA – DNA do cinetoplasto

kGy – kilo-gray

Limma – *Linear Models for Microarray Data*

mRNA – RNA mensageiro

NaCl – Cloreto de sódio

NaH<sub>2</sub>PO<sub>4</sub> – Fosfato monobásico de sódio

NaOH – Hidróxido de sódio

NCBI – *National Center for Biotechnology Information*

NHEJ – *Non-Homologous End-Joining* (Recombinação não-homóloga)

nt - nucleotídeo

O<sub>2</sub><sup>•-</sup> - Radical superóxido

pb – Pares de bases

PFGRC – *Pathogen Functional Genomics Resource Center*

qRT-PCR – *Real Time quantitative Reverse Transcription PCR* (PCR transcriptase reversa em tempo real)

RHS – proteínas retrotransposon *hot spot*

RMA – *Robust Multichip Average*

ROS - *Reactive Oxygen Species* (Espécies reativas de oxigênio)

SDS – Sódio dodecil sulfato

SSB – *Single Strand Break* (Quebra de fita simples de DNA)

SSC – Cloreto de sódio-Fosfato de sódio – tampão EDTA

TA – Temperatura ambiente

TIGR – *The Institute for Genomic Research*

## RESUMO

---

Entre os vários tipos de danos causados pela radiação ionizante, as lesões de maior importância biológica são as quebras de fita dupla de DNA. A tolerância a esse tipo de dano requer um eficiente aparato de reconhecimento e reparo de lesões e tais características definem o grau de resistência dos vários organismos a este tipo de estresse. O *Trypanosoma cruzi* é altamente resistente a raios gama, pois verificou-se que, após uma dose de 500 Gy de raios gama, o DNA genômico do parasito é totalmente fragmentado e o cariótipo é gradualmente reparado, restaurando o padrão de bandas cromossômicas em menos de 48 horas. Neste trabalho, objetivamos, através da técnica de microarranjos, comparar o padrão de expressão gênica de epimastigotas provenientes de duas réplicas biológicas diferentes da cepa CL Brener de *T. cruzi* irradiadas ou não com uma dose sub-letal de 500 Gy. Os resultados mostraram que a irradiação causou paralisação no crescimento das células em cultura, mas não afetou a integridade das moléculas de RNA, ao contrário do que foi relatado para as moléculas de DNA. A expressão gênica foi afetada de forma tempo-dependente, sendo que o pico de genes subexpressos (composto em sua maioria por genes de função conhecida) ocorreu após 4 horas, e a superexpressão (pico composto majoritariamente por genes de função desconhecida), ocorreu após 96 horas. Quatro genes do maxicírculo e os pertencentes à família de proteínas com *hot-spot* para inserção de retroelementos (RHSs) foram fortemente induzidos a partir de 48 horas, já os genes relacionados às funções metabólicas basais apresentaram expressão diminuída. Não houve indução na expressão dos genes de reparo, com exceção do gene tirosil-DNA fosfodiesterase -1.

## ABSTRACT

---

Among the various types of damages caused by ionizing radiation, the lesions of greatest biological importance are the double-stranded DNA breaks. The tolerance for this type of damage requires an efficient apparatus for recognition and repair of lesions and these characteristics define the degree of resistance of various organisms to this type of stress. *Trypanosoma cruzi* is highly resistant to gamma rays, considering that a 500 Gy dose induces a genomic DNA fragmentation, but the karyotype is gradually repaired and the chromosomal bands pattern is restored in less than 48 hours. In this study, we aimed to compare *T. cruzi* epimastigotes gene expression from two biological replicates irradiated or not with a sub-lethal dose of 500 Gy, through microarray experiments. Our results showed that irradiation caused an arrest in cell growth, but did not significantly affect the integrity of RNA molecules, contrary to what was seen in DNA molecules. The gene expression was affected in a time-dependent manner, because the peak of down-regulated genes (composed mostly of genes with known function) and up-regulated genes (composed mostly of genes of unknown function) occurred after 4 and 96 hours, respectively. Four maxicircle genes and those coding for retrotransposons hot spot protein (RHSs) were strongly induced after 48 hours. However, genes related to basal metabolic functions presented a decrease in the expression levels. There was no induction in the expression of DNA repair genes, excepted for the tyrosil-DNA phosphodiesterase 1 gene.

# 1 - INTRODUÇÃO

---

## 1.1. O *Trypanosoma cruzi* e a Doença de Chagas: aspectos históricos, epidemiológicos e biológicos

---

O *T. cruzi* é um protozoário flagelado, digenético, pertencente à ordem Kinetoplastida e à família Trypanosomatidae. Este protozoário possui uma mitocôndria única e alongada que contém uma estrutura chamada de cinetoplasto. O parasito apresenta diversas formas evolutivas (amastigotas, tripomastigotas e epimastigotas), que se alternam nos hospedeiros vertebrados e invertebrados dentro de um ciclo de vida complexo. Considerando a transmissão pelo inseto vetor (barbeiros triatomíneos), a infecção pelo *T. cruzi* inicia quando tripomastigotas metacíclicas eliminadas pelas fezes e urina do barbeiro vetor penetram no local da picada e invadem células do hospedeiro, onde se transformam em amastigotas, e se multiplicam por divisão binária simples. As amastigotas se diferenciam em tripomastigotas sanguíneas, e são liberadas no espaço intersticial podendo invadir outras células de qualquer órgão ou tecido iniciando um novo ciclo de reprodução assexuada ou são destruídas por mecanismos imunes. Alternativamente, as formas tripomastigotas sanguíneas podem ser ingeridas pelos triatomíneos vetores, transformando-se em formas epimastigotas em seu intestino posterior, onde se multiplicam por divisão binária e geram as formas tripomastigotas metacíclicas, as formas infectantes para o hospedeiro vertebrado (De Souza, 2002).

A doença de Chagas constitui um grave problema médico-social e uma das principais causas de morte súbita em pessoas ainda em fase produtiva. De acordo com a Organização Mundial de Saúde, há 10 milhões de pessoas infectadas, e pela primeira vez, assumiu-se que



a transmissão da doença não está mais confinada somente no continente americano (principalmente na América Latina), com transmissão confirmada em outros continentes. Mais de 25 milhões de pessoas estão sob o risco de infecção e estima-se que mais de 10.000 pessoas morreram vítimas da doença em 2010 (Who, 2010).

No Brasil, medidas de controle vetorial proporcionaram a eliminação da transmissão da doença pelo barbeiro *Triatoma infestans* (principal espécie vetora) e a transmissão transfusional sofreu uma importante redução após a implantação de medidas que visam o controle de doadores de sangue e de hemoderivados. Atualmente, predominam os casos crônicos da doença (estima-se que existam entre dois e três milhões de pessoas infectadas), e as formas de transmissão vertical (congênita) e oral ganharam importância epidemiológica, sendo responsáveis pela maior parte dos novos casos identificados (Ministério da Saúde, 2010).

A doença de Chagas apresenta duas fases bastante distintas. A fase inicial aguda é caracterizada por alta parasitemia, manifestações locais (chagoma de inoculação) e sintomas variados (febre, edema localizado e generalizado, poliadenia, hepatomegalia, esplenomegalia). A infecção avança para uma fase crônica, quando os níveis de parasitemia decaem. Os sobreviventes entram em uma fase assintomática (forma indeterminada) que pode durar toda a vida do paciente. Entretanto, em cerca de 30% dos pacientes, a doença evolui para uma fase sintomática que podem caracterizar a forma cardíaca, digestiva, cardiodigestiva ou nervosa da doença (Tanowitz *et al.*, 1992). Os pacientes permanecem infectados pelo resto da vida, e as formas clínicas apresentam características regionais distintas: em alguns países como o Panamá e a Venezuela, a infecção chagásica se manifesta exclusivamente na forma cardíaca, enquanto que no Brasil, Argentina e Chile, as manifestações clínicas são bastante variadas (Coura *et al.*, 1983).

## 1.2 – O projeto genoma de *T. cruzi*: características da cepa CL Brener

---

Em 2005, a seqüência completa do genoma de um clone da cepa CL Brener de *T. cruzi* foi publicada (El-Sayed *et al.*, 2005a). CL Brener, uma linhagem híbrida classificada como pertencente grupo *T. cruzi* VI (Zingales *et al.*, 2009) foi selecionada como referência para o projeto genoma por ser uma cepa muito bem caracterizada biologicamente e experimentalmente, como também para facilitar comparações com os dados gerados no projeto de EST (*Expressed Sequence Tag*) de CL Brener que estava em andamento na época (Brandao *et al.*, 1997; Verdun *et al.*, 1998).

A anotação da seqüência completa do genoma do clone da cepa CL Brener indicou a existência de 22.570 genes codificando proteínas, sendo que desses, 6.159 representam alelos presentes no haplótipo do tipo Esmeraldo, 6.043 representam alelos provenientes de haplótipo do tipo não-Esmeraldo, e 10.368 representam seqüências que não puderam ser relacionadas a nenhum haplótipo em particular. Deste total, foi possível sugerir a função de 50,8% das seqüências codificadoras com base na literatura e em resultados de similaridade com proteínas já caracterizadas, ou devido à presença de domínios funcionais característicos (El-Sayed *et al.*, 2005a). Elementos repetitivos consistindo de grandes famílias gênicas de proteínas de superfície, retrotransposons, e repetições subteloméricas respondem por pelo menos 50% do genoma da cepa CL Brener (El-Sayed *et al.*, 2005a). A finalização do genoma do *T. cruzi* e as análises comparativas com os genomas de *Trypanosoma brucei* (Berriman *et al.*, 2005) e *Leishmania major* (Ivens *et al.*, 2005) complementaram os conhecimentos sobre o genoma desses organismos.

Possivelmente por terem divergido cedo da linhagem evolutiva que originou os demais eucariotos (Vickerman, 1994), os kinetoplastídeos apresentam características não usuais, ou características biológicas únicas relacionadas à expressão gênica. Um número

muito reduzido de promotores gênicos foi caracterizado e, de fato, o controle da expressão gênica ocorre principalmente ao nível pós-transcricional. Os genomas dos tripanosomatídeos são organizados em longos grupos ou *clusters* gênicos direcionais, os quais são co-transcritos em longas unidades policistrônicas (Martinez-Calvillo *et al.*, 2003; El-Sayed *et al.*, 2005b). A geração de mRNA monocistrônicos maduros em tripanosomatídeos envolve dois eventos de clivagem que resultam na adição de uma seqüência de 39 nt denominada mini-éxon ou *spliced-leader* contendo o CAP na extremidade 5' e adição de cauda poli-A na extremidade 3' dos transcritos (Clayton, 2002; Teixeira e Da Rocha, 2003). Um dos mecanismos de controle da expressão gênica em tripanosomatídeos mais bem caracterizado envolve a presença de seqüências regulatórias nas regiões não traduzidas dos mRNAs que modulam a estabilidade dos transcritos e/ou eficiência de tradução (Nozaki e Cross, 1995; Teixeira *et al.*, 1995; Abuin *et al.*, 1999; Weston *et al.*, 1999; Di Noia *et al.*, 2000; Bartholomeu *et al.*, 2002).

### 1.3 – O efeito da radiação ionizante e o processo de reparo de quebra de fita dupla nas moléculas de DNA

---

A radiação ionizante é caracterizada por ondas eletromagnéticas e partículas capazes de ionizar, ou seja, remover um elétron de um átomo ou de uma molécula presente no meio em que ela está sendo propagada (Unsear, 2000). A radiação ionizante pode ser emitida como consequência do processo natural de decaimento de alguns átomos com núcleo instável, ou em função da excitação dos átomos e de seus núcleos por reatores nucleares, ciclotrons, máquinas de raios-X, ou outros instrumentos (Unsear, 2000). Há vários tipos de radiação ionizante, dentre elas: as partículas alfa, as partículas beta, os nêutrons, os raios-X e os raios gama, sendo o último, o de interesse neste trabalho.

A radiação ionizante é capaz de causar dano oxidativo às bases nucleotídicas, assim como quebras de fita simples (SSB) ou dupla (DSB) no DNA (revisado por Slupphaug *et al.*, 2003). As DSBs são letais, exceto quando reparadas (Pfeiffer *et al.*, 2000). A molécula de DNA pode ser danificada devido ao contato direto com os raios ionizantes (ação direta), ou indiretamente, pelas espécies reativas de oxigênio geradas pela radiação (Figura 1).

As espécies reativas de oxigênio (ROS) mais importantes são o radical superóxido ( $O_2^{\cdot-}$ ), gerado pela remoção de um elétron do oxigênio, o peróxido de hidrogênio ( $H_2O_2$ ), gerado pela remoção de dois elétrons quando o oxigênio se encontra totalmente protonado, e o radical hidroxil ( $HO^{\cdot}$ ), o mais reativo e menos seletivo de todos gerado por uma redução adicional do oxigênio (Cuzzocrea *et al.*, 2001). Por muito tempo acreditou-se que a toxicidade atribuída aos radicais superóxidos era causada pela interação direta com moléculas biológicas. Hoje se sabe que muitos dos efeitos teciduais da ação do oxigênio é resultado da sua ação direta somado com a formação secundária de espécies reativas de oxigênio que juntos interagem com vários alvos moleculares, como os lipídeos (Aikens e Dix, 1991), catecolaminas (Macarthur *et al.*, 2000), e DNA (Dix *et al.*, 1996).

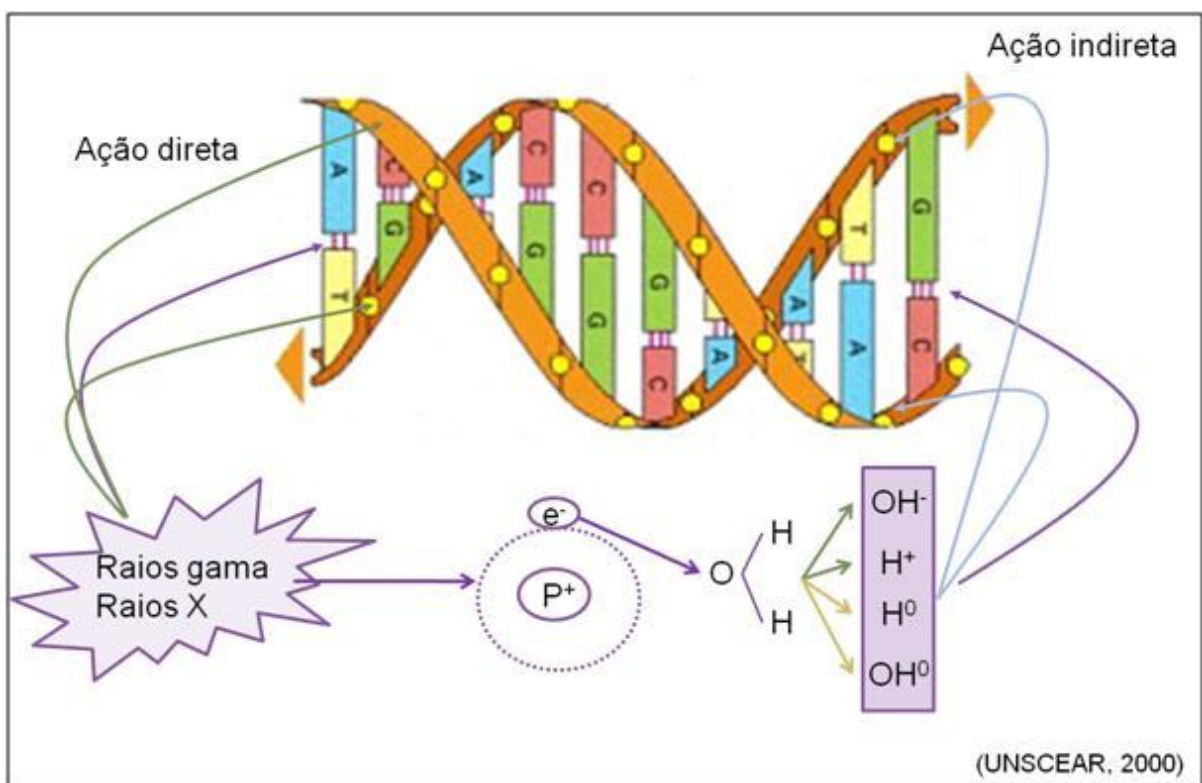


Figura 1: O efeito da radiação ionizante sobre a molécula de DNA.

## 1.4 – As vias de reparo de DNA

---

As principais vias de reparo de DNA são por excisão de base (BER), por excisão de nucleotídeo (NER), por erro de pareamento, por via direta, além da síntese translesão e reparo de quebras de fita dupla de DNA. Todas elas estão presentes em *T. cruzi* (revisado por Passos-Silva *et al.*, 2010). O processo de reparo de DSBs já é bastante conhecido em diversos organismos eucariotos, e ocorre tanto por mecanismos de recombinação homóloga (*homologous recombination*) quanto não-homóloga (*non-homologous end-joining*) (Pastink *et al.*, 2001).

A recombinação homóloga (HR) é um tipo de rearranjo genético que ocorre através da quebra e junção de moléculas de DNA em regiões idênticas, ou muito similares. A HR é a principal via de reparo de DSB em eucariotos inferiores (Bhattacharyya *et al.*, 2004) e contribui com a geração de variabilidade genética nos tripanosomatídeos (Machado *et al.*, 2006). Os genes que codificam para as proteínas MRE11, Rad50, NSB1 e RPA, participantes da etapa de inicial da HR, foram encontrados nos tripanosomatídeos (Berriman *et al.*, 2005; El-Sayed *et al.*, 2005a; Ivens *et al.*, 2005). O passo central da HR é caracterizado pela busca por homologia, pareamento de DNA homólogo, e pela reação de troca de fitas. Esta última etapa é mediada por recombinases como a RAD51 e DMC1, ambas também presentes nos tripanosomatídeos (Berriman *et al.*, 2005; El-Sayed *et al.*, 2005a; Ivens *et al.*, 2005). Por outro lado, a recombinação não-homóloga (NHEJ) conecta as pontas da molécula de DNA sem a necessidade de existir qualquer similaridade entre as partes, sendo considerada não-conservativa (Moore e Haber, 1996). No entanto, NHEJ aparentemente é ausente em *T. cruzi*, uma vez que apenas quatro de oito genes envolvidos na via (Mre11, Rad50, KU70 e KU80) foram identificados nos tripanosomatídeos (Berriman *et al.*, 2005; El-Sayed *et al.*, 2005a; Ivens *et al.*, 2005). Além disso, estudos em *T. brucei* mostraram que estes genes atuam

principalmente na manutenção do telômero e mutantes para os genes KU70 e KU80 não mostraram sensibilidade aumentada quando tratados com agentes causadores de danos ao DNA (Conway *et al.*, 2002; Janzen *et al.*, 2004). Por fim, não foram encontrados nos tripanosomatídeos os genes codificadores para a DNA ligase IV e XRCC4/Lif1, que fazem parte da NHEJ (Williams e Lustig, 2003). Uma diferença importante entre os dois mecanismos de reparo é que o HR é essencialmente livre de erro, enquanto que o NHEJ pode eventualmente gerar inserção ou deleção de alguns nucleotídeos (revisado por Slupphaug *et al.*, 2003).

#### 1.5 – O efeito da radiação ionizante em *T. cruzi* e em outros organismos

---

Há diversos indícios que mostram que o *T. cruzi* é um dos organismos eucarióticos mais resistentes à irradiação ionizante. Estudos da década de 80, que tinham como objetivo desenvolver protocolos de atenuação de parasitos, já indicavam que *T. cruzi* apresentava uma resistência não usual a raios gama (Takeda *et al.*, 1986). De fato, formas epimastigotas são capazes de tolerar doses de irradiação semelhantes às toleradas por leveduras (800 Gy) (Bennett *et al.*, 2001), doses 10 vezes maiores às toleradas por *E. coli* (Bennett *et al.*, 2001) e aproximadamente 50-100 vezes maiores que as toleradas por raízes de plantas (Jiang *et al.*, 1997), *Plasmodium* (Geoghegan *et al.*, 2001), células de mamíferos (Takata *et al.*, 2001; Yonetani *et al.*, 2005) e por *Trypanosoma brucei*, o organismo mais próximo filogeneticamente de *T. cruzi* em que se estudou a tolerância aos raios gama (Tan *et al.*, 2002). Gray (Gy) é a unidade internacional de medida de absorção de dose de radiação proveniente de fontes de radiação ionizante. Um Gy é a absorção de um joule de energia por quilograma do meio absorvedor. *Leishmania major*, outro kinetoplastídeo, é similarmente

resistente aos raios gama, embora seja um pouco mais sensível do que *T. cruzi*, uma vez que células são capazes de retomar o crescimento após uma dose de 300 Gy (Seo *et al.*, 1993).

Há vários procariotos que apresentam resistência maior que *T. cruzi*, como a bactéria *Pyrococcus furiosus*, que tolera doses de 2,5 kGy (Williams *et al.*, 2007), a cepa NRC1 de *Halobacterium sp.* e a cianobactéria fixadora de nitrogênio *Anabaena sp.* (cepa L-31), uma vez que ambos sobrevivem a doses de 5 kGy (Kottemann *et al.*, 2005; Singh *et al.*, 2010), diversas espécies do gênero *Deinococcus*, especialmente a *D. radiodurans*, que não perde a viabilidade mesmo após doses de 15 kGy (Liu *et al.*, 2003), e *Kineococcus radiotolerans*, cuja curva de crescimento é retomada aproximadamente 100 horas após uma dose de 20 kGy (Bagwell *et al.*, 2010). Em um estudo com *D. radiodurans*, através da técnica de microarranjos e validação pela técnica de PCR quantitativa em tempo real (qRT-PCR), foi possível observar que nos primeiros momentos após a irradiação, os genes envolvidos na replicação do DNA, no reparo, na recombinação, no metabolismo da parede celular, no transporte celular, e vários outros codificadores de proteínas hipotéticas foram induzidos (Liu *et al.*, 2003). Em outro estudo, foi visto que três genes que codificam proteínas com função desconhecida e que parecem fazer parte nos processos de reparo de DNA apresentam expressão aumentada após a irradiação e consequente fragmentação do DNA, além de uma recombinase A (*recA*), uma gliceraldeído 3-fosfato desidrogenase (*gap*), uma proteína de reparo de quebra de fita-dupla Rad52/22 e uma proteína de reparo de dano de DNA (Tanaka *et al.*, 2004).

O grupo do Dr. Carlos Renato Machado e Dra. Santuza Teixeira do Departamento de Bioquímica e Imunologia do ICB, UFMG, tem trabalhando na caracterização de genes do parasito envolvidos nos mecanismos de reparo do DNA danificado (Augusto-Pinto *et al.*, 2001; Augusto-Pinto *et al.*, 2003; Machado *et al.*, 2006; Regis-Da-Silva *et al.*, 2006; Lopes *et al.*, 2008; Machado-Silva *et al.*, 2008; De Moura *et al.*, 2009; Rajao *et al.*, 2009; Campos *et*



*al.*, 2010). Foi verificado por Regis-da-Silva e colaboradores em 2006 que as bandas cromossômicas dos parasitos submetidos a 500 Gy de irradiação gama sofrem extensa fragmentação (sobrevivência de aproximadamente 50%). Porém, 48 horas após o processo, o padrão de bandas é restabelecido, evidenciando o reparo do DNA (figura 2). Além disso, foi demonstrado que em organismos da cepa CL Brener que superexpressam o gene TcRAD51, codificador de uma proteína com papel importante no reparo de quebra dupla de DNA, o reparo das duplas-fitas ocorreu mais rapidamente, sendo possível visualizar as bandas reconstituídas 24 horas após a irradiação, como também pode ser visto na figura 2. O crescimento celular começa a ser evidenciado 240 horas após o estresse (Regis-Da-Silva *et al.*, 2006). Outro estudo demonstrou que a superexpressão de TcPol $\kappa$ , uma DNA polimerase kappa de baixa fidelidade, que possui a capacidade de ignorar diversos tipos de lesões, também foi capaz de aumentar a resistência de *T. cruzi* frente à radiação gama e zeocina, outro agente capaz de causar quebras de fita-dupla (Rajao *et al.*, 2009). No entanto, nenhum estudo em larga escala foi realizado até o momento para investigar o padrão de expressão gênica em resposta à radiação gama em *T. cruzi*.

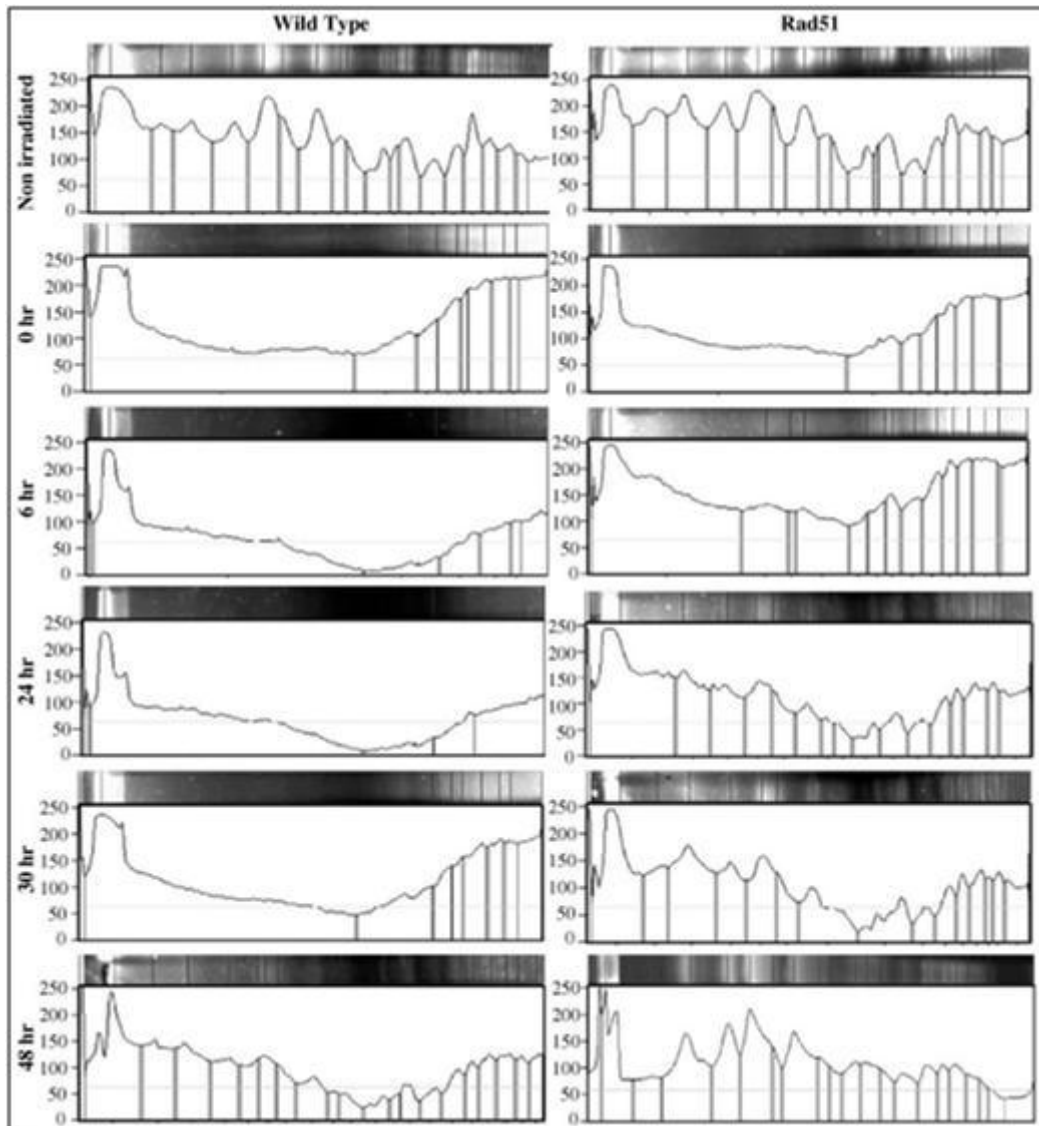


Figura 2: Resposta de *T. cruzi* à radiação ionizante. Gel de *Pulsed Field Gel Eletrophoresis* (PFGE) corado com brometo de etídeo dos cromossomos de *T. cruzi* após exposição à dose de 500 Gy de radiação gama (em epimastigotas selvagens ou superexpressando o gene RAD51). Para cada tempo é mostrado a análise densitométrica (linhas verticais) das bandas (figura retirada do trabalho de Regis-Da-Silva et al., 2006)

## 1.6 - A técnica de Microarranjos

---

Nos últimos 15 anos temos testemunhado um grande avanço nas pesquisas biológicas em relação à quantidade de informação disponível com o advento dos projetos genoma. No entanto, o simples conhecimento das sequências das bases de DNA muitas vezes não nos permite inferir funções dos genes. Apenas um subconjunto de todos os genes em uma determinada célula e em um dado momento é expresso, sendo que os níveis e tipos de genes expressos governam o desenvolvimento, diferenciação, função e fisiologia celular. A alteração do transcriptoma é uma das mais rápidas e versáteis respostas de um organismo frente a mudanças do seu ambiente ou a estresses induzidos. Um dos grandes objetivos de estudos de genômica funcional é permitir, com a utilização de novas tecnologias, investigar em larga escala o padrão espacial e temporal da expressão de milhares de genes e até mesmo de genomas inteiros.

O microarranjo, cuja técnica permite o monitoramento dos níveis de expressão de milhares de genes simultaneamente em uma célula, é utilizado para elucidar uma pergunta biológica cuja resposta é obtida através da incubação de uma lâmina contendo seqüências de DNA imobilizadas (sondas) junto com uma solução contendo cDNAs ou RNA marcados que são gerados/obtidos a partir da amostra biológica em estudo (figura 3). A idéia central por trás do método é de que os cDNAs ou RNAs da amostra biológica hibridizarão às seqüências de DNA imobilizados na lâmina pelo princípio da complementaridade (Draghici, 2003).

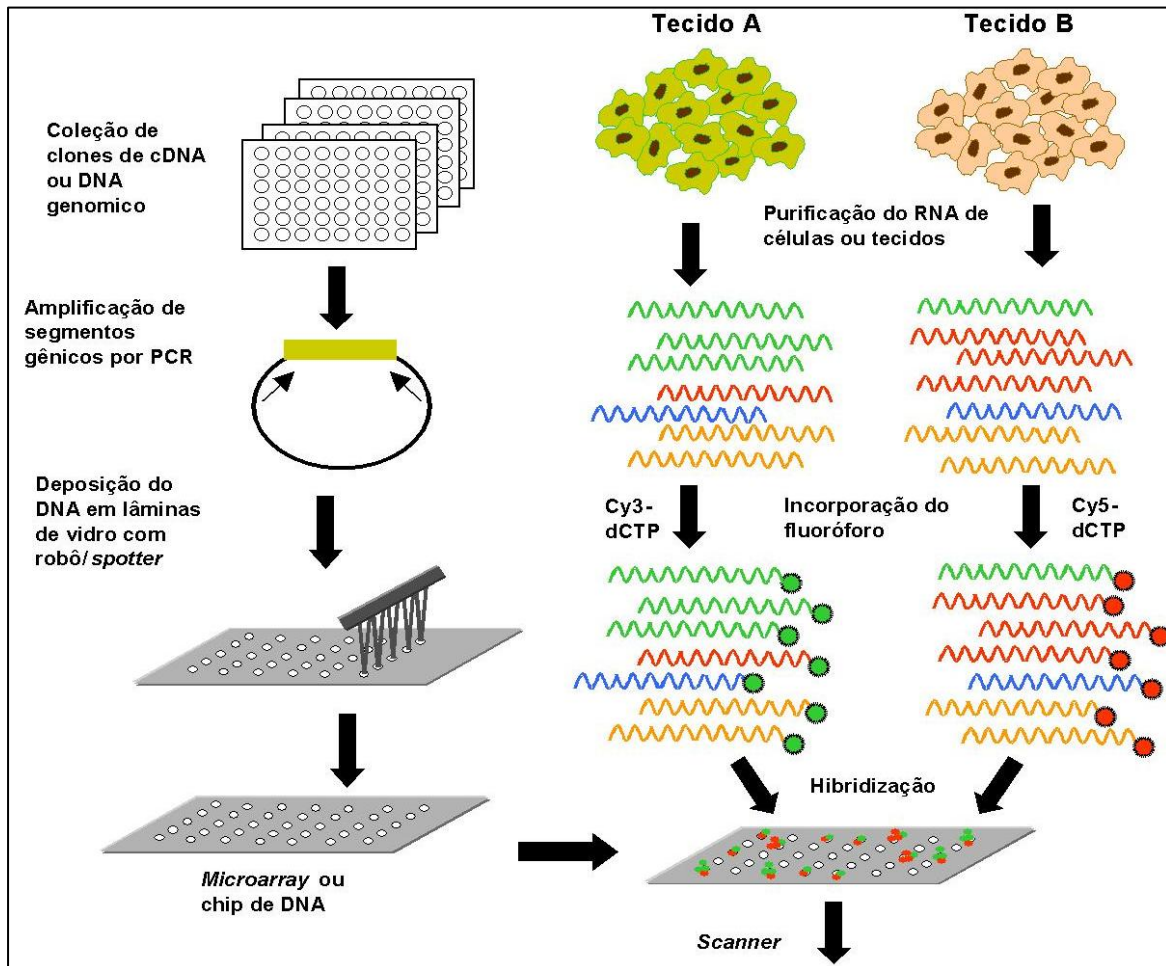


Figura 3: Esquematização de um experimento de microarranjo. As lâminas são produzidas através da impressão de sondas que podem ser segmentos genômicos amplificados ou oligonucleotídeos sintetizados. Elas recebem mRNA marcado com diferentes fluoróforos extraídos a partir de amostras biológicas, sendo uma considerada controle e a outra, teste. Figura retirada do livro *Genômica* (Ed. Atheneu, 2004).

Em um experimento clássico em microarranjos de cDNA ou de oligonucleotídeos, duas amostras (amostra controle e amostra teste) são marcadas com fluoróforos diferentes (por exemplo, cianina 3 ou Cy3 e cianina 5 ou Cy5) e estas são usadas ao mesmo tempo em um processo de hibridização competitiva (figura 3). Desta maneira, a razão dos valores de intensidade obtidos pela leitura do laser de cada cor para cada elemento informa sobre as diferenças relativas de expressão gênica entre duas amostras quaisquer.

Muitos são os desafios encontrados pelos pesquisadores que utilizam a técnica de microarranjos. Devido à sua natureza, os ruídos são introduzidos a cada passo, em cada procedimento: preparação do mRNA (tipos de tecidos, marcas de kits e procedimentos variados), transcrição reversa (variações inerentes à reação e enzimas), marcação (tipo e idade dos reagentes, e variações na taxa de incorporação dos fluoróforos), amplificação do mRNA (variação na temperatura ambiente), cobertura química da lâmina, umidade relativa do ar, tipo de impressão e volume das sondas a serem impressas na lâmina (pode haver variação até na mesma agulha usada para imprimir), heterogeneidade das lâminas (variações na produção), fixação das sondas, parâmetros de hibridização (tempo, temperatura, solução utilizada), hibridização inespecífica (mRNA marcado hibridado em áreas onde não houve pareamento perfeito), escaneamento (configurações do *PMT gain*, alinhamento entre-canais), segmentação (separação entre sinal e *background*) e quantificação dos *spots* (Schuchhardt *et al.*, 2000; Draghici, 2003).

Para exemplificar uma das formas de ruídos que inevitavelmente são inseridos durante os procedimentos laboratoriais, uma fonte de variação sistemática vem da associação das diferentes sondas fluorescentes. Isso pode ser verificado em experimentos onde duas amostras de mRNA idênticas são marcadas com diferentes fluoróforos e subsequentemente hibridizadas numa mesma lâmina (Dudoit, 2002). Nesse instante, é raro obter intensidade dos fluoróforos iguais por todos os *spots* entre as duas amostras. A variação dos fluoróforos pode

ser devido a diversos fatores, incluindo propriedades físicas de cada um (sensibilidade ao calor e a luz, tempo de meia vida), eficiência de incorporação de cada fluoróforo, variabilidade experimental nos procedimentos e processos de hibridização ou configuração do *scanner* na etapa de coleta de dados. Entretanto, os níveis de expressão gênica relativa das réplicas experimentais podem ter variações nas amostras devido a diferentes condições de experimento. Muitos desses fatores, internos ou externos às amostras alvo, dificultam a distinção entre genes diferentemente e constantemente expressos (Scharpf *et al.*, 2007). O fato é que o ruído é um fenômeno real e é impossível evitá-lo, restando aos pesquisadores recorrerem às réplicas técnicas do experimento (repetição) e aplicação de testes estatísticos para neutralizá-lo.

Para que os efeitos inerentes a técnica sejam minimizados são necessárias correção de intensidade de *background* (ruído) e normalização dos dados. Softwares de análise de imagens retornam a intensidade do sinal (*foreground*), e intensidade de fundo (*background*) de cada *spot*. O sinal é uma medida global de intensidade de cada *spot*, e assume-se que é a soma entre o sinal verdadeiro e o *background* local, enquanto que *background* é uma medida de ruído do ambiente (que no caso dos experimentos de microarranjos, é a medida do sinal ao redor do *spot*). A fluorescência de *background* pode surgir de diversas fontes, como ligações não específicas das sondas na superfície dos *arrays*, efeitos de processamento como depósito de sondas após a lavagem ou ruído óptico do *scanner*. Já a proposta da normalização é ajustar os efeitos advindos tanto das diferenças biológicas entre as amostras de mRNA quanto entre as sondas utilizadas. O sinal das sondas pode variar de acordo com a posição espacial na lâmina, e isso pode ocorrer em função das diferenças entre as agulhas na impressão do *array*, variações no curso da corrida, a não uniformidade das hibridações, ou artefatos na superfície da lâmina que afeta uma cor mais que a outra. Finalmente, as variações entre os *arrays* podem surgir das diferenças na qualidade da impressão, das diferenças nas condições

ambientais onde as lâminas são processadas, ou simplesmente por mudanças nas configurações do *scanner* no momento da leitura do sinal. Portanto, a normalização dentro e entre as lâminas são cruciais (Smyth e Speed, 2003).

Vários estudos de padrão temporal de expressão e a associação entre os genes e seus produtos têm sido publicados usando experimentos do tipo *time-course*, incluindo periodicidade (Filkov *et al.*, 2002; Wichert *et al.*, 2004), detecção de co-expressão através do cálculo do coeficiente de correlação e suas variantes (Schafer e Strimmer, 2005; Zhu *et al.*, 2005), agrupamentos gênicos (Peddada *et al.*, 2003; Schliep *et al.*, 2003; Song *et al.*, 2007), desenvolvimento de testes de casualidade para construção de redes de interação gênica (Mukhopadhyay e Chatterjee, 2007), e aplicação de periodograma de Laplace para detecção de expressão gênica periódica para conjunto de dados contendo *outliers* e os intervalos entre os tempos não são uniformes (Liang *et al.*, 2009).

### 1.7 - Aplicação da técnica de microarranjo de DNA em estudos com o *T. cruzi*

---

Diversos estudos utilizando a técnica de microarranjos já foram realizados na tentativa de elucidar diferentes aspectos biológicos do *T. cruzi*. Parte desses trabalhos visou avaliar os padrões de expressão gênica do hospedeiro (murino ou humano) frente à infecção por este parasito. A resposta genética do hospedeiro murino (Garg *et al.*, 2003; Mukherjee *et al.*, 2003; Mukherjee *et al.*, 2008; Soares *et al.*, 2010) e humano (Nde *et al.*, 2010) frente à miocardite induzida pela infecção com o parasito em diferentes estágios da doença forneceu novas informações sobre o desenvolvimento da miocardite chagásica. Investigou-se também a habilidade do *T. cruzi* em modular a expressão dos genes da via apoptótica TNFR1 em camundongos Balb/c (Moore-Lai e Rowland, 2004), e as alterações nos padrões de expressão gênica durante a formação intracelular de ninhos de amastigotas em fibroblastos (Imai *et al.*,

2005), em células humanas da linhagem HeLa (Shigihara *et al.*, 2008) e macrófagos (Zhang *et al.*, 2010).

Outros trabalhos buscaram contribuir para o conhecimento sobre a biologia do parasito. Utilizando a tecnologia Affymetrix<sup>TM</sup>, mudanças que ocorrem durante a transição do estágio de tripomastigota para amastigota foram caracterizadas (Minning *et al.*, 2003). O primeiro trabalho com lâminas contendo cDNAs da cepa CL Brener mostrou que foi possível utilizar esses microarranjos em estudos de genômica comparativa e análise da expressão gênica em *T. cruzi* (Baptista *et al.*, 2004). Dessa forma, foi publicado um trabalho caracterizando a alteração da expressão de certos genes em parasitos tratados com o agente terapêutico para a doença de Chagas, o benzonidazole (Murta *et al.*, 2006). Foi realizada também a comparação entre parasitos isolados de humanos assintomáticos e de pacientes cardíacos na tentativa de compreender de forma mais aprofundada o motivo que leva 20-30% dos indivíduos com infecção crônica a desenvolverem os sintomas da doença de Chagas. Os autores demonstraram que o gene que codifica a subunidade NADH desidrogenase do cinetoplasto (kDNA) apresenta expressão 30 vezes maior nas cepas cardíacas do que nas cepas assintomáticas (Baptista *et al.*, 2006). Dallagiovanna e colaboradores (Dallagiovanna *et al.*, 2008) buscaram identificar os mRNAs que se associavam a *TcPUF6*, proteína que possui a função predita de regular a estabilidade e tradução do mRNA. As conclusões desse estudo sugerem que o gene *TcPUF6* regula a meia vida dos seus próprios transcritos associados através de diferentes formas de associação da proteína TcPUF6 a complexos protéicos de degradação de mRNA durante o ciclo de vida do parasito. Em 2009 foi publicado um estudo onde a expressão gênica nos diferentes estágios evolutivos do parasito foi avaliada (Minning *et al.*, 2009). A principal conclusão deste trabalho foi que, ao contrário do que se acreditava e se evidenciara em estudos genômicos e proteômicos (Avila *et al.*, 2003; Minning *et al.*, 2003; Baptista *et al.*, 2004; Paba *et al.*, 2004a; Paba *et al.*, 2004b; Parodi-Talice *et al.*, 2004;



Atwood *et al.*, 2005; Atwood *et al.*, 2006; Parodi-Talice *et al.*, 2007; Koumandou *et al.*, 2008), mais de 50% dos genes detectados pelos microarranjos foram significativamente alterados durante as mudanças de fase do ciclo (Minning *et al.*, 2009).

Nesta tese, aplicamos a tecnologia de microarranjos para analisar as mudanças na expressão gênica de epimastigotas submetidas a doses sub-letais de radiação gama. Processos biológicos como interações bioquímicas e atividades regulatórias envolvem relações de dependência complexas entre os genes. Dessa forma, experimentos de microarranjos do tipo *time course* fornece uma forma inovadora de estudar a natureza dinâmica da dependência entre os genes (Gao *et al.*, 2009).

## 2 - OBJETIVOS

---

### 2.1 – Objetivo Geral

---

Acompanhar e caracterizar os padrões de expressão gênica em clones selvagens da cepa CL Brener submetidas ou não a uma dose de 500 Gy de radiação gama por cinco tempos diferentes: logo após, 4, 24, 48, e 96 horas após a irradiação utilizando para isso a técnica de microarranjos.

### 2.2 – Objetivos específicos

---

- Comparar os métodos comumente usados para a correção de *background* e para a normalização em cada lâmina em dados de experimentos de microarranjo de DNA de *Trypanosoma cruzi* irradiados com radiação gama em diferentes tempos;
- Identificar conjuntos específicos de genes que apresentam perfis semelhantes de expressão gênica ao longo dos tempos analisados;
- Identificar processos biológicos e funções moleculares que foram alterados, ou seja, induzidos ou reprimidos, frente à radiação gama;
- Validar os níveis de expressão gênica em uma amostragem de genes utilizando a PCR quantitativa em tempo real

### 3 – MATERIAL E MÉTODOS

---

#### 3.1 – Cultura de *T. cruzi*

---

Epimastigotas da cepa CL Brener foram cultivadas em meio LIT (*Liver Infusion Tryptone - liver digest neutralized* 5g/L, NaCl 4g/L, bactotriptose 5g/L, KCl 0,4g/L, NaH<sub>2</sub>PO<sub>4</sub> 8g/L, dextrose 2g/L, hemin 0,2% 10ml em NaOH 0,05M, pH 7,3) contendo 10% de soro fetal bovino (Cultilab) e os antibióticos penicilina (100U/ml) e estreptomicina (100µg/ml) (Invitrogen). A cultura foi mantida a 28°C em frascos apropriados e em sua fase exponencial de crescimento, através de repiques semanais. Para contagem do número de parasitos, esses foram diluídos em PBS (0,15 M pH 7,2) e eritrosina (4% em PBS) e posteriormente, os parasitos vivos foram contados em câmara citométrica.

#### 3.2 – Irradiação das células com radiação gama

---

As garrafas de cultura contendo células epimastigotas na concentração de  $2,0 \times 10^7$  células/mL de meio LIT (ou  $5,0 \times 10^8$  células totais) foram submetidas a uma dose de 500 Gy de radiação gama proveniente de uma fonte de cobalto (<sup>60</sup>Co), disponível no CNEN/CDTN, um instituto de pesquisa do Ministério de Ciência e Tecnologia localizado no *campus* da UFMG. A dose no valor de 500 Gy foi determinada por estudos anteriores do nosso grupo (Regis-Da-Silva *et al.*, 2006). Essas doses foram obtidas através da exposição a uma taxa de 1.548 Gy/h por 20 min. Após a irradiação, em tempos determinados, os parasitos foram contados em câmara citométrica e submetidos à extração de RNA.

### 3.3 – Extração, purificação (*Clean-up*) e amplificação de RNA

---

As células, suspensas em 20 mL de meio LIT, foram transferidas para um tubo falcon de 50 mL, e submetidas à centrifugação a 3.000g, por 10 minutos, a 4°C. O sobrenadante foi descartado, e 20 mL de PBS 1X estéril foram adicionados ao sedimento. Seguiu-se mais uma rodada de centrifugação nas mesmas condições e o sobrenadante foi novamente descartado. O RNA foi extraído utilizando o reagente Trizol®, seguindo as instruções do fabricante. Os RNAs foram quantificados no aparelho NanoDrop ND-100 UV/Vis (NanoDrop Technologies, USA) e a qualidade foi avaliada usando um gel de agarose desnaturante.

Para experimentos de análise de expressão gênica, é muito importante que o RNA esteja livre de impurezas e contaminantes, como moléculas de DNA e compostos protéicos. Para isso, a purificação das amostras de RNA foi realizada utilizando o RNeasy® MiniElute™ Cleanup Kit, da marca Qiagen (Alemanha), seguindo as instruções do fabricante (protocolo para limpeza de até 45 µg). As amostras limpas foram quantificadas com o aparelho NanoDrop ND-100 UV/Vis (NanoDrop Technologies, USA) e visualizadas em gel de agarose desnaturante 1,2% a fim de verificar a qualidade do RNA total. 2 µg de RNA total foram amplificados de cada amostra (*pool*) usando o kit AminoAllyl MessageAmp II (Ambion, Inc, USA), segundo instruções do fabricante. Brevemente, moléculas de mRNA foram retrotranscritas com o iniciador oligo(dT) T7 em cDNA, e seguiu-se para a conversão de cDNA fita-simples em fita-dupla contendo o promotor T7, que atuou como molde para a reação de transcrição *in vitro*. Esta etapa caracteriza-se pela síntese de milhares de cópias (amplificação) de aRNA modificado com aminoalil que, após etapas de purificação, foi usado para as hibridizações.

### 3.4 – Design experimental

---

Para avaliar o efeito da radiação gama em células epimastigotas de *T. cruzi*, dois experimentos idênticos, porém independentes, foram realizados. Cada experimento (réplica biológica) foi feito em triplicata, onde epimastigotas foram cultivadas em sete grupos de três garrafas de cultura (figura 4). O tempo de intervalo entre as duas réplicas biológicas foi de um mês. As garrafas do grupo C-I não foram irradiadas, e suas células não foram usadas para os experimentos de microarranjo de DNA, servindo apenas para o controle da curva de crescimento da cultura de células epimastigotas. As garrafas do grupo C-II também não foram irradiadas, mas suas células tiveram o seu RNA extraído, compondo a amostra referência (controle) dos experimentos de microarranjo de DNA. As garrafas dos grupos 1 a 5 foram irradiadas, e o RNA das células foi extraído em vários intervalos de tempo após a irradiação (logo após e 4, 24, 48 e 96 horas após a irradiação). O RNA total extraído das três réplicas de cada conjunto de amostras foi juntado em um mesmo *pool* no momento da etapa de *clean up*, onde foi tomado o cuidado de colocar quantidades idênticas, ou muito próximas de RNA total de cada amostra. Assim cada *pool* foi formado por 1/3 de RNA total de cada amostra. Todas foram subsequentemente purificadas (*clean-up*). Cada amostra de cada réplica biológica (dos dois experimentos idênticos independentes) foi hibridizada contra a amostra controle da sua réplica biológica respectiva duas vezes, na forma de *dye-swap*, ou seja, em uma lâmina o controle foi marcado com Cy3 e a amostra teste foi marcada com Cy5, e na outra lâmina, o controle foi marcado com Cy5, e a amostra teste, com Cy3 (figura 5). Dessa forma, para um dos cinco tempos do experimento, totalizamos quatro lâminas: duas lâminas (*dye-swap*) de cada réplica biológica.

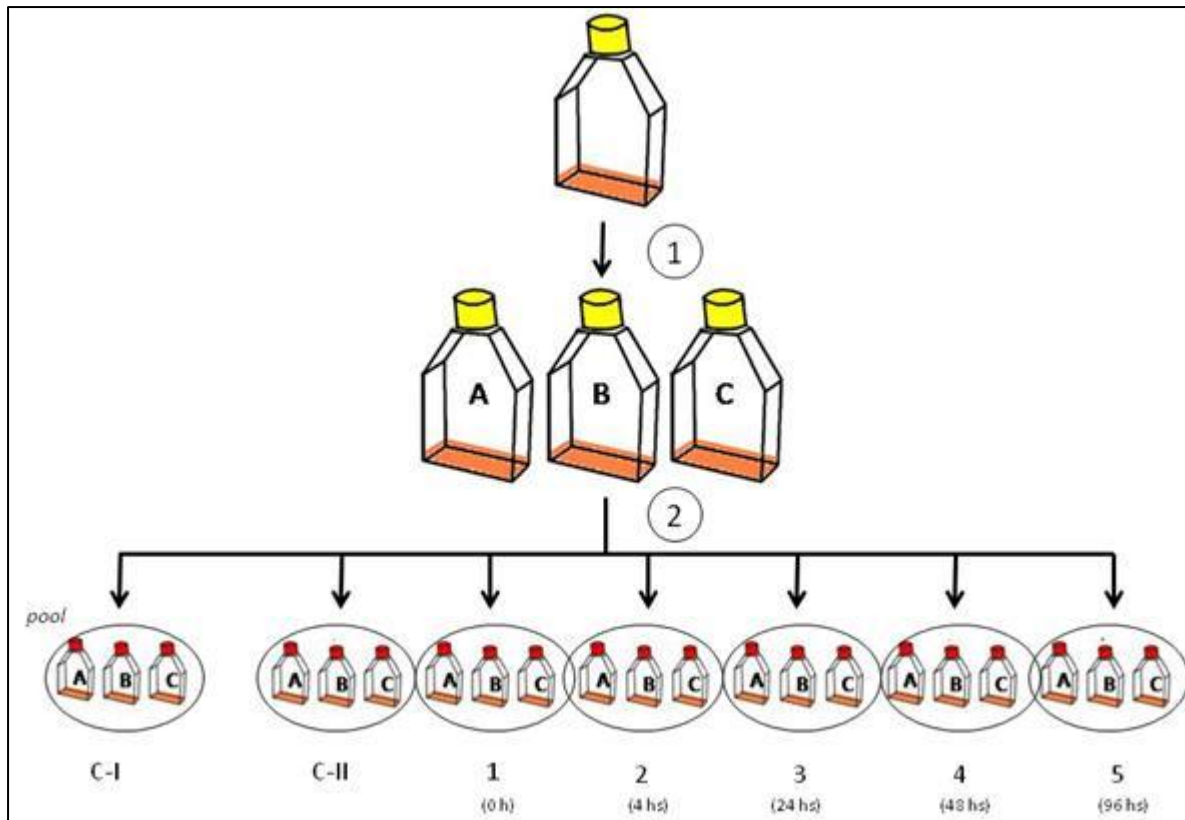


Figura 4: Desenho experimental. Para cada réplica biológica foi realizado o seguinte protocolo: 1) A partir de uma garrafa de cultivo de 150 cm<sup>3</sup> de capacidade (tampa amarela), contendo  $2 \times 10^7$  células em 270 mL de meio LIT, as epimastigostas foram divididas em três garrafas, cada uma com  $7,2 \times 10^6$  células em 90 mL (ou  $8 \times 10^4$  células/mL). Cada garrafa (A, B e C) representa uma réplica técnica do experimento, cujo cultivo foi posteriormente dividido em sete garrafas de cultivo de 50 cm<sup>3</sup> de capacidade (tampa vermelha) com  $5 \times 10^8$  células totais em cada. O conjunto de garrafas C-I foi usado como controle da curva de crescimento das células, não sendo irradiadas. O conjunto de garrafas C-II foi usado como controle do experimento de microarranjo. Elas não foram irradiadas, e o RNA das células de cada garrafa do conjunto C-II foi extraído juntamente com o conjunto 1. Os conjuntos de garrafas de número 1 a 5 foram irradiadas no mesmo momento. No entanto, os RNAs das células foram extraídos em momentos diferentes, sendo eles em ordem: logo após a irradiação (l.a.i.), 4, 24, 48 e 96 horas após a irradiação.

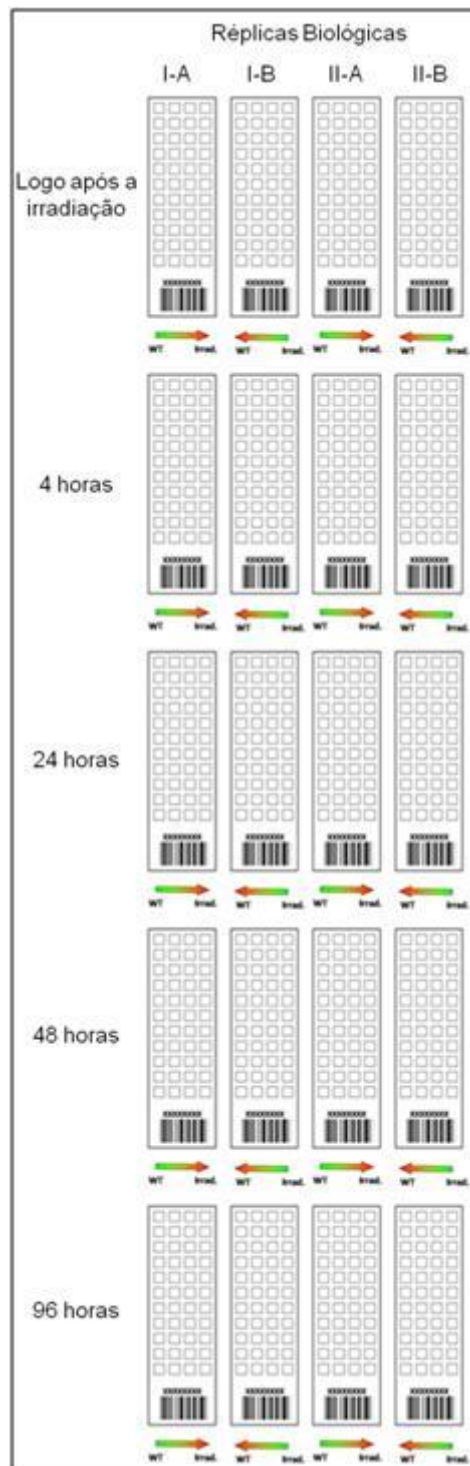


Figura 5: Esquema de hibridização das lâminas. Para cada tempo, quatro lâminas foram hibridizadas, sendo duas de cada réplica biológica (I e II) em forma de *dye-swap* (A e B). As setas com pontas vermelhas indicam amostras hibridizadas com a Cy5, enquanto os finais de setas verdes indicam amostras hibridizadas com Cy3.

### 3.5 – Marcação e hibridização das amostras

---

mRNAs amplificados (aRNA) com moléculas de aminoalil UTP incorporadas foram marcados utilizando os fluoróforos Cy3 e Cy5 por marcação indireta. Todas as etapas foram realizadas no ambiente mais escuro possível para evitar a degradação dos fluoróforos (moléculas fluorescentes). Para a marcação, foi usado o protocolo fornecido pelo kit AminoAllyl MessageAmp II (Ambion, Inc, USA). Brevemente, tubos individuais contendo cianinas liofilizadas foram hidratadas com 11  $\mu\text{L}$  de DMSO e incubadas por 1 hora no escuro a TA. Paralelamente, 9  $\mu\text{L}$  de *Coupling buffer* foram adicionados em 8  $\mu\text{g}$  de aRNA previamente seco no *Speed Vac*. As amostras reidratadas foram adicionadas nos tubos contendo as cianinas já hidratadas para o acoplamento, e as soluções foram incubadas por 30 min no escuro a TA. 4,5  $\mu\text{L}$  de hidroxilamina 4M foram adicionados às soluções e incubou-se por 15 min a TA para bloqueio da reação de acoplamento. Por fim, 5,5  $\mu\text{L}$  de  $\text{H}_2\text{O}$  livre de nuclease foram adicionados.

