

UNIVERSIDADE FEDERAL DO PARANÁ

MICHEL BATISTA

**Efeito da Depleção de MAP-Kinases em
*Trypanosoma brucei***

**CURITIBA
2013**

MICHEL BATISTA

**Efeito da Depleção de MAP-Kinases em
*Trypanosoma brucei***

Tese apresentada ao Programa de Pós-Graduação em Biologia Celular e Molecular da Universidade Federal do Paraná, como parte das exigências para obtenção do título de Doutor em Biologia Celular e Molecular.

**Orientador: Dr. Marco Aurélio Krieger
Co-orientador: Dr. Fabricio Klerynton Marchini**

**CURITIBA
2013**

UNIVERSIDADE FEDERAL DO PARANÁ.
SISTEMA DE BIBLIOTECAS.
BIBLIOTECA CENTRAL.
COORDENAÇÃO DE PROCESSOS TÉCNICOS.
Ficha catalográfica

B333 Batista, Michel
Efeito da Depleção de MAP-Quinases em Trypanosoma
brucei / Michel Batista .-- Curitiba, PR , 2013.
140f. : il., gráfs., tabs.

Orientador : Marco Aurélio Krieger
Co-Orientador : Fabricio Kleryton Marchini
Tese(doutorado)- Universidade Federal do Paraná, Setor
de Ciências Biológicas, Programa de Pós-Graduação em
Biologia Celular e Molecular.
Inclui referências

1. Trypanosoma brucei. Doença do sono. 2. MAP Qui-
nases. 3. Proteínas quinases. I. Krieger, Marco Aurélio.
II. Marchini, Fabricio Kleryton. III. Universidade
Federal do Paraná. Setor de Ciências Biológicas. Programa
de Pós-Graduação em Biologia Celular e Molecular. IV.
Título.

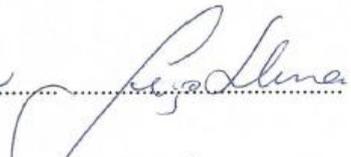
CDD 22.ed. 616.9363

PROGRAMA DE PÓS-GRADUAÇÃO EM BIOLOGIA CELULAR E MOLECULAR

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

ATA DA TRECENTÉSIMA QUADRAGÉSIMA SEGUNDA SESSÃO PÚBLICA DE DEFESA DO PROGRAMA DE PÓS-GRADUAÇÃO EM BIOLOGIA CELULAR E MOLECULAR

Aos vinte e nove dias do mês de Agosto de dois mil e treze, às 14h00min, nas dependências do Instituto Carlos Chagas (ICC), reuniu-se a banca examinadora designada pelo Colegiado do Programa de Pós-Graduação em Biologia Celular e Molecular, composta pelos Profs: Drs. Adriana Franco Paes Leme (Laboratório Nacional de Biociências - LNBio), Vanessa Santos Sotomaior (Pontifícia Universidade Católica do Paraná - PUCPR), Sérgio Schenkman (Universidade Federal de São Paulo - UNIFESP), Wanderson Duarte da Rocha (Universidade Federal do Paraná - UFPR), e Marco Aurélio Krieger (Orientador do Instituto Carlos Chagas - ICC), com a finalidade de julgar a Tese do candidato **Michel Batista**, intitulada: "Efeito da depleção de MAP-quinases sobre o proteoma e fosfoproteoma de *Trypanosoma brucei*", para obtenção do grau de Doutor em Biologia Celular e Molecular. O candidato teve em torno de 45 (quarenta e cinco) minutos para a apresentação e cada examinador, teve um tempo máximo de arguição de 30 (trinta) minutos, seguido de mais 30 (trinta) minutos para resposta do candidato ou de 60 (sessenta) minutos, quando houve diálogo na arguição. O desenvolvimento dos trabalhos seguiu o roteiro de sessão de defesa, estabelecido pela coordenação do Programa, com abertura, condução e encerramento da sessão solene de defesa feito pelo orientador, Dr. Marco Aurélio Krieger. Após analisar a tese e arguido o candidato, os membros da banca examinadora deliberaram pela "**APROVADO**", habilitando-o ao título de Doutor em Biologia Celular e Molecular, área de concentração em Biologia Celular e Molecular. A emissão do título está condicionada à implementação das correções sugeridas pelos membros da banca examinadora, bem como ao cumprimento integral das exigências estabelecidas no Regimento interno deste Programa de Pós-Graduação.

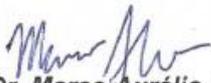
<u>Examinador/Afiliação</u>	<u>Parecer</u>	<u>Assinatura</u>
Prof. Dr. Marco Aurélio Krieger Orientador e presidente da banca Instituto Carlos Chagas - ICC	APROVADO	
Profa. Dra. Adriana Franco Paes Leme Laboratório Nacional de Biociências (LNBio)	APROVADO	
Profa. Dra. Vanessa Santos Sotomaior Pontifícia Universidade Católica do Paraná - PUCPR	APROVADO	
Prof. Dr. Sergio Schenkman Universidade Federal de São Paulo - UNIFESP	APROVADO	
Prof. Dr. Wanderson Duarte da Rocha Universidade Federal do Paraná - UFPR	APROVADO	

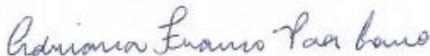
PROGRAMA DE PÓS-GRADUAÇÃO EM BIOLOGIA CELULAR E MOLECULAR

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

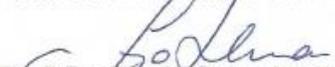
PARECER

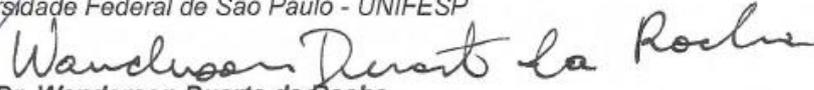
A banca examinadora, instituída pelo colegiado do Programa de Pós-Graduação em Biologia Celular e Molecular, do Setor de Ciências Biológicas, da Universidade Federal do Paraná, composta por:


Prof. Dr. Marco Aurélio Krieger
Orientador e presidente da banca
Instituto Carlos Chagas - ICC


Profa. Dra. Adriana Franco Paes Leme
Laboratório Nacional de Biotecnologia (LNBio)


Profa. Dra. Vanessa Santos Sotomaior
Pontifícia Universidade Católica do Paraná - PUCPR.


Prof. Dr. Sérgio Schenkman
Universidade Federal de São Paulo - UNIFESP


Prof. Dr. Wanderson Duarte da Rocha
Universidade Federal do Paraná - UFPR

E tendo como suplentes,

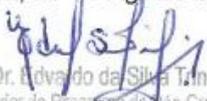
Prof. Dr. Alejandro Correa Dominguez
Instituto Carlos Chagas - ICC

Prof. Dr. Nilson Ivo Zanchin
Instituto Carlos Chagas - ICC

Após arguir o candidato **Michel Batista**, em relação ao seu trabalho intitulado: "Efeito da depleção de MAP-quinases sobre o proteoma e fosfoproteoma de *Trypanosoma brucei*", são de parecer favorável à **APROVAÇÃO** do acadêmico, habilitando-o ao título de Doutor em Biologia Celular e Molecular, área de concentração Biologia Celular e Molecular.

A obtenção do título está condicionada à implementação das correções sugeridas pelos membros da banca examinadora, bem como ao cumprimento integral das exigências estabelecidas no Regimento Interno deste Programa de Pós-Graduação.

Curitiba, 29 de Agosto de 2013


Prof. Dr. Eduardo da Silva Trindade
Coordenador do Programa de Pós-Graduação
em Biologia Celular e Molecular - UFPR
Câmp. 155795

À minha querida e adorável esposa, Kilma, e à nossa filha, Mirela.
Aos meus pais, Lucila e Alcides (*in memoriam*).
Ao meu irmão, Murillo.

AGRADECIMENTOS

Ao meu orientador Dr. Marco Krieger por ter me dado a oportunidade de iniciar a carreira científica e pela ajuda durante o trabalho.

Ao meu co-orientador Dr. Fabricio Marchini, pela ajuda, pelos ensinamentos, pela paciência e pela ativa participação neste trabalho.

A doutoranda Fernanda Kugeratski, pela parceria durante o projeto.

Ao Dr. Christian Probst pela ajuda nas análises de duplo-híbrido e de RNAseq.

Ao Dr. Henrique Preti pela ajuda nos ensaios de duplo híbrido

Ao doutorando Rafael Kessler pela ajuda nos ensaios de RNAseq.

Ao biólogo Paulo Arauco pela ajuda nos ensaios de RNAseq.

A doutoranda Carla Vanessa de Paula Lima pela ajuda durante o projeto.

A Dra. Lyris Martins Franco de Godoy pelas discussões sobre o projeto.

A toda a equipe do Nilson José Fidencio, o Silvio, a Sibelle, a Tania e a Vanessa pelo suporte durante todo o projeto.

A minha querida e paciente esposa Kilma que me deu apoio e suporte durante todo o meu doutorado.

A minha filha Mirela pelo amor e carinho que me deram forças.

A minha mãe Dona Lucila, pelo apoio e por sempre torcer por mim.

Ao meu pai Alcides, por todo o suporte e carinho.

Ao meu irmão Murillo, ao Noah e a Renata pela companhia, pelo espírito jovem, pelas risadas...

A Dona Silvia, Dona Julieta, pelo incentivo e ajuda.

Ao Dorneles, Juliana, Matheus e Murilo por acreditarem no meu propósito.

A família Silva, Valdecir, Adriana, Thiago e Rebeca, pelos importantes momentos de descontração.

Aos meus amigos do Instituto Carlos Chagas e IBMP, pelo carinho e apoio.

RESUMO

Os Tripanossomatídeos possuem um ciclo de vida complexo, alternando entre os ambientes distintos encontrados nos seus hospedeiros. Os mecanismos envolvidos na resposta do parasita às mudanças do ambiente ainda são pouco conhecidos, apesar de muito estudados. Dentre estes mecanismos, a fosforilação de proteínas possui um papel fundamental nas adaptações do parasita ao meio. Alguns dados, como a ausência de tirosino-quinases no genoma, a super representação de famílias de quinases, a existência de proteíno-quinases únicas, entre outros, evidenciam o processo de fosforilação como um tópico de grande interesse em Tripanossomatídeos. O protozoário *Trypanosoma brucei* possui 176 proteíno-quinases, divididas em diversas famílias, sendo que uma das mais bem estudadas, devido à sua posição central em processos biológicos essenciais, é o das quinases ativadas por mitógeno (MAPK). Tem sido descrita a participação destas proteínas em processos como divisão e diferenciação celular, resposta ao estresse celular, entre outros. No entanto, pouco é conhecido sobre a influência destas quinases na sinalização celular dos Tripanossomatídeos. Diante disso, o objetivo deste trabalho foi avaliar a influência da depleção de MAPKs sobre *T. brucei*. Dentre os resultados obtidos está a realização de curvas de crescimento, onde se verificou diminuição na proliferação celular em quase metade dos *knockdowns* analisados. O proteoma, fosfoproteoma e transcriptoma de um dos *knockdowns* (Tb927.10.10870) foi avaliado quantitativamente por proteômica baseada em espectrometria de massas combinada com marcação metabólica por SILAC e por RNA-Seq. No total, 1.511 proteínas, 1.510 sítios de fosforilação e cerca de 12 milhões de *reads* foram obtidos. Apesar da forte diminuição na proliferação celular deste *knockdown*, foi evidenciada pouca modulação no proteoma, fosfoproteoma e transcriptoma, indicando uma robustez na via de sinalização em estudo, corroborando com estudos publicados. Dentre as proteínas que sofreram modulação, aproximadamente um terço está relacionado à funções mitocondriais. Os sítios de fosforilação modulados após depleção da MAPK em estudo pertencem a proteínas com diferentes funções, no entanto com uma predominância de proteínas associadas à estabilização / degradação de mRNAs. A fim de se avaliar a interação entre os membros da via das MAPKs de *T. brucei*, fez-se a análise das interações através de duplo híbrido. Das 1.024 interações testadas, 22 foram positivas. A comparação com as interações entre os ortólogos destas proteínas, descritas em organismos modelo, não demonstrou sobreposição de nenhuma das interações positivas. Os resultados deste trabalho podem conduzir novos experimentos a fim de melhor compreender a via de sinalização das MAPKs no Tripanossomatídeos.

ABSTRACT

Trypanosomatids have a complex life cycle switching between different hosts. In this way, they find distinct environments where quick adaptation is needed. Among the cell signaling mechanisms, protein phosphorylation, driven by protein kinases and protein phosphatases, is an important process with potential to supply this necessity. Some finds on Trypanosomatids' genomes, such as, the absence of dedicated tyrosine or tyrosine kinase-like receptors, the over-representation of some kinase groups, the presence of unique kinases, and other, highlight the phosphorylation in Trypanosomatids as a hot spot topic. The flagellate parasite *Trypanosoma brucei* possesses 176 protein kinases classified in many families, among them, the mitogen activated protein kinase (MAPKs) family. Several publications have been documented the participation of this family in processes like cellular division and differentiation, in response to stress conditions, and other. Nevertheless little is known about its function in Trypanosomatids. In this sense, this work aimed to evaluate the effects of MAPK depletion on *T. brucei*. As a start point, the participation of MAPKs in the cellular proliferation was addressed. The results showed that almost a half of MAPK knockdowns had the cell proliferation rate affected. One of these knockdowns (Tb927.10.10870) was also quantitatively analyzed at the level of proteins, phosphoproteins, and mRNA transcripts by mass spectrometry based proteomics combined with metabolic labeling (SILAC) and by RNA-Seq. In total, 1,511 proteins, 1,510 phosphorylation sites, and nearly of 12 million of reads were identified. Despite the cell proliferation rate was strongly affected in this knockdown, few modulations in the proteome, phosphoproteome or transcriptome was observed. In consonance with other works, this could represent the robustness of MAPK signaling pathway in Trypanosomatids. Between the few modulations is worthy of note some mitochondrial proteins and phosphorylation sites on proteins related to mRNA stability control. As the MAPKs act in phosphorylation cascades, we also evaluate the interaction between the members of this signaling pathway by two hybrid assay. Among the 1,024 tested interactions, 22 were positive, showing no overlap with the interactions described for their ortologues in model organisms. Take together, the finds of this work have the potential to drive future projects in order to well understand the MAPK signaling pathway in Trypanosomatids.

SUMÁRIO

1. Introdução	15
1.1 <i>Trypanosoma brucei</i>	17
1.1.1 Ciclo de Vida de <i>Trypanosoma brucei</i>	18
1.2 Fosforilação de Proteínas em Tripanossomatídeos	19
1.3 Proteíno-Quinases Ativadas por Mitógeno.....	23
1.3.1 MAPKs em <i>T. brucei</i>	26
1.4 Metodologias de Trabalho	28
1.4.1 RNA de Interferência	29
1.4.2 Proteômica Baseada em Espectrometria de Massas	31
1.4.3 RNA-seq	42
1.4.4 Duplo-híbrido	42
1.4.5 Combinação de Metodologias de Trabalho	43
1.5 Proteômica e Fosfoproteômica em Tripanossomatídeos	44
2. OBJETIVO.....	46
3. JUSTIFICATIVA	47
4. MATERIAL E MÉTODOS.....	48
4.1 Desenho dos Oligonucleotídeos para RNAi	48
4.2 Amplificação dos Alvos para RNAi	48
4.3 Clonagem dos Alvos	49
4.4 Transfecção em <i>T. brucei</i>	50
4.5 Curvas de Crescimento	51
4.6 Avaliação dos Níveis de mRNA por QPCR	52
4.7 Meio de Cultivo e Incorporação dos Aminoácidos Pesados.....	53
4.8 Preparo das Amostras para Proteômica	53

4.9	Fracionamento da Amostra	54
4.10	Enriquecimento dos Fosfopeptídeos.....	55
4.10.1	Dióxido de Titânio – Protocolo 1.....	55
4.10.2	Dióxido de Titânio – Protocolo 2.....	56
4.10.3	Dióxido de Titânio – Protocolo 3.....	56
4.10.4	Dióxido de Titânio – Protocolo 4.....	57
4.10.5	Dióxido de Titânio – Protocolos 5 e 6.....	57
4.11	Análise no Espectrômetro de Massas.....	58
4.12	Análise dos Dados de Proteômica e Fosfoproteômica	58
4.13	RNA-seq	60
4.14	Duplo-Híbrido.....	61
5.	Resultados e Discussão	64
5.1	Padronizações	64
5.1.1	Elaboração do Meio de Cultivo para SILAC	64
5.1.2	Diminuição das Concentrações de Lisina e Arginina em SDM-79.....	66
5.1.3	Marcação Metabólica de Proteínas por SILAC	67
5.1.4	Fracionamento da amostra.....	70
5.1.5	Enriquecimento de fosfopeptídeos	70
5.2	Confirmação da Depleção de RNA Mensageiro Após RNAi	71
5.3	Efeito da Depleção das MAPKs na Proliferação Celular	73
5.4	Avaliação do transcriptoma após depleção da MAPKs A.....	79
5.4.1	Avaliação geral do experimento de RNAseq	79
5.5	Análise Proteômica	89
5.5.1	Fracionamento da amostra.....	89
5.5.2	Perfil geral do proteoma	91
5.5.3	Proteínas com expressão modulada	94

5.5.4	Proteínas diminuídas após depleção da MAPK-A	96
5.5.5	Proteínas aumentadas após a depleção da MAPK-A.....	97
5.6	Análise Fosfoproteômica.....	99
5.6.1	Perfil das análises.....	99
5.6.2	Perfil geral dos fosfossítios.....	103
5.6.3	Seleção dos fosfossítios modulados	108
5.6.4	Fosfossítios diminuídos após depleção da MAPK-A	111
5.6.5	Fosfossítios aumentados após depleção da MAPK-A.....	113
5.7	Avaliação da Interação entre proteínas da via das MAPKs.....	115
6.	Conclusão	121
7.	ANEXOS.....	122
8.	Referências Bibliográficas.....	131

LISTA DE FIGURAS

Figura 1: Ciclo de vida de <i>T. brucei</i>	19
Figura 2: Subdomínios das proteíno-quinases.....	24
Figura 3: Desenho esquemático da via de sinalização das MAPKs.....	25
Figura 4: Substratos e funções biológicas das MAPKAPKs.....	26
Figura 5: Desenho esquemático do sistema de RNAi.....	30
Figura 6: Sobreposição entre os sítios de fosforilação identificados por diferentes métodos de enriquecimento.....	41
Figura 7: Diagrama esquemático do perfil de fosforilação após depleção de uma MAPK.....	46
Figura 8: Vetor P2T7-177 modificado.....	50
Figura 9: Proliferação celular em diferentes concentrações de soro.....	64
Figura 10 - Curva de crescimento de <i>T. brucei</i> 29-13 para validação do meio SDM-79.....	66
Figura 11: Avaliação do crescimento de <i>T. brucei</i> em quantidades reduzidas de lisina e arginina.....	67
Figura 12: Avaliação da incorporação de Arg6 e Lys8 em <i>T. brucei</i>	69
Figura 13: Quantificação relativa dos mRNAs das MAPKs depletadas.....	73
Figura 14: Proliferação celular afetada após depleção das MAPKs.....	76
Figura 15: Proliferação celular pouco afetada após depleção das MAPKs.....	78
Figura 16: Comparação entre os níveis de mRNA de cada quinase nos <i>knockdowns</i>	81
Figura 17: Genes diminuídos após depleção da MAPK-A.....	85
Figura 18: Genes com expressão aumentada após depleção da MAPK-A.....	88
Figura 19: Correlação entre as intensidades NI/I das proteínas quantificadas entre as réplicas biológicas.....	92
Figura 20: Relações entre as intensidades das proteínas no ensaio de proteômica.....	93
Figura 21: Correlação geral entre a proteômica e o RNAseq.....	99
Figura 22: Correlação entre as intensidades NI/I dos fosfossítios nas 3 réplicas biológicas.....	100

Figura 23: Fosfossítios modulados normalizados pela proteômica e não normalizados.	103
Figura 24: Fosfossítios modulados normalizados pela proteômica, não normalizados e normalizados pelos peptídeos não fosforilados.....	103
Figura 25: Sobreposição dos fosfossítios identificados entre diferentes trabalhos..	107
Figura 26: Sobreposição das fosfoproteínas identificadas entre diferentes trabalhos.	108
Figura 27: Modulação nos fosfossítios de NOT5 quantificados..	112
Figura 28: Interações entre MAPKs detectadas por Y2H.....	119
Figura 29: União das redes de interação envolvendo ortólogos de proteínas das vias das MAPKs.....	120
Figura 30: Avaliação dos níveis de mRNA e do efeito da depleção das MAPKs na proliferação celular de <i>T. brucei</i>	130

LISTA DE TABELAS

Tabela 1: Classificação das quinases eucarióticas.	20
Tabela 2.: Comparação do repertório de quinases e fosfatases em diferentes genomas.....	22
Tabela 3: MAPKs identificadas em <i>T. brucei</i>	28
Tabela 4: Identificação das MAPKs e tamanho das sequências codificantes e proteínas.	49
Tabela 5: Parâmetros selecionados no algoritmo MaxQuant.	58
Tabela 6: Dosagem dos peptídeos pós digestão e pós fracionamento.	70
Tabela 7: Fosfossítios identificados durante a padronização do enriquecimento.	71
Tabela 8: Proporção dos reads mapeados por RNAseq.	80
Tabela 9: Número de genes modulados após depleção da MAPK-A.....	82
Tabela 10: Número de peptídeos identificados na proteômica nas diferentes frações do pré-fracionamento.	91
Tabela 11: Número de peptídeos e proteínas únicos por réplica quantificados na proteômica.....	94
Tabela 12: Proteínas com moduladas na proteômica.	95
Tabela 13: Resumo do número de proteínas, peptídeos e fosfossítios identificados.	100
Tabela 14: Fosfossítios quantificados por fração e por passo de enriquecimento dos fosfopeptídeo.....	101
Tabela 15: Número de fosfossítios normalizados pela proteômica.	102
Tabela 16: Classificação das proteíno-quinases identificadas na fosfoproteômica.	104
Tabela 17: Sobreposição dos fosfoproteomas identificados em diferentes trabalhos.	106
Tabela 18: Fosfossítios diminuídos no <i>knockdown</i> normalizados ou não normalizados com modulação mínima de 25%.	109
Tabela 19: Fosfossítios aumentados no <i>knockdown</i> normalizados ou não normalizados com modulação mínima de 25%.	110
Tabela 20: Interações detectadas e auto-ativação por duplo-híbrido.....	116
Tabela 21: Autoativações obtidas no ensaio de duplo híbrido.	117

Tabela 22: Proteínas não avaliadas por Y2H.....	117
Tabela 23: Oligonucleotídeos usados no RNAi e na QPCR.....	122
Tabela 24: Protocolo de preparo do meio SDM-79.	123
Tabela 25: Oligonucleotídeos utilizados nas ampliações dos alvos para duplo híbrido	125

1. INTRODUÇÃO

Os avanços científicos no século XX e início do século XXI certamente trouxeram benefícios sem precedentes para a ciência e, conseqüentemente, para a sociedade, disparando uma série de novas descobertas que vem aguçando a curiosidade da comunidade científica mundial. A compreensão a respeito da biologia, da física e da química passou por mudanças de paradigma, por exemplo, com a descoberta das partículas subatômicas, do raio laser, com a invenção dos microprocessadores e com a engenharia genética, entre outros vários avanços. O campo da medicina também tem seus méritos, diante da revolução no diagnóstico, prevenção e tratamento de agravos à saúde. Estes avanços somados contribuíram e vem contribuindo para o aumento não só da expectativa de vida das pessoas e também de vários animais, mas também da qualidade de vida. O mistério acerca do genoma começou a ser desvendado na década de 90, com o seu auge no ano 2000, quando a sequência do genoma humano foi publicada por cientistas americanos em uma parceria público-privada. A explosão que se seguiu no número de organismos com o seu genoma sequenciado é espantosa. As novas tecnologias de sequenciamento estão permitindo um aumento exponencial na velocidade com que os genomas são sequenciados, e uma redução nos custos, talvez na mesma magnitude.

Diante de tanta informação, ferramentas computacionais se tornaram indispensáveis para os pesquisadores da área da genômica. Tal necessidade vem sendo suprida pelos avanços na informática e pela interdisciplinaridade, talvez ainda inicial, cada vez mais necessária no campo da bioinformática. Toda a informação que vem sendo gerada pela genômica certamente tem contribuído enormemente para o conhecimento biológico, ao mesmo tempo em que também trouxe inúmeras perguntas acerca dos produtos gênicos, os RNAs e as proteínas. O primeiro grupo, principalmente os RNAs mensageiros, vem sendo abordados pela genômica funcional através de várias metodologias, com destaque para os microarranjos de DNA e mais recentemente para a metodologia de RNA-Seq. O estudo das proteínas, em geral os executores no contexto celular, abordando suas funções e interações representa um grande desafio e vem sendo explorado de várias maneiras.

Conhecendo-se a sequência de DNA das células, tornou-se possível iniciar o estudo a partir de uma proteína de interesse em sentido à investigação da sua função, este fluxo recebe o nome de genética reversa, em contraposição à genética clássica, onde a partir de fenótipos induzidos faz-se a busca pela proteína responsável por tal fenótipo. Contribuindo com a genética reversa, a engenharia genética permitiu a expressão heteróloga de proteínas, utilizando sistemas de expressão em bactérias, leveduras e células de mamíferos. A possibilidade de se obter proteínas purificadas favoreceu a obtenção de anticorpos como ferramenta na pesquisa, a realização de ensaios bioquímicos e de interação *in vitro*, a determinação da estrutura tridimensional das proteínas, entre outras metodologias, além de servir como insumo para tratamentos de algumas doenças.

Deixando este campo ainda mais interessante, no contexto celular, a atividade de muitas proteínas é controlada por modificações causadas por outras proteínas. Vários são os tipos de modificações pós traducionais (PTMs), como por exemplo, fosforilação, ubiquitilação, glicosilação, sumoilação, entre outras. O entendimento de quais proteínas são modificadas, em qual momento e por quais proteínas, pode contribuir para o esclarecimento dos vários mecanismos moleculares das células. As células possuem vias de sinalização, onde, por exemplo, um ligante externo pode ativar uma via de sinalização, culminando em algum tipo de resposta, divisão, diferenciação, entre outros, de acordo com o tipo e quantidade do ligante. Ou ainda, a ausência de uma determinada proteína pode ser necessária para a célula em determinado momento, então esta proteína pode ser marcada por ubiquitilação e, citando apenas um exemplo de mecanismo desta modificação, ser encaminhada para degradação.

Estudos de modificações pontuais vêm contribuindo para o entendimento de vários mecanismos celulares, no entanto, diante do vasto repertório proteico das células, surge a necessidade de estudos mais globais a respeito da influência das PTMs sobre as células. Várias ferramentas têm sido desenvolvidas no sentido de suprir tal necessidade, com destaque para a espectrometria de massas (MS). O desenvolvimento de novos equipamentos vem possibilitando a análise global do perfil proteico celular, incluindo as PTMs. Esta análise, realizada no extrato proteico total, fracionado ou enriquecido, trabalha com peptídeos oriundos da digestão

proteica, feita por proteases específicas e é conhecida como proteômica *shotgun* ou *bottom up*. Outra abordagem é a análise de proteínas intactas (*top down*), ou seja, proteínas não digeridas por proteases. Esta técnica tendo sido menos aplicada, devido ainda a algumas limitações técnicas. A proteômica, além de informar qualitativamente o perfil proteico de um extrato, também pode fornecer dados quantitativos a respeito das proteínas e suas PTMs, permitindo assim uma avaliação da dinâmica celular em consequência à regulação da expressão gênica ou em resposta a alguma perturbação no sistema. O entendimento das vias de sinalização e seus componentes, além de incrementar a ciência básica, pode indicar diferenças entre organismos, por exemplo, micro-organismos e mamíferos, possibilitando o desenho racional de fármacos para o tratamento de moléstias causadas por micro-organismos patogênicos. Diante de tantas possibilidades, a proteômica, aliada à genômica, interatômica e metabolômica pode fornecer informações sobre o funcionamento global das células, fomentando o campo da biologia de sistemas. Na tentativa de contribuir para a grande área da biologia de sistemas, aqui neste projeto alguns aspectos relacionados à dinâmica da fosforilação celular foram abordados. Tal estudo foi realizado usando-se como modelo o protozoário *Trypanosoma brucei*, cuja interação com mamíferos resulta em um agravo de impacto negativo para a saúde humana e animal, além de resultar em prejuízos econômicos para a sociedade, principalmente no continente africano.

1.1 *Trypanosoma brucei*

Descrito em 1902 pelo cientista escocês David Bruce, *Trypanosoma brucei* é um protozoário flagelado pertencente à Ordem Kinetoplastida, Família Trypanosomatidae. Sua interação com o homem pode resultar na tripanossomiase africana ou doença do sono africana, que acomete milhares de pessoas no continente africano. Sua transmissão ocorre através da saliva de insetos vetores contaminados, da Ordem Diptera, Família Glossinidae, conhecidos como tsé-tsé.

São conhecidas três subespécies de *T. brucei*, as quais são virtualmente idênticas bioquímica e morfológicamente, no entanto diferindo no tipo de hospedeiros e na virulência. *Trypanosoma brucei gambiense* e *Trypanosoma brucei*

rhodesiense causam as formas aguda e crônica da doença do sono em humanos, enquanto *Trypanosoma brucei brucei* não causa doença em humanos, mas sim em bovinos e outros mamíferos [1].