Para a purificação das amostras, ou seja, para a eliminação dos fluoróforos não incorporados em nenhuma molécula de aminoalil UTP, utilizaram-se colunas Microcon YM-30 *Centrifugal Filter Unit*, da empresa Millipore® (USA), segundo instruções do fabricante. De forma sucinta, as amostras marcadas com as cianinas foram misturadas, formando uma solução única, e 400  $\mu\text{L}$  de  $\text{H}_2\text{O}$  deionizada foram adicionados. A solução foi transferida para a coluna e submetida à centrifugação (13.000 xg, 6 min). O líquido residual foi descartado, e repetiu-se a lavagem por mais três vezes (adição de 400  $\mu\text{L}$  de  $\text{H}_2\text{O}$  deionizada e centrifugação), sendo que a última centrifugação durou 10 min. A coluna foi invertida em um tubo coletor novo, e aplicou-se um *spin* para recolher a solução purificada. O aRNA marcado foi completamente seco no *Speed Vac*, e acrescentou-se 30  $\mu\text{L}$  de tampão de hibridização (40% Formamida, 5X SSC, 0,1% SDS, 0,1 mM DTT). Essa solução foi aquecida a 95 °C por



3 minutos e sofreu um choque térmico (gelo) por 30 seg. A solução foi então pipetada na lâmina.

Brevemente, as lâminas foram pré-hibridizadas em solução de pré-hibridização (5X SSC, 0,1% SDS, 1% BSA) por 1 hora em banho-maria a 42 °C. As lâminas foram então lavadas em dois Beckers contendo água destilada através de 10 imersões em cada, seguido por 3 imersões em álcool isopropílico. As lâminas foram colocadas dentro de tubos tipo Falcon de 50mL forradas com folhas de papel absorventes, e centrifugadas por 5 minutos à 1500g. As amostras foram hibridizadas sob lamínulas e dentro de câmaras de hibridização Corning® (Corning, USA), mantidas em banho-maria a 42 °C por 14 horas no escuro. Após esse processo, as lâminas foram lavadas duas vezes por 5 minutos cada em solução de lavagem de baixa estrigência (2X SSC, 0,1% N-lauroylsarcosine), sendo a primeira lavagem realizada a 42 °C, e a segunda, à temperatura ambiente (TA). As lavagens subsequentes foram realizadas a TA. As lâminas foram lavadas 2 vezes por 5 minutos em solução de média estrigência, (0,1X SSC, 0,1% N-lauroylsarcosine), e 2 vezes por 5 minutos em solução de alta estrigência (0,1X SSC). Imediatamente após as lavagens, as lâminas foram secadas por centrifugação (5 minutos à 1500xg dentro de tubos falcon 50 mL) e lidas usando o *scanner* a laser com dois canais (ScanExpress Lite da PerkinElmer®, USA) numa resolução de 10 µm, potência do laser em 100% e PMT ajustado para se obter distribuições iguais de sinais verde e vermelho. Durante a leitura, um laser verde (543 nm) e um laser vermelho (633 nm) excitam os fluoróforos Cy3 e Cy5 respectivamente, e a luz emitida foi registrada para cada ponto na superfície da lâmina.

As amostras foram hibridizadas em lâminas de oligonucleotídeos gentilmente doados pelo *Pathogen Functional Genomics Resource Center* – PFGRC, divisão pertencente ao *The Craig Venter Institute*, antigo *The Institute for Genomic Research* – TIGR. A lâmina consiste

de 12.288 oligonucleotídeos de 70-mer, impressos duas vezes (réplicas), além de 500 controles negativos (oligonucleotídeos 70-mer de *Arabidopsis thaliana*).

### 3.6 – Análises das imagens

---

Dados brutos de uma lâmina de microarranjo de DNA consistem de duas imagens escaneadas de 16 bits. O primeiro passo na análise da imagem do *array* é sobrepor um *grid* retangular em cima dos pixels de forma a posicionar cada *spot* dentro de uma célula. Esse processo geralmente é semi ou totalmente automatizado. O software *ScanArray Express*, da PerkinElmer®, gera automaticamente o *grid* baseado no arquivo do tipo *.gal*. Arquivos *.gal* são arquivos de texto delimitados por *tab*, sendo que as linhas correspondem às sequências das sondas impressas e as colunas contém coordenadas e anotações. É nesse arquivo que o software utilizado encontra o *layout* do *array*, usado para criar o *grid*. Uma inspeção manual se faz necessária, pois diversas células devem ser reposicionadas manualmente, em função de micro-variações na posição da impressão das sondas que podem gerar dados falso-negativos (ausência de sinal no lugar de valores de intensidade) (figura 6). Além disso, células que se sobrepõem a artefatos devem ser marcadas como de má qualidade, para que o software não crie intensidades falso-positivas. Esta etapa é bastante laboriosa, porém fundamental para a garantia da qualidade e confiabilidade dos dados a serem analisados.

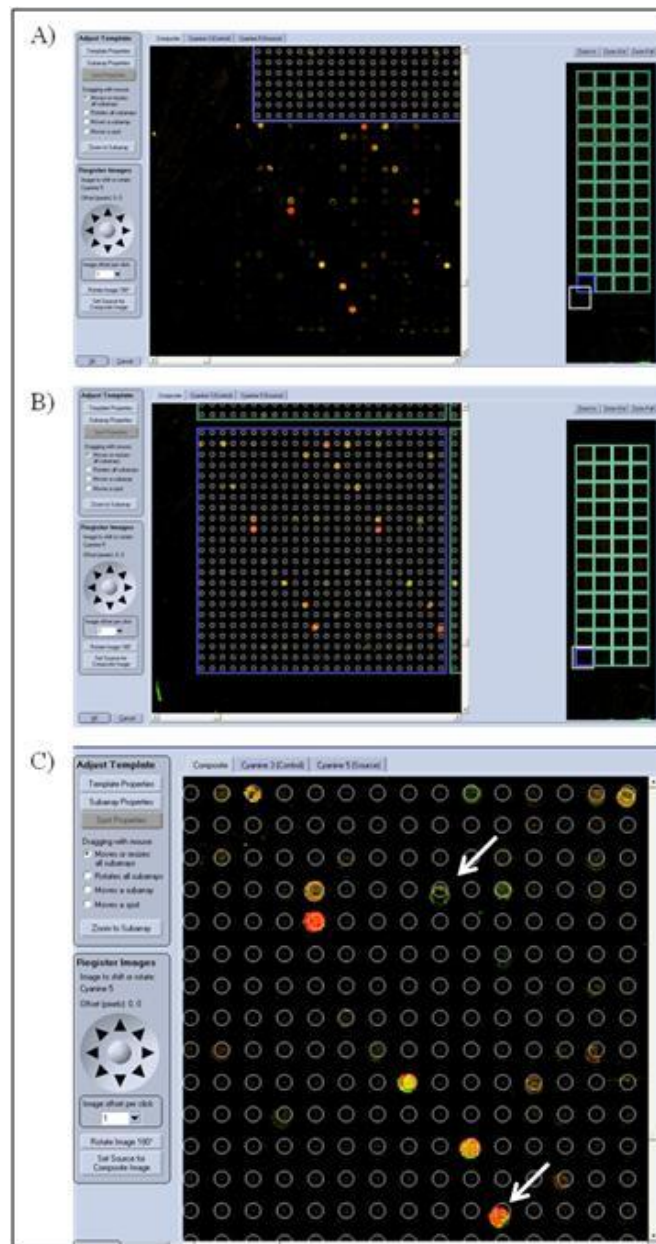


Figura 6: Imagens do programa *ScanArray Express*, usado para a leitura e análise das imagens das lâminas. A) Ao chamar o *grid* (montado de acordo com as informações contidas no arquivo .gal), ele aparece na tela fora do limite da lâmina (como pode ser visto pelos *spots* soltos na imagem). B) Acerto do *grid* de acordo com as posições dos *spots*. C) Detalhe do *grid* já alinhado com os *spots*, mostrando a necessidade da inspeção visual para acertar as células do *grid* nos *spots* que por ventura ficaram deslocados (alguns exemplos evidenciados pelas setas brancas).

### 3.7 – Correção do *background*, normalização dos dados e aplicação de modelo linear para geração da tabela de genes diferentemente expressos.

---

Para cada um dos cinco tempos, a análise de expressão gênica foi feita com informações obtidas de quatro lâminas, sendo um par (em forma de *dye-swap*) para cada uma das duas réplicas biológicas. Dados brutos de intensidade foram analisados usando a linguagem estatística *open source* R. Os dados foram inspecionados em busca de variações espaciais nos canais verde e vermelho, viés de agulhas de impressão e de intensidade-dependente, utilizando, para tais fins, os pacotes do *Bioconductor* Limma (*Linear Models for Microarray Data*) e *marray* (*Exploratory analysis for two-color spotted microarray data*).

Dois métodos distintos para correção de *background* foram testados com o intuito de padronizar um *pipeline* de análise para os dados deste trabalho: o método padrão *subtract*, que subtrai os valores de intensidade de *background* a partir dos valores de intensidade do *foreground* (sinal) em cada *spot* (Kooperberg *et al.*, 2002), e o modelo convolucional *NormExp* (offset=50), modelo matemático baseado no algoritmo RMA (*Robust Multichip Average*) para dados de Affymetrix adaptado para o uso em experimentos com dois canais (Ritchie *et al.*, 2007). Este algoritmo assume que as intensidades são resultantes da soma dos componentes de *background* e sinal, sendo que o *background* assume uma função densidade normal, e o sinal, uma função densidade exponencial. A função do pacote Limma usa a estimação por máxima verossimilhança para enquadrar o modelo *NormExp* nas intensidades dos sinais corrigidos.

A normalização dos valores de intensidade dos *spots* dentro da lâmina é uma etapa essencial. Ela busca ajustar os dados para os efeitos que tiveram como origem variações de razões técnicas, como os ajustes do laser, variações na quantidade de aRNA hibridados, diferenças entre os fluoróforos, dentre outros (Simon *et al.*, 2003). O processo permite que

valores de intensidades correspondentes a um determinado gene possa ser comparado com valores de intensidade do mesmo gene, porém, de outra lâmina (Draghici, 2003). Quatro métodos foram testados: *loess*, *print tip loess*, *median*, e *robust spline*. Os dois primeiros são baseados em regressão linear local (Yang *et al.*, 2002), sendo que o segundo aplica o método em cada agulha usada para imprimir as sondas separadamente. O método *median* subtrai a mediana ponderada do valor do  $\log_2$  do sinal vermelho –  $\log_2$  do sinal verde em cada lâmina. E o *robust spline* aplica a regressão curva (*spline*), que ajusta a melhor a função “linear-curva” nos dados (Smyth, 2004).

Para fins de padronização desta etapa, cada método de correção de *background* (*Subtract* e *NormExp*) foi testado com cada um dos quatro métodos de normalização dos dados dentro de cada lâmina para a escolha da melhor associação de algoritmos para os dados deste trabalho.

Um problema a ser considerado é o fato de que os fluoróforos apresentarem propriedades bioquímicas sutilmente diferentes que podem afetar os valores de intensidade dos genes. Para isso, métodos de normalização entre as lâminas fazem-se necessários. Uma característica que diferencia a normalização entre as lâminas da normalização dentro da lâmina é que o foco está nos valores de intensidade individual dos canais verde e vermelho, ao invés de focar apenas nas razões do *log* ou *log-ratios* (Smyth, 2004). O método utilizado foi o *aquantile*.

As análises foram feitas usando um modelo linear que incorpora replicação biológica (duas réplicas biológicas), replicação técnica (*dye-swap*) e replicação dentro das lâminas (dois oligonucleotídeos impressos para cada gene em cada lâmina). Essa abordagem requer duas matrizes a serem especificadas: a matriz do design experimental, que representa as diferentes amostras e a forma como elas foram hibridizadas e uma matriz de contraste, que indica como as comparações entre as lâminas devem ser realizadas.

A hipótese testada foi se um determinado gene de cada amostra irradiada foi significativamente mais ou menos expresso que o mesmo gene na amostra controle. Para isso, aplicou-se a função *lmFit* do pacote *Limma*, que adapta múltiplos modelos lineares para cada gene dado uma série de lâminas de microarranjo (Smyth, 2004). Os dados de entrada são os *log-ratios* dos valores de intensidade. O modelo é baseado no método dos mínimos quadrados, uma técnica de otimização matemática que busca encontrar o melhor ajustamento para um determinado conjunto de dados. Por último, aplica-se em paralelo uma estatística empírica bayesiana (B) e calcula-se para cada gene o valor p ajustado para múltiplos testes, utilizando o método “BH”, que trabalha com a proporção esperada de resultados falso-positivos - FDR, ou *False Discovery Rate* (Benjamini e Hochberg, 1995), com o objetivo de aumentar o poder estatístico e simultaneamente reduzir o risco de falsos positivos (Smyth, 2004).

O resultado final dessa seqüência de análises foram cinco listas, uma para cada tempo avaliado, contendo as seguintes informações: a) sobre cada sonda presente na lâmina: localização física, o identificador e seqüência de oligonucleotídeos lâmina; e b) sobre cada gene associado a uma sonda: o número de acesso e nome do produto codificado por cada gene, o valor do *fold-change* em log de base 2 (*logFC*), o valor médio de expressão, o valor-p, o valor p ajustado e o valor B. Neste estudo, utilizamos para as análises subseqüentes os valores de *logFC* e o valor p-ajustado. O *pipeline* de análise está esquematizado na figura 7.

O *script* final desenvolvido na linguagem estatística R, destacando os métodos *NormExp* (offset=50) para a correção de *background* e *Robust Spline* para normalização entre as lâminas, além do design experimental e esquema de contrastes desenvolvido para esta análise está evidenciado no Anexo 1.

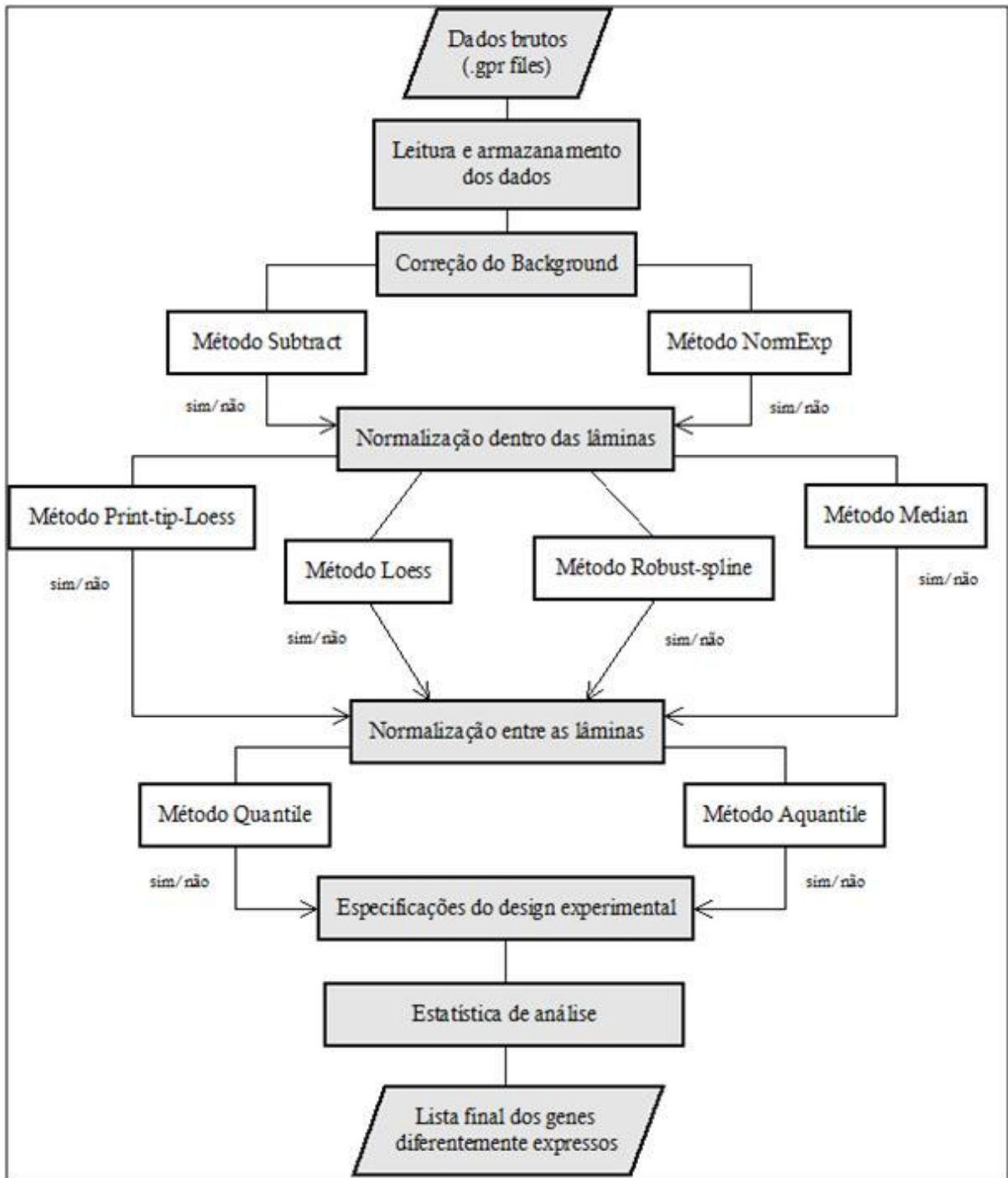


Figura 7: *Pipeline* de análise dos dados de microarranjo após a leitura das lâminas. Blocos cinza claro: passos essenciais da análise das lâminas. Blocos brancos: momentos de decisão entre os métodos a serem aplicados.

### 3.8 – Avaliação dos perfis de expressão gênica ao longo do tempo

---

Para visualizar a variação da expressão gênica nos cinco tempos avaliados, uma abordagem de agrupamento em função i) dos valores de *fold-change* e ii) ao longo do tempo foi aplicada (*bicluster*) foi aplicada. Para isso, um gráfico do tipo *heatmap*, que se baseia em distância euclidiana e em agrupamento hierárquico, foi gerado através da função *heatmap.2* do pacote *gplots* - *Various R programming tools for plotting data*. Os dados de entrada foram os valores de *fold-change* apresentados por cada gene nuclear que foi diferentemente expresso em pelo menos um dos tempos, ao longo de todos os tempos analisados. Os valores de *fold-change* foram plotados de forma linear e com tons de verde ou vermelho atribuídos para cada valor, sendo que tons de verde indicam genes subexpressos (com valores de *fold-change* negativos) e tons de vermelho indicam os superexpressos (com valores de *fold-change* positivos). O algoritmo agrupa horizontalmente genes com valores de *fold-change* similares ao longo do tempo, e verticalmente, as relações entre os tempos. Este resultado serviu de base para identificar grupos de genes com perfis semelhantes de expressão gênica e gráficos de tendência em relação ao tempo foram gerados com *script* em R.

### 3.9 – Análises das sequências obsoletas e identificação de sequências mitocondriais diferentemente expressas

---

As sequências obsoletas são oligonucleotídeos inicialmente desenhados para regiões supostamente codificadoras de proteínas. No entanto, após a montagem final do genoma, ficou evidente que estes oligonucleotídeos foram desenhados para regiões não codificadoras de proteínas, e suas anotações foram subsequentemente modificadas para obsoletas. Estas sequências foram alinhadas contra os genomas nuclear e mitocondrial (kDNA) utilizando os



algoritmos blastn para a busca contra o banco de dados de nucleotídeos do NCBI (NR) e contra as sequências específicas dos maxicírculos e minicírculos do kDNA.

As sequências obsoletas diferentemente expressas foram mapeadas no genoma nuclear anotado de *T. cruzi* utilizando o *software* Artemis (Carver *et al.*, 2008) com o objetivo de verificar a localização de cada uma delas dentro do contexto genômico.

### 3.10 – Atribuições de categorias de ontologia gênica (*Gene Ontology* – GO) e análise funcional dos dados

---

Para que os dados extraídos de experimentos de microarranjos possam fazer algum sentido, a análise funcional dos genes diferentemente expressos deve ser realizada. Assim, o pesquisador consegue visualizar quais os principais processos biológicos e quais funções moleculares foram estimulados ou reprimidos durante o tratamento. Para que isso seja possível, a atribuição de categorias de GO (*gene ontology*, ou ontologia gênica) se faz necessária. Para isso, usamos a ferramenta GOanna, disponível no site do banco de dados AgBase (<http://agbase.msstate.edu/cgi-bin/tools/GOanna.cgi>), que realiza a anotação de GO baseado em similaridade de sequências entre o arquivo de entrada (arquivo multi-fasta contendo as sequências de interesse) e as sequências disponíveis em diversos bancos de dados através do algoritmo blastp (Mccarthy *et al.*, 2006). Para este trabalho, selecionamos as bases de dados UniProt e SwissProt.

Neste trabalho, três arquivos diferentes multi-fasta de proteínas foram usados: o primeiro formado com os genes subexpressos, o segundo com os superexpressos, e o terceiro com todos os genes presentes na lâmina que possuem identificação oficial do tipo Tc00. As sequências de aminoácidos usadas para a criação dos três arquivos foram recuperadas através da ferramenta *Sequence retrieval* do banco de dados TriTrypDB, o *Kinetoplastid Genome*

*Resource* (<http://tritypdb.org/tritypdb/>). O GOanna retorna uma tabela Excel contendo a anotação dos termos GO para cada uma das sequências onde foi possível encontrar similaridade de sequências (parâmetros padrão da ferramenta) e um arquivo de entrada para a ferramenta *GOSlim Viewer*. As categorias *GO-slim* podem ser criadas para transformar a análise de dados de determinadas espécies mais apropriada. Assim, já há categorias GO-slim específicas para plantas, leveduras e bactérias. Como não há nada específico para protozoários, o conjunto de GO-slim utilizado nesse trabalho foi o genérico.

O resultado obtido para os conjuntos de genes sub e superexpressos foram curados manualmente com dois objetivos: 1) eliminar redundância de produtos gênicos (com identificadores idênticos) nos agrupamentos de termos GO-Slim, e 2) gerar manualmente categorias mais específicas, porém de importância biológica e funcional para este estudo. A redundância de produtos gênicos ocorre porque para um único gene, pode ter sido atribuído mais de um termo GO associado a ele dentro da mesma ontologia Slim, e acaba contando duas ou mais vezes ao ser agrupado para um nível hierárquico maior. O resultado final foi usado para gerar um gráfico de barras agrupadas para as ontologias de processos biológicos e função molecular onde foi possível comparar de forma direta os número de GOs associados a um termo específico entre os processos e funções induzidas ou reprimidas pela radiação gama.

### 3.11 – PCR quantitativa em tempo real (qRT-PCR)

---

Um subconjunto de genes diferentemente expressos foi selecionado para ser validado por qRT-PCR. Os iniciadores foram desenhados usando os softwares GenScript Real-time PCR (TaqMan) Primer Design (<https://www.genscript.com/ssl-bin/app/primer>) e Primer3 (<http://frodo.wi.mit.edu/primer3/>). O algoritmo blastn foi aplicado para alinhar os pares de

iniciadores contra o banco de dados de sequências de nucleotídeos do NCBI para excluir aqueles que, por ventura, se alinharam em mais de uma região genômica do *T. cruzi*, ou que tivesse alto grau de similaridade com outros organismos. No total, 18 pares de iniciadores foram sintetizados e padronizados inicialmente com DNA genômico para confirmar a capacidade de amplificação e o tamanho do produto gerado (em torno de 100 pb).

Os RNAs totais extraídos e purificados foram tratados com DNase e usados para a síntese subsequente de cDNA usando o kit Superscript III cDNA Synthesis (®Invitrogen), seguindo as instruções do fabricante. Para fins de padronização, testou-se nove diferentes combinações de concentrações de iniciadores (300/300, 300/600, 300/900, 600/300, 600/600, 600/900, 900/300, 900/600, 900/900 nM) e quatro quantidades iniciais de cDNA (10 ng, 1 ng, 0,1 ng e 0,01 ng) para a geração das curvas de dissociação e padrão, valores de *slope*, taxa de eficiência (que deve apresentar valores próximos de 2) e valores de co-eficiência  $R^2$  para cada iniciador. Assim, o mix para a reação de qRT-PCR foi composto por 10 ng de cDNA, 600 nM de cada iniciador, 2,5  $\mu$ L de Sybr Green PCR Master Mix (Applied Biosystems, Foster City, CA), e água deionizada, totalizando 5  $\mu$ L de volume final de reação. As amostras (controle e irradiadas) foram aplicadas em placas de 384 wells (Applied Biosystems, Foster City, CA) em triplicata, além do controle negativo (branco) e controle positivo (amplificação do gene gliceraldeído 3-fosfato desidrogenase - GAPDH, usado para a posterior normalização. As placas foram lidas no equipamento AB7900 (Applied Biosystems). A normalização foi realizada aplicando o método  $-2^{\Delta\Delta Ct}$ .

O teste de Friedman (Conover, 1980), que é um teste não-paramétrico (para dados que não apresentam distribuição normal) que compara três ou mais grupos pareados, e o teste de Dunn (Conover, 1980), que realiza múltiplas comparações entre pares de colunas (que neste caso, são os tempos analisados) foram aplicados para questionar se o *fold-change* em cada um dos tempos obtido pela qRT-PCR era significativamente diferente do valor de expressão

da amostra controle. O coeficiente de classificação de Spearman e regressão linear foram aplicados para verificar a correlação entre os resultados do microarranjo de DNA e qRT-PCR. Toda a parte estatística foi realizada utilizando o software GraphPad Prism V5.04.

### 3.12 – Análise dos genes RHS

---

Para testar se a irradiação induziu de forma direcionada a superexpressão de genes RHS, o teste de qui-quadrado duas-caudas foi aplicado em cada um dos tempos analisados para testar se a frequência dos genes RHS superexpressos era similar à frequência dos genes superexpressos restantes (hipótese nula:  $f1 = f2$ ), ou não (hipótese alternativa:  $f1 \neq f2$ ), de acordo com as equações 1 e 2 abaixo:

$$(1) \quad f1 = \frac{\# RHS \text{ superexpressos}}{\# total RHS}$$

$$(2) \quad f2 = \frac{\# genes \text{ superexpressos} - \# RHS \text{ superexpressos}}{\# total genes - \# total RHS},$$

sendo que i) RHS superexpressos = número de genes RHS significativamente superexpressos; ii) genes superexpressos = número total de genes significativamente superexpressos, iii) total RHS = número de sondas impressas na lâmina para o genes RHS (379) e iv) total genes = número total de sondas para genes de *T. cruzi* impressas na lâmina (12.288).

## 4 - RESULTADOS

---

---

### 4.1 – Curva de crescimento após a irradiação

---

O *T. cruzi* é um organismo altamente resistente à radiação ionizante, sugerindo que este parasito tenha um eficiente sistema de reparo de quebra de fita dupla. Foi visto que culturas de células CL Brener do tipo selvagem irradiadas com 500 Gy de radiação gama param de crescer, levando em média 240 horas para retornar às mesmas características (densidade celular e curva de crescimento logarítmica) quando comparadas com a cultura de células não irradiadas (Regis-Da-Silva *et al.*, 2006). Resultados semelhantes foram encontrados durante a padronização de cultivo e irradiação para este trabalho. Apesar de não termos reproduzido o experimento completo (até o crescimento exponencial das culturas que é de 240 horas), verificamos que células irradiadas diminuem sua taxa de crescimento celular de forma considerável, permanecendo aparentemente sem crescimento até 96 horas após a irradiação gama, tempo máximo observado nesse estudo (figura 8).

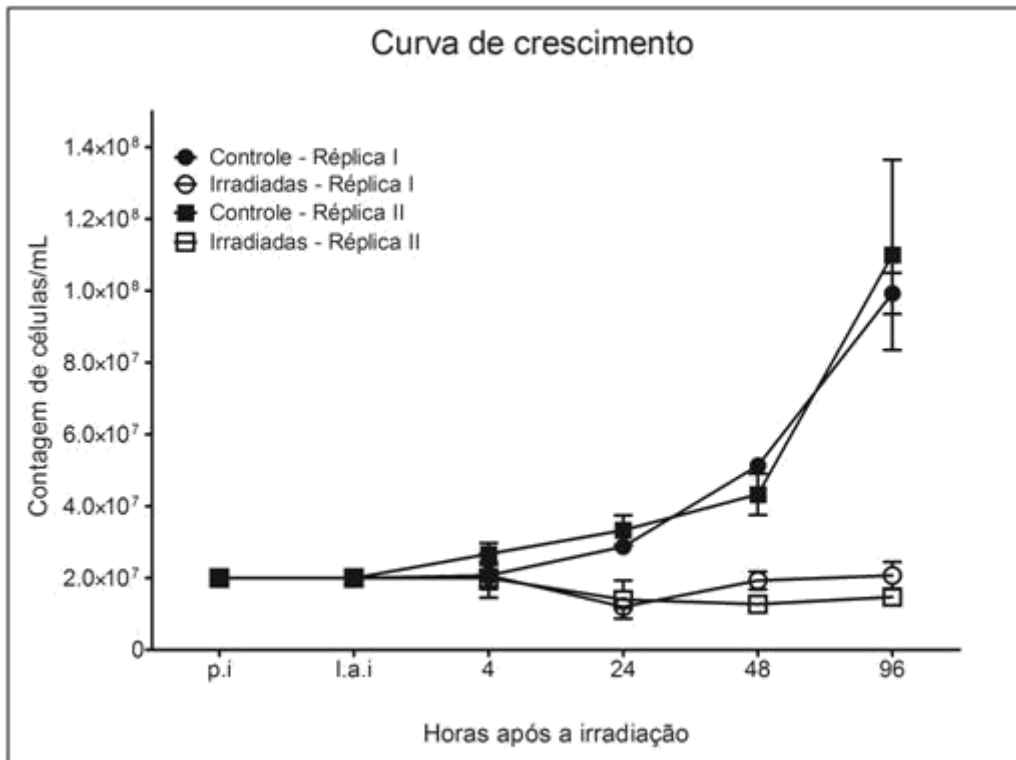


Figura 8: Curvas de crescimento de células epimastigotas de *T. cruzi* (cepa CL-Brener).

●/■ = células controles. ○/□ = células irradiadas. ○/● = Réplica biológica I. □/■ =

Réplica biológica II. p.i = pré-irradiação. l.a.i = logo após a irradiação.

#### 4.2 – Avaliação da integridade do RNA de *T. cruzi* após a irradiação

---

O próximo passo foi verificar se após o estresse causado pela alta dose de radiação, que certamente fragmenta as moléculas de DNA (Regis-Da-Silva *et al.*, 2006), seria possível extrair RNA total com qualidade suficiente para marcarmos e hibridizarmos as lâminas. Não sabíamos se este tratamento fragmentaria as moléculas de RNA, obtendo material somente quando surgissem os primeiros novos transcritos. Os resultados sugerem que os raios gama fragmentaram de forma considerável as moléculas de RNA ribossomal. Na figura 9 é possível perceber o padrão de três bandas referentes aos RNAs ribossômicos (usual em RNA de *T. cruzi*). Visualmente, a qualidade do RNA é semelhante em todas as amostras, tanto no controle, quanto em todos os tempos analisados.

O processo de extração de RNA total foi bem sucedido, rendendo quantidades significativas de material, e desprovido de contaminação com produtos protéicos (260/280 > 1,8) (tabela 1). Foi então realizado a etapa de *clean-up* dos *pools* RNA total, conforme explicado na metodologia.

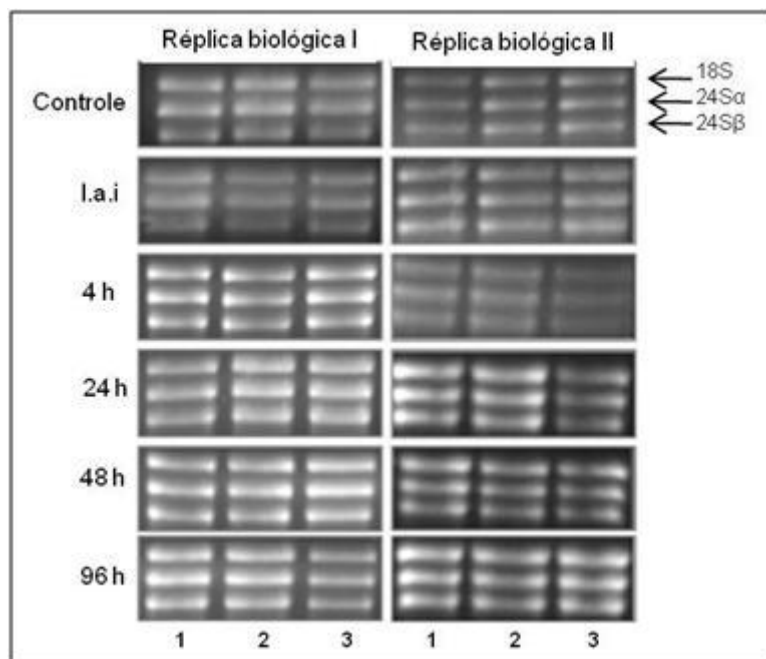


Figura 9: Avaliação da integridade do RNA de *T. cruzi* após a irradiação. RNAs das amostras das réplicas biológicas I e II (em triplicatas, conforme indicado pelos números 1, 2 e 3) foram submetidas a uma eletroforese em géis de agarose 1,2% desnaturantes e posteriormente corados com brometo de etídeo. 1.a.i = logo após a irradiação. Em destaque, estão mostradas as três bandas de rRNA de *T. cruzi*.



Tabela 1: Quantificação do RNA total antes e após o *clean-up* de RNA das amostras.

	Extração do RNA total			RNA após <i>Clean-up</i>		
	Amostra	$\mu\text{g}/\mu\text{L}$	A260/A280	Amostra	$\mu\text{g}/\mu\text{L}$	A260/A280
<b>Réplica Biológica I</b>	Controle - I	1,4	2,13	Controle	2,4	2,20
	Controle - II	2,5	2,09			
	Controle - III	2,1	2,16			
	l.a.i. - I	1,4	2,12	l.a.i	2,4	2,18
	l.a.i. - II	2,5	2,10			
	l.a.i. - III	2,0	2,10			
	4 horas - I	1,6	2,07	4 horas	2,8	2,17
	4 horas - II	2,9	2,06			
	4 horas - III	2,8	2,07			
	24 horas - I	1,9	2,09	24 horas	2,4	2,18
	24 horas - II	2,8	2,06			
	24 horas - III	2,8	2,07			
	48 horas - I	1,5	2,11	48 horas	1,8	2,20
	48 horas - II	3,0	2,03			
	48 horas - III	2,4	2,09			
96 horas - I	2,1	2,09	96 horas	2,3	2,20	
96 horas - II	2,7	2,08				
96 horas - III	2,3	2,09				
<b>Réplica Biológica II</b>	Controle - I	1,7	2,02	Controle	1,9	2,19
	Controle - II	1,8	2,05			
	Controle - III	2,4	2,09			
	l.a.i. - I	2,0	2,09	l.a.i	2,3	2,17
	l.a.i. - II	1,9	2,10			
	l.a.i. - III	1,9	2,05			
	4 horas - I	2,3	2,12	4 horas	2,0	2,18
	4 horas - II	2,1	2,09			
	4 horas - III	1,7	2,09			
	24 horas - I	2,6	2,10	24 horas	2,3	2,19
	24 horas - II	1,8	2,06			
	24 horas - III	2,4	2,12			
	48 horas - I	2,1	2,10	48 horas	2,4	2,19
	48 horas - II	2,8	2,09			
	48 horas - III	2,6	2,09			
96 horas - I	2,2	2,05	96 horas	1,4	2,24	
96 horas - II	2,0	2,10				
96 horas - III	2,2	2,09				

l.a.i. = logo após a irradiação. Valores (após a quantificação) em  $\mu\text{g}/\mu\text{L}$

### 4.3 – Amplificação, marcação e hibridização do RNA

---

A quantidade substancial de RNA necessária para análise de expressão é um fator limitante para a tecnologia de microarranjos. Tentou-se, por muitos meses, marcar as amostras sem o processo de amplificação de mRNA, o que culminou em tentativas frustradas, com perda total do mRNA no processo de purificação das amostras marcadas. Quando não havia perda de material, o acoplamento das cianinas na molécula de aminoalil era ineficiente, sendo impossível detectar sinal após leitura da lâmina no *scanner*.

Dessa forma, optou-se por adicionar uma etapa de amplificação de mRNA, conhecido como método de Eberwine (Phillips e Eberwine, 1996), cujos passos gerais envolvem uma reação de transcriptase reversa. O método é robusto, reprodutível, e o rendimento do RNA antisense amplificado é alto e de boa qualidade, como pode ser visto na tabela 2, que mostra a concentração do mRNA amplificado (aRNA) e a relação 260/280. A partir da quantidade inicial de 2 µg de RNA total, obteve-se, em média, 1 µg/µL de mRNA amplificado (aRNA), em um volume total de 100 µL. Considerando que em 2 µg de RNA total espera-se encontrar no máximo 0,1 µg de mRNA (estima-se que 2 a 5% do RNA total é composto por mRNAs), o método amplifica em até 1000 vezes essa quantidade. De posse de quantidades suficientes para a marcação das amostras com os fluoróforos, as hibridações foram realizadas.

Tabela 2: Quantificação do mRNA amplificado (aRNA) e análise do seu grau de contaminação com proteínas.

Amostra		$\mu\text{g}/\mu\text{L}$	A260/A280		$\mu\text{g}/\mu\text{L}$	A260/A280
Controle		1,0	2,01		1,1	2,00
l.a.i.		1,3	1,99		1,1	2,00
4 horas	<b>Réplica</b>	1,2	2,01	<b>Réplica</b>	1,0	2,02
24 horas	<b>Biológica I</b>	1,1	2,01	<b>Biológica II</b>	0,8	2,03
48 horas		1,1	2,02		0,9	2,03
96 horas		0,9	1,98		1,0	2,00

l.a.i = logo após a irradiação.

#### 4.4 – Análises de microarranjos: correção do *background* e normalização dos dados

---

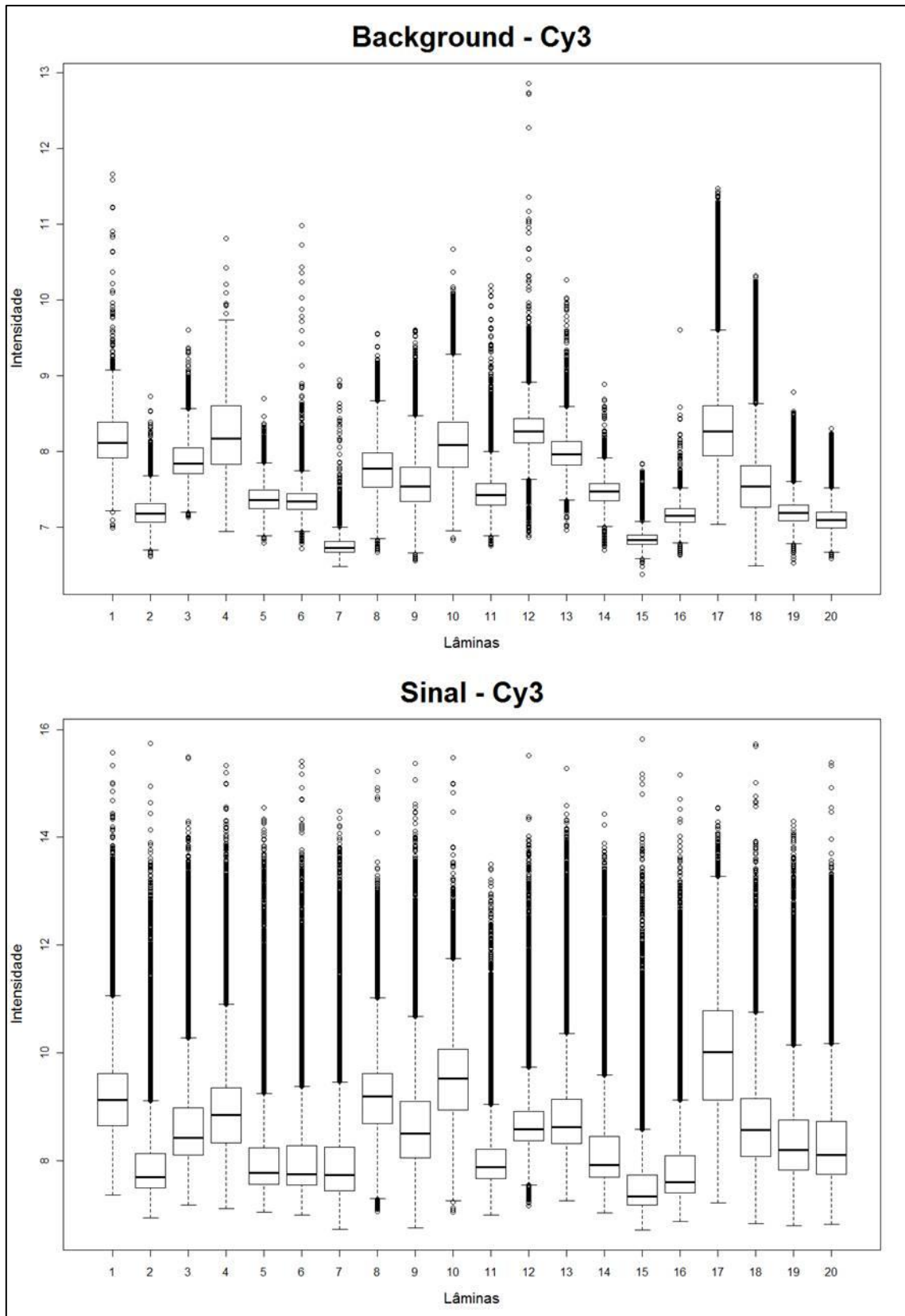
A padronização da análise foi realizada utilizando os pacotes Limma e marray do Bioconductor - BioC (projeto de código aberto e de desenvolvimento de softwares para análise de dados genômicos), que são baseados primariamente na linguagem de programação R. Oito diferentes combinações entre métodos de correção de *background* e normalização dos dados foram testados, e os resultados completos encontram-se no Anexo 2.

##### 4.4.1 – Correção do *background*: teste de dois diferentes algoritmos

As lâminas usadas neste trabalho apresentaram valores altos de *background* por dois motivos: o primeiro foi em função da alta aderência dos fluoróforos presentes na solução de hibridização na superfície sólida da lâmina. O segundo motivo foi em função do baixo número de *spots* acesos em cada sub-*array* da lâmina. Consequentemente a lâmina apresentou grandes áreas sem sinal, e que acabou sendo englobado no valor total de *background*. Dessa forma, uma metodologia robusta de correção de *background* foi

necessária para gerar dados concisos para os passos de normalização dentro de cada lâmina e entre as lâminas.

O método *subtract* não corrigiu de forma satisfatória a intensidade de *background* por usar um modelo simplista de correção. Em um conjunto de dados com valores altos de *background*, esse método gera valores de sinal negativos (representado pelos pequenos círculos na posição de ordenada “zero” nos gráficos das figuras 10 e 11. O método de correção de *background* que melhor corrigiu os dados foi o *NormExp*, onde houve o tratamento dos *spots* com valores baixos de sinal, deixando como *outliers* somente os *spots* com sinal muito forte. Os *Boxplots* abaixo (figuras 10 e 11) ilustram bem como esse modelo tratou de maneira mais satisfatória os valores de *background*, corrigindo melhor os valores de baixa intensidade.



(continua na próxima página)

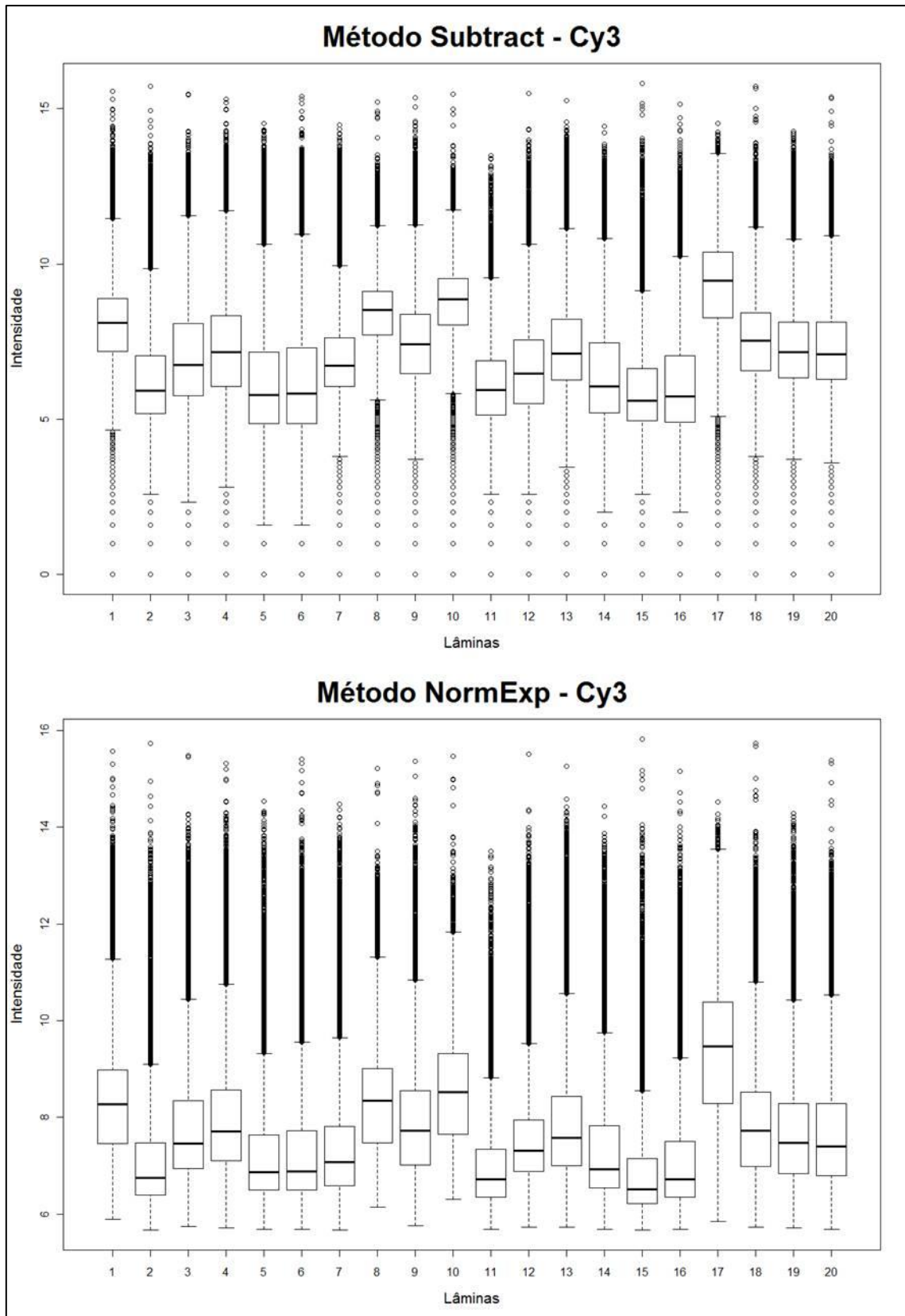
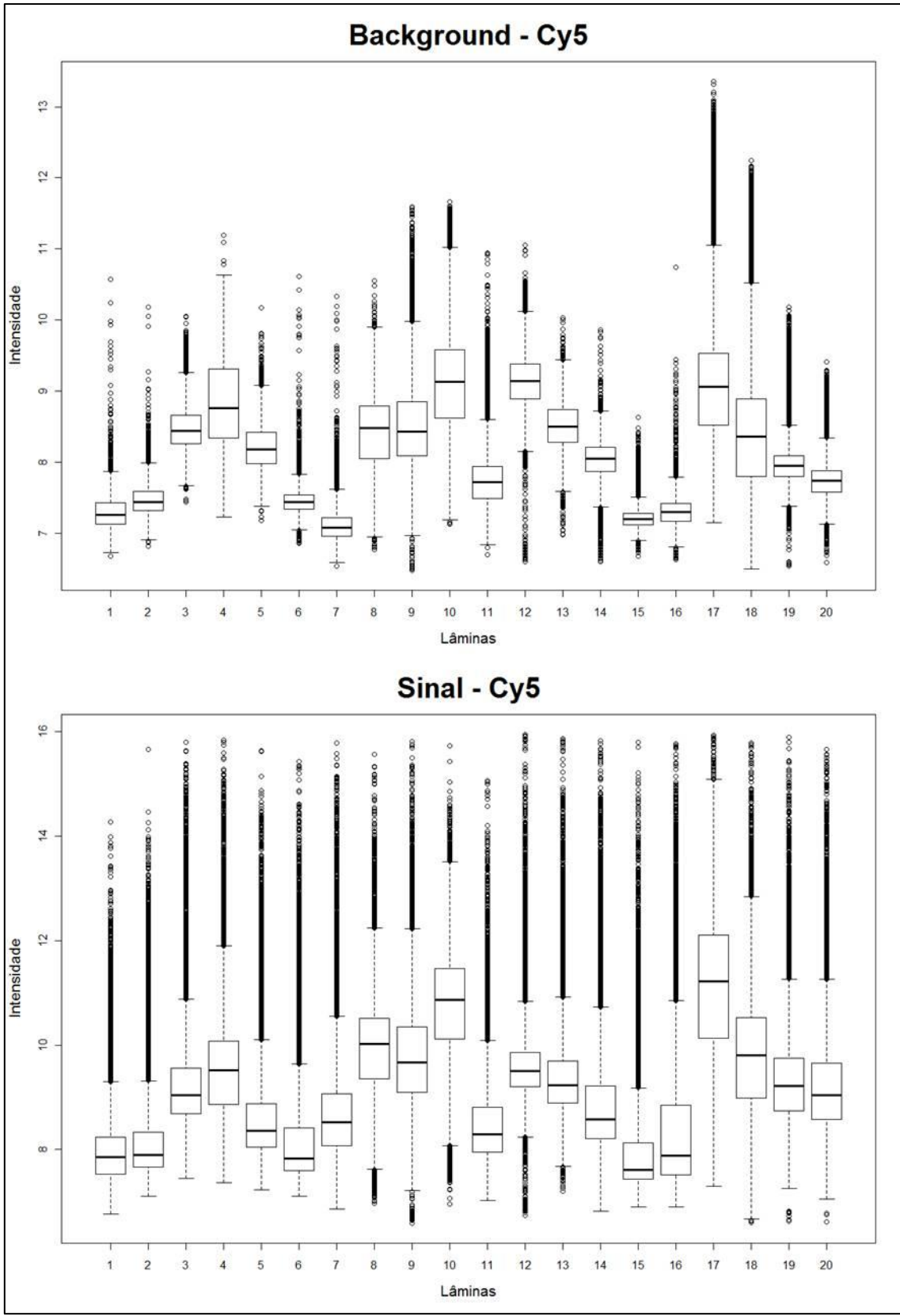


Figura 10: Boxplots de cada lâmina ilustrando os valores de *background* e sinal brutos, e os valores do sinal corrigidos pelo método *Subtract* ou *NormExp* para o canal verde (Cy3)



(continua na próxima página)

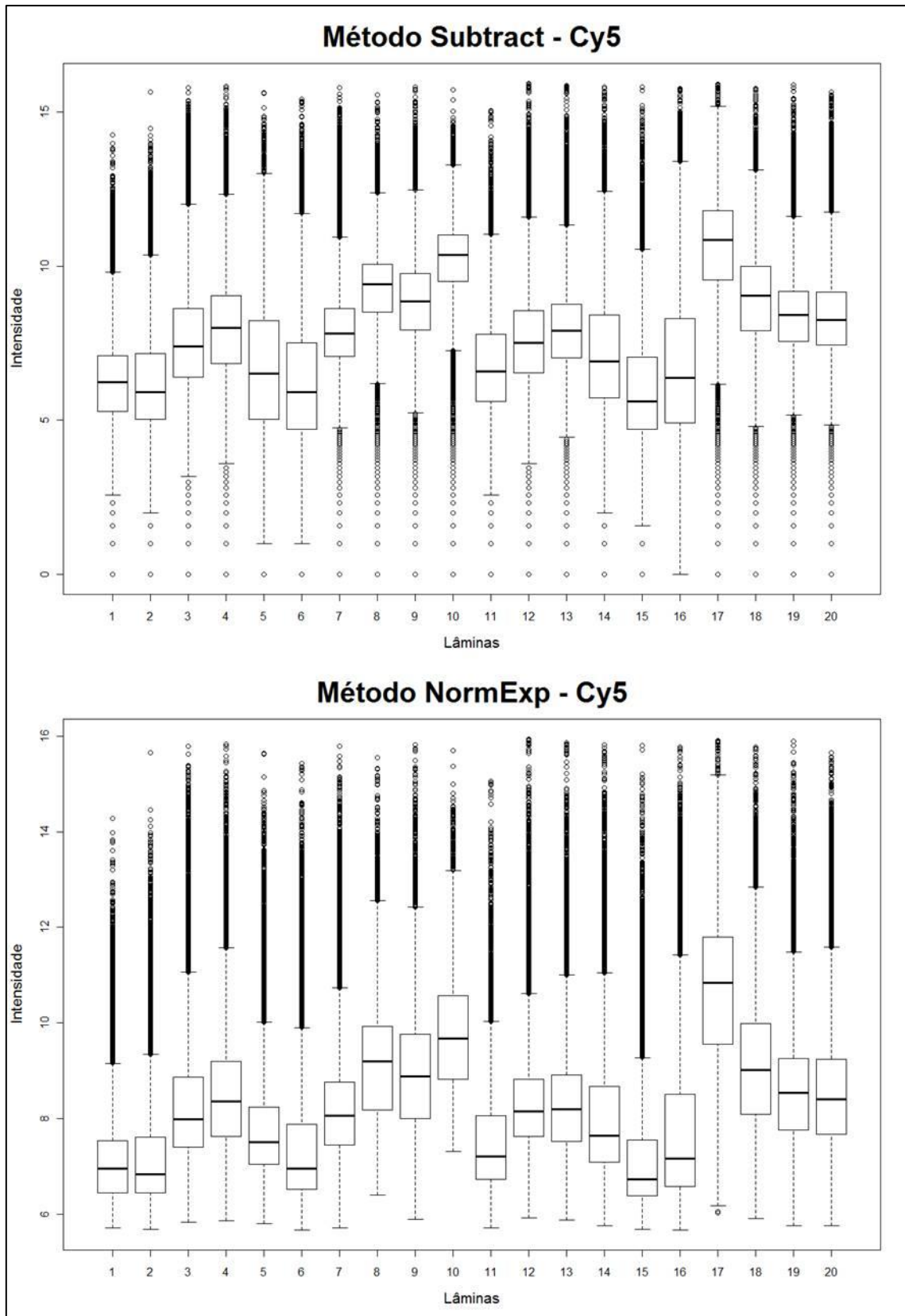


Figura 11: Boxplots de cada lâmina ilustrando os valores de *background* e sinal brutos, e os valores do sinal corrigidos pelo método *Subtract* ou *NormExp* para o canal vermelho (Cy5)



#### 4.4.2 – Normalização dentro e entre as lâminas

Diversos desafios são encontrados ao analisar os dados obtidos pela técnica de microarranjos, pois variações de diversas naturezas são introduzidas ao longo de cada etapa do experimento. Para uma quantificação acurada e precisa das variações da expressão gênica, é importante considerar as variações aleatórias (experimentais) (Schuchhardt *et al.*, 2000) e as sistemáticas, como por exemplo, a variação de taxas de incorporação entre as duas cianinas usadas, Cy3 e Cy5. Tal variação pode ser originada de uma variedade de fatores, incluindo propriedades físicas dos fluoróforos (sensibilidade à temperatura e à luz e meia-vida relativa), eficiência da incorporação da *dye*, variabilidade experimental durante a hibridização ou de configurações diferentes do *scanner* no momento da leitura dos sinais (Yang *et al.*, 2002). A idéia de normalizar os dados é para minimizar as variações sistemáticas nos níveis de expressão gênica medidos de duas amostras co-hibridizadas, permitindo, assim, que as variações biológicas dentro de uma lâmina se destaquem, e que seja possível comparar os níveis de expressão entre as lâminas (Yang *et al.*, 2002).

O método que melhor normalizou os dados dentro de cada lâmina foi o *robust spline* (Smyth e Speed, 2003; Smyth, 2004), como pode-se perceber tomando como exemplo os gráficos denominados MA-plots para a lâmina 1 (figura 12), onde  $M$  = razão logarítmica da intensidade entre os canais verde e vermelho ( $\log_2(R/G)$ ), e  $A$  = intensidade média dos canais verde e vermelho ( $(\log_2(R \times G)/2)$ ). O método *Loess*, que é o mais difundido entre os métodos testados, também normalizou os dados, mas percebe-se que as linhas coloridas não permaneceram próximas de zero nos dados de valores mais baixos de intensidade. O método *Print tip loess* não foi eficiente para normalizar os dados, assim como o método *Median* (falhou em corrigir os dados com valores baixos de intensidade). Percebe-se que o método *Robust spline* foi que melhor centralizou os dados, ou seja, o valor médio da razão entre a intensidade de cada canal tornou-se próximo de zero para todas as agulhas (*print tips*). O

motivo pelo qual nos leva a esperar esse comportamento dos dados é que um dos pressupostos estatísticos aplicado na análise de dados de microarranjos é que a maioria dos genes (pelo menos 95%) não é diferentemente expressa em duas situações diferentes. Assim, o  $\log_2$  vermelho/verde ( $\log_2 R/G$ ) da maior parte dos genes deve ser igual ou próximo de zero:

$$\log_2 \frac{R}{G} = 0 \rightarrow 2^0 = \frac{R}{G} \rightarrow 2^0 = 1. \text{ Assim, } \frac{R}{G} = 1$$

Os gráficos para as outras 19 lâminas (antes e após a normalização com o método *Robust spline*) encontram-se no material anexo. A normalização dessas 19 lâminas pelos outros métodos também foi realizada, mas não foram mostradas aqui. Como o método usado para normalizar os dados utiliza os valores de intensidade de cada *sub-array*, é interessante verificar a distribuição dos valores de M ao longo das 12 agulhas usadas para fabricar cada lâmina. Pode-se perceber que os valores de M para cada agulha ficam próximos de zero após a normalização (figura 13).

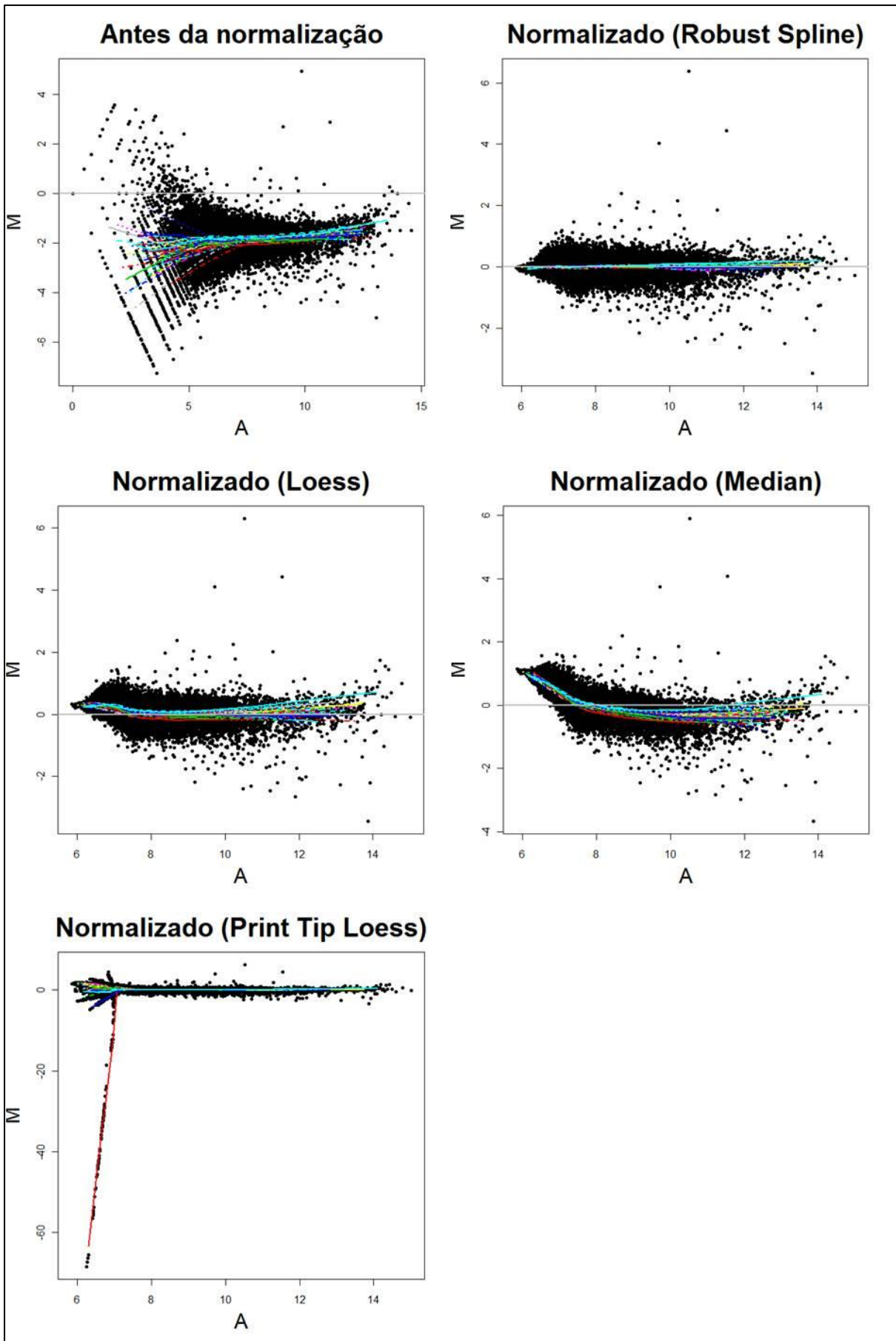


Figura 12: MA *plots* para a lâmina 1 com diferentes métodos de normalização. O M representa o  $\log_2$  da razão entre os sinais do canal vermelho (R) e verde (G) ( $\log_2(R/G)$ ) e o A representa a média da soma dos  $\log_2$  dos canais vermelho e verde ( $\log_2(R+G)/2$ ). O primeiro gráfico representa os dados antes da normalização. Os gráficos seguintes representam os dados normalizados. O segundo foi normalizado com o método *robust spline*. O terceiro, com o método *loess*, o quarto com o método *print tip loess*, e o quinto, com o *median*.

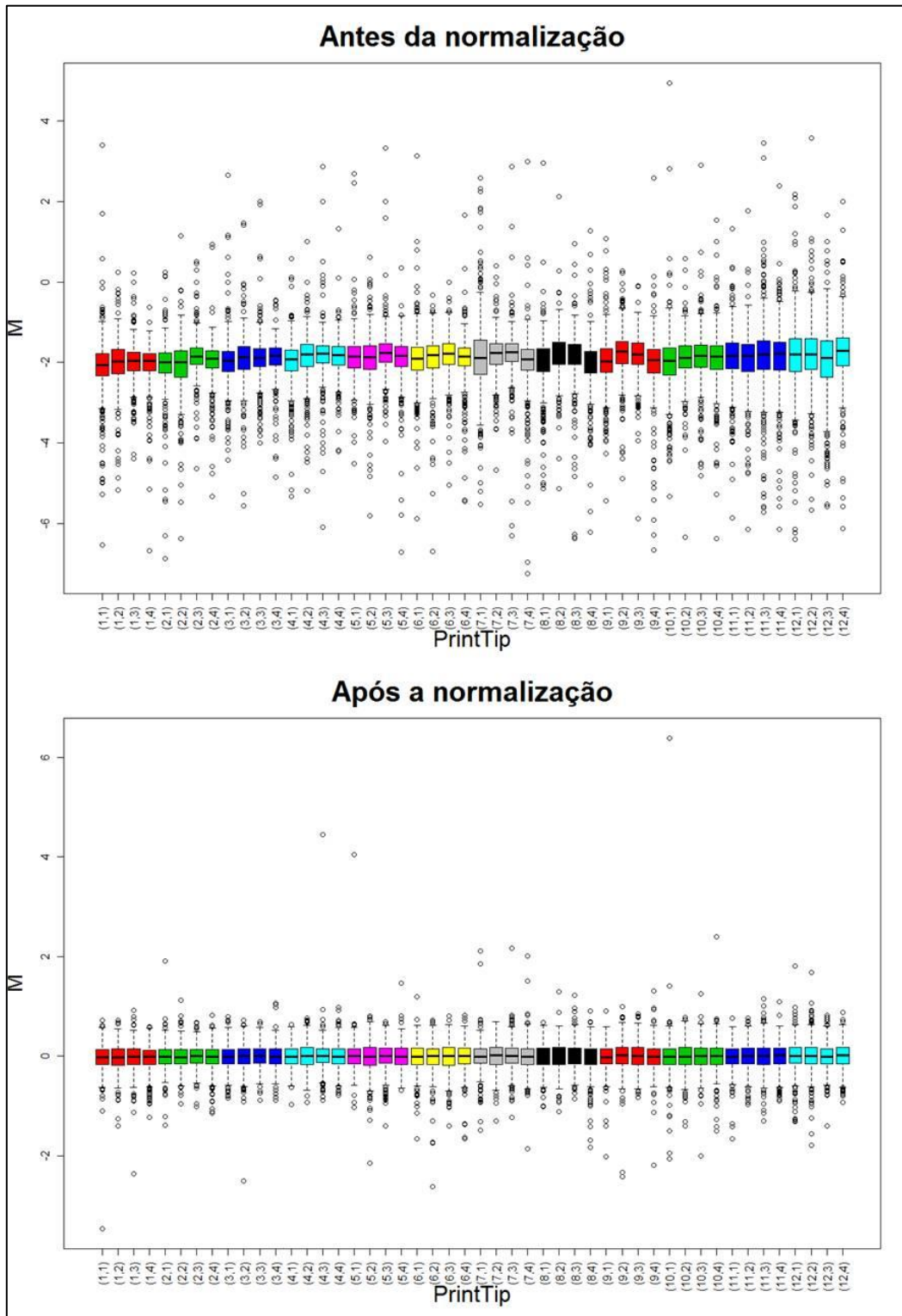


Figura 13: *Boxplots* representando os valores de  $M$  ( $\log_2(R/G)$ ) para cada agulha de impressão. Valores pré (primeiro gráfico) e pós-normalização pelo método *Robust spline* (segundo gráfico) para a lâmina 1.

Conforme esperado, a normalização dos dados entre as lâminas se mostrou necessária em função da variação das intensidades dos dois canais (figura 14-1), mesmo após a correção do *background* (figura 14-2). A distribuição empírica suavizada (*smoothed empirical densities*) para os canais verde e vermelho são consideravelmente diferentes antes da normalização entre as lâminas, pois é possível perceber que os sinais de cada fluoróforo apresentam variações entre as lâminas (figura 14-3). Isso inviabiliza a análise dos dados, pois os valores de intensidade de cada gene de cada lâmina não são comparáveis entre si. Após a aplicação do método de normalização entre as lâminas *aquantile*, as intensidades dos sinais de cada lâmina tornam-se praticamente iguais, o que demonstra que o método foi eficaz na correção de variações devido à diferença entre os marcadores (figura 14-4). No material anexo encontram-se as distribuições empíricas suavizadas para os dados corrigidos com os dois métodos de correção de *background* associados com os quatro métodos de normalização de dados dentro de cada lâmina.

Por fim, a figura 15 que representa os *boxplots* dos valores de M para todas as lâminas mostra o resultado final do processamento dos dados. Essa figura mostra dois aspectos: as medianas do valor M ficaram centralizadas no zero, indicando que os dados foram adequadamente normalizados, e o intervalo de valores da ordenada (M), que diminuiu de -9 a 12 para -5 a 8 (valores aproximados).

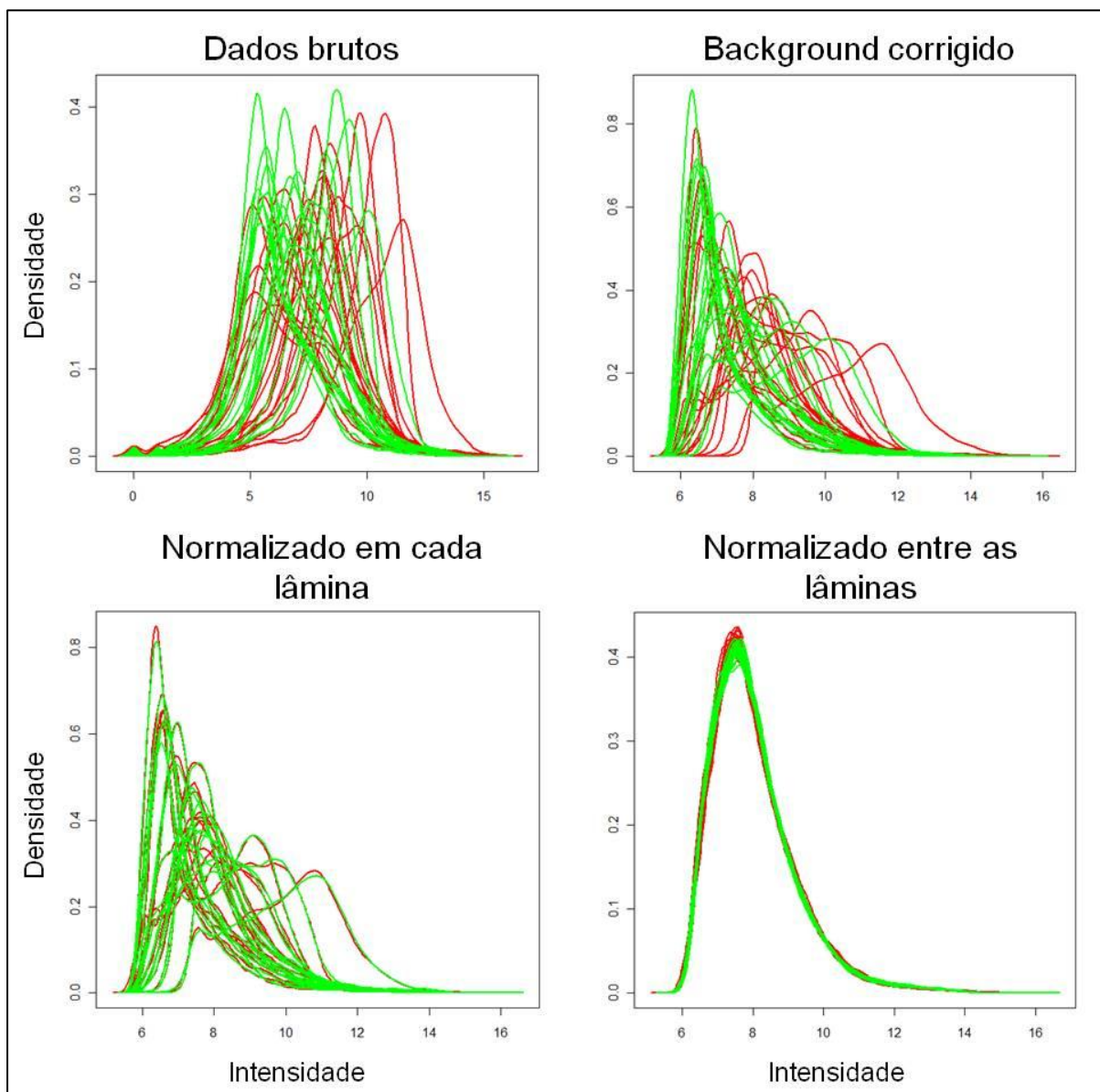


Figura 14: Distribuição empírica suavizada dos sinais dos canais verde e vermelho para todas as lâminas. 1- Dados brutos. 2- Dados com o *background* corrigido pelo método *NormExp*. 3- Dados normalizados dentro de cada lâmina pelo método *Robust spline*. 4- Dados normalizados dentro e entre as lâminas pelo método *Aquantile*.

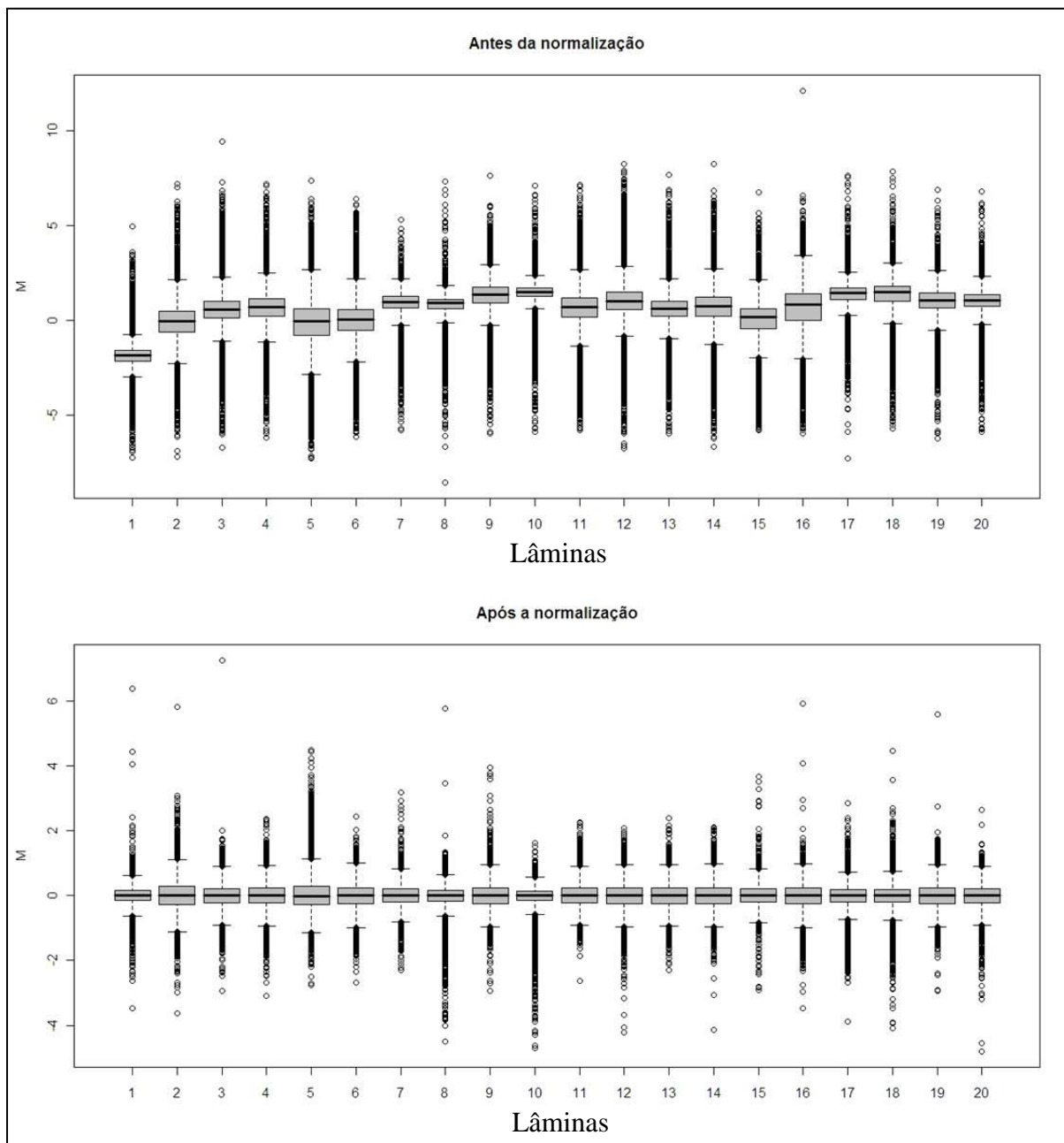


Figura 15: *Boxplots* representado os valores de  $M (\log_2(R/G))$  antes (primeira figura) e após a normalização (segunda figura). O eixo X representa cada uma das 20 lâminas usadas nesse trabalho. O eixo Y representa os valores de  $M$  na escala de  $\log_2$ . O valor de  $M = 0$  representa uma razão de  $R/G=1$ .



## 4.5 – Avaliação das alterações na expressão gênica induzida por radiação

---

### 4.5.1 – Genes diferentemente expressos e sua distribuição ao longo do tempo

Foram considerados genes diferentemente expressos aqueles cujo valor de p-ajustado foi abaixo de 0,05. No total 273 genes foram considerados diferentemente expressos em, pelo menos, um dos tempos analisados, sendo 160 superexpressos e 113 subexpressos. A tabela 3 mostra o número de genes sub ou superexpressos em cada tempo. A soma final não totaliza os números acima citados, pois há genes que são diferentemente expressos em mais de um tempo analisado e foram contados mais de uma vez. A figura 16-A e tabela 3 mostram de forma clara que a resposta à radiação foi tempo dependente. Ao final das quatro primeiras horas, o número de genes subexpressos significativos é maior que o número de genes superexpressos. Porém, após 24 horas, o cenário se inverte e se mantém até o último tempo analisado (96 horas), conforme ilustra também a figura 16-B. Os valores de *fold-change* estão expressos em valores logarítmicos na base 2. Isso significa que, no eixo das ordenadas, o valor 2 corresponde a genes 4x mais expressos e o valor 4 corresponde a genes 8 vezes mais expressos. Da mesma forma, o valor -2 corresponde a genes 4x menos expressos.

Tabela 3: Quantidade de genes diferentemente expressos em cada tempo analisado.

	l.a.i	4 horas	24 horas	48 horas	96 horas
Subexpressos	25	59	46	26	6
Superexpressos	2	16	76	67	93

l.a.i. = logo após a irradiação

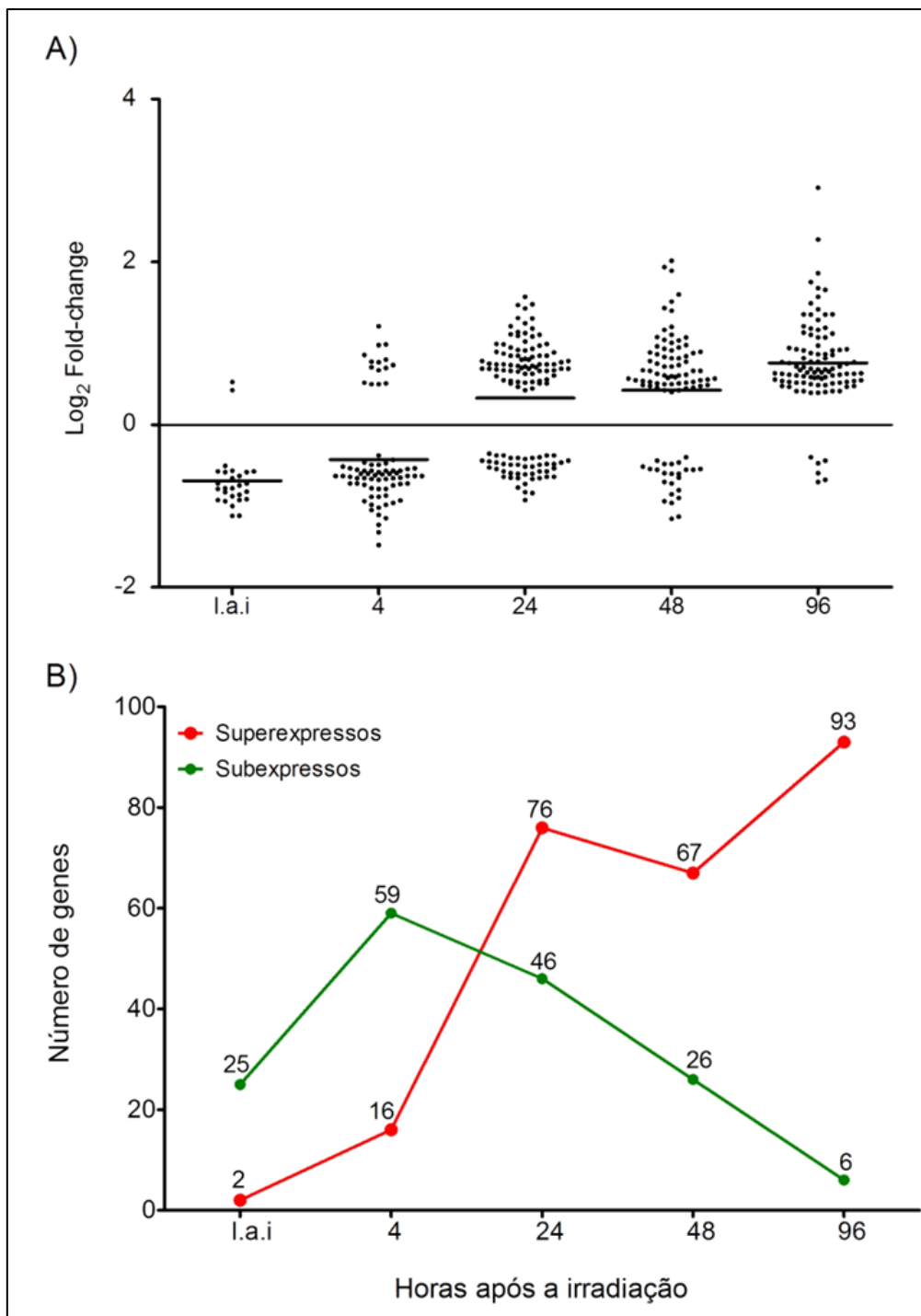


Figura 16: Número de genes diferentemente expressos nos tempos definidos após a irradiação. A) Gráfico de dispersão indicando a distribuição dos valores de *fold-change* dos genes significativamente diferentemente expressos ao longo dos tempos analisados. B) Número de genes que foram diferentemente expressos ao longo dos tempos. ○ = superexpresso, □ = subexpresso. l.a.i = logo após a irradiação.

Os gráficos do tipo vulcão (figura 17) mostram os valores da intensidade dos sinais em um sistema de dois eixos. O eixo x representa o  $\log_2$  da razão R/G (*fold-change*, ou valor de M), que indica quantas vezes o gene está mais expresso ou não e o eixo y corresponde ao valor de p-ajustado (*q-value*), que é a probabilidade que certo gene seja diferentemente expresso. As áreas cinza claro contêm os genes que possuem 95% de chance (valor p-ajustado  $\leq 0,05$ ) de serem pelo menos 1,5X mais expressos ( $\log_2 R/G \geq 0,6$  ou  $\leq -0,6$ ). A área cinza escuro representa genes que tiveram valor p-ajustado significativo, mas que não possuem diferença de expressão maior do que 1,5X entre os grupos. Os pontos localizados no lado esquerdo (lado negativo) são aqueles menos expressos, enquanto os pontos localizados no lado direito (lado positivo) são os mais expressos nas amostras testes (logo após e 4, 24, 48, e 96 horas após a irradiação). Esses gráficos ilustram os baixos valores de *fold-change* apresentados pelos genes diferentemente expressos. Conforme já explicado, o valor de *fold-change* indica o número de vezes que certo gene é mais ou menos expresso em duas situações diferentes.

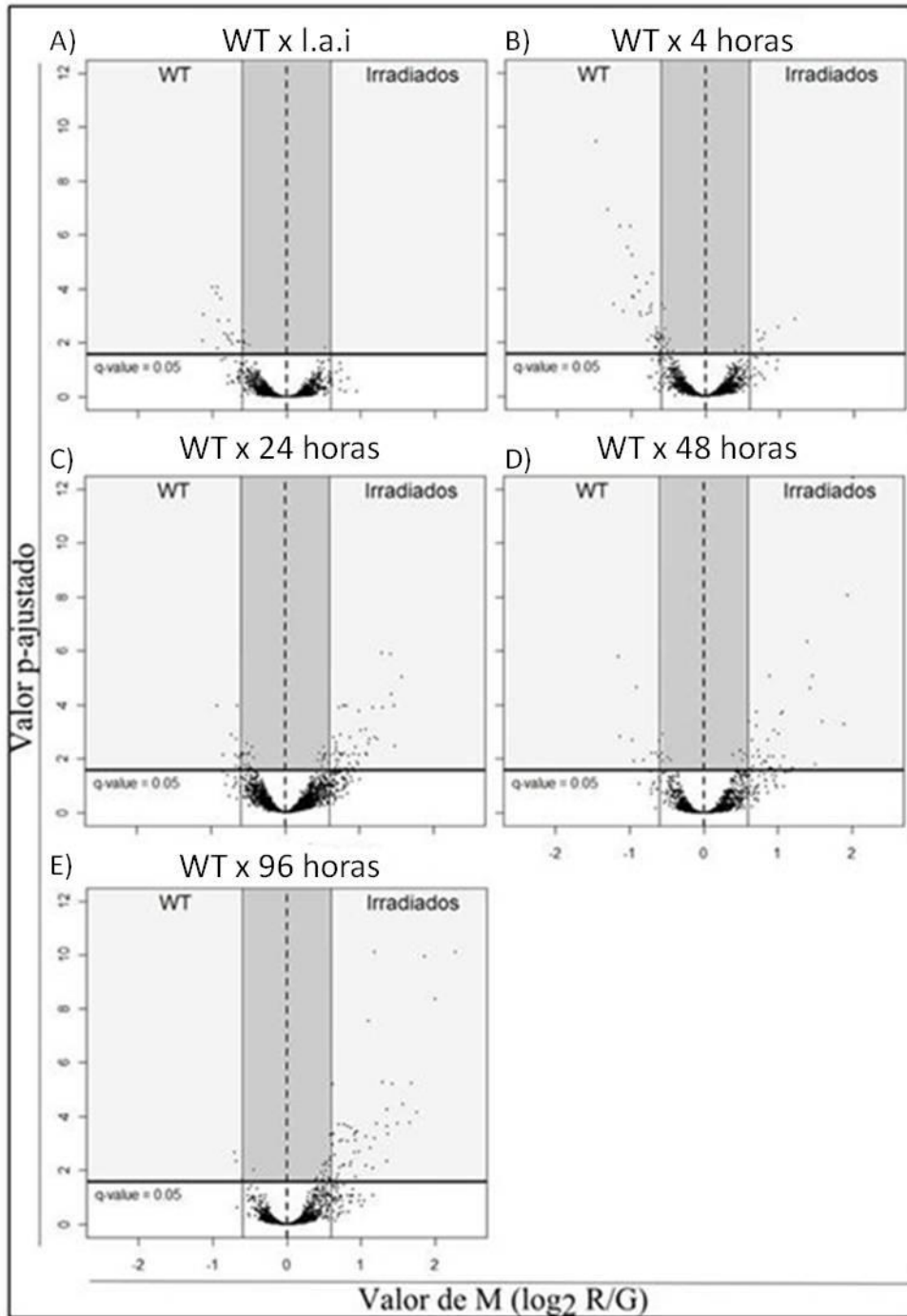


Figura 17: Gráficos do tipo vulcão para os tempos logo após a irradiação – l.a.i (A), 4 horas (B), 24 horas (C), 48 horas (D) e 96 horas (E) após a irradiação, comparados com a amostra não irradiada (WT). O eixo x representa os valores de M ( $\log_2 R/G$ ), e o eixo y representa os valores de p-ajustado. O lado esquerdo (valores abaixo de zero) representa os genes com maior expressão na amostra controle (WT), e o lado direito (valores positivos), nas amostras testes (irradiadas).

#### 4.5.2 – Categorias de genes diferentemente expressos

Genes (sub ou superexpressos) que codificam para proteínas com funções conhecidas e/ou possuem domínios preditos conservados correspondem a 41,4% dos genes presentes na lista dos diferentemente expressos. Os genes que codificam para proteínas com *hot-spot* para inserção de retroelementos (RHS), por serem parte de uma família multigênica foram considerados com uma categoria à parte e correspondem 7,7% dos genes da lista. Os genes remanescentes foram anotados como proteínas hipotéticas conservadas ou não (35,5%) e obsoletas (15,4%), que serão explicadas no item 4.6. Essa análise trouxe um panorama interessante para este trabalho, e ele pode ser visualizado na figura 18, que mostra a distribuição dessas quatro categorias de genes sub ou superexpressos ao longo do tempo. Genes com função predita (laranja) são os mais prevalentes entre os subexpressos. No entanto, a categoria dos genes superexpressos é composta principalmente por aqueles que codificam para proteínas hipotéticas (verde). As sequências obsoletas (vermelho) são bem distribuídas ao longo dos vários tempos, independentemente de serem sub ou superexpressas. Outro fato interessante é que os RHSs (amarelo) só estão presentes entre os genes superexpressos.

A tabela 4 mostra os valores de *fold-change* de cada gene que foi diferentemente expresso em pelo menos um dos tempos analisados. Os valores em negrito são os estatisticamente significativos. Nesta tabela, os genes que codificam para proteínas com funções desconhecidas (hipotéticas), RHSs, e sequências obsoletas foram excluídos, pois serão tratados em seções posteriores.

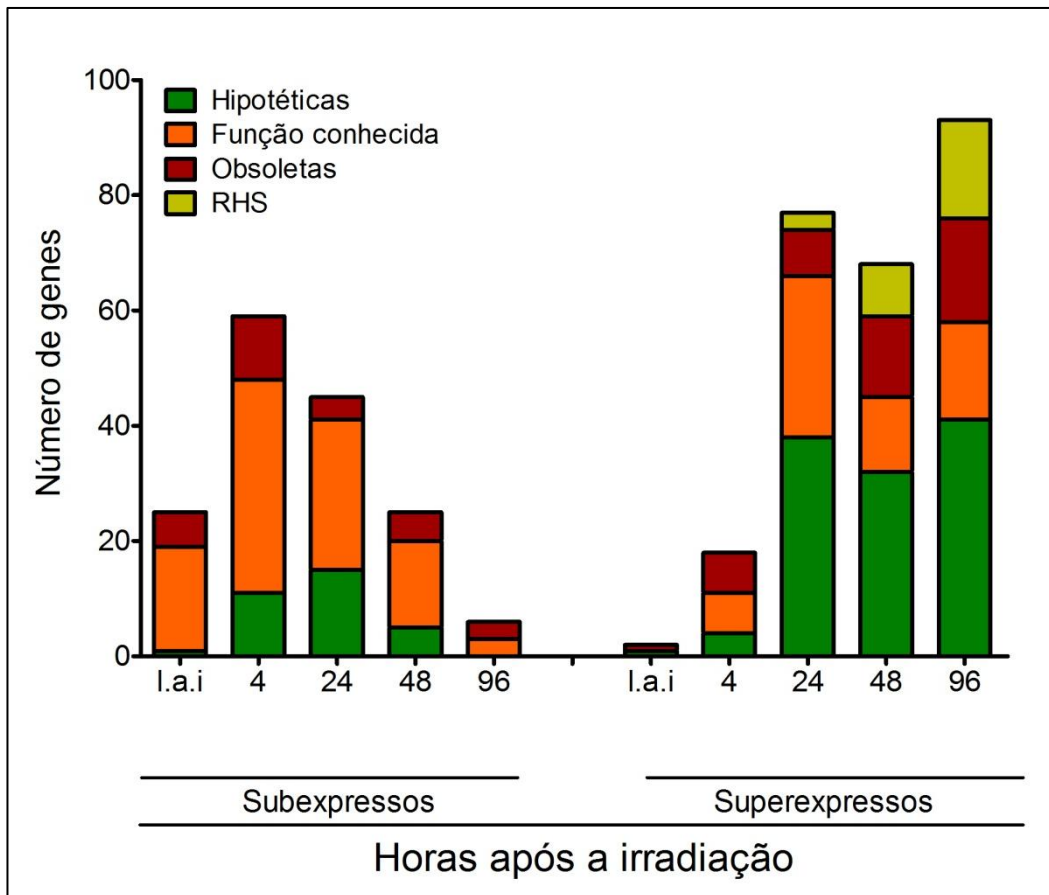


Figura 18: Categorias dos genes diferentemente expressos e suas distribuições ao longo do tempo. Verde = genes que codificam para proteínas hipotéticas. Laranja = genes que codificam proteínas com funções conhecidas. Vermelho = sequências obsoletas. Amarelo = genes RHS. l.a.i = logo após a irradiação.

Tabela 4: Valores de *fold-change* dos genes diferentemente expressos que codificam proteínas de função conhecida.

Gene ID	Nome do produto	<i>Fold-change</i>				
		l.a.i	4 hs	24 hs	48 hs	96 hs
<b>Subexpressos</b>						
Tc00.1047053507671.30	25 kDa translation elongation factor 1-beta	-1,61	<b>-1,55</b>	-1,09	-1,50	-1,34
Tc00.1047053504221.20	26S proteasome regulatory non-ATPase subunit	-1,05	-1,20	<b>-1,44</b>	-1,24	-1,10
Tc00.1047053509693.100	2-aminoethylphosphonate:pyruvateamino-transferase-like protein	<b>-1,42</b>	<b>-1,51</b>	-1,14	-1,19	-1,24
Tc00.1047053504427.70	3-oxo-5-alpha-steroid 4-dehydrogenase	-1,10	-1,12	<b>-1,43</b>	-1,05	-1,06
Tc00.1047053506679.150	40S ribosomal protein S10	-1,04	-1,12	-1,16	<b>-1,36</b>	-1,17
Tc00.1047053510101.430	40S ribosomal protein S21	1,03	-1,10	-1,31	<b>-1,51</b>	-1,22
Tc00.1047053509353.30	40S ribosomal protein S3	1,00	-1,09	-1,17	<b>-1,47</b>	-1,20
Tc00.1047053506297.150	40S ribosomal protein S5	-1,06	-1,12	<b>-1,58</b>	-1,30	-1,23
Tc00.1047053510425.19	40S ribosomal protein SA	-1,04	-1,03	<b>-1,36</b>	<b>-1,27</b>	-1,23
Tc00.1047053503719.20	40S ribosomal protein SA	-1,03	-1,04	<b>-1,34</b>	-1,20	-1,05
Tc00.1047053505977.26	60S acidic ribosomal protein P2	-1,06	-1,12	<b>-1,30</b>	-1,27	-1,27
Tc00.1047053507251.20	60S ribosomal protein L21E	-1,06	-1,09	<b>-1,28</b>	<b>-1,46</b>	-1,18
Tc00.1047053508461.490	60S ribosomal protein L23	1,06	-1,11	-1,35	<b>-1,46</b>	<b>-1,51</b>
Tc00.1047053506297.270	60S ribosomal protein L28	-1,06	-1,13	<b>-1,51</b>	-1,35	-1,23
Tc00.1047053511211.120	Activated protein kinase C receptor	-1,16	-1,17	-1,32	<b>-1,32</b>	-1,11
Tc00.1047053411235.9	Alpha tubulin	<b>-1,89</b>	<b>-1,85</b>	-1,43	-1,33	-1,32
Tc00.1047053510655.120	Aminopeptidase P	-1,20	<b>-1,55</b>	1,17	-1,29	-1,02
Tc00.1047053506563.40	Beta tubulin	<b>-1,65</b>	<b>-1,68</b>	-1,19	-1,24	-1,07
Tc00.1047053504153.160	Carboxypeptidase	-1,50	<b>-1,58</b>	-1,19	-1,29	-1,02
Tc00.1047053507873.20	Cell differentiation protein	-1,14	-1,19	<b>-1,30</b>	-1,13	-1,16
Tc00.1047053506247.50	Chaperonin	-1,28	<b>-1,45</b>	-1,11	-1,17	-1,15
Tc00.1047053507641.280	Chaperonin HSP60, mitochondrial precursor	<b>-1,92</b>	<b>-1,85</b>	-1,18	-1,40	-1,23
Tc00.1047053510187.420	Chaperonin HSP60, mitochondrial precursor (pseudogene)	-1,27	-1,19	-1,10	<b>-1,38</b>	-1,20
Tc00.1047053510187.270	COP-coated vesicle membrane protein erv25 precursor	-1,10	<b>-1,48</b>	-1,07	-1,09	-1,11
Tc00.1047053506529.550	Cysteine peptidase	<b>-2,00</b>	<b>-2,50</b>	-1,38	<b>-2,23</b>	-1,35
Tc00.1047053508317.10	Cysteine proteinase	-1,35	<b>-1,58</b>	-1,06	-1,13	-1,26
Tc00.1047053511391.160	Cytochrome c1, heme protein, mitochondrial precursor	-1,42	-1,33	-1,13	<b>-1,63</b>	-1,39
Tc00.1047053510099.120	D-isomer specific 2-hydroxyacid dehydrogenase protein	-1,07	-1,22	<b>-1,37</b>	-1,30	-1,21
Tc00.1047053510119.9	Elongation factor 1-alpha (EF-1-alpha)	<b>-1,84</b>	<b>-2,22</b>	-1,30	<b>-1,57</b>	<b>-1,63</b>
Tc00.1047053508153.730	Elongation initiation factor 2 alpha subunit	1,03	<b>-1,38</b>	-1,13	-1,06	-1,23
Tc00.1047053463269.10	Eukaryotic translation initiation factor 1A	-1,17	-1,31	<b>-1,45</b>	-1,16	-1,07
Tc00.1047053506943.160	Eukaryotic translation initiation factor 3 subunit 7-like protein	-1,06	-1,00	<b>-1,30</b>	1,05	1,01
Tc00.1047053506679.70	Eukaryotic translation initiation factor 6 (eIF-6)	<b>-1,58</b>	<b>-1,92</b>	-1,33	-1,23	-1,14
Tc00.1047053511823.70	Farnesyl pyrophosphate synthase	-1,43	<b>-1,55</b>	-1,32	-1,26	-1,21
Tc00.1047053511075.9	Fatty acid desaturase	-1,41	<b>-1,59</b>	-1,11	-1,12	1,01
Tc00.1047053506661.30	Fatty acid elongase	<b>-1,72</b>	<b>-1,65</b>	-1,40	-1,18	-1,30
Tc00.1047053507891.47	Flagellar calcium-binding protein	-1,04	-1,02	-1,10	<b>-1,75</b>	-1,12
Tc00.1047053507547.90	Glycosomal phosphoenolpyruvate carboxykinase	<b>-1,50</b>	<b>-1,98</b>	-1,18	<b>-1,43</b>	-1,06
Tc00.1047053508179.70	GPR1/FUN34/yaaH family	-1,31	-1,29	<b>-1,56</b>	-1,05	-1,01
Tc00.1047053503539.30	GTP-binding nuclear protein rtb2	<b>-2,18</b>	-1,37	1,16	-1,03	1,66
Tc00.1047053510131.40	Haloacid dehalogenase-like hydrolase	-1,34	<b>-1,55</b>	<b>-1,90</b>	<b>-1,47</b>	<b>-1,60</b>
Tc00.1047053511745.10	Heat shock 70 kDa protein, mit.precursor	1,08	1,02	-1,16	<b>-1,26</b>	-1,18
Tc00.1047053511041.40	Hexose transporter	<b>-1,78</b>	<b>-1,98</b>	-1,02	-1,31	1,07
Tc00.1047053507943.40	Histone H4	-1,11	<b>-1,35</b>	-1,25	-1,05	-1,09
Tc00.1047053509793.10	Kinetoplast DNA-associated protein	-1,11	-1,19	<b>-1,46</b>	1,00	1,11
Tc00.1047053510667.14	Membrane transporter protein	-1,09	-1,09	<b>-1,57</b>	1,03	-1,03
Tc00.1047053436521.9	Mevalonate kinase	-1,19	-1,36	<b>-1,66</b>	-1,14	-1,19
Tc00.1047053436521.9	Mevalonate kinase	-1,22	<b>-1,45</b>	-1,36	-1,18	-1,35
Tc00.1047053510155.20	Mitochondrial RNA editing ligase 1	-1,18	<b>-1,48</b>	-1,22	-1,20	-1,19
Tc00.1047053508173.100	Monoxygenase	-1,01	-1,16	<b>-1,40</b>	-1,22	-1,14

Tc00.1047053511817.40	NADH-cytochrome B5 reductase	<b>-1,90</b>	-1,49	1,21	1,06	-1,05
Tc00.1047053510645.20	Nuclear transcription factor	-1,38	<b>-1,59</b>	-1,42	-1,14	-1,22
Tc00.1047053508173.180	Nuclear transport factor 2 protein (NFT2)	-1,10	-1,31	<b>-1,52</b>	-1,27	-1,19
Tc00.1047053511573.58	Nucleolar protein	-1,09	-1,17	<b>-1,78</b>	-1,07	-1,03
Tc00.1047053510859.17	Nucleolar RNA-binding protein	-1,46	<b>-1,53</b>	-1,22	-1,17	-1,14
Tc00.1047053508707.200	Nucleoside diphosphate kinase	-1,06	-1,02	-1,29	<b>-1,65</b>	-1,20
Tc00.1047053506773.50	Nucleoside transporter-like	-1,33	<b>-1,52</b>	-1,09	-1,03	1,21
Tc00.1047053511355.30	Phosphatidic acid phosphatase protein	-1,18	-1,21	<b>-1,79</b>	-1,02	-1,15
Tc00.1047053507617.9	Prostaglandin F2alpha synthase	-1,22	<b>-1,66</b>	-1,08	-1,12	-1,01
Tc00.1047053508461.80	Prostaglandin F2alpha synthase	<b>-1,72</b>	<b>-1,91</b>	-1,03	-1,32	-1,18
Tc00.1047053511635.40	Protein tyrosine phosphatase	<b>-1,55</b>	<b>-1,67</b>	<b>-1,38</b>	-1,09	-1,12
Tc00.1047053511715.100	Pumilio/PUF RNA binding protein 7	1,06	1,04	<b>-1,71</b>	-1,06	1,02
Tc00.1047053510091.80	Pyruvate dehydrogenase E1 beta subunit	<b>-1,65</b>	<b>-1,73</b>	<b>-1,59</b>	-1,34	-1,24
Tc00.1047053508479.340	Succinyl-CoA synthetase alpha subunit	<b>-1,73</b>	<b>-2,16</b>	<b>-1,56</b>	-1,39	-1,38
Tc00.1047053504045.60	Thermostable carboxypeptidase 1	-1,26	<b>-1,52</b>	1,04	-1,08	-1,20
Tc00.1047053506855.260	Thymidine kinase	-1,53	<b>-1,83</b>	-1,17	-1,35	-1,15
Tc00.1047053503555.30	Trypanothione reductase	-1,33	<b>-1,39</b>	<b>-1,33</b>	-1,15	-1,16
Tc00.1047053508421.30	Tryptophanyl-tRNA synthetase (pseudogene)	<b>-1,50</b>	<b>-1,73</b>	-1,33	-1,35	-1,24
Tc00.1047053503487.50	UDP-Gal or UDP-GlcNAc-dependent glycosyltransferase	1,12	-1,12	<b>-1,53</b>	1,05	-1,03
Tc00.1047053503929.10	V-type ATPase, A subunit	<b>-1,68</b>	<b>-1,58</b>	-1,22	-1,14	-1,14
Tc00.1047053509669.40	Zn-finger protein	-1,23	<b>-1,65</b>	1,11	-1,20	1,05

### Superexpressos

Tc00.1047053509941.100	2,4-dienoyl-coa reductase FADH1	1,10	1,29	<b>1,67</b>	1,33	1,32
Tc00.1047053507611.290	acetyltransferase	-1,04	-1,01	<b>1,42</b>	<b>1,40</b>	1,32
Tc00.1047053506357.50	alcohol dehydrogenase	-1,20	-1,17	<b>1,42</b>	1,16	1,11
Tc00.1047053504071.110	ama1 protein	1,01	1,17	1,41	1,44	<b>1,64</b>
Tc00.1047053508737.194	ARP2/3 complex subunit	1,03	1,24	<b>1,58</b>	1,40	1,33
Tc00.1047053508375.30	aspartate carbamoyltransferase	1,00	-1,13	<b>1,99</b>	-1,12	1,04
Tc00.1047053508183.4	aspartyl aminopeptidase	-1,09	1,09	<b>1,59</b>	1,05	-1,02
Tc00.1047053505997.70	ATPase	1,19	1,41	1,36	<b>1,59</b>	<b>1,50</b>
Tc00.1047053508903.100	ATPase	1,29	1,37	1,27	<b>1,52</b>	<b>1,88</b>
Tc00.1047053504149.20	ATP-binding cassette transporter ABCA1	<b>1,34</b>	1,24	1,06	1,06	1,07
Tc00.1047053506563.210	calpain-like cysteine peptidase	1,03	1,19	1,43	<b>1,46</b>	1,23
Tc00.1047053506563.170	calpain-like cysteine peptidase (pseudo.)	-1,06	1,10	<b>1,38</b>	1,10	<b>1,31</b>
Tc00.1047053508357.80	cation transporter protein	1,20	<b>1,74</b>	<b>1,75</b>	<b>1,82</b>	<b>1,55</b>
Tc00.1047053507929.20	co-chaperone GrpE	1,09	1,16	<b>1,64</b>	-1,01	-1,03
Tc00.1047053509455.140	cyclin	1,03	1,15	<b>1,67</b>	1,26	1,34
Tc00.1047053511421.110	developmentally regulated phosphoprotein	1,04	1,26	<b>1,59</b>	1,28	<b>1,59</b>
Tc00.1047053509585.10	dynein heavy chain	1,18	1,32	1,37	1,31	<b>1,56</b>
Tc00.1047053510687.10	dynein heavy chain (pseudogene)	1,07	-1,15	1,03	1,19	<b>1,47</b>
Tc00.1047053510897.6	dynein light chain lc6, flagellar outer arm	1,09	1,21	<b>1,34</b>	1,32	<b>1,45</b>
Tc00.1047053506391.30	EF-hand protein 5	-1,05	1,15	<b>1,88</b>	1,37	<b>1,61</b>
Tc00.1047053507483.20	EF-hand protein 5	-1,29	1,13	<b>2,15</b>	1,46	<b>1,56</b>
Tc00.1047053506529.508	glucose-6-phosphate isomerase, glycosomal	1,19	<b>1,66</b>	<b>2,69</b>	<b>2,06</b>	<b>1,92</b>
Tc00.1047053506529.508	glucose-6-phosphate isomerase, glycosomal	-1,03	1,34	<b>1,74</b>	<b>1,64</b>	1,32
Tc00.1047053508457.30	glycine cleavage system H protein	-1,11	1,15	<b>1,42</b>	1,15	1,33
Tc00.1047053510199.10	GP85-like protein	1,16	1,23	1,16	<b>1,32</b>	<b>1,54</b>
Tc00.1047053510659.240	lactoylglutathione lyase-like protein	1,04	1,35	1,57	<b>1,71</b>	1,36
Tc00.1047053510743.70	lactoylglutathione lyase-like protein	-1,23	1,32	<b>1,58</b>	1,32	1,11
Tc00.1047053473111.10	lathosterol oxidase	-1,11	1,65	<b>2,79</b>	1,98	<b>2,56</b>
Tc00.1047053506341.10	N-acetylglucosamine-6-P deacetylase-like prot.	1,23	<b>1,42</b>	1,03	-1,19	-1,17
Tc00.1047053508461.400	nucleoside diphosphate kinase	1,06	1,15	<b>1,67</b>	1,10	1,29
Tc00.1047053504113.10	protein kinase	1,65	<b>1,62</b>	<b>1,60</b>	<b>1,42</b>	1,27
Tc00.1047053511671.80	protein kinase	1,29	1,38	1,12	<b>1,48</b>	1,41
Tc00.1047053511269.50	protein kinase A catalytic subunit	1,05	-1,11	<b>1,45</b>	1,34	<b>1,49</b>
Tc00.1047053508215.9	protein kinase C substrate protein, heavy chain	-1,09	1,14	<b>1,61</b>	1,39	1,24
Tc00.1047053510121.130	serine/threonine protein kinase	1,41	1,07	1,29	<b>2,05</b>	1,77
Tc00.1047053509907.60	trans-sialidase (pseudogene)	1,14	1,08	1,04	1,24	<b>1,41</b>
Tc00.1047053503899.119	trypanothione/tryparedoxin dependent peroxidase 2	1,05	<b>1,33</b>	1,18	1,03	-1,06
Tc00.1047053508649.5	tryparedoxin peroxidase	-1,03	1,21	<b>1,46</b>	1,24	1,21
Tc00.1047053509069.30	tubulin binding cofactor A-like protein	-1,02	-1,02	1,52	1,78	<b>2,10</b>
Tc00.1047053506619.40	tyrosyl-DNA phosphodiesterase (Tdp1)	-1,00	1,01	<b>1,86</b>	1,29	1,32
Tc00.1047053510105.100	UDP-glucose dehydrogenase	1,06	-1,02	<b>1,44</b>	1,16	1,17
Tc00.1047053508479.290	vacuolar sorting protein	1,07	-1,03	<b>1,89</b>	1,12	1,15

Valores em negrito: valores estatisticamente significativos ( $p \leq 0,05$ ). I.a.i = logo após a irradiação



#### 4.6 – O efeito da irradiação na expressão das sequências obsoletas

---

A lâmina de microarranjos contém 1.672 (dos 12.288 oligos válidos) anotados como sequências obsoletas. No entanto, resultados mostraram que 42 destas sequências são diferentemente expressas, sendo 15 subexpressas e 27 superexpressas. Dessa forma, alguns questionamentos foram levantados: i) o que elas realmente são? ii) onde elas se localizam no genoma? iii) O que elas codificam? iv) Por que algumas apresentam valores tão altos de *fold-change*, como a sequência #5737, que está 7,53 vezes mais expressa quando comparada com o não irradiado?

Para começar a responder essas perguntas, uma busca por similaridade contra o banco de dados de nucleotídeos do NCBI foi a estratégia escolhida. Assim, todas as 1.672 sequências dos oligonucleotídeos anotados como obsoletas foram usadas como arquivo de entrada utilizando o algoritmo blastx. Quatro alinhamentos foram perfeitos, revelando que havia quatro genes do maxicírculo (componente do cinetoplasto, o DNA mitocondrial de *T. cruzi*, ou kDNA) presentes na lâmina (oligos #5737, 8141, 10153 e 14569), e todos foram diferentemente expressos.

O maxicírculo possui no total 18 genes (Simpson *et al.*, 1987), e um busca extra por similaridade entre as sequências obsoletas e as sequências de nucleotídeos do maxi e minicírculo, utilizando o algoritmo blastn, na tentativa de identificar novos genes, mas não houveram novos resultados positivos. As sequências obsoletas identificadas correspondem aos genes codificadores das proteínas citocromo oxidase I (COI), NADH-desidrogenase 1 (ND-1), NADH-desidrogenase 5 (ND-5) e *Mitochondrial undefined reading frame I* (MURF-1). Os valores de *fold-change* foram traçados em um gráfico que revelou que os genes possuem perfis semelhantes de expressão ao longo do tempo (figura 19), com exceção na ND-5. Mais curioso ainda, os genes apresentam um

drástico aumento de expressão após 48 e principalmente 96 horas da irradiação. Os valores de *fold-change* estão expressos em valores logarítmicos na base 2. Isso significa que, no eixo das ordenadas, o valor 1 corresponde a genes 2x mais expressos, o valor 2 corresponde a genes 4x mais expressos, e assim sucessivamente.

O próximo passo foi localizar as sequências diferentemente expressas (excluindo as do kDNA) no genoma de *T. cruzi*, utilizando o software Artemis (Carver *et al.*, 2008). Os resultados dessa análise estão resumidos na tabela 5: i) 17 sequências foram localizadas no genoma em regiões fora de regiões codificadoras (CDS), ii) oito foram localizados em mais de uma região do genoma e iii) 11 não foram localizadas em lugar algum.

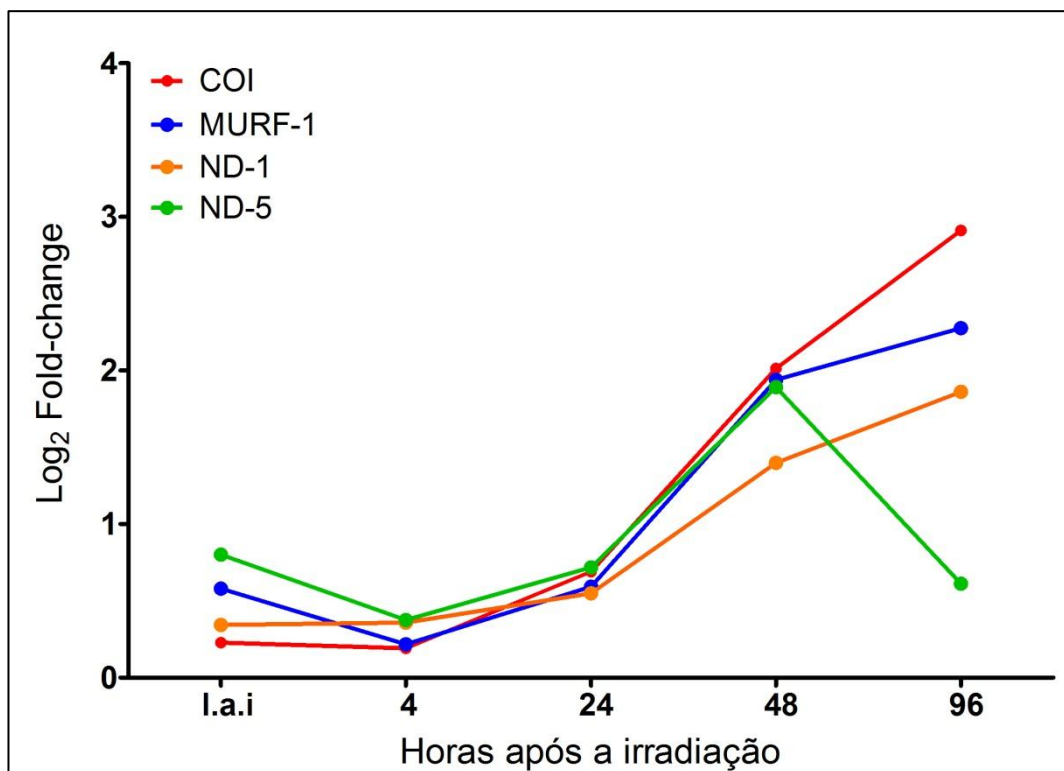


Figura 19: Perfil de expressão gênica de quatro genes do maxicículo após a irradiação. COI = Citocromo oxidase I, ND-1 = NADH-desidrogenase 1, ND-5 = NADH-desidrogenase 5, MURF-1 = *Mitochondrial undefined reading frame 1*. l.a.i = logo após a irradiação.