A tripanossomiase africana em humanos ocorre em dois estágios. No primeiro, geralmente é observado febre, dor de cabeça, prurido severo, lesões de pele, edema de face e extremidades. No segundo estágio, ocorrem desordens neuropsiquiátricas e endócrinas. São observadas alterações no ritmo circadiano influenciando a qualidade de sono do paciente. A fase terminal da doença é acompanhada de demência e o paciente fica susceptível a doenças oportunistas, o que pode levá-lo à morte [2].

A tripanossomiase em animais, denominada de nagana, na sua fase aguda pode causar febre, cansaço, perda de peso, perda de pelo, anemia e paralisia. A progressão da doença debilita progressivamente o animal, deixando-o inapto para o trabalho, no caso de animais destinados a esta finalidade. O termo nagana (de origem Zulu, “N’gana”) significa “sem força”, “sem uso” [3].

Além de representar um grave problema para a saúde pública, *T. brucei* possui algumas características biológicas bastante interessantes comuns aos Tripanossomatídeos. Dentre elas, a aparente ausência de controle transcricional em genes codificadores de proteínas, a presença de *trans-splicing* e a transcrição policistrônica constituem alguns exemplos de processos característicos desta família.

1.1.1 Ciclo de Vida de *Trypanosoma brucei*

O protozoário *T. brucei* apresenta um ciclo de vida no qual passa por dois hospedeiros, um inseto e um mamífero. Durante o ciclo de vida a célula necessita se adaptar às condições do ambiente, sendo que tais condições são causadas, por exemplo, pelo sistema imune do hospedeiro mamífero, pelo aporte de nutrientes distinto entre os hospedeiros, entre outros. Uma forma utilizada pelo parasita para evadir ao sistema imune do hospedeiro mamífero é a expressão de uma capa antigênica formada pelas glicoproteínas variáveis de superfície (*variable surface glycoproteins* – VSGs) [4].

Na Figura 1 está esquematizado o ciclo de vida de *T. brucei*, com informação adicional de algumas características das diferentes formas. Aspectos como a expressão de VSGs, atividade mitocondrial, localização do cinetoplasto e proliferação celular também estão contidos na Figura 1. No intestino do inseto, as formas ingeridas se diferenciam em tripomastigotas procíclicas, que após três semanas migram para a glândula salivar do inseto e se diferenciam nas formas tripomastigotas metacíclicas. Estas estão presentes na saliva do inseto e são injetadas no mamífero, onde se diferenciam rapidamente nas formas sanguíneas *slender*, as quais se replicam na corrente sanguínea. No hospedeiro mamífero pode existir uma população pleomórfica formada pelas formas sanguíneas *slender* e formas sanguíneas mais curtas e não replicativas chamadas *stumpy*, que ao serem ingeridas pelo inseto irão se diferenciar nas formas procíclicas fechando o ciclo.

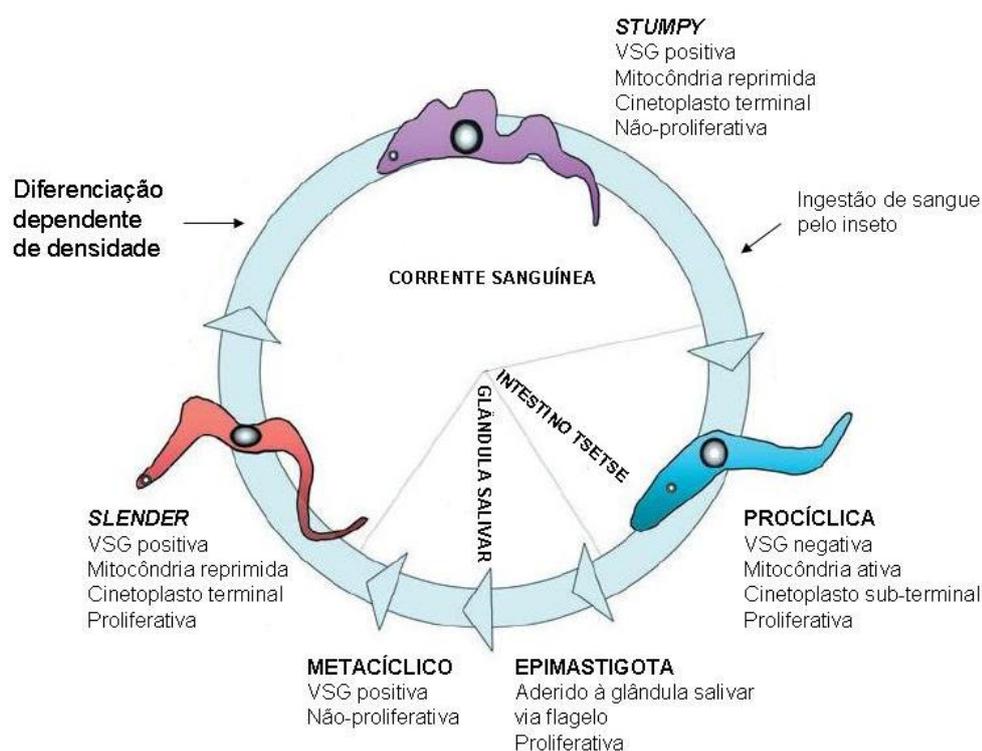


Figura 1: Ciclo de vida de *T. brucei*. Adaptado de Matthews, 2005 [5].

1.2 Fosforilação de Proteínas em Tripanossomatídeos

As respostas celulares aos sinais presentes nos meios extra e intracelular envolvem mecanismos diversos, podendo desencadear vários processos, como por

exemplo, expressão gênica, divisão celular, diferenciação celular, apoptose, entre outros. A regulação da expressão gênica de forma pós-transcricional é um conjunto de mecanismos capaz de modular processos celulares de forma rápida. As modificações pós-traducionais de proteínas (*Protein post-Translational Modifications* - PTMs) são exemplos destes mecanismos. Dentre as PTMs, pode-se citar a ubiquitilação, a glicosilação, a fosforilação, entre outras. A fosforilação proteica, foco deste trabalho, é a transferência de um grupamento fosfato (98 Da) de uma molécula de ATP para um resíduo de aminoácido específico. Tais aminoácidos geralmente são serina, treonina e tirosina. A fosforilação é coordenada pela ação de proteíno-quinases (quinases) e proteíno-fosfatases (fosfatases).

A superfamília das quinases pode ser classificada em grupos de acordo com a similaridade dos domínios catalítico e regulatório. As quinases eucarióticas podem ser classificadas em oito grupos convencionais (ePKs) e quatro grupos atípicos (aPKs) [6], [7] de acordo com a tabela Tabela 1.

Tabela 1: Classificação das quinases eucarióticas.

GRUPO	DESCRIÇÃO
CONVENCIONAIS	
AGC	quinases dependentes de nucleotídeo cíclico e dependentes de cálcio-fosfolípido, quinases que fosforilam proteína ribossomal S6, quinases acopladas à proteína G.
CAMK	quinases reguladas por calmodulina.
CK1	quinase caseína 1.
CMGC	quinases dependentes de ciclina (CDK), quinases ativadas por mitógeno (MAPK), quinases glicogênio sintase, quinases relacionadas às CDKs (CDKL).
RGC	receptor guanilato ciclase.
STE	envolvidas nas cascatas das MAPKs.
TK	tirosino-quinases.
TKL	quinases relacionadas às tirosino-quinases.
ATÍPICAS	
quinase-alpha	quinase da cadeia pesada da miosina em <i>Dictyostelium discoideum</i> , por exemplo.
PIKK	quinases relacionadas às quinases fosfatidil-inositol 3'.
RIO	participam da biogênese ribossomal.
PHDK	quinases piruvato desidrogenase.

Os Tripanossomatídeos patogênicos *Trypanosoma brucei*, *Trypanosoma cruzi* e *Leishmania major*, denominados pelos pesquisadores da área como *Trityps*, durante seus ciclos de vida necessitam se adaptar de forma rápida às condições distintas encontradas nos seus hospedeiros, como já mencionado anteriormente. Desta forma, a sinalização por fosforilação proteica é de grande importância para estes parasitas. Em 2005, em artigo publicado na *BMC Genomics*, Parsons e colaboradores [8] fizeram uma comparação entre os conjuntos de quinases (quinomas) de *T. brucei*, *T. cruzi* e *L. major*. Este trabalho demonstrou a presença de um grande número de quinases nestes organismos (176 em *T. brucei*, 190 em *T. cruzi*, 199 em *L. major*), sendo possível através da análise das suas sequências a organização da maioria das quinases em sete grupos, incluindo um grupo de quinases não usuais, as quais podem constituir importantes alvos terapêuticos. Alguns anos depois, Nett e colaboradores [9] analisaram o fosfoproteoma de *T. brucei*, e ainda gerando dados adicionais sobre a classificação do quinoma de *T. brucei*. O número de representantes de quinases e fosfatases nos *Trityps*, comparando com *Plasmodium falciparum* e *Homo sapiens* está resumido na Tabela 2.

A análise das quinases por similaridade de sequência fornece um panorama geral do quinoma dos Tripanossomatídeos, mas somente com este estudo não é possível a atribuição de função a estas proteínas. Sendo assim, é necessária a utilização de outras ferramentas para o entendimento das interações envolvendo as quinases, o que pode nos fornecer indícios sobre as suas funções biológicas. Alguns exemplos destas ferramentas incluem estudos por genética reversa, *knockdown* e *knockout* gênico.

Tabela 2.: Comparação do repertório de quinases e fosfatases em diferentes genomas.

Grupo/Família	<i>T. brucei</i>	<i>T. cruzi</i>	<i>L. major</i>	<i>P. falciparum</i>	<i>H. sapiens</i>
Proteíno-quinases					
AGC	12(8) ^a , 22(13) ^b	12(7)	11(6)	14(16)	82(17)
CAMK	14(9) ^a , 28(16) ^b	13(8)	16(9)	20(23)	95(20)
CK I	5(3) ^a , 5(3) ^b	8(5)	7(4)	1(1)	12(3)
CMGC	42(27) ^a , 47(28) ^b	42(25)	45(25)	19(22)	68(14)
Outras	39(25) ^a , 20(12) ^b	42(25)	40(22)	27(31)	21(4)
STE	25(16) ^a , 28(16) ^b	31(18)	34(19)	1(1)	61(13)
Tirosino-quinase	0(0) ^a , 0(0) ^b	0(0)	0(0)	0(0)	91(19)
TKL	0(0) ^a , 0(0) ^b	0(0)	0(0)	5(6)	48(10)
Exclusivas*	19(12) ^a	23(13)	26(15)	-	-
Total	156^a, 170^b	171	179	87	581
Proteíno-fosfatases					
Proteíno tirosino fosfatase	2(2)	2(3)	3(3)	3(9)	43(31)
Fosfatase com dupla especificidade	20(23)	19(24)	22(25)	1(3)	43(31)
Lipídeo fosfatase	7(8)	3(4)	4(5)	0(0)	21(15)
Serino/Treonino fosfatase	56(65)	54(69)	58(66)	30(88)	33(24)
Total	86	78	88	34	140

Quinases: [8] (*T. brucei*^a, *T. cruzi* e *L. major*), [9] (*T. brucei*^b, *P. falciparum* e *H. sapiens*). Fosfatases: [10] (*T. brucei*, *T. cruzi* e *L. major*), [11] (*P. falciparum*), [12], [13], [14], [15] (*H. sapiens*). *Não foram encontrados genes ortólogos aos membros deste grupo. Em parênteses, a porcentagem representativa de cada grupo.

Em relação aos grupos de quinases encontrados nos *Tritryps*, os grupos CAMK e AGC possuem poucos genes representantes, enquanto o grupo das tirosino-quinases não possui representantes no genoma dos *Tritryps* [8]. Apesar disso, fosforilações em resíduos de tirosina tem sido documentadas na literatura

[16]. Em contrapartida, os grupos CMGC, STE e NEK estão relativamente bem representados dentro do quinoma destes parasitas quando comparados com humanos, ou seja, a relação entre o número de genes dos grupos mencionados e o total de quinases é maior do que esta mesma relação feita com os dados do genoma humano. Todas as famílias CMGC identificadas em humanos estão também representadas em Tripanossomatídeos [8]. Outro dado relevante é o grande número de genes relacionados às quinases ativadas por mitógeno (MAPK) presente em Tripanossomatídeos. Este fato pode refletir o papel das cascatas de MAPK nas respostas coordenadas às condições do ambiente [8]. Além dos grupos de quinases clássicos, os *Trityps* apresentam um grupo de quinases únicas, as quais não puderam ser agrupadas com nenhuma quinase dos organismos analisados por Parsons e colaboradores (*Homo sapiens*, *Drosophila melanogaster*, *Caenorhabditis elegans* e *Saccharomyces cerevisiae*).

1.3 Proteíno-Kinases Ativadas por Mitógeno

A atividade das proteíno-quinases é catalisar a transferência de um grupo fosforil, proveniente do fosfato gama da molécula de ATP, para o grupo hidroxil de resíduos de serina, treonina ou tirosina em um substrato proteico, representado no esquema a seguir:



As proteíno-quinases possuem um domínio de quinase característico composto por onze subdomínios (Figura 2). Alguns subdomínios e resíduos de aminoácidos específicos são requeridos para a atividade catalítica das quinases, tais como o subdomínio I, responsável pela ligação ao ATP, e resíduos como lisina no subdomínio II e ácido aspártico nos subdomínios VI e VII [17].

Uma propriedade fundamental das células é responder às mudanças ambientais e estímulos diversos. Tais respostas necessitam de mecanismos moleculares para se efetivarem. A cascata de fosforilação das proteíno-quinases ativadas por mitógeno (MAPKs) é um destes mecanismos. As MAPKs pertencem ao

grupo de quinases CMGC (*Cyclin dependent kinases (CDK)*, *Mitogen-activated protein kinases*, *Glycogen synthase kinases* e *CDK-like kinases*).

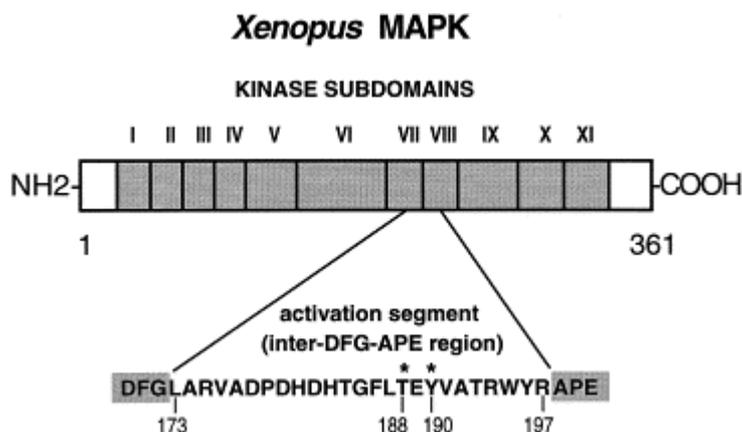


Figura 2: Subdomínios das proteíno-quinases. Fonte: [18]

Uma característica das MAPKs é a sua ativação pelas MAP2Ks (grupo STE) através da fosforilação em um *loop* de ativação, normalmente em tirosina e treonina mutuamente. As MAP2Ks, por sua vez, são ativadas por MAP3Ks.

As proteíno-quinases classificadas como *extracellular signal-regulated kinase* 1 e 2 (ERK1, ERK2) foram as primeiras MAPKs clonadas e caracterizadas de mamíferos [19, 20]. Estas quinases são ativadas por fatores de crescimento, tais como fator de crescimento derivado de plaquetas (PDGF), fator de crescimento epidermal (EGF), fator de crescimento do nervo (NGF), em resposta à insulina, por receptores associados à proteína G, citocinas, estresse osmótico e desorganização dos microtúbulos [20, 21]. Entre as atividades de ERK1/ERK2 está a participação no controle da progressão celular, mais especificamente na transição entre G1 e S. As MAPKs ERK1/ERK2 fazem parte das MAPKs convencionais, juntamente com ERK5, p38 (α , β , γ , δ) e JNK1/2/3. As MAPKs atípicas são representadas por ERK3/4/7/8 e NLK. Tanto as MAPKs convencionais quanto as atípicas ativam por fosforilação uma série de proteínas chamadas proteíno-quinases ativadas por MAPK (MAPKAPK) (Figura 3).

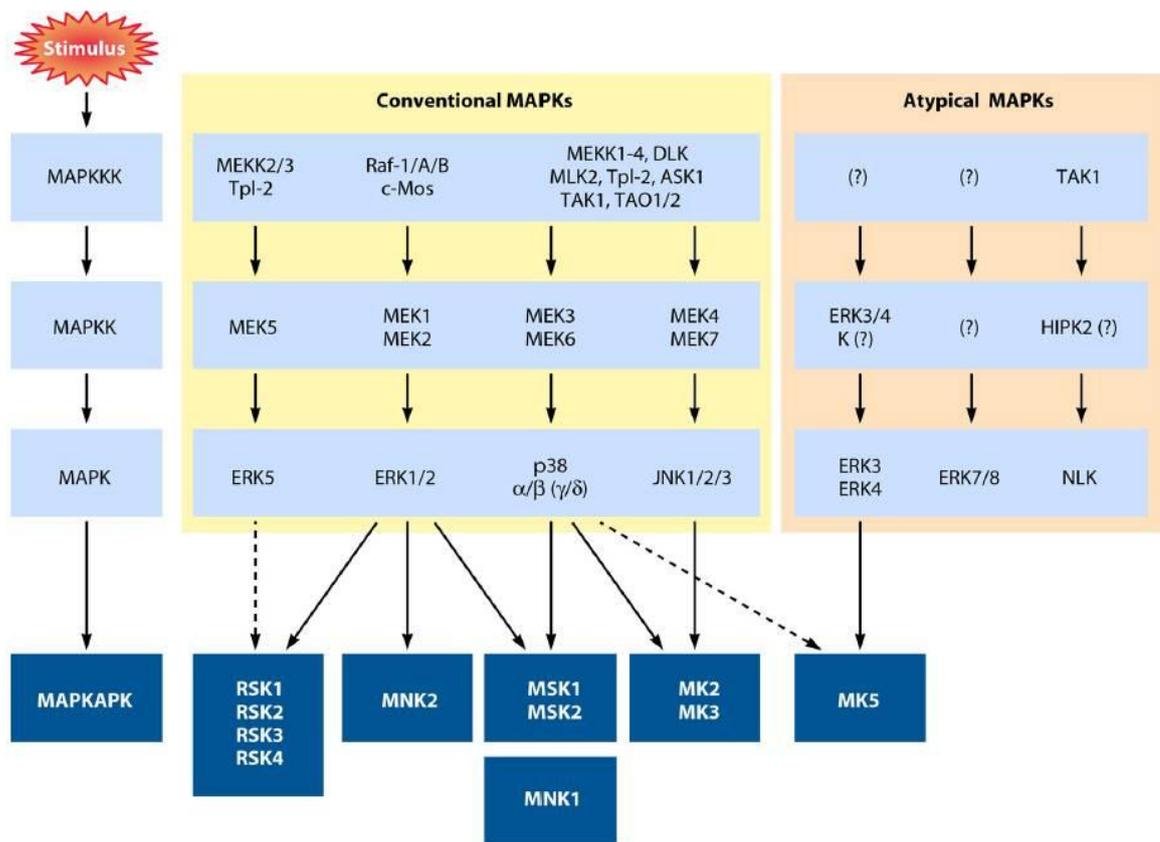


Figura 3: Desenho esquemático da via de sinalização das MAPKs [22].

As MAPKAPKs possuem diversos substratos, estes por sua vez associados a uma variedade de funções na célula. Um resumo das principais funções e substratos das MAPKAPKs pode ser visualizado na

RSK1-4	MSK1/2	MNK1/2	MK2/3	MK5
<p>Gene transcription SRF, Nur77, c-Fos, ER81, TIF-1A, ERα, p65, ATF4</p> <p>Cell proliferation c-Fos, p27^{kip1}, MAD1, ILK1, Myt1, Emi2, GSK3</p> <p>Cell growth rpS6, TSC2, Raptor, eIF4B, LKB1, GSK3, eEF2 kinase</p> <p>Cell survival C/EBPβ, DAPK</p> <p>Others NHE-1, Filamin A, RanBP3, YB-1, nNOS, AS160</p>	<p>Gene transcription CREB, ATF1, p65, ER81, p53</p> <p>Nucleosomal response Histone H3, HMG1</p>	<p>mRNA translation eIF4E, eIF4G</p> <p>Others hnRNP A1, PSF, cPLA₂, Spry2</p>	<p>Actin remodeling and cell migration Hsp25/27, p16-Arc, LSP-1, CapZIP, vimentin, $\alpha\beta$-crystallin</p> <p>Cytokine production hnRNP A0, TTP, PABP1, HuR, BRF1</p> <p>Gene transcription BMI1, SRF, ER81, E47</p> <p>Cell cycle control CDC25B, CDC25C, HDM2, TSC2</p>	<p>Tumor suppression p53?</p> <p>Actin remodeling Hsp25/27?</p>

Figura 4: Substratos e funções biológicas das MAPKAPKs [22].

As MAPKs são bastante estudadas em *Saccharomyces cerevisiae*, onde está descrita a sua participação em diversas vias de transdução de sinal. Entre elas pode-se citar a via de resposta a feromônios (Fus3), a via de ajuste às condições nutricionais limitantes (Kss1), a via de sobrevivência em condições hiperosmóticas (Hog1), a via de reparo a danos na parede celular (Sit2/Mpk1) e via de montagem da parede em esporos durante a meiose e esporulação (Smk1, Ime2) [23].

1.3.1 MAPKs em *T. brucei*

Em *T. brucei* foram identificados 15 genes relacionados às MAPKs, 5 às MAP2Ks, 14 às MAP3Ks, 2 às MAP4Ks e 4 não relacionados com os anteriores. Dentre as 15 MAPKs, uma parte dos genes é mais próxima às MAPKs e outra mais próxima às famílias CDKL e RCK (Tabela 3). Algumas MAPKs já foram estudadas em *T. brucei*, como por exemplo, KFR1, TbMAPK2 e TbECK1. Estes genes estão relacionados à proliferação celular, diferenciação e ciclo de vida, respectivamente [24][25][26][27].

Hua e wang mostraram a homologia de KFR1 com os genes de levedura *FUS3/KSS1* (40% de identidade) e com *ERK1/ERK2* de *Xenopus* e de mamíferos. Neste trabalho foi analisada a atividade de KFR1, sendo descrita maior atividade *in vitro*, maior resposta à privação de soro e maior resposta após estímulo com

interferon-gama na forma sanguínea em comparação com a forma procíclica. O nocaute do gene homólogo de KFR1 em *Leishmania mexicana* levou a uma diminuição na proliferação celular na forma amastigota, tanto em macrófago quanto em cultivo livre de célula [28].

Em 2002, Muller e colaboradores demonstraram a participação da proteína TbMAPK2 tanto na manutenção da cinética de diferenciação da forma sanguínea para a forma procíclica de *T. brucei* quanto na proliferação após diferenciação. Neste trabalho não foi possível o nocaute do segundo alelo diretamente na forma procíclica. Este fato também foi relatado por Hua e Wang, que, da mesma forma que Muller, atribuíram este fato à importância das quinases KFR1 e TbMAPK2 para a forma procíclica.

Ellis e colaboradores descreveram, através de expressão ectópica, a influência da extensão C-terminal de TbECK1 no crescimento celular e no cariótipo. Transfectantes contendo TbECK1 sem sua extensão C-terminal apresentaram diminuição no crescimento celular e células com número de núcleo e cinetoplasto variáveis. Em contrapartida, transfectantes contendo TbECK1 íntegra não apresentaram fenótipo distinto das células selvagens.

A participação de pelo menos uma MAPK (TbMAPK5) na diferenciação da forma *slender* (proliferativa) para a forma *stumpy* (não proliferativa) de *T. brucei* sanguíneo foi observada. O nocaute de TbMAPK5 causou diferenciação prematura tanto em camundongo quanto em cultivo celular [27].

Além dos estudos utilizando *T. brucei* como modelo, trabalhos abordando MAPKs de *Leishmania* também foram publicados. Além do trabalho já citado de 1998, Wiese e colaboradores [29], através de PCR utilizando oligos degenerados, descreveram oito homólogos de MAPKs em *L. mexicana*. Uma característica comum a estes homólogos é a presença de uma extensão C-terminal contendo domínios similares aos domínios SH3, os quais possuem a característica de mediar interações proteicas. Dois anos mais tarde, foi demonstrado, também em *L. mexicana*, o envolvimento de MAPKs na morfogênese do flagelo (LmxMPK9) e com importância em ambos os estágios amastigota e promastigota (LmxMPK4), [30, 31].

Apesar dos vários trabalhos caracterizando MAPKs em Tripanossomatídeos, a descrição dos alvos moleculares destas enzimas e seus respectivos ativadores

ainda é um campo a ser explorado. A única interação descrita entre MAPKs em Tripanossomatídeos foi entre uma MAP2K (LmxMKK5) e uma MAPK (LmxMPK4) de *Leishmania*. Esta interação foi observada *in vitro* através da co-expressão destas proteínas em *E. coli* [32].

Tabela 3: MAPKs identificadas em *T. brucei*.

GRUPO/FAMÍLIA	COMENTÁRIOS	IDENTIFICAÇÃO	NOME EM <i>T. brucei</i>
CMGC	MAPK- <i>like</i>	Tb927.10.10870	
CMGC/MAPK	falta sequência regulatória de MAPK	Tb927.10.3230	
CMGC/MAPK		Tb927.8.3770	
CMGC/MAPK		Tb927.8.3550	MAPK3
CMGC/MAPK		Tb927.10.16030	MAPK2
CMGC/MAPK		Tb927.6.1780	
CMGC/MAPK		Tb927.6.4220	MAPK5
CMGC/MAPK		Tb927.10.12040	
CMGC/MAPK		Tb927.10.5140	
CMGC/MAPK		Tb927.10.7780	KFR1, MAPK1
CMGC/CDKL	MAPK- <i>like</i>	Tb927.11.8170	
CMGC/CDKL	MAPK- <i>like</i>	Tb927.11.16790	ECK1
CMGC/RCK	MAPK- <i>like</i>	Tb927.10.14800	MAPK9
CMGC/RCK	MAPK- <i>like</i>	Tb927.3.690	
CMGC/RCK-MAPK		Tb927.9.9320	MOK

Parsons et al., 2005 – material suplementar.

1.4 Metodologias de Trabalho

Diversas são as metodologias de trabalho dentro da biologia molecular para o estudo funcional dos genes, o que possibilita estudos sobre a sinalização celular. Dentre estas opções estão as abordagens por genética reversa, onde, a partir da sequência de um determinado gene é possível avaliar os fenótipos nos quais existe a participação deste gene. A genética reversa necessita do conhecimento prévio da

sequência nucleotídica dos genes, logo, com o crescimento do número de genomas sequenciados, tornou-se possível a sua aplicação no estudo de um grande número de organismos. Alguns exemplos de metodologias envolvendo genética reversa são o silenciamento gênico através de nocaute ou RNA de interferência (RNAi), a expressão homóloga e heteróloga de genes para diversas finalidades, tais como dominante negativo, fusão com etiquetas moleculares, obtenção de proteínas para produção de anticorpos ou estudos bioquímicos de atividade.

Técnicas de estudo em larga-escala, também dependentes do conhecimento prévio do genoma do organismo a ser estudado, tais como sequenciamento de RNAs para estudos de genômica funcional, espectrometria de massas para estudos de proteômica, duplo-híbrido em leveduras para estudos de interação proteica tem sido empregadas na busca pelo conhecimento dos sistemas biológicos. No trabalho aqui apresentado foram empregadas metodologias de genética reversa utilizando RNAi como ferramenta para o silenciamento gênico, espectrometria de massas e RNAseq para a avaliação dos efeitos de silenciamentos gênicos específicos sobre o proteoma / fosfoproteoma e transcriptoma celular, respectivamente e duplo-híbrido para a avaliação de interações proteicas envolvendo as proteínas aqui estudadas. Na sequência serão abordados alguns aspectos sobre estas metodologias.

1.4.1 RNA de Interferência

O RNA de interferência (RNAi) é um mecanismo de defesa em células eucarióticas contra retrotransposons. A descrição inicial do fenômeno de RNA de interferência ocorreu no nematódeo *Caenorhabditis elegans* [33] e em seguida em outros organismos, bem como no organismo foco deste trabalho, *T. brucei* [34]. A degradação de um RNA mensageiro (mRNA) específico pode ser induzida pela presença deste mesmo RNA, no entanto na conformação fita dupla. Tal conformação é reconhecida pelo sistema que envolve duas ribonucleases, Dicer e Slicer (AGO1). Pela presença do RNA fita dupla, o RNA fita simples, ou seja, o mRNA é degradado, causando o silenciamento do gene em questão (Figura 5).