Tabela 5: Valores de *fold-change* das sequências obsoletas diferentemente expressas

Gene ID	Localização no genoma	<i>Fold-change</i>				
		l.a.i	4 hs	24 hs	48 hs	96 hs
<b>Subexpressos</b>						
4010 (ID - lâmina)	Não localizada	-1.34	<b>-1.47</b>	-1.05	-1.13	-1.12
4685 (ID - lâmina)	Não localizada	-1.10	-1.13	<b>-1.32</b>	<b>-1.40</b>	<b>-1.32</b>
8330 (ID - lâmina)	Não localizada	-1.06	<b>-1.55</b>	-1.52	-1.00	-1.13
9101 (ID - lâmina)	Não localizada	-1.10	-1.20	-1.36	<b>-1.52</b>	-1.33
23161 (ID - lâmina)	Não localizada	-1.18	<b>-1.48</b>	-1.18	1.03	1.00
23282 (ID - lâmina)	Não localizada	<b>-1.77</b>	<b>-1.95</b>	-1.03	-1.12	1.13
24338 (ID - lâmina)	Não localizada	<b>-1.49</b>	<b>-1.65</b>	-1.06	-1.11	-1.04
5692 (ID - lâmina)	Fora de CDS	-1.28	<b>-1.52</b>	-1.23	-1.20	-1.16
21242 (ID - lâmina)	Fora de CDS	<b>-1.90</b>	<b>-2.79</b>	<b>-1.49</b>	<b>-1.87</b>	<b>-1.39</b>
3577 (ID - lâmina)	Multilocal	-1.06	-1.07	<b>-1.41</b>	-1.24	-1.09
5161 (ID - lâmina)	Multilocal	-1.05	-1.10	<b>-1.56</b>	-1.19	<b>-1.36</b>
6316 (ID - lâmina)	Multilocal	-1.46	<b>-2.07</b>	-1.09	-1.37	-1.20
6794 (ID - lâmina)	Multilocal	<b>-1.49</b>	<b>-2.03</b>	-1.18	<b>-1.40</b>	-1.27
22884 (ID - lâmina)	Multilocal	<b>-1.82</b>	<b>-2.35</b>	-1.41	<b>-2.19</b>	-1.16
23089 (ID - lâmina)	Multilocal	<b>-1.48</b>	<b>-1.46</b>	-1.23	-1.13	1.01
<b>Superexpressos</b>						
2117 (ID - lâmina)	Fora de CDS	1,12	1,31	1,23	1,32	<b>1,71</b>
9204 (ID - lâmina)	Fora de CDS	1,27	1,11	1,35	1,66	<b>2,97</b>
11747 (ID - lâmina)	Fora de CDS	1,10	<b>1,81</b>	<b>2,20</b>	<b>1,86</b>	<b>2,28</b>
14862 (ID - lâmina)	Fora de CDS	1,66	<b>1,97</b>	1,50	<b>1,97</b>	<b>2,24</b>
21056 (ID - lâmina)	Fora de CDS	1,18	1,18	1,28	1,25	<b>1,70</b>
26458 (ID - lâmina)	Fora de CDS	<b>1,44</b>	1,23	1,10	-1,15	<b>1,44</b>
5737 (ID - lâmina)	kDNA	1,17	1,14	<b>1,61</b>	<b>4,04</b>	<b>7,53</b>
8041 (ID - lâmina)	kDNA	1,49	1,16	1,51	<b>3,83</b>	<b>4,84</b>
10153 (ID - lâmina)	kDNA	1,27	1,28	1,46	<b>2,64</b>	<b>3,63</b>
14569 (ID - lâmina)	kDNA	1,74	1,30	1,65	<b>3,71</b>	1,53
6361 (ID - lâmina)	Não localizada	1,36	<b>1,43</b>	1,02	1,07	1,05
6364 (ID - lâmina)	Não localizada	1,01	<b>1,71</b>	<b>2,48</b>	<b>1,93</b>	<b>1,71</b>
11281 (ID - lâmina)	Não localizada	-1,01	1,20	1,07	<b>1,41</b>	<b>2,27</b>
13999 (ID - lâmina)	Não mapeado*	1,16	1,38	1,46	<b>1,58</b>	1,44
21269 (ID - lâmina)	Não mapeado*	1,14	1,09	<b>1,70</b>	-1,13	1,06
5884 (ID - lâmina)	Fora de CDS	-1,10	1,37	1,56	<b>2,30</b>	<b>3,15</b>
7785 (ID - lâmina)	Fora de CDS	1,06	1,34	1,47	<b>1,88</b>	<b>3,20</b>
13664 (ID - lâmina)	Fora de CDS	1,09	1,10	<b>2,18</b>	1,10	-1,04
13828 (ID - lâmina)	Fora de CDS	1,36	<b>1,70</b>	1,34	1,27	1,52
16329 (ID - lâmina)	Fora de CDS	1,28	<b>1,41</b>	<b>1,45</b>	<b>1,39</b>	1,03
18850 (ID - lâmina)	Fora de CDS	-1,03	1,34	<b>1,80</b>	1,08	1,08
19423 (ID - lâmina)	Fora de CDS	1,07	1,28	1,57	1,52	<b>1,70</b>
21055 (ID - lâmina)	Fora de CDS	1,20	1,33	1,29	1,22	<b>1,40</b>
23196 (ID - lâmina)	Fora de CDS	1,15	<b>1,59</b>	<b>1,63</b>	1,43	<b>1,90</b>
1404 (ID - lâmina)	Multilocal	-1,16	1,11	1,05	<b>1,45</b>	<b>1,76</b>
1497 (ID - lâmina)	Multilocal	-1,05	-1,06	1,04	<b>1,58</b>	1,22
18649 (ID - lâmina)	Multilocal	1,17	1,14	1,23	1,42	<b>1,82</b>

Valores em negrito: valores estatisticamente significativos ( $p \leq 0,05$ ). l.a.i = logo após a irradiação. \*Em contigs não utilizados na montagem final do genoma. CDS = *coding sequence*

#### 4.7 – O efeito da irradiação na expressão dos genes RHS

---

RHSs são genes pertencentes a uma família gênica numerosa, com 752 membros identificados, sendo 557 anotados como pseudo-genes (El-Sayed *et al.*, 2005a). São caracterizadas pela presença de sítio de inserção de RIME/*ingi*, e se localizam preferencialmente nas regiões subteloméricas (especificamente entre a sequência associada ao telômero e o primeiro gene com sequência não-repetitiva) (Bringaud *et al.*, 2002).

A radiação gama foi capaz de induzir a expressão desses genes. Dos 379 RHSs presentes na lâmina, 21 foram considerados superexpressos (tabela 6). Não há nenhum gene codificador para RHS subexpresso. A localização de todos os genes diferentemente expressos foi pesquisada, e o resultado também se encontra na tabela 6: dez se encontram nas pontas dos cromossomos, quatro no meio, porém em regiões repetitivas, e sete em contigs não utilizados para a montagem final do genoma.

Nós nos questionamos se a radiação gama induziu a expressão desses genes, ou se o resultado observado foi ao acaso. Para isso, foi realizado um teste qui-quadrado (duas-caudas) para cada tempo analisado. Os resultados foram significativos nos tempos 48 e 96 horas ( $p < 0,0001$ ), o que indica o efeito da radiação sobre a expressão destes genes (figura 20). Dessa forma, levantamos a hipótese de que essa família gênica possa estar relacionada de forma direta ou indireta na resposta à reparação dos danos causados no DNA após radiação gama. Os genes RHS se localizam agrupados no genoma com outras proteínas codificadas por grandes famílias multigênicas do parasito, tais como trans-sialidases, família com 1430 membros, e DGF-1 (*Dispersed gene family*), com 565 membros (El-Sayed *et al.*, 2005a). No entanto, apenas duas proteínas do tipo trans-sialidase (trans-sialidase (pseudogene) e gp-85) foram superexpressas.

Tabela 6: Valores de *fold-change* dos genes RHS diferentemente expressos

Gene ID	Localização no genoma	l.a.i	4 hs	<i>Fold-change</i>		
				24 hs	48 hs	96 hs
<b>Superexpressos</b>						
Tc00.1047053445777.10	Na ponta do cromossomo (TcChr40-P)	1,17	1,08	1,03	1,21	<b>1,38</b>
Tc00.1047053508071.90	Último gene do contig Tcruzi_7589	-1,03	1,20	1,20	1,29	<b>1,33</b>
Tc00.1047053509259.180	Na ponta do cromossomo (TcChr25-P)	1,18	1,17	1,31	1,21	<b>1,56</b>
Tc00.1047053506537.10	Na ponta do cromossomo (TcChr23-P)	1,06	1,09	1,26	1,29	<b>1,39</b>
Tc00.1047053507975.10	Na ponta do cromossomo (TcChr17-S)	1,06	1,31	1,38	1,35	<b>1,45</b>
Tc00.1047053506951.10	TcChr24-S – próximo de trans-sialidasas, MASPs, DGFs-1	1,21	1,25	1,23	1,26	<b>1,50</b>
Tc00.1047053504099.70	Na ponta do cromossomo (TcChr26-S)	1,07	1,21	1,18	<b>1,34</b>	<b>1,44</b>
Tc00.1047053508483.40	Último gene do contig Tcruzi_7746	1,14	1,19	1,29	1,31	<b>1,33</b>
Tc00.1047053509765.60	Na ponta do cromossomo (TcChr17-S)	1,08	1,18	1,19	1,45	<b>1,57</b>
Tc00.1047053508877.30	TcChr41-P - próximo de trans-sialidasas, MASPs, DGFs-1	1,11	1,17	1,23	<b>1,47</b>	1,30
Tc00.1047053505915.10	Um dos dois genes presentes no contig Tcruzi_6735	1,07	1,15	1,07	<b>1,39</b>	<b>1,42</b>
Tc00.1047053506683.210	Na ponta do cromossomo (TcChr12-S)	1,02	1,14	<b>1,47</b>	<b>1,48</b>	1,46
Tc00.1047053506751.70	TcChr41-P - próximo de trans-sialidasas, MASPs, DGFs-1	1,10	1,13	1,15	1,22	<b>1,43</b>
Tc00.1047053504285.10	Único gene presente no contig Tcruzi_6244	1,08	1,04	1,18	<b>1,53</b>	<b>1,69</b>
Tc00.1047053507427.30	Último gene do contig Tcruzi_7332	1,02	-1,16	<b>1,54</b>	1,13	1,03
Tc00.1047053507777.30	Na ponta do cromossomo (TcChr21-P)	-1,00	1,20	1,38	<b>1,44</b>	<b>1,43</b>
Tc00.1047053506349.83	Último gene do contig Tcruzi_6908	-1,06	-1,12	<b>1,61</b>	1,02	1,05
Tc00.1047053506129.80	Na ponta do cromossomo (TcChr6-S)	1,12	1,20	1,17	<b>1,49</b>	<b>1,54</b>
Tc00.1047053506561.20	Na ponta do cromossomo (TcChr5-P)	1,14	1,15	1,24	<b>1,52</b>	<b>1,50</b>
Tc00.1047053509559.20	TcChr12-P - próximo de trans-sialidasas, MASPs, DGFs-1	-1,05	-1,13	1,16	1,23	<b>1,33</b>
Tc00.1047053413293.30	Último gene do contig Tcruzi_12248	1,07	1,02	1,15	<b>1,34</b>	<b>1,40</b>

Valores em negrito: valores estatisticamente significativos ( $p \leq 0,05$ ). l.a.i = logo após a irradiação

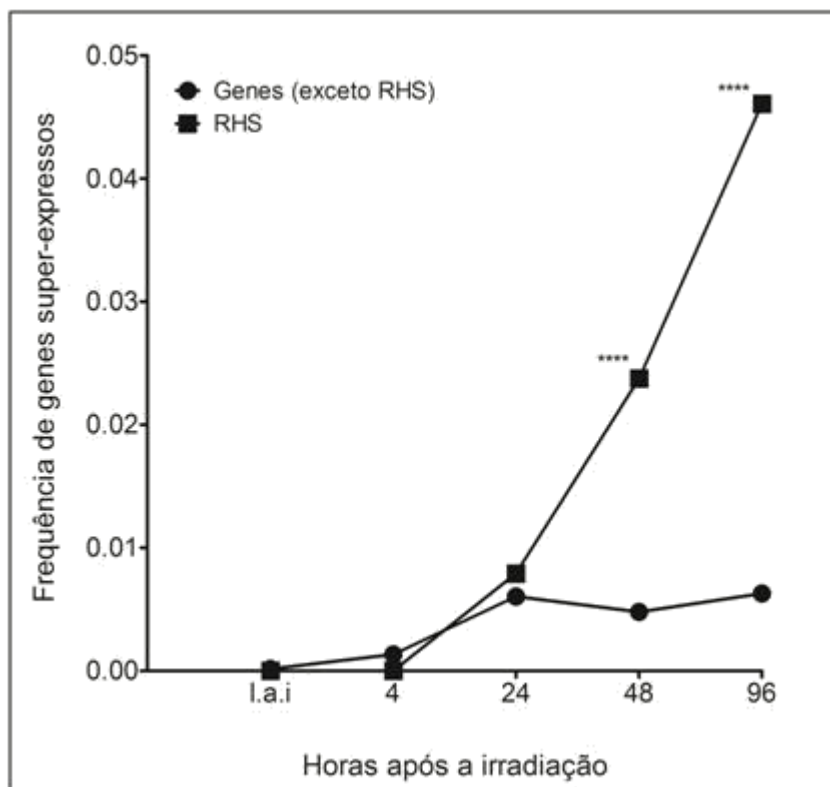


Figura 20: Frequência dos genes RHSs e não-RHS superexpressos em resposta à radiação gama ao longo do tempo. Quadrados pretos = Frequência de genes RHS superexpressos. Círculos pretos = Frequência de genes superexpressos exceto RHS. l.a.i = logo após a irradiação. \*\*\*\*  $p < 0,0001$

#### 4.8 – O efeito da irradiação na expressão de genes hipotéticos

---

Experimentos de microarranjos são capazes de fornecer um panorama bastante abrangente dos processos biológicos que foram estimulados ou reprimidos após algum tratamento aplicado em amostras testes. Proteínas de funções ainda desconhecidas geralmente figuram nas listas de genes diferentemente expressos e estas podem ser futuramente caracterizadas bioquímica e molecularmente. Porém, o cenário se torna bastante nebuloso quando o número de genes diferentemente expressos que codificam para proteínas de função desconhecida é maior do que o número de genes com função conhecida, situação que ocorreu neste trabalho. Os resultados indicaram 27 genes subexpressos e 70 superexpressos que codificam para proteínas hipotéticas (tabela 7). Sete genes subexpressos e nove superexpressos (em negrito) apresentaram regiões codificadoras de domínios protéicos conservados e foram incluídos nas análises funcionais posteriores.



Tabela 7: Valores de *fold-change* dos genes diferentemente expressos que codificam proteínas hipotéticas

Gene ID	Nome do produto	l.a.i	<i>Fold-change</i>			
			4 hs	24 hs	48 hs	96 hs
<b>Subexpressos</b>						
Tc00.1047053447255.10	hypothetical protein, conserved	1.36	-1.17	<b>-1.41</b>	-1.23	-1.12
Tc00.1047053503395.20	hypothetical protein, conserved	1.19	-1.11	<b>1.65</b>	-1.03	-1.02
Tc00.1047053503687.30	hypothetical protein, conserved	1.09	<b>-1.41</b>	-1.15	-1.23	-1.06
Tc00.1047053503703.30	hypothetical protein, conserved	1.29	<b>-1.60</b>	-1.29	-1.23	-1.09
Tc00.1047053503809.120	hypothetical protein, conserved	1.00	-1.24	<b>-1.52</b>	1.02	-1.11
Tc00.1047053503897.120	hypothetical protein, conserved	1.33	<b>-1.50</b>	<b>-1.49</b>	-1.14	-1.21
Tc00.1047053504057.80	hypothetical protein, conserved	1.22	-1.09	<b>-1.33</b>	-1.06	-1.20
<b>Tc00.1047053506219.40</b>	hypothetical protein, conserved	1.43	<b>-1.72</b>	-1.58	-1.16	-1.24
Tc00.1047053506885.30	hypothetical protein, conserved	1.25	-1.07	-1.15	<b>-1.81</b>	1.00
<b>Tc00.1047053506925.310</b>	hypothetical protein, conserved	1.17	<b>-1.48</b>	<b>1.72</b>	1.27	1.23
Tc00.1047053507053.180	hypothetical protein, conserved	1.18	-1.29	<b>-1.49</b>	<b>-1.47</b>	-1.24
Tc00.1047053507221.30	hypothetical protein, conserved	-1.03	-1.19	<b>-1.40</b>	-1.26	-1.10
<b>Tc00.1047053507509.40</b>	hypothetical protein, conserved	<b>-2.17</b>	<b>-1.57</b>	-1.40	-1.26	-1.03
Tc00.1047053507747.94	hypothetical protein	-1.41	<b>-1.43</b>	-1.19	-1.22	-1.01
<b>Tc00.1047053507993.380</b>	hypothetical protein, conserved	-1.01	-1.13	<b>-1.40</b>	-1.25	-1.15
Tc00.1047053508173.264	hypothetical protein, conserved	-1.19	<b>-1.30</b>	<b>-1.43</b>	-1.23	-1.14
Tc00.1047053508257.180	hypothetical protein, conserved	-1.03	-1.06	-1.06	<b>-1.92</b>	-1.09
Tc00.1047053508307.90	hypothetical protein, conserved	1.18	-1.12	<b>-1.41</b>	-1.09	-1.09
Tc00.1047053508719.30	hypothetical protein, conserved	-1.07	1.01	<b>-1.36</b>	-1.19	-1.15
Tc00.1047053509141.40	hypothetical protein, conserved	1.10	1.05	-1.06	<b>-1.95</b>	1.10
<b>Tc00.1047053509267.40</b>	hypothetical protein, conserved	-1.14	-1.38	-1.45	<b>-1.52</b>	-1.43
<b>Tc00.1047053509769.60</b>	hypothetical protein, conserved	-1.40	<b>-1.54</b>	-1.46	-1.24	-1.12
Tc00.1047053510099.100	hypothetical protein, conserved	-1.12	-1.16	<b>-1.30</b>	-1.26	-1.10
Tc00.1047053510543.80	hypothetical protein	-1.04	-1.05	<b>-1.33</b>	-1.07	1.01
Tc00.1047053510857.30	hypothetical protein, conserved	1.13	-1.32	<b>-1.58</b>	-1.09	1.08
<b>Tc00.1047053511071.50</b>	hypothetical protein, conserved	-1.13	<b>-1.41</b>	-1.28	-1.33	-1.20
Tc00.1047053511517.120	hypothetical protein, conserved	-1.43	<b>-1.56</b>	-1.31	-1.27	-1.32
Tc00.1047053447255.10	hypothetical protein, conserved	-1.36	-1.17	<b>-1.41</b>	-1.23	-1.12
<b>Superexpressos</b>						
Tc00.1047053478463.10	hypothetical protein	1.21	1.20	1.19	<b>1.36</b>	<b>1.63</b>
Tc00.1047053503395.20	hypothetical protein, conserved	-1.19	-1.11	<b>1.65</b>	-1.03	-1.02
Tc00.1047053503415.40	hypothetical protein	-1.09	1.19	<b>1.99</b>	<b>1.51</b>	1.21
Tc00.1047053503419.54	hypothetical protein, conserved	1.12	1.03	<b>1.67</b>	-1.04	-1.06
<b>Tc00.1047053503839.19</b>	hypothetical protein, conserved	1.11	1.05	<b>2.47</b>	1.04	1.03
<b>Tc00.1047053503841.20</b>	hypothetical protein, conserved	1.08	-1.04	<b>1.83</b>	1.01	-1.02
Tc00.1047053503911.30	hypothetical protein, conserved	1.14	1.43	<b>2.12</b>	1.46	1.45
Tc00.1047053503925.6	hypothetical protein, conserved	-1.22	1.20	<b>2.27</b>	1.43	<b>1.76</b>
Tc00.1047053503975.100	hypothetical protein, conserved	1.11	1.17	<b>1.66</b>	<b>1.37</b>	1.21
Tc00.1047053504005.54	hypothetical protein, conserved	-1.08	1.07	<b>1.36</b>	<b>1.41</b>	<b>1.52</b>
Tc00.1047053504037.10	hypothetical protein, conserved	-1.13	-1.02	<b>1.67</b>	<b>1.64</b>	<b>1.55</b>
Tc00.1047053504227.10	hypothetical protein, conserved	1.05	1.19	<b>1.44</b>	-1.03	-1.01
Tc00.1047053504433.20	hypothetical protein, conserved	1.42	1.53	<b>2.77</b>	<b>1.76</b>	<b>2.17</b>
Tc00.1047053506133.179	hypothetical protein	1.13	1.19	1.20	<b>1.85</b>	<b>2.15</b>
Tc00.1047053506195.290	hypothetical protein, conserved	1.03	<b>1.41</b>	1.20	1.03	1.07
Tc00.1047053506713.24	hypothetical protein, conserved	1.10	1.21	1.27	1.22	<b>1.33</b>
<b>Tc00.1047053506739.99</b>	hypothetical protein, conserved	1.06	1.18	<b>1.78</b>	<b>1.58</b>	<b>1.51</b>
Tc00.1047053506773.104	hypothetical protein, conserved	-1.08	1.27	1.42	<b>1.50</b>	<b>1.53</b>
Tc00.1047053506789.270	hypothetical protein, conserved	1.18	1.03	<b>1.98</b>	-1.10	-1.04
Tc00.1047053506825.200	hypothetical protein	1.03	1.18	1.36	<b>1.54</b>	1.45
Tc00.1047053506829.90	hypothetical protein, conserved	-1.13	1.59	<b>2.97</b>	<b>2.70</b>	<b>2.55</b>
Tc00.1047053506831.63	hypothetical protein, conserved	-1.04	1.23	<b>1.68</b>	1.52	1.48
Tc00.1047053506925.310	hypothetical protein, conserved	-1.17	<b>-1.48</b>	<b>1.72</b>	1.27	1.23
<b>Tc00.1047053506947.50</b>	hypothetical protein, conserved	1.25	1.16	1.06	1.19	<b>1.46</b>
Tc00.1047053507053.140	hypothetical protein, conserved	1.00	-1.03	<b>-1.38</b>	-1.09	-1.07
<b>Tc00.1047053507083.10</b>	hypothetical protein, conserved	1.13	<b>1.63</b>	<b>2.03</b>	<b>1.72</b>	<b>1.88</b>
Tc00.1047053507083.70	hypothetical protein, conserved	1.03	1.13	1.14	1.04	<b>1.64</b>

Tc00.1047053507159.30	hypothetical protein, conserved	1.09	1.12	<b>1.52</b>	1.26	1.30
<b>Tc00.1047053507165.30</b>	hypothetical protein, conserved	-1.03	1.46	1.42	1.67	<b>1.84</b>
Tc00.1047053507275.24	hypothetical protein, conserved	1.08	1.17	1.15	1.21	<b>1.31</b>
Tc00.1047053507491.140	hypothetical protein, conserved	-1.06	1.28	<b>1.85</b>	1.27	1.24
Tc00.1047053507521.60	hypothetical protein, conserved	1.00	1.27	1.30	1.21	<b>1.60</b>
Tc00.1047053507611.270	hypothetical protein, conserved	-1.03	-1.12	1.39	1.42	<b>1.82</b>
Tc00.1047053507875.30	hypothetical protein, conserved	-1.05	-1.01	<b>1.93</b>	-1.08	-1.04
<b>Tc00.1047053508059.50</b>	hypothetical protein, conserved	-1.04	1.04	<b>1.54</b>	<b>1.68</b>	<b>1.61</b>
Tc00.1047053508061.10	hypothetical protein	1.08	1.11	1.06	1.23	<b>1.54</b>
Tc00.1047053508153.364	hypothetical protein, conserved	-1.06	1.19	<b>1.80</b>	1.25	1.29
Tc00.1047053508293.130	hypothetical protein	1.18	1.10	1.08	<b>2.01</b>	<b>1.60</b>
Tc00.1047053508319.10	hypothetical protein	-1.05	1.37	1.60	<b>2.15</b>	<b>2.55</b>
Tc00.1047053508479.330	hypothetical protein, conserved	-1.01	1.38	<b>1.51</b>	<b>1.44</b>	<b>1.43</b>
Tc00.1047053508707.300	hypothetical protein, conserved	1.00	1.12	1.31	<b>1.40</b>	<b>1.52</b>
<b>Tc00.1047053508741.390</b>	hypothetical protein, conserved	1.09	1.19	1.32	1.25	<b>1.38</b>
Tc00.1047053508879.10	hypothetical protein, conserved	-1.31	-1.16	<b>1.80</b>	1.02	1.06
Tc00.1047053508909.160	hypothetical protein, conserved	-1.08	-1.02	<b>2.37</b>	1.38	1.18
Tc00.1047053509001.30	hypothetical protein, conserved	-1.01	1.25	<b>1.62</b>	<b>1.37</b>	1.25
Tc00.1047053509029.20	hypothetical protein, conserved	1.32	<b>1.98</b>	<b>2.13</b>	<b>1.79</b>	<b>2.67</b>
Tc00.1047053509073.60	hypothetical protein, conserved	1.14	1.37	1.63	<b>3.03</b>	<b>3.37</b>
Tc00.1047053509245.29	hypothetical protein, conserved	-1.21	-1.24	1.33	<b>1.66</b>	1.26
Tc00.1047053509393.10	hypothetical protein, conserved	-1.02	1.26	<b>1.58</b>	1.41	1.22
Tc00.1047053509571.30	hypothetical protein, conserved	1.18	1.25	1.36	<b>1.76</b>	<b>1.96</b>
Tc00.1047053509601.140	hypothetical protein	-1.03	1.35	<b>1.61</b>	<b>1.42</b>	<b>1.68</b>
Tc00.1047053509627.10	hypothetical protein, conserved	1.52	1.58	1.58	1.63	<b>1.90</b>
Tc00.1047053509755.89	hypothetical protein	1.02	-1.13	-1.33	1.05	<b>1.39</b>
Tc00.1047053510173.120	hypothetical protein, conserved	1.66	1.74	<b>2.16</b>	<b>2.24</b>	<b>2.19</b>
Tc00.1047053510381.40	hypothetical protein, conserved	-1.00	1.24	<b>1.61</b>	1.35	1.23
Tc00.1047053510479.30	hypothetical protein	1.15	1.02	1.05	1.30	<b>2.31</b>
Tc00.1047053510535.10	hypothetical protein, conserved	1.44	<b>2.31</b>	<b>2.31</b>	<b>2.11</b>	<b>2.56</b>
Tc00.1047053510599.40	hypothetical protein, conserved	1.19	1.22	1.26	<b>1.43</b>	<b>1.48</b>
Tc00.1047053510835.20	hypothetical protein, conserved	1.47	1.32	1.22	<b>1.95</b>	<b>2.81</b>
Tc00.1047053511121.30	hypothetical protein	1.25	1.43	1.36	<b>1.84</b>	<b>2.10</b>
Tc00.1047053511201.40	hypothetical protein, conserved	1.36	1.47	<b>1.74</b>	<b>2.10</b>	<b>2.44</b>
Tc00.1047053511239.60	hypothetical protein, conserved	1.06	1.21	1.14	1.48	<b>1.81</b>
<b>Tc00.1047053511391.110</b>	hypothetical protein, conserved	1.12	1.27	<b>1.57</b>	1.34	1.24
Tc00.1047053511421.90	hypothetical protein, conserved	1.10	1.23	<b>1.80</b>	1.43	1.18
Tc00.1047053511467.60	hypothetical protein, conserved	1.00	1.11	<b>1.91</b>	1.17	1.32
Tc00.1047053511517.130	hypothetical protein, conserved	-1.05	1.02	<b>1.43</b>	1.25	<b>1.53</b>
Tc00.1047053511523.50	hypothetical protein, conserved	1.02	1.05	1.01	1.32	<b>1.89</b>
Tc00.1047053511611.30	hypothetical protein	1.06	1.43	1.29	<b>1.91</b>	1.52
Tc00.1047053511643.80	hypothetical protein, conserved	1.17	1.08	1.03	1.21	<b>1.38</b>
Tc00.1047053511671.140	hypothetical protein	1.93	1.42	1.49	<b>2.85</b>	2.23
Tc00.1047053511727.220	hypothetical protein, conserved	1.03	1.14	1.07	<b>1.46</b>	1.16
Tc00.1047053511755.19	hypothetical protein, conserved	-1.05	1.05	<b>1.96</b>	1.52	1.65

Valores em negrito: valores estatisticamente significativos ( $p \leq 0,05$ ). IDs em negrito: genes com domínios conservados. I.a.i = logo após a irradiação

#### 4.9 – Agrupamento de genes baseado no perfil de expressão gênica ao longo do tempo

---

Um dos propósitos de abordagens do tipo *time course* é poder agrupar genes que apresentam mudanças nos níveis de expressão ao longo do tempo. Para isso, foi gerado um gráfico do tipo *heatmap* com 269 genes nucleares diferentemente expressos. Os genes mitocondriais foram excluídos dessa análise, uma vez que um dos objetivos dessa estratégia é buscar por genes que sejam co-regulados, ou que estejam em locais próximos no genoma. O *heatmap* é capaz de agrupar os genes tanto pela variação dos valores de *fold-change* (agrupamento horizontal), quanto pela comparação tempo a tempo (agrupamento vertical), agrupando os tempos mais próximos entre si. Alguns dos resultados já tinham sido evidenciados antes, como o fato de os dois primeiros tempos, por conter basicamente genes subexpressos, serem mais parecidos entre si quando comparados com os três últimos. O algoritmo foi capaz de agrupar os genes em 12 perfis diferentes de expressão (figura 21). Os valores dos *fold-changes* foram inseridos em gráficos de tendência em relação ao tempo. Como alguns agrupamentos foram formados por um número muito grande de genes, os seus nomes estão especificados no Anexo 3. A linha vermelha criada manualmente representa a média dos valores dos pontos em cada tempo (figura 21).

A análise individual de cada agrupamento mostrou que nenhum é formado por genes que fazem parte de uma mesma categoria funcional, processo regulatório ou via metabólica, exceto os genes codificadores de elementos estruturais dos ribossomos, que fazem parte do grupo 11. A análise também revelou que não há nenhuma relação entre a localização cromossômica e o perfil de expressão gênica (avaliado através do *software* Artemis), indicando a ausência de operons.

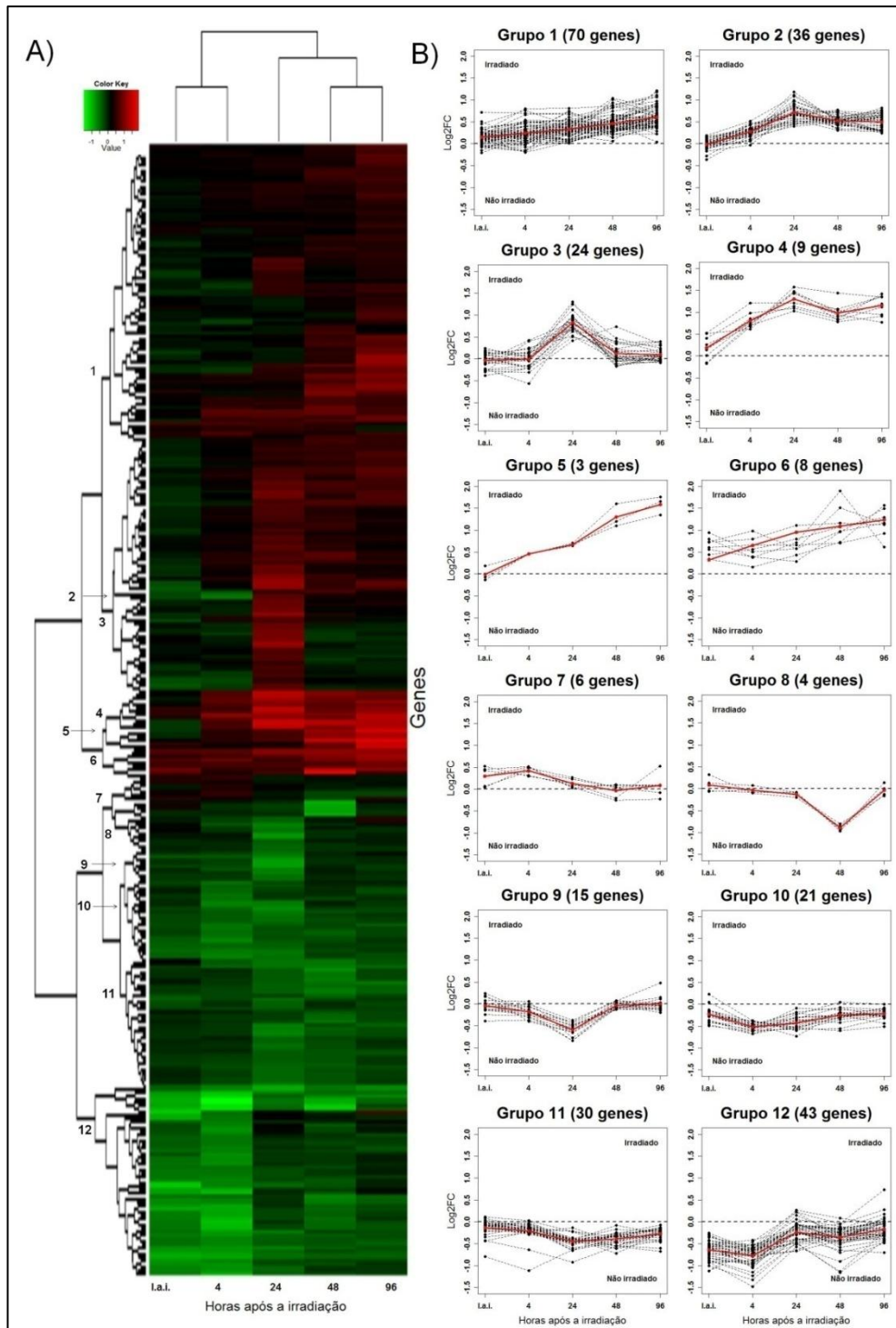


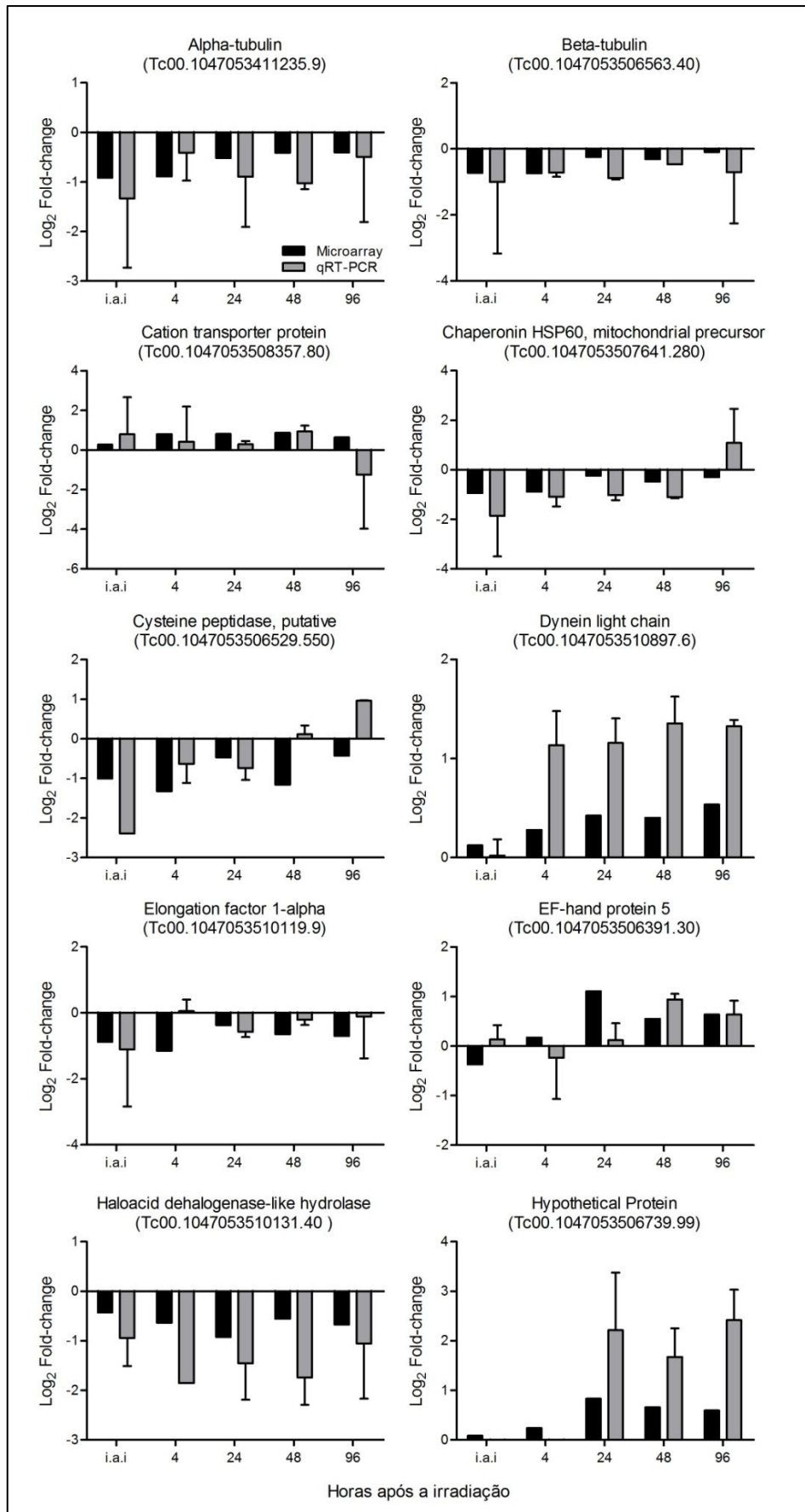
Figura 21: A) *Heatmap* de genes nucleares de *T. cruzi* diferentemente expressos. Tons verdes indicam genes subexpressos, e tons de vermelho, superexpressos. B) Agrupamentos de genes baseados nos perfis similares de expressão ao longo do tempo. Linha vermelha: média de valores de *fold-change*. A listagem completa dos genes que compuseram cada grupo encontra-se no anexo 4.

#### 4.10 – Validação dos experimentos de microarranjos por PCR quantitativa em tempo real (qRT-PCR).

---

Experimentos de microarranjos são passíveis de acúmulo de erros em todas as suas etapas, erros estes que podem prejudicar a acurácia dos resultados. Para diminuir ao máximo a chance de se obter resultados incorretos, uma série de procedimentos foi utilizada neste trabalho: aplicação de réplicas técnicas e biológicas, sendo esta última de forma independente; a utilização de uma lâmina contendo duplicatas dos oligonucleotídeos; a curadoria manual da análise das imagens (correção do *grid* e eliminação de artefatos que poderiam ser confundidos com sinal de expressão); padronização dos métodos de correção de *background* e normalização; e aplicação de testes estatísticos apropriados para os dados, conforme já explicado em itens anteriores. Ainda sim, a chance de obtermos resultados incorretos é grande, e validá-los com alguma técnica laboratorial independente é fundamental para a confiabilidade dos dados.

Dessa forma, 18 genes sendo 13 codificadores de proteínas de função conhecida, três codificadores de proteínas de função desconhecida e duas regiões genômicas que continham sequências obsoletas foram validados por qRT-PCR (figura 22). Vários genes validados por qRT-PCR serão citados nas próximas seções. Os resultados de correlação de Spearman entre os dados de microarranjo de DNA e qRT-PCR (figura 23) foram: 0,6228 ( $p=0,0022$ ) logo após a irradiação, 0,4298 ( $p=0,0331$ ) 4 horas após a irradiação, 0,7185 ( $p=0,0003$ ) 24 horas após a irradiação, 0,6789 ( $p=0,0007$ ) 48 horas após a irradiação, e 0,6088 ( $p=0,0028$ ) 96 horas após a irradiação. Os valores de eficiência de cada par de iniciador usado encontram-se na Tabela 8 (incluindo as sequências dos iniciadores).



(continua na próxima página)

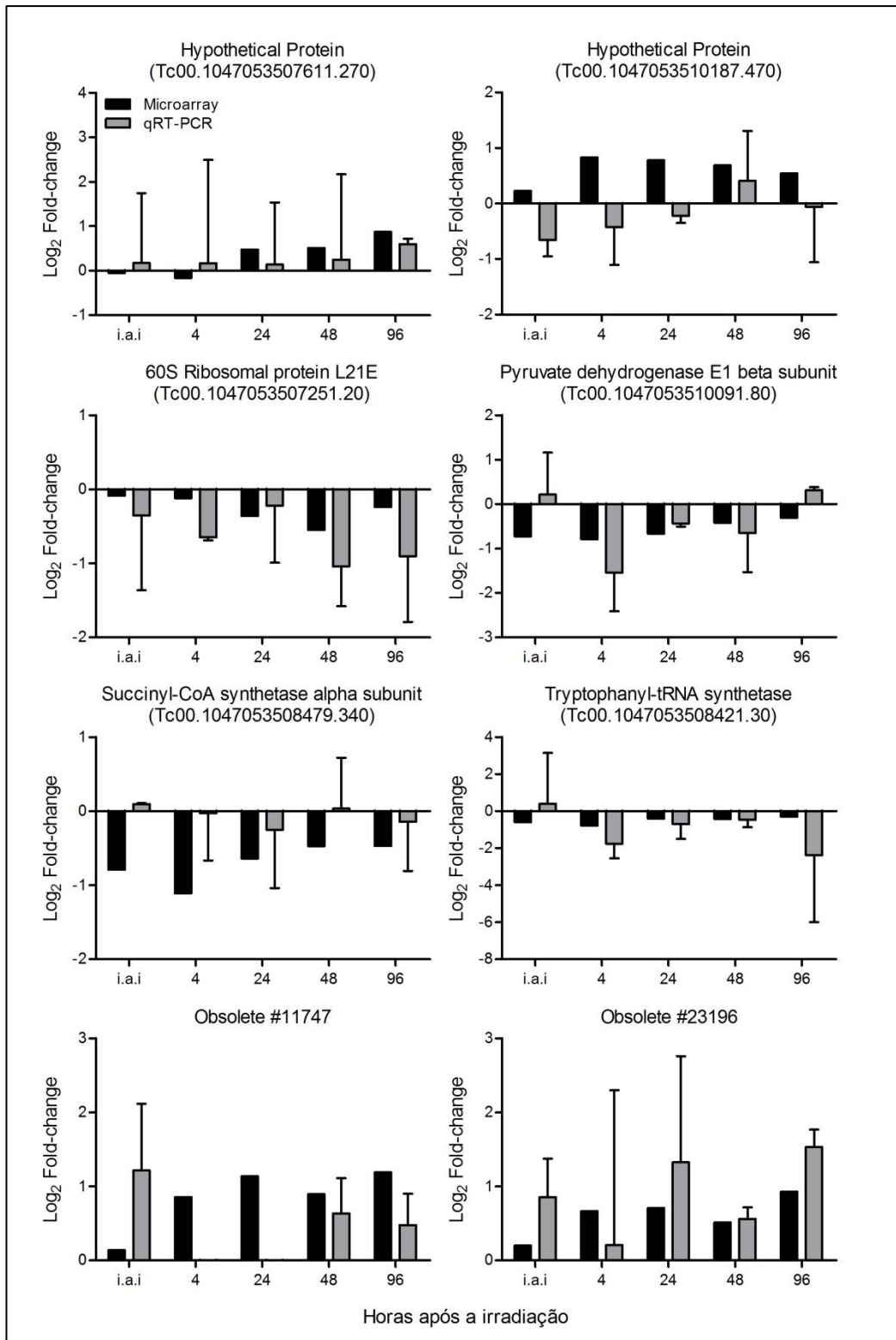


Figura 22: Comparação entre os valores de *fold-change* obtidos por microarranjo (colunas pretas) e qRT-PCR (colunas cinzas).

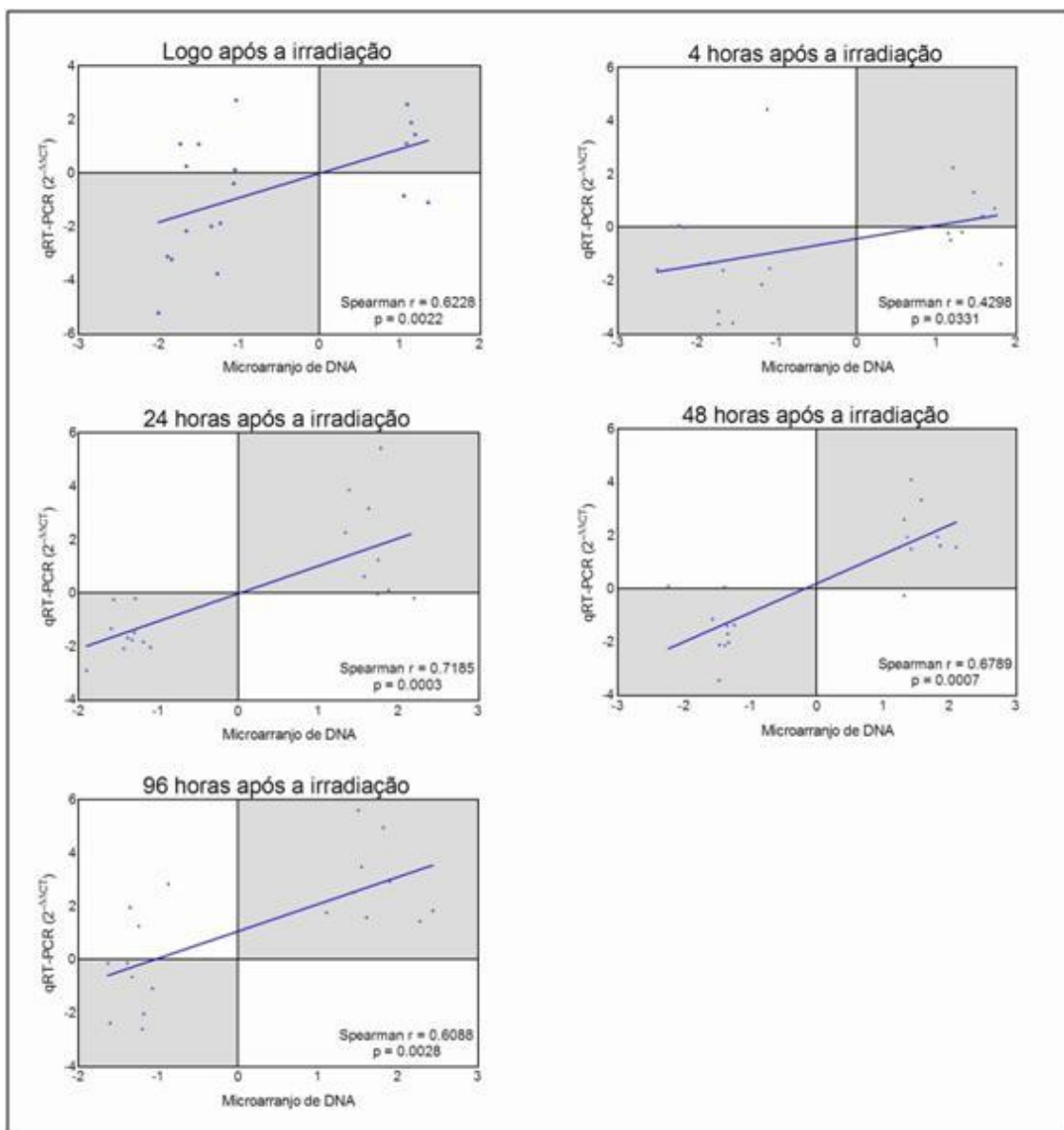


Figura 23: Correlação de Spearman entre os resultados de microarranjo e qRT-PCR. As áreas cinza são as de correlação entre as duas técnicas, onde os valores de *fold-change* obtidos por ambas as técnicas são igualmente negativos ou positivos.



Tabela 8: Sequências dos pares de iniciadores usados para a PCR quantitativa em tempo real, e os valores respectivos de slope, R<sup>2</sup> e eficiência

Nome do gene	Gene ID	Sequência (5'-3')	slope	R <sup>2</sup>	Efic.
Alpha tubulin	Tc00.1047053411235.9	F: GTTCTGTCTGGAGCACGGTA/R: CTCCGAGAAGAACGTGTGTA	-3,54	0,989	1,91
Beta tubulin	Tc00.1047053506563.40	F: AAAGAACATGATGCAGGCT/R: GTTCTGCACGTTGAGCATCT	-3,37	0,988	1,98
Cation transporter protein	Tc00.1047053508357.80	F:TGGGTGGACATCCATATGAC/R: AATGAGGCCATGCACAAATA	-4,32	0,942	1,70
Chaperonin HSP60, mitoch. precursor	Tc00.1047053507641.280	F: AGGTGTGCAACAAGACCAAC/R: CTTCATGTCAATTGGGTTCG	-3,92	0,996	1,78
Cysteine peptidase	Tc00.1047053506529.550	F: CGTGCGACGACACTAACTTT/R: GGGTAGCTTTCCTCCGTGTA	-3,26	0,984	2,02
Dynein light chain lc6	Tc00.1047053510897.6	F: GCCAAAAGACATGGAGAAC/R: GGGGCCGTA CTCTTATCAA	-3,38	0,976	1,98
EF-hand protein 5	Tc00.1047053506391.30	F: GATGCAAACGTGGATGGTAA/R :GTTGTAGTGGCATTGGCTTG	-3,41	0,969	1,96
Elongation factor 1-alpha	Tc00.1047053510119.9	F: GTCCTTGTCATTGCGTCATC/R: AGCACACAACCATCTGCTTC	-3,78	0,990	1,83
Haloacid dehalogenase-like hydrolase	Tc00.1047053510131.40	F: GAACGAGCATCACAGCATCT/R: ACATCGAAATAGGCACGAC	-3,29	0,979	2,01
Hypothetical Protein	Tc00.1047053511201.40	F: TCACAGCAACCTCAACAACA/R: CTCTTCAGGCAATCCCTTGT	-3,40	0,989	1,97
Hypothetical Protein	Tc00.1047053507611.270	F: CCGGACCAACTGGAAGTATT/R: CGGCGTGAAGTATCTCCATA	-3,62	0,981	1,89
Hypothetical Protein	Tc00.1047053506739.99	F: AGTTCTCTGAACGCCGTCTT/R: GTCTCTTCCCGGTGTTGTT	-3,41	0,990	1,96
Obsolete # 11747	No gene ID	F: TTCACCGATTTCAGCAACTC/R: AGTACCCACGGGAATCTTTG	-4,00	0,983	1,75
Obsolete # 23196	No gene ID	F: CTATGGCCGGTTACTGTTC/R: CCTGCCGTTTCTTGTCTTT	-3,37	0,986	1,98
Pyruvate dehydrogenase E1 beta subunit	Tc00.1047053510091.80	F: GTCATGGAAAGTGAGGCATT/R: GCTACGGCAAGAACATCTGA	-3,20	0,989	2,05
Ribosomal protein L21E (60S)	Tc00.1047053507251.20	F: ATCTGTGAGCACCATCCTGA/R: CAACACCACGAGGAGTAACG	-3,10	0,995	2,10
Succinyl-CoA synthetase alpha subunit	Tc00.1047053508479.340	F: ATCAAGAAGCCGGTGGTATC/R: TGTACCCGTGCCACCAGAAA	-3,53	0,993	1,92
Tryptophanyl-tRNA synthetase	Tc00.1047053508421.30	F: ACTGCTGCACACAAAGTTCC /R: CCGTGTGCTGTAGTGTGATG	-3,39	0,981	1,97

Os valores de slope devem ficar entre -3.1 and -3.6. Valores de co-eficiência R<sup>2</sup> devem estar o mais próximo possível de 1. Efic. (eficiência da reação):  $E = 10^{(-1/\text{slope})} - 1$ . Valores devem estar o mais próximo possível de 2 *fold-change*, o que significa que a eficiência da PCR está entre 90% e 110% e o amplicon dobrou de quantidade a cada ciclo.

A comparação entre os resultados dos experimentos de microarranjos e qRT-PCR com relação à significância podem ser conferidos na tabela 9. As células coloridas indicam os tempos onde houve diferença significativa na expressão gênica, sendo que o cinza escuro representa os resultados de microarranjos e cinza claro, qRT-PCR.

Tabela 9: Comparação entre os resultados significativos entre as técnicas de microarranjo e qRT-PCR

Genes	Tempos					Técnica
	l.a.i	4h	24h	48h	96h	
Alpha-tubulin						Microarranjo
						qRT-PCR
Beta-tubulin						Microarranjo
						qRT-PCR
Cation transporte protein						Microarranjo
						qRT-PCR
Chaperonin HSP60, mitochondrial precursor						Microarranjo
						qRT-PCR
Cysteine peptidase						Microarranjo
						qRT-PCR
Dynein light chain						Microarranjo
						qRT-PCR
Elongation factor 1-alpha						Microarranjo
						qRT-PCR
EF-hand protein 5						Microarranjo
						qRT-PCR
Haloacid dehalogenase-like hydrolase						Microarranjo
						qRT-PCR
Hypothetical protein #1						Microarranjo
						qRT-PCR
Hypothetical protein #2						Microarranjo
						qRT-PCR
Hypothetical protein #3						Microarranjo
						qRT-PCR
60S ribossomal protein L21E						Microarranjo
						qRT-PCR
Pyruvate dehydrogenase E1 beta subunit						Microarranjo
						qRT-PCR
Succinyl-CoA synthetase alpha subunit						Microarranjo
						qRT-PCR
Tryptophanyl-tRNA synthetase						Microarranjo
						qRT-PCR
Obsolete #11747						Microarranjo
						qRT-PCR
Obsolete #23196						Microarranjo
						qRT-PCR

Células cinza escuro = diferença significativa pelo experimento de microarray. Células cinza claro = diferença significativa pelo experimento de qRT-PCR.

#### 4.11 - Análise funcional dos genes diferentemente expressos

---

A lista de genes diferentemente expressos gerado pela técnica de microarranjos faz pouco sentido se não for traduzida de forma a melhorar a compreensão dos fenômenos biológicos observados. Com o objetivo de verificar quais processos biológicos (BP, de *biological process*) e funções moleculares (MF, de *molecular function*) estão reprimidos ou induzidos em consequência da radiação ionizante, a análise de ontologia gênica (*gene ontology*, ou GO) foi realizada. Neste trabalho, utilizamos *GO-slims*, que são subconjuntos de categorias GO gerados a partir do agrupamento de categorias muito específicas e produzindo categorias mais genéricas. A razão disso é que, segundo levantamento feito em abril de 2010, o GO continha 18.903 termos para processos biológicos, 8.713 termos para funções moleculares, e 2.734 termos para componentes celulares. O enorme número de termos torna muito difícil a comparação e análise dos dados anotados com os termos GO, e dificulta a seleção dos termos mais interessantes para serem anotados pelo pesquisador (Davis *et al.*, 2010). Assim, a anotação em termos GO-slim foca nos termos mais relevantes para um dado problema biológico ou conjunto de dados. Foi verificado, inclusive, que o uso de GO-slim torna a anotação mais robusta e os dados tornam-se mais apropriados para a aplicação de testes estatísticos (Rhee *et al.*, 2008). Esse procedimento foi essencial para nos dar uma visão panorâmica dos efeitos da radiação gama nas células epimastigotas. Os resultados estão compilados na figura 24. A listagem dos genes que compõem cada processo biológico ou função molecular indicados na figura 24 encontra-se na tabela 10.

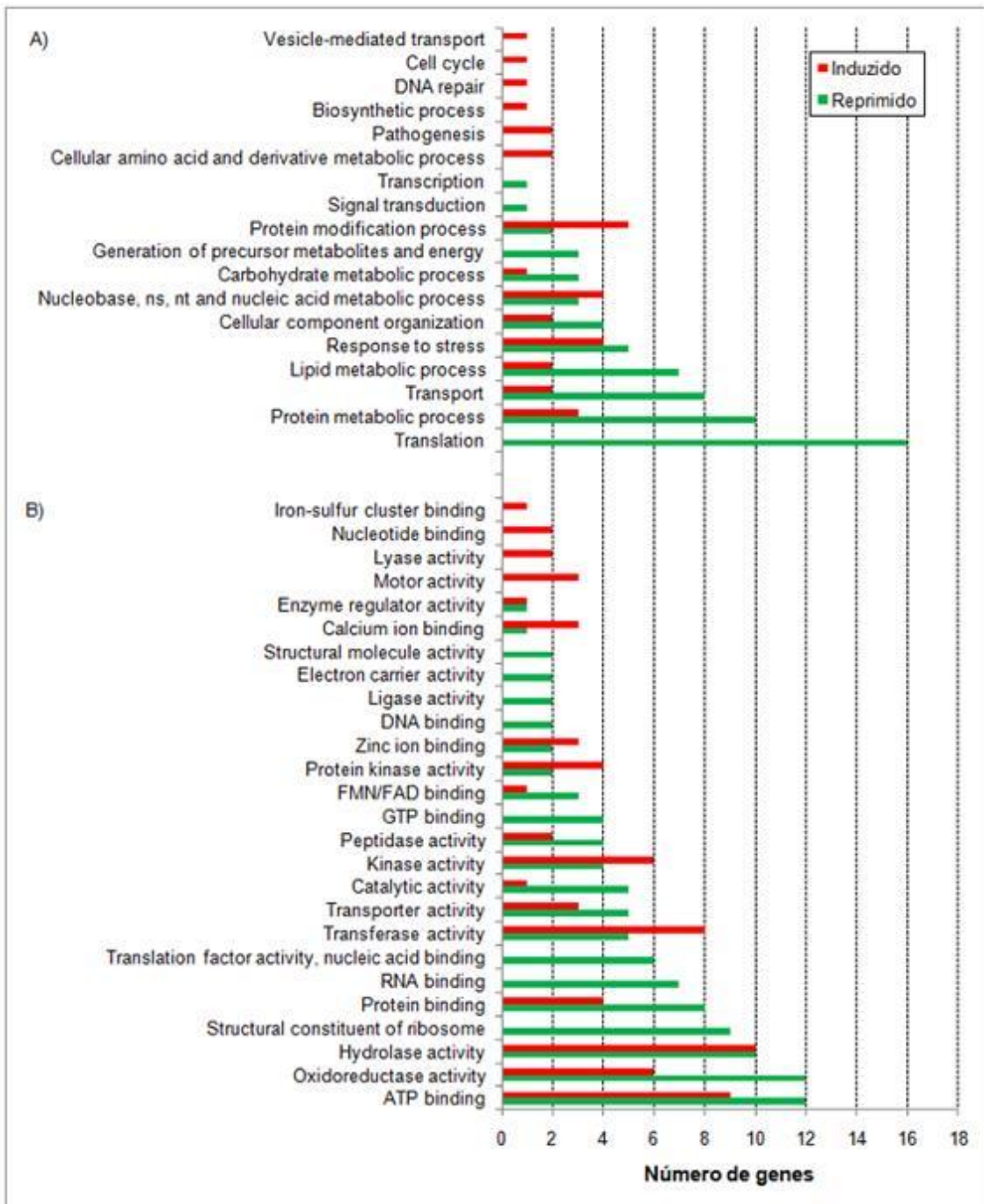


Figura 24: Categorias funcionais Slim de ontologia gênica dos genes diferentemente expressos. A) Processos Biológicos (BP) e B) Função Molecular (MF). Barras verdes: categorias BP e MF reprimidas após a irradiação. Barras vermelhas: categorias BP e MF induzidas após a irradiação.

Tabela 10: Termos GO-Slim para as categorias de processos biológicos e função molecular; e os respectivos genes

ID	Categoria <i>GO-Slim</i>	Grupo	Identificador	Produto
<b>A) Processos Biológicos – Genes subexpressos</b>				
GO:0005975	Carbohydrate metabolic process	12	Tc00.1047053510091.80	pyruvate dehydrogenase E1 beta subunit, putative
		12	Tc00.1047053511041.40	hexose transporter, putative
		12	Tc00.1047053507547.90	glycosomal phosphoenolpyruvate carboxykinase, putative
GO:0006091	Generation of precursor metabolites and energy	12	Tc00.1047053503929.10	V-type ATPase, A subunit, putative
		11	Tc00.1047053508479.340	succinyl-CoA synthetase alpha subunit, putative
		10	Tc00.1047053509769.60	hypothetical protein, conserved (2-iron 2-sulfer protein, component of succinate dehydrogenase complex)
GO:0006139	Nucleobase, nucleoside, nucleotide and nucleic acid metabolic process	12	Tc00.1047053506773.50	nucleoside transporter-like, putative
		11	Tc00.1047053508707.200	nucleoside diphosphate kinase, putative
		12	Tc00.1047053506855.260	thymidine kinase, putative
GO:0006350	Transcription	10	Tc00.1047053510645.20	nuclear transcription factor, putative
GO:0006412	Translation	12	Tc00.1047053507671.30	25 kDa translation elongation factor 1-beta
		11	Tc00.1047053510101.430	40S ribosomal protein S21, putative
		11	Tc00.1047053509353.30	40S ribosomal protein S3, putative
		11	Tc00.1047053506297.150	40S ribosomal protein S5, putative
		11	Tc00.1047053510425.19	40S ribosomal protein SA, putative
		11	Tc00.1047053503719.20	40S ribosomal protein SA, putative
		11	Tc00.1047053505977.26	60S acidic ribosomal protein P2, putative
		11	Tc00.1047053508461.490	60S ribosomal protein L23, putative
		11	Tc00.1047053506297.270	60S ribosomal protein L28, putative
		12	Tc00.1047053510119.9	elongation factor 1-alpha (ef-1-alpha), putative
		10	Tc00.1047053508153.730	elongation initiation factor 2 alpha subunit, putative
		10	Tc00.1047053463269.10	eukaryotic translation initiation factor 1A, putative
		9	Tc00.1047053506943.160	eukaryotic translation initiation factor 3 subunit 7-like protein, putative
		12	Tc00.1047053506679.70	eukaryotic translation initiation factor 6 (eIF-6), putative
		11	Tc00.1047053507251.20	ribosomal protein L21E (60S), putative
12	Tc00.1047053508421.30	tryptophanyl-tRNA synthetase (pseudogene), putative		
GO:0006464	Protein modification process	9	Tc00.1047053503487.50	UDP-Gal or UDP-GlcNAc-dependent glycosyltransferase, putative
		12	Tc00.1047053507509.40	hypothetical protein, conserved (kinase protein)
		9	Tc00.1047053504427.70	3-oxo-5-alpha-steroid 4-dehydrogenase, putative
GO:0006629	Lipid metabolic process	12	Tc00.1047053511823.70	Farnesyl diphosphate synthase
		12	Tc00.1047053511075.9	fatty acid desaturase, putative
		10	Tc00.1047053436521.9	mevalonate kinase, putative
		12	Tc00.1047053507617.9	prostaglandin F2alpha synthase
		12	Tc00.1047053508461.80	prostaglandin F2alpha synthase

		10	Tc00.1047053511071.50	hypothetical protein, conserved (pseudogene) (HMG-CoA synthase)
		10	Tc00.1047053510187.270	COP-coated vesicle membrane protein erv25 precursor, putative,ER--golgi transport protein erv25 precursor, putative
		12	Tc00.1047053503539.30	GTP-binding nuclear protein rtb2, putative
		12	Tc00.1047053511041.40	hexose transporter, putative
GO:0006810	Transport	9	Tc00.1047053510667.14	membrane transporter protein, putative
		10	Tc00.1047053508173.180	nuclear transport factor 2 protein(NFT2), putative
		12	Tc00.1047053506773.50	nucleoside transporter-like, putative
		12	Tc00.1047053503929.10	V-type ATPase, A subunit, putative
		11	Tc00.1047053507993.380	hypothetical protein, conserved (signal recognition particle component)
		12	Tc00.1047053507641.280	chaperonin Hsp60, mitochondrial precursor,groELprotein,heat shock protein 60
GO:0006950	Response to stress	11	Tc00.1047053510187.420	chaperonin HSP60, mitochondrial precursor (pseudogene),groELprotein
		11	Tc00.1047053508173.100	Monoxygenase, putative
		10	Tc00.1047053503555.30	trypanothione reductase, putative
		11	Tc00.1047053511745.10	heat shock 70 kDa protein, mitochondrial precursor, putative
GO:0007165	Signal transduction	12	Tc00.1047053503539.30	GTP-binding nuclear protein rtb2, putative
		12	Tc00.1047053411235.9	alpha tubulin, putative
GO:0016043	Cellular component organization	12	Tc00.1047053506563.40	beta tubulin, putative
		12	Tc00.1047053506679.70	eukaryotic translation initiation factor 6 (eIF-6), putative
		10	Tc00.1047053507943.40	histone H4, putative
		11	Tc00.1047053504221.20	26S proteasome regulatory non-ATPase subunit, putative
		12	Tc00.1047053510655.120	aminopeptidase P, putative,metallo-peptidase, Clan MG, Family M24, putative
		12	Tc00.1047053504153.160	carboxypeptidase, putative,metallo-peptidase, clan MA(E), family M32, putative
		12	Tc00.1047053507641.280	chaperonin Hsp60, mitochondrial precursor,groELprotein,heat shock protein 60
GO:0019538	Protein metabolic process	11	Tc00.1047053510187.420	chaperonin HSP60, mitochondrial precursor (pseudogene),groELprotein, degenerate,heat shock protein 60, degenerate,chaperonin HSP60, mitochondrial precursor, degenerate
		12	Tc00.1047053506247.50	chaperonin, putative,T-complex protein 1, theta subunit, putative
		12	Tc00.1047053506529.550	cysteine peptidase, putative,cysteine peptidase, clan CA, family C1, cathepsin L-like, putative
		12	Tc00.1047053508317.10	cysteine proteinase, putative
		11	Tc00.1047053511745.10	heat shock 70 kDa protein, mitochondrial precursor, putative
		10	Tc00.1047053509267.40	hypothetical protein, conserved (ARM motif)
<b>B) Processos Biológicos – Genes superexpressos</b>				
GO:0005975	Carbohydrate metabolic process	4	Tc00.1047053506529.508	glucose-6-phosphate isomerase, glycosomal, putative
		3	Tc00.1047053508375.30	aspartate carbamoyltransferase, putative
GO:0006139	Nucleobase, nucleoside, nucleotide and nucleic acid metabolic process	1	Tc00.1047053508461.400	nucleoside diphosphate kinase, putative
		1	Tc00.1047053507611.290	acetyltransferase, putative
		4	Tc00.1047053507083.10	hypothetical protein, conserved
GO:0006281	DNA repair	3	Tc00.1047053506619.40	tyrosyl-DNA Phosphodiesterase (Tdp1), putative
GO:0006464	Protein modification process	1	Tc00.1047053510121.130	serine/threonine protein kinase, putative
		1	Tc00.1047053507611.290	acetyltransferase, putative

		1	Tc00.1047053511269.50	protein kinase A catalytic subunit, putative
		1	Tc00.1047053504113.10	protein kinase, putative
		1	Tc00.1047053511671.80	protein kinase, putative
GO:0006519	cellular amino acid and derivative metabolic process	3	Tc00.1047053508375.30	aspartate carbamoyltransferase, putative
		1	Tc00.1047053508457.30	glycine cleavage system H protein, putative
GO:0006629	lipid metabolic process	4	Tc00.1047053473111.10	lathosterol oxidase, putative
		1	Tc00.1047053508741.390	hypothetical protein, conserved (Phospholipase C-like phosphodiesterase)
GO:0006810	transport	1	Tc00.1047053508357.80	cation transporter protein, putative
		1	Tc00.1047053506947.50	hypothetical protein, conserved (major facilitator superfamily conserved domain)
		3	Tc00.1047053506619.40	tyrosyl-DNA Phosphodiesterase (Tdp1), putative
GO:0006950	response to stress	2	Tc00.1047053508649.5	tryparedoxin peroxidase, putative
		3	Tc00.1047053507929.20	co-chaperone GrpE, putative
		7	Tc00.1047053503899.119	trypanothione/tryparedoxin dependent peroxidase 2, putative
GO:0007049	cell cycle	1	Tc00.1047053509455.140	cyclin, putative
GO:0009405	pathogenesis	1	Tc00.1047053510199.10	GP85-like protein, putative
		1	Tc00.1047053509907.60	trans-sialidase (pseudogene), putative
GO:0009058	biosynthetic process	1	Tc00.1047053507083.70	hypothetical protein, conserved
GO:0016043	cellular component organization	2	Tc00.1047053508737.194	ARP2/3 complex subunit, putative
		4	Tc00.1047053509069.30	tubulin binding cofactor A-like protein, putative
		3	Tc00.1047053508183.4	aspartyl aminopeptidase, putative
GO:0019538	protein metabolic process	2	Tc00.1047053506563.210	calpain-like cysteine peptidase, putative,cysteine peptidase, Clan CA, family C2, putative
		3	Tc00.1047053507929.20	co-chaperone GrpE, putative
GO:0016193	vesicle-mediated transport	3	Tc00.1047053508479.290	vacuolar sorting protein, putative

### C) Função Molecular – Genes subexpressos

		12	Tc00.1047053507641.280	chaperonin Hsp60, mitochondrial precursor,groELprotein,heat shock protein 60
		11	Tc00.1047053510187.420	chaperonin HSP60, mitochondrial precursor (pseudogene),groELprotein, degenerate,heat shock protein 60, degenerate,chaperonin HSP60, mitochondrial precursor, degenerate
		12	Tc00.1047053506247.50	chaperonin, putative,T-complex protein 1, theta subunit, putative
		12	Tc00.1047053507547.90	glycosomal phosphoenolpyruvate carboxykinase, putative
GO:0005524	ATP binding	10	Tc00.1047053436521.9	mevalonate kinase, putative
		10	Tc00.1047053510155.20	mitochondrial RNA editing ligase 1, putative
		11	Tc00.1047053508707.200	nucleoside diphosphate kinase, putative
		12	Tc00.1047053506855.260	thymidine kinase, putative
		12	Tc00.1047053508421.30	tryptophanyl-tRNA synthetase (pseudogene), putative
		12	Tc00.1047053503929.10	V-type ATPase, A subunit, putative
		11	Tc00.1047053511745.10	heat shock 70 kDa protein, mitochondrial precursor, putative
		12	Tc00.1047053507509.40	hypothetical protein, conserved (kinase protein)
		12	Tc00.1047053411235.9	alpha tubulin, putative
GO:0005526	GTP binding	12	Tc00.1047053510119.9	elongation factor 1-alpha (ef-1-alpha), putative
		12	Tc00.1047053503539.30	GTP-binding nuclear protein rtb2, putative

		12	Tc00.1047053506563.40	beta tubulin, putative
GO:0010181	FMN/FAD binding	12	Tc00.1047053507617.9	prostaglandin F2alpha synthase
		10	Tc00.1047053503555.30	trypanothione reductase, putative
		12	Tc00.1047053508461.80	prostaglandin F2alpha synthase
		10	Tc00.1047053507943.40	histone H4, putative
GO:0003677	DNA binding	9	Tc00.1047053509793.10	kinetoplast DNA-associated protein, putative
GO:0008270	Zinc ion binding	10	Tc00.1047053510645.20	nuclear transcription factor, putative
		12	Tc00.1047053509669.40	Zn-finger protein, putative
GO:0003723	RNA binding	12	Tc00.1047053510859.17	nucleolar RNA-binding protein, putative
		11	Tc00.1047053509353.30	40S ribosomal protein S3, putative
		11	Tc00.1047053506297.150	40S ribosomal protein S5, putative
		10	Tc00.1047053508153.730	elongation initiation factor 2 alpha subunit, putative
		10	Tc00.1047053463269.10	eukaryotic translation initiation factor 1A, putative
		9	Tc00.1047053511715.100	pumilio/PUF RNA binding protein 7, putative
		11	Tc00.1047053507993.380	hypothetical protein, conserved (signal recognition particle component)
GO:0003824	Catalytic activity	11	Tc00.1047053510099.120	D-isomer specific 2-hydroxyacid dehydrogenase-protein, putative
		11	Tc00.1047053510131.40	haloacid dehalogenase-like hydrolase, putative
		9	Tc00.1047053511355.30	phosphatidic acid phosphatase protein, putative
		11	Tc00.1047053508479.340	succinyl-CoA synthetase alpha subunit, putative
		10	Tc00.1047053511071.50	hypothetical protein, conserved (pseudogene) (HMG-CoA synthase)
GO:0016491	Oxidoreductase activity	9	Tc00.1047053504427.70	3-oxo-5-alpha-steroid 4-dehydrogenase, putative
		12	Tc00.1047053511391.160	cytochrome c1, heme protein, mitochondrial precursor, putative
		11	Tc00.1047053510099.120	D-isomer specific 2-hydroxyacid dehydrogenase-protein, putative
		12	Tc00.1047053511075.9	fatty acid desaturase, putative
		11	Tc00.1047053508173.100	Monooxygenase, putative
		12	Tc00.1047053511817.40	NADH-cytochrome B5 reductase, putative
		12	Tc00.1047053507617.9	prostaglandin F2alpha synthase
		12	Tc00.1047053508461.80	prostaglandin F2alpha synthase
		12	Tc00.1047053510091.80	pyruvate dehydrogenase E1 beta subunit, putative
		10	Tc00.1047053503555.30	trypanothione reductase, putative
		10	Tc00.1047053509769.60	hypothetical protein, conserved (2-iron 2-sulfur protein, component of succinate dehydrogenase complex)
GO:0016874	ligase activity	12	Tc00.1047053506219.40	hypothetical protein, conserved (short chain dehydrogenase)
		10	Tc00.1047053510155.20	mitochondrial RNA editing ligase 1, putative
		12	Tc00.1047053508421.30	tryptophanyl-tRNA synthetase (pseudogene), putative
GO:0004672	protein kinase activity	12	Tc00.1047053507509.40	hypothetical protein, conserved (kinase protein)
		11	Tc00.1047053511211.120	activated protein kinase C receptor, putative, guanine nucleotide-binding protein beta subunit-like protein, putative
GO:0003735	structural constituent of ribosome	11	Tc00.1047053510101.430	40S ribosomal protein S21, putative
		11	Tc00.1047053509353.30	40S ribosomal protein S3, putative



		11	Tc00.1047053506297.150	40S ribosomal protein S5, putative
		11	Tc00.1047053510425.19	40S ribosomal protein SA, putative
		11	Tc00.1047053503719.20	40S ribosomal protein SA, putative
		11	Tc00.1047053505977.26	60S acidic ribosomal protein P2, putative
		11	Tc00.1047053507251.20	60S ribosomal protein L21E, putative
		11	Tc00.1047053508461.490	60S ribosomal protein L23, putative
		11	Tc00.1047053506297.270	60S ribosomal protein L28, putative
GO:0005198	Structural molecule activity	12	Tc00.1047053411235.9	alpha tubulin, putative
		12	Tc00.1047053506563.40	beta tubulin, putative
GO:0005215	transporter activity	12	Tc00.1047053511391.160	cytochrome c1, heme protein, mitochondrial precursor, putative
		12	Tc00.1047053511041.40	hexose transporter, putative
		9	Tc00.1047053510667.14	membrane transporter protein, putative
		12	Tc00.1047053506773.50	nucleoside transporter-like, putative
		12	Tc00.1047053503929.10	V-type ATPase, A subunit, putative
GO:0005509	calcium ion binding	8	Tc00.1047053507891.47	flagellar calcium-binding protein, putative
GO:0005515	protein binding	12	Tc00.1047053507641.280	chaperonin Hsp60, mitochondrial precursor, groEL protein, heat shock protein 60
		11	Tc00.1047053507873.20	cell differentiation protein, putative
		11	Tc00.1047053510187.420	chaperonin HSP60, mitochondrial precursor (pseudogene), groEL protein, degenerate, heat shock protein 60, degenerate, chaperonin HSP60, mitochondrial precursor, degenerate
		12	Tc00.1047053506247.50	chaperonin, putative, T-complex protein 1, theta subunit, putative
		12	Tc00.1047053503539.30	GTP-binding nuclear protein rtb2, putative
		11	Tc00.1047053511745.10	heat shock 70 kDa protein, mitochondrial precursor, putative
		10	Tc00.1047053509267.40	hypothetical protein, conserved (ARM motif)
GO:0008135	translation factor activity, nucleic acid binding	12	Tc00.1047053507671.30	25 kDa translation elongation factor 1-beta
		12	Tc00.1047053510119.9	elongation factor 1-alpha (ef-1-alpha), putative
		10	Tc00.1047053508153.730	elongation initiation factor 2 alpha subunit, putative
		10	Tc00.1047053463269.10	eukaryotic translation initiation factor 1A, putative
		9	Tc00.1047053506943.160	eukaryotic translation initiation factor 3 subunit 7-like protein, putative
		12	Tc00.1047053506679.70	eukaryotic translation initiation factor 6 (eIF-6), putative
GO:0008233	peptidase activity	12	Tc00.1047053510655.120	aminopeptidase P, putative, metallo-peptidase, Clan MG, Family M24, putative
		12	Tc00.1047053504153.160	carboxypeptidase, putative, metallo-peptidase, clan MA(E), family M32, putative
		12	Tc00.1047053506529.550	cysteine peptidase, putative, cysteine peptidase, clan CA, family C1, cathepsin L-like, putative
		12	Tc00.1047053508317.10	cysteine proteinase, putative
GO:0009055	electron carrier activity	12	Tc00.1047053511391.160	cytochrome c1, heme protein, mitochondrial precursor, putative
		10	Tc00.1047053509769.60	hypothetical protein, conserved (2-iron 2-sulfur protein, component of succinate dehydrogenase complex)
GO:0016301	kinase activity	12	Tc00.1047053507547.90	glycosomal phosphoenolpyruvate carboxykinase, putative
		10	Tc00.1047053436521.9	mevalonate kinase, putative
		11	Tc00.1047053508707.200	nucleoside diphosphate kinase, putative
		12	Tc00.1047053506855.260	thymidine kinase, putative

GO:0016740	transferase activity	12	Tc00.1047053509693.100	2-aminoethylphosphonate:pyruvateaminotransferase- likeprotein,putative
		12	Tc00.1047053511823.70	Farnesyl diphosphate synthase
		9	Tc00.1047053503487.50	UDP-Gal or UDP-GlcNAc-dependent glycosyltransferase, putative
		12	Tc00.1047053507509.40	hypothetical protein, conserved (kinase protein)
		10	Tc00.1047053511071.50	hypothetical protein, conserved (pseudogene) (HMG-CoA synthase)
GO:0016787	hydrolase activity	12	Tc00.1047053411235.9	alpha tubulin, putative
		12	Tc00.1047053510655.120	aminopeptidase P, putative,metallo-peptidase, Clan MG, Family M24, putative
		12	Tc00.1047053506563.40	beta tubulin, putative
		12	Tc00.1047053504153.160	carboxypeptidase, putative,metallo-peptidase, clan MA(E), family M32, putative
		12	Tc00.1047053506529.550	cysteine peptidase, putative,cysteine peptidase, clan CA, family C1, cathepsin L-like, putative
		12	Tc00.1047053508317.10	cysteine proteinase, putative
		12	Tc00.1047053510119.9	elongation factor 1-alpha (ef-1-alpha), putative
		12	Tc00.1047053503539.30	GTP-binding nuclear protein rtb2, putative
		11	Tc00.1047053510131.40	haloacid dehalogenase-like hydrolase, putative
GO:0030234	enzyme regulator activity	12	Tc00.1047053503929.10	V-type ATPase, A subunit, putative
		11	Tc00.1047053504221.20	26S proteasome regulatory non-ATPase subunit, putative
<b>D) Função Molecular – Genes superexpressos</b>				
GO:0010181/ GO:0050660	FMN/FAD binding	2	Tc00.1047053509941.100	2,4-dienoyl-coa reductase fadh1, putative
GO:0005524	ATP binding	7	Tc00.1047053504149.20	ATP-binding cassette transporter ABCA1, putative
		2	Tc00.1047053511421.110	developmentally regulated phosphoprotein, putative
		1	Tc00.1047053510687.10	dynein heavy chain (pseudogene), putative
		1	Tc00.1047053509585.10	dynein heavy chain, putative
		1	Tc00.1047053508461.400	nucleoside diphosphate kinase, putative
		1	Tc00.1047053510121.130	serine/threonine protein kinase, putative
		1	Tc00.1047053511269.50	protein kinase A catalytic subunit, putative
		1	Tc00.1047053504113.10	protein kinase, putative
		1	Tc00.1047053511671.80	protein kinase, putative
GO:0000166	Nucleotide binding	1	Tc00.1047053510105.100	UDP-glucose dehydrogenase, putative
GO:0003774	Motor activity	3	Tc00.1047053507929.20	co-chaperone GrpE, putative
		1	Tc00.1047053510687.10	dynein heavy chain (pseudogene), putative
		1	Tc00.1047053509585.10	dynein heavy chain, putative
GO:0016491	Oxidoreductase activity	1	Tc00.1047053510897.6	Dynein light chain LC6, flagellar outer arm, putative
		2	Tc00.1047053509941.100	2,4-dienoyl-coa reductase fadh1, putative
		4	Tc00.1047053473111.10	lathosterol oxidase, putative
		1	Tc00.1047053510105.100	UDP-glucose dehydrogenase, putative
		7	Tc00.1047053503899.119	trypanothione/tryparedoxin dependent peroxidase 2, putative
GO:0016829	Lyase activity	2	Tc00.1047053508649.5	tryparedoxin peroxidase, putative
		3	Tc00.1047053506357.50	alcohol dehydrogenase, putative
		2	Tc00.1047053510659.240	lactoylglutathione lyase-like protein, putative

		3	Tc00.1047053510743.70	lactoylglutathione lyase-like protein, putative
GO:0003824	Catalytic activity	3	Tc00.1047053503841.20	hypothetical protein, conserved (CBS conserved domain)
		1	Tc00.1047053510121.130	serine/threonine protein kinase, putative
GO:0004672	Protein kinase activity	1	Tc00.1047053511269.50	protein kinase A catalytic subunit, putative
		1	Tc00.1047053504113.10	protein kinase, putative
		1	Tc00.1047053511671.80	protein kinase, putative
		1	Tc00.1047053508357.80	cation transporter protein, putative
GO:0005215	Transporter activity	3	Tc00.1047053508479.290	vacuolar sorting protein, putative
		1	Tc00.1047053506947.50	hypothetical protein, conserved (major facilitator superfamily conserved domain)
GO:0051536	Iron-sulfur cluster binding	1	Tc00.1047053507083.70	hypothetical protein, conserved
		3	Tc00.1047053508183.4	aspartyl aminopeptidase, putative
GO:0008270	Zinc ion binding	2	Tc00.1047053506739.99	hypothetical protein, conserved
		3	Tc00.1047053503839.19	hypothetical protein, conserved (B-box zinc-finger conserved domain)
		2	Tc00.1047053506391.30	ef-hand protein 5, putative
GO:0005509	Calcium ion binding	2	Tc00.1047053507483.20	ef-hand protein 5, putative
		1	Tc00.1047053507165.30	hypothetical protein, conserved (EF-hand conserved domain)
		4	Tc00.1047053509069.30	tubulin binding cofactor A-like protein, putative
GO:0005515	Protein binding	2	Tc00.1047053508215.9	protein kinase C substrate protein, heavy chain, putative
		3	Tc00.1047053507929.20	co-chaperone GrpE, putative
		2	Tc00.1047053508059.50	hypothetical protein, conserved
GO:0008233	Peptidase activity	3	Tc00.1047053508183.4	aspartyl aminopeptidase, putative
		2	Tc00.1047053506563.210	calpain-like cysteine peptidase, putative,cysteine peptidase, Clan CA, family C2, putative
		1	Tc00.1047053508461.400	nucleoside diphosphate kinase, putative
		1	Tc00.1047053510121.130	serine/threonine protein kinase, putative
GO:0016301	Kinase activity	1	Tc00.1047053511269.50	protein kinase A catalytic subunit, putative
		2	Tc00.1047053508215.9	protein kinase C substrate protein, heavy chain, putative
		1	Tc00.1047053504113.10	protein kinase, putative
		1	Tc00.1047053511671.80	protein kinase, putative
		3	Tc00.1047053508375.30	aspartate carbamoyltransferase, putative
		1	Tc00.1047053508457.30	glycine cleavage system H protein, putative
GO:0016740	Transferase activity	1	Tc00.1047053508461.400	nucleoside diphosphate kinase, putative
		1	Tc00.1047053511269.50	protein kinase A catalytic subunit, putative
		1	Tc00.1047053504113.10	protein kinase, putative
		1	Tc00.1047053511671.80	protein kinase, putative
		1	Tc00.1047053507611.290	acetyltransferase, putative
		2	Tc00.1047053511391.110	hypothetical protein, conserved (histone-lysine N-methyltransferase, putative)
GO:0016787	Hydrolase activity	3	Tc00.1047053508183.4	aspartyl aminopeptidase, putative
		7	Tc00.1047053504149.20	ATP-binding cassette transporter ABCA1, putative
		2	Tc00.1047053506563.210	calpain-like cysteine peptidase, putative,cysteine peptidase, Clan CA, family C2, putative
		1	Tc00.1047053510687.10	dynein heavy chain (pseudogene), putative

		1	Tc00.1047053509585.10	dynein heavy chain, putative
		1	Tc00.1047053510199.10	GP85-like protein, putative
		3	Tc00.1047053506619.40	tyrosyl-DNA Phosphodiesterase (Tdp1), putative
		1	Tc00.1047053509907.60	trans-sialidase (pseudogene), putative
		7	Tc00.1047053506341.10	N-acetylglucosamine-6-phosphate deacetylase-like protein, putative
		1	Tc00.1047053508741.390	hypothetical protein, conserved ( (Phospholipase C-like phosphodiesterase)
GO:0030234	Enzyme regulator activity	3	Tc00.1047053507929.20	co-chaperone GrpE, putative

A e C: genes subexpressos. B e D: genes superexpressos. A coluna “grupo” indica o número do grupo ilustrado na figura 21B.

#### 4.11.1 – Efeitos da irradiação na expressão dos genes envolvidos na atividade motora e ciclo celular

A análise da curva de crescimento do *T. cruzi* após a irradiação revelou que as células de ambas as réplicas biológicas sobreviveram, mas não se dividiram durante o período avaliado nesse trabalho. A falta de crescimento celular pode ser parcialmente justificada pela subexpressão de alfa e beta-tubulinas (*cellular component organization* – BP e *GTP binding* - MF), cujos valores de *fold-change* foram significativos nos dois primeiros tempos, ou seja, logo após e 4 horas após a irradiação. Essas proteínas fazem parte da estrutura dos microtúbulos, que por sua vez são componentes do citoesqueleto, que participa, dentre vários processos, da mitose e citocinese. Esses resultados foram confirmados por qRT-PCR (Figura 22). Inclusive, há evidências de que alfa-tubulina continuou significativamente subexpressa até 48 horas após a irradiação, e beta-tubulina, além dos dois primeiros tempos, estava subexpressa no tempo de 96 horas.

Apesar do dano físico sofrido pelas células, elas ficavam bastante ativas e se movimentavam mais que as células controle durante a contagem no microscópio óptico (constatação visual). Surpreendentemente, três genes associados com atividade motora de microtúbulos foram superexpressos (*motor activity* – MF). As dineínas de cadeia leve (do tipo lc6) e pesada (do tipo 2) são componentes estruturais dos flagelos (Asai e Wilkes, 2004; Dibella *et al.*, 2005), responsáveis pelos movimentos flagelares. Os níveis de expressão da dineína de cadeia leve (lc6) foi avaliada por qRT-PCR, e os resultados foram concordantes (Figura 22).

#### 4.11.2 – Efeito da irradiação na expressão dos genes envolvidos com a síntese e processamento de proteínas

O primeiro processo biológico alterado que se destacou foi o de tradução protéica (termo GO-Slim: *translation*). Há no total 16 genes subexpressos (em pelo menos um tempo analisado) e nenhum superexpresso, o que indica que células irradiadas podem não estar traduzindo proteínas, ou esse processo está consideravelmente prejudicado. Mesmo com *fold-change* não significativo, a tendência foi de subexpressão nos tempos restantes (tabela 1). Em termos de funções moleculares associadas a este processo, podemos citar as atividades de fator de tradução e ligação de ácidos nucléicos (termo GO-Slim: *translation factor activity, nucleic acid binding*) com seis genes subexpressos, a atividade de ligação de RNA (termo GO-Slim: *RNA binding*), com sete genes subexpressos, e elementos estruturais ribossomais (termo GO-Slim: *structural constituent of ribosome*), com nove genes subexpressos. Em todas essas categorias acima citadas, também não há nenhum gene superexpresso. O teste qui-quadrado foi realizado e mostrou que a redução da tradução protéica está associada com os efeitos da radiação ( $p = 0,0073$ ). Os níveis de expressão gênica dos genes fator de alongação *1-alpha* e proteína ribossomal 60S L21E foram avaliados por qRT-PCR (figura 22).

A categoria de processos metabólicos protéicos (*protein metabolic process - BP*) também foi formada principalmente por genes subexpressos, indicando que as proteínas podem não estar sendo enoveladas corretamente e não estar sofrendo proteólise, sendo que esse último processo pode estar diminuído exatamente porque a síntese está comprometida. Como alguns exemplos de genes subexpressos desta categoria, podemos citar as proteínas *heat-shock* hsp70 e hsp60 (validada por qRT-PCR), complexo T da proteína chaperonina, aminopeptidase, carboxipeptidase, cisteína proteinase, cisteína

peptidase e uma subunidade regulatória do proteassomo 26S. No entanto, há três proteínas superexpressas: co-chaperona GrpE, aspartil aminopeptidase e uma peptidase do tipo calpaína. Interessante foi o fato de que a categoria *protein modification process* (BP), formado principalmente por genes que possuem atividade quinase (*kinase activity* – MF) como as proteínas quinase A, serina/treonina, e outras proteínas quinases, e por genes com atividade transferase, como a acetiltransferase, está induzida. A indução na expressão de proteínas quinase, em associação com genes que possuem função de ligação de íons cálcio (como três genes que codificam para a proteína EF-hand 5 e uma cisteína peptidase do tipo calpaína), como também uma proteína hipotética semelhante a uma fosfodiesterase do tipo C fosfolipase sugere que *T. cruzi* está respondendo ao estresse causado pela radiação com ativando vias de transdução de sinal que envolvem cálcio e lipídeos de membrana. A expressão do gene que codifica para a proteína transportadora de cátions também foi validado por qRT-PCR.

#### 4.11.3 – Efeito da irradiação na expressão dos genes envolvidos com a síntese de ATP

Após a radiação, *T. cruzi* sabidamente repara o DNA fragmentado e retorna o seu crescimento após 240 horas (Regis-Da-Silva *et al.*, 2006). Dessa forma, a geração de ATP é um processo essencial para a manutenção do metabolismo basal e para o reparo de DNA. Porém, para a nossa grande surpresa, a análise funcional revelou que genes que participam de diversas vias envolvidas em metabolismo de carboidratos e de precursores de metabólitos e energia (*carbohydrate metabolic process* e *generation of precursor of metabolites and energy* - BP) estão subexpressos. Alguns exemplos incluem: transportador de hexose, subunidade beta da piruvato desidrogenase E1 fosfoenolpiruvato, carboxiquinase glicosomal, uma proteína hipotética conservada que parece ser um componente protéico do tipo 2-ferro-2-enxofre do complexo succinato

desidrogenase, succinil-CoA sintetase, e uma ATPase do tipo V. Apenas um gene, que codifica a glicose-6 fosfato isomerase, está superexpresso. Assim, transcritos que codificam para proteínas dos processos metabólicos basais como a glicólise, gliconeogênese, ciclo do ácido cítrico, e produção de ATP estão diminuídos e indicam que estas vias estão reprimidas em algum grau.

Considerando a cadeia respiratória, o gene nuclear citocromo c1 está subexpresso, mas lembrando os resultados obtidos após a análise das sequências obsoletas, três genes mitocondriais conhecidos (Citocromo oxidase I - COI, NADH desidrogenase subunidade I – ND1, e 5 – ND5) estão superexpressos.

#### 4.11.4 – Efeito da irradiação na expressão dos genes envolvidos com a resposta ao estresse oxidativo

A radiação gama é capaz de causar estresse oxidativo nas células através da produção de espécies reativas de oxigênio (ROS), consequência dos efeitos dos raios ionizantes sobre as moléculas de água presentes do organismo. ROS é capaz de causar quebras de fita simples e dupla nas moléculas de DNA, caracterizando dessa forma o efeito indireto da radiação ionizante. Os resultados indicaram que a radiação gama induziu dois genes relacionados com a remoção de ROS: peroxidase 2 dependente de tripanotiona/triparedoxina e triparedoxina peroxidase. Esses genes foram classificados como genes de resposta ao stress, dentro dos processos biológicos (*Response to stress - BP*), e possuem atividade de oxidoreductase (*oxidoreductase activity – MF*). Esses genes fazem parte do único mecanismo antioxidante encontrado nos tripanosomatídeos (Muller *et al.*, 2003). No entanto, outra proteína participante da via de detoxificação, a tripanotiona redutase, está subexpressa (tabela 5).



#### 4.11.5 – Efeito da irradiação na expressão dos genes envolvidos com o reparo de DNA

Conforme já foi dito anteriormente, *T. cruzi* é bastante resistente à radiação ionizante, sendo capaz de sobreviver e reparar as quebras de fita dupla ocorridas em suas moléculas de DNA. Já se sabe também que o gene TcRAD51, envolvido no reparo de quebras de fita dupla através do mecanismo de recombinação homóloga, apresenta superexpressão 4 horas após o desafio com 500 Gy de raios gama (Regis-Da-Silva *et al.*, 2006).

*T. cruzi* apresenta diversos outros genes envolvidos no reparo de DNA e eles foram compilados em uma tabela suplementar (S8) publicado junto com o genoma deste organismo (El-Sayed *et al.*, 2005a). Os valores de *fold-change* dos genes cuja intensidade de expressão foi mensurada foram reunidos na tabela 10. Porém, muitos genes que atuam no reparo não tiveram nenhum valor de expressão associado. A lista completa pode ser consultada no material anexo 5. Percebe-se com clareza que os valores de *fold-change* variaram muito pouco, o que pode indicar que a expressão destes genes manteve-se em níveis constitutivos. No entanto, se voltarmos na figura 24, pode-se ver que há um gene relacionado com reparo de DNA (*DNA repair* – BP). Este gene se chama tirosil-DNA fosfodiesterase (Tdp-1) e foi significativamente superexpresso 24 horas após a irradiação, provavelmente quando o processo de reparo já estava ativo. No entanto, no momento da publicação do genoma de *T. cruzi*, essa enzima o mecanismo de ação dessa enzima era pouco conhecido, e o gene não foi listado como gene de reparo pelo consórcio que publicou o genoma (El-Sayed *et al.*, 2005a).

Tabela 11: Valores de *fold-change* de genes participantes de vias de reparo de DNA

Via de reparo e nome do gene	ID	<i>Fold-change</i>				
		l.a.i	4hs	24hs	48hs	96hs
<b>Direct repair</b>						
O-6 methyl-guanine alkyl transferase, putative (MGMT)	Tc00.1047053508347.20	-1,02	-1,00	-1,03	-1,00	-1,00
<b>Oxidative repair</b>						
alkylated DNA repair protein, putative (alkB)	Tc00.1047053510687.140	-1,04	-1,10	-1,13	-1,03	-1,05
<b>Base excision repair - DNA glycosylase</b>						
MutY/Nth/A-G glycosylase	Tc00.1047053511803.20	-1,10	1,15	1,12	1,03	1,01
uracil-DNA glycosylase	Tc00.1047053511277.330	-1,01	1,01	-1,05	-1,00	-1,10
<b>Base excision repair - DNA polymerase</b>						
mitochondrial DNA polymerase beta, putative	Tc00.1047053503955.20	-1,07	-1,08	1,07	-1,04	-1,09
<b>Nucleotide excision repair –Open complex</b>						
TFIIH-XPD/RAD3	Tc00.1047053511075.30	-1,12	-1,11	1,09	1,08	1,04
<b>Homologous recombination – Initiation</b>						
meiosis recombination protein SPO11, putative	Tc00.1047053509537.60	1,10	1,01	1,05	1,18	1,23
<b>Homologous recombination - Strand exchange</b>						
meiotic recombination protein DMC1, putative	Tc00.1047053506885.310	-1,00	1,21	-1,00	1,14	1,10
DNA repair protein RAD51, putative	Tc00.1047053503801.30	-1,27	1,06	-1,10	1,14	1,03
DNA repair protein, putative (RAD51-4)	Tc00.1047053503613.30	1,01	-1,00	-1,07	-1,04	-1,09
<b>Homologous recombination - Meiosis-specific recombination enzymes</b>						
meiotic recombination protein DMC1, putative	Tc00.1047053506885.310	-1,00	1,21	-1,00	1,14	1,10
meiosis recombination protein SPO11, putative	Tc00.1047053509537.60	1,10	1,01	-1,05	-1,18	-1,23

l.a.i = logo após a irradiação

## 5 - DISCUSSÃO

---

*T. cruzi* é um organismo extremamente resistente à radiação ionizante, sendo capaz de sobreviver a doses muito altas de radiação gama. O nosso grupo relatou previamente que este organismo é capaz de sobreviver a doses de 1000 Gy. Além disso, doses de 500 Gy de raios gama são capazes de causar extensa fragmentação dos cromossomos e, de forma impressionante, *T. cruzi* é capaz de reparar o DNA danificado em menos de 48 horas (Regis-Da-Silva *et al.*, 2006). Em aproximadamente 10 dias, as células tornam a se multiplicar, entrando novamente na fase logarítmica de crescimento. Pouco se sabe sobre o processo que modela esse mecanismo, que deve incluir o reparo de fitas-dupla de DNA através da via de recombinação homóloga com a participação do gene TcRAD-51 (Regis-Da-Silva *et al.*, 2006). Dessa forma, buscamos através da técnica de microarranjos avaliar as mudanças na expressão gênica ao longo das 96 primeiras horas após a irradiação com o objetivo de incorporar novos conhecimentos que contribuam para a elucidação do processo de recuperação dos danos no DNA causados pela radiação ionizante.

Experimentos de microarranjos podem ser problemáticos, uma vez que a cada etapa acumulam-se erros que podem comprometer os resultados finais. A manipulação das amostras e aplicação de algoritmos corretos para tratamento dos dados são fatores fundamentais para a robustez dos resultados finais. A etapa de hibridização é especialmente importante, pois a captação de sinal de boa qualidade é o pilar da técnica. Para esta fase, o experimento ideal é aquele que utiliza mRNAs que não sofreram amplificação. No entanto, a não amplificação do RNA mostrou-se bastante ineficaz neste projeto, e o resultado final foram lâminas escaneadas que não produziram sinal

algum. Dessa forma, a solução foi incluir a etapa de amplificação das moléculas de mRNA através do método de Eberwine (Phillips e Eberwine, 1996). A inclusão desta etapa é motivo de discussão. Vários estudos compararam os níveis de expressão dos genes cujos mRNAs foram, ou não amplificados. A introdução do viés é inevitável e, foi visto que as plataformas comerciais da Agilent (plataforma antiga) e Affymetrix foram as mais afetadas, pois entre 1 e 9% das sondas tiveram suas intensidades aumentadas quando comparadas com as intensidades de sondas marcadas com mRNA não-amplificado (Kerkhoven *et al.*, 2008). Plataformas de duas-cores domésticas não foram avaliadas. No entanto, diversos autores já haviam concluído que, apesar de os níveis de correlação entre hibridizações com amostras amplificadas ou não nem sempre se manterem, o processo é extremamente útil para a caracterização de genes com baixos níveis de expressão (Wang *et al.*, 2000; Hu *et al.*, 2002; Puskas *et al.*, 2002; Scheidl *et al.*, 2002; Zhao *et al.*, 2002; Nygaard *et al.*, 2003), que é uma das características das amostras deste trabalho.

A padronização dos métodos de correção de *background* e normalização (pré-processamento dos dados) foi um passo fundamental. O pré-processamento refere-se ao procedimento utilizado para converter dados brutos de intensidade em valores analisáveis de expressão gênica e posteriormente utilizados nas análises estatísticas. É importante salientar que a escolha de um método em detrimento a outro pode ter significantes efeitos nos resultados de medida de expressão e sua utilidade nas análises subseqüentes. O ideal é ter medidas de expressão que sejam precisas e acuradas, ou seja, que apresentem baixo viés e pequena variação.

O primeiro passo do pré-processamento dos dados foi a correção do *background* e o método *NormExp*, baseado em um modelo de convolução (Ritchie *et al.*, 2007), foi o escolhido. No entanto, o método de correção de *background* mais utilizado é o

*Subtract*. A aplicação deste algoritmo gerou muitos valores negativos, pois as lâminas apresentaram valores altos de *background*. Isso aconteceu porque menos da metade dos *spots* de cada sub-*array* acendeu, o que tornou o *background* médio muito alto. Somase o fato de que os valores médios de intensidade foram baixos, pela própria característica de controle da expressão gênica do *T. cruzi* (transcrição policistrônica com controle de expressão gênica principalmente pós-transcricional). Assim, o método *Subtract* mostrou ser uma solução imperfeita para redução do viés espacial (causado por não-uniformidade da hibridização) por não conseguir lidar com valores altos de *background* e baixos de sinal (Do e Choi, 2006). Caso tivéssemos optado por esse método, dados válidos acabariam sendo desconsiderados na etapa de normalização, pois o valor final do sinal deve ser obrigatoriamente positivo.

O método *Robust spline* foi o escolhido para normalizar os dados dentro de cada lâmina. Este é um método empírico bayesiano intermediário entre o método *loess* (ou *lowess*) aplicado por agulha de impressão ou global. O *Loess* (tanto o global quanto o por agulhas) funciona muito bem para lâminas que possuem muitos pontos acesos em cada sub-*array*. Esse método consegue voltar os valores para mediana zero através de uma função de regressão linear capaz de trabalhar com todos os pontos do gráfico, uma vez que possua informação para isso. Porém, como foi mencionado acima, menos da metade dos *spots* apresentaram sinais de intensidade e achar uma função linear que representasse todos os pontos foi inviável. Por isso testou-se outros métodos disponíveis no Limma e o *Robust spline* mostrou ser mais apropriado, pois ele não procura uma função linear, e sim, uma curva que se ajusta melhor aos valores, o que permite a geração de listas de genes diferentemente expressos robustas para as análises subsequentes.

As listas de genes diferentemente expressos (uma para cada tempo analisado) sugerem que as mudanças na expressão gênica que ocorreram após a irradiação com raios gama foram dependentes do tempo. Durante as quatro primeiras horas o cenário foi caracterizado por uma proporção maior de genes que apresentaram expressão reprimida e não induzida. No entanto, o cenário mudou após 24 horas, considerando que o número de genes superexpressos foi maior que o número de genes subexpressos, tendência que se manteve até o último tempo analisado. Dessa forma, é possível que *T. cruzi* inicie a recuperação frente à radiação em algum momento entre 4 e 24 horas após o tratamento. Em estudos onde a resposta dos organismos *D. radiodurans* e *Sacharomicys cereviseae* frente à radiação gama foi avaliada, os autores não conseguiram inferir um momento específico quando o número de genes induzidos supera o número de genes reprimidos (Liu *et al.*, 2003; Kimura *et al.*, 2006).

As análises indicaram 273 genes diferentemente expressos ao longo das 96 horas, sendo 113 subexpressos, e 160 superexpressos. Os genes foram classificados em quatro categorias distintas: genes com função conhecida ou predita, genes cujo produto é anotado como proteína hipotética (genes de função ainda desconhecida), genes que codificam para RHS (*Retrotransposon hot spot protein*), e sequências obsoletas. Os genes que codificam para proteínas de função desconhecida representam mais de 60% dos genes superexpressos, quando as sequências obsoletas e os genes RHSs são desconsiderados. A presença de genes de função desconhecida em lista de genes diferentemente expressos é situação corriqueira nesse tipo de estudo, mas uma proporção tão grande torna a análise biológica bastante obscura. A anotação funcional por GO-Slim atribuiu termos de ontologia gênica para alguns destes genes hipotéticos a partir da similaridade com domínios protéicos. Ainda assim, as informações incorporadas pouco ajudaram. Nem mesmo os agrupamentos de genes baseados em

padrões de expressão gênica ajudaram nessa tarefa, ao contrário do que aconteceu com a bactéria *D. radiodurans*, onde os autores foram capazes de inferir funções para genes de acordo com o contexto do padrão de expressão gênica de genes que faziam parte de um mesmo operon (Liu *et al.*, 2003). De fato, estruturas clássicas de operons nunca foram identificadas neste parasito, mesmo os genes sendo transcritos em unidades policistrônicas. Essas unidades podem conter genes de vias não relacionadas que, após o processo de *trans-splicing*, origina transcritos monocistrônicos maduros com níveis de expressão variados. E por fim, nenhum promotor clássico para a enzima RNA polimerase II foi encontrado no genoma de *T. cruzi* (El-Sayed *et al.*, 2005b).

É bem sabido que os níveis de mRNAs de tripanosomatídeos são ditados principalmente pela estabilidade do mRNA, sendo o controle pós-transcricional. No entanto, os níveis de proteínas também dependem de mecanismos adicionais que controlam de forma eficiente o início e alongamento da tradução (seleção de tradução) e o *turnover* da proteína (Teixeira e Da Rocha, 2003; Clayton e Shapira, 2007; Haile e Papadopoulou, 2007; Horn, 2008). Utilizando microarranjos para avaliar os perfis de transcrição nos diferentes estágios de diferenciação, Queiroz e colaboradores demonstraram que genes de *Trypanosoma brucei* formam regulons pós-transcricionais nos quais mRNAs codificadores de proteínas que participam de determinadas vias, ou que codificam componentes de complexos protéicos, apresentam expressão co-regulada. Os autores chamaram a atenção para transcritos que codificam proteínas envolvidas na biogênese de ribossomos e flagelos, assim como componentes de vias de transdução de sinal e complexos protéicos necessários para vias metabólicas que apresentaram uma notável co-regulação (Queiroz *et al.*, 2009). Co-regulação de níveis de transcritos podem ser alcançados através da ligação de proteínas regulatórias específicas, ou de RNAs pequenos a elementos na região 3' não traduzida (3'-UTR) dos

mRNAs. Esses elementos podem controlar a tradução, ou o decaimento do mRNA (Clayton e Shapira, 2007; Haile e Papadopoulou, 2007). Neste trabalho, tentou-se, sem sucesso, identificar motivos conservados na região 3'-UTR dos transcritos diferentemente expressos que formavam um mesmo agrupamento gênico (usando o software MEME Suite) e que poderiam atuar como elementos de controle do decaimento do mRNA (dados não mostrados).

Foi relatado nos resultados que os valores de *fold-change* dos genes diferentemente expressos raramente foram maiores que 2X. Geralmente tais valores baixos seriam excluídos das análises subsequentes em estudos com outros organismos. No entanto, para organismos da ordem Kinetoplastida, artigos publicados tanto com espécies do gênero *Trypanosoma* quanto *Leishmania* relataram mudanças bastante modestas na abundância de moléculas de mRNA detectadas por microarranjos, mas que foram validados pela técnica de qRT-PCR (Minning *et al.*, 2003; Saxena *et al.*, 2003). Estes trabalhos mostraram que genes com valores baixos de *fold-change*, porém com significância estatística, são de fato diferentemente expressos.

Assim como já havia sido descrito, a curva de crescimento das células irradiadas estagna (Regis-Da-Silva *et al.*, 2006) indicando que, pelo menos durante 96 horas, não há multiplicação aparente. Genes associados com a divisão celular e motilidade (alfa e beta-tubulinas), com metabolismo de lipídeo (sendo alguns genes pertencentes à via de biossíntese de ergosterol) e com atividade proteassômica (subunidade não-ATPase regulatória do proteassomo 26S) estão subexpressos. Considerando que as células não estão se dividindo, é razoável pensar na possibilidade de que a formação de citoesqueleto e a biossíntese de membranas celulares sejam processos desnecessários nos primeiros dias após a irradiação, quando o organismo provavelmente está focando as suas energias no reparo do DNA danificado. O complexo proteassomo 26S, que faz



parte da via proteolítica, controla, entre diversas funções, a progressão do ciclo celular (Pagano, 1997). Foi visto que a inibição do proteassomo através do tratamento com lactacistina inibe a diferenciação (células epimastigotas não se diferenciaram em tripomastigotas metacíclicas) e multiplicação celular, sugerindo que este complexo possui um papel importante na divisão celular (Cardoso *et al.*, 2008). A diminuição da divisão celular já foi descrita para outros organismos após o estresse induzido pela radiação ionizante (Liu *et al.*, 2003; Ricaud *et al.*, 2007). No entanto, não foram encontrados trabalhos que relataram a diminuição nos níveis de expressão gênica de enzimas envolvidas no metabolismo de lipídeo que poderiam estar associados com o efeito da radiação ionizante.

Três genes associados com a atividade motora de microtúbulos foram superexpressos: uma dineína de cadeia leve (lc6) e duas de cadeia pesada, que são componentes do flagelo (Asai e Wilkes, 2004; Dibella *et al.*, 2005) e responsáveis pelo movimento celular das células de *T. cruzi*. De fato, recentemente foi demonstrado que o silenciamento de uma dineína de cadeia pesada do braço interno do flagelo resultou na imobilidade flagelar das células de *T. brucei* (Springer *et al.*, 2011). Mesmo apresentando uma parada no crescimento celular, células epimastigotas apresentaram intenso movimento flagelar durante todo o período estudado, e isso pode ser explicado, em parte, pelo aumento nos níveis dos constituintes do microtúbulo flagelar.

Quarenta e duas sequências obsoletas apresentaram expressão diferenciada, sendo 15 subexpressas, e 27 superexpressas. Dentro do grupo das superexpressas, quatro chamaram a atenção pelos valores altos de *fold-change* (uma delas foi quase 9X mais expressa), e nos instigou a estudá-las mais detalhadamente. De forma inesperada (uma vez que a lâmina foi projetada para conter sondas de genes nucleares), os resultados mostraram que quatro de 18 genes conhecidos do maxicírculo, componente

do cinetoplasto de *T. cruzi* (kDNA) foram induzidos frente à irradiação: citocromo oxidase subunidade 1 (COI), NADH desidrogenase subunidades 1 e 5 (ND-1 e ND-5) e proteína não identificada (MURF-1). Os genes restantes do kDNA não foram impressos na lâmina e consequentemente não foram avaliados. Interessantemente, os quatro genes apresentaram perfis semelhantes de expressão gênica, com exceção de 96 horas quando o gene ND-5 apresentou uma queda nos níveis de transcritos. Essa queda pode ser real, ou seja, de fato este gene começa a voltar para níveis mais basais de expressão após 96 horas, ou ocorreu algum problema na quantificação da intensidade deste gene (como por exemplo, problema de hibridização na região onde está localizado o *spot*). De qualquer forma, os níveis de expressão semelhantes ao longo das horas dos quatro genes nos levou a considerar a possibilidade dos outros 14 genes não avaliados estarem igualmente superexpressos após 24 horas. Esses resultados indicam que o kDNA foi responsivo à radiação ionizante, talvez como consequência do estresse oxidativo. Em um estudo com uma linhagem celular de glioblastoma humano radioresistente (T98G), os autores demonstraram que os genes citocromo c oxidase (subunidades 1 e 2) e NADH desidrogenase subunidade 4 (ND-4) foram estimulados após tratamento com radiação ionizante com aumento na quantidade dos seus transcritos, e que este fato havia sido mediado pelo estresse oxidativo (Gong *et al.*, 1998). Nenhum dos genes do kDNA foi incluído para ser validado por qRT-PCR. O cinetoplasto é formado por diversas cópias de maxicírculos e minicírculos. Porém, o gene utilizado como normalizador foi o gliceraldeído 3-fosfato desidrogenase (GAPDH), de localização nuclear, e seria impossível comparar níveis de expressão de genes nucleares com genes mitocondriais. Vale ressaltar que os valores de *fold-change* de fato refletem o aumento na quantidade de transcritos destes genes, uma vez que os valores de intensidade são obtidos através da razão dos *spots* nas amostras testes sobre os *spots* referentes do

mesmo gene nas amostras controle. Ou seja, são valores de intensidade de expressão de genes do kDNA das amostras irradiadas sobre valores de expressão de genes do kDNA das amostras não-irradiadas.