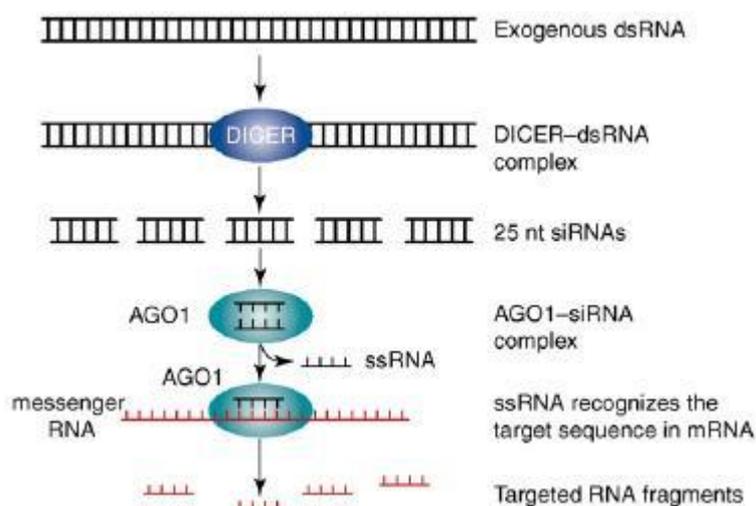


Figura 5: Desenho esquemático do sistema de RNAi [35].

Este mecanismo pode ser utilizado intencionalmente para diminuir a expressão de um determinado gene, possibilitando assim a inferência na sua função, que pode ser avaliada por análise fenotípica, proteômica, transcriptômica, entre outras. Nos últimos anos, o RNAi aliado à disponibilidade da sequência do genoma [34], alavancou os estudos envolvendo *T. brucei* [36-39]. O RNAi não é funcional em *T. cruzi* nem em *L. major*, o que pode ser explicado pela perda desta via nestes parasitas durante a evolução [40]. Pela dificuldade na obtenção de nocaute gênico nestes parasitas, têm se usado como estratégia alternativa o estudo de genes ortólogos em *T. brucei* empregando RNAi.

Algumas facilidades, tais como a possibilidade de utilização de vetores de expressão induzida por tetraciclina [41], a disponibilidade de cepas com expressão constitutiva do repressor de tetraciclina e da T7 RNA polimerase [42] e a utilização de algoritmos para a análise de sequências gênicas possibilitam a utilização em larga escala deste sistema. Um exemplo de algoritmo é o RNAit [43], que auxilia na determinação da melhor região do gene a ser utilizada para o RNAi. Sendo assim pode-se empregar RNAi no estudo desde genes isolados ou pertencentes a uma mesma organela ou estrutura, por exemplo genes que codificam proteínas flagelares [44] até genes de um cromossomo inteiro [45].

1.4.2 *Proteômica Baseada em Espectrometria de Massas*

As proteínas desempenham uma série de funções nas células, tanto estruturais, como na regulação de diversos processos. Na era pós-genômica, ficou ainda mais evidente a necessidade de investimento em esforços para a caracterização funcional das proteínas. Dentre as diversas metodologias com esta finalidade, tem sido cada vez mais frequente a associação entre tais metodologias e proteômica baseada em espectrometria de massas. Avanços nas técnicas de cromatografia e de espectrometria de massas possibilitaram análises globais do proteoma celular, com a identificação e quantificação de milhares de proteínas em diferentes modelos biológicos [46] [47] [48].

A técnica conhecida por *shotgun* é comumente a abordagem de escolha nos experimentos de proteômica associada à espectrometria de massas. Nesta abordagem, as proteínas de uma amostra são digeridas por uma protease específica, como por exemplo tripsina ou Lys-C, e os peptídeos provenientes desta digestão são analisados por espectrometria de massas. Através de algoritmos, os dados gerados são então confrontados com bancos de dados teóricos, sendo possível a predição dos peptídeos, desde que se conheça a especificidade da protease. Por exemplo, se um determinado extrato proteico for digerido com tripsina, cuja especificidade é clivar ligações peptídicas após resíduos de lisina e arginina, os peptídeos gerados são denominados peptídeos trípticos, os quais serão confrontados com um banco de dados teórico, também de peptídeos trípticos.

Outras abordagens de proteômica, nas quais, ou as proteínas não são digeridas (proteômica *top-down*) ou são digeridas por proteases que reconhecem sítios raros, gerando assim peptídeos maiores que os peptídeos trípticos (proteômica *middle-down*) possuem aplicações mais específicas. Estas abordagens normalmente exigem espectrômetros com alta resolução e são utilizadas em amostras menos complexas, para diferenciar isoformas proteicas ou identificar modificações pós-traducionais [49]. A seguir serão apresentadas as etapas de análise proteômica por espectrometria de massas.

CROMATOGRAFIA LÍQUIDA

Normalmente em amostras proteicas complexas, tais como extrato total, é feita uma separação que precede à espectrometria de massas em si de maneira acoplada (*on line*). Em amostras mais simples, como *spots* de um gel bidimensional por exemplo, os peptídeos não são previamente separados. A técnica de separação de escolha é a cromatografia líquida (LC) de alta pressão. Na separação de peptídeos por LC, o método de escolha é a cromatografia de fase reversa (RP). Para peptídeos, normalmente a fase estacionária da coluna analítica é C18, composta por cadeias alquila de 18 carbonos ligadas a uma resina de sílica. A fase móvel é uma mistura de água com um solvente orgânico, que pode ser metanol ou acetonitrilo.

Durante a cromatografia, a fase móvel passa pela fase estacionária com uma pequena porcentagem de solvente na mistura água/solvente, que vai aumentando ao longo da corrida. Os peptídeos interagem com a C18 por hidrofobicidade, sendo assim, ao longo da corrida cromatográfica, os peptídeos menos hidrofóbicos são eluídos antes, ou seja, com menor porcentagem de solvente. Já os peptídeos mais hidrofóbicos são eluídos posteriormente com o aumento da porcentagem de solvente na fase móvel. Esta separação contribui para a identificação de um número maior de peptídeos, visto que se todos os peptídeos de uma amostra fossem injetados ao mesmo tempo no espectrômetro, apenas os mais abundantes seriam detectados, o que seria uma pequena fração do universo de peptídeos em uma amostra complexa.

IONIZAÇÃO DOS PEPTÍDEOS

Para serem analisados no espectrômetro de massa, os peptídeos precisam ser ionizados. Os métodos de ionização comumente utilizados para peptídeos são nano *electro spray ionisation* (nano ESI ou *nanospray*) e *matrix assisted laser desorption ionisation* (MALDI). Na ionização por *nanospray*, a amostra recebe uma corrente elétrica sendo então formado um *spray* composto por gotículas contendo íons, positivos no caso de peptídeos. O solvente presente nas gotículas evapora e, então, os íons são direcionados para o espectrômetro de massas. Na ionização por

MALDI, os peptídeos são misturados com uma matriz (ácido 2,5-dihidroxibenzoico - DHB, por exemplo) e depositados em *spots* sobre uma placa. A seguir, os peptídeos são ionizados por um laser e lançados para o espectrômetro. A cromatografia líquida é geralmente acoplada com a ionização por ESI, enquanto MALDI é comumente empregada em amostras provenientes de *spots* de gel 2D.

Os peptídeos apresentam características diferentes quanto à ionização. Alguns são mais facilmente ionizáveis em relação a outros. Este seria um dos motivos para o insucesso na identificação de algumas proteínas, por exemplo as que geram peptídeos de difícil ionização. Outra característica é o número de cargas em cada peptídeo, o qual pode variar no caso da ionização por ESI. Em geral, são analisados no espectrômetro peptídeos com no mínimo 2 cargas devido à presença de no mínimo uma carga em uma lisina ou arginina em um dos lados do peptídeo (decorrente da digestão com tripsina) e outra carga na porção amino.

MS1 (*FULL SCAN*)

Após a ionização, os peptídeos são direcionados para o espectrômetro de massas. Um espectrômetro é composto de uma fonte de ionização, um analisador de massas e um detector. Na medida em que são eluídos da coluna cromatográfica, os íons são separados no analisador de massas de acordo com as suas relações massa sobre carga (m/z). A separação e detecção dos íons dá origem a espectros. Uma das formas de visualização destes espectros é através de um cromatograma, onde cada espectro representa a soma das intensidades dos íons (eixo y) eluídos da coluna cromatográfica por um determinado tempo (eixo x). Ao selecionar-se um período de tempo de retenção específico, os íons podem ser visualizados de acordo com a intensidade (eixo y) e m/z (eixo x). Estes cromatogramas gerados representam o MS1 ou *full scan*.

Um cromatograma pode conter íons (peptídeos) com diferentes cargas, o que poderia dificultar a determinação correta da massa destes íons, pois é necessário conhecer o número de cargas (z) de um determinado íon para se determinar a sua massa. A maneira de contornar este problema é a utilização dos próprios dados gerados no MS1. Entre os peptídeos existe a ocorrência natural de isótopos, ou seja,

peptídeos contendo átomos com massa maior, como por exemplo carbono 13 ou nitrogênio 15. Sendo assim, os peptídeos são representados por um conjunto de íons com diferença de massa correspondente a aproximadamente um Dalton, chamado de envelope isotópico. Dentro de um envelope isotópico, normalmente a intensidade dos íons com maior massa (contendo isótopos de massa maior) diminui, devido às probabilidades de ocorrência dos diferentes isótopos na natureza. Como a diferença de massa entre os isótopos é conhecida (aproximadamente um Dalton), é possível saber o número de cargas dos íons de um envelope isotópico. Por exemplo, se a diferença de relação m/z entre os isótopos do envelope é 0,5, pode-se concluir que o número de cargas destes íons é dois ($m/z = \frac{1}{2} = 0,5$). Com isso, é possível a determinação das massas dos íons detectados no MS1.

FRAGMENTAÇÃO DE ÍONS E MS2 (MS/MS)

A fragmentação de íons geralmente ocorre por colisão com gases inertes ou reagentes específicos e os diferentes tipos podem formar fragmentos distintos, favorecendo diferentes tipos de análises. Um método de fragmentação bastante empregado é a dissociação induzida por colisão com gases inertes - *collisionally induced dissociation* (CID). No caso de peptídeos, a fragmentação por CID pode gerar íons característicos nomeados de acordo com a ligação rompida no peptídeo.

Na identificação de peptídeos, os íons mais informativos são os gerados pela fragmentação das ligações peptídicas. De acordo com a nomenclatura proposta por Roepstorff–Fohlmann–Biemann [50, 51] estes fragmentos são nomeados de íons-b, quando a carga está presente na porção amino ou íons-y, quando a carga está presente na porção carbóxi. A identificação de um peptídeo pode ocorrer pela identificação completa ou parcial da sequência de resíduos de aminoácidos que o compõe. Um determinado peptídeo é representado por diversas moléculas idênticas, considerando que a ionização ocorreu igualmente nestas moléculas. Sendo assim, em uma condição ideal, cada molécula pode gerar pelo menos um íon-b ou -y. As diferenças entre as massas dos diferentes íons-b ou -y gerados (contendo diferentes números de resíduos de aminoácidos) correspondem às massas dos resíduos de aminoácidos que compõe o peptídeo em questão. Este processo, chamado de MS2

ou MS/MS é uma metodologia poderosa para identificação de proteínas por espectrometria de massas.

ESPECTRÔMETROS DE MASSAS

Os espectrômetros de massas são compostos por um ou mais analisadores de massa (iguais ou diferentes) e um ou mais detectores. Os analisadores basicamente separam os íons de acordo com a sua m/z . Cada analisador possui seu princípio para executar esta tarefa. O analisador do tipo quadrupolo é formado por quatro bastões dispostos paralelamente para formar um “canal” para os íons. Através de alternância de radiofrequência (RF) e de corrente elétrica, determinados íons são selecionados (diferentes relações entre corrente e RF selecionam diferentes m/z) e passam entre os bastões, podendo alcançar o detector. O analisador quadrupolo possui algumas variações, como por exemplo, o *linear trap quadrupole* (LTQ), também chamado de *ion trap*. No analisador *time of flight* (TOF), o tempo de voo dos íons dentro de um tubo contendo vácuo é medido. Este tempo varia de acordo com a relação m/z do íon. No *Fourier transform ion cyclotron resonance* (FTICR), os íons oscilam em um campo magnético e esta oscilação varia de acordo com a relação m/z . Esta oscilação é convertida em sinal por equações (transformada de Fourier). O analisador orbitrap também se baseia na oscilação dos íons, os quais orbitam ao redor de um eletrodo central, no entanto, diferentemente do FTICR, sem campo magnético. Os analisadores podem ser combinados, por exemplo, triplo quadrupolo, TOF-TOF, Q-TOF, LTQ-FTICR e LTQ-Orbitrap.

A combinação de dois ou mais analisadores é algo comum nas configurações de espectrômetros de massa. Um analisador também pode ser utilizado como filtro de íons ou como câmara de fragmentação, como é o caso do triplo quadrupolo, onde os íons são fragmentados no segundo quadrupolo. Cada analisador possui suas características, tais como resolução, rapidez, sensibilidade e range dinâmico (diferença entre as concentrações das moléculas mais abundantes e das menos abundantes). A escolha do analisador dependerá da natureza das moléculas a serem analisadas e do tipo de experimento a ser realizado. Por exemplo, para a quantificação de compostos químicos conhecidos, uma boa opção é o triplo

quadrupolo, que permite a monitoração de íons específicos na amostra no modo SRM (*single reaction monitoring*). Os analisadores TOF são bastante rápidos e não possuem limite na relação m/z detectada. O LTQ possui alta sensibilidade, porém com baixa resolução. O FTICR e o Orbitrap apresentam alta resolução, porém são mais lentos e menos sensíveis que o LTQ. Uma configuração bastante eficiente para a proteômica ocorre ao se combinar os analisadores LTQ e Orbitrap. Esta configuração é encontrada nos equipamentos da Thermo da série LTQ Orbitrap e foi utilizada neste projeto.

IDENTIFICAÇÃO DE PROTEÍNAS EM BANCOS DE DADOS

Com o início do sequenciamento dos genomas de vários organismos, tornou-se possível a criação de bancos de dados de proteínas e consequentemente de peptídeos. Os dados de MS1 podem ser usados para identificação de proteínas por uma metodologia chamada *peptide mass finger-printing* (PMF) [52] [53] [54] [55] [56]. Nesta metodologia, as massas dos peptídeos obtidas experimentalmente são comparadas com as massas teóricas de peptídeos trípticos para cada proteína em um banco de dados. A identificação por PMF normalmente é usada para MALDI. Uma metodologia de identificação mais sensível e específica pode ser obtida com os dados de MS2. Frequentemente estes dados não cobrem toda a sequência dos peptídeos analisados, no entanto podem ser suficientes para o pareamento inequívoco destes peptídeos com um banco de dados. Diversos algoritmos têm sido desenvolvidos a fim de fazer o pareamento entre os dados obtidos e os dados teóricos preditos. Alguns, bastante conhecidos, são o Sequest [57] e o Mascot (Matrix Science). Estes algoritmos possuem suas peculiaridades, mas basicamente fazem a sobreposição dos espectros teóricos de MS2 com os espectros obtidos experimentalmente. Esta sobreposição pode encontrar mais de um *match* possível entre um par de peptídeos comparados (teórico vs experimental). Após a sobreposição dos espectros, os *matches* são então classificados com *scores*. Estes algoritmos oferecem opções que podem aumentar a correlação entre os espectros teóricos e obtidos. Por exemplo, pode-se informar a presença de algumas

modificações nos peptídeos obtidos, tais como oxidação em metionina, alquilação em cisteína, entre outras. Isto facilita a busca nos bancos de dados teóricos.

Com o aumento dos dados gerados por MS2, a determinação da qualidade destes dados está tornando-se cada vez mais importante. Por exemplo, pode-se estimar a taxa de falsos positivos em uma busca com dados de MS2. Isto pode ser feito por uma metodologia chamada *decoy database*. Nesta metodologia, as sequências das proteínas de um banco de dados são invertidas e os peptídeos gerados são comparados com os dados de MS2 obtidos experimentalmente [58]. Os *matches* encontrados nesta situação representam falsos positivos (*false discovery rate* - FDR) e a sua taxa máxima pode ser estipulada, por exemplo, em no máximo 1%, ou seja, um dos parâmetros que irá determinar o nível de estringência na identificação dos peptídeos será o valor máximo de FDR.

PROTEÔMICA QUANTITATIVA

Alguns algoritmos fornecem a opção de análise de dados quantitativos gerados por metodologias como *isotope-coded affinity tag* (ICAT), *isobaric tags for relative and absolute quantitation* (ITRAQ), *stable isotope labeling by amino acids in cell culture* (SILAC), entre outras [59] [60] [61]. Nas duas primeiras, a marcação é feita quimicamente em proteínas nos resíduos de cisteína (ICAT) ou em peptídeos nos grupamentos amino ou cadeias laterais de lisina (ITRAQ). No ICAT, uma limitação seria a marcação somente de proteínas contendo resíduos de cisteína, com a posterior purificação dos peptídeos marcados, podendo-se assim ocorrer perda de informações, tais como peptídeos contendo PTMs. O ITRAQ, ao contrário do ICAT e do SILAC, possui a vantagem de não duplicar a complexidade da mistura de peptídeos proveniente de diferentes amostras. Isto porque os pares de peptídeos serão isobáricos, ou seja, possuirão a mesma massa. Nesta metodologia, a diferenciação entre os pares de peptídeos ocorre no MS2, quando são liberados por fragmentação íons repórteres, os quais possuem massas diferentes, de acordo com a marcação do peptídeo precursor. No ITRAQ, a mistura das amostras é feita em nível de peptídeos, sendo assim, as interferências causadas por manipulação podem prejudicar a qualidade da quantificação.

Na metodologia SILAC, ocorre a marcação metabólica das proteínas celulares. As células são cultivadas em um meio específico, o deve possuir algum aminoácido marcado, de acordo com a enzima utilizada na digestão das proteínas, por exemplo, lisina e arginina quando se usa tripsina. Os aminoácidos marcados possuem isótopos pesados de carbono e/ou nitrogênio, por exemplo, Arg6 (6 carbonos 13) e Lys8 (6 carbonos 13 + 2 nitrogênios 15). Desta forma, ao utilizar estes aminoácidos como fonte para síntese de proteínas, a célula faz a incorporação dos mesmos. Após aproximadamente 5 divisões celulares, a taxa de incorporação dos aminoácidos marcados deve ser de mais 95%. Nesta metodologia é feito o cultivo em paralelo de duas ou mais culturas em condições biológicas distintas, sendo uma marcada (pesada) e outra não marcada (leve). Antes da extração das proteínas, as células são misturadas 1:1 e submetidas ao mesmo tratamento subsequentes. Esta é uma das vantagens desta metodologia para quantificação, pois diminui as possíveis interferências decorrentes da manipulação individual das amostras. Uma desvantagem desta metodologia é o aumento na complexidade da amostra pelo fato de duplicar ou triplicar (no caso de três condições biológicas) o número de proteínas na amostra.

A análise dos dados provenientes de SILAC é realizada pelo algoritmo MaxQuant [62]. Este programa pode trabalhar com dados de alta resolução de quantificação gerados pelo Orbitrap. Simplificadamente, seu princípio se baseia na determinação das áreas tridimensionais dos picos de cada peptídeo (dimensões: intensidade, m/z e tempo de retenção), obtidos no MS1, esta área possui correlação com a quantidade dos peptídeos analisados. Este programa identifica os pares de peptídeos (leve e pesado – *SILAC pair*) em uma amostra contendo uma mistura de proteínas marcadas e não marcadas. Assim é possível determinar a intensidade relativa dos peptídeos representantes de uma determinada proteína e a soma destes representando a quantidade desta proteína. Esta análise é feita nos dados de MS1, sendo que os dados de MS2 são utilizados para a identificação dos peptídeos.

Além das diferentes opções citadas, a quantificação de proteínas por MS pode ser realizada sem marcação metabólica ou química das proteínas. A simples análise da contagem espectral entre diferentes amostras, ou seja, a comparação do número de espectros de MS2 obtidos para cada proteína entre as amostras

analisadas possui correlação com a abundância destas proteínas nas respectivas amostras. Esta metodologia livre de marcação (*label-free quantification*) possui a vantagem de ser em teoria aplicável a qualquer tipo celular, o que não ocorre com a marcação metabólica, onde há a necessidade do uso de protocolos específicos de cultivo celular, os quais nem sempre estão estabelecidos. Apesar do seu potencial de aplicabilidade, a metodologia *label-free* ainda é um campo a ser aprimorado, sobretudo na análise de PTMs por MS, onde a manipulação das amostras pode introduzir variações na quantificação relativa de amostras.

FOSFOPROTEÔMICA

Após serem sintetizadas, as proteínas podem ser modificadas dependendo das suas funções na célula. Vários são os tipos de modificação que existem e apenas alguns serão comentados aqui. Por exemplo, proteínas de membrana podem receber a adição de carboidratos, os quais conferem proteção para a célula e intermedeiam interações com outras moléculas. No interior da célula, modificações nas proteínas, como a ubiquitilação, marcam tais proteínas destinando-as a determinados processos. Uma via bastante conhecida é a degradação, pelo proteassomo, de proteínas ubiquitiladas. Uma modificação pós-traducional (PTM) bastante estudada é a fosforilação de proteínas, cujas funções já foram comentadas no tópico 1.2.

As PTMs acima citadas são feitas pela adição de grupamentos químicos ou pequenos polipeptídeos a proteínas específicas. Isto gera um aumento na massa destas proteínas, o que torna possível a detecção de tais PTMs por MS. As PTMs são dinâmicas, principalmente no caso das fosforilações. Além disso, apenas parte das proteínas consideradas potenciais alvos de fosforilação estão fosforiladas em um determinado momento. A detecção de tais fosforilações (sítios de fosforilação) e a quantificação destes sítios entre diferentes amostras podem fornecer indícios da participação dos sítios modulados em funções específicas na célula. Várias técnicas para análise do estado de fosforilação de uma única proteína isolada têm sido utilizadas [63]. No entanto, a análise em larga escala do perfil de fosforilação de uma célula representa um desafio. Algumas das dificuldades encontradas são a baixa

abundância de fosfoproteínas [63] e as características particulares de cada aminoácido fosforilado [64].

Devido à baixa abundância de fosfoproteínas, as amostras para fosfoproteômica necessitam de um enriquecimento prévio. Existem várias técnicas para tal procedimento, dentre elas pode-se citar o enriquecimento por fosforamidoato (*phosphoramidate chemistry* – PAC) [65], cromatografia de afinidade por metal imobilizado (*immobilized metal affinity chromatography* - IMAC) [66, 67] e dióxido de titânio (TiO₂) [68, 69].

O enriquecimento por PAC, assim como os outros citados, permite a seleção de peptídeos com fosforilação nos resíduos de serina, treonina e tirosina. Os fosfopeptídeos são acoplados a um suporte de fase sólida e, após lavagem, são eluídos em condições ácidas. Os métodos baseados em afinidade para isolamento de fosfopeptídeos incluem IMAC utilizando íons metálicos (gálio, ferro, entre outros) ou resinas contendo TiO₂ ou dióxido de zircônio.

Outro desafio encontrado quando se deseja obter o fosfoproteoma de uma célula é a representatividade dos fosfopeptídeos isolados em cada método de enriquecimento, ou seja, existe uma dificuldade em se conseguir o fosfoproteoma total de uma célula. Um trabalho publicado em 2007 [70] fez a comparação de três métodos para enriquecimento de fosfopeptídeos (PAC, IMAC e TiO₂) em células Kc167 de *D. melanogaster*. Nos resultados obtidos, houve sobreposição parcial dos fosfopeptídeos encontrados pelos três métodos de enriquecimento utilizados (Figura 6), sendo que cada método apresentou especificidade distinta para certos tipos de fosfopeptídeos. Este resultado demonstra a importância da utilização do maior número possível de métodos de enriquecimento quando se deseja obter uma representatividade maior do fosfoproteoma celular.

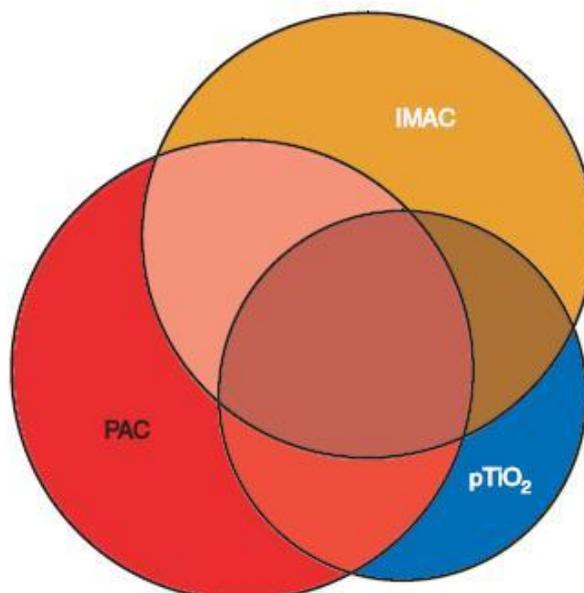


Figura 6: Sobreposição entre os sítios de fosforilação identificados por diferentes métodos de enriquecimento. Fonte: Bodenmiller et al., 2007.

Tendo a amostra sido processada para o enriquecimento dos fosfopeptídeos, a etapa seguinte é a análise por espectrometria de massas. Os fosfopeptídeos contendo um grupamento fosfato possuem um incremento de 98 Da na massa, 80 (HPO_3) + 18 (H_2O). Na fragmentação por CID, o grupamento fosfato é perdido ocasionando uma perda neutra (*neutral loss*), ou seja, a m/z do peptídeo é alterada após a perda de massa, no entanto a carga se mantém. Através da metodologia chamada *Data Dependent Neutral Loss MS3* (DDNLMS3) [71], o íon que sofreu a perda do grupamento fosfato é selecionado e fragmentado em um evento de MS3. Alternativamente, em uma metodologia denominada *multistage activation* (MSA) [72] a fragmentação do íon precursor que sofreu a perda neutra ocorre sem a necessidade de um isolamento adicional (MS3), além disso, conservando os fragmentos gerados no MS2, o que contribui para a identificação da sequência do fosfopeptídeo.

A confiabilidade na localização dos fosfosítios é calculada por um algoritmo que analisa os quatro íons mais intensos por janela de 100 m/z . São feitas então as combinações da massa destes íons com os três resíduos de aminoácidos possíveis (Ser, Thr e Tyr), e são então atribuídos *scores*. A partir destes, são geradas as probabilidades da localização dos fosfosítios estarem corretas [73].

PRÉ-FRACIONAMENTO DA AMOSTRA

Para a realização de proteômica e fosfoproteômica é possível a utilização de uma amostra de extrato proteico total, no entanto a complexidade deste tipo de amostra associada ao limite de amplitude de detecção dos espectrômetros de massa pode limitar o número de proteínas identificadas. O uso de técnicas de pré-fracionamento da amostra pode aumentar consideravelmente a identificação de proteínas diante da análise de frações do extrato, com menor complexidade que o extrato total. Vale ainda ressaltar a importância na escolha do método de fracionamento da amostra (proteínas ou peptídeos) antes de submetê-la à espectrometria de massas. Dentre os métodos de fracionamento existentes, pode-se citar o fracionamento de proteínas por SDS-PAGE e de peptídeos ou proteínas por cromatografia de troca iônica (*strong cation exchange chromatography* - SCX) ou por focalização isoelétrica.

1.4.3 RNA-seq

As técnicas para análise transcriptômica tem se mostrado ferramentas poderosas nos estudos de genômica funcional. Desde os microarranjos de DNA, nos quais milhares de sondas de DNA são depositadas em uma lâmina para posterior hibridização comparativa de amostras de RNA. As novas tecnologias de sequenciamento de DNA, como é caso do SOLID (Applied Biosystems), empregam ligase no seu princípio de funcionamento sendo possível a quantificação de RNA de uma amostra.

Os *reads* gerados pelo SOLID possuem tamanho médio de 50 nucleotídeos, sendo possível a geração de milhões de *reads* por corrida. O SOLID fornece maior acurácia e reprodutibilidade comparado ao microarranjo de DNA.

1.4.4 Duplo-híbrido

O duplo híbrido é uma técnica *in vivo* para detecção de interações binárias, através da transcrição de genes repórteres, sendo realizada em levedura ou bactéria. Dentre as metodologias disponíveis para análise de interações proteicas, o

duplo híbrido é a que permite mais facilmente a execução em larga escala nos sistemas de análise *in vivo*. Com a disponibilidade de vetores plasmidiais para fusão do gene aos domínios de ativação (AD) e de ligação ao DNA (BD), contendo o sistema Gateway de clonagem (Invitrogen), esta plataforma se torna ainda mais atraente. Como característica desta técnica está o alto número de falsos positivos e negativos. O uso de diferentes cepas de levedura e diferentes composições de meio para o teste fenotípico auxilia na diminuição destes eventos. Esta metodologia é muito explorada, sendo eficiente e conveniente para se utilizar na detecção e exploração de interações individuais ou de interatomos inteiros, sendo a mais utilizada até o momento.