A análise funcional sugere que a tradução e processamento de proteínas foram prejudicados após a irradiação, uma vez que os níveis de expressão de sete fatores de iniciação e alongamento da tradução, assim como nove componentes estruturais ribossomais e a triptofanil-tRNA sintetase foram reprimidos. E o mais interessante, nenhum gene com tais categorias atribuídas foi superexpresso. Sabe-se que os kinetoplastídeos controlam a expressão gênica pós-transcricionalmente e traducionalmente. De fato, poucos fatores de transcrição foram identificados nestes organismos (Palenchar e Bellofatto, 2006). Assim, os níveis de transcrição dos genes (em geral) não variam muito (o que novamente justifica os baixos valores de *fold-change* observados). Para os organismos eucariotos que controlam a expressão dos genes principalmente durante a transcrição, a radiação ionizante provoca um impacto global na maquinaria e nos fatores de transcrição. Foi relatado que a radiação gama diminui a atividade transcricional em leveduras (Kimura *et al.*, 2006) e *Arabidopsis thaliana* (Tanaka *et al.*, 2004; Ricaud *et al.*, 2007). No entanto, foi visto que a indução da expressão gênica frente à radiação ionizante em *D. radiodurans*, um organismo procarioto, ocorre em genes relacionados à síntese e endereçamento de proteínas (Tanaka *et al.*, 2004). É importante ressaltar que a recuperação de danos ao DNA após a radiação gama ocorre de forma extraordinariamente rápida nesta bactéria, sendo possível observar o aumento da expressão desses genes uma hora após o tratamento (Tanaka *et al.*, 2004).

É sabido que a indução de situações de estresse pode levar os organismos a reagirem expressando diversas classes de proteínas, sendo que as *heat-shock* são as

mais conhecidas. Estas proteínas desempenham um importante papel no enovelamento e conformação de proteínas. No entanto, a redução na taxa de síntese e processamento de proteínas pode ter levado à subexpressão de proteínas chaperonas hsp60 e hsp70. Esses resultados foram bastante diferentes dos encontrados na literatura, onde a radiação gama induziu a expressão de proteínas *heat-shock* em duas bactérias gram-negativas (*E. coli* e *Salmonella* sorotipo *Typhimurium*) e em duas gram-positivas (*Staphylococcus aureus* e *Listeria monocytogenes*) (Caillet *et al.*, 2008), em camundongos (Nogami *et al.*, 1993) e em células humanas (Calini *et al.*, 2003). Por outro lado, três genes relacionados a processos de metabolismo de proteínas (co-chaperona GrpE, aspartil aminopeptidase e cisteína peptidase do tipo calpaína), e diversos outros genes que atuam nos processos de modificações de proteínas que apresentam atividades de quinase, transferase, ligação de íons cálcio e transporte de cátions foram superexpressos, o que nos leva a sugerir que *T. cruzi* está respondendo ao estresse através da ativação de vias de transdução de sinal dependente de cálcio ou lipídeos de membranas. Os nossos resultados referentes à cisteína peptidase do tipo calpaína estão de acordo com um estudo prévio onde foi relatado que este gene apresenta expressão aumentada em células epimastigotas sob três diferentes tipos de estresse (nutricional, temperatura e pH), através da avaliação da expressão gênica com microarranjos e qRT-PCR, o que levou os autores a sugerir que este gene possui uma função geral na resposta à estresses, apresentando um possível papel na transdução de sinal em função de sua atividade proteolítica (Probst *et al.*, resultados não publicados; Giese *et al.*, 2008).

O *T. cruzi* possui diversos genes que sabidamente atuam no processo de reparo do DNA (El-Sayed *et al.*, 2005a). No entanto os níveis de expressão destes genes não foram afetados ou detectados neste estudo. É possível que a expressão destes genes seja

constitutiva, como já foi citado a archaee hipertermofílica *Pirococcus* (Williams *et al.*, 2007). Neste trabalho, genes sabidamente responsivos à radiação ionizante não foram diferentemente expressos quando analisados por microarranjos. Ainda assim, esse resultado foi inesperado, uma vez que o nosso grupo já havia descrito que o TcRAD51, envolvido no reparo de quebras de fita dupla através da recombinação homóloga, aumenta seus níveis de transcrito aproximadamente 4X e 2X após 4 e 24 horas após a irradiação (respectivamente) de acordo com resultados de qRT-PCR (Regis-Da-Silva *et al.*, 2006). O trabalho do nosso grupo pode ter sido capaz de perceber as mudanças nos transcritos de TcRAD51 porque usou uma técnica sabidamente mais sensível para medir transcritos (qRT-PCR). Apesar disso, o gene codificador de uma enzima de reparo que nem consta na relação dos genes relacionados ao reparo de DNA (El-Sayed *et al.*, 2005a) foi superexpresso 24 horas após a irradiação. A tirosil-DNA fosfodiesterase I (tdp-1) é capaz de liberar o DNA da topoisomerase I (topo-I) através da clivagem de ligações 3'-fosfotirosil entre a extremidade 3' e a tirosina do tipo-I, liberando o DNA com uma ponta 3'-fosfato (Yang *et al.*, 1996). Esse processo prepara a fita de DNA com quebra-dupla para o reparo através da atuação da DNA ligase III (El-Khamisy *et al.*, 2005). Esse gene foi primeiramente caracterizado no núcleo (Yang *et al.*, 1996), mas recentemente a sua presença foi verificada na mitocôndria de células de mamíferos, e o mecanismo de reparo por excisão de base (BER), necessário para um reparo eficiente contra o estresse oxidativo, depende dele (Das *et al.*, 2010). Alguns aspectos dessa enzima já foram caracterizados em *Leishmania donovani*, um parasito da mesma ordem de *T. cruzi* (Kinetoplastida). Os autores demonstraram que a tdp-1 possivelmente também atua no reparo de DNA mediado pela clivagem da topoisomerase 1 que se encontra presa à molécula de DNA, tanto no núcleo quanto no kDNA. Além disso, a superexpressão desse gene protegeu as células promastigotas

contra a fragmentação do DNA quando tratadas com peróxido de hidrogênio (H<sub>2</sub>O<sub>2</sub>), aumentando a sua sobrevivência. No entanto, não foi proposto nenhum mecanismo de ação, nem mesmo quais seriam as proteínas que co-participariam deste processo, considerando que os kinetoplastídeos não possuem ortólogos de DNA ligase III (Banerjee *et al.*, 2010). Assim, visto que tdp-1 possui um papel frente ao estresse oxidativo gerado por H<sub>2</sub>O<sub>2</sub>, e que foi responsivo ao estresse causado por raios gama, mais estudos são necessários para desvendar o real papel desta enzima no processo de reparo de DNA, e o mecanismo, incluindo os genes que atuariam em conjunto.

*T. cruzi* é naturalmente exposto às espécies reativas de oxigênio (ROS) produzidas tanto por ele mesmo quanto pelo sistema imune do hospedeiro vertebrado durante o seu ciclo de vida (Docampo, 1990), e possui um sistema completo para combater este estresse baseado no metabolismo do tripanotiona-tiol (Krauth-Siegel e Comini, 2008). Os resultados mostraram que dois genes pertencentes a esta via (e que possuem ação oxidoreductase) estão superexpressos: uma peroxidase do tipo 2 dependente de tripanotiona/triparedoxina e uma triparedoxina peroxidase. No entanto, a tripanotiona redutase, uma enzima chave que reduz a tripanotiona dissulfeto para o seu estado ativo (para tripanotiona), está subexpressa. Esses resultados sugerem que *T. cruzi* está respondendo às espécies reativas de oxigênio em algum grau. Corroborando com os nossos resultados, foi observado que células de *E. coli* transfectadas com gene da tripanotiona de *T. cruzi* apresentaram um aumento de 4,6X na proteção do DNA genômico e 4,3X na resistência contra as ações dos raios gama, além de uma redução nos níveis intracelulares de ROS (Fitzgerald *et al.*, 2010).

A questão que nos deixa mais perplexos é como o DNA é consistentemente reparado se as bandas cromossômicas são extensivamente fragmentadas (Regis-Da-Silva *et al.*, 2006). Nós mostramos que a radiação gama induziu de forma positiva a

expressão dos genes/pseudogenes RHS, uma vez que eles estão super-representados ( $p < 0,0001$ ) dentro dos grupos de genes superexpressos 48 e 96 horas após a irradiação quando comparadas com o restante dos genes (Figura 20). Coincidentemente, 48 horas é o momento onde acredita-se que o reparo de DNA foi alcançado, pois é quando os cromossomos se reestruturaram através da recombinação homóloga (Regis-Da-Silva *et al.*, 2006). Os genes que codificam para os RHSs estão localizados em regiões cromossômicas bastante repetitivas. No entanto, outros genes presentes em tais regiões como os genes das famílias das transalidasas, glicoproteínas, proteínas de superfície preditas DGF-1, ou *dispersed gene family-1* e MASP não foram induzidas da mesma forma como RHS. Infelizmente não foi possível demonstrar uma correlação direta entre o reparo de DNA e a explosão na expressão de genes RHS. Corroborando com os nossos resultados, danos genômicos causados por radiação gama em *S. cerevisiae* induziram a expressão de transcritos de Ty1 (retroelementos do tipo LTR), alcançando o impressionante valor de 86X após a dose de 800 Gy, quando comparados com células não irradiadas, e o fenômeno de retrotransposição foi ativado de forma dose-dependente (Sacerdot *et al.*, 2005). Foi visto também que a retrotransposição de elementos móveis longos intercalados (*Long Interspersed Elements* - LINE-1s) presentes no genoma humano também aumentou após exposição a raios gama, e isso causou diretamente instabilidade genômica (Farkash *et al.*, 2006).

*T. cruzi* não é naturalmente exposto à radiação ionizante, mas mesmo assim, é extremamente tolerante. Se analisarmos alguns aspectos evolutivos do ciclo de vida deste parasito, é bem razoável pensar que a resistência à radiação existe provavelmente devido a uma combinação de estresses sofridos pelo *T. cruzi* durante a fase no hospedeiro invertebrado (barbeiros triatomíneos). Podemos citar algumas fontes de estresse: temperatura, variações no pH e na osmolaridade, produtos presentes na saliva

e na digestão sanguínea (Kollien e Schaub, 2000). Além disso, a digestão do sangue acaba sendo a fonte de grupos heme (moléculas que contém íons de ferro) fundamentais para a reação química de Fenton. Esta reação é altamente prejudicial porque gera uma quantidade muito grande de radicais hidroxil que, por sua vez, causam estresse oxidativo. Para incrementar a lista de possíveis estresses que células podem sofrer dentro do intestino dos barbeiros, logo após uma refeição, a água presente no sangue é excretada pelos túbulos de Malpighi, e o sangue desidratado é estocado na cavidade intestinal (Kappagoda *et al.*, 1979), o que pode causar uma situação de dessecação. Tudo isso indica que a passagem pelo sistema intestinal do inseto vetor é um desafio extraordinário para a integridade genômica do *T. cruzi*. Corroborando com esse ponto de vista, organismos altamente resistentes aos raios gama, como *D. radiodurans*, também não são naturalmente expostos à radiação ionizante, mas são principalmente caracterizados por uma altíssima tolerância à dessecação, ao invés do estresse oxidativo. Foi demonstrado que cepas de *D. radiodurans* sensíveis à radiação ionizante foram também suscetíveis à dessecação causada por longos períodos de desidratação (que também são capazes de causar quebras de fita dupla no DNA), e que a sobrevivência é diretamente dependente da habilidade deste organismo de reparar o dano (Mattimore e Battista, 1996).



## 6 - CONCLUSÕES

---

Esta tese estudou os efeitos da radiação gama na expressão gênica de *T. cruzi*.

As principais conclusões deste trabalho foram:

- O número de genes diferentemente expressos ao longo de 96 horas (em cinco tempos diferentes analisados) foram 273, sendo 160 superexpressos e 113 subexpressos.
- A radiação gama afetou a expressão gênica de *T. cruzi* de forma tempo dependente, quando as células apresentam uma queda geral dos transcritos nas quatro primeiras horas e um aumento geral dos transcritos após 24 horas.
- Conforme esperado, os genes não apresentaram altos valores de *fold-change*, conseqüência do processo de transcrição policistrônica constitutiva e ausência de fatores de transcrição
- A maioria dos genes subexpressos possui função conhecida e a maioria dos genes superexpressos são os que codificam para proteínas hipotéticas
- Genes do maxicírculo estão consideravelmente induzidos, e este fato pode estar correlacionado com uma possível ação mitocondrial contra espécies reativas de oxigênio.
- Os agrupamentos formados pelo *heatmap* revelaram que genes classificados em uma mesma categoria funcional, processo regulatório ou via metabólica não apresentaram cinética similar de expressão ao longo das horas. A exceção são os genes ribossomais, que se agruparam, apesar da presença também de outros genes não relacionados no mesmo agrupamento.
- Os genes RHS foram substancialmente induzidos, e a razão disto mantém-se obscura

- Os genes de reparo não foram induzidos de forma significativa, com exceção do gene tirosil-DNA fosfodiesterase -1

## 7 – REFERÊNCIAS BIBLIOGRÁFICAS

---

- Abuin, G., Freitas-Junior, L.H., Colli, W., *et al.* Expression of trans-sialidase and 85-kDa glycoprotein genes in *Trypanosoma cruzi* is differentially regulated at the post-transcriptional level by labile protein factors. *J Biol Chem*, 1999, 274 (19):13041-13047
- Aikens, J. and Dix, T.A. Peroxy radical (HOO.) initiated lipid peroxidation. The role of fatty acid hydroperoxides. *J Biol Chem*, 1991, 266 (23):15091-15098
- Asai, D.J. and Wilkes, D.E. The dynein heavy chain family. *J Eukaryot Microbiol*, 2004, 51 (1):23-29
- Atwood, J.A., 3rd, Minning, T., Ludolf, F., *et al.* Glycoproteomics of *Trypanosoma cruzi* trypomastigotes using subcellular fractionation, lectin affinity, and stable isotope labeling. *J Proteome Res*, 2006, 5 (12):3376-3384
- Atwood, J.A., 3rd, Weatherly, D.B., Minning, T.A., *et al.* The *Trypanosoma cruzi* proteome. *Science*, 2005, 309 (5733):473-476
- Augusto-Pinto, L., Bartholomeu, D.C., Teixeira, S.M., *et al.* Molecular cloning and characterization of the DNA mismatch repair gene class 2 from the *Trypanosoma cruzi*. *Gene*, 2001, 272 (1-2):323-333
- Augusto-Pinto, L., Teixeira, S.M., Pena, S.D., *et al.* Single-nucleotide polymorphisms of the *Trypanosoma cruzi* MSH2 gene support the existence of three phylogenetic lineages presenting differences in mismatch-repair efficiency. *Genetics*, 2003, 164 (1):117-126
- Avila, A.R., Dallagiovanna, B., Yamada-Ogatta, S.F., *et al.* Stage-specific gene expression during *Trypanosoma cruzi* metacyclogenesis. *Genet Mol Res*, 2003, 2 (1):159-168

- Bagwell, C.L., Moloney, M.G. and Yaqoob, M. Oxazolomycins: natural product lead structures for novel antibacterials by click fragment conjugation. *Bioorg Med Chem Lett*, 2010, 20 (7):2090-2094
- Banerjee, B., Roy, A., Sen, N., *et al.* A tyrosyl DNA phosphodiesterase 1 from kinetoplastid parasite *Leishmania donovani* (LdTdp1) capable of removing topo I-DNA covalent complexes. *Mol Microbiol*, 2010, 78 (1):119-137
- Baptista, C.S., Vencio, R.Z., Abdala, S., *et al.* Differential transcription profiles in *Trypanosoma cruzi* associated with clinical forms of Chagas disease: Maxicircle NADH dehydrogenase subunit 7 gene truncation in asymptomatic patient isolates. *Mol Biochem Parasitol*, 2006, 150 (2):236-248
- Baptista, C.S., Vencio, R.Z., Abdala, S., *et al.* DNA microarrays for comparative genomics and analysis of gene expression in *Trypanosoma cruzi*. *Mol Biochem Parasitol*, 2004, 138 (2):183-194
- Bartholomeu, D.C., Silva, R.A., Galvao, L.M., *et al.* *Trypanosoma cruzi*: RNA structure and post-transcriptional control of tubulin gene expression. *Exp Parasitol*, 2002, 102 (3-4):123-133
- Benjamini, Y. and Hochberg, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B*, 1995, 57 289-300
- Bennett, C.B., Lewis, L.K., Karthikeyan, G., *et al.* Genes required for ionizing radiation resistance in yeast. *Nat Genet*, 2001, 29 (4):426-434
- Berriman, M., Ghedin, E., Hertz-Fowler, C., *et al.* The genome of the African trypanosome *Trypanosoma brucei*. *Science*, 2005, 309 (5733):416-422

- Bhattacharyya, M.K., Norris, D.E. and Kumar, N. Molecular players of homologous recombination in protozoan parasites: implications for generating antigenic variation. *Infect Genet Evol*, 2004, 4 (2):91-98
- Brandao, A., Urmenyi, T., Rondinelli, E., *et al.* Identification of transcribed sequences (ESTs) in the *Trypanosoma cruzi* genome project. *Mem Inst Oswaldo Cruz*, 1997, 92 (6):863-866
- Bringaud, F., Biteau, N., Melville, S.E., *et al.* A new, expressed multigene family containing a hot spot for insertion of retroelements is associated with polymorphic subtelomeric regions of *Trypanosoma brucei*. *Eukaryot Cell*, 2002, 1 (1):137-151
- Caillet, S., Millette, M., Dussault, D., *et al.* Effect of gamma radiation on heat shock protein expression of four foodborne pathogens. *J Appl Microbiol*, 2008, 105 (5):1384-1391
- Calini, V., Urani, C. and Camatini, M. Overexpression of HSP70 is induced by ionizing radiation in C3H 10T1/2 cells and protects from DNA damage. *Toxicol In Vitro*, 2003, 17 (5-6):561-566
- Campos, P.C., Silva, V.G., Furtado, C., *et al.* *Trypanosoma cruzi* MSH2: Functional analyses on different parasite strains provide evidences for a role on the oxidative stress response. *Mol Biochem Parasitol*, 2010,
- Cardoso, J., Soares, M.J., Menna-Barreto, R.F., *et al.* Inhibition of proteasome activity blocks *Trypanosoma cruzi* growth and metacyclogenesis. *Parasitol Res*, 2008, 103 (4):941-951
- Carver, T., Berriman, M., Tivey, A., *et al.* Artemis and ACT: viewing, annotating and comparing sequences stored in a relational database. *Bioinformatics*, 2008, 24 (23):2672-2676

- Clayton, C. and Shapira, M. Post-transcriptional regulation of gene expression in trypanosomes and leishmanias. *Mol Biochem Parasitol*, 2007, 156 (2):93-101
- Clayton, C.E. Life without transcriptional control? From fly to man and back again. *EMBO J*, 2002, 21 (8):1881-1888
- Conway, C., Proudfoot, C., Burton, P., *et al.* Two pathways of homologous recombination in *Trypanosoma brucei*. *Mol Microbiol*, 2002, 45 (6):1687-1700
- Coura, J.R., Anunziato, N. and Willcox, H.P. [Chagas' disease morbidity. I - Study of cases originating in various states of Brazil, observed in Rio de Janeiro]. *Mem Inst Oswaldo Cruz*, 1983, 78 (3):363-372
- Cuzzocrea, S., Riley, D.P., Caputi, A.P., *et al.* Antioxidant therapy: a new pharmacological approach in shock, inflammation, and ischemia/reperfusion injury. *Pharmacol Rev*, 2001, 53 (1):135-159
- Dallagiovanna, B., Correa, A., Probst, C.M., *et al.* Functional genomic characterization of mRNAs associated with TcPUF6, a pumilio-like protein from *Trypanosoma cruzi*. *J Biol Chem*, 2008, 283 (13):8266-8273
- Das, B.B., Dexheimer, T.S., Maddali, K., *et al.* Role of tyrosyl-DNA phosphodiesterase (TDP1) in mitochondria. *Proc Natl Acad Sci U S A*, 2010, 107 (46):19790-19795
- Davis, M.J., Sehgal, M.S. and Ragan, M.A. Automatic, context-specific generation of Gene Ontology slims. *BMC Bioinformatics*, 2010, 11 498
- De Moura, M.B., Schamber-Reis, B.L., Passos Silva, D.G., *et al.* Cloning and characterization of DNA polymerase eta from *Trypanosoma cruzi*: roles for translesion bypass of oxidative damage. *Environ Mol Mutagen*, 2009, 50 (5):375-386
- De Souza, W. Basic cell biology of *Trypanosoma cruzi*. *Curr Pharm Des*, 2002, 8 (4):269-285

- Di Noia, J.M., D'orso, I., Sanchez, D.O., *et al.* 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*, 2000, 275 (14):10218-10227
- Dibella, L.M., Gorbatyuk, O., Sakato, M., *et al.* Differential light chain assembly influences outer arm dynein motor function. *Mol Biol Cell*, 2005, 16 (12):5661-5674
- Dix, T.A., Hess, K.M., Medina, M.A., *et al.* Mechanism of site-selective DNA nicking by the hydrodioxyl (perhydroxyl) radical. *Biochemistry*, 1996, 35 (14):4578-4583
- Do, J.H. and Choi, D.K. Normalization of microarray data: single-labeled and dual-labeled arrays. *Mol Cells*, 2006, 22 (3):254-261
- Docampo, R. Sensitivity of parasites to free radical damage by antiparasitic drugs. *Chem Biol Interact*, 1990, 73 (1):1-27
- Dudoit, S., Yang, Y H, Callow, M J, Speed, T P. Statistical methods for identifying differentially expressed genes in replicated cDNA microarrays experiments. *Statistica Sinica*, 2002, 12 111-139
- El-Khamisy, S.F., Saifi, G.M., Weinfeld, M., *et al.* Defective DNA single-strand break repair in spinocerebellar ataxia with axonal neuropathy-1. *Nature*, 2005, 434 (7029):108-113
- El-Sayed, N.M., Myler, P.J., Bartholomeu, D.C., *et al.* The genome sequence of *Trypanosoma cruzi*, etiologic agent of Chagas disease. *Science*, 2005a, 309 (5733):409-415
- El-Sayed, N.M., Myler, P.J., Blandin, G., *et al.* Comparative genomics of trypanosomatid parasitic protozoa. *Science*, 2005b, 309 (5733):404-409

- Farkash, E.A., Kao, G.D., Horman, S.R., *et al.* Gamma radiation increases endonuclease-dependent L1 retrotransposition in a cultured cell assay. *Nucleic Acids Res*, 2006, 34 (4):1196-1204
- Filkov, V., Skiena, S. and Zhi, J. Analysis techniques for microarray time-series data. *J Comput Biol*, 2002, 9 (2):317-330
- Fitzgerald, M.P., Madsen, J.M., Coleman, M.C., *et al.* Transgenic biosynthesis of trypanothione protects *Escherichia coli* from radiation-induced toxicity. *Radiat Res*, 2010, 174 (3):290-296
- Gao, X., Pu, D.Q. and Song, P.X. Transition dependency: a gene-gene interaction measure for times series microarray data. *EURASIP J Bioinform Syst Biol*, 2009, 535869
- Garg, N., Popov, V.L. and Papaconstantinou, J. Profiling gene transcription reveals a deficiency of mitochondrial oxidative phosphorylation in *Trypanosoma cruzi*-infected murine hearts: implications in chagasic myocarditis development. *Biochim Biophys Acta*, 2003, 1638 (2):106-120
- Geoghegan, D.S., Skinner-Adams, T. and Davis, T.M. *Plasmodium falciparum*: isolate-specific radiosensitivity. *Exp Parasitol*, 2001, 99 (2):108-110
- Giese, V., Dallagiovanna, B., Marchini, F.K., *et al.* *Trypanosoma cruzi*: a stage-specific calpain-like protein is induced after various kinds of stress. *Mem Inst Oswaldo Cruz*, 2008, 103 (6):598-601
- Gong, B., Chen, Q. and Almasan, A. Ionizing radiation stimulates mitochondrial gene expression and activity. *Radiat Res*, 1998, 150 (5):505-512
- Haile, S. and Papadopoulou, B. Developmental regulation of gene expression in trypanosomatid parasitic protozoa. *Curr Opin Microbiol*, 2007, 10 (6):569-577



- Horn, D. Codon usage suggests that translational selection has a major impact on protein expression in trypanosomatids. *BMC Genomics*, 2008, 9 2
- Hu, L., Wang, J., Baggerly, K., *et al.* Obtaining reliable information from minute amounts of RNA using cDNA microarrays. *BMC Genomics*, 2002, 3 (1):16
- Imai, K., Mimori, T., Kawai, M., *et al.* Microarray analysis of host gene-expression during intracellular nests formation of *Trypanosoma cruzi* amastigotes. *Microbiol Immunol*, 2005, 49 (7):623-631
- Ivens, A.C., Peacock, C.S., Worthey, E.A., *et al.* The genome of the kinetoplastid parasite, *Leishmania major*. *Science*, 2005, 309 (5733):436-442
- Janzen, C.J., Lander, F., Dreesen, O., *et al.* Telomere length regulation and transcriptional silencing in KU80-deficient *Trypanosoma brucei*. *Nucleic Acids Res*, 2004, 32 (22):6575-6584
- Jiang, C.Z., Yen, C.N., Cronin, K., *et al.* UV- and gamma-radiation sensitive mutants of *Arabidopsis thaliana*. *Genetics*, 1997, 147 (3):1401-1409
- Kappagoda, C.T., Knapp, M.F., Linden, R.J., *et al.* Diuresis from left atrial receptors: effect of plasma on the secretion of the Malpighian tubules of *Rhodnius prolixus*. *J Physiol*, 1979, 291 381-391
- Kerkhoven, R.M., Sie, D., Nieuwland, M., *et al.* The T7-primer is a source of experimental bias and introduces variability between microarray platforms. *PLoS One*, 2008, 3 (4):e1980
- Kimura, S., Ishidou, E., Kurita, S., *et al.* DNA microarray analyses reveal a post-irradiation differential time-dependent gene expression profile in yeast cells exposed to X-rays and gamma-rays. *Biochem Biophys Res Commun*, 2006, 346 (1):51-60

- Kollien, A.H. and Schaub, G.A. The development of *Trypanosoma cruzi* in triatominae. *Parasitol Today*, 2000, 16 (9):381-387
- Kooperberg, C., Fazio, T.G., Delrow, J.J., *et al.* Improved background correction for spotted DNA microarrays. *J Comput Biol*, 2002, 9 (1):55-66
- Kottemann, M., Kish, A., Iloanusi, C., *et al.* Physiological responses of the halophilic archaeon *Halobacterium* sp. strain NRC1 to desiccation and gamma irradiation. *Extremophiles*, 2005, 9 (3):219-227
- Koumandou, V.L., Natesan, S.K., Sergeenko, T., *et al.* The trypanosome transcriptome is remodelled during differentiation but displays limited responsiveness within life stages. *BMC Genomics*, 2008, 9 298
- Krauth-Siegel, R.L. and Comini, M.A. Redox control in trypanosomatids, parasitic protozoa with trypanothione-based thiol metabolism. *Biochim Biophys Acta*, 2008, 1780 (11):1236-1248
- Liang, K.C., Wang, X. and Li, T.H. Robust discovery of periodically expressed genes using the laplace periodogram. *BMC Bioinformatics*, 2009, 10 15
- Liu, Y., Zhou, J., Omelchenko, M.V., *et al.* Transcriptome dynamics of *Deinococcus radiodurans* recovering from ionizing radiation. *Proc Natl Acad Sci U S A*, 2003, 100 (7):4191-4196
- Lopes, D.O., Chamber-Reis, B.L., Regis-Da-Silva, C.G., *et al.* Biochemical studies with DNA polymerase beta and DNA polymerase beta-PAK of *Trypanosoma cruzi* suggest the involvement of these proteins in mitochondrial DNA maintenance. *DNA Repair (Amst)*, 2008, 7 (11):1882-1892
- Macarthur, H., Westfall, T.C., Riley, D.P., *et al.* Inactivation of catecholamines by superoxide gives new insights on the pathogenesis of septic shock. *Proc Natl Acad Sci U S A*, 2000, 97 (17):9753-9758

- Machado-Silva, A., Teixeira, S.M., Franco, G.R., *et al.* Mismatch repair in *Trypanosoma brucei*: heterologous expression of MSH2 from *Trypanosoma cruzi* provides new insights into the response to oxidative damage. *Gene*, 2008, 411 (1-2):19-26
- Machado, C.R., Augusto-Pinto, L., Mcculloch, R., *et al.* DNA metabolism and genetic diversity in Trypanosomes. *Mutat Res*, 2006, 612 (1):40-57
- Martinez-Calvillo, S., Yan, S., Nguyen, D., *et al.* Transcription of *Leishmania major* Friedlin chromosome 1 initiates in both directions within a single region. *Mol Cell*, 2003, 11 (5):1291-1299
- Mattimore, V. and Battista, J.R. Radioresistance of *Deinococcus radiodurans*: functions necessary to survive ionizing radiation are also necessary to survive prolonged desiccation. *J Bacteriol*, 1996, 178 (3):633-637
- Mccarthy, F.M., Wang, N., Magee, G.B., *et al.* AgBase: a functional genomics resource for agriculture. *BMC Genomics*, 2006, 7 229
- Minning, T.A., Bua, J., Garcia, G.A., *et al.* Microarray profiling of gene expression during trypomastigote to amastigote transition in *Trypanosoma cruzi*. *Mol Biochem Parasitol*, 2003, 131 (1):55-64
- Minning, T.A., Weatherly, D.B., Atwood, J., 3rd, *et al.* The steady-state transcriptome of the four major life-cycle stages of *Trypanosoma cruzi*. *BMC Genomics*, 2009, 10 370
- Moore-Lai, D. and Rowland, E. Microarray data demonstrate that *Trypanosoma cruzi* downregulates the expression of apoptotic genes in BALB/c fibroblasts. *J Parasitol*, 2004, 90 (4):893-895

- Moore, J.K. and Haber, J.E. Cell cycle and genetic requirements of two pathways of nonhomologous end-joining repair of double-strand breaks in *Saccharomyces cerevisiae*. *Mol Cell Biol*, 1996, 16 (5):2164-2173
- Mukherjee, S., Belbin, T.J., Spray, D.C., *et al.* Microarray analysis of changes in gene expression in a murine model of chronic chagasic cardiomyopathy. *Parasitol Res*, 2003, 91 (3):187-196
- Mukherjee, S., Nagajyothi, F., Mukhopadhyay, A., *et al.* Alterations in myocardial gene expression associated with experimental *Trypanosoma cruzi* infection. *Genomics*, 2008, 91 (5):423-432
- Mukhopadhyay, N.D. and Chatterjee, S. Causality and pathway search in microarray time series experiment. *Bioinformatics*, 2007, 23 (4):442-449
- Muller, S., Liebau, E., Walter, R.D., *et al.* Thiol-based redox metabolism of protozoan parasites. *Trends Parasitol*, 2003, 19 (7):320-328
- Murta, S.M., Krieger, M.A., Montenegro, L.R., *et al.* Deletion of copies of the gene encoding old yellow enzyme (TcOYE), a NAD(P)H flavin oxidoreductase, associates with in vitro-induced benznidazole resistance in *Trypanosoma cruzi*. *Mol Biochem Parasitol*, 2006, 146 (2):151-162
- Nde, P.N., Johnson, C.A., Pratap, S., *et al.* Gene network analysis during early infection of human coronary artery smooth muscle cells by *Trypanosoma cruzi* and Its gp83 ligand. *Chem Biodivers*, 2010, 7 (5):1051-1064
- Nogami, M., Huang, J.T., James, S.J., *et al.* Mice chronically exposed to low dose ionizing radiation possess splenocytes with elevated levels of HSP70 mRNA, HSC70 and HSP72 and with an increased capacity to proliferate. *Int J Radiat Biol*, 1993, 63 (6):775-783

- Nozaki, T. and Cross, G.A. Effects of 3' untranslated and intergenic regions on gene expression in *Trypanosoma cruzi*. *Mol Biochem Parasitol*, 1995, 75 (1):55-67
- Nygaard, V., Loland, A., Holden, M., *et al.* Effects of mRNA amplification on gene expression ratios in cDNA experiments estimated by analysis of variance. *BMC Genomics*, 2003, 4 (1):11
- Paba, J., Ricart, C.A., Fontes, W., *et al.* Proteomic analysis of *Trypanosoma cruzi* developmental stages using isotope-coded affinity tag reagents. *J Proteome Res*, 2004a, 3 (3):517-524
- Paba, J., Santana, J.M., Teixeira, A.R., *et al.* Proteomic analysis of the human pathogen *Trypanosoma cruzi*. *Proteomics*, 2004b, 4 (4):1052-1059
- Pagano, M. Cell cycle regulation by the ubiquitin pathway. *FASEB J*, 1997, 11 (13):1067-1075
- Palenchar, J.B. and Bellofatto, V. Gene transcription in trypanosomes. *Mol Biochem Parasitol*, 2006, 146 (2):135-141
- Parodi-Talice, A., Duran, R., Arrambide, N., *et al.* Proteome analysis of the causative agent of Chagas disease: *Trypanosoma cruzi*. *Int J Parasitol*, 2004, 34 (8):881-886
- Parodi-Talice, A., Monteiro-Goes, V., Arrambide, N., *et al.* Proteomic analysis of metacyclic trypomastigotes undergoing *Trypanosoma cruzi* metacyclogenesis. *J Mass Spectrom*, 2007, 42 (11):1422-1432
- Passos-Silva, D.G., Rajao, M.A., Nascimento De Aguiar, P.H., *et al.* Overview of DNA Repair in *Trypanosoma cruzi*, *Trypanosoma brucei*, and *Leishmania major*. *J Nucleic Acids*, 2010, 2010 840768
- Pastink, A., Eeken, J.C. and Lohman, P.H. Genomic integrity and the repair of double-strand DNA breaks. *Mutat Res*, 2001, 480-481 37-50

- Peddada, S.D., Lobenhofer, E.K., Li, L., *et al.* Gene selection and clustering for time-course and dose-response microarray experiments using order-restricted inference. *Bioinformatics*, 2003, 19 (7):834-841
- Pfeiffer, P., Goedecke, W. and Obe, G. Mechanisms of DNA double-strand break repair and their potential to induce chromosomal aberrations. *Mutagenesis*, 2000, 15 (4):289-302
- Phillips, J. and Eberwine, J.H. Antisense RNA Amplification: A Linear Amplification Method for Analyzing the mRNA Population from Single Living Cells. *Methods*, 1996, 10 (3):283-288
- Puskas, L.G., Zvara, A., Hackler, L., Jr., *et al.* RNA amplification results in reproducible microarray data with slight ratio bias. *Biotechniques*, 2002, 32 (6):1330-1334, 1336, 1338, 1340
- Queiroz, R., Benz, C., Fellenberg, K., *et al.* Transcriptome analysis of differentiating trypanosomes reveals the existence of multiple post-transcriptional regulons. *BMC Genomics*, 2009, 10 495
- Rajao, M.A., Passos-Silva, D.G., Darocha, W.D., *et al.* DNA polymerase kappa from *Trypanosoma cruzi* localizes to the mitochondria, bypasses 8-oxoguanine lesions and performs DNA synthesis in a recombination intermediate. *Mol Microbiol*, 2009, 71 (1):185-197
- Regis-Da-Silva, C.G., Freitas, J.M., Passos-Silva, D.G., *et al.* Characterization of the *Trypanosoma cruzi* Rad51 gene and its role in recombination events associated with the parasite resistance to ionizing radiation. *Mol Biochem Parasitol*, 2006, 149 (2):191-200
- Rhee, S.Y., Wood, V., Dolinski, K., *et al.* Use and misuse of the gene ontology annotations. *Nat Rev Genet*, 2008, 9 (7):509-515

- Ricaud, L., Proux, C., Renou, J.P., *et al.* ATM-mediated transcriptional and developmental responses to gamma-rays in Arabidopsis. *PLoS One*, 2007, 2 (5):e430
- Ritchie, M.E., Silver, J., Oshlack, A., *et al.* A comparison of background correction methods for two-colour microarrays. *Bioinformatics*, 2007, 23 (20):2700-2707
- Sacerdot, C., Mercier, G., Todeschini, A.L., *et al.* Impact of ionizing radiation on the life cycle of *Saccharomyces cerevisiae* Ty1 retrotransposon. *Yeast*, 2005, 22 (6):441-455
- Saúde, M.D. Doenças Infecciosas e Parasitárias. 2010,
- Saxena, A., Worthey, E.A., Yan, S., *et al.* Evaluation of differential gene expression in *Leishmania major* Friedlin procyclics and metacyclics using DNA microarray analysis. *Mol Biochem Parasitol*, 2003, 129 (1):103-114
- Schafer, J. and Strimmer, K. An empirical Bayes approach to inferring large-scale gene association networks. *Bioinformatics*, 2005, 21 (6):754-764
- Scharpf, R.B., Iacobuzio-Donahue, C.A., Sneddon, J.B., *et al.* When should one subtract background fluorescence in 2-color microarrays? *Biostatistics*, 2007, 8 (4):695-707
- Scheidl, S.J., Nilsson, S., Kalen, M., *et al.* mRNA expression profiling of laser microbeam microdissected cells from slender embryonic structures. *Am J Pathol*, 2002, 160 (3):801-813
- Schliep, A., Schonhuth, A. and Steinhoff, C. Using hidden Markov models to analyze gene expression time course data. *Bioinformatics*, 2003, 19 Suppl 1 i255-263
- Schuchhardt, J., Beule, D., Malik, A., *et al.* Normalization strategies for cDNA microarrays. *Nucleic Acids Res*, 2000, 28 (10):E47

- Seo, M., Chun, D.K., Hong, S.T., *et al.* Influence of heat shock, drugs, and radiation on karyotype of *Leishmania major*. *Korean J Parasitol*, 1993, 31 (3):277-283
- Shigihara, T., Hashimoto, M., Shindo, N., *et al.* Transcriptome profile of *Trypanosoma cruzi*-infected cells: simultaneous up- and down-regulation of proliferation inhibitors and promoters. *Parasitol Res*, 2008, 102 (4):715-722
- Simpson, L., Neckelmann, N., De La Cruz, V.F., *et al.* Comparison of the maxicircle (mitochondrial) genomes of *Leishmania tarentolae* and *Trypanosoma brucei* at the level of nucleotide sequence. *J Biol Chem*, 1987, 262 (13):6182-6196
- Singh, H., Fernandes, T. and Apte, S.K. Unusual radioresistance of nitrogen-fixing cultures of *Anabaena* strains. *J Biosci*, 2010, 35 (3):427-434
- Slupphaug, G., Kavli, B. and Krokan, H.E. The interacting pathways for prevention and repair of oxidative DNA damage. *Mutat Res*, 2003, 531 (1-2):231-251
- Smyth, G.K. Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Stat Appl Genet Mol Biol*, 2004, 3 Article3
- Smyth, G.K. and Speed, T. Normalization of cDNA microarray data. *Methods*, 2003, 31 (4):265-273
- Soares, M.B., De Lima, R.S., Rocha, L.L., *et al.* Gene expression changes associated with myocarditis and fibrosis in hearts of mice with chronic chagasic cardiomyopathy. *J Infect Dis*, 2010, 202 (3):416-426
- Song, J.J., Lee, H.J., Morris, J.S., *et al.* Clustering of time-course gene expression data using functional data analysis. *Comput Biol Chem*, 2007, 31 (4):265-274
- Springer, A.L., Bruhn, D.F., Kinzel, K.W., *et al.* Silencing of a putative inner arm dynein heavy chain results in flagellar immotility in *Trypanosoma brucei*. *Mol Biochem Parasitol*, 2011, 175 (1):68-75



- Takata, M., Sasaki, M.S., Tachiiri, S., *et al.* Chromosome instability and defective recombinational repair in knockout mutants of the five Rad51 paralogs. *Mol Cell Biol*, 2001, 21 (8):2858-2866
- Takeda, G.K., Campos, R., Kieffer, J., *et al.* [Effect of gamma rays on blood forms of *Trypanosoma cruzi*. Experimental study in mice]. *Rev Inst Med Trop Sao Paulo*, 1986, 28 (1):15-18
- Tan, K.S., Leal, S.T. and Cross, G.A. *Trypanosoma brucei* MRE11 is non-essential but influences growth, homologous recombination and DNA double-strand break repair. *Mol Biochem Parasitol*, 2002, 125 (1-2):11-21
- Tanaka, M., Earl, A.M., Howell, H.A., *et al.* Analysis of *Deinococcus radiodurans*'s transcriptional response to ionizing radiation and desiccation reveals novel proteins that contribute to extreme radioresistance. *Genetics*, 2004, 168 (1):21-33
- Tanowitz, H.B., Kirchhoff, L.V., Simon, D., *et al.* Chagas' disease. *Clin Microbiol Rev*, 1992, 5 (4):400-419
- Teixeira, S.M. and Da Rocha, W.D. Control of gene expression and genetic manipulation in the Trypanosomatidae. *Genet Mol Res*, 2003, 2 (1):148-158
- Teixeira, S.M., Kirchhoff, L.V. and Donelson, J.E. Post-transcriptional elements regulating expression of mRNAs from the amastin/tuzin gene cluster of *Trypanosoma cruzi*. *J Biol Chem*, 1995, 270 (38):22586-22594
- Verdun, R.E., Di Paolo, N., Urmenyi, T.P., *et al.* Gene discovery through expressed sequence Tag sequencing in *Trypanosoma cruzi*. *Infect Immun*, 1998, 66 (11):5393-5398
- Vickerman, K. The evolutionary expansion of the trypanosomatid flagellates. *Int J Parasitol*, 1994, 24 (8):1317-1331

- Wang, E., Miller, L.D., Ohnmacht, G.A., *et al.* High-fidelity mRNA amplification for gene profiling. *Nat Biotechnol*, 2000, 18 (4):457-459
- Weston, D., La Flamme, A.C. and Van Voorhis, W.C. Expression of *Trypanosoma cruzi* surface antigen FL-160 is controlled by elements in the 3' untranslated, the 3' intergenic, and the coding regions. *Mol Biochem Parasitol*, 1999, 102 (1):53-66
- Who. Chagas disease (American trypanosomiasis) fact sheet (revised in June 2010). *Wkly Epidemiol Rec*, 2010, 85 (34):334-336
- Wichert, S., Fokianos, K. and Strimmer, K. Identifying periodically expressed transcripts in microarray time series data. *Bioinformatics*, 2004, 20 (1):5-20
- Williams, B. and Lustig, A.J. The paradoxical relationship between NHEJ and telomeric fusion. *Mol Cell*, 2003, 11 (5):1125-1126
- Williams, E., Lowe, T.M., Savas, J., *et al.* Microarray analysis of the hyperthermophilic archaeon *Pyrococcus furiosus* exposed to gamma irradiation. *Extremophiles*, 2007, 11 (1):19-29
- Yang, S.W., Burgin, A.B., Jr., Huizenga, B.N., *et al.* A eukaryotic enzyme that can disjoin dead-end covalent complexes between DNA and type I topoisomerases. *Proc Natl Acad Sci U S A*, 1996, 93 (21):11534-11539
- Yang, Y.H., Dudoit, S., Luu, P., *et al.* Normalization for cDNA microarray data: a robust composite method addressing single and multiple slide systematic variation. *Nucleic Acids Res*, 2002, 30 (4):e15
- Yonetani, Y., Hochegger, H., Sonoda, E., *et al.* Differential and collaborative actions of Rad51 paralog proteins in cellular response to DNA damage. *Nucleic Acids Res*, 2005, 33 (14):4544-4552

- Zhang, S., Kim, C.C., Batra, S., *et al.* Delineation of diverse macrophage activation programs in response to intracellular parasites and cytokines. *PLoS Negl Trop Dis*, 2010, 4 (3):e648
- Zhao, H., Hastie, T., Whitfield, M.L., *et al.* Optimization and evaluation of T7 based RNA linear amplification protocols for cDNA microarray analysis. *BMC Genomics*, 2002, 3 (1):31
- Zhu, D., Hero, A.O., Qin, Z.S., *et al.* High throughput screening of co-expressed gene pairs with controlled false discovery rate (FDR) and minimum acceptable strength (MAS). *J Comput Biol*, 2005, 12 (7):1029-1045
- Zingales, B., Andrade, S.G., Briones, M.R., *et al.* A new consensus for *Trypanosoma cruzi* intraspecific nomenclature: second revision meeting recommends TcI to TcVI. *Mem Inst Oswaldo Cruz*, 2009, 104 (7):1051-1054

## 8 - ANEXOS

---

### Anexo 1: Script em linguagem estatística R desenvolvido para as análises das lâminas de microarranjo usadas neste trabalho

---

```
#Carregamento das bibliotecas usadas para as análises
library(limma)
library(marray)
library(convert)
library(DAAG)

#Carregando os arquivos de entrada. Estrutura do arquivo (tabular):

targets <- readTargets("Targets.txt")

>targets

      SlideNumber  Name  FileName  Cy3  Cy5  Date
1      890      N1    C_Cy3x0hs_Cy5_RI.gpr  WT_RI  0hs_RI  05/Dez/2008
2      888      N2    0hs_Cy3xC_Cy5_RI.gpr  0hs_RI  WT_RI  05/Dez/2009
3      872      N3    C_Cy3x4hs_Cy5_RI.gpr  WT_RI  4hs_RI  05/Fev/2009
4      873      N4    4hs_Cy3XC_Cy5_RI.gpr  4hs_RI  WT_RI  05/Fev/2009
5      871      N5    C_Cy3X24hs_Cy5_RI.gpr  WT_RI  24hs_RI  06/Fev/2009
6      876      N6    24hs_Cy3XC_Cy5_RI.gpr  24hs_RI  WT_RI  06/Fev/2009
7      8423     N7    C_Cy3X48hs_Cy5_RI.gpr  WT_RI  48hs_RI  27/Mai/2009
8      8424     N8    48hs_Cy3xC_Cy5_RI.gpr  48hs_RI  WT_RI  29/Mai/2009
9      8357     N9    C_Cy3X96hs_Cy5_RI.gpr  WT_RI  96hs_RI  09/Jun/2009
10     835     N10    96hs_Cy3xC_Cy5_RI.gpr  96hs_RI  WT_RI  09/Jun/2009
11     881     N11    C_Cy3X0hs_Cy5_RII.gpr  WT_RI  0hs_RII  28/Jan/2009
12     456     N12    0hs_Cy3XC_Cy5_RII.gpr  0hs_RII  WT_RII  28/Jan/2009
13     877     N13    C_Cy3X4hs_Cy5_RII.gpr  WT_RII  4hs_RII  29/Jan/2009
14     880     N14    4hs_Cy3XC_Cy5_RII.gpr  4hs_RII  WT_RII  29/Jan/2009
15     868     N15    C_Cy3X24hs_Cy5_RI.gpr  WT_RII  24hs_RII  12/Fev/2009
16     869     N16    24hs_Cy3XC_Cy5_RII.gpr  24hs_RII  WT_RII  12/Fev/2009
17     8355     N17    C_Cy3X48hs_Cy5_RII.gpr  WT_RII  48hs_RII  04/Jun/2009
18     8356     N18    48hs_Cy3XC_Cy5_RII.gpr  48hs_RII  WT_RII  04/Jun/2009
19     8359     N19    C_Cy3X96hs_Cy5_RII.gpr  WT_RII  96hs_RII  10/Jun/2009
20     8360     N20    96hs_Cy3XC_Cy5_RII.gpr  96hs_RII  WT_RII  10/Jun/2009

#Selecionando dados analisáveis (dados com valor de flag acima de 0)

f <- function(x) as.numeric(x$Flags >= 0)

#Leitura das intensidades

RG <- read.maimages(targets$FileName, columns = list(Rf = "F633 Mean", Gf = "F543
Mean", Rb = "B633 Median", Gb = "B543 Median"), source= "genepix", sep="\t",
annotation = c("Block", "Row", "Column", "ID", "Name"), wt.fun=f)

Read Controle_Cy3x0horas_Cy5_RepI_05Dez2008.gpr
Read 0horas_Cy3xControle_Cy5_RepI_05Dez2008.gpr
Read Controle_Cy3x4horas_Cy5_RepI_05Fev2009.gpr
Read 4horas_Cy3XControle_Cy5_RepI_05Fev2009.gpr
Read Controle_Cy3X24horas_Cy5_RepI_06Fev2009.gpr
Read 24horas_Cy3XControle_Cy5_RepI_06Fev2009.gpr
```

```

Read Controle_Cy3X48horas_Cy5_RepI_27Maio2009.gpr
Read 48horas_Cy3xControle_Cy5_RepI_29Maio2009.gpr
Read Controle_Cy3X96horas_Cy5_RepI_09Jun2009.gpr
Read 96horas_Cy3xControle_Cy5_RepI_09Jun2009.gpr
Read Controle_Cy3X0horas_Cy5_RepII_28Jan2009.gpr
Read 0horas_Cy3XControle_Cy5_RepII_28Jan2009.gpr
Read Controle_Cy3X4horas_Cy5_RepII_29Jan2009.gpr
Read 4horas_Cy3XControle_Cy5_RepII_29Jan2009.gpr
Read Controle_Cy3X24horas_Cy5_RepII_12Fev2009.gpr
Read 24horas_Cy3XControle_Cy5_RepII_12Fev2009.gpr
Read Controle_Cy3X48horas_Cy5_RepII_04Jun2009.gpr
Read 48horas_Cy3XControle_Cy5_RepII_04Jun2009.gpr
Read Controle_Cy3X96horas_Cy5_RepII_10Jun2009.gpr
Read 96horas_Cy3XControle_Cy5_RepII_10Jun2009.gpr

#Leitura do arquivo .gal (arquivo de anotação da lâmina)

RG$genes <- readGAL()

#Determinando o layout da lâmina

RG$printer <- getLayout(RG$genes)

$ngrid.r
[1] 12 #Número de linhas
$ngrid.c
[1] 4 #Número de colunas
$nspt.r
[1] 23 # Número de linhas em um sub-array
$nspt.c
[1] 24 # Número de colunas em um sub-array
attr(,"class")
[1] "PrintLayout"

#Convertendo para um objeto do pacote marray

data.marray <- as(RG, "marrayRaw")

#Correção do background

RGbc <- backgroundCorrect (RG, method="normexp", offset=50)

#Normalização em cada array, método "robustpline"

MA <- normalizeWithinArrays(RGbc, RG$printer, method="robustspline", maxit=27)

#Normalização entre os arrays, método "Aquantile"

MA.aq <- normalizeBetweenArrays (MA, method="Aquantile")

#Convetendo para um objeto do pacore marray

MAaq.marray <- as(MA.aq, "marrayNorm")

###Análise estatística
##Design experimental

design <- cbind(wt1vsmu1 = c(1,-1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0),
wt2vsmu2 = c(0,0,1,-1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0),
wt3vsmu3 = c(0,0,0,0,1,-1,0,0,0,0,0,0,0,0,0,0,0,0,0,0),
wt4vsmu4 = c(0,0,0,0,0,0,1,-1,0,0,0,0,0,0,0,0,0,0,0,0),
wt5vsmu5=c(0,0,0,0,0,0,0,0,1,-1,0,0,0,0,0,0,0,0,0,0),
wt6vsmu6=c(0,0,0,0,0,0,0,0,0,0,1,-1,0,0,0,0,0,0,0,0),
wt7vsmu7=c(0,0,0,0,0,0,0,0,0,0,0,0,1,-1,0,0,0,0,0,0),
wt8vsmu8=c(0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,-1,0,0,0,0),
wt9vsmu9=c(0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,-1,0,0),
wt10vsmu10=c(0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,-1))

```

```

#Matriz de contrastes

wtvs0horas <- makeContrasts (wtvs0hs = (wt1vsu1+wt6vsu6)/2, levels=design)
wtvs4horas <- makeContrasts (wtvs4hs = (wt2vsu2+wt7vsu7)/2, levels=design)
wtvs24horas <- makeContrasts (wtvs24hs = (wt3vsu3+wt8vsu8)/2, levels=design)
wtvs48horas <- makeContrasts (wtvs48hs = (wt4vsu4+wt9vsu9)/2, levels=design)
wtvs96horas<- makeContrasts (wtvs96hs = (wt5vsu5+wt10vsu10)/2, levels=design)

#Determinando as replicas biológicas

biolrep<- c(1,1,1,1,1,1,1,1,1,1,2,2,2,2,2,2,2,2,2,2)
#Correlação entre os spots

corfit <- duplicateCorrelation(MA.aq, design, ndups=2, spacing=12)

#Aplicando o modelo linear

fit <- lmFit(MA.aq, design, ndups=2, spacing=12, cor=corfit$consensus)

#CLB x 0horas
fit0hs <- contrasts.fit(fit, wtvs0horas)
fit0hs.eb <- eBayes(fit0hs)
ordinary.t_0horas = fit0hs.eb$coef/fit0hs.eb$stdev.unscaled/fit$sigma
topt_0hs = order(abs(ordinary.t_0horas), decreasing = T)[1:200]
toptable_results_0horas <- topTable
(fit0hs.eb,coef=NULL,number=30000,genelist=fit$genes,adjust.method="BH",sort.by="B"
, resort.by=NULL,p.value=1,lfc=0)
write.table(toptable_results_0horas, file="CLBx0horas.xls", sep="\t")

#CLB x 4horas
fit4hs <- contrasts.fit(fit, wtvs4horas)
fit4hs.eb <- eBayes(fit4hs)
ordinary.t_4horas = fit4hs.eb$coef/fit4hs.eb$stdev.unscaled/fit$sigma
topt_4hs = order(abs(ordinary.t_4horas), decreasing = T)[1:200]
toptable_results_4horas <- topTable
(fit4hs.eb,coef=NULL,number=30000,genelist=fit$genes,adjust.method="BH",sort.by="B"
, resort.by=NULL,p.value=1,lfc=0)
write.table(toptable_results_4horas, file="CLBx4horas.xls", sep="\t")

#CLB x 24horas
fit24hs <- contrasts.fit(fit, wtvs24horas)
fit24hs.eb <- eBayes(fit24hs)
ordinary.t_24horas = fit24hs.eb$coef/fit24hs.eb$stdev.unscaled/fit$sigma
topt_24hs = order(abs(ordinary.t_24horas), decreasing = T)[1:200]
toptable_results_24horas <- topTable
(fit24hs.eb,coef=NULL,number=30000,genelist=fit$genes,adjust.method="BH",sort.by="B"
, resort.by=NULL,p.value=1,lfc=0)
write.table(toptable_results_24horas, file="CLBx24horas.xls", sep="\t")

#CLB x 48horas
fit48hs <- contrasts.fit(fit, wtvs48horas)
fit48hs.eb <- eBayes(fit48hs)
ordinary.t_48horas = fit48hs.eb$coef/fit48hs.eb$stdev.unscaled/fit$sigma
topt_48hs = order(abs(ordinary.t_48horas), decreasing = T)[1:200]
toptable_results_48horas <- topTable
(fit48hs.eb,coef=NULL,number=30000,genelist=fit$genes,adjust.method="BH",sort.by="B"
, resort.by=NULL,p.value=1,lfc=0)
write.table(toptable_results_48horas, file="CLBx48horas.xls", sep="\t")

#CLB x 96horas
fit96hs <- contrasts.fit(fit, wtvs96horas)
fit96hs.eb <- eBayes(fit96hs)
ordinary.t_96horas = fit96hs.eb$coef/fit96hs.eb$stdev.unscaled/fit$sigma
topt_96hs = order(abs(ordinary.t_96horas), decreasing = T)[1:200]
toptable_results_96horas <- topTable
(fit96hs.eb,coef=NULL,number=30000,genelist=fit$genes,adjust.method="BH",sort.by="B"
, resort.by=NULL,p.value=1,lfc=0)
write.table(toptable_results_96horas, file="CLBx96horas.xls", sep="\t")

```

Anexo 2: Escolha da melhor associação de métodos para a correção do *background* e normalização dos dados.

---

Conforme explicado na metodologia deste trabalho, testamos dois métodos de correção de *background* e quatro métodos para normalização dos dados dentro de cada lâmina. A decisão final sobre a melhor associação de métodos foi em função da análise visual das imagens de gráficos MA-plot e boxplots. Foram testadas as seguintes associações:

Normalização \ <i>Background</i>	Loess (L)	Print-tip Loess (P)	Median (M)	Robust spline (R)
<i>Subtract</i> (S)	SL	SP	SM	SR
<i>NormExp</i> (N)	NL	NP	NM	NR

Os resultados estão sumarizados nas figuras abaixo:

A) MA-plots para as intensidades da lâmina #1, mostrando a dispersão dos pontos dos dados brutos e após a aplicação das oito combinações. Detalhe para as linhas coloridas centrais. Quanto mais próximo do eixo zero horizontal, melhor foi o processo de normalização.

B) MA-plots das 20 lâminas usadas neste trabalho antes e após a associação dos métodos *NormExp* + Robust spline para a normalização dos dados.

C) Distribuição empírica suavizada dos sinais dos canais verde e vermelho para todas as lâminas dos dados brutos e após a aplicação das oito combinações.