1.4.5 Combinação de Metodologias de Trabalho

Uma abordagem de estudo que vem sendo bastante utilizada é a análise quantitativa por proteômica, fosfoproteômica e transcriptômica associada com a produção de interferências ou de estímulos na célula, como RNAi, nocaute gênico, drogas, estresse celular, fatores de crescimento, entre outros. Tais interferências podem causar mudanças na expressão proteica da célula. Comparando-se a amostra submetida à interferência com uma amostra controle (sem interferência), as mudanças na expressão gênica ou em PTMs podem ser detectadas e quantificadas. Alguns trabalhos abrangendo estas combinações de metodologias tem sido publicados [47, 74]. Em artigo publicado em 2008, Bonaldi e colaboradores [75] utilizaram RNAi para *knockdown* de uma subunidade de um complexo de remodelamento de cromatina (ISWI) em *Drosophila melanogaster*. O perfil proteico das células submetidas ao RNAi foi obtido por proteômica quantitativa (marcação por SILAC) e os resultados comparados com os dados de microarranjo, gerados por outro trabalho [76]. Esta comparação mostrou discordância entre microarranjo e proteômica na relação entre vários genes e proteínas. Um exemplo disto foi observado para uma proteína parceira de ISWI (Acf-1), a qual teve seu nível reduzido, enquanto o nível de mRNA manteve-se estável após o *knockdown*. Este resultado reforça a utilização de proteômica quantitativa, principalmente nos casos onde a regulação da expressão gênica ocorre com o controle dos níveis proteicos e não dos níveis de mRNA.

Em outro estudo, também em células de *Drosophila*, avaliou-se o efeito da depleção por RNAi de uma tirosina fosfatase (Ptp61F) sobre o proteoma e fosfoproteoma celular. Os resultados demonstraram que os efeitos da depleção de Ptp61F sobre o fosfoproteoma foram maiores que os efeitos sobre o proteoma celular. Isto mostra uma atuação preferencial desta fosfatase sobre a rede de fosforilação comparando-se com a atuação sobre a expressão gênica [77].

1.5 Proteômica e Fosfoproteômica em Tripanossomatídeos

Existem poucos estudos que abordam proteômica baseada em espectrometria de massas para Tripanossomatídeos. Alguns exemplos são o estudo do proteoma e secretoma de *Leishmania*, analisando proteínas estágio-dependente diferencialmente expressas ou diferenças entre os proteomas de cepas resistentes ou susceptíveis à drogas [78-81]. Adicionalmente, foi publicada uma análise comparativa do proteoma dos estágios de desenvolvimento de *T. cruzi* [82].

O primeiro trabalho a realizar fosfoproteômica baseada em espectrometria de massas em Tripanossomatídeos foi publicado em 2009 [9]. Neste trabalho foram identificadas mais de 400 fosfoproteínas e cerca de 1200 fosfossítios, relacionados com diferentes processos biológicos, tais como processamento de DNA e RNA e transdução de sinal. Ainda em 2009, Nakayasu e colaboradores analisaram o fosfoproteoma de *T. cruzi*, mapeando 220 fosfossítios neste parasita.

Recentemente, nosso grupo fez a identificação de mais de 2.500 fosfossítios ao longo da metaciclogênese de *T. cruzi* [16]. Foram analisadas formas epimastigotas em fase exponencial de crescimento, epimastigotas submetidas a estresse nutricional, formas aderidas e tripomastigotas metacíclicas. Dentre os dados encontrados, 30% das fosfotirosinas identificadas foram atribuídas a proteínó-quinases. A análise de ontologia gênica dos fosfossítios ao longo da diferenciação de *T. cruzi* mostrou o enriquecimento de proteínas relacionados aos processos de ligação a íons cálcio, de atividade quinase e de modificação de proteínas. Também foi constatada a presença de sítios de ativação clássicos em quinases do tipo MAPK, formados por resíduos de tirosina e treonina. Este resultado corrobora com os dados do trabalho de Parsons e colaboradores [8] descrevendo a presença de domínios de

ativação clássicos em MAPKs de *T. cruzi*. Na sequência, os dados de proteômica quantitativa durante a metaciclogênese foram publicados pelo mesmo grupo [83]. Quase 3.000 proteínas foram quantificadas neste trabalho, sendo que as proteínas diferencialmente expressas mostraram enriquecimento para alguns processos celulares, tais como citoesqueleto, atividade motora, ribossomo, tradução, entre outros.

Em um experimento quantitativo, Urbaniak e colaboradores realizaram a comparação entre os proteomas dos dois estágios de vida de *T. brucei*, procíclico e sanguíneo. Usando SILAC e pré-fracionamento da amostra por SCX e SDS-PAGE foram quantificadas 3.553 proteínas. As principais diferenças foram evidenciadas em proteínas que participam do metabolismo energético e em proteínas de superfície, o que está em concordância com a biologia deste parasita e com as respostas aos diferentes hospedeiros deste parasita [84]. No ano seguinte, os fosfoproteomas destas mesmas formas foram comparados, obtendo-se a quantificação de mais de 8.000 fosfossítios. Dentre os dados obtidos, foi observada a funcionalidade da via das MAPKs em ambas as formas de *T. brucei*, além da extensiva fosforilação em proteínas de ligação a RNA, fato que pode estar relacionado à participação destas no controle pós-transcricional da expressão gênica. Foram encontrados 631 fosfossítios aumentados no mínimo 10 vezes ou exclusivos na forma sanguínea em relação à forma procíclica. Dentre eles pode-se citar sítios presentes em proteínas hipotéticas e em proteínas sugeridas como participantes nas diferenças fisiológicas entre as duas formas. Os sítios exclusivos ou aumentados na forma procíclica (pelo menos 10 vezes) somaram 614. Dentre estes, estão vários sítios pertencentes à proteínas hipotéticas e a proteínas cuja expressão também está aumentada na forma procíclica [85].

Diante do cenário apresentado, é possível identificar vários esforços no sentido de compreender os processos de sinalização intracelular por fosforilação nos Tripanossomatídeos. Sendo assim, na tentativa de diminuir esta lacuna, aqui neste trabalho focou-se no entendimento da participação de MAPKs de *T. brucei* nos processos de sinalização intracelular através da identificação de fosfossítios, proteínas e transcritos de RNA modulados após a depleção destas MAPKs.

2. OBJETIVO

Tendo sido feita uma revisão bibliográfica sobre MAPKs e fosfoproteômica em Tripanossomatídeos, é possível constatar que vários esforços vêm sendo empregados com o intuito de entender as vias de sinalização por fosforilação em Tripanossomatídeos. O fato de estes protozoários possuírem regulação da expressão gênica pós-transcricional, aumenta o interesse nos estudos envolvendo PTMs. No entanto, pouco se sabe sobre os alvos moleculares das proteíno-quinases nestas células. Aqui neste trabalho, busca-se analisar quantitativamente o proteoma, o fosfoproteoma e o transcriptoma de *T. brucei* após a depleção de proteínas da família das MAPKs. Esta análise pode identificar proteínas-alvo diretos e indiretos das MAPKs. Além disso, as modificações encontradas podem fornecer indícios para a correlação de processos celulares com cada MAPK avaliada. A Figura 7 exemplifica parte do objetivo deste projeto através de um esquema resumido da mudança no perfil de fosforilação após a depleção de uma MAPK.

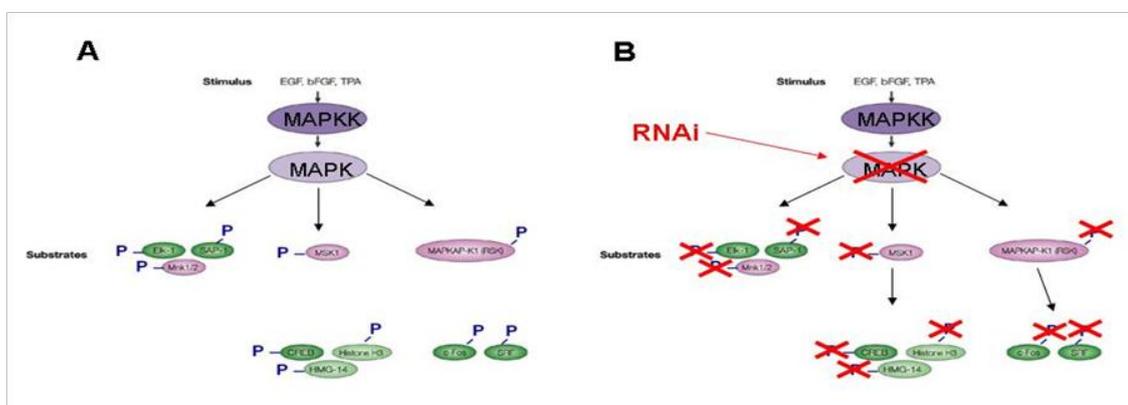


Figura 7: Diagrama esquemático do perfil de fosforilação após depleção de uma MAPK. Em A, situação pré-indução do RNAi, em B, condição pós-indução do RNAi. Adaptado de Hazzalin & Mahadevan, 2002.

3. JUSTIFICATIVA

T. brucei é um importante patógeno para humanos e outros mamíferos, tais como bovinos, por exemplo. É também considerado um organismo modelo de estudo dentre os Tripanossomatídeos e, apesar de ser um protozoário bastante estudado, ainda existem inúmeras questões em aberto a respeito da sua biologia. Com o genoma sequenciado e com a disponibilidade de ferramentas, tais como, vetores para expressão induzida, RNAi, novas tecnologias de sequenciamento e espectrometria de massas, entre outras, o estudo dos processos celulares em *T. brucei* pode ser ampliado. Dentre os processos celulares vitais, vale à pena ressaltar a sinalização por fosforilação de proteínas, principalmente para os Tripanossomatídeos, que necessitam se adaptar às drásticas mudanças ambientais durante o ciclo de vida. Entre as quinases identificadas em *T. brucei*, as MAPKs estão entre as famílias mais bem representadas neste parasita, podendo refletir a importante presença das cascatas envolvendo MAPKs nas respostas às condições ambientais. O entendimento dos processos celulares depende do conhecimento das interações moleculares, principalmente entre proteínas, sendo assim, na tentativa de compreender os processos de fosforilação, torna-se necessária a identificação dos alvos das proteíno-quinases. O conhecimento destas interações permite a construção de mapas de sinalização, o que possibilita um entendimento de forma mais global a respeito da biologia celular e molecular dos Tripanossomatídeos.

4. MATERIAL E MÉTODOS

4.1 Desenho dos Oligonucleotídeos para RNAi

Inicialmente foi necessária a busca das sequências nucleotídicas das 15 MAPKs de *T. brucei* TREU 927 em um banco de dados (GeneDB). Com estes dados em mãos foi feita a escolha das regiões alvo para o RNAi e o desenho dos oligonucleotídeos necessários para a amplificação destas regiões do genoma por PCR (ANEXOS - Tabela 23). Este desenho dos oligos foi executado com o auxílio da ferramenta RNAit [43]. Para isto, foi usada a seguinte configuração:

- *database: Trypanosoma brucei 927*
- *PCR product size range: 400 – 600 bp*
- *perfect match exclusion length: 20*
- *primer Tm: 60 °C*
- *blastn stringency: 80 – 90%*

Neste trabalho, as MAPKs de *T. brucei* receberam uma identificação de A a O, conforme a Tabela 4, a fim de facilitar os experimentos e apresentação dos dados.

4.2 Amplificação dos Alvos para RNAi

Os alvos para RNAi foram amplificados por PCR utilizando-se o seguinte protocolo: 2 mM de MgSO₄, 0,4 mM de cada oligo, tampão *Platinum Taq*, 1 U de *Platinum Taq DNA Polymerase High Fidelity* (Invitrogen), 10 mM de dNTPs, 100 ng de DNA genômico de *T. brucei* Lister 427 em um volume final de 50 µl. A programação do termociclador foi a seguinte:

- 94 °C – 2 min
- 94 °C – 30 s
- 55 °C – 30 s
- 68 °C – 1 min
- 30 ciclos
- 68 °C – 5 min

Tabela 4: Identificação das MAPKs e tamanho das sequências codificantes e proteínas.

Id	Id interna	Tamanho CDS (pb)	Tamanho proteína (kDa)	Nome
Tb927.10.10870	A	1.713	64,473	
Tb927.10.3230	B	1.122	42,207	
Tb927.8.3770	C	1.230	46,938	
Tb927.8.3550	D	1.116	42,673	MAPK3
Tb927.10.16030	E	1.098	41,808	MAPK2
Tb927.6.1780	F	1.236	46,120	
Tb927.6.4220	G	1.140	43,691	MAPK5
Tb927.10.12040	H	1.242	47,390	
Tb927.10.5140	I	1.329	49,742	
Tb927.10.7780	J	1.053	39,359	KFR1, MAPK1
Tb927.11.8170	K	1.464	54,125	
Tb927.11.16790	L	1.974	72,144	ECK1
Tb927.10.14800	M	1.164	42,750	MAPK9
Tb927.3.690	N	1.752	63,828	
Tb927.9.9320	O	1.095	41,757	MOK

4.3 Clonagem dos Alvos

Após a amplificação, os produtos de PCR foram purificados através do kit *Wizard PCR Preps DNA Purification System* (Promega), de acordo com as instruções do fabricante. Os produtos purificados foram então clonados no vetor p2T7-177 [86], modificado a fim de permitir clonagem TA e *screening* dos clones por

lacZ (Figura 8). O vetor foi inicialmente digerido na seguinte reação: 1 ug de vetor, 20 U de *XcmI*, 1x de tampão NEB 2, em um volume final de 30 μ l. Esta reação foi incubada por 3 h a 37 °C.



Figura 8: Vetor P2T7-177 modificado. 177 – região sub-telomérica do genoma de *T. brucei*. T7 term – sequência terminadora da T7 RNA polimerase. T7P – região promotora da T7 RNA polimerase. Act 3' – região intergênica 3' do gene da actina.

Após digestão, a enzima foi inativada incubando-se a mistura a 65 °C por 20 min. Cinquenta ng do vetor digerido foram usados para cada reação de ligação: 1 U de T4 DNA ligase (Invitrogen), tampão T4 DNA ligase, 50 ng do vetor digerido, inserto com excesso de 10 vezes, em um volume final de 10 μ l, incubada por 16 h a 16 °C. Adicionalmente foi construído um controle, extraído-se GFP do vetor p2T7-177 original. As ligações foram transformadas em *Escherichia coli* DH5 alpha cálcio competente. As transformações foram em seguida plaqueadas em meio LB contendo IPTG X-GAL e 25 μ g/ml de ampicilina

A busca pelos clones foi por análise em gel de agarose do DNA das colônias bacterianas. As colônias foram lisadas com o seguinte tampão: 50 mM de NaOH, 0,5% de SDS, 5% de Glicerol, 5 mM de EDTA, aproximadamente 1 mg de azul de bromofenol. Quinze microlitros deste tampão foram adicionados para cada colônia analisada e incubados por 10 min a 65 °C. O conteúdo desta lise foi analisado por eletroforese em gel de agarose, submergindo o gel após a entrada do material na matriz. Os clones positivos foram propagados e o DNA plasmidial isolado e purificado pelo kit *QIAprep Spin Miniprep kit* da Qiagen. As minipreps dos clones foram digeridas com *XbaI* e *XhoI* para liberação do inserto e confirmação dos clones.

4.4 Transfecção em *T. brucei*

Os clones obtidos foram linearizados na seguinte reação: dois tubos contendo cada um 15 U de *NotI*, tampão D (Amershan), 100 μ g/ml de BSA, 5 μ g de DNA plasmidial, em um volume final de 30 μ l, incubados por 16 h a 37 °C. Este DNA foi

transfectado em *T. brucei* Lister 427, forma procíclica (PC), clone 29-13 com o seguinte protocolo: 4×10^7 células foram baixadas por centrifugação 5 min a $6.000 \times g$ e $4^\circ C$, lavadas com tampão de eletroporação (NaCl 129 mM, KH_2PO_4 1,5 mM, KCl 8 mM, NaH_2PO_4 8 mM, $MgCl_2$ 1,5 mM, $CaCl_2$ 0,09 mM, CH_3COONa 2,4 mM, pH 7,0) e ressuspensas em 400 μl deste mesmo tampão. Esta suspensão foi transferida para uma cubeta de eletroporação de 0,4 cm de *gap* (Bio-Rad) e, em seguida, 10 μg do DNA linearizado foram adicionados, a cubeta foi então deixada no gelo por 10 min. Adicionalmente, dois controles foram feitos O primeiro, controle da transfecção, no qual não foi adicionado DNA, e o segundo, contendo o vetor p2T7-177 sem inserto como controle para os experimentos posteriores. Após a incubação em gelo, as células foram submetidas a 2 pulsos no eletroporador (Gene Pulse II – Bio-Rad) na seguinte configuração: 1,6 kV, 25 μF , *time constant* de aproximadamente 0,8 s. As células foram em seguida transferidas para um tubo Falcon de 15 ml contendo 3,6 ml de meio SDM-79, suplementado com 10% de soro fetal bovino (SFB), 15 $\mu g/ml$ de G418, e 50 $\mu g/ml$ de higromicina, e incubadas a $28^\circ C$ em B.O.D. Após 24 h, adicionou-se 5 $\mu g/ml$ de fleomicina à cultura. Após 5 dias da transfecção, a cultura foi diluída 1:4 em um volume de 4 ml e cultivada até a detecção de crescimento celular (aproximadamente 10 dias). Alíquotas das células transfectantes obtidas foram estocadas em nitrogênio líquido em 20% de glicerol e 40% de SFB para posteriores análises.

4.5 Curvas de Crescimento

O teste de crescimento de *T. brucei* na ausência de soro e a validação do meio produzido no Instituto Carlos Chagas foram realizados em tubo falcon de 15 ml, com contagem do número de células e repique de três em três dias. Na avaliação do crescimento após RNAi, os parasitas transfectantes foram cultivados em placas de 24 poços em estufa com 5% de tensão de CO_2 a $28^\circ C$. O inóculo inicial foi de 10^6 células/ml, sendo que cada cultivo foi dividido em dois poços, em um deles adicionou-se diariamente 1 $\mu g/ml$ de tetraciclina para indução do RNAi. A contagem do número de células foi feito em dois formatos. No primeiro, as culturas foram avaliadas diariamente até a fase estacionária. No segundo, a avaliação foi feita a

cada dois dias e as culturas foram refrescadas para o inóculo inicial de 10^6 células/ml e contadas por aproximadamente 8 dias. Neste formato foram realizadas curvas de crescimento em meio com SFB dializado e em meio com SFB padrão. As contagens foram feitas em câmara de Neubauer para o primeiro formato e em contador automatizado (Beckman Coulter Z2) para o segundo formato. O experimento foi realizado em triplicata biológica.

4.6 Avaliação dos Níveis de mRNA por QPCR

Após a seleção dos transfectantes, os níveis de mRNA das MAPKs foram avaliados. Para isto, foram extraídas amostras de RNA total de culturas no 3º e no 4º dia após a indução do RNAi (total de 64 amostras) utilizando-se o kit *RNeasy* da Qiagen. Aproximadamente 1×10^8 células foram usadas para cada extração, rendendo cerca de 30 µg de RNA total. A cultura controle (transfectada com o vetor sem inserto) foi tratada com tetraciclina da mesma forma que as culturas induzidas. O RNA total foi retrotranscrito *in vitro* para obtenção do cDNA na seguinte reação: 1 µg de RNA e 1 µM de oligo dT foram misturados e incubados por 10 min a 70 °C. Então, 4 µl de Improm-II buffer (Promega), $MgCl_2$ 3 mM, dNTP 0,5 mM cada, 40 U de RNaseOUT (Invitrogen) e 2 µl de Improm-II Reverse Transcriptase (Promega) foram misturados em um volume final de 20 µl e incubados por 2 h a 42°C. O produto desta reação foi então purificado em Microcon YM-30 (Millipore) e ressuspenso em água na concentração de 2 ng/µl. As reações de QPCR incluíram 10 ng ou 0,12-30,00 ng (curva padrão) de DNA simples fita como molde, 0,25 µmol de cada oligo e SYBR Green PCR Master Mix (Applied Biosystems) em um volume final de 20 µl. Um branco, sem adição da amostra, foi realizado para cada par de oligos. A fim de fazer a normalização entre as amostras, foram usados três genes referência de *T. brucei*, actina, proteína paraflagelar (PFR) e transcriptase reversa do telômero (TERT) [87]. As reações foram feitas em triplicata técnica no equipamento da Applied Biosystems 7500 Real-Time PCR System. Os dados foram adquiridos através do programa Real-Time PCR System Detection Software v1.4 (Applied Biosystems). As análises foram realizadas usando-se a média de 3 quantificações para cada amostra, usando-se como guia o seguinte manual: *Guide to Performing*

Relative Quantitation of Gene Expression Using Real-Time Quantitative PCR da Applied Biosystems.

4.7 Meio de Cultivo e Incorporação dos Aminoácidos Pesados

O meio SDM-79 possui 70 componentes, sendo alguns presentes em baixa quantidade. Sendo assim, estes componentes foram previamente solubilizados em solventes em concentrações adequadas, e então utilizados na elaboração do meio (ANEXOS -

Tabela 24). Os reagentes foram adicionados um a um para que fosse possível a exclusão de lisina e arginina. Para a validação do meio, realizou-se uma curva de crescimento com *T. brucei* clone 29-13, comparando-se o meio em questão com um meio comercial. Durante a curva foram realizados repiques de 10^6 células/ml a cada três dias. Para avaliar a possibilidade de uso do SDM-79 sem SFB, cultivou-se *T. brucei* selvagem na presença de 10%, 5%, 1% ou na ausência de SFB. As células foram contadas a cada 3 dias e repicadas para 10^6 células por ml.

A metodologia de marcação por SILAC (*Stable Isotope Labelling by Amino acids in Cell culture*) foi escolhida para este projeto. Como esta metodologia ainda não estava estabelecida para Tripanossomatídeos, foi necessário o estabelecimento do tempo e eficiência na incorporação dos aminoácidos com isótopos pesados. Para isto, a partir de uma cultura de *T. brucei* 427 selvagem foi feito um inóculo em meio contendo Lys8 (6 carbonos 13 + 2 nitrogênios 15) e Arg6 (6 carbonos 13) da empresa Cambridge Isotope Laboratories. Após 6 dias de cultivo (com repiques a cada 2 dias), as células foram lisadas e processadas para espectrometria de massas. A amostra foi analisada por LC-MS/MS e os dados gerados foram analisados pelo algoritmo MaxQuant versão 1.3.0.5.

4.8 Preparo das Amostras para Proteômica

Para o início do experimento, foram utilizados dois cultivos, ambos cultivados por pelo menos 6 dias ou em meio pesado (Arg6 / Lys8) ou em meio leve (Arg0 / Lys0), suplementados com 10% de SFB dializado. No dia 0, foram feitos dois inóculos (leve e pesado) de 25 ml contendo 10^6 células /ml cada. Um deles foi induzido com 2 μ g/ml de tetraciclina no dia 0 (após, 1 μ g/ml diariamente). Após dois

dias, as culturas foram repicadas aumentando-se o volume do inóculo para 225 ml. Após dois dias de cultivo, as células dos dois cultivos foram misturadas (1:1 em número de células), baixadas por centrifugação 5 min a 6.000 x g e 4 °C, lavadas duas vezes com PBS e ressuspensas em tampão de lise para proteômica (N-octilglicosídeo 1%, DTT 1 mM, inibidor de protease – Roche, uréia 6 M, tiouréia 1 M, fluoreto de sódio 50 mM, pirofosfato de sódio 10 mM, ortovanadato de sódio 1 mM, β-glicerofosfato 10 mM, MgCl₂ 1 mM, 75 U benzonase (Novagen) / mg de proteína). Uma quantidade de 2 x 10⁷ células pesadas foi separada antes da mistura e processada em paralelo para verificação da taxa de incorporação dos aa pesados.

Estes extratos foram incubados por 30 min a temperatura ambiente (TA) em agitador orbital e armazenados a -80 °C até posterior processamento. Após descongelamento, adicionou-se SDS ficando em uma concentração de 4% no extrato e incubou-se com agitação (500 rpm) por 20 min, a seguir os debris foram retirados por centrifugação a 14.000 x g por 10 min. A dosagem de proteínas na amostra foi baseada em dosagem feita por BCA (Sigma) em extrato solubilizado no mesmo tampão, sem a adição dos inibidores de fosfatase, pois estes interferem no método de BCA. O sobrenadante (aproximadamente 10 mg de proteína) foi tratado de acordo com o protocolo FASP (*Filter Aided Sample Preparation*) [88]. O extrato (3 ml) foi misturado com 8 ml de uréia 8 M, 0,1 M Tris-HCl, pH 8,5 (UA) e transferido para um filtro de Amicon Ultra 15 (Millipore). Em seguida centrifugou-se a 3.000 x g por 40 min a 20 °C, lavou-se com 10 ml de UA centrifugando-se nas mesmas condições, alquilou-se as proteínas com 2 ml de solução de iodacetamida 55 mM por 30 min no escuro TA. Em seguida, o filtro foi lavado duas vezes com 10 ml de UA, duas vezes com 10 ml de bicarbonato de amônio 50 mM (ABC) e as proteínas digeridas por 19 horas a 37 °C com 130 ug de tripsina (relação de 1:100, tripsina:amostra em massa) (Promega cat. V5113). Os peptídeos foram eluídos por centrifugação, seguida de lavagem com 1 ml de ABC e 1 ml de NaCl 0,5 M.

4.9 Fracionamento da Amostra

Após eluição dos peptídeos, os mesmos foram fracionados por cromatografia de fase reversa em pH básico (RPHP), usando-se 10 cartuchos de C18 (Sep-Pak

C18 Plus Light Cartridge Waters). Para isto, aos peptídeos eluídos (4,5 ml), adicionou-se 95,5 ml de formiato de amônio (FA) 20 mM pH 10, diluído 1:10 de uma solução 200 mM (hidróxido de amônio 0,8%, ácido fórmico 0,16%). Estes 100 ml foram distribuídos nas 10 colunas de C18, previamente ativadas com 10 ml de metanol e equilibradas com 10 ml de FA 20 mM. Em seguida lavou-se as colunas com 10 ml de FA 20 mM e fez-se a eluição em cinco passos sequenciais de 1 ml, usando-se cinco soluções diferentes contendo FA 20 mM e diferentes concentrações de acetonitrilo (10, 14, 18, 26 e 40%). As cinco frações foram concentradas em speed vac por aproximadamente 6 horas. O equivalente a 10% da amostra foi purificado em tips de C18 e destinado à proteômica. Para isso, uma ponteira (2-200 µl) foi preenchida com uma porção de membrana de C18 (Empore SPE Disks C18), a qual foi ativada com 100 µl de metanol e lavada com 200 µl de ácido fórmico 0,1% (solução A) por centrifugação a 1.000 x g. A amostra foi então passada na ponteira por centrifugação e lavada com 200 µl de A. A eluição dos peptídeos de cada uma das cinco frações foi feita com acetonitrilo 80%, ácido fórmico 0,1%. O eluído foi concentrado em speed vac e ressuspenso em uma concentração de 0,5 µg/µl de solução A. Foram injetados 5 microgramas no espectrômetro de massas. Os outros 90% da amostra foram destinados ao enriquecimento de fosfopeptídeos.

4.10 Enriquecimento dos Fosfopeptídeos

4.10.1 *Dióxido de Titânio – Protocolo 1*

Os fosfopeptídeos foram enriquecidos por dióxido de titânio (TiO₂) seguindo-se o protocolo de Bodenmiller e colaboradores [89]. Cada fração contendo aproximadamente 0,5 mg de peptídeos (para os testes de enriquecimento utilizou-se 1 mg de extrato total digerido), previamente dessalinizados em C18 (Waters, cat. WAT023501), foi ressuspenso em 1 ml de TFA 2,5%, acetonitrilo (ACN) 80% e 70 mg/ml de ácido ftálico. Em seguida, 4 mg de esferas de TiO₂ equilibradas em 100 µl de TFA 2,5%, ACN 80% foram adicionadas à amostra e incubadas por 15 min com homogeneização por inversão. As esferas foram baixadas a 1.000 x g e o sobrenadante incubado com nova alíquota de TiO₂. Foram realizadas 3 incubações sequenciais (nos testes fez-se apenas uma incubação). As esferas foram então

lavadas duas vezes com 500 µl de TFA 2,5%, ACN 80%, 70 mg/ml ácido ftálico, duas vezes com 500 µl de TFA 0,1%, ACN 80% e duas vezes com 500 µl de TFA 0,1%. Os peptídeos foram eluídos com 100 µl de NH₄OH 0,3 M, secos em *speed vac* (aproximadamente 1 h), ressuspensos em 25 µl (15 µl nos testes de enriquecimento) de TFA 1%, ACN 5%, centrifugados a 14.000 x g por 3 min para eliminar beads carregadas na eluição. Para cada análise por LC-MS/MS foram usados 10 µl.

4.10.2 Dióxido de Titânio – Protocolo 2

A quantia de 1 mg de extrato digerido dessalinizado em C18 (Waters, cat. WAT023501) foi ressuspensa em 1 ml de TFA 5%, 80% acetonitrilo (ACN) e 76 mg/ml de ácido glicólico. Em seguida, 4 mg de esferas de TiO₂ foram adicionadas à amostra e incubadas por 15 min com homogeneização por inversão. As esferas foram baixadas a 1.000 x g e o sobrenadante descartado. As esferas foram então lavadas com 100 µl de TFA 5%, ACN 80%, 76 mg/ml ácido ftálico, em seguida com 100 µl de TFA 1%, ACN 80% e então com 100 µl de TFA 0,1%, ACN 10%. As beads foram secas deixando-se o tubo aberto na bancada por 10 min. Os peptídeos foram eluídos com 100 µl de NH₄OH 1,12%, pH 11,3. O eluato foi passado por um tip de fase reversa de C8, seco em *speed vac* e ressuspensa em 15 µl de TFA 1%, ACN 5%. Para cada análise por LC-MS/MS foram usados 10 µl.

4.10.3 Dióxido de Titânio – Protocolo 3

Este protocolo foi executado de acordo com Larsen e colaboradores [69] com algumas modificações. Um mg de extrato digerido (1,3 ml) foi misturado com 260 µl de tampão de incubação – LB - (TFA 1%, acetonitrilo (ACN) 80% e ácido dihidroxibenzóico 30 mg/ml). Em seguida, 4 mg de esferas de TiO₂ em 20 µl de LB foram adicionadas à amostra e incubadas por 30 min com homogeneização por inversão. As esferas foram baixadas a 1.000 x g e o sobrenadante descartado. As esferas foram então lavadas com 1,5 ml de TFA 3%, ACN 30%, em seguida com 1,5 ml de TFA 0,1%, ACN 80%. Os peptídeos foram eluídos com 50 µl de NH₄OH 3,5%, ACN 20%, pH 11 e com 50 µl de NH₄OH 3,5%, ACN 40%, pH 11, secos em *speed vac* e ressuspensos em 15 µl de 1% TFA, 5% ACN. A seguir centrifugou-se os tubos

a 14.000 x g por 3 min para eliminar beads carregadas na eluição Para cada análise por LC-MS/MS foram usados 10 µl.

4.10.4 Dióxido de Titânio – Protocolo 4

Neste teste utilizou-se uma adaptação do protocolo de Lundby e colaboradores [90]. A quantidade de 1 mg de extrato digerido dessalinizado em C18 (Waters, cat. WAT023501) foi ressuspensa em 1 ml de tampão de incubação contendo ácido acético 0,5%, acetonitrilo (ACN) 80% e 70 mg/ml de ácido ftálico. Em seguida, 4 mg de esferas de TiO₂ equilibradas em 100 µl do tampão de incubação foram adicionadas à amostra e incubadas por 15 min com homogeneização por inversão. As esferas foram baixadas a 1.000 x g e o sobrenadante descartado. As esferas foram então lavadas com 100 µl de KH₂PO₄ 5 mM, ACN 30%, KCl 350 mM, seguida por lavagem com 100 µl de ACN 40%, ácido acético 0,5%, TFA 0,05% e ressuspensas em 50 µl de ACN 80%, ácido acético 0,5%. As esferas foram transferidas para um tip contendo C8, pré-equilibrado com ACN 80%, ácido acético 0,5% e lavadas com este mesmo tampão. A eluição foi feita com 2 x 10 µl de NH₄OH 5% e 2 x 10 µl de NH₄OH 10%, ACN 25%. O eluato foi seco em *speed vac* e ressuspenso em 100 µl de TFA 1%, ACN 5% e passado por um tip contendo C18 pré-condicionado com 20 µl de metanol, 20 µl de ACN 80%, ácido acético 0,5%, 2 x 20 µl de TFA 1%, ACN 3%. Este tip foi lavado com 2 x 20 µl de ACN 8%, ácido acético 0,5% e 1 x com 50 µl de ácido acético 0,5%. Os peptídeos foram eluídos com ACN 80%, ácido fórmico 0,1%, secos em *speed vac* e ressuspensos em 15 µl de TFA 1%, ACN 5%. Para cada análise por LC-MS/MS foram usados 10 µl.

4.10.5 Dióxido de Titânio – Protocolos 5 e 6

Este protocolo foi executado de acordo com a metodologia usada no Laboratório Nacional de Luz Síncrotron (LNLS) com algumas adaptações. Um mg de extrato digerido foi misturado com 4 mg (protocolo 5) ou 10 mg (protocolo 6) de esferas de TiO₂ e com TFA para uma concentração final de 0,1 %. Esta mistura foi incubada por 15 min com homogeneização por inversão. As esferas foram baixadas a 1.000 x g e o sobrenadante descartado. As esferas foram então lavadas com 500

μl de TFA 0,2%, ACN 80% e os peptídeos eluídos com 100 μl de NH_4OH 0,5%, ACN 40%, pH 10,5, secos em *speed vac* e ressuspensos em 15 μl de TFA 1%, ACN 5%. A seguir centrifugou-se os tubos a 14.000 x g por 3 min para eliminar beads carregadas na eluição Para cada análise por LC-MS/MS foram usados 10 μl .

4.11 Análise no Espectrômetro de Massas

Foram injetados 10 μl (5 μg) de amostra no cromatógrafo nano LC da Eksigent. A separação dos peptídeos foi feita com gradiente linear de 120 min (fosfoproteômica) ou 240 min (proteômica) de 5 a 40% de ACN em coluna analítica de 15 cm, com diâmetro interno de 75 μm e partículas de C18 de 3 μm . Os peptídeos foram ionizados por nano eletro spray (voltagem de 2,3 kV) e injetados no espectrômetro de massas LTQ Orbitrap XL ETD (Thermo Scientific). O método de análise foi o seguinte: *scan* inicial no Orbitrap com resolução de 15.000, seguido de seleção dos 10 íons mais intensos, os quais foram fragmentados por CID e analisados no *ion trap*. Paralelamente ao MS2, foi realizado um *full scan* no Orbitrap com uma resolução de 60.000. Na seleção dos íons foi usada uma lista de exclusão de 90 segundos. Nas corridas de fosfoproteômica, o isolamento e fragmentação dos precursores contendo grupamento fosfato foi feita utilizando-se *multistage activation*. A opção *lock mass* [91] foi utilizada para obtenção de melhor acurácia – erro abaixo de 0,5 p.p.m. - de massa dos peptídeos tripticos precursores detectados por MS.

4.12 Análise dos Dados de Proteômica e Fosfoproteômica

Os arquivos gerados (.RAW) foram analisados no algoritmo MaxQuant versões 1.3.0.5 e 1.4.0.8 [62]. Os parâmetros usados foram Maximum miss-cleavage sites = 2; maximum charge state = 7; enzima tripsina/P; modo específico e os descritos na

Tabela 5.

Tabela 5: Parâmetros selecionados no algoritmo MaxQuant.

Parâmetro	Valor	Parâmetro	Valor
Version	1.4.0.8	Peptides used for protein quantification	Razor

Fixed modifications	Carbamidomethyl (C)	Discard unmodified counterpart peptides	TRUE
Decoy mode	revert	Min. ratio count	2
Special AAs	KR	Site quantification	Use least modified peptide
Include contaminants	TRUE	Re-quantify	TRUE
MS/MS tol. (FTMS)	20 ppm	Use delta score	FALSE
Top MS/MS peaks per 100 Da. (FTMS)	10	iBAQ	FALSE
MS/MS deisotoping (FTMS)	TRUE	iBAQ log fit	FALSE
MS/MS tol. (ITMS)	0.5 Da	MS/MS recalibration	FALSE
Top MS/MS peaks per 100 Da. (ITMS)	8	Match between runs	TRUE
MS/MS deisotoping (ITMS)	FALSE	Matching time window [min]	1
MS/MS tol. (TOF)	0.1 Da	Alignment time window [min]	20
Top MS/MS peaks per 100 Da. (TOF)	10	Find dependent peptides	FALSE
MS/MS deisotoping (TOF)	TRUE	Fasta file	C:\Databases\TbruceiTreu927AnnotatedProteins_TriTrypDB-4.2.fasta
MS/MS tol. (Unknown)	0.5 Da	Labeled amino acid filtering	TRUE
Top MS/MS peaks per 100 Da. (Unknown)	10	Site tables	Oxidation (M)Sites.txt;Phospho (STY)Sites.txt
MS/MS deisotoping (Unknown)	FALSE	Cut peaks	TRUE
PSM FDR	0.01	Decoy mode	Revert
Protein FDR	0.01	Special AAs	KR
Site FDR	0.01	Include contaminants	TRUE
Use Normalized Ratios For Occupancy	TRUE	RT shift	FALSE
Min. peptide Length	7	Advanced ratios	FALSE
Min. score for unmodified peptides	0	AIF correlation	0.47
Min. score for modified peptides	40	First pass AIF correlation	0.8
Min. delta score for unmodified peptides	0	AIF topx	20
Min. delta score for modified peptides	17	AIF min mass	0
Min. unique peptides	0	AIF SIL weight	4
Min. razor peptides	1	AIF ISO weight	2
Min. peptides	1	AIF iterative	TRUE
Use only unmodified peptides and	TRUE	AIF threshold FDR	0.01
Modifications included in protein quantification	Acetyl (Protein N-term);Oxidation (M)		

As tabelas geradas pelo MaxQuant foram analisadas no algoritmo Perseus versão 1.4.0.11. Os contaminantes e as entradas identificadas no banco de dados reverso foram retirados da tabela. Para os fosfósílios, foram usados somente os com probabilidade de localização igual ou superior a 0,75 e com Delta Score maior ou igual a 5. As razões H/L obtidas para os fosfósílios foram normalizados pelas razões

H/L das respectivas proteínas. Tanto para proteômica quanto fosfoproteômica foram filtradas as entradas encontradas em apenas uma das réplicas técnicas. O teste estatístico aplicado foi o teste-t usando-se os valores de p menores ou iguais a 0,05.

4.13 RNA-seq

As amostras para RNAseq foram extraídas de culturas em fase exponencial de crescimento da seguinte maneira: foram inóculos iniciais de 10^6 células / ml, indução com tetraciclina, repique no segundo dia após indução e extração de RNA no quarto dia após indução. O RNA total foi purificado com o kit RNeasy (Qiagen) de acordo com as especificações do fabricante. O preparo das amostras foi realizado segundo o protocolo SOLiD *Total RNA Seq Kit*. O RNA mensageiro foi obtido e purificado a partir do kit *PolyATtract mRNA Isolation Systems* (Promega) e do produto *RiboMinus Concentration Module* (Invitrogen), e teve sua qualidade verificada pelo equipamento Bioanalyzer.

A ligação dos adaptadores e a transcrição reversa do RNA foi feita com os kits SOLiD *Small RNA Expression*. Após o término da reação de transcrição reversa, o material era purificado com o kit *MinElute PCR Purification* (Qiagen) de acordo com as instruções do fabricante.

O material obtido da transcrição reversa foi fracionado em gel de poliacrilamida 6% desnaturante (sistema Novex - Invitrogen). A banda de gel contendo fragmentos de 100-200 foi excisada. O material contido no fragmento de gel excisado foi amplificado por PCR também com o kit SOLiD *Small RNA Expression*. Este material amplificado foi então purificado com o kit *PureLink™ PCR Micro* (Invitrogen). A concentração e qualidade deste material foram analisadas com os equipamentos NanoDrop e Bioanalyzer.

As PCRs em emulsão (ePCR) foram realizadas segundo protocolo SOLiD™ 4 *System Template Bead Preparation Guide* (ePCR em mini-escala, produção de 75-150 milhões de beads). As beads da ePCR foram enriquecidas e combinadas com as beads de sequenciamento. As lâminas foram preparadas e o sequenciamento realizado segundo o protocolo SOLiD™ 4 *System Instrument Operation Guide*.

Na análise dos dados, fez-se o mapeamento dos reads gerados no genoma de *T. brucei* 927. Os reads foram alinhados nas condições Unique30, Unique40, Random30 e Random40. Foram aplicados FDRs de 20, 10, 5 e 1 e 0,1%.

4.14 Duplo-Híbrido

As sequências codificantes (CDSs) dos genes participantes da cascata das MAPKs de *T. brucei* foram amplificados por PCR, totalizando 40 genes (ANEXOS - Tabela 25). Nestas amplificações utilizou-se a enzima de alta fidelidade *AccuPrime Taq DNA Polymerase, High Fidelity* (Life Technologies). Nestas amplificações, foram inseridas as sequências para recombinação pelo sistema de clonagem Gateway (Invitrogen). Os produtos de PCR foram purificados e clonados no vetor de entrada pDONR 221 de acordo com manual Clonase II (Invitrogen). A seguir os clones resultantes (clones de entrada) foram recombinados com os vetores para duplo híbrido pAD (pDEST22) e pBD (pDEST32), ambos linearizados utilizando a enzima de restrição *SmaI*.

Após esta etapa, as recombinações foram transformadas em *E. coli* DH5-alfa cálcio competente. Os clones foram confirmados por análise de DNA das colônias bacterianas (conforme descrito no tópico 4.3). A extração dos plasmídeos foi realizada através do kit CosMCPrep (Agencourt).

Os plasmídeos recombinantes pAD foram transformados através de choque térmico nas cepas MAV203 e Y8800 de *Saccharomyces cerevisiae*, enquanto os plasmídios pBD foram inseridos nas cepas MAV103 e Y8930. Os clones pAD e pBD foram selecionados em meio na ausência triptofano ou de leucina, respectivamente. Cada cepa foi estriada em placa de meio sólido YPD, a qual foi incubada entre 48 e 72 horas a 30 °C.

Após o crescimento, preparou-se o pré-inóculo para cada cepa inoculando 20 ml de meio YPD com 10 colônias, permanecendo a 30 °C por 18 horas sob agitação de 220 rpm. A OD600 deve estar entre 0,4 e 0,6 e, após a medição, o pré-inóculo foi diluído em 100 ml de meio YPD para a OD600 0.1. Incubou-se então sob agitação até alcançar a OD600 entre 0,6 e 0,8 (período entre quatro a seis horas). Após, centrifugou-se as culturas por 5 min a 800 x g, descartou-se o sobrenadante e

ressuspendeu-se as células em água estéril, centrifugando-as nas mesmas condições anteriores. O sobrenadante foi então descartado e as células lavadas em 10 ml da solução TE/LiAc (Tris-HCl 10 mM (pH 8.0), EDTA 1 mM (pH 8.0) e acetato de lítio 100 mM).

Após a lavagem, as células foram ressuspensas em 2 ml de TE/LiAc. Em seguida, acrescentou-se 10 ml de TE/LiAc/PEG (8 volumes de PEG3350 44% (w/v), 1 volume de TE 10x pH 8 e 1 volume de acetato de lítio 1 M) e 200 microlitros de DNA de esperma de salmão (Sigma-D9156). A suspensão foi homogeneizada por inversão e distribuída em uma placa de 96 poços (120 µl por poço). Dez microlitros de plasmídeo foram adicionados e as células foram incubadas a 30 °C por trinta minutos. Após esta etapa, realizou-se o choque térmico a 42 °C por quinze minutos, centrifugou-se as placas por 5 min a 800 x g, removeu-se o sobrenadante e acrescentou-se 100 µl de água estéril. Centrifugou-se nas condições anteriores e, em seguida, retirou-se 90 µl do sobrenadante e ressuspendeu-se as células com o restante da água. Cinco microlitros foram depositados em placas de meio seletivo sólido. As placas foram incubadas por 72 horas a 30°C. Após o crescimento, fez-se estoques em glicerol.

O processo de *mating* foi auxiliado com o uso do robô Tecan Freedom 200. As células foram descongeladas e feito inóculo em meio líquido seletivo. Foram utilizadas placas de 96 poços contendo 1 ml de cultura (em YPD), as quais foram incubadas sobre agitação a 30 °C por 72 horas. Então, fez-se a transferência manual dos clones contendo o vetor pDEST22 para uma placa de 384 poços. Estes, foram transferidos pelo robô em placa de meio sólido YPD. Em seguida, fez-se a transferência da cultura de um clone BD sobre cada um dos clones AD, propiciando assim a mistura AD / BD. Em seguida, as placas foram incubadas a 30 °C por 48 horas. Durante esta etapa as células diplóides são formadas através do processo de *mating*. Finalmente, as células foram carimbadas em placas de meio seletivo sólido sem triptofano e sem leucina e cultivadas a 30°C por 72 horas para a seleção dos diplóides.

Após a seleção dos diplóides, fez-se a transferência das culturas das placas de seleção para as placas de testes de interações. As colônias foram inoculadas nos seguintes meios fenotípicos: a) sem uracila (cepa MAV), b) sem histidina com 100

mM de 3AT (cepa MAV), c) sem histidina com 1 mM de 3AT (cepa Y), d) sem histidina com 3 mM de 3AT (cepa Y), e) sem adenina (cepa Y). Após 10 dias de incubação os resultados foram coletados e as interações proteicas obtidas identificadas.

A análise dos dados foi feita comparando-se os resultados obtidos com os observados nos bancos de dados de interações Intact (<http://www.ebi.ac.uk/intact/>) e Biogrid (<http://thebiogrid.org/>) contendo 314.020 interações. Para isso foram selecionados os grupos de ortólogos através do banco de dados do OrthoMCL. Foram selecionados os ortólogos dos 40 membros da via das MAPKs presentes nos organismos *Homo sapiens*, *Drosophila melanogaster*, *Saccharomyces cerevisiae*, *Mus musculus*, *Trypanosoma cruzi*, *Leishmania major*, *Arabidopsis thaliana* e *Caenorhabditis elegans*.

As Ids dos ortólogos encontrados foram convertidas para as Ids do Uniprot para posterior busca nos bancos de dados de interações proteicas. As redes de sinalização foram montadas no programa cytoscape versão 2.8.3.

5. RESULTADOS E DISCUSSÃO

5.1 Padronizações

5.1.1 *Elaboração do Meio de Cultivo para SILAC*

Para que haja uma marcação eficiente por SILAC utilizando aminoácidos contendo isótopos pesados, existe a necessidade de se eliminar do meio de cultivo os aminoácidos correspondentes não marcados. O meio utilizado neste trabalho é o SDM-79 [92], ao qual devem ser adicionados 10% de soro fetal bovino (SFB) para o cultivo dos parasitas. A presença do SFB diminui a eficiência na marcação por SILAC por conter aminoácidos livres. Por este motivo, algumas estratégias foram avaliadas para contornar esta questão. A estratégia mais simples seria a utilização de meio sem SFB. A fim de se testar esta possibilidade, foi utilizada uma cultura selvagem, cultivada em meio suplementado com diferentes quantidades de SFB. O que pode ser observado foi a morte dos parasitas na cultura sem soro (dia 7 - Figura 9) e um crescimento semelhante nas na presença de 10%, 5% e 1% de soro. Isto demonstra que apesar da complexidade do meio SDM-79, devem existir no SFB fatores indispensáveis para a proliferação celular.

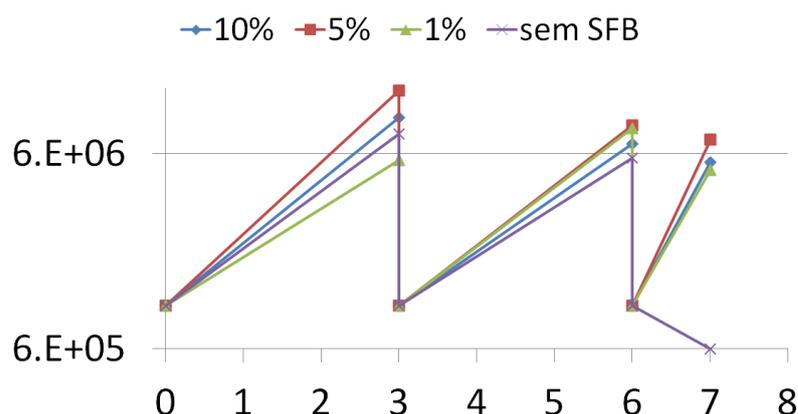


Figura 9: Proliferação celular em diferentes concentrações de soro. No eixo X estão representados os dias de cultivo, enquanto no eixo Y está o número de células por ml em log de 10.

Outra tentativa usada para eliminar os aminoácidos não marcados, foi a utilização de 10% de SFB dializado (SFBd). Este soro passa por um tratamento onde são retiradas moléculas com peso molecular menor que 10.000. Em meio suplementado com este soro, os parasitas cresceram de maneira similar ao meio com SFB normal. A utilização de SFB dializado para *T. brucei* tem sido relatada na literatura [93] [84, 85].

Com a possibilidade de se utilizar SFBd nos ensaios de marcação com SILAC, começou-se o levantamento dos componentes para confecção deste meio. Os reagentes deveriam ser adicionados um a um para que fosse possível a substituição de lisina e arginina. O meio SDM-79 possui 70 componentes, sendo alguns presentes em baixa quantidade. Tal fator, aliado à insolubilidade de alguns componentes em água, gerou uma dificuldade adicional na confecção deste meio. Os componentes em baixa quantidade foram previamente solubilizados em solventes e concentrações adequados, e então utilizados na elaboração do meio. No primeiro lote produzido, metade da quantidade do meio foi deixada sem os aminoácidos lisina e arginina, para posterior adição de isótopos estáveis destes aminoácidos. A outra metade continha lisina e arginina e foi usada para os testes iniciais. Para a validação do meio, realizou-se uma curva de crescimento com *T. brucei* clone 29-13, comparando-se o meio em questão com um meio comercial usando SFB normal. Durante a curva foram realizados repiques de 10^6 células / ml a cada três dias. Na Figura 10, os pontos demonstrados representam a densidade celular no terceiro dia após os repiques. Os parasitas cultivados no meio produzido no Instituto Carlos Chagas (ICC) apresentaram crescimento semelhante aos cultivados em meio comercial (Figura 10).

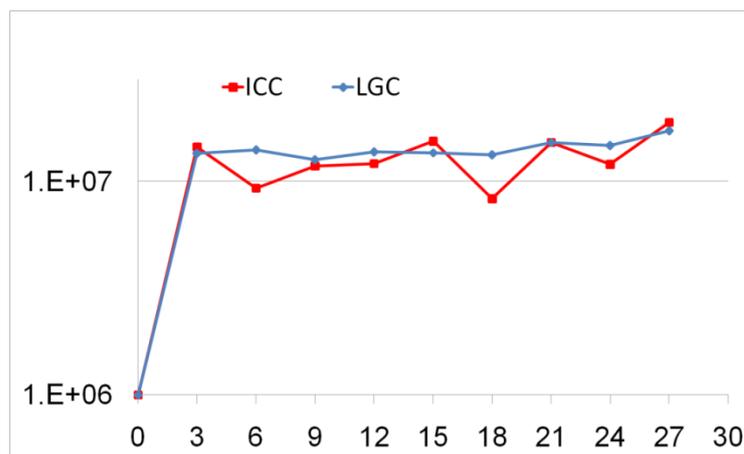


Figura 10 - Curva de crescimento de *T. brucei* 29-13 para validação do meio SDM-79. O número de células por ml está em escala logarítmica (ordenadas) e o tempo (abscissas) em dias. A linha em azul representa a cultura crescida em SDM-79 produzido no ICC, enquanto a linha em vermelho representa a cultura crescida em SDM-79 comercial (empresa LGC).

5.1.2 Diminuição das Concentrações de Lisina e Arginina em SDM-79

A marcação metabólica por SILAC é feita com a utilização de isótopos pesados de lisina e arginina (com carbono 13 e/ou nitrogênio 15). Estes aminoácidos possuem um custo elevado, onerando projetos em larga escala. A fim de diminuir os custos neste projeto, avaliou-se o crescimento de formas procíclicas de *T. brucei* 427 em SDM-79 suplementado com 10% de SFB dializado contendo diferentes concentrações de lisina (Lys) e arginina (Arg). As células foram cultivadas na concentração original de Lys e Arg e em 50%, 25%, 12,5%, e 6,25% da concentração original. As concentrações de Lys e Arg foram testadas separadamente, ou seja, enquanto se manteve a concentração original de Lys, se diminuía a concentração de Arg, e vice-versa. Também foram realizados cultivos sem Lys ou Arg. As culturas foram acompanhadas por 18 dias com repique e contagem a cada 3 dias.

As médias de crescimento após 3 dias de cultivo estão representadas na Figura 11. Na concentração de 25% de Lys ou Arg não houve diferença significativa no crescimento celular. A fim de avaliar o crescimento na presença de 25% de Lys e Arg na mesma cultura, realizou-se uma curva de crescimento comparando-se com uma cultura contendo a concentração original de Lys e Arg em SFB dializado e não-

dializado. As culturas foram acompanhadas até a fase estacionária com contagem diária do número de células. A figura mostra que não houve diferença significativa de crescimento nas culturas com 25% ou 100% de Lys e Arg, com SFB dializado ou não-dializado. Este experimento demonstra a possibilidade de utilização de quantidades menores dos isótopos de Lys e Arg na marcação por SILAC diminuindo os custos do projeto.

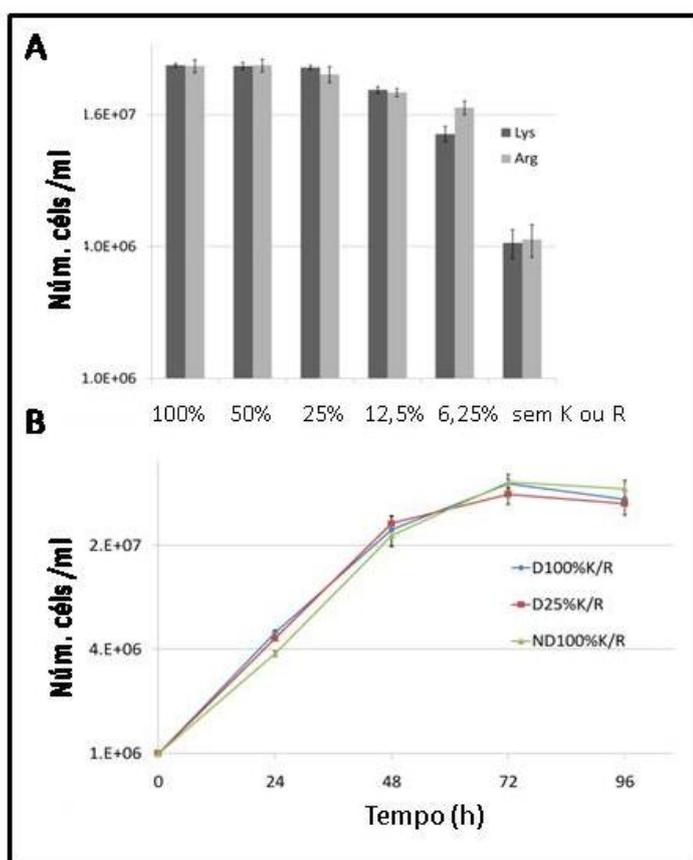


Figura 11: Avaliação do crescimento de *T. brucei* em quantidades reduzidas de lisina e arginina. Em A, o eixo das abscissas representa as concentrações de lisina ou arginina em relação às suas concentrações originais em SDM-79. Em B, D = SFB dializado, ND = SFB não-dializado. Em A e B, o eixo das ordenadas está em escala logarítmica.

5.1.3 Marcação Metabólica de Proteínas por SILAC

Experimentos de proteômica quantitativa relativa envolvem comparação entre populações celulares, como por exemplo, células tratadas e não tratadas com um

determinado fármaco, ou células induzidas e não induzidas para RNAi. As amostras podem ser comparadas por espectrometria de massas sem o emprego de marcação ou após a marcação diferencial entre duas ou mais amostras. Dentre as diversas metodologias existentes para marcação de amostras de proteínas ou peptídeos, a marcação metabólica através de aminoácidos com isótopos pesados possui características desejáveis para este projeto. A possibilidade de marcar *in vivo* as proteínas de uma determinada população celular e misturá-la com uma população não marcada diminui as eventuais interferências causadas pelo processamento individualizado das amostras. Esta característica se torna ainda mais relevante em estudos de fosfoproteômica quantitativa, nos quais a etapa de enriquecimento dos fosfopeptídeos feita individualmente pode interferir na quantificação dos sítios de fosforilação.

Diante do exposto, a metodologia de marcação por SILAC foi escolhida para este projeto. Esta metodologia ainda não estava estabelecida para *Tripanossomatídeos* no início do projeto, assim, foi necessária a elaboração do meio de cultivo para *T. brucei* compatível com SILAC e ainda o estabelecimento do tempo e eficiência na incorporação dos aminoácidos com isótopos pesados. O uso de Lys e Arg contendo isótopos pesados em experimentos com marcação por SILAC se deve ao fato da presença de pelo menos um resíduo destes aminoácidos nos peptídeos tripticos (tripsina cliva após resíduos de lisina e arginina). Nos meios de cultivo para SILAC, geralmente é usado soro dializado (SFBd). A ausência ou pelo menos diminuição da quantidade de aminoácidos livres no meio aumenta a eficiência de incorporação dos aminoácidos contendo isótopos pesados.

A incorporação dos aminoácidos ocorre ao longo do cultivo celular. Considerando-se que a cada ciclo celular ocorra a síntese de proteínas para a divisão celular, são necessários pelo menos cinco ciclos para se alcançar mais de 95% de incorporação de aminoácidos. Três cultivos de *T. brucei* 427 selvagem foram testados em meio onde se fez a substituição de Lys0/Arg/0 por Lys8/Arg6. As concentrações usadas destes aminoácidos foram 25%, 50% e 100% das respectivas concentrações presentes na formulação original do SDM-79. A ideia foi testar a utilização de menor quantidade destes aminoácidos na marcação metabólica das células, seguindo o mesmo raciocínio do experimento onde se avaliou o crescimento

do parasita em concentrações menores de lisina e arginina (item 5.1.2). As taxas de incorporação obtidas foram de 97,5%, 96,8% e 97,6%, para 25, 50 e 100 % de lisina e arginina, respectivamente. Diante dos resultado das curvas de crescimento (item 5.1.2) deste resultado, como as taxas foram semelhantes entre si e acima de 96% e considerado aqui um valor aceitável para as análises quantitativas, optou-se nos ensaios seguintes pelo uso de 25% de Lys8 e Arg6.

Na Figura 12-A está representado um par de peptídeos (SILAC *pair*) contendo 6 Dalton de diferença na relação massa/carga, extraído de uma mistura 1:1 de células marcadas e não marcadas de *T. brucei*. Esta diferença é compatível com a incorporação de uma Arg contendo carbono 13 no peptídeo usado como exemplo. Na Figura 12-B estão plotadas as intensidades das proteínas identificadas no experimento provenientes dos dois cultivos. Foi obtida uma alta correlação entre as amostras, resultado da taxa de incorporação alcançada. Estes dados, juntamente com os dados de incorporação, demonstram o estabelecimento bem sucedido desta elegante metodologia, o que possibilitou a sua utilização na busca pelos objetivos deste projeto.

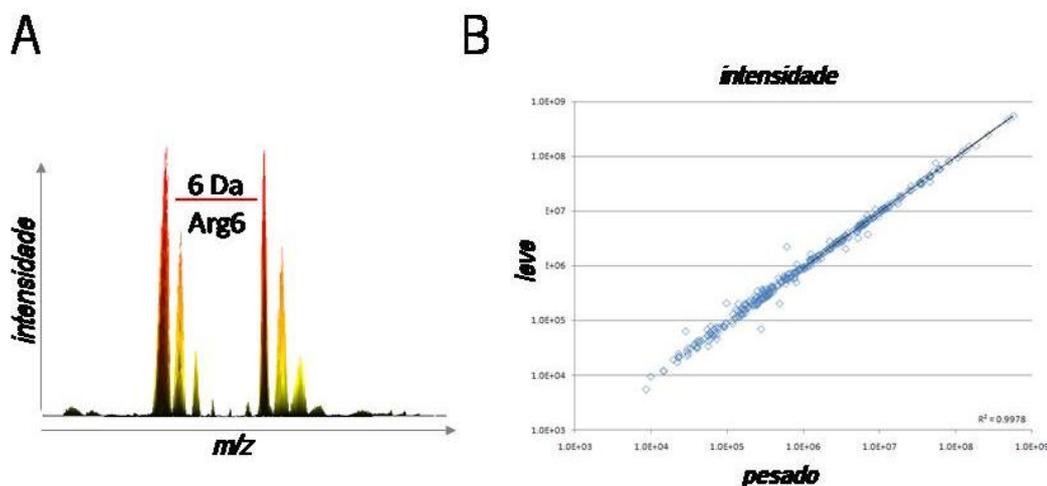


Figura 12: Avaliação da incorporação de Arg6 e Lys8 em *T. brucei*. Na figura A estão representados dois picos referentes a peptídeos de mesma composição, com exceção da troca de uma Arg0 por Arg6 no pico da direita. Da = Dalton, m/z = massa/carga. Na figura B o valor de R^2 foi de 0,9978.

5.1.4 *Fracionamento da amostra*

Os peptídeos pós digestão foram fracionados por cromatografia de fase reversa em pH básico (RPHP). Após o fracionamento, fez-se a dosagem dos peptídeos por absorvância em 280 nm (coeficiente de extinção molar de 1100 e massa média dos peptídeos de 1400 Da) para análise relativa entre as frações eluídas (Tabela 6).

Tabela 6: Dosagem dos peptídeos pós digestão e pós fracionamento.

Amostra	Réplica 1		Réplica 2		Réplica 3	
	Concentração (mg/ml)	total (mg)	Concentração (mg/ml)	total (mg)	Concentração (mg/ml)	Total (mg)
Pós-FASP	1.005	4.52	0.650	2.93	0.947	4.26
Não ligado	0.019	1.90	0.027	2.67	0.078	7.76
E 10% ACN	0.068	0.68	0.043	0.43	0.062	0.62
E 14% ACN	0.095	0.95	0.017	0.17	0.093	0.93
E 18% ACN	0.054	0.54	0.045	0.45	0.059	0.59
E 26% ACN	0.050	0.50	0.071	0.71	0.070	0.70
E 40% ACN	0.046	0.46	0.167	1.67	0.036	0.36
Total Eluído		3.13		3.42		3.19

ACN – acetonitrilo, FASP – *filter aided sample preparation*

5.1.5 *Enriquecimento de fosopeptídeos*

Devido à baixa quantidade de peptídeos fosforilados nos extratos proteicos em geral, é necessária uma etapa de enriquecimento dos fosfopeptídeos presentes na amostra. Neste trabalho, optou-se pelo enriquecimento por micro-esferas de dióxido de titânio (TiO₂), as quais possuem afinidade pelos fosfopeptídeos em condições ácidas. Peptídeos acídicos também possuem afinidade pelo TiO₂, por isso, os protocolos geralmente utilizam um ácido competidor com afinidade por intermediária por TiO₂, entre os fosfopeptídeos e os peptídeos acídicos. Aqui neste trabalho foram testados protocolos com diferentes ácidos (ácido dihidroxibenzoico, ácido ftálico e ácido glicólico - Tabela 7).

Tabela 7: Fosfossítios identificados durante a padronização do enriquecimento.

	proteínas	fosfoproteínas	%fosfoproteínas	peptídeos	fosfopeptídeos	%fosfopeptídeos	Fosfossítios	LP>0,75	%fosfopeptídeos LP>0,75
1	347	252	73	492	379	77	589	447	76
2	346	251	73	461	355	77	429	338	79
3	139	89	64	203	106	52	252	216	86
4	412	56	14	655	62	9	82	63	77
5	231	36	16	448	25	6	37	25	68
6	169	23	14	274	14	5	27	22	81

Protocolos extraídos de (1) [89], (2) Ácido glicólico, (3) [69], (4) [90], (5) e (6) LNLS. LP>0,75 – probabilidade superior a 75% de correta localização do resíduo fosforilado.

O protocolo que apresentou maior identificação de fosfossítios (589) e fosfoproteínas (252) foi o número 1, sendo que nas análises posteriores empregou-se o mesmo no enriquecimento de fosfopeptídeos.

5.2 Confirmação da Depleção de RNA Mensageiro Após RNAi

Após a seleção dos transfectantes, os níveis de RNA mensageiro (mRNA) das MAPKs (identificadas aqui de A a O, Tabela 4) foram avaliados nos 15 transfectantes e no controle (vetor sem inserto). Para isto, avaliou-se por QPCR as culturas no 3º e no 4º dia após indução do RNAi, comparando-se com as respectivas culturas não-induzidas (total de 64 amostras). Os iniciadores para as QPCRs foram desenhados de forma a amplificarem regiões distintas das usadas para o RNAi, evitando assim eventuais interferências dos RNAs dupla fita gerados no *knockdowns*. Tais iniciadores também foram obtidos pelo programa RNAit, seguindo-se os mesmos critérios em relação à especificidade do alvo (item 4.1). A fim de

normalizar as amostras, foram usados três genes referência de *T. brucei*, actina, proteína paraflagelar (PFR) e transcriptase reversa do telômero (TERT).

Os resultados demonstraram em média 80% de diminuição nos níveis de mRNA dos transfectantes induzidos em comparação com os não-induzidos (Figura 13). A diminuição de mRNA foi mais intensa no 4º dia em relação ao 3º tendo um resultado similar para os três genes referência utilizados escolhidos com base no trabalho de Brenndörfer e Boshart (2010). Os níveis de mRNA detectados no controle em geral foram menores que os detectados nas culturas não-induzidas, porém maiores que os detectados nas culturas induzidas, essencialmente no 4º dia após a indução do silenciamento (dado não mostrado).

A decisão pela extração no 3º e 4º dias após indução do RNAi foi baseada no desenho experimental para os ensaios de proteômica e fosfoproteômica. Para estes experimentos são necessárias pelo menos 3×10^8 células por réplica experimental. Diante disto, a fim de avaliar os níveis de mRNA das MAPKs e o proteoma e fosfoproteoma após o mesmo tempo de indução, a extração de RNA foi realizada nos dias em que a cultura apresenta maior densidade celular. Os resultados obtidos demonstram forte inibição 96 horas após indução.

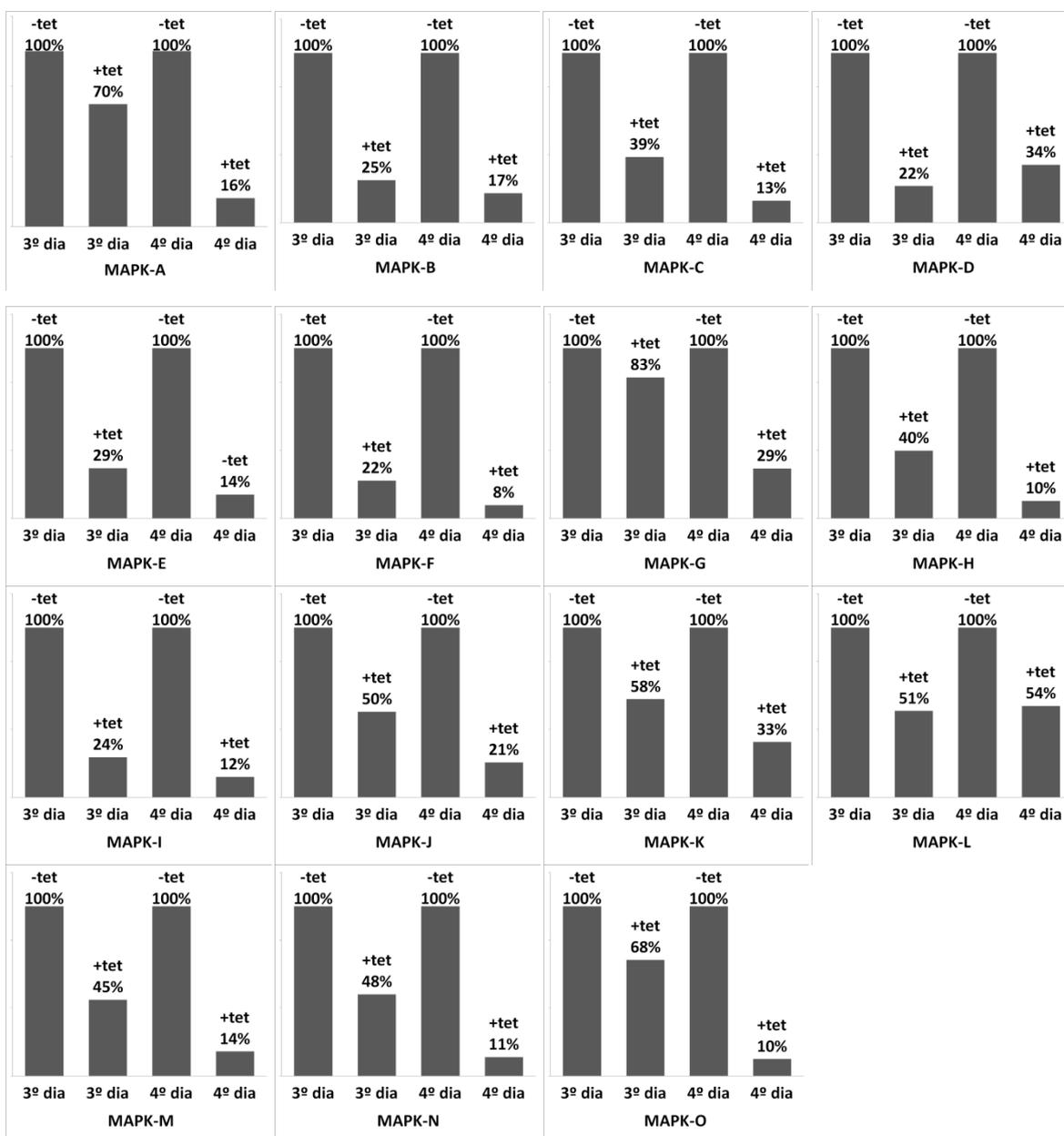
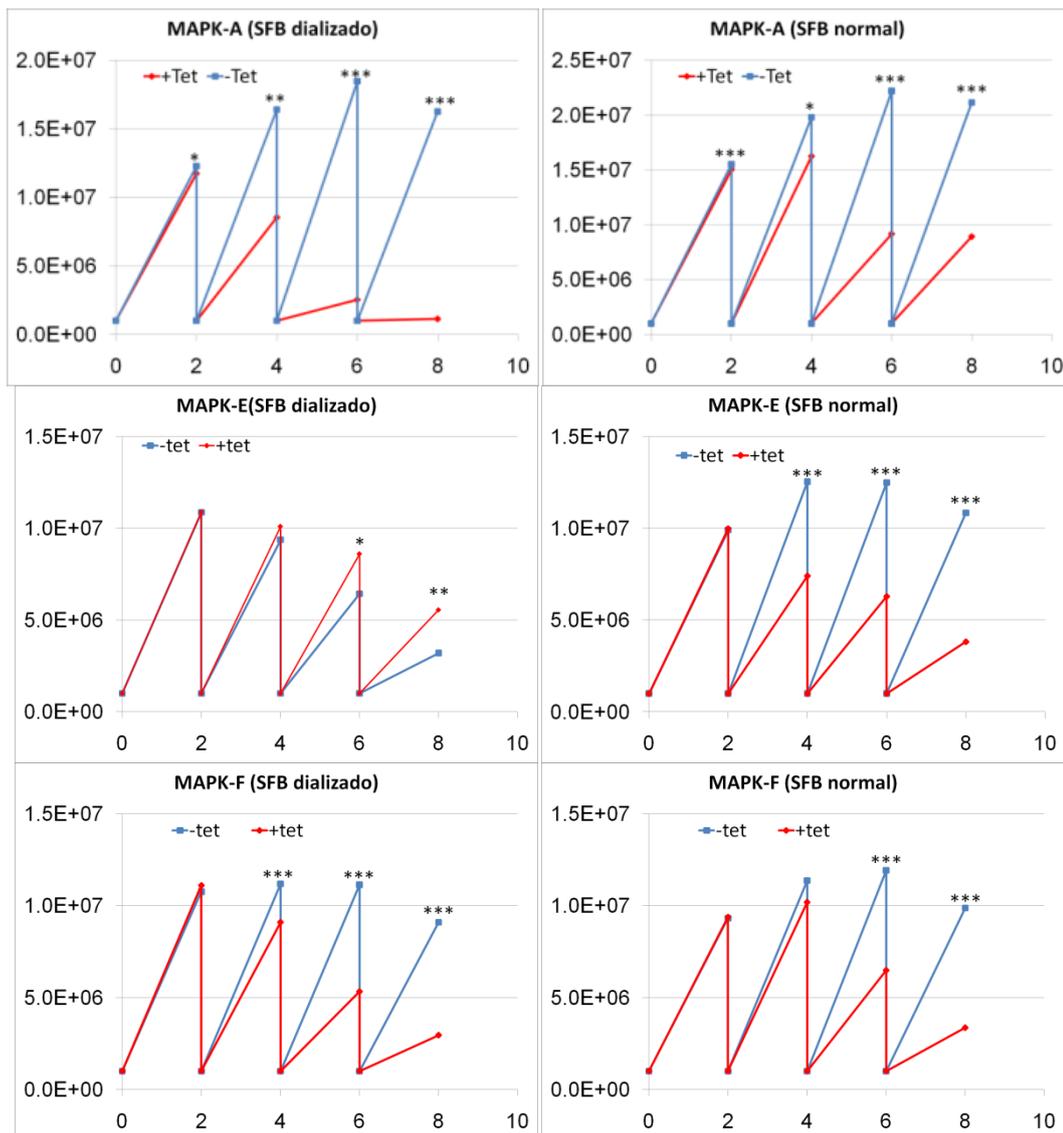


Figura 13: Quantificação relativa dos mRNAs das MAPKs depletadas. As barras representam a porcentagem do mRNA analisado nas amostras induzida (+tet) e não-induzida (-tet) para RNAi no 3º e 4º dias após indução.

5.3 Efeito da Depleção das MAPKs na Proliferação Celular

Com a finalidade de avaliar o efeito da depleção das MAPKs na proliferação celular de *T. brucei*, foram realizadas curvas de crescimento com os transfectantes

induzidos e não-induzidos utilizando-se meio com SFB dializado e com SFB normal. O crescimento celular foi avaliado em triplicata biológica em dois modelos: 1) com contagem diária do número de células até o início da fase estacionária; 2) com contagem e repique das células a cada dois dias. Utilizando-se o modelo 1, para a maioria dos transfectantes não foi observada diferença significativa no crescimento (ANEXO - Figura 30). Com o uso do modelo 2, alguns cultivos não apresentaram diferença no crescimento após indução do RNAi (Figura 15). No entanto, para alguns transfectantes houve diminuição do crescimento nas culturas com indução do RNAi (Figura 14). Os pontos das curvas foram analisados estatisticamente pelo Teste T, onde a significância dos resultados é indicada pelos valores de p. No caso da MAPK-J, na cultura induzida houve diminuição do crescimento nos pontos 48 e 72 horas. Hua e Wang (1997) demonstraram a indução *in vitro* desta quinase por interferon-gama em formas sanguíneas de *T. brucei*. Isto sugere a possibilidade de participação desta quinase na proliferação celular dos parasitas diante da secreção de interferon-gama pelo hospedeiro, embora isto necessite ser avaliado. A baixa interferência da depleção das MAPKs na proliferação celular observada no modelo 1 poderia estar relacionada ao desenho experimental. Como houve forte diminuição dos mRNAs das MAPKs somente no 4º dia após indução do RNAi (início da fase estacionária de crescimento), as respectivas proteínas poderiam ainda se apresentar em quantidades suficientes para manutenção das suas atividades durante a fase exponencial de crescimento. Esta questão foi avaliada realizando-se curvas mais prolongadas com repiques a cada dois dias. Com isto, considerando-se que os níveis proteicos das MAPKs diminuíram ao longo da curva, foi possível avaliar o crescimento celular nas primeiras 48 horas com um silenciamento gênico mais expressivo.



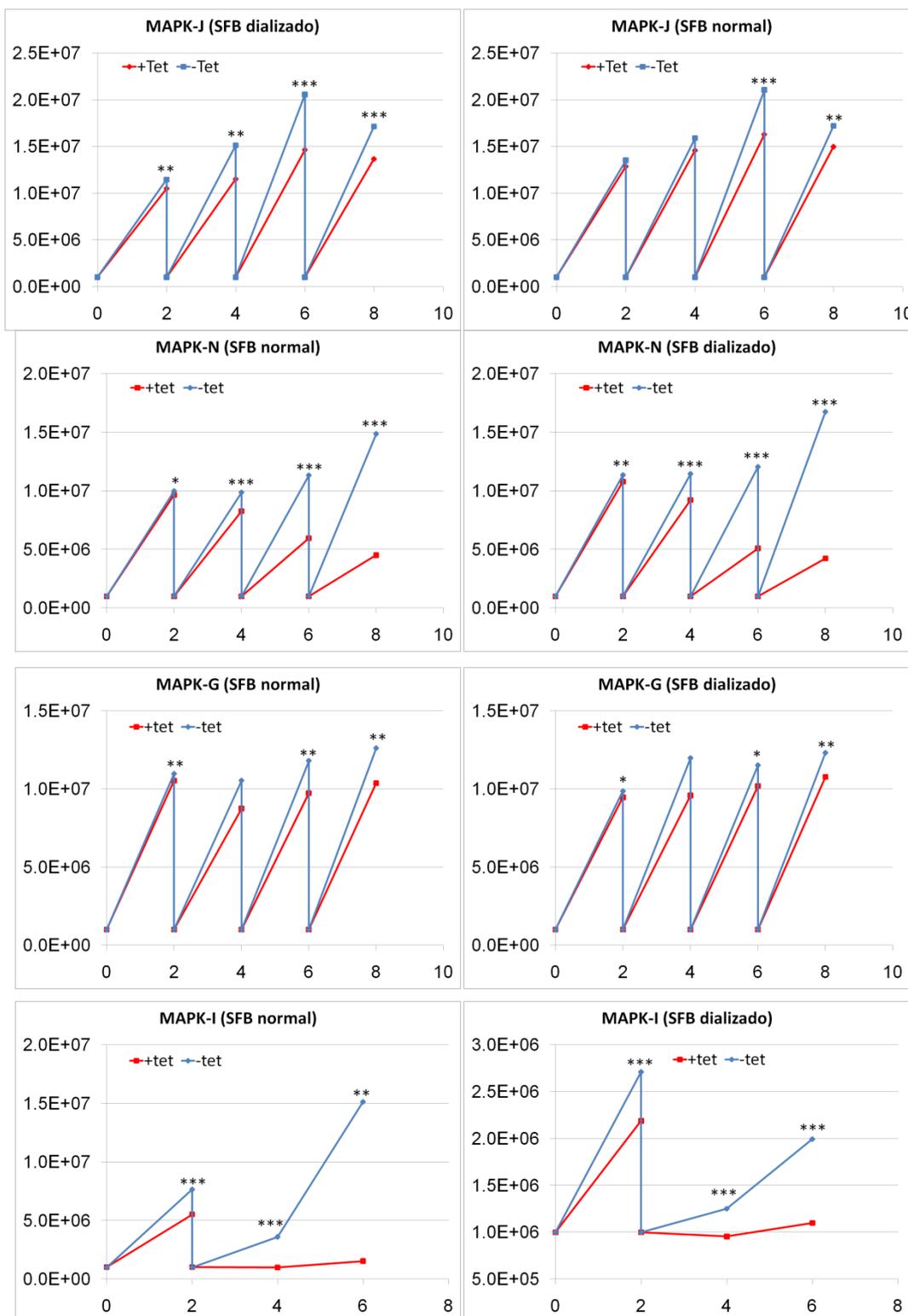
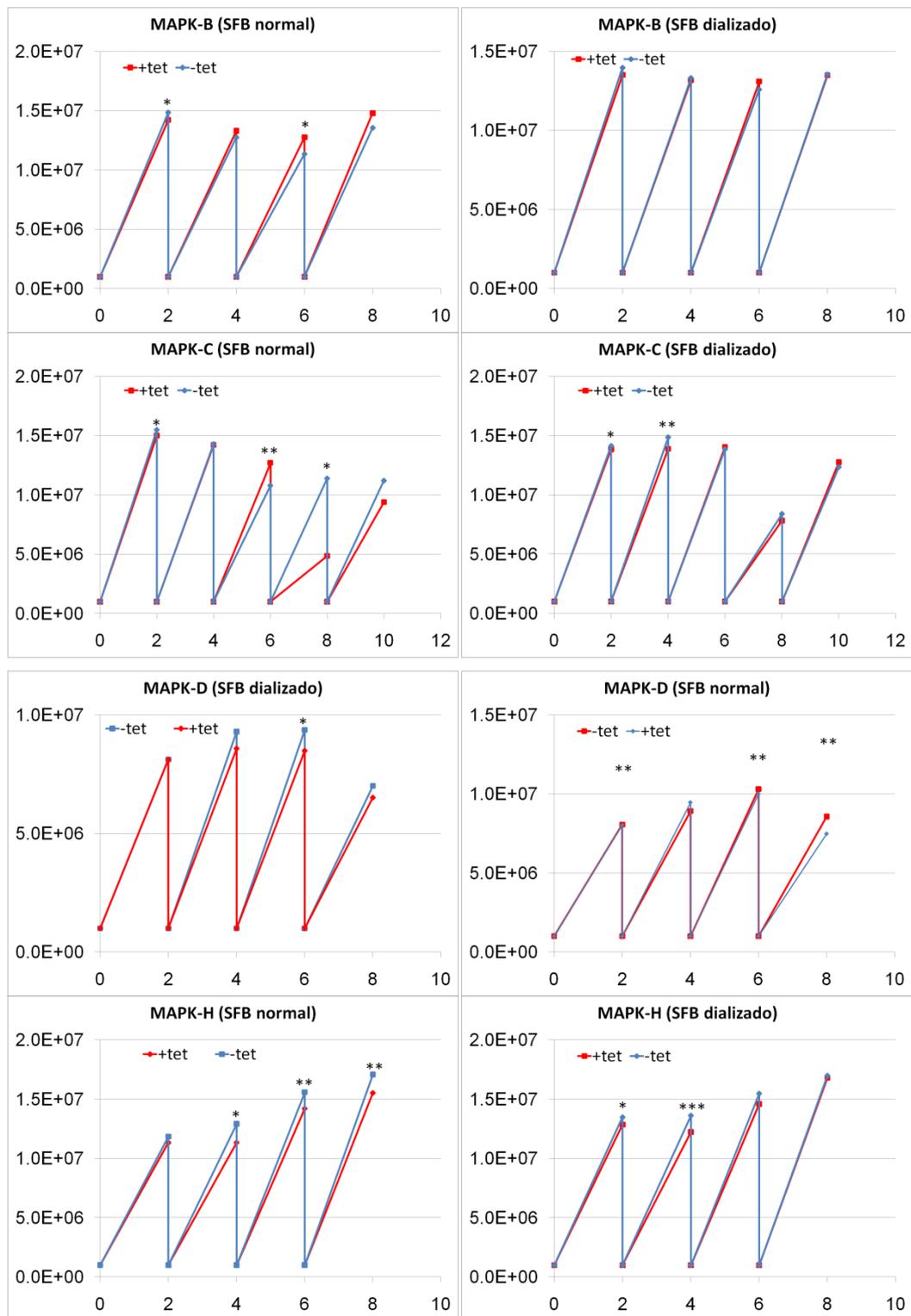


Figura 14: Proliferação celular afetada após depleção das MAPKs. No eixo X estão representados os dias de cultivo, enquanto no eixo Y está o número de células por ml em log₂. * 0,01 < pvalue < 0,05; ** 0,001 < pvalue < 0,01; *** = pvalue < 0,001.



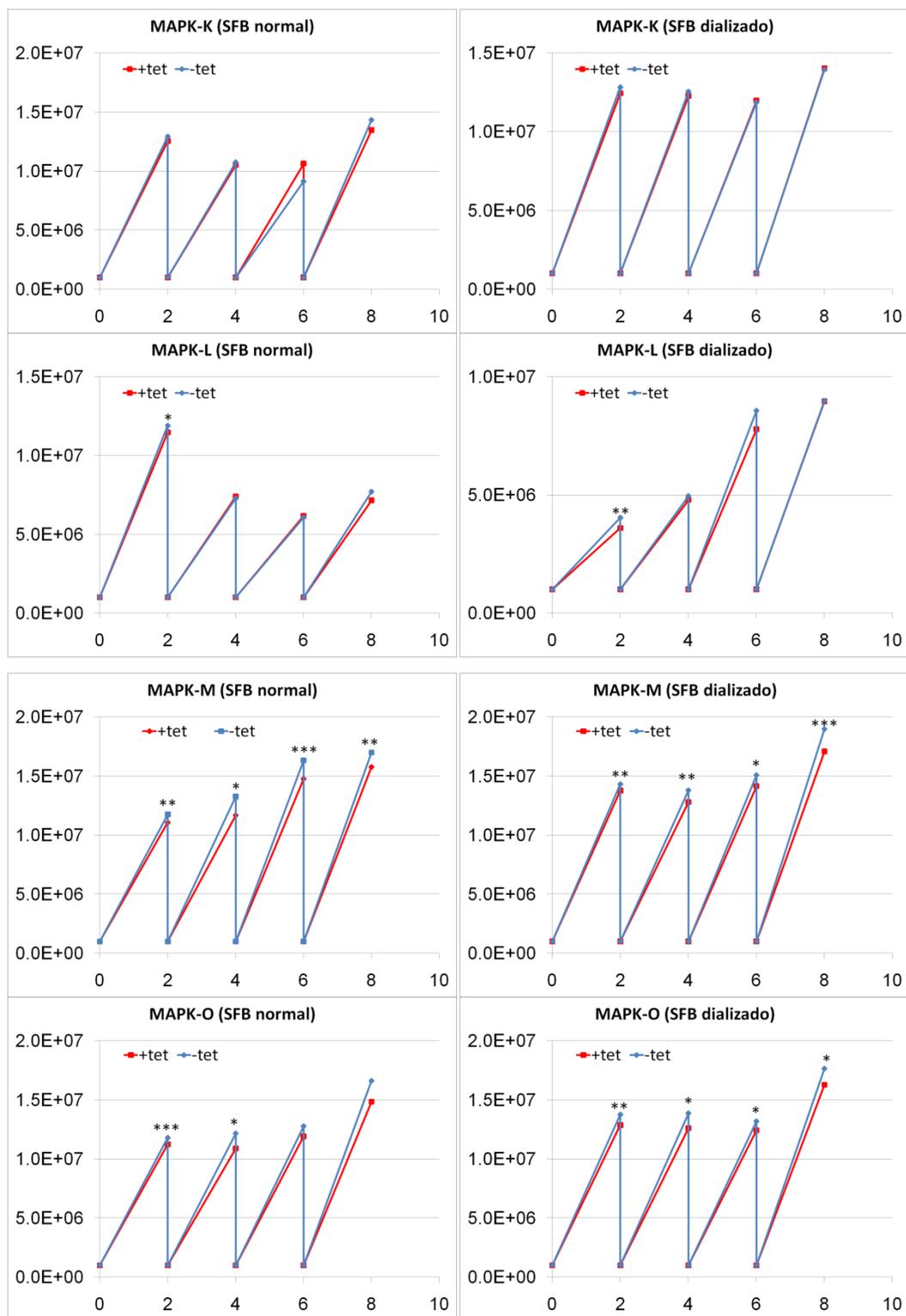


Figura 15: Proliferação celular pouco afetada após depleção das MAPKs. No eixo X estão representados os dias de cultivo, enquanto no eixo Y está o número de células por ml em log₂. * 0,01 < pvalue < 0,05; ** 0,001 < pvalue < 0,01; *** = pvalue < 0,001.