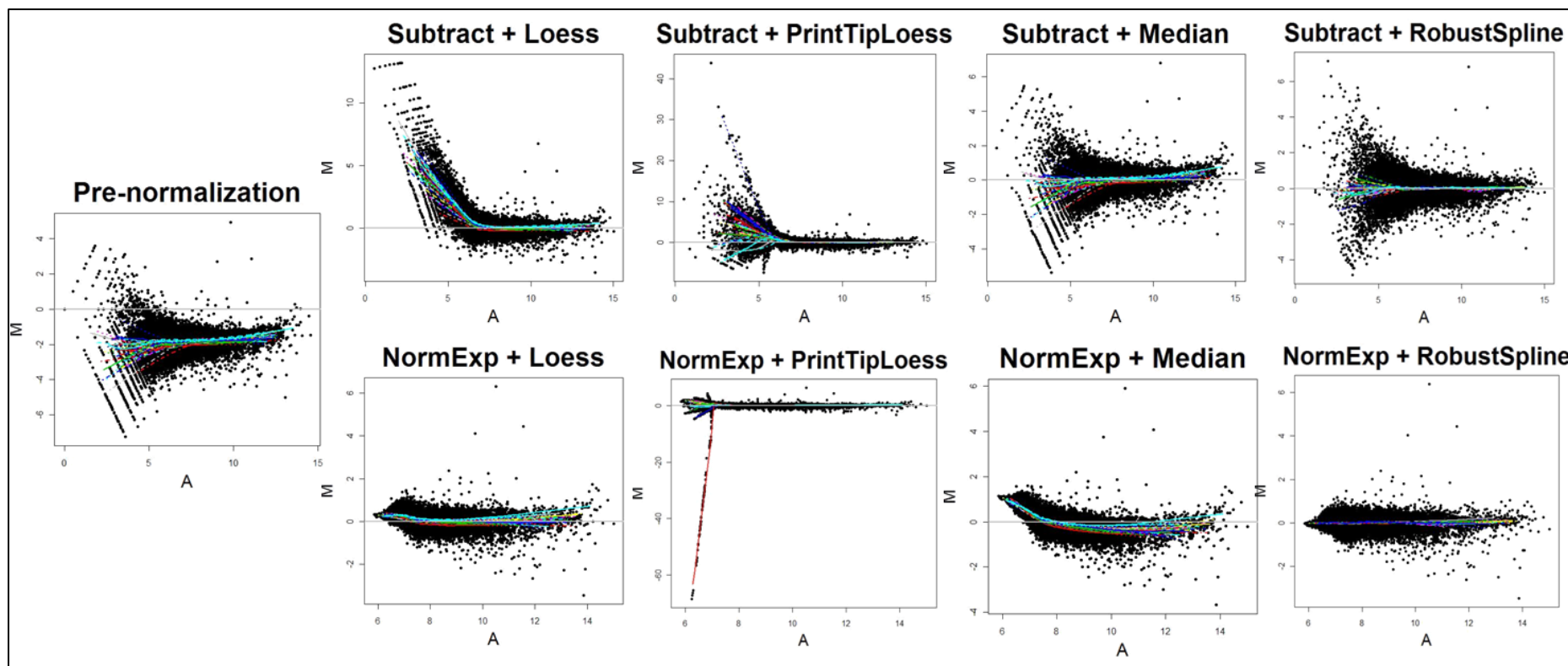


Figura A1: *MA plots* para a lâmina 1 com diferentes métodos de normalização. O M representa o  $\log_2$  da razão entre os sinais do canal vermelho (R) e verde (G) ( $\log_2(R/G)$ ) e o A representa a média da soma dos  $\log_2$  dos canais vermelho e verde ( $(\log_2R + \log_2G)/2$ ). As linhas coloridas representam a mediana da intensidade de cada agulha. O primeiro gráfico representa os dados antes da normalização. Os gráficos seguintes representam os dados normalizados com os métodos *Loess*, *Print tip loess*, *Median*, e o *Robust Spline*, respectivamente.



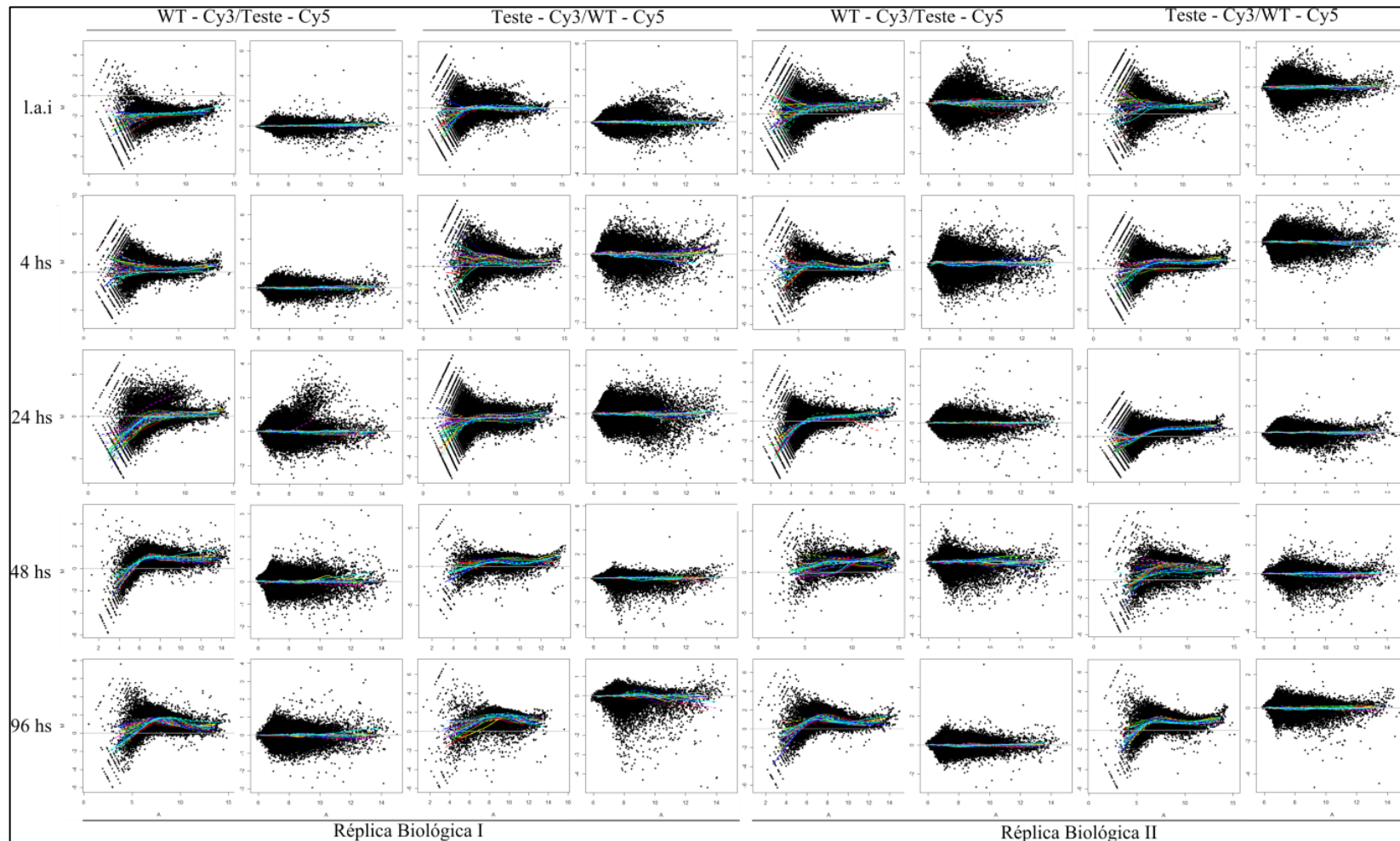


Figura A2: Comparação entre dados brutos e normalizados com o método *Robust-spline*. *MA plots* para as 20 lâminas usadas neste trabalho. 1.a.i = logo após a irradiação. WT = células controles (não-irradiadas).

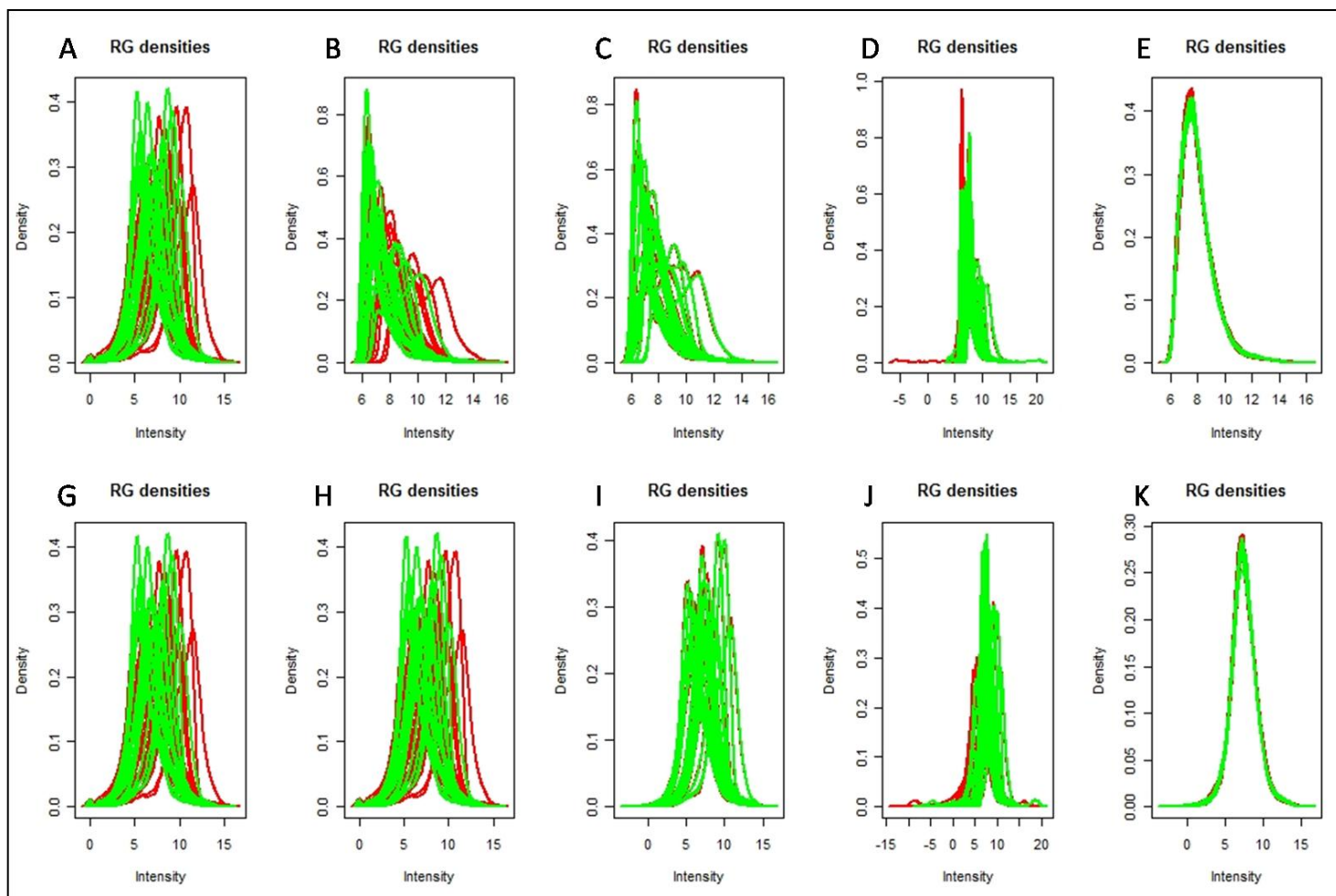


Figura A3: Gráfico de densidades de canais verdes e vermelhos. A e G) Dados brutos; B) Dados com correção de *background* com *NormExp*; C) *NormExp* + *Robust spline*; D) *NormExp* + *Loess*; E) *NormExp* + *Robust spline* + *Aquantile*; H) Dados com correção de *background* com *Subtract*; I) *Subtract* + *Robust spline*; J) *Subtract* + *Loess*; K) *Subtract* + *Robust spline* + *Aquantile*

Anexo 3 – Lista dos genes formadores dos clusters definidos pelo *heatmap*

Tabela A1: Genes pertencentes a cada um dos grupos baseados nos perfis de expressão gênica da **Erro! Fonte de referência não encontrada.B**

#	Genes (with ID number)
Grupo 1	<b>GP85</b> (Tc00.1047053510199.10), <b>acetyltransferase</b> (Tc00.1047053507611.290), <b>ATPase, putative</b> (Tc00.1047053505997.70 e Tc00.1047053508903.100), <b>calpain-like cysteine peptidase (pseudogene)</b> (Tc00.1047053506563.170), <b>cation transporter protein, putative</b> (Tc00.1047053508357.80), <b>cyclin, putative</b> (Tc00.1047053509455.140), <b>dynein heavy chain</b> (Tc00.1047053509585.10 e Tc00.1047053510687.10), <b>dynein light chain (lc6)</b> (Tc00.1047053510897.6), <b>glycine cleavage system H protein, putative</b> , (Tc00.1047053508457.30), <b>nucleoside diphosphate kinase, putative</b> (Tc00.1047053508461.400), <b>protein kinase, putative</b> (Tc00.1047053504113.10 e Tc00.1047053511671.80), <b>protein kinase A catalytic subunit, putative</b> (Tc00.1047053511269.50), <b>serine/threonine protein kinase, putative</b> (Tc00.1047053510121.130), <b>trans-sialidase</b> (Tc00.1047053509907.60), <b>UDP-glucose dehydrogenase, putative</b> (Tc00.1047053510105.100), <b>hypothetical proteins</b> (Tc00.1047053511121.30, Tc00.1047053508061.10, Tc00.1047053508293.130, Tc00.1047053510479.30, Tc00.1047053511611.30, Tc00.1047053506133.179, Tc00.1047053507521.60, Tc00.1047053509571.30, Tc00.1047053510599.40, Tc00.1047053506713.24, Tc00.1047053511239.60, Tc00.1047053508741.390, Tc00.1047053507275.24, Tc00.1047053506947.50, Tc00.1047053511727.220, Tc00.1047053507083.70, Tc00.1047053507159.30, Tc00.1047053511523.50, Tc00.1047053511517.130, Tc00.1047053507611.270, Tc00.1047053511643.80, Tc00.1047053507165.30, Tc00.1047053478463.10), <b>obsolete sequences</b> (21055, 2117, 11281, 21056, 18649, 1404, 1497, 13828, 23196, 16329, 13999), <b>RHS</b> (Tc00.1047053506537.10, Tc00.1047053509765.60, Tc00.1047053508877.30, Tc00.1047053506683.210, Tc00.1047053506751.70, Tc00.1047053506129.80, Tc00.1047053506561.20, Tc00.1047053507975.10, Tc00.1047053509559.20, Tc00.1047053413293.30, Tc00.1047053506951.10, Tc00.1047053445777.10, Tc00.1047053506349.83, Tc00.1047053508071.90, Tc00.1047053509259.180, Tc00.1047053504099.70, Tc00.1047053508483.40)
Grupo 2	<b>amal protein, putative</b> (Tc00.1047053504071.110), <b>ARP2/3 complex subunit, putative</b> (Tc00.1047053508737.194), <b>calpain-like cysteine peptidase, putative</b> (Tc00.1047053506563.210), <b>developmentally regulated phosphoprotein, putative</b> (Tc00.1047053511421.110), <b>EF-hand protein 5, putative</b> (Tc00.1047053506391.30 e Tc00.1047053507483.20), <b>2,4-dienoyl-coa reductase FADH1, putative</b> (Tc00.1047053509941.100), <b>glucose-6-phosphate isomerase, glycosomal, putative</b> (Tc00.1047053506529.508), <b>lactoylglutathione lyase-like protein, putative</b> (Tc00.1047053510659.240), <b>protein kinase C substrate protein, heavy chain, putative</b> (Tc00.1047053508215.9), <b>tryparedoxin peroxidase, putative</b> (Tc00.1047053508649.5), <b>hypothetical proteins</b> (Tc00.1047053509601.140, Tc00.1047053503415.40, Tc00.1047053506825.200, Tc00.1047053503911.30, Tc00.1047053508479.330, Tc00.1047053507491.140, Tc00.1047053506773.104, Tc00.1047053511391.110, Tc00.1047053509393.10, Tc00.1047053509001.30, Tc00.1047053510381.40, Tc00.1047053511421.90, Tc00.1047053506831.63, Tc00.1047053503925.6, Tc00.1047053508153.364, Tc00.1047053506739.99, Tc00.1047053503975.100, Tc00.1047053508707.300, Tc00.1047053504005.54, Tc00.1047053511755.19, Tc00.1047053508059.50, Tc00.1047053504037.10), <b>obsolete sequence</b> (19423), <b>RHS</b> (Tc00.1047053505915.10, Tc00.1047053507427.30)
Grupo 3	<b>Alcohol dehydrogenase, putative</b> (Tc00.1047053506357.50), <b>aspartate carbamoyltransferase, putative</b> (Tc00.1047053508375.30), <b>aspartyl aminopeptidase, putative</b> (Tc00.1047053508183.4), <b>co-chaperone GrpE, putative</b> (Tc00.1047053507929.20), <b>lactoylglutathione</b>

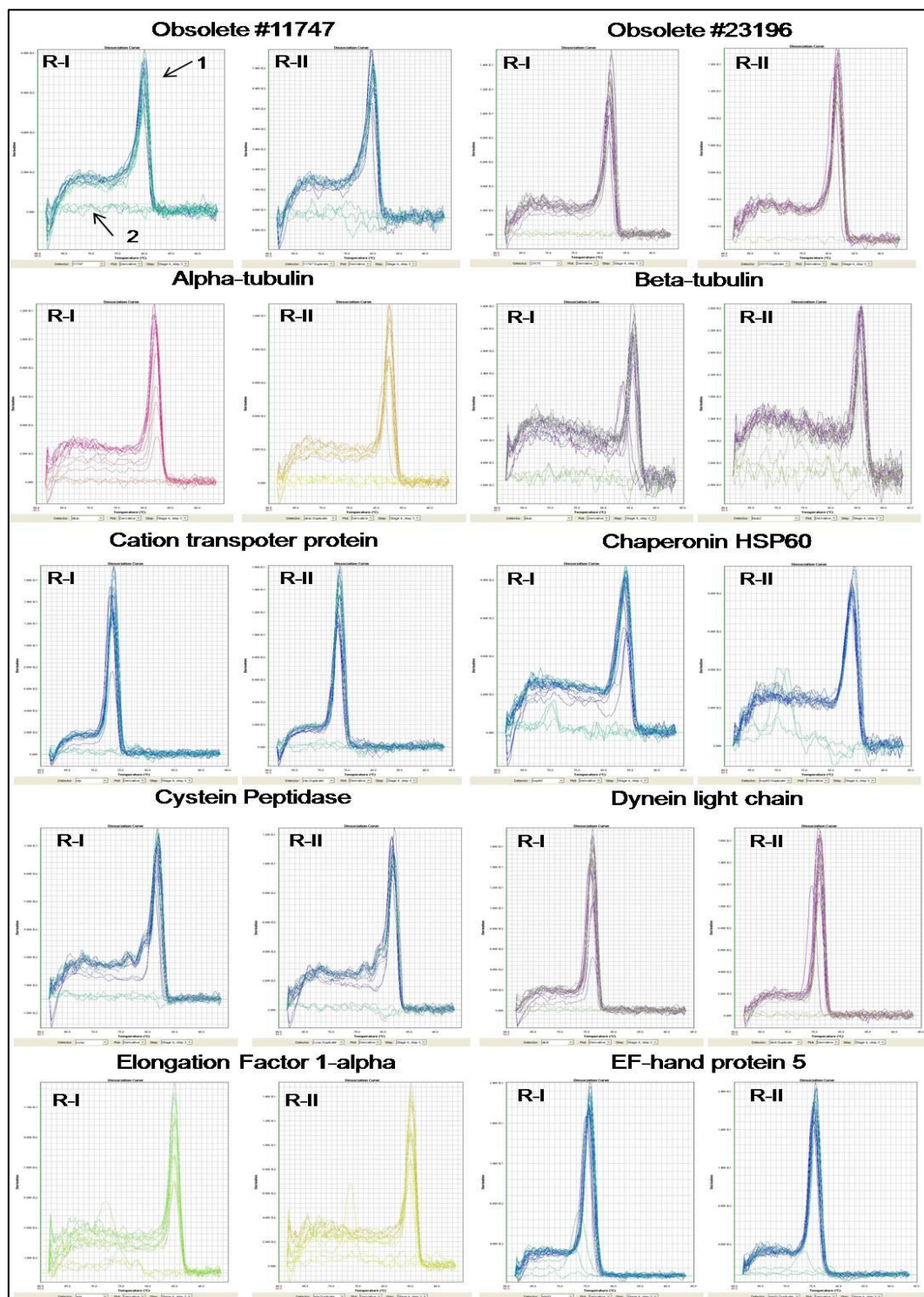
	<p><b>lyase-like protein, putative</b> (Tc00.1047053510743.70), <b>tyrosyl-DNA phosphodiesterase, putative</b> (Tc00.1047053506619.40), <b>vacuolar sorting protein, putative</b> (Tc00.1047053508479.290), <b>hypothetical proteins</b> (Tc00.1047053504227.10, Tc00.1047053511467.60);</p> <p style="text-align: right;">(Continuação)</p> <p>Tc00.1047053503839.19, Tc00.1047053503419.54, Tc00.1047053506789.270, Tc00.1047053507875.30, Tc00.1047053508909.160, Tc00.1047053503841.20, Tc00.1047053503395.20, Tc00.1047053508879.10, Tc00.1047053509245.29, Tc00.1047053506925.310), <b>obsolete sequences</b> (18850, 13664, 21269), <b>RHS</b> (Tc00.1047053504285.10, Tc00.1047053507777.30)</p>
Grupo 4	<p><b>Glucose-6-phosphate isomerase, glycosomal, putative</b> (Tc00.1047053506529.508), <b>lathosterol oxidase, putative</b> (Tc00.1047053473111.10), <b>tubulin binding cofactor A-like protein, putative</b> (Tc00.1047053509069.30), <b>hypothetical proteins</b> (Tc00.1047053510535.10, Tc00.1047053509029.20, Tc00.1047053507083.10, Tc00.1047053506829.90, Tc00.1047053504433.20), <b>obsolete sequences</b> (11747, 6364)</p>
Grupo 5	<p><b>Hypothetical proteins</b> (Tc00.1047053508319.10, Tc00.1047053509073.60), <b>obsolete sequence</b> (5884)</p>
Grupo 6	<p><b>Hypothetical proteins</b> (Tc00.1047053510835.20, Tc00.1047053510173.120, Tc00.1047053511201.40, Tc00.1047053509627.10, Tc00.1047053511671.140), <b>obsolete sequences</b> (9204, 14862, 14569)</p>
Grupo 7	<p><b>ATP-binding cassette transporter ABCA1, putative</b> (Tc00.1047053504149.20), <b>N-acetylglucosamine-6-phosphate deacetylase-like protein, putative</b> (Tc00.1047053506341.10), <b>trypanothione/tryparedoxin dependent peroxidase 2, putative</b> (Tc00.1047053503899.119), <b>hypothetical protein</b> (Tc00.1047053506195.290), <b>obsolete sequence</b> (6361, 26458)</p>
Grupo 8	<p><b>flagellar calcium-binding protein, putative</b> (Tc00.1047053507891.47), <b>hypothetical proteins</b> (Tc00.1047053509141.40, Tc00.1047053506885.30, Tc00.1047053508257.180)</p>
Grupo 9	<p><b>3-oxo-5-alpha-steroid 4-dehydrogenase, putative</b> (Tc00.1047053504427.70), <b>eukaryotic translation initiation factor 3 subunit 7-like protein, putative</b> (Tc00.1047053506943.160), <b>GPR1/FUN34/yaaH family, putative</b> (Tc00.1047053508179.70), <b>kinetoplast DNA-associated protein, putative</b> (Tc00.1047053509793.10), <b>membrane transporter protein, putative</b> (Tc00.1047053510667.14), <b>nucleolar protein, putative</b> (Tc00.1047053511573.58), <b>phosphatidic acid phosphatase protein, putative</b> (Tc00.1047053511355.30), <b>pumilio/PUF RNA binding protein 7, putative</b> (Tc00.1047053511715.100), <b>UDP-Gal or UDP-GlcNAc-dependent glycosyltransferase, putative</b> (Tc00.1047053503487.50), <b>hypothetical proteins</b> (Tc00.1047053510543.80, Tc00.1047053509755.89, Tc00.1047053507053.140, Tc00.1047053508307.90, Tc00.1047053503809.120, Tc00.1047053510857.30)</p>
Grupo 10	<p><b>COP-coated vesicle membrane protein erv25 precursor, putative</b> (Tc00.1047053510187.270), <b>elongation initiation factor 2 alpha subunit, putative</b> (Tc00.1047053508153.730), <b>eukaryotic translation initiation factor 1A, putative</b> (Tc00.1047053463269.10), <b>mevalonate kinase, putative (2 cópias) (Tc00.1047053436521.9)</b>, <b>mitochondrial RNA editing ligase 1, putative</b> (Tc00.1047053510155.20), <b>nuclear transport factor 2 protein(NFT2), putative</b> (Tc00.1047053508173.180), <b>nuclear transcription factor, putative</b> (Tc00.1047053510645.20), <b>trypanothione reductase, putative</b> (Tc00.1047053503555.30), <b>hypothetical proteins</b> (Tc00.1047053507943.40, Tc00.1047053507053.180, Tc00.1047053508173.264, Tc00.1047053509267.40, Tc00.1047053503687.30, Tc00.1047053503897.120, Tc00.1047053509769.60, Tc00.1047053503703.30, Tc00.1047053511071.50), <b>obsolete sequences</b> (23161, 5692, 8330)</p>

(Continuação)

Grupo 11 **26S proteasome regulatory non-ATPase subunit, putative** (Tc00.1047053504221.20), **40S ribosomal protein S10** (Tc00.1047053506679.150), **40S ribosomal protein S21** (Tc00.1047053510101.430), **40S ribosomal protein S3** (Tc00.1047053509353.30), **40S ribosomal protein S5** (Tc00.1047053506297.150), **40S ribosomal protein SA** (Tc00.1047053510425.19 e Tc00.1047053503719.20), **60S ribosomal protein L21E** (Tc00.1047053507251.20), **60S ribosomal protein L23** (Tc00.1047053508461.490), **60S ribosomal protein L28** (Tc00.1047053506297.270), **60S ribosomal protein P2** (Tc00.1047053505977.26), **cell differentiation protein, putative** (Tc00.1047053507873.20), **haloacid dehalogenase-like hydrolase, putative** (Tc00.1047053510131.40), **heat shock 70 kDa protein, mitochondrial precursor, putative** (Tc00.1047053511745.10), **chaperonin HSP60, mitochondrial precursor (pseudogene)** (Tc00.1047053510187.420), **D-isomer specific 2-hydroxyacid dehydrogenase-protein, putative** (Tc00.1047053510099.120), **monooxygenase, putative** (Tc00.1047053508173.100), **nucleoside diphosphate kinase, putative** (Tc00.1047053508707.200), **activated protein kinase C receptor, putative** (Tc00.1047053511211.120), **succinyl-CoA synthetase alpha subunit, putative** (Tc00.1047053508479.340), **hypothetical proteins** (Tc00.1047053508719.30, Tc00.1047053504057.80, Tc00.1047053507993.380, Tc00.1047053510099.100, Tc00.1047053447255.10, Tc00.1047053507221.30), **obsolete sequences** (3577, 5161, 4685)

Grupo 12 **2- aminoethylphosphonate:pyruvateaminotransferase-like protein, putative** (Tc00.1047053509693.100), **aminopeptidase P, putative** (Tc00.1047053510655.120), **alpha-tubulin, putative** (Tc00.1047053411235.9), **beta-tubulin, putative** (Tc00.1047053506563.40), **carboxypeptidase, putative** (Tc00.1047053504153.160), **chaperonin, putative** (Tc00.1047053506247.50), **cysteine peptidase, putative** (Tc00.1047053506529.550 e Tc00.1047053508317.10), **cytochrome c1, heme protein, mitochondrial precursor, putative** (Tc00.1047053511391.160), **elongation factor 1-alpha (EF-1-alpha), putative** (Tc00.1047053510119.9), **eukaryotic translation initiation factor 6 (eIF-6), putative** (Tc00.1047053506679.70), **farnesyl pyrophosphate synthase, putative** (Tc00.1047053511823.70), **fatty acid desaturase, putative** (Tc00.1047053511075.9), **fatty acid elongase, putative** (Tc00.1047053506661.30), **glycosomal phosphoenolpyruvate carboxykinase, putative** (Tc00.1047053507547.90), **hexose transporter, putative** (Tc00.1047053511041.40), **chaperonin HSP60, mitochondrial precursor** (Tc00.1047053507641.280), **NADH-cytochrome B5 reductase, putative** (Tc00.1047053511817.40), **nucleoside transporter-like, putative** (Tc00.1047053506773.50), **pyruvate dehydrogenase E1 beta subunit, putative** (Tc00.1047053510091.80), **prostaglandin F2alpha synthase** (Tc00.1047053507617.9 e Tc00.1047053508461.80), **nucleolar RNA-binding protein, putative** (Tc00.1047053510859.17), **GTP-binding nuclear protein rtb2, putative** (Tc00.1047053503539.30), **25 kDa translation elongation factor 1-beta** (Tc00.1047053507671.30), **thermostable carboxypeptidase 1, putative** (Tc00.1047053504045.60), **thymidine kinase, putative** (Tc00.1047053506855.260), **tryptophanyl-tRNA synthetase (pseudogene)** (Tc00.1047053508421.30), **protein tyrosine phosphatase, putative** (Tc00.1047053511635.40), **V-type ATPase, A subunit, putative** (Tc00.1047053503929.10), **Zn-finger protein, putative** (Tc00.1047053509669.40), **hypothetical proteins** (Tc00.1047053507747.94, Tc00.1047053511517.120, Tc00.1047053507509.40, Tc00.1047053506219.40), **obsolete sequences** (23089, 4010, 24338, 23282, 6794, 6316, 22884, 21242)

Anexo 4 – Curvas de dissociação dos iniciadores dos genes usados nos experimentos de qRT-PCR.



(Continua na próxima página)

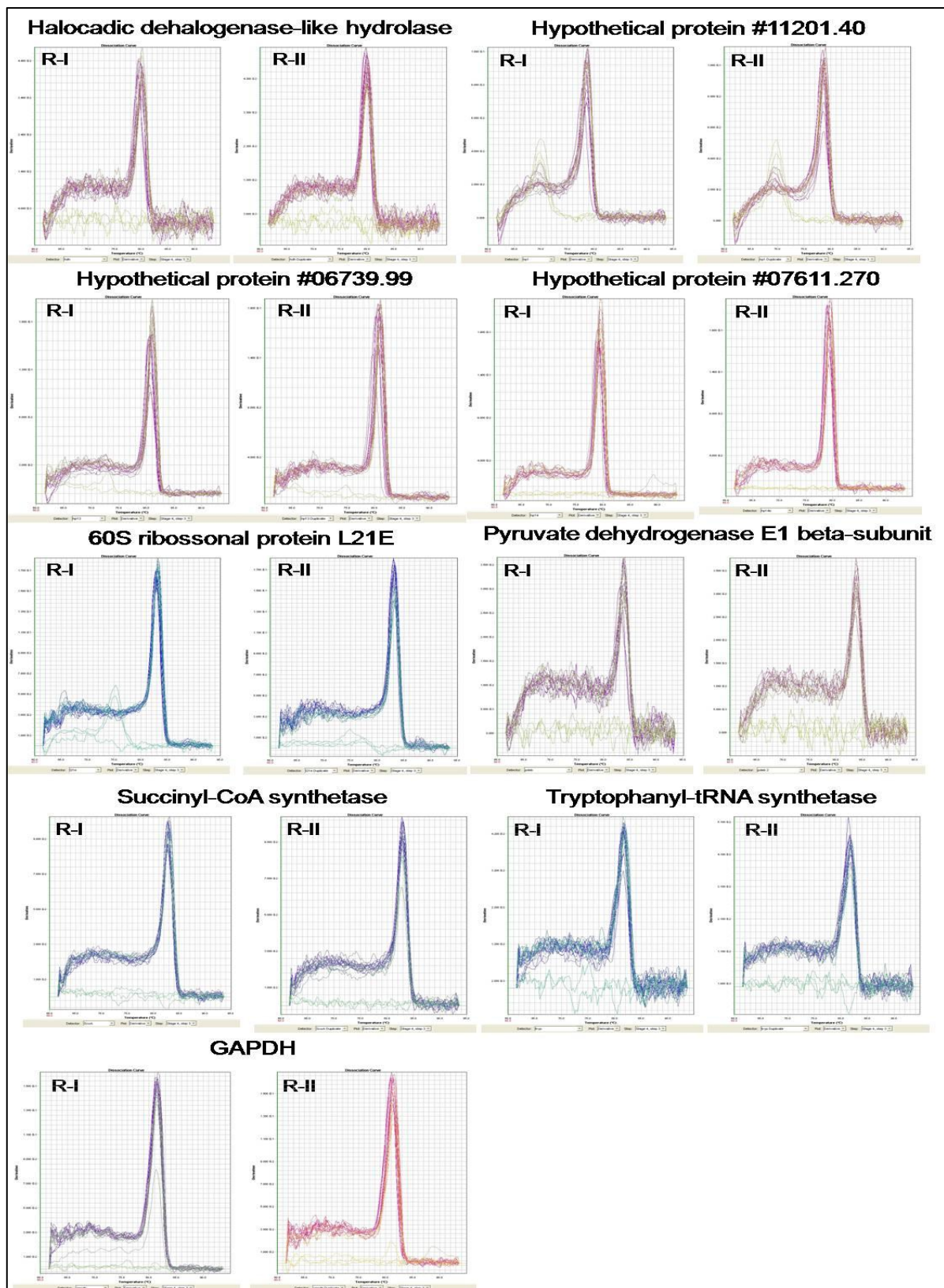


Figura A4: Curvas de dissociação dos iniciadores utilizados nos experimentos de qRT-PCR  
R-1 = Réplica Biológica I. R-II = Réplica Biológica II. 1 = Pico de amplificação. 2 = controle negativo (onde não ocorreu amplificação)

Anexo 5 – Genes associados com processos de reparo de DNA que não apresentam valores de intensidade

---

**Direct repair:** DNA photolyase, putative (Tc00.1047053509027.10)

**Base excision repair - DNA glycosylase:** 8-oxoguanine DNA glycosylase, putative (Tc00.1047053506559.170), MutY/Nth/A-G glycosylase (Tc00.1047053504005.10)

**Base excision repair - Flap endonuclease:** flap endonuclease-1 (FEN-1), putative (Tc00.1047053511867.110)

**Base excision repair - Apurinic/aprimidinic endonuclease:** APE-1/exonucleaseIII (Tc00.1047053507083.30), APE-2/APN2/exonucleaseIII (Tc00.1047053504827.50)

**Base excision repair - DNA polymerase:** mitochondrial DNA polymerase beta-PAK, putative (Tc00.1047053507063.100)

**Nucleotide excision repair - Recognition, global genome repair:** RAD16 (Tc00.1047053506925.140), UV excision repair RAD23-like protein, putative (Tc00.1047053506833.60)

**Nucleotide excision repair – Recognition, transcription coupled repair:** CSB/RAD26 (Tc00.1047053508675.20)

**Nucleotide excision repair –Open complex:** Replication factorA, RFA1 (Tc00.1047053510901.60), Replication factorA, RFA2 (Tc00.1047053510821.50), TFIIH-SSL1 (Tc00.1047053511907.300), TFIIH-TFB2 (Tc00.1047053506009.80), TFIIH-XPB/RAD25 (Tc00.1047053510149.50), TFIIH-XPB/RAD25 (Tc00.1047053511527.20), XPG/RAD2 (Tc00.1047053510667.40)

**Nucleotide excision repair –Incision:** RAD1/XPF (Tc00.1047053509779.10), RAD10/ERCC1 (Tc00.1047053510165.20), XPG/RAD2 (Tc00.1047053510667.40)



**Mismatch repair** MLH1(Tc00.1047053504035.140), MSH2 (Tc00.1047053509643.80), MSH3 (Tc00.1047053508277.180), MSH4 (Tc00.1047053507233.110), MSH5 (Tc00.1047053509617.30), MSH6 (Tc00.1047053510187.430), PMS1 (Tc00.1047053510761.10);

**Homologous recombination – Initiation:** MRE11 (Tc00.1047053509099.70), NBS1 (Tc00.1047053506743.180), RAD50 (Tc00.1047053508817.70), RECQ (Tc00.1047053506941.90), RECQ (Tc00.1047053507989.40), SPO11 (Tc00.1047053503619.10);

**Homologous recombination - Strand Exchange:** RAD51-3 (Tc00.1047053504153.220), RAD51-5 (Tc00.1047053511837.50), RAD51-6 (Tc00.1047053508075.20), RAD54 (Tc00.1047053505183.40), Replication factorA, RFA1 (Tc00.1047053510901.60), Replication factorA, RFA2 (Tc00.1047053510821.50)

**Homologous recombination – Resolution:** MUS81 (Tc00.1047053510065.20)

**Homologous recombination - Meiosis-specific recombination enzymes:** MSH4 (Tc00.1047053507233.110), MSH5 (Tc00.1047053509617.30), SPO11(Tc00.1047053503619.10)

**Homologous recombination – Topoisomerases:** Topoisomerase IIA (Tc00.1047053508277.370), Topoisomerase IIA (Tc00.1047053509203.70), DNA polymerase Kappa (Tc00.1047053510131.70), DNA polymerase Zeta, catalytic subunit (Tc00.1047053509769.130), Rev1 (Tc00.1047053508625.90); **Other repair genes:** PARP (Tc00.1047053510173.90)

**Non-homologous end-joining:** KU70 (Tc00.1047053503643.10), KU80 (Tc00.1047053511491.50).

Anexo 6 – Manuscrito referente a este trabalho submetido para a revista internacional indexada “PLoS Neglected Tropical Diseases”

---

*Trypanosoma cruzi* gene expression in response to gamma radiation

Priscila Grynberg<sup>1</sup>, Danielle Gomes Passos-Silva<sup>1</sup>, Marina de Moraes Mourão<sup>3</sup>, Roberto Hirata-Jr<sup>4</sup>, Andrea Mara Macedo<sup>1</sup>, Carlos Renato Machado<sup>1</sup>, Daniella Castanheira Bartholomeu<sup>2</sup>, Gloria Regina Franco<sup>1</sup>

<sup>1</sup>Departamento de Bioquímica e Imunologia and <sup>2</sup>Departamento de Parasitologia, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais\*; <sup>3</sup>Grupo de Genômica e Biologia Computacional and Instituto Nacional de Ciência e Tecnologia em Doenças Tropicais, Centro de Pesquisas René Rachou, FIOCRUZ; <sup>4</sup>Instituto de Matemática e Estatística, USP.

\*Av. Antônio Carlos, 6627, Caixa Postal 486, Belo Horizonte, 30210-970, MG, Brazil

E-mail; [gfranco@icb.ufmg.br](mailto:gfranco@icb.ufmg.br)

Telephone: +55 31 3409-2537

Fax: +55 31 3409-2984

Running title: *T. cruzi* gene expression in response to gamma rays

Key-words: *Trypanosoma cruzi*, gamma radiation, microarray, gene expression, RHS proteins, kDNA genes

## **Abstract**

*Trypanosoma cruzi* is an organism highly resistant to ionizing radiation. Following a dose of 500 Gy of gamma radiation, the fragmented genomic DNA is gradually reconstructed and the pattern of chromosomal bands is restored in less than 48 hours. Cell growth arrests after irradiation but, while DNA is completely fragmented, RNA maintain its integrity. In this work we compared the transcriptional profiles of irradiated and non-irradiated epimastigotes at different time points after irradiation using microarray. In total, 273 genes were differentially expressed, from these, 160 were up-regulated and 113 down-regulated. We found that genes with predicted functions are the most prevalent in the down-regulated gene category. Translation and protein metabolic processes, as well as generation of precursor of metabolites and energy pathways were affected. In contrast, the up-regulated category was mainly composed by obsolete sequences (which include some genes of the kinetoplast DNA), genes coding for hypothetical proteins, and Retrotransposon Hot Spot genes. Finally, the tyrosyl-DNA phosphodiesterase 1, a gene involved in double-strand DNA break repair process, was up-regulated. Our study demonstrated the peculiar response to ionizing radiation, raising questions about how this organism changes its gene expression to manage such a harmful stress.

## Introduction

*Trypanosoma cruzi*, the etiologic agent of Chagas disease, is a kinetoplastid organism highly resistant to DNA damage caused by ionizing radiation [1,2]. It is likely that this parasite has an efficient double-strand DNA break repair apparatus considering that after 500 Gy dose of gamma radiation the fragmented DNA is gradually repaired and the pattern of chromosomal bands is restored in less than 48 hours [2]. Remarkably, this dose is approximately 50-100 times higher than that tolerated by plant roots [3], *Plasmodium* [4], mammalian cells [5] and *Trypanosoma brucei*, the *T. cruzi* phylogenetic closest organism [6]. *Leishmania major*, another kinetoplastid, is similarly resistant to gamma rays since it is capable of growing by subculture after 300 Gy dose of gamma radiation, though slightly more sensitive when compared to *T. cruzi*, [7]. A plausible hypothesis underlying this extraordinary recovery is the role played by the TcRAD51, a component of the homologous recombination machinery, which is induced following irradiation [2].

Time-course studies have been performed analyzing the overall gene expression in response to gamma rays in yeast [8], *Arabidopsis thaliana* [9], in a Archaea organism [10], in human fibroblasts [11] and in *Deinococcus radiodurans*, an extremophile bacterium capable to endure doses of gamma radiation higher than 15,000 Gy [12]. Moreover, little is known about global changes in gene expression during DNA repair and cell recovery mechanisms following ionizing irradiation in *T. cruzi*.

To address this question we used microarray experiments to explore global gene expression alterations in the first 96 hours after gamma radiation, when DNA repair has already been achieved. Amongst the gene expression patterns observed, we found an interesting prevalence in the up-regulated gene categories of members of the retrotransposon hot spot (RHS) gene family and kinetoplast mitochondrial DNA genes

(kDNA). Remarkably, these genes increased their expression at 48 hours, when *T. cruzi* chromosomes are known to be fully reassembled [2]. In addition, translation and protein degradation processes are repressed. XX

## **Methods**

### ***T. cruzi* epimastigotes cell culture and gamma radiation**

Epimastigote cells were cultivated at 28 °C in LIT medium (*Liver Infusion Tryptone - liver digest neutralized*) [13] supplemented with complement-inactivated 10% fetal bovine serum, streptomycin sulfate (0.2 g/L) and penicillin (200,000 units/L). Cultures with  $5 \times 10^8$  parasites in 20 mL of LIT medium ( $2 \times 10^7$  cells/mL) were exposed to a dose of 500 Gy (1578 Gy/h per 20 minutes) in a cobalt ( $^{60}\text{Co}$ ) irradiator located at Centro de Desenvolvimento da Tecnologia Nuclear (CDTN), Belo Horizonte, Brazil. After irradiation, cells were counted in a cytometric chamber in specific time points and submitted to RNA extraction.

### **RNA extraction, purification and amplification**

Epimastigotes cells were subjected to RNA extraction using Trizol reagent (Invitrogen Life Technologies, USA) and RNA samples were purified using RNeasy® MiniElute™ Cleanup Kit (Qiagen, German) according to the manufacturers' instructions. Total RNA was quantified using a Nanodrop ND-100 UV/Vis spectrophotometer (NanoDrop Technologies, USA) and the overall RNA quality was assessed by denaturing gel electrophoresis [14]. Total RNA (2 µg) was amplified using the Amino Allyl MessageAmp II kit (Ambion, USA), according to the manufacturer's specifications.

### **Experimental design**

To evaluate the gamma radiation effect on *T. cruzi* epimastigotes cells, two identical and independent experiments (biological replicates) were performed in triplicates. Each

triplicate was composed by seven samples: cell culture growth control (not used in the microarray experiments), non-irradiated cells (used as reference sample in the microarray analysis) and five irradiated samples that had their RNA extracted immediately after irradiation (or i.a.i.), at 4, 24, 48 and 96 hours post-irradiation. Triplicates were pooled after the RNA extraction and before the purification/amplification steps. Each pool corresponding to one of the five time points was hybridized twice against the reference pool from each experiment, applying a reference time-course dye-swap design. Thus, four microarray slides for each time point were produced, two for each biological replicate, totalizing 20 slides.

The glass slide spotted with synthesized oligonucleotide DNA was kindly provided by Pathogen Functional Genomics Resource Center (PFGRC), a division of J. Craig Venter Institute, Rockville, MD. The arrays were coated with aminosilane. The arrays contained 12,288 oligonucleotides (70-mers), 500 *Arabidopsis thaliana* oligonucleotides (70-mers) controls spotted in duplicate and 920 empty spots totalizing 26,496 spots. Considering the *T. cruzi* oligonucleotides, there were 10,616 annotated genes, including 5,791 genes encoding hypothetical or conserved hypothetical proteins and 1,672 obsolete sequences.

### **aRNA labeling**

Aminoallyl amplified RNA was labeled with Cy3 and Cy5 according to a modified version of the AminoAllyl MessageAmp II Kit (Ambion, USA) and TIGR's standard operational procedure – SOP #M008. Briefly, we followed the manufacturer's instructions for the labeling step, but the initial amount of amplified RNA was changed to 8 µg. The Cy3 and Cy5 labeled samples were then combined. Labeled RNA was purified away from unincorporated dyes using YM-30 Microcon columns following manufacturer's specifications (Millipore®, USA). The final sample was dried again and

resuspended in 30  $\mu$ L of hybridization buffer (50% formamide, 5 $\times$  SSC, 0.1% SDS, 0.1M DDT and 6% salmon sperm as blocking agent) according to TIGR's SOP #M008. The solution was heated to 95  $^{\circ}$ C during 3 minutes and placed in ice for 30 seconds. After a brief centrifugation, the solution was dispensed onto the slide surface and covered by a coverslip.

### **Slide hybridization and scanning**

Slide pre-hybridization and hybridization steps were done as described elsewhere [15] with minor modifications. Briefly, slides were pre-hybridized by placing them in coupling jars containing pre-hybridization solution (5 $\times$  SSC, 0.1% SDS, 1% BSA) at 42 $^{\circ}$ C for one hour. Slides were washed twice by immersing 10 times in a beaker containing MilliQ water followed by dipping three times in isoamyl alcohol and subsequently spun dry. The slides containing 30  $\mu$ L of samples in hybridization buffer were hybridized for 14 hours in a water bath at 42 $^{\circ}$ C in the dark under cover slips inside Corning<sup>®</sup> hybridization chambers (Corning, USA). Slides were washed two times for five minutes each in a low stringency wash solution (2 $\times$  SSC, 0.1% N-lauroylsarcosine) at 42 $^{\circ}$ C (first wash) and RT (second wash), followed by two washes of five min in medium stringency wash (0.1 $\times$  SSC, 0.1% N-lauroylsarcosine) at RT and two washes for five minutes each in high stringency wash solution (0.1 $\times$  SSC) at RT. Slides were spun dry and scanned using a microarray dual channel laser scanner (ScanExpress Lite da PerkinElmer<sup>®</sup>, USA) at 10  $\mu$ m resolution, 100% laser power and PMT levels which were adjusted in order to obtain similar distributions of red and green signal intensities.

### **Background correction, normalization and statistical analysis**

For each time point, gene expression analysis was done based on information obtained from four slides, one dye-swap pair for each biological replicate. ScanArray Express (PerkinElmer<sup>®</sup>) software was used to generate the slides images and raw intensity data

which were then analyzed using specific packages from the R statistical language [16]. Only spots of good quality (positive flag value) were considered for analysis. Data were inspected for spatial biases on both red and green channels (background and signal), for print-tip bias, dye bias and bias dependent of intensity using the LIMMA [17] and marray [18] packages. Background correction was done with normexp method (off-set = 50) (Ritchie et al., 2007). Robust spline and aquantile methods were used for normalization within and between arrays, respectively. A linear model that incorporates biological and technical replication was used for statistical analysis. A list of differentially expressed genes was generated by applying adjusted p-values for multiple tests using 'BH' method, which works with the expected proportion of false-positives (FDR- False Discovery Rate) among the rejected hypothesis (Benjamini and Hochberg, 1995). Graphics were drawn using GraphPad version 5.03. Heatmap and cluster graphics were created using gplots package (Various R programming tools for plotting data) and personal scripts (R statistical language).

### **GO functional analysis and enrichment**

For the differentially expressed genes, functional assignment analysis for down and up-regulated genes was performed using GOanna tool (<http://agbase.msstate.edu/cgi-bin/tools/GOanna.cgi>). GO terms returned by GOanna tool were used as inputs for GO-slim Viewer tool, generic subset. The GO-slim result was manually cured to remove redundancy (two or more equal gene products associated with the same GO-slim category) and to generate more specific categories when a category was quite generic.

### **Obsolete sequences analysis**

The *T. cruzi* microarray (version 2) slide was designed based in the first genome release of this species. Obsolete oligonucleotide sequences differentially expressed were



aligned to the nuclear and kinetoplast genomes using blastn and were mapped into the annotated genome using Artemis software [19] in order to determine the localization.

### **Analysis of RHS genes**

A two-tail chi-square test was applied in each of the analyzed times to test whether  $f1 = f2$  (null hypothesis) or  $f1 \neq f2$  (alternative hypothesis) according to the equations 1 and 2 below:

$$(1) f1 = \frac{\# \text{ upregulated RHS}}{\# \text{ total RHS}}$$

$$(2) f2 = \frac{\# \text{ upregulated genes} - \# \text{ upregulated RHS}}{\# \text{ total genes} - \# \text{ total RHS}},$$

where i) up-regulated RHS = RHS genes significantly up-regulated; ii) up-regulated genes = total number of significantly up-regulated genes; iii) total RHS = number of RHS probes spotted on the slide (379); iv) total genes = number of gene probes spotted on the slide (12,288).

### **qRT-PCR**

A subset of differentially expressed genes was selected to be validated by qRT-PCR. Primers were designed using GenScript Real-time PCR (TaqMan) Primer Design (<https://www.genscript.com/ssl-bin/app/primer>) and Primer3 (<http://frodo.wi.mit.edu/primer3/>). Sequences are available at table S1. In order to test primer specificity, blastn was performed to align pairs of primers against NCBI nucleotide sequences database. The purified RNAs were treated with RQ1 RNase-Free DNase (Promega®, USA) and subsequently used for synthesis of cDNA using Superscript III cDNA Synthesis kit (Invitrogen™, USA), following the manufacturers' instructions. The qRT-PCR mix consisted of 10 ng of cDNA, 600 nM of each primer, 2.5 µL of Sybr Green PCR Master Mix (Applied Biosystems, Foster City, CA) and deionized H<sub>2</sub>O, totalizing 5 µL of reaction. Plates with 384 wells were read at 7900

7900HT Fast Real Time PCR System (Applied Biosystems, USA). Primers for *T. CRUZI* glyceraldehyde 3-phosphate dehydrogenase (GAPDH) were used as endogenous normalization control in all samples as reference applying the  $-2^{\Delta\Delta Ct}$  method [20]. Friedman and Dunn (post-test) tests were applied to verify the significance of the results. Spearman's rank coefficient was applied to verify the correlation between microarray and qRT-PCR results.

## **Results**

### **Effect of gamma radiation in *T. cruzi* overall gene expression**

Analysis of *T. cruzi* growth behavior in response to gamma-radiation revealed that, in both biological replicates, epimastigotes survived but did not grow during the whole time period analyzed, which is in accordance to previously published results [2]. However, despite the physical damage suffered by the cells (e.g. DNA fragmentation), they presented intense flagellar movement (visual observation) and remained viable (Fig. S1).

To visualize how gamma radiation affects *T. cruzi* overall gene expression we performed time-course microarray experiments in non-irradiated controls and in irradiated epimastigotes immediately after irradiation (i.a.i.), and at 4, 24, 48 and 96 hours after treatment. A dispersion graphic was generated using the fold-change values regarding the 273 genes found as being differentially expressed (adjusted p-value <0.05) in at least one time point. This result revealed a change in the expression pattern over time (Fig. 1A). At the first 4 hours, we observed a peak in the number of down-regulated genes (Fig. 1B).

It was also important to determine whether the overall decrease in gene expression in the first 4 hours after irradiation was due to a possible RNA degradation

caused by gamma radiation and that could compromise the labeling and hybridization steps. However, denaturing agarose gel electrophoresis revealed no difference on the rRNA (18S, 24S $\alpha$  and 24S $\beta$ ) band patterns (Fig. S2), demonstrating the high quality of total RNA in controls and irradiated samples (in triplicates) in both biological replicates.

Next, a subset of 18 differentially expressed genes was selected for validation using qRT-PCR. Of those, 13 corresponded to known genes, three to hypothetical proteins, and two sequences were classified as obsoletes. Obsolete sequences are those that, on a previous genome draft release, were initially considered to be derived from coding regions and therefore the corresponding oligos were spotted in the slide. However, after the release of the annotated *T. cruzi* CL Brener genome sequence [21], these oligonucleotides were, in fact, derived from non coding regions [22]. Spearman's rank coefficient was applied to verify the correlation between microarray and qRT-PCR results.  $R^2$  ranged from 0.4298 to 0.7185 at 4 and 24 hours post irradiation, respectively (Fig. S3) and all p-values were significant. The comparisons between microarray and qRT-PCR results from 8 genes are shown in Fig. 2. Data for the remaining 10 qRT-PCR results can be found in the supporting information (Fig. S4). Primers sequences, slope,  $R^2$  and efficiency are shown in Table S1.

Genes up or down-regulated coding for proteins with known function and/or predicted conserved domains corresponded to 41.4% of the total differentially expressed genes. RHS genes, a large family whose members are located preferentially at subtelomeric regions of *T. cruzi* chromosomes and present a RIME/ingi insertion site [23], corresponded to 7.7% of the total. The 50.9% remaining genes were found to be annotated as coding for hypothetical proteins, conserved or not (35.5%) or obsolete sequences (15.4%). Fig. 3 shows the distribution of these four categories for down and up-regulated genes over time. Genes with predicted function (dark gray) were the most

prevalent in the down-regulated gene category. In contrast, the up-regulated gene category was mainly composed by genes coding for hypothetical proteins (white), obsolete sequences (light gray) and RHSs (black). This last category was observed only at 24 hours post-irradiation.

Notably, differentially expressed genes usually showed low fold-change values, i.e. usually less than 2-fold variation in gene expression in comparison to non-irradiated cells (Fig. 1A, Table 1 and Table S2). Usually, such low values would be excluded from further analysis in other organisms. However, for parasites of Kinetoplastida order, previous microarray studies for *Trypanosoma* and *Leishmania* reported modest changes in the abundance of mRNA molecules which were confirmed by qRT-PCR [24,25]. They established that differentially expressed genes with low fold-change values, but with significant adjusted p-values for multiple tests are significantly differentially expressed.

### **Gamma rays effect on the expression of obsolete sequences, genes of the maxi-circle kinetoplast DNA and genes coding for RHS**

Of the 1,672 obsolete sequences spotted on the microarray slide, 42 (2.5%) were differentially expressed in, at least, one time point (adjusted p-value <0.05). qRT-PCR confirmed the increase in gene expression of the obsolete sequence #23196 in all time points analyzed (Fig. 2). This result prompted us to find out where in the genome these obsolete sequences were located. Thus, we used the Artemis software [19] to locate the oligonucleotides (70-mer) referring to obsolete sequences into the *T. cruzi* annotated genome (<http://tritrypdb.org/common/downloads/release-2.5/Tcruzi/>). We found that some sequences were located outside ORFs (17), some were located in contigs that could not be assigned to any particular chromosome (2), some were found in, at least, two different genomic regions (multilocal sequences) (9), and others were not located

anywhere in the nuclear genome (14). Curiously, four sequences of this last category presented high values of fold-change (up to eight times), in particular at 48 and 96 hours post-irradiation (Fig. 4). A search against the non-redundant (nr) NCBI database using the blastx program identified these sequences as genes of the maxicircle kinetoplast mitochondrial genome (kDNA), which is composed by 18 genes [26]. The over-expressed kDNA genes code for three proteins of the mitochondrial respiratory chain, the subunits ND1 and ND5 of NADH dehydrogenase and the cytochrome oxidase subunit one (COI), and for a mitochondrial unidentified reading frame (MURF-1). All four sequences presented a highly similar expression pattern, showing a substantial and significant increase in the transcript levels at 48 hours post-irradiation. The only exception was ND-5, which presented a decrease in gene expression 96 hours post-irradiation (Fig. 4). The remaining 14 mitochondrial genes were not represented in the microarray slide.

The *T. cruzi* genome project has identified 752 members of the RHS gene family, of which 557 are pseudogenes, and reported the presence of RHS orthologs in *T. brucei* [21]. We observed that 21 different RHS genes out of the 379 spotted on the microarray slide were up-regulated (Table S2): three at 24 hours, nine at 48 hours (one is the same as 24 hours), and 17 at 96 hours post-irradiation (of which eight were expressed only at this time point). In contrast, RHS genes were not significantly down-regulated. Using the genome browser of the database TriTrypDB, we retrieved the genomic location of all 21 differentially expressed RHS genes. Ten are located on the edges of chromosomes, four in the central but repetitive chromosomal regions, and seven in contigs not embedded in the final genome assembly (Table S2). Two-sided Chi-square's tests were independently performed for each time point and detected a

significant association between RHS gene expression and the effect of gamma radiation ( $p < 0.0001$ ) at 48 and 96 hours (Fig. 5).