5.4 Avaliação do transcriptoma após depleção da MAPKs A

5.4.1 Avaliação geral do experimento de RNAseq

Este ensaio foi realizado na plataforma de sequenciamento de segunda geração (SOLiD, Applied Biosystems). O desenho experimental incluiu a comparação entre os transcriptomas das células não induzidas para RNAi com as respectivas células onde houve indução do RNAi e consequente depleção das quinase MAPK-A, MAPK-J e uma MAP2K (Tb927.8.5950), pertencente à família STE7-like de *T. brucei*, estudada em um trabalho paralelo [126].

A escolha destes *knockdowns* foi devido à influência da depleção dos mRNAs das respectivas proteíno-quinases na proliferação celular dos mesmos. O número de reads obtidos entre os três experimentos variou de cerca de 6.300.000 a 13.400.000. Para o mapeamento destes reads no genoma foram aplicados diferentes condições de alinhamento. Na condição unique, somente os reads mapeados em uma única região do genoma foram considerados, enquanto em random, os reads podem estar mapeados em mais de uma região do genoma. A condição unique exclui famílias multigênicas, já que estas estão presentes em múltiplas cópias no genoma. As condições 30 e 40, se referem ao tamanho mínimo do alinhamento obtido no genoma, sendo assim a condição unique 40 a de maior estringência. No presente trabalho, optou-se por analisar os reads mapeados na condição unique 30, fornecendo assim um nível intermediário de estringência. A proporção de reads mapeados no genoma de *T. brucei* 927 de acordo com o modelo de alinhamento aplicado está demonstrado na Tabela 8.

Tabela 8: Proporção dos reads mapeados por RNAseq.

Proporção dos reads mapeados										
		Sem correção					Com correção			
Tb927.10.7780 (MAPK-J)										
Amostra	Rép	Reads	Únique 40	Random 40	Únique 30	Random 30	Únique 40	Random 40	Únique 30	Random 30
Não Induzido	A	6868807.00	20.54%	34.47%	37.88%	61.97%	25.92%	43.59%	39.39%	65.53%
Induzido	A	9945740.00	14.47%	35.92%	28.21%	64.82%	18.79%	45.65%	29.58%	69.19%
Não Induzido	B	13383568.00	14.60%	23.55%	33.48%	53.36%	19.94%	32.44%	35.15%	57.11%
Induzido	B	8876903.00	20.00%	33.48%	37.30%	60.84%	25.43%	42.57%	38.90%	64.46%
Não Induzido	C	7755972.00	15.09%	25.48%	33.56%	55.60%	20.44%	34.65%	35.11%	59.36%
Induzido	C	7584083.00	13.56%	30.39%	28.07%	58.95%	17.78%	39.47%	29.55%	63.66%
Tb927.10.10870 (MAPK-A)										
Não Induzido	A	12060397.00	18.21%	30.82%	36.13%	59.03%	23.49%	39.71%	37.89%	62.91%
Induzido	A	10911055.00	19.28%	34.21%	35.97%	61.05%	24.61%	43.25%	37.63%	64.83%
Não Induzido	B	8708582.00	19.50%	31.53%	37.38%	58.69%	25.28%	40.63%	39.24%	62.56%
Induzido	B	11371126.00	20.51%	33.43%	38.60%	61.93%	26.16%	42.90%	40.25%	65.59%
Não Induzido	C	7516908.00	18.29%	30.39%	35.93%	57.74%	23.66%	39.24%	37.73%	61.68%
Induzido	C	8064015.00	17.45%	32.01%	33.47%	58.99%	22.37%	40.81%	35.12%	62.92%
Tb927.8.5950 (MAP2K)										
Não Induzido	A	11829056.00	20.27%	30.31%	39.76%	58.76%	26.17%	39.61%	41.61%	62.78%
Induzido	A	6304881.00	20.58%	29.33%	43.29%	61.13%	27.51%	39.46%	44.85%	64.36%
Não Induzido	B	6974000.00	20.56%	29.87%	40.77%	58.89%	27.00%	39.63%	42.71%	62.98%
Induzido	B	12307044.00	19.92%	28.25%	42.44%	59.87%	26.87%	38.36%	44.08%	63.21%
Não Induzido	C	6297042.00	20.81%	30.03%	42.59%	60.99%	27.38%	39.88%	44.40%	64.74%
Induzido	C	6874001.00	20.22%	28.95%	40.82%	58.28%	26.41%	38.34%	42.66%	62.11%

A análise por RNAseq foi realizada para três *knockdowns*, no entanto serão analisados somente os genes modulados após a depleção da MAPK-A. Esta escolha é justificada pela obtenção de dados de proteômica e fosfoproteômica somente para esta quinase. Desta forma será focado na análise em conjunto dos dados de RNA, proteína e sítios de fosforilação. O *knockdown* desta quinase apresentou nos dados de RNAseq uma diminuição de 52% no nível do mRNA correspondente a esta proteína, contra 20,4% para MAPK-J e 24,6% para MAP2K (Figura 16 e Figura 17). Na figura Figura 16 pode-se observar o nível de cada mRNA analisado nos três *knockdowns* sequenciados em três réplicas.

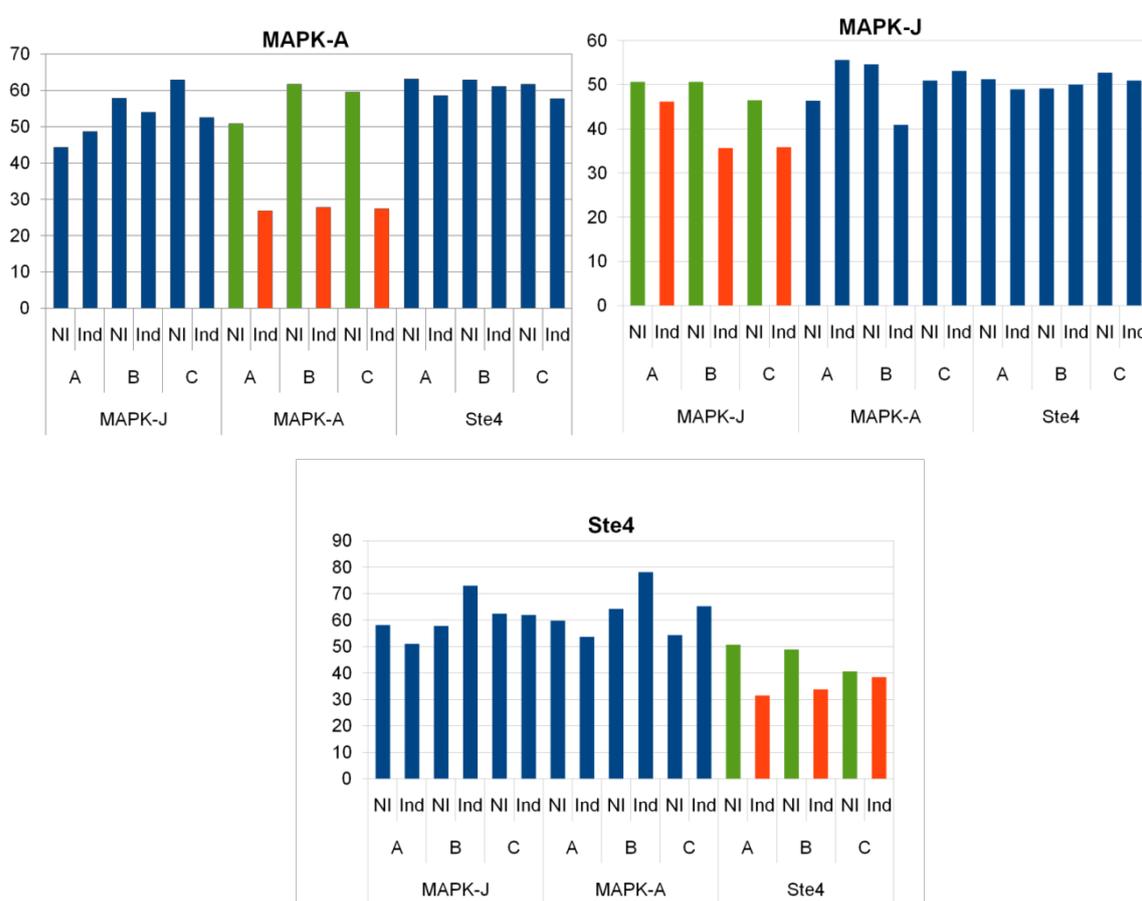


Figura 16: Comparação entre os níveis de mRNA de cada quinase nos *knockdowns*. NI – amostra não induzido, Ind – amostra induzida. No eixo Y está representado o número de reads mapeados. No eixo X estão representadas as diferentes réplicas biológicas (A, B e C) e os diferentes transfectantes (MAPK-J, MAPK-A e Ste4).

Com relação as modulações encontradas, avaliando-se os dados obtidos com unique30 versus unique40, optou-se por avaliar os genes modulados em unique30,

e FDR de 5% (Tabela 9). Poucos genes sofreram modulação na sua expressão após a depleção da MAPK-A. Foram encontrados 6 genes com expressão aumentada e 9 com expressão diminuída.

Tabela 9: Número de genes modulados após depleção da MAPK-A.

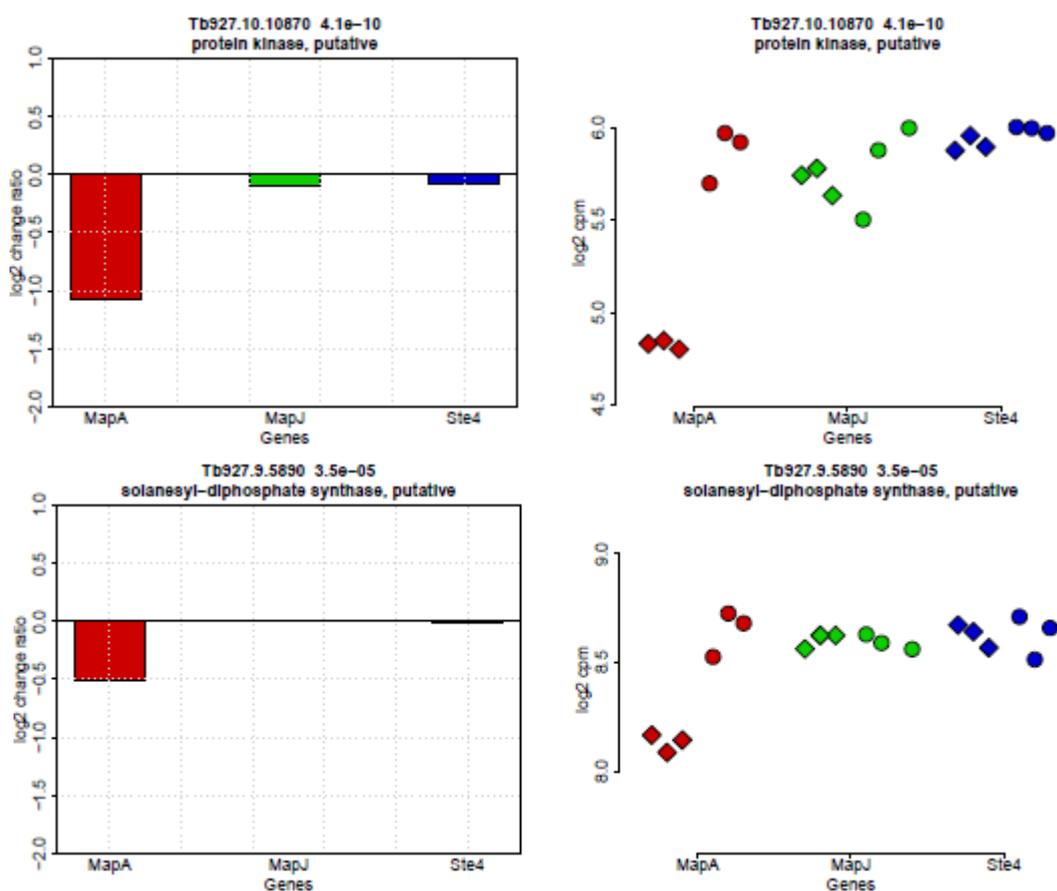
	FDR 20%		FDR 10%		FDR 5%		FDR 1%		FDR 0.1%	
	Up	Down	Up	Down	Up	Down	Up	Down	Up	Down
Unique30	15	16	7	12	6	9	3	6	1	2
Unique40	9	9	2	6	1	5	1	4	1	2

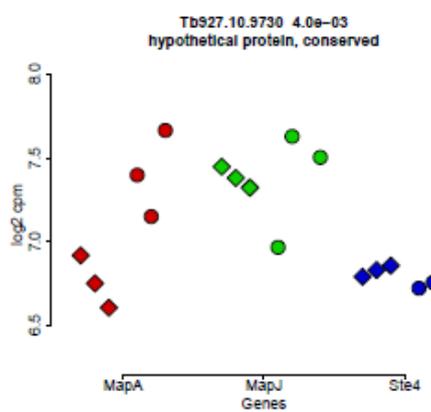
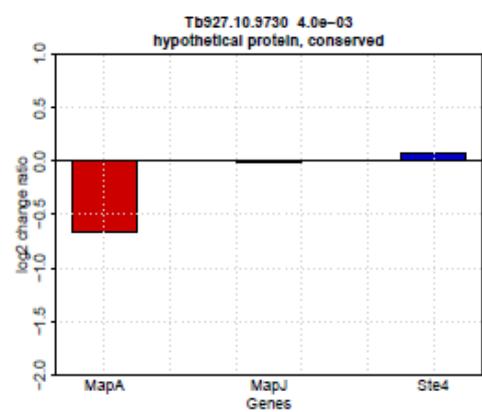
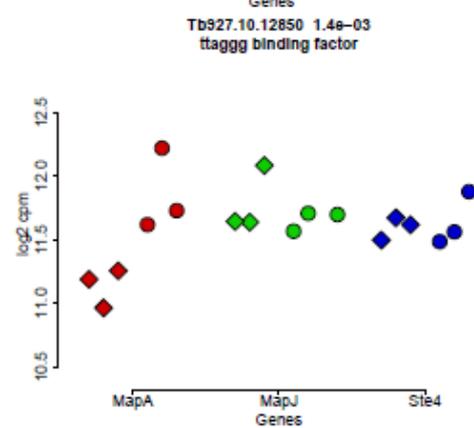
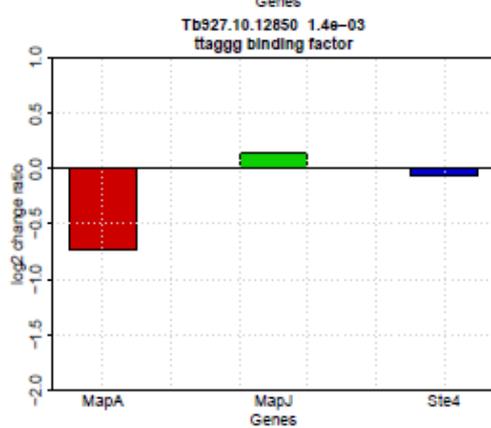
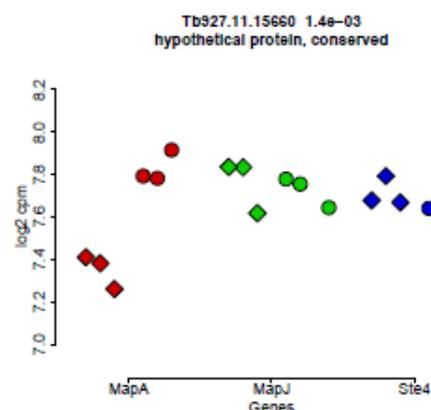
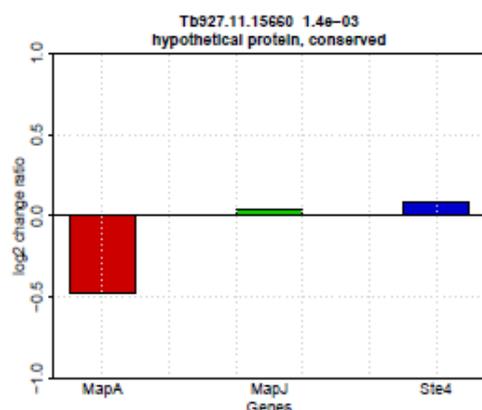
A análise em conjunto dos três *knockdowns* permite uma melhor visualização das modulações na expressão de alguns genes como foi o caso de seis genes que tiveram sua expressão diminuída no *knockdown* (Tb927.9.5890, Tb927.11.15660, Tb927.10.12850, Tb927.7.610, Tb927.1.2820). Comparando-se a amostra induzida com a amostra não induzida e também com os outros *knockdowns* que apresentaram uma expressão constante destes genes, foi observada uma discreta diminuição (*fold change* aproximado de 1,5) (Figura 17). Analisando as características destes genes não existe uma relação clara de função entre os mesmos. O gene com maior diminuição foi a própria MAPK-A, com um *fold change* próximo de 2. Entre os outros está uma proteína *solaneyl-diphosphate synthase, putative*. Esta proteína está relacionada com a via de síntese de ubiquinona, a qual atua no transporte de elétrons para produção de ATP na célula. Ensaios de imunofluorescência e de fusão com tag GFP e TY1 localizaram a *solaneyl-diphosphate synthase* na mitocôndria de *T. brucei* [94]. Esta proteína foi localizada no glicosomo em *T. cruzi* [95] e possui diferentes localizações em eucariotos, como retículo endoplasmático e plastídeos em *A. thaliana*. Outro mRNA entre os modulados é o Fator ligador de ttaggg, que possui homólogos em *T. cruzi* e *Leishmania major*, possuindo similaridade com TRF2 de vertebrados. Este fator é essencial para estabilização dos telômeros (*G-overhang*) [96].

Ainda entre os genes com expressão diminuída está uma *mitochondrial DNA ligase homolog, LIG k-alpha*, localizada no cinetoplasto de *T. brucei*, tendo sua função relacionada com a ligação dos minicírculos antes da sua divisão. A depleção desta proteína causa parada na proliferação celular [97]. A forte inibição da

proliferação celular após a depleção da MAPK-A poderia estar associada com a diminuição desta ligase, fato que merece ser melhor investigado. Uma proteína, *pteridine transporter, putative*, com um domínio da família BT1, presente em um transportador de biopterina de *Leishmania*, cujo nocaute leva à diminuição da infectividade neste parasita [98] também foi encontrada entre os genes modulados.

Os outros três genes no grupo dos genes com expressão diminuída (proteínas hipotéticas), apesar de apresentarem diferenças entre induzido e não induzido, não apresentaram um padrão claro de diminuição em comparação com Ste4 e MAPK-J. A busca por domínios no banco de dados Pfam não retornou nenhum domínio significativo para estas proteínas.





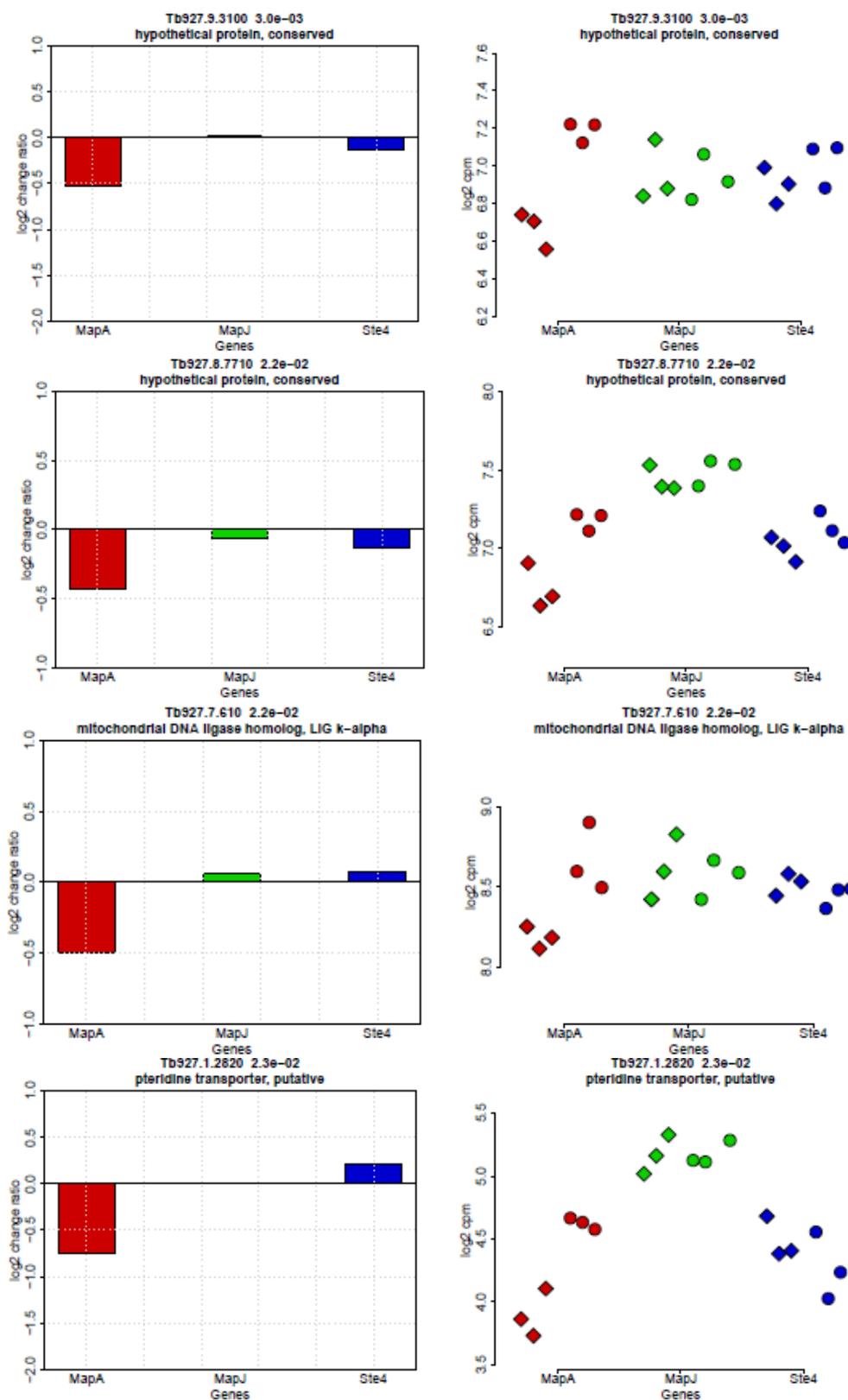
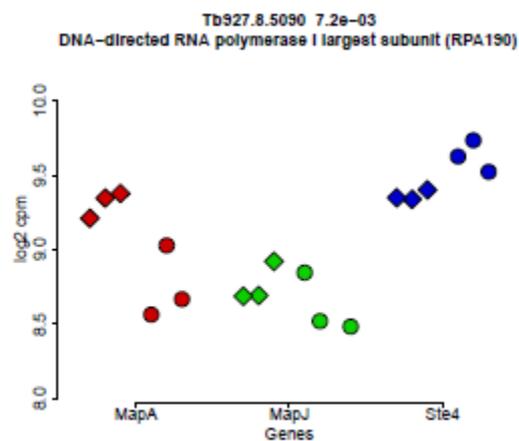
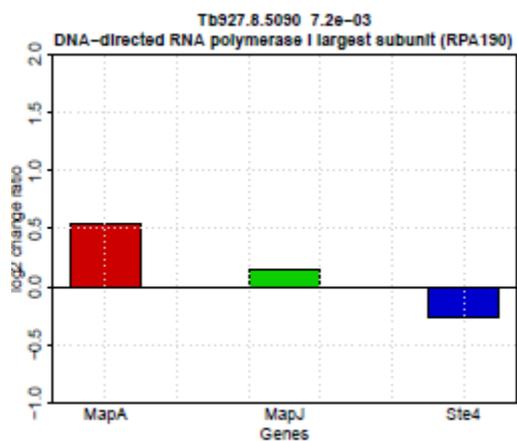
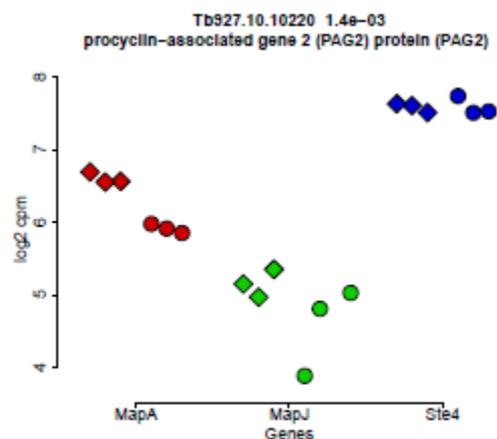
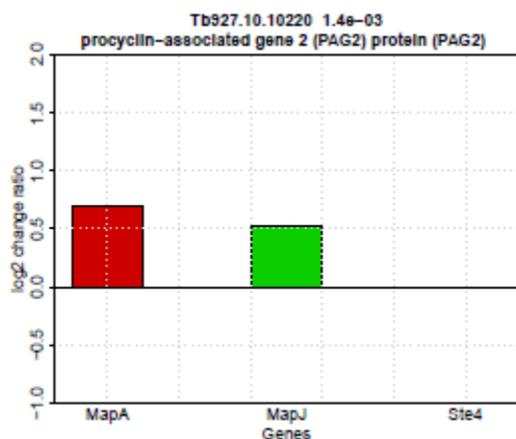
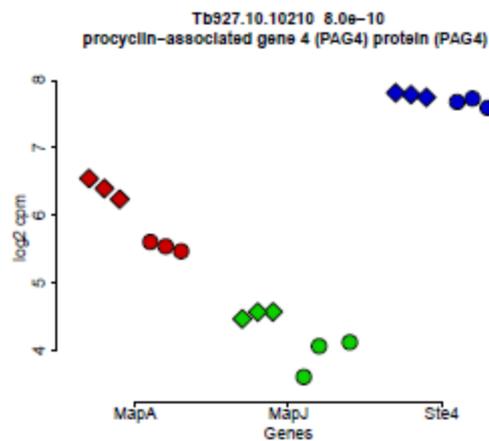
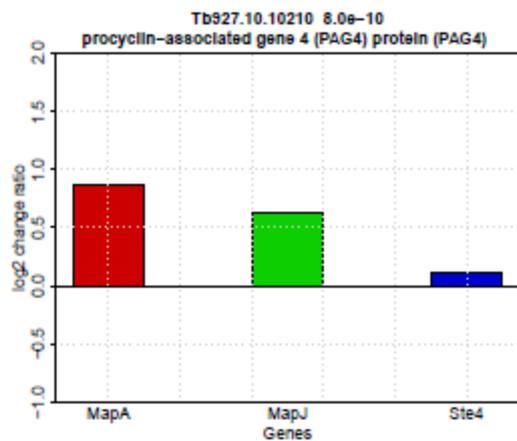


Figura 17: Genes diminuídos após depleção da MAPK-A. Os losangos representam as amostras com indução do RNAi e os círculos representam as amostras não induzidas. Vermelho – MAPK-A, verde – MAPK-J e azul – Ste4.

Dentre os seis genes aumentados, quatro deles pertencem à família dos genes associados às prociclinas (PAGs). Esta família de genes está localizada em arranjos de repetição em tandem dos genes das prociclinas, que são glicoproteínas de superfície características das formas procíclicas de *T. brucei*. Interessantemente, o padrão de modulação observado para estes genes parece ser dependente do transfectante analisado, sendo diferente entre os transfectantes de MAPK-A, MAPK-J e Ste4 e não variando entre induzido e não induzido (Figura 18).

Entre os genes aumentados, uma proteína anotada como *leucine-rich repeat protein*, possui domínios LRR sem significância resultantes da busca no banco de dados do Pfam. Uma *DNA-directed RNA polymerase I largest subunit* (RPA190), detectada por proteômica na fração glicossomal de *T. brucei* forma sanguínea, de acordo com a literatura [99], também está entre os genes aumentados.



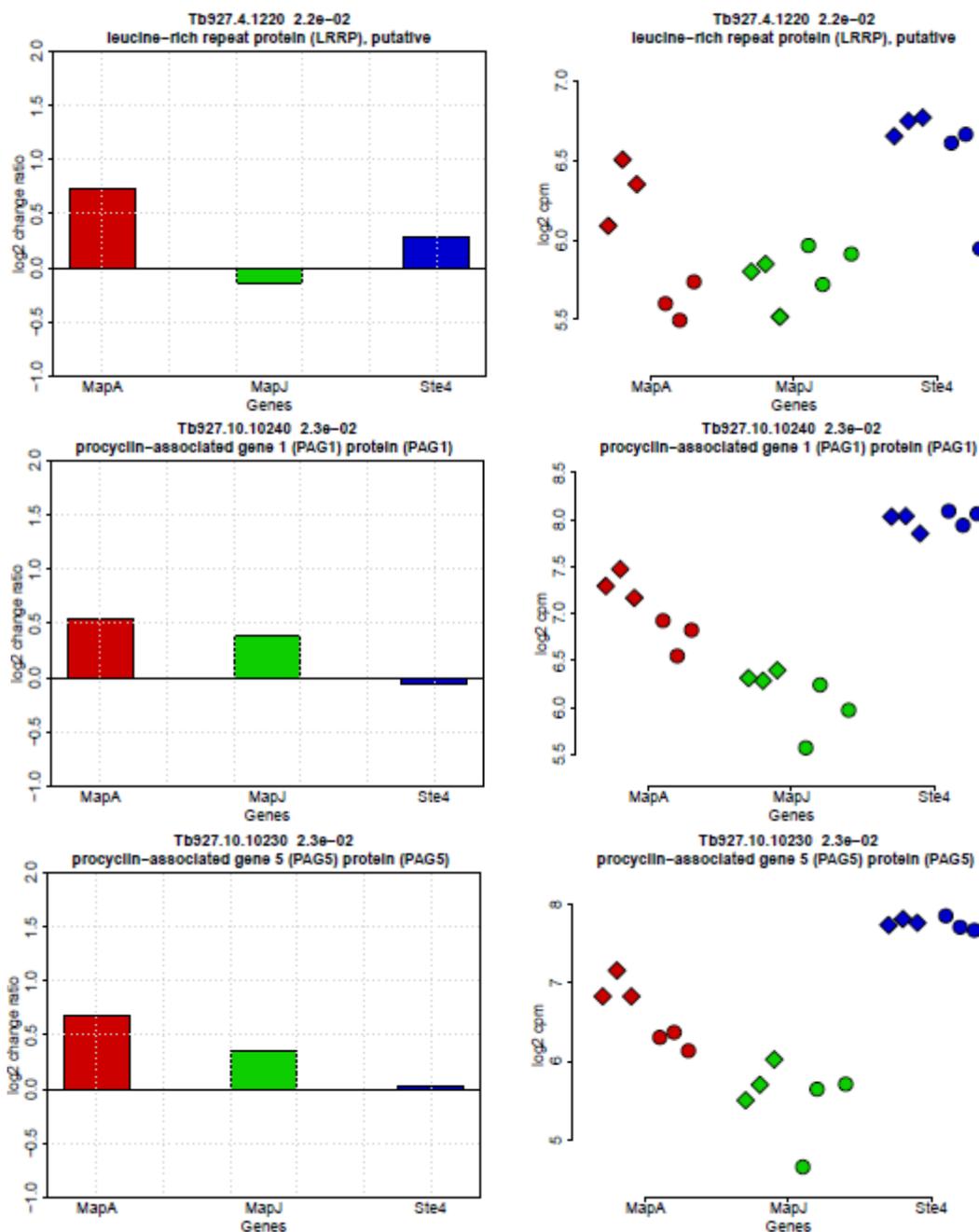


Figura 18: Genes com expressão aumentada após depleção da MAPK-A. Os losangos representam as amostras com indução do RNAi e os círculos representam as amostras não induzidas. Vermelho – MAPK-A, verde – MAPK-J e azul – Ste4.

Mesmo com a forte inibição da proliferação celular observada após a depleção da MAPK-A, foram encontrados poucos genes modulados em baixa magnitude na análise por RNA-seq. Isto pode em parte ser reflexo do mecanismo de regulação da expressão gênica predominantemente pós-transcricional em *T. brucei*. Apesar de existirem mecanismos de regulação da expressão por degradação de

mRNAs, o que poderia causar modulações detectáveis por RNAseq, a estocagem de mRNA e a associação ou não a polissomos também poderia contribuir com tais modulações, no entanto podendo dificultar a detecção por metodologias que analisam o mRNA total da célula, como é o caso deste trabalho.

Também não se pode deixar de especular a participação da MAPK-A apenas em processos de fosforilação ou interações proteicas, cujo resultado direto não estaria relacionado à modulação dos níveis de mRNA.

5.5 Análise Proteômica

O objetivo inicial deste trabalho era analisar quantitativamente o proteoma e fosfoproteoma de *T. brucei* após a depleção de cada uma das 15 proteínas classificadas como MAPK [8]. No entanto, por questões de prazo, visto que as padronizações feitas neste trabalho foram longas, optou-se por analisar apenas o *knockdown* de uma das MAPKs. A escolha foi baseada nos dados das curvas de crescimento e na quantificação dos mRNAs após indução do RNAi. A depleção dos mRNAs da MAPK-A (identificação interna), com Id Tb927.10.10870, mostrou uma forte influência na proliferação celular, praticamente inibindo o crescimento após 8 dias de indução do RNAi (Figura 14). A diminuição nos níveis de mRNA foi de mais de 80% no quarto dia de indução. Estes resultados foram motivadores para a escolha desta quinase, visto que ela poderia estar ativada nas formas celulares estudadas e que parecia não haver uma sobreposição de função na sua ausência.

5.5.1 Fracionamento da amostra

Neste trabalho, obteve-se o extrato proteico de células de *T. brucei*, forma procíclica, as quais tiveram depleção por RNAi do mRNA de uma MAPK (Tb927.10.10870) deste parasita. Este extrato foi digerido com tripsina e os peptídeos originados foram pré-fracionados a fim de se diminuir a complexidade da amostra, podendo-se assim aumentar o número de proteínas identificadas. A etapa de pré-fracionamento de proteínas ou peptídeos vem sendo feita em diversos trabalhos que analisam extrato proteico total [9]. Neste trabalho, o extrato proteico analisado possui complexidade aumentada devido à marcação metabólica por

SILAC. Células marcadas metabolicamente foram misturadas em igual proporção com células não marcadas, sendo assim, cada proteína presente no extrato terá sua versão marcada (pesada) e não marcada (leve).

A fim de pré-fracionar os peptídeos ou proteínas foram testados diferentes métodos, tais como SDS-PAGE, *strong cation exchange* (SCX) e *high pH reverse phase chromatography* (RPHP). O fracionamento por SDS-PAGE se mostrou um bom método com a identificação de 1.310 proteínas em um experimento em paralelo usando-se SILAC para comparação entre células de *T. brucei* cultivadas em meio contendo soro dializado ou não dializado. Alguns limitantes para o uso do SDS-PAGE são o aumento no custo e no tempo das etapas de bancada, além da impossibilidade de fracionamento de grandes quantidades de proteína, por exemplo alguns miligramas, quantidade geralmente usada nos experimentos de fosfoproteômica. O pré-fracionamento por troca iônica tem sido amplamente utilizado em análises proteômicas e fosfoproteômicas. A combinação de espectrômetros de massa de última geração com sistemas cromatográficos de alta resolução tem possibilitado a identificação de mais de 8.400 proteínas em apenas um dia de análise por LC-MS/MS [100]. Apesar da boa eficiência do SCX, no presente trabalho conseguiu-se mais rapidamente a padronização do pré-fracionamento por RPHP, o qual possui a vantagem de não requerer etapa preliminar para remoção de sal, como é o caso do SCX. A metodologia de RPHP combinada com fase reversa (RP) em pH ácido tem se demonstrado uma boa opção para a separação 2D [101] [102].

O pré-fracionamento dos peptídeos foi realizado no modo *off line* por cromatografia reversa em pH básico segundo a metodologia demonstrada por Gilar e colaboradores [101]. O número de peptídeos identificados por fração está representado na Tabela 10. Foi possível observar maior sobreposição dos peptídeos presentes em frações adjacentes em relação às frações mais distantes dentro do gradiente de eluição. Este fato pode ser devido à eluição ineficiente em cada fração ou à baixa ortogonalidade da separação. Uma melhora na ortogonalidade da separação combinando RPHP e RP em pH ácido foi demonstrada por Yang e colaboradores, através de pré-fracionamento em múltiplas frações na primeira dimensão da cromatografia, seguido de concatenação de frações não adjacentes para a segunda dimensão da cromatografia [102].

Tabela 10: Número de peptídeos identificados na proteômica nas diferentes frações do pré-fracionamento.

Fração	Peptídeos	Fração	Peptídeos
ACN10	1099	ACN14	828
ACN10+ACN14	707	ACN14+ACN18	514
ACN10+ACN14+ACN18	159	ACN14+ACN18+ACN26	249
ACN10+ACN14+ACN18+ACN26	35	ACN14+ACN18+ACN26+ACN40	16
ACN10+ACN14+ACN18+ACN26+ACN40	4	ACN14+ACN18+ACN40	3
ACN10+ACN14+ACN18+ACN40	9	ACN14+ACN26	6
ACN10+ACN14+ACN26	0	ACN14+ACN26+ACN40	1
ACN10+ACN14+ACN26+ACN40	0	ACN14+ACN40	2
ACN10+ACN14+ACN40	5	ACN18	597
ACN10+ACN18	0	ACN18+ACN26	569
ACN10+ACN18+ACN26	0	ACN18+ACN26+ACN40	42
ACN10+ACN18+ACN26+ACN40	0	ACN18+ACN40	1
ACN10+ACN18+ACN40	0	ACN26	1193
ACN10+ACN26	1	ACN26+ACN40	110
ACN10+ACN26+ACN40	0	ACN40	82
ACN10+ACN40	1		

Os peptídeos identificados nas frações únicas e adjacentes estão sombreados.

5.5.2 Perfil geral do proteoma

Fazendo-se a correlação de Pearson das relações de intensidade NI/I entre as réplicas 1 e 2, foi obtido um valor de 0,48 para as 1.270 proteínas presentes em ambas as réplicas. Esta baixa correlação sobe para 0,82 quando se exclui 104 proteínas com valor de p maior que 0,05 (teste de Benjamini Hochberg, FDR 0,05). Das 104 proteínas excluídas, 70 foram quantificadas apenas por um peptídeo, enquanto na fração com p menor ou igual a 0,05, 301 das 1.166 proteínas foram quantificadas apenas por um peptídeo. A quantificação de proteínas por um único peptídeo poderia justificar a maior variabilidade entre as réplicas na quantificação de algumas proteínas. Outro fator poderia ser a quantificação de proteínas por peptídeos compartilhados por mais de uma proteína, mas que não foi o caso das proteínas aqui quantificadas.

Semelhante ao descrito acima, na réplica 1 versus 3, a correlação de 0,54 (811 proteínas) alterou-se para 0,78 (698 proteínas) com a exclusão de 113 proteínas pelo mesmo critério citado acima, com 60 destas proteínas sendo quantificadas por apenas um peptídeo. A correlação entre as réplicas 2 e 3 subiu de 0,61 (817 proteínas) para 0,80 (718 proteínas) após a exclusão de 99 proteínas, sendo 62 destas quantificadas apenas por um peptídeo. Na Figura 19, pode-se notar a baixa modulação na expressão proteica, com raras proteínas chegando próximas à relação de +/- 1 (\log_2).

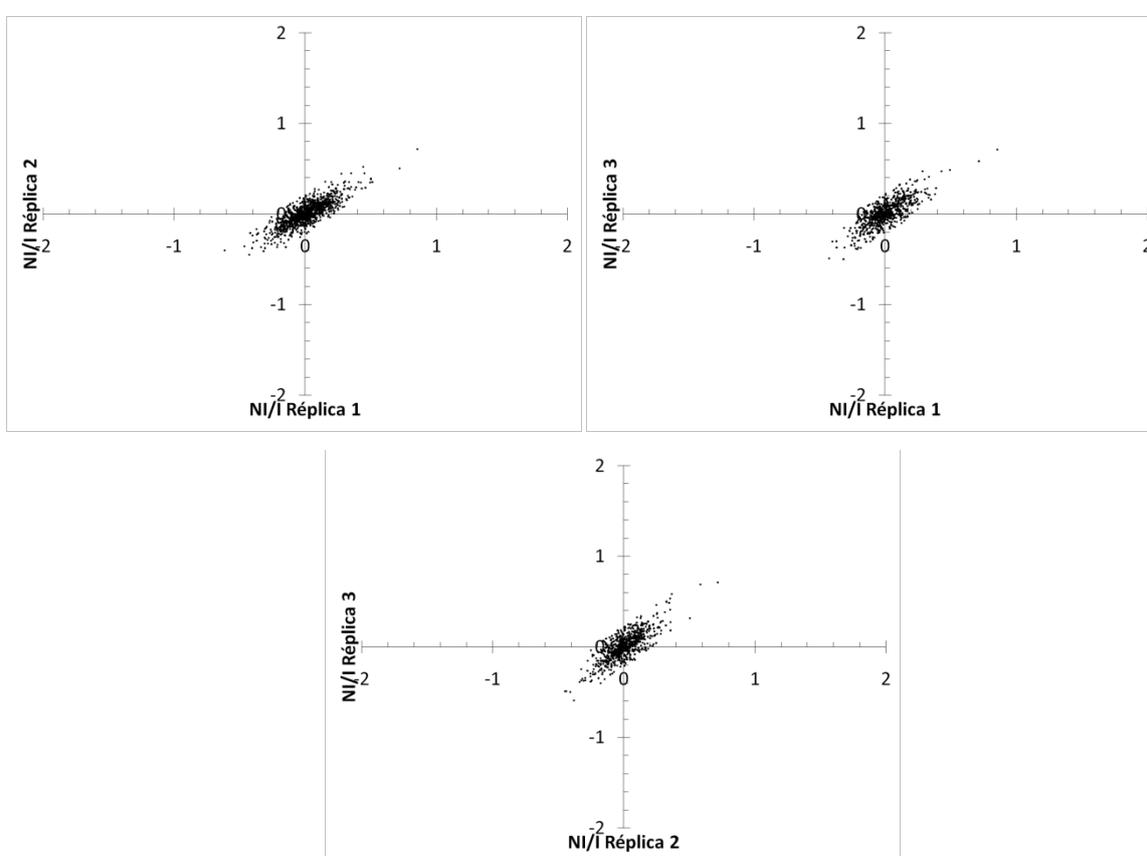


Figura 19: Correlação entre as intensidades NI/I das proteínas quantificadas entre as réplicas biológicas. Eixos x e y \log_2 , NI – não induzido, I – induzido.

A plotagem das intensidades das proteínas nas diferentes réplicas analisadas mostrou maior correlação entre as réplicas 1 e 2 (Pearson = 0,96) (Figura 20). Quando se combina a réplica 1 com a 3 (0,87) ou a 2 com a réplica 3 (0,88), há uma menor correlação das intensidades. As relações das intensidades NI/I tem menor variação quando analisada a réplica 3 contra 1 ou 2 em comparação com as

correlações das respectivas intensidades. Apesar dessa flutuação, as relações NI/I se mantiveram.

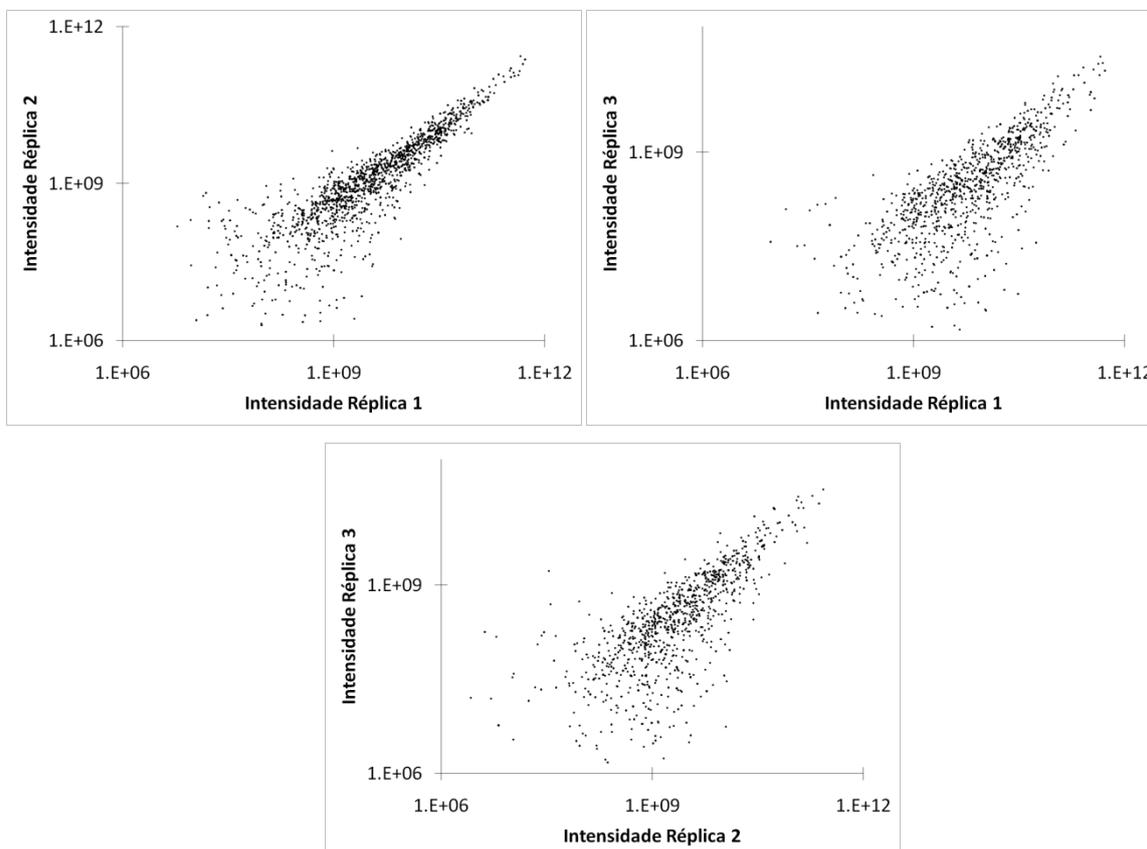


Figura 20: Relações entre as intensidades das proteínas no ensaio de proteômica. Os eixos estão em escala logarítmica.

Na análise proteômica foram identificadas 1.511 proteínas, das quais 1.498 foram quantificadas. A Tabela 11 detalha o número de peptídeos e proteínas quantificados únicos dentro de cada réplica biológica ao longo do fracionamento. Entre as 3 réplicas, na terceira foram quantificados menos peptídeos e menos proteínas, enquanto as duas primeiras apresentaram números maiores de peptídeos e proteínas quantificados. No último passo de eluição do pré-fracionamento por RPHP (ACN40) foram identificados poucos peptídeos e proteínas, sendo que a porcentagem de peptídeos únicos por réplica não passou de 1.3% nas réplicas 1 e 2, demonstrando que a grande maioria dos peptídeos são eluídos com 26% de

acetonitrilo (ACN26). Na réplica 3, o passo ACN40 representou 12,7% devido ao baixo número de peptídeos identificados nos outros passos desta réplica.

Tabela 11: Número de peptídeos e proteínas únicos por réplica quantificados na proteômica.

Fração	Peptídeos Quantificados						Proteínas Quantificadas					
	Réplica 1		Réplica 2		Réplica 3		Réplica 1		Réplica 2		Réplica 3	
ACN10	2020	32.4%	2287	28.8%	1612	82.8%	820	61.4%	900	69.9%	718	92.9%
ACN14	1619	26.0%	1119	14.1%	81	4.2%	292	21.9%	165	12.8%	26	3.4%
ACN18	1209	19.4%	1215	15.3%	5	0.3%	131	9.8%	126	9.8%	3	0.4%
ACN26	1303	20.9%	1099	13.8%	1	0.1%	77	5.8%	96	7.5%	0	0.0%
ACN40	82	1.3%	15	0.26%	248	12.7%	16	1.2%	0	0.0%	26	3.4%
Total por Réplica	6233	100%	5735	100%	1947	100%	1336	100.0%	1287	100.0%	773	100.0%
Réplicas 1 e 2	3924						1159					
Réplicas 1 e 3	1563						743					
Réplicas 2 e 3	1439						738					
Réplicas 1, 2, 3	1314						719					
Total	8294						1498					

5.5.3 Proteínas com expressão modulada

Cada réplica analisada foi originada da mistura de células marcadas metabolicamente com células não marcadas. A cultura não marcada foi induzida com tetraciclina (amostra I) para depleção do mRNA alvo, enquanto a cultura marcada não foi induzida (amostra NI) e foi utilizada como controle. Analisando-se no geral as relações de intensidade I/NI foi possível observar pouca modulação. Apenas 16 proteínas tiveram modulação com *fold-change* igual ou superior a 1,25 x (+/- 0,32 em \log_2) para mais ou para menos, com valor de *p* menor ou igual a 0,05 (test-t), sendo que a maior modulação foi uma diminuição de 1,7 x na expressão da enzima aconitase (Tb927.10.14000) no *knockdown*. Além do *fold-change* e do valor de *p*, outro critério usado para seleção destas proteínas foi a quantificação das mesmas em pelo menos duas réplicas biológicas (Tabela 12).

Tabela 12: Proteínas com moduladas na proteômica.

Id	Proteína	I/NI mínimo 2 réplicas	t-test p value 1;2;3
Tb927.10.14000	aconitase	-0.76	0.004
Tb927.2.280	retrotransposon hot spot protein 2 (RHS2), putative	-0.56	0.033
Tb927.10.13430	citrate synthase, putative	-0.51	0.047
Tb927.9.2520	microtubule- associated protein	-0.47	0.037
Tb927.9.6560	protein kinase, putative	-0.45	0.002
Tb927.1.190	retrotransposon hot spot protein (RHS, pseudogene), putative	-0.44	0.011
Tb927.7.5480	dihydrofolate reductase-thymidylate synthase	-0.37	0.004
Tb927.10.14090	transporter, putative	-0.33	0.006
Tb927.7.6270	peptidase t, putative, aminotripeptidase, putative	-0.33	0.046
Tb927.11.12820	ribonucleoside-diphosphate reductase small chain	-0.33	0.010
Tb927.10.7570	dihydrolipoamide acetyltransferase E2 subunit, putative	0.33	0.003
Tb927.10.6950	sterol 24-c-methyltransferase, putative	0.34	0.005
Tb927.3.3890	hypothetical protein, conserved	0.35	0.023
Tb927.7.1790	Adenine phosphoribosyltransferase, putative	0.41	0.017
Tb927.11.2690	succinyl-coA:3-ketoacid-coenzyme A transferase, mitochondrial precursor, putative	0.43	0.038
Tb927.6.2790	L-threonine 3-dehydrogenase, putative	0.45	0.002

Ids em negrito correspondem às proteínas quantificadas nas 3 réplicas biológicas. I/NI = relação da média entre as intensidades das proteínas na amostra induzida e não-induzida, valor em log₂. Valor de P obtido após aplicação de teste t.

Dentre as proteínas moduladas, 11 foram quantificadas nas 3 réplicas analisadas e 5 delas somente em 2 réplicas. A análise entre as réplicas demonstrou a robustez da metodologia SILAC, visto que aproximadamente apenas 1% das proteínas quantificadas (16) tiveram modulação igual ou superior a 25% para mais ou para menos, Tabela 12. Interessantemente, a forte influência da MAPK-A na proliferação celular, observada nas curvas de crescimento é acompanhada de uma baixa modulação na expressão proteica, pelo menos nas proteínas quantificadas no presente estudo. Um resultado semelhante foi observado em uma análise proteômica e fosfoproteômica em células de *Drosophila* após depleção de uma proteína fosfatase [77]. Neste trabalho, das 3.434 proteínas quantificadas, apenas 22 estavam moduladas pelo menos 1,5 x.

Dentre as 1.498 proteínas quantificadas aqui, 540 (36%) são hipotéticas ou hipotéticas conservadas, uma proporção aproximada aos 43,9% encontrados em uma análise de proteômica quantitativa durante a metacicloênese em *T. cruzi* [83].

A análise de enriquecimento de GO foi feita usando-se como entrada as 16 proteínas classificadas como moduladas neste experimento. Destas proteínas, foi encontrada anotação para apenas sete entradas. Os termos retornados desta busca não apresentaram e-value significativo, ficando acima de 0,05. O baixo número de entradas contribuiu para tal resultado, sendo que a associação significativa a algum termo, com um baixo número de entradas, dependeria de um enriquecimento maior que o observado de 14,3%.

5.5.4 Proteínas diminuídas após depleção da MAPK-A

Dentre as proteínas moduladas existe uma proteína hipotética, para a qual buscou-se por domínios proteicos anotados no banco de dados Pfam, Prosite e Interpro, no entanto, para esta proteína não foram encontrados domínios significativos.

Entre as proteínas diminuídas, esta um aconitase, pertencente à subfamília IRP-1, cuja proteína ortóloga em mamíferos se localiza no citoplasma e atua como sensor de ferro e regulador pós-transcricional e como transdutor de sinal no estresse oxidativo. Em *T. brucei*, esta enzima foi localizada 30% na mitocôndria, onde sugere-se que participe no ciclo do ácido cítrico e 70% no citoplasma, onde sua função é desconhecida. Esta enzima é diferencialmente expressa durante o ciclo de vida de *T. brucei*, no entanto com pequenas modulações no nível de mRNA, sugerindo-se uma modulação pós transcricional [103]. Outra enzima entre as moduladas é a *citrate synthase, putative*, a qual teve sua localização inferida como mitocondrial por análise de sequência [104].

Foram encontradas duas proteínas associadas à síntese de nucleotídeos, *dihydrofolate reductase-thymidylate synthase* e *ribonucleoside-diphosphate reductase small chain*. A primeira participa da via de síntese de timidilato, necessária para a síntese de ácidos nucleicos. Células duplo nocaute para este gene não sobrevivem na ausência de timidina, sendo esta estudada como potencial alvo terapêutico [105]. A segunda proteína está anotada como *ribonucleoside-diphosphate reductase small chain*. Juntamente com a subunidade maior, participa da redução de nucleotídeos para o respectivo desoxirribonucleotídeo.

Um elemento de transposição foi encontrado modulado, *retrotransposon hot spot protein 2 (RHS2), putative*. As proteínas encontradas anotadas como RHS pertencem a uma família multigênica associada com elementos móveis no genoma. A localização subcelular desta família foi evidenciada no núcleo e na região perinuclear [106]. Outra classe de proteína encontrada está relacionada com o citoesqueleto, *microtubule-associated protein*. Esta proteína possui repetições na sua sequência, sugerindo-se a interação destas repetições com os microtúbulos, e foi localizada na parte posterior do parasita *T. brucei*. Sua função pode estar relacionada com a elongação dos microtúbulos ou com a ligação destes com a membrana celular [107]. Um transportador (*transporter, putative*) foi encontrado com 25% de diminuição. Este possui um domínio MFS (*Major Facilitator Superfamily*). Esta família está envolvida com o transporte de pequenas moléculas em gradientes quimiosmóticos de íons, tais como metabólitos, íons, oligossacarídeos, drogas, inositois entre outros [108].

Com igual modulação, uma *peptidase t, putative, aminotripeptidase, putative*, foi selecionada entre as proteínas diferencialmente expressas. Esta proteína possui um domínio peptidase M20, relacionado à glutamato carboxipeptidases. As peptidases atuam clivando ligações peptídicas nas proteínas. Suas funções estão relacionadas com fator de virulência e interação parasita – hospedeiro em *T. cruzi* [127]. Apenas uma proteíno-quinase, *putative*, está entre as proteínas moduladas detectadas na proteômica. Esta proteína foi classificada no trabalho de Parsons e colaboradores como Outras, sendo provavelmente não catalítica [8].

5.5.5 Proteínas aumentadas após a depleção da MAPK-A

As proteínas selecionadas como aumentadas em no mínimo 25% pertencem a vias metabólicas, com exceção de uma proteína hipotética, para a qual não foram encontrados domínios conservados por busca no banco Pfam, Interpro e Prosite, realizada neste trabalho. Uma subunidade proteica, *dihydrolipoamide acetyltransferase E2 subunit, putative*, associada com o complexo mitocondrial piruvato desidrogenase, que participa da descarboxilação oxidativa de substratos do ácido oxo-2 a CO₂ e derivados de acetil-CoA [109] foi encontrada aumentada, tendo uma relação I/NI de 0,33 (log₂).

As proteínas *sterol 24-c-methyltransferase, putative, Adenine phosphoribosyltransferase, putative, succinyl-coA:3-ketoacid-coenzyme A transferase, mitochondrial precursor, putative* e *L-threonine 3-dehydrogenase, putative* completam a lista das proteínas aumentadas no *knockdown*, as duas últimas tendo sua localização celular atribuída à mitocôndria por proteômica e por análise computacional de sequência, respectivamente [110] [104].

Em uma análise geral da modulação observada no proteoma de *T. brucei* após depleção da MAPK-A, pouca modulação foi evidenciada. É possível que modulações mais importantes e, conseqüentemente, responsáveis pelo fenótipo de inibição da proliferação celular observado, estejam presentes entre as proteínas menos abundantes na célula e não detectadas neste experimento, como pode ser o caso da própria MAPK-A. A proteômica associada à espectrometria de massas é capaz de “enxergar” principalmente as proteínas mais abundantes da célula. Os avanços que vem ocorrendo na área resultaram em um aumento do número de identificações, já que houve um incremento de sensibilidade causado por melhorias nos equipamentos e na resolução da cromatografia. Entre as modulações observadas, um terço é representado por proteínas descritas na literatura como mitocondriais. Proteínas da via das MAPKs (ERK1/ERK2, MEK e Raf) têm sido relacionadas com função mitocondriais em resposta a estímulos, tendo papel no processo de apoptose [111]. Foi evidenciada a translocação da MAP3K Raf-1 para a mitocôndria quando fosforilada em resíduos específicos [112]. Dados como a localização celular, análise de complexos proteicos e ensaios bioquímicos poderiam contribuir para o endereçamento desta observação.

As proteínas correspondentes aos genes modulados na análise de RNAseq não foram identificadas na proteômica, não sendo possível uma comparação, destas modulações, entre proteína e mRNA. A distribuição dos mRNAs conforme a abundância está representada na Figura 21. Tal correlação se altera quando são plotados somente os mRNAs correspondentes às proteínas identificadas na proteômica. As proteínas identificadas na proteômica se distribuem entre mRNAs com diferentes níveis de expressão, com uma maior representatividade entre os mRNAs mais abundantes (Figura 21).