### **Functional analysis of differentially expressed genes in epimastigote cells after irradiation**

To identify the biological processes and molecular functions that were repressed or induced in response to ionizing radiation, Gene Ontology (GO) functional analysis was performed. GO-slim terms for biological processes (BP) and molecular function (MF) were assigned to differentially expressed genes and this procedure was essential to provide an overview of the effect of gamma rays in epimastigotes cells (Fig. 6 and Table S3 for complete results). Concomitantly, a heatmap using all differentially expressed nuclear genes (269 genes) was generated to verify whether genes that participate in the same BP or MF showed similar gene expression profiles over time and could be grouped in the same cluster. Twelve different clusters were identified (Fig. 7). All clusters failed to group genes mapping at closer chromosomal locations (data not shown), and were formed by genes from different BP or MF categories. Interestingly, all genes from the MF category “structural constituent of ribosome” were clustered together (#11), along with genes from other MF and BP categories (Table S4).

Several ribosomal proteins, translation initiation and elongation factors and a tryptophanyl-tRNA synthetase were significantly down-regulated in at least one time point (Fig. 6 and Table S3). Even if the remaining fold-changes were not statistically significant the overall decreasing tendency was similar (Table 1). The two-sided Chi-square's test indicated that the number of genes down-regulated associated with protein translation was not a random event ( $p = 0.0073$ ). Gene expression levels of elongation factor 1-alpha and 60S ribosomal protein L21E were assessed by qRT-PCR, which confirmed the microarray results (Fig. 2).

Lipid metabolic process (BP) is also noticeable repressed, as genes coding for enzymes involved in isoprenoid and sterol synthesis - mevalonate kinase, 3-oxo-5-alpha-steroid 4-dehydrogenase and farnesyl pyrophosphate synthase and genes coding for enzymes required for the synthesis of long and unsaturated fatty acids such as fatty acid desaturase, fatty acid elongase and prostaglandin F2alpha synthase were also down-regulated (Fig 6, Table 1 and Table S3).

Additionally, the GO category “protein metabolic process”-BP was also noticeably down-regulated. Examples of down-regulated genes from this category were the heat shock proteins Hsp70 and Hsp60 (the expression profile of the later was also confirmed by qRT-PCR– Figure 2), chaperonin T-complex protein 1, aminopeptidase, carboxipeptidase, cysteine proteinase, cysteine peptidase and a regulatory subunit of the 26S proteasome (Fig. 6 and Table S3). Interestingly, this category presented three up-regulated genes, (co-chaperone GrpE, aspartyl aminopeptidase and calpain-like cysteine peptidase). Noteworthy is the finding that the GO category “protein modification process”-BP was up-regulated, mainly by “protein kinase activity”-MF (e.g. serine/threonine protein kinase, protein kinase A and other protein kinases) and “transferase activity”-MF (e.g. acetyltransferase). An increase in levels of transcripts for protein kinases, along with up-regulation of two representatives of the EF-hand protein 5 and a hypothetical protein containing an EF-hand conserved domain, a cation transporter protein, a calpain-like cysteine peptidase and a hypothetical protein similar to phospholipase C-like phosphodiesterase may suggest that *T. cruzi* is responding to radiation stress by activating signal transduction pathways involving calcium or membrane lipids (Fig. 6 and Table S3). Our results regarding the calpain-like cystein peptidase agrees with a previous study that indicated a possible role in stress response in *T. cruzi*. This gene was strongly responsive to three different types of stress (nutritional,

temperature and acidic pH) and presented no proteolytic activity which is an indicative of signal transduction activity (Giese *et al.*, 2008).

In order to recover from irradiation, *T. cruzi* needs to produce energy. Our functional analysis revealed, however, that some genes participating in pathways involved in “carbohydrate metabolic process” and in “generation of precursor of metabolites and energy” (BP categories) were down regulated, such as a hexose transporter, pyruvate dehydrogenase E1 beta subunit, glycosomal phosphoenolpyruvate carboxykinase, one conserved hypothetical protein which seems to be a 2Fe-2S protein component of succinate dehydrogenase complex, succinyl-CoA synthetase, and V-type ATPase (Fig. 6 and Table S3). Transcripts coding for proteins involved in basic metabolic processes such as glycolysis, gluconeogenesis, acid citric cycle and ATP production were diminished. In contrast, only one gene (glucose-6-phosphate isomerase) from the “carbohydrate metabolic process”-BP category was over-expressed (Fig. 6). We also investigated nuclear genes coding for proteins from the respiratory chain but found that only cytochrome c1 was down-regulated (Table 1).

Contrary to the expected, we found that gamma radiation did not induce a massive expression of key antioxidant genes. Indeed, only two important genes (trypanothione/tryparedoxin dependent peroxidase 2 and tryparedoxin peroxidase) related to ROS removal were up-regulated (Table 1). However trypanothione reductase, a key enzyme of the trypanothione-based thiol redox metabolism that reduces trypanothione disulfide back to its active state trypanothione, was down-regulated (Table 1).

Besides TcRAD51, *T. cruzi* possesses several other genes involved in DNA repair [21] but we found that only the DNA repair enzyme tyrosyl-DNA



phosphodiesterase (Tdp-1) was significantly up-regulated 24 hours after irradiation, probably when the DNA repair process was already active. Curiously, this gene had not even been annotated as a DNA repair gene by the time the *T. cruzi* genome was published [21].

## **Discussion**

An efficient DNA repair system may explain *T. cruzi* impressive resistance to ionizing radiation as it is able to reconstitute a completely fragmented karyotype in less than 48 hours after gamma radiation [2]. Our time course microarray analysis pointed to 273 statistically significant differentially expressed genes after parasite irradiation. These genes were classified into four different subsets: genes coding for proteins with known function (attributed to GO categories), genes for hypothetical proteins, genes for RHS and obsolete sequences. More than 60% of the genes which are up-regulated code for proteins with unknown function. The remaining up-regulated genes comprised RHS and obsolete sequences. However, unlike a study conducted for *D. radiodurans*, where authors could infer function for uncharacterized genes based on operon context [12], analysis of the 12 *T. cruzi* gene clusters presenting similar expression profile demonstrated no correlation between chromosome location and pattern of gene expression over time. This indicates that *T. cruzi* genes are not transcribed as operon units. Indeed, no classical operons have been so far identified in this parasite, despite the fact that Trypanosomatid genes are transcribed as polycistronic units. These units may contain genes from unrelated pathways which, after trans-splicing processing, originate monocistronic mature transcripts with widely different expression levels [27]. Also no canonical RNA polymerase II promoter was found in the genomes of trypanosomes.

It is well known that Trypanosomatid mRNAs steady state is dictated primarily post-transcriptionally by mRNA stability, but the protein levels also depend on additional mechanisms to efficiently control translation initiation and elongation (translational selection) and protein turnover [27,28,29,30]. Using a microarray transcriptomic analysis at different stages of differentiation to detected transcripts that were induced only during differentiation, Queiroz and collaborators have demonstrated that *T. brucei* genes form post-transcriptional regulons, in which mRNAs encoding proteins playing roles in particular pathways, or encoding components of protein complexes, show a co-regulated expression. Co-regulation of transcript levels may be achieved by binding of specific regulatory protein or RNA factors capable of controlling translation and mRNA decay [28,29] to elements in the 3'-untranslated regions (3'-UTR). In this work, using the MEME suite software, we attempted to discover motifs at 3'-UTR of differentially expressed transcripts that could act as elements to control mRNA decay. Our results revealed no conserved motifs harbored by sequences from the same gene cluster (data not show). This indicates that, if these clusters are composed of co-regulated genes (post-transcription regulons), their expression are not coordinated by 3'-UTR elements. Indeed, the only genes related to the same protein functional category that clustered together were the ones coding for ribosomal proteins (cluster 11).

Our functional analysis suggests that protein translation and processing are reduced after irradiation, given that the expression levels of seven translation initiation and elongation factors, as well as nine ribosomal constituents and a tryptophanyl-tRNA synthetase are down-regulated and none were up-regulated because Kinetoplastids have very few transcription factors identified in these organisms and therefore control their gene expression at post-transcriptional and translational level [31], it is likely that ionizing radiation does not affect transcriptional rate to a great extent. This effect is

remarkably different in other organisms. Indeed, eukaryotes control their gene expression mainly at initiation of transcription, and ionizing radiation causes an overall impact in the constituents of the transcriptional apparatus and transcription factors as reported for yeast [8] and *Arabidopsis thaliana* [9]. The opposite was found on genes from *D. radiodurans* related to protein synthesis and protein fate [32]. It is important to note that DNA damage recovery after gamma radiation in this species is remarkably fast, being possible to observe genes up-regulated just 1 hour after treatment [32]. The current study suggests that the putatively reduction in the rate of protein synthesis and processing may also lead to the down-regulation of the chaperone proteins hsp60 and hsp70. Such results are remarkably different from those reported in the literature which showed that gamma radiation induces over-expression of heat shock proteins [33,34,35].

Previous work has shown that epimastigotes growth halt after irradiation for at least for 96 hours without cell division [2]. The arrest in cell growth may be explained by a decreased expression of  $\alpha$ - and  $\beta$ -tubulin genes during all time points analyzed. These proteins assemble as the components of the cytoskeleton microtubules which are involved in mitosis and cytokinesis. Similarly, depletion of cell division and motility has been described for other organisms after stress induced by ionizing radiation [9,12]. Nevertheless, in spite of their growth arrest, epimastigotes were always active with an intense flagellar movement. A possible explanation may be related to the observation that genes associated with microtubule motor activity and components of the parasite swimming flagellum were up-regulated: the dynein light (lc6) and two dynein heavy chains [36,37]. In fact, it was recently demonstrated that silencing of a putative inner arm dynein heavy chain resulted in flagellar immotility of *T. brucei* cells [38].

Noticeably is the fact that genes involved in lipid metabolism were also down-regulated. Perhaps the growth arrest reflects the observed reduction in the levels of complex lipids as they are compounds of membranes and proteins, second messengers or even cell mediators and hormones.

Gamma radiation induces oxidative stress caused directly or indirectly via reactive oxygen species (ROS) production. ROS can also cause DNA single and double-strand breaks and base damage. *T. cruzi* is naturally exposed to ROS produced by itself or by the host response to infection (Docampo, 1990), and this parasite has a full system based on trypanothione-thiol metabolism to deal with oxidative stress (Krauth-Siegel & Comini, 2008). It was observed that *E. coli* cells expressing the *T. cruzi* trypanothione showed an increased protection against genomic DNA damage (4.6-fold) and resistance to gamma rays (4.3-fold), as well as a decrease in intracellular levels of ROS (Fitzgerald *et al.*, 2010). We think that *T. cruzi* is experiencing oxidative stress and dealing with ROS species, although results did not indicate a strong stimulation in the expression of genes involved in oxidative damage repair. Interestingly, expression levels of most of the known *T. cruzi* DNA repair genes [13] were not noticeably affected (Table S5).

Genes participating in different energy uptake pathways were down-regulated, which made us speculate how this organism is getting energy for its survival. Perhaps *T. cruzi* basal metabolism is maintained at a minimum level after irradiation with a decrease in the transcripts quantity. Additionally, proteins involved in catabolic processes to generate energy are possibly saved from degradation. On the other hand, mitochondrial genes may be over-expressed as consequence of a direct effect of the ionizing radiation in the mitochondrion, an organelle highly sensitive to oxidative stress (Slupphaug *et al.*, 2003). Reinforcing this hypothesis is the finding that in the radioresistant human glioblastoma cell line T98G the levels of transcripts from

cytochrome c oxidase subunits 1 and 2 and NADH dehydrogenase subunit 4 were raised after treatment with ionizing radiation, as consequence of oxidative stress (Gong *et al.*, 1998).

It was reported that ionizing radiation induces the expression of TcRAD51 in *T. cruzi*. Moreover, epimastigotes overexpressing this gene halved the time necessary for full recovery after irradiation [2]. Surprisingly, the tyrosyl-DNA phosphodiesterase I (tdr-1), a DNA repair enzyme that had not even been annotated as a DNA repair gene by the time the *T. cruzi* genome was published [20], is up-regulated mainly 24 hours post-irradiation. In the nucleus Tdr-1 hydrolyze the linkage among the tyrosine of trapped topoisomerase 1 and the 3' end of DNA, through the cleavage of 3'-phosphotyrosyl bonds [39]. This process, which is DNA ligase III alpha dependent [40], prepares DNA double-strand breaks for DNA ligation and, consequently, DNA repair. Tdr-1 is also localized in the mitochondria and the base excision repair (BER) mechanism necessary for a efficient oxidative damage repair depends on its action [41]. Recently, this enzyme was characterized in *Leishmania donovani*. The authors demonstrated that tdp-1 has a possible role in topoisomerase I-mediated DNA repair pathway both in the nucleus and in the kDNA. However none mechanism of action and possible interactive partners was proposed so far for Trypanosomatids, considering that DNA ligase III is absent in kinetoplastids [42]. In any case, more studies will be needed to reveal which extent Tdp-1 participates in the cell recovery process: whether it acts in concert, in parallel or independently of other genes.

One of the most intriguing question of this study is how DNA is consistently reconstructed considering that chromosomal bands are extensively fragmented [2]. We showed that gamma radiation induces the up-regulated the expression RHS genes/pseudogenes at 48 and 96 hours after irradiation. Coincidentally, 48 hours is the

time period when DNA repair is thought to be completed [2]. RHSs can be found preferentially in the repetitive subtelomeric region, but also in the center of chromosomes. However, other repetitive gene families such trans-sialidase-like family, MASP, gp63 and the putative surface protein DGF-1 (dispersed gene family-1) were not over-expressed such as RHS. Our results suggest that the expression of members of this gene family may be related to the irradiation stress although it is not possible at this point establish whether directly or indirectly. Similarly, in *S. cerevisiae*, genomic damage after exposure to gamma radiation induces Ty1 (LTR-retroelements) transcripts which reaches to the impressive value of 86-fold after a dose of 800 Gy gamma rays when compared to non-irradiated cells. This increase in Ty1 transcripts activates retrotransposition in a dose-dependent manner, [43]. Retrotransposition of Long Interspersed Elements (LINE-1s), mobile elements present in the human genome, is also increased after exposure to gamma rays and is responsible for genomic instability [44].

Focusing on aspects of *T. cruzi* life cycle, it is reasonable to think that the mechanisms underlying the parasite radiation resistance may be part of the responses prepared to the stresses the organism will face throughout its life cycle such as vector intestine, temperature, pH and osmolarity changes, as well as saliva and blood digestion products [45]. Furthermore, blood digestion can provide heme molecules sub-products required for the Fenton's chemical reaction, a harmful free radical reaction that produces hydroxyl radical and causes oxidative stress. And last, just after a meal, the water present in the blood is excreted by the Malphigian tubules in the intestine of the bug and the desiccated blood is stored in the crop [46]. Taking all together, the pathway though the vector's gut is a challenge for *T. cruzi* genome integrity.

In conclusion, we showed that gamma rays affect the *T. cruzi* gene expression in a time-dependent manner. Expression of transcripts related to basal metabolic functions

was reduced although apparently maintained at a minimum operation rate to meet the organism needs. Intriguingly, RHS, a category of proteins found preferentially in the repetitive subtelomeric region and in the center of chromosomes genes were substantially induced although it is not clear whether or not they play a direct role in cell recovery and what role would that be. Further studies regarding the role of some of the genes found to be most potentially up or down regulated following irradiation as well as genes known to be directly involved in DNA repair such as tdr-1 gene will shed light into the repair/re-arrangement process after DNA fragmentation caused by the ionizing radiation. XX

### **Acknowledgments**

We would like to thank the Pathogen Functional Genomics Resource Center (PFGRC), a division of J. Craig Venter Institute, Rockville, MD for kindly provide the glass slides. We also would like to thank Neuza Antunes Rodrigues for technical assistance, the Brazilian funding agency CNPq for grant support (Process number 476782/2007-4) and CAPES for Priscila Grynberg's scholarship.

### **References**

1. Takeda GK, Campos R, Kieffer J, Moreira AA, Amato Neto V, et al. (1986) [Effect of gamma rays on blood forms of *Trypanosoma cruzi*. Experimental study in mice]. *Rev Inst Med Trop Sao Paulo* 28: 15-18.
2. Regis-da-Silva CG, Freitas JM, Passos-Silva DG, Furtado C, Augusto-Pinto L, et al. (2006) Characterization of the *Trypanosoma cruzi* Rad51 gene and its role in

- recombination events associated with the parasite resistance to ionizing radiation. *Mol Biochem Parasitol* 149: 191-200.
3. Jiang CZ, Yen CN, Cronin K, Mitchell D, Britt AB (1997) UV- and gamma-radiation sensitive mutants of *Arabidopsis thaliana*. *Genetics* 147: 1401-1409.
  4. Geoghegan DS, Skinner-Adams T, Davis TM (2001) *Plasmodium falciparum*: isolate-specific radiosensitivity. *Exp Parasitol* 99: 108-110.
  5. Yonetani Y, Hochegger H, Sonoda E, Shinya S, Yoshikawa H, et al. (2005) Differential and collaborative actions of Rad51 paralog proteins in cellular response to DNA damage. *Nucleic Acids Res* 33: 4544-4552.
  6. Tan KS, Leal ST, Cross GA (2002) *Trypanosoma brucei* MRE11 is non-essential but influences growth, homologous recombination and DNA double-strand break repair. *Mol Biochem Parasitol* 125: 11-21.
  7. Seo M, Chun DK, Hong ST, Lee SH (1993) Influence of heat shock, drugs, and radiation on karyotype of *Leishmania major*. *Korean J Parasitol* 31: 277-283.
  8. Kimura S, Ishidou E, Kurita S, Suzuki Y, Shibato J, et al. (2006) DNA microarray analyses reveal a post-irradiation differential time-dependent gene expression profile in yeast cells exposed to X-rays and gamma-rays. *Biochem Biophys Res Commun* 346: 51-60.
  9. Ricaud L, Proux C, Renou JP, Pichon O, Fochesato S, et al. (2007) ATM-mediated transcriptional and developmental responses to gamma-rays in *Arabidopsis*. *PLoS One* 2: e430.
  10. Williams E, Lowe TM, Savas J, DiRuggiero J (2007) Microarray analysis of the hyperthermophilic archaeon *Pyrococcus furiosus* exposed to gamma irradiation. *Extremophiles* 11: 19-29.



11. Tachiiri S, Katagiri T, Tsunoda T, Oya N, Hiraoka M, et al. (2006) Analysis of gene-expression profiles after gamma irradiation of normal human fibroblasts. *Int J Radiat Oncol Biol Phys* 64: 272-279.
12. Liu Y, Zhou J, Omelchenko MV, Beliaev AS, Venkateswaran A, et al. (2003) Transcriptome dynamics of *Deinococcus radiodurans* recovering from ionizing radiation. *Proc Natl Acad Sci U S A* 100: 4191-4196.
13. Castellani O RL, Fernandes JF (1967) Differentiation of *Trypanosoma cruzi* in culture. *J Protozool* 14: 447-451.
14. Sambrook J FE, Maniatis T (1989) *Molecular cloning: a laboratory manual*. Cold Spring Harbor: Cold Spring Harbor Laboratory.
15. Waisberg M, Lobo FP, Cerqueira GC, Passos LK, Carvalho OS, et al. (2007) Microarray analysis of gene expression induced by sexual contact in *Schistosoma mansoni*. *BMC Genomics* 8: 181.
16. Team RDC (2010) *R: A language and environment for statistical computing*. R Foundation for Statistical Computing.
17. Smyth GK (2004) Linear models and empirical Bayes methods for assessing differential expression in microarray experiments. *Statistical Applications in Genetics and Molecular Biology* 3.
18. Wang J, Nygaard V, Smith-Sorensen B, Hovig E, Myklebost O (2002) MArray: analysing single, replicated or reversed microarray experiments. *Bioinformatics* 18: 1139-1140.
19. Carver T, Berriman M, Tivey A, Patel C, Bohme U, et al. (2008) Artemis and ACT: viewing, annotating and comparing sequences stored in a relational database. *Bioinformatics* 24: 2672-2676.

20. Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the  $2^{-\Delta\Delta C(T)}$  Method. *Methods* 25: 402-408.
21. El-Sayed NM, Myler PJ, Bartholomeu DC, Nilsson D, Aggarwal G, et al. (2005) The genome sequence of *Trypanosoma cruzi*, etiologic agent of Chagas disease. *Science* 309: 409-415.
22. Minning TA, Weatherly DB, Atwood J, 3rd, Orlando R, Tarleton RL (2009) The steady-state transcriptome of the four major life-cycle stages of *Trypanosoma cruzi*. *BMC Genomics* 10: 370.
23. Bringaud F, Biteau N, Melville SE, Hez S, El-Sayed NM, et al. (2002) A new, expressed multigene family containing a hot spot for insertion of retroelements is associated with polymorphic subtelomeric regions of *Trypanosoma brucei*. *Eukaryot Cell* 1: 137-151.
24. Minning TA, Bua J, Garcia GA, McGraw RA, Tarleton RL (2003) Microarray profiling of gene expression during trypomastigote to amastigote transition in *Trypanosoma cruzi*. *Mol Biochem Parasitol* 131: 55-64.
25. Saxena A, Worthey EA, Yan S, Leland A, Stuart KD, et al. (2003) Evaluation of differential gene expression in *Leishmania major* Friedlin procyclics and metacyclics using DNA microarray analysis. *Mol Biochem Parasitol* 129: 103-114.
26. Simpson L, Neckelmann N, de la Cruz VF, Simpson AM, Feagin JE, et al. (1987) Comparison of the maxicircle (mitochondrial) genomes of *Leishmania tarentolae* and *Trypanosoma brucei* at the level of nucleotide sequence. *J Biol Chem* 262: 6182-6196.

27. Teixeira SM, da Rocha WD (2003) Control of gene expression and genetic manipulation in the Trypanosomatidae. *Genet Mol Res* 2: 148-158.
28. Clayton C, Shapira M (2007) Post-transcriptional regulation of gene expression in trypanosomes and leishmanias. *Mol Biochem Parasitol* 156: 93-101.
29. Haile S, Papadopoulou B (2007) Developmental regulation of gene expression in trypanosomatid parasitic protozoa. *Curr Opin Microbiol* 10: 569-577.
30. Horn D (2008) Codon usage suggests that translational selection has a major impact on protein expression in trypanosomatids. *BMC Genomics* 9: 2.
31. Queiroz R, Benz C, Fellenberg K, Hoheisel JD, Clayton C (2009) Transcriptome analysis of differentiating trypanosomes reveals the existence of multiple post-transcriptional regulons. *BMC Genomics* 10: 495.
32. Palenchar JB, Bellofatto V (2006) Gene transcription in trypanosomes. *Mol Biochem Parasitol* 146: 135-141.
33. Tanaka M, Earl AM, Howell HA, Park MJ, Eisen JA, et al. (2004) Analysis of *Deinococcus radiodurans*'s transcriptional response to ionizing radiation and desiccation reveals novel proteins that contribute to extreme radioresistance. *Genetics* 168: 21-33.
34. Caillet S, Millette M, Dussault D, Shareck F, Lacroix M (2008) Effect of gamma radiation on heat shock protein expression of four foodborne pathogens. *J Appl Microbiol* 105: 1384-1391.
35. Nogami M, Huang JT, James SJ, Lubinski JM, Nakamura LT, et al. (1993) Mice chronically exposed to low dose ionizing radiation possess splenocytes with elevated levels of HSP70 mRNA, HSC70 and HSP72 and with an increased capacity to proliferate. *Int J Radiat Biol* 63: 775-783.

36. Calini V, Urani C, Camatini M (2003) Overexpression of HSP70 is induced by ionizing radiation in C3H 10T1/2 cells and protects from DNA damage. *Toxicol In Vitro* 17: 561-566.
37. Asai DJ, Wilkes DE (2004) The dynein heavy chain family. *J Eukaryot Microbiol* 51: 23-29.
38. DiBella LM, Gorbatyuk O, Sakato M, Wakabayashi K, Patel-King RS, et al. (2005) Differential light chain assembly influences outer arm dynein motor function. *Mol Biol Cell* 16: 5661-5674.
39. Das BB, Dexheimer TS, Maddali K, Pommier Y (2010) Role of tyrosyl-DNA phosphodiesterase (TDP1) in mitochondria. *Proc Natl Acad Sci U S A* 107: 19790-19795.
40. Sacerdot C, Mercier G, Todeschini AL, Dutreix M, Springer M, et al. (2005) Impact of ionizing radiation on the life cycle of *Saccharomyces cerevisiae* Ty1 retrotransposon. *Yeast* 22: 441-455.
41. Farkash EA, Kao GD, Horman SR, Prak ET (2006) Gamma radiation increases endonuclease-dependent L1 retrotransposition in a cultured cell assay. *Nucleic Acids Res* 34: 1196-1204.
42. Kollien AH, Schaub GA (2000) The development of *Trypanosoma cruzi* in triatominae. *Parasitol Today* 16: 381-387.

Table 1: Fold-change of differentially expressed genes coding for proteins with known function.

Gene ID	Product Name	Fold-change				
		i.a.i	4h	24h	48h	96h
<b>Down-regulated</b>						
Tc00.1047053509693.100	2- aminoethylphosphonate:pyruvateamino-transferase-like protein	<b>-1.42</b>	<b>-1.51</b>	-1.14	-1.19	-1.24
Tc00.1047053507671.30	25 kDa translation elongation factor 1-beta	-1.61	<b>-1.55</b>	-1.09	-1.50	-1.34
Tc00.1047053504221.20	26S proteasome regulatory non-ATPase subunit	-1.05	-1.20	<b>-1.44</b>	-1.24	-1.10
Tc00.1047053504427.70	3-oxo-5-alpha-steroid 4-dehydrogenase	-1.10	-1.12	<b>-1.43</b>	-1.05	-1.06
Tc00.1047053506679.150	40S ribosomal protein S10	-1.04	-1.12	-1.16	<b>-1.36</b>	-1.17
Tc00.1047053510101.430	40S ribosomal protein S21	1.03	-1.10	-1.31	<b>-1.51</b>	-1.22
Tc00.1047053509353.30	40S ribosomal protein S3	1.00	-1.09	-1.17	<b>-1.47</b>	-1.20
Tc00.1047053506297.150	40S ribosomal protein S5	-1.06	-1.12	<b>-1.58</b>	-1.30	-1.23
Tc00.1047053510425.19	40S ribosomal protein SA	-1.04	-1.03	<b>-1.36</b>	<b>-1.27</b>	-1.23
Tc00.1047053503719.20	40S ribosomal protein SA	-1.03	-1.04	<b>-1.34</b>	-1.20	-1.05
Tc00.1047053505977.26	60S acidic ribosomal protein P2	-1.06	-1.12	<b>-1.30</b>	-1.27	-1.27
Tc00.1047053508461.490	60S ribosomal protein L23	1.06	-1.11	-1.35	<b>-1.46</b>	<b>-1.51</b>
Tc00.1047053506297.270	60S ribosomal protein L28	-1.06	-1.13	<b>-1.51</b>	-1.35	-1.23
Tc00.1047053511211.120	activated protein kinase C receptor	-1.16	-1.17	-1.32	<b>-1.32</b>	-1.11
Tc00.1047053411235.9	alpha tubulin	<b>-1.89</b>	<b>-1.85</b>	-1.43	-1.33	-1.32
Tc00.1047053510655.120	aminopeptidase P	-1.20	<b>-1.55</b>	1.17	-1.29	-1.02
Tc00.1047053506563.40	beta tubulin	<b>-1.65</b>	<b>-1.68</b>	-1.19	-1.24	-1.07
Tc00.1047053504153.160	carboxypeptidase	-1.50	<b>-1.58</b>	-1.19	-1.29	-1.02
Tc00.1047053507873.20	cell differentiation protein	-1.14	-1.19	<b>-1.30</b>	-1.13	-1.16
Tc00.1047053507641.280	chaperonin HSP60, mitochondrial precursor	<b>-1.92</b>	<b>-1.85</b>	-1.18	-1.40	-1.23
Tc00.1047053510187.420	chaperonin HSP60, mitochondrial precursor (pseudogene)	-1.27	-1.19	-1.10	<b>-1.38</b>	-1.20
Tc00.1047053506247.50	chaperonin	-1.28	<b>-1.45</b>	-1.11	-1.17	-1.15
Tc00.1047053510187.270	COP-coated vesicle membrane protein erv25 precursor	-1.10	<b>-1.48</b>	-1.07	-1.09	-1.11

Tc00.1047053506529.550	cysteine peptidase	<b>-2.00</b>	<b>-2.50</b>	-1.38	<b>-2.23</b>	-1.35
Tc00.1047053508317.10	cysteine proteinase	-1.35	<b>-1.58</b>	-1.06	-1.13	-1.26
Tc00.1047053511391.160	cytochrome c1, heme protein, mitochondrial precursor	-1.42	-1.33	-1.13	<b>-1.63</b>	-1.39
Tc00.1047053510099.120	D-isomer specific 2-hydroxyacid dehydrogenase-protein	-1.07	-1.22	<b>-1.37</b>	-1.30	-1.21
Tc00.1047053510119.9	elongation factor 1-alpha (EF-1-alpha)	<b>-1.84</b>	<b>-2.22</b>	-1.30	<b>-1.57</b>	<b>-1.63</b>
Tc00.1047053508153.730	elongation initiation factor 2 alpha subunit	1.03	<b>-1.38</b>	-1.13	-1.06	-1.23
Tc00.1047053463269.10	eukaryotic translation initiation factor 1A	-1.17	-1.31	<b>-1.45</b>	-1.16	-1.07
Tc00.1047053506943.160	eukaryotic translation initiation factor 3 subunit 7-like protein	-1.06	-1.00	<b>-1.30</b>	1.05	1.01
Tc00.1047053506679.70	eukaryotic translation initiation factor 6 (eIF-6)	<b>-1.58</b>	<b>-1.92</b>	-1.33	-1.23	-1.14
Tc00.1047053511823.70	farnesyl pyrophosphate synthase	-1.43	<b>-1.55</b>	-1.32	-1.26	-1.21
Tc00.1047053511075.9	fatty acid desaturase	-1.41	<b>-1.59</b>	-1.11	-1.12	1.01
Tc00.1047053506661.30	fatty acid elongase	<b>-1.72</b>	<b>-1.65</b>	-1.40	-1.18	-1.30
Tc00.1047053507891.47	flagellar calcium-binding protein	-1.04	-1.02	-1.10	<b>-1.75</b>	-1.12
Tc00.1047053507547.90	glycosomal phosphoenolpyruvate carboxykinase	<b>-1.50</b>	<b>-1.98</b>	-1.18	<b>-1.43</b>	-1.06
Tc00.1047053508179.70	GPR1/FUN34/yaaH family	-1.31	-1.29	<b>-1.56</b>	-1.05	-1.01
Tc00.1047053503539.30	GTP-binding nuclear protein rtb2	<b>-2.18</b>	-1.37	1.16	-1.03	1.66
Tc00.1047053510131.40	haloacid dehalogenase-like hydrolase	-1.34	<b>-1.55</b>	<b>-1.90</b>	<b>-1.47</b>	<b>-1.60</b>
Tc00.1047053511041.40	hexose transporter	<b>-1.78</b>	<b>-1.98</b>	-1.02	-1.31	1.07
Tc00.1047053507943.40	histone H4	-1.11	<b>-1.35</b>	-1.25	-1.05	-1.09
Tc00.1047053509793.10	kinetoplast DNA-associated protein	-1.11	-1.19	<b>-1.46</b>	1.00	1.11
Tc00.1047053510667.14	membrane transporter protein	-1.09	-1.09	<b>-1.57</b>	1.03	-1.03
Tc00.1047053436521.9	mevalonate kinase	-1.19	-1.36	<b>-1.66</b>	-1.14	-1.19
Tc00.1047053436521.9	mevalonate kinase	-1.22	<b>-1.45</b>	-1.36	-1.18	-1.35
Tc00.1047053510155.20	mitochondrial RNA editing ligase 1	-1.18	<b>-1.48</b>	-1.22	-1.20	-1.19
Tc00.1047053508173.100	monooxygenase	-1.01	-1.16	<b>-1.40</b>	-1.22	-1.14
Tc00.1047053511817.40	NADH-cytochrome B5 reductase	<b>-1.90</b>	-1.49	1.21	1.06	-1.05
Tc00.1047053510645.20	nuclear transcription factor	-1.38	<b>-1.59</b>	-1.42	-1.14	-1.22
Tc00.1047053508173.180	nuclear transport factor 2 protein(NFT2)	-1.10	-1.31	<b>-1.52</b>	-1.27	-1.19

Tc00.1047053511573.58	nucleolar protein	-1.09	-1.17	<b>-1.78</b>	-1.07	-1.03
Tc00.1047053510859.17	nucleolar RNA-binding protein	-1.46	<b>-1.53</b>	-1.22	-1.17	-1.14
Tc00.1047053508707.200	nucleoside diphosphate kinase	-1.06	-1.02	-1.29	<b>-1.65</b>	-1.20
Tc00.1047053506773.50	nucleoside transporter-like	-1.33	<b>-1.52</b>	-1.09	-1.03	1.21
Tc00.1047053511355.30	phosphatidic acid phosphatase protein	-1.18	-1.21	<b>-1.79</b>	-1.02	-1.15
Tc00.1047053507617.9	prostaglandin F2alpha synthase	-1.22	<b>-1.66</b>	-1.08	-1.12	-1.01
Tc00.1047053508461.80	prostaglandin F2alpha synthase	<b>-1.72</b>	<b>-1.91</b>	-1.03	-1.32	-1.18
Tc00.1047053511635.40	protein tyrosine phosphatase	<b>-1.55</b>	<b>-1.67</b>	<b>-1.38</b>	-1.09	-1.12
Tc00.1047053510091.80	pyruvate dehydrogenase E1 beta subunit	<b>-1.65</b>	<b>-1.73</b>	<b>-1.59</b>	-1.34	-1.24
Tc00.1047053507251.20	ribosomal protein L21E (60S)	-1.06	-1.09	<b>-1.28</b>	<b>-1.46</b>	-1.18
Tc00.1047053508479.340	succinyl-CoA synthetase alpha subunit	<b>-1.73</b>	<b>-2.16</b>	<b>-1.56</b>	-1.39	-1.38
Tc00.1047053504045.60	thermostable carboxypeptidase 1	-1.26	<b>-1.52</b>	1.04	-1.08	-1.20
Tc00.1047053506855.260	thymidine kinase	-1.53	<b>-1.83</b>	-1.17	-1.35	-1.15
Tc00.1047053503555.30	trypanothione reductase	-1.33	<b>-1.39</b>	<b>-1.33</b>	-1.15	-1.16
Tc00.1047053508421.30	tryptophanyl-tRNA synthetase (pseudogene)	<b>-1.50</b>	<b>-1.73</b>	-1.33	-1.35	-1.24
Tc00.1047053503487.50	UDP-Gal or UDP-GlcNAc-dependent glycosyltransferase	1.12	-1.12	<b>-1.53</b>	1.05	-1.03
Tc00.1047053503929.10	V-type ATPase, A subunit	<b>-1.68</b>	<b>-1.58</b>	-1.22	-1.14	-1.14
Tc00.1047053511745.10	heat shock 70 kDa protein, mitochondrial precursor	1.08	1.02	-1.16	<b>-1.26</b>	-1.18
Tc00.1047053511715.100	pumilio/PUF RNA binding protein 7	1.06	1.04	<b>-1.71</b>	-1.06	1.02
Tc00.1047053509669.40	Zn-finger protein	-1.23	<b>-1.65</b>	1.11	-1.20	1.05

### Up-regulated

Tc00.1047053509941.100	2,4-dienoyl-coa reductase FADH1	1.10	1.29	<b>1.67</b>	1.33	1.32
Tc00.1047053504071.110	ama1 protein	1.01	1.17	1.41	1.44	<b>1.64</b>
Tc00.1047053508737.194	ARP2/3 complex subunit	1.03	1.24	<b>1.58</b>	1.40	1.33
Tc00.1047053508375.30	aspartate carbamoyltransferase	1.00	-1.13	<b>1.99</b>	-1.12	1.04
Tc00.1047053508183.4	aspartyl aminopeptidase	-1.09	1.09	<b>1.59</b>	1.05	-1.02
Tc00.1047053505997.70	ATPase	1.19	1.41	1.36	<b>1.59</b>	<b>1.50</b>
Tc00.1047053508903.100	ATPase	1.29	1.37	1.27	<b>1.52</b>	<b>1.88</b>
Tc00.1047053504149.20	ATP-binding cassette transporter ABCA1	<b>1.34</b>	1.24	1.06	1.06	1.07
Tc00.1047053506563.170	calpain-like cysteine peptidase (pseudo.)	-1.06	1.10	<b>1.38</b>	1.10	<b>1.31</b>

Tc00.1047053506563.210	calpain-like cysteine peptidase	1.03	1.19	1.43	<b>1.46</b>	1.23
Tc00.1047053508357.80	cation transporter protein	1.20	<b>1.74</b>	<b>1.75</b>	<b>1.82</b>	<b>1.55</b>
Tc00.1047053509455.140	cyclin	1.03	1.15	<b>1.67</b>	1.26	1.34
Tc00.1047053511421.110	developmentally regulated phosphoprotein	1.04	1.26	<b>1.59</b>	1.28	<b>1.59</b>
Tc00.1047053510687.10	dynein heavy chain (pseudogene)	1.07	-1.15	1.03	1.19	<b>1.47</b>
Tc00.1047053509585.10	dynein heavy chain	1.18	1.32	1.37	1.31	<b>1.56</b>
Tc00.1047053510897.6	dynein light chain lc6, flagellar outer arm	1.09	1.21	<b>1.34</b>	1.32	<b>1.45</b>
Tc00.1047053506391.30	EF-hand protein 5	-1.05	1.15	<b>1.88</b>	1.37	<b>1.61</b>
Tc00.1047053507483.20	EF-hand protein 5	-1.29	1.13	<b>2.15</b>	1.46	<b>1.56</b>
Tc00.1047053506529.508	glucose-6-phosphate isomerase, glycosomal	1.19	<b>1.66</b>	<b>2.69</b>	<b>2.06</b>	<b>1.92</b>
Tc00.1047053506529.508	glucose-6-phosphate isomerase, glycosomal	-1.03	1.34	<b>1.74</b>	<b>1.64</b>	1.32
Tc00.1047053508457.30	glycine cleavage system H protein	-1.11	1.15	<b>1.42</b>	1.15	1.33
Tc00.1047053510199.10	GP85-like protein	1.16	1.23	1.16	<b>1.32</b>	<b>1.54</b>
Tc00.1047053510659.240	lactoylglutathione lyase-like protein	1.04	1.35	1.57	<b>1.71</b>	1.36
Tc00.1047053510743.70	lactoylglutathione lyase-like protein	-1.23	1.32	<b>1.58</b>	1.32	1.11
Tc00.1047053473111.10	lathosterol oxidase	-1.11	1.65	<b>2.79</b>	1.98	<b>2.56</b>
Tc00.1047053508461.400	nucleoside diphosphate kinase	1.06	1.15	<b>1.67</b>	1.10	1.29
Tc00.1047053510105.100	UDP-glucose dehydrogenase	1.06	-1.02	<b>1.44</b>	1.16	1.17
Tc00.1047053508479.290	vacuolar sorting protein	1.07	-1.03	<b>1.89</b>	1.12	1.15
Tc00.1047053509069.30	tubulin binding cofactor A-like protein	-1.02	-1.02	1.52	1.78	<b>2.10</b>
Tc00.1047053506619.40	tyrosyl-DNA phosphodiesterase (Tdp1)	-1.00	1.01	<b>1.86</b>	1.29	1.32
Tc00.1047053503899.119	trypanothione/tryparedoxin dependent peroxidase 2	1.05	<b>1.33</b>	1.18	1.03	-1.06
Tc00.1047053508649.5	tryparedoxin peroxidase	-1.03	1.21	<b>1.46</b>	1.24	1.21
Tc00.1047053509907.60	trans-sialidase (pseudogene)	1.14	1.08	1.04	1.24	<b>1.41</b>
Tc00.1047053510121.130	serine/threonine protein kinase	1.41	1.07	1.29	<b>2.05</b>	1.77
Tc00.1047053511269.50	protein kinase A catalytic subunit	1.05	-1.11	<b>1.45</b>	1.34	<b>1.49</b>
Tc00.1047053508215.9	protein kinase C substrate protein, heavy chain	-1.09	1.14	<b>1.61</b>	1.39	1.24
Tc00.1047053504113.10	protein kinase	1.65	<b>1.62</b>	<b>1.60</b>	<b>1.42</b>	1.27
Tc00.1047053511671.80	protein kinase	1.29	1.38	1.12	<b>1.48</b>	1.41
Tc00.1047053507611.290	acetyltransferase	-1.04	-1.01	<b>1.42</b>	<b>1.40</b>	1.32
Tc00.1047053506357.50	alcohol dehydrogenase	-1.20	-1.17	<b>1.42</b>	1.16	1.11



Tc00.1047053507929.20	co-chaperone GrpE	1.09	1.16	<b>1.64</b>	-1.01	-1.03
Tc00.1047053506341.10	N-acetylglucosamine-6-phosphate deacetylase-like protein	1.23	<b>1.42</b>	1.03	-1.19	-1.17

---

Bolded fold-changes: statistically significant values. i.a.i.: immediately after irradiation

### Figure Legends:

Fig. 1: Number of differentially expressed genes in definite time points after *T. cruzi* irradiation. A) Scatter chart showing the distribution of fold-change values for statistically significant up-regulated genes over time points. B) Number of genes differentially expressed over time. ○: up-regulated, □:down-regulated. i.a.i: immediately after irradiation.

Fig. 2: Comparisons of microarray and qRT-PCR fold-change values for selected genes. Black columns: microarray results. Gray columns:qRT-PCR. \*:statistically significant results according to Linear model test for microarray and Friedman in association with Dunn tests for RT-PCR. i.a.i :immediately after irradiation.

Fig. 3: Categories of differentially expressed genes and their distributions along the time after *T. cruzi* irradiation. White: genes coding for hypothetical proteins. Light gray :obsolete sequences. Dark gray :genes coding for proteins with known function. Black: RHS genes. i.a.i: immediately after irradiation.

Fig. 4: Time course expression of kDNA genes after *T. cruzi* cells irradiation. COI: Cytochrome oxidase I; ND-1:NADH-dehydrogenase 1; ND-5: NADH-dehydrogenase 5; MURF-1 = Mitochondrial undefined reading frame 1. i.a.i: immediately after irradiation.

Fig. 5: Frequency of RHS genes and non-RHS genes up-regulated in response to treatment with gamma rays over time. White squares:Frequency of up-regulated RHS

genes. Black circles: Frequency of up-regulated genes, excluding RHS genes. i.a.i: immediately after irradiation. \*\*\*\*  $p < 0,0001$

Fig. 6: Gene Ontology Slim functional categories of differentially expressed genes. A) Biological Processes (BP) ,and B) Molecular Function (MF). Green bars: BP and MF categories repressed after gamma rays treatment. Red bars: BP and MF categories induced after gamma rays treatment.

Fig. 7: A) Heatmap of nuclear genes differentially expressed after *T. cruzi* irradiation. Shades of green indicate down-regulated and shades of red up-regulated genes. B) Clusters of genes based on similar expression profiles over time. Red-lines: average fold-change values.

Figure 1:

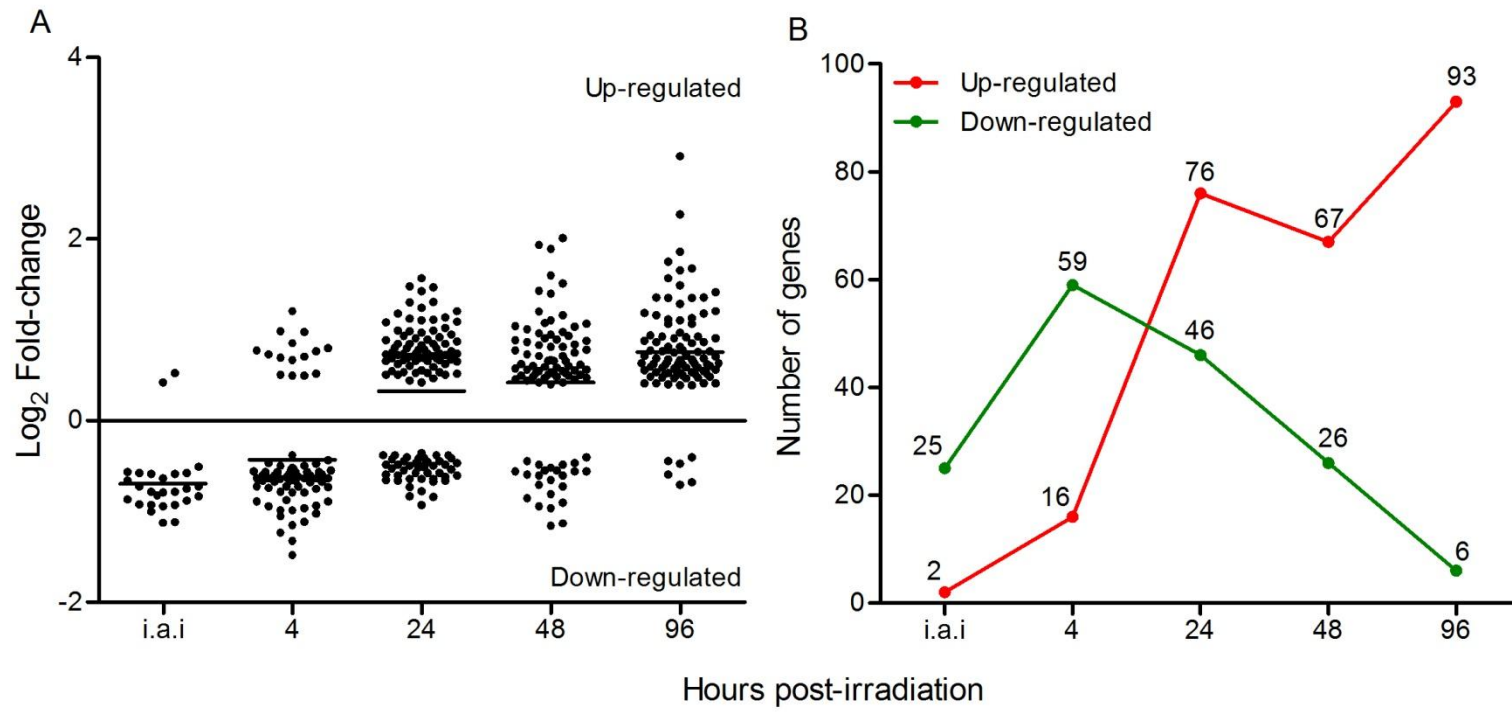


Figure 2:

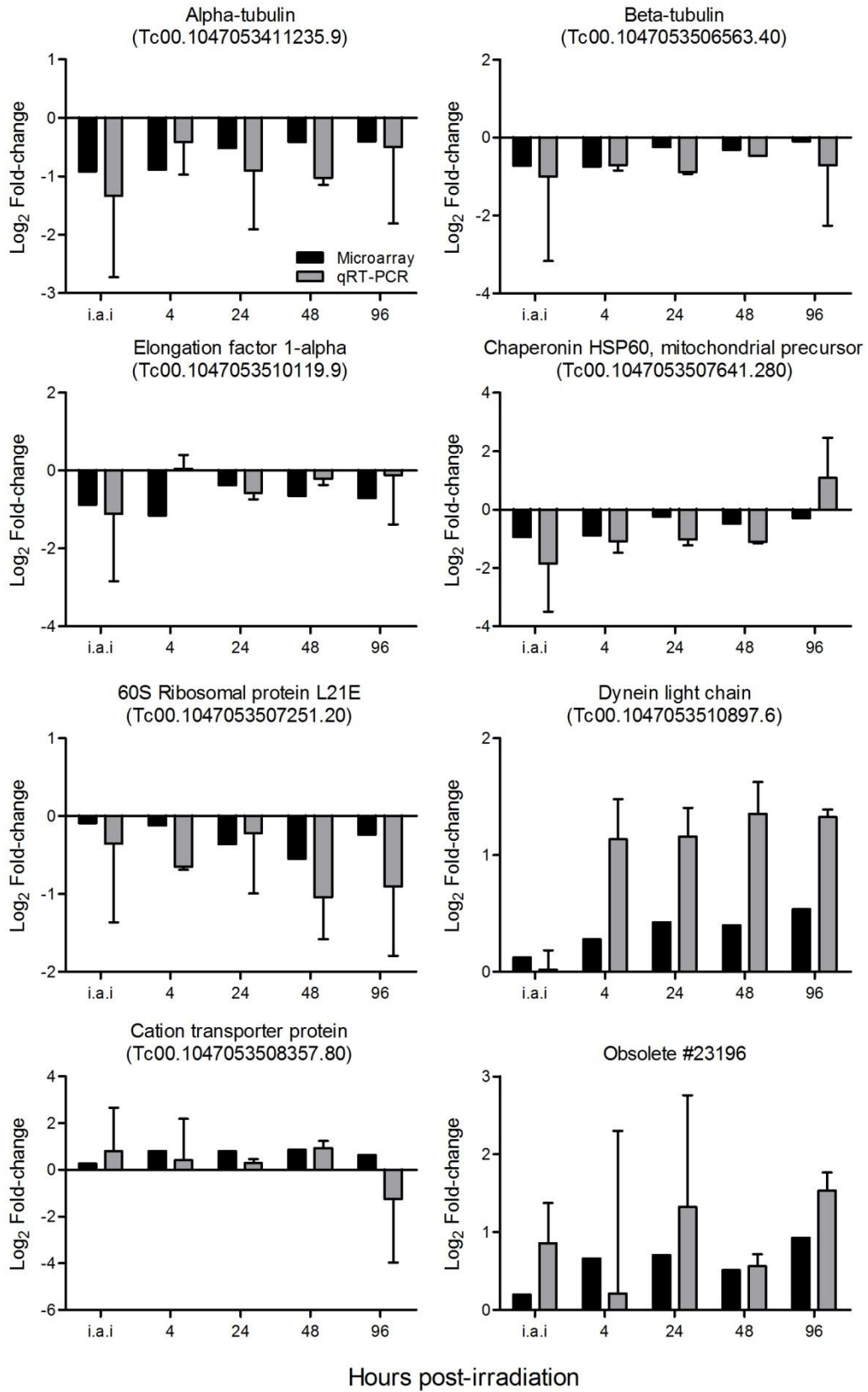


Figure 3:

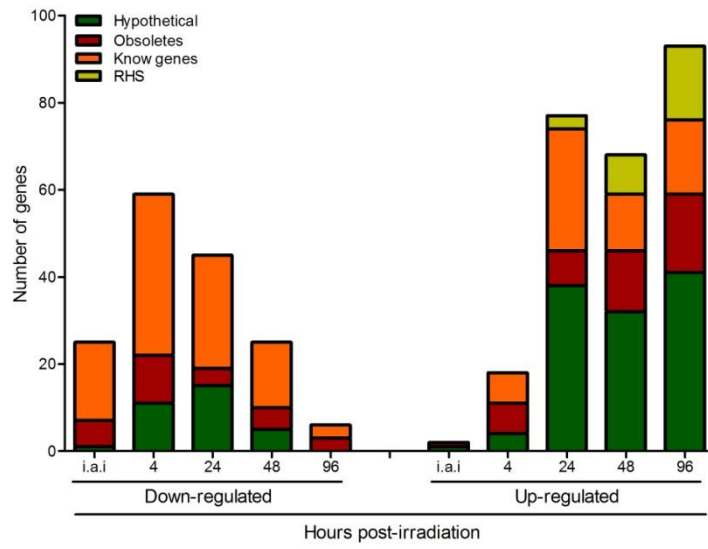


Figure 4:

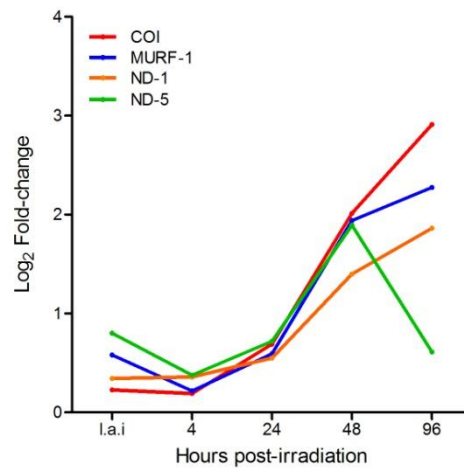


Figure 5:

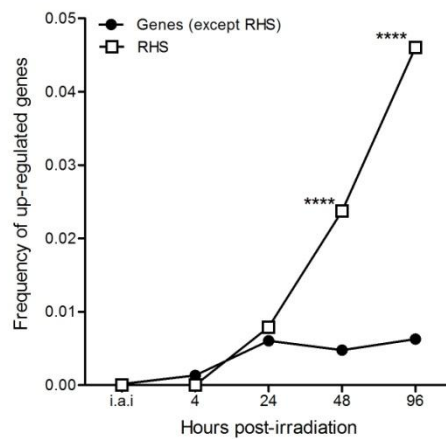


Figure 6:

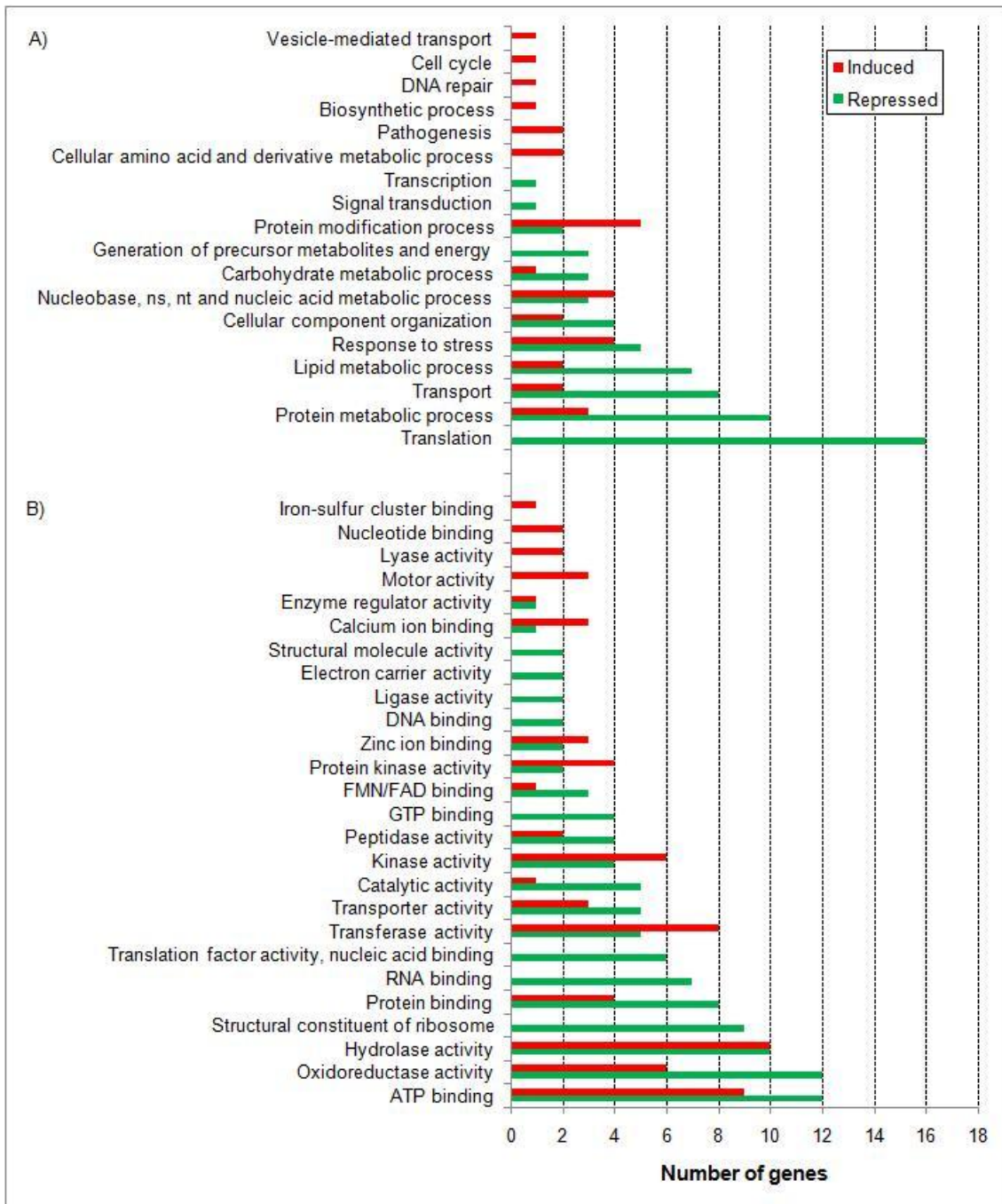


Figure 7:

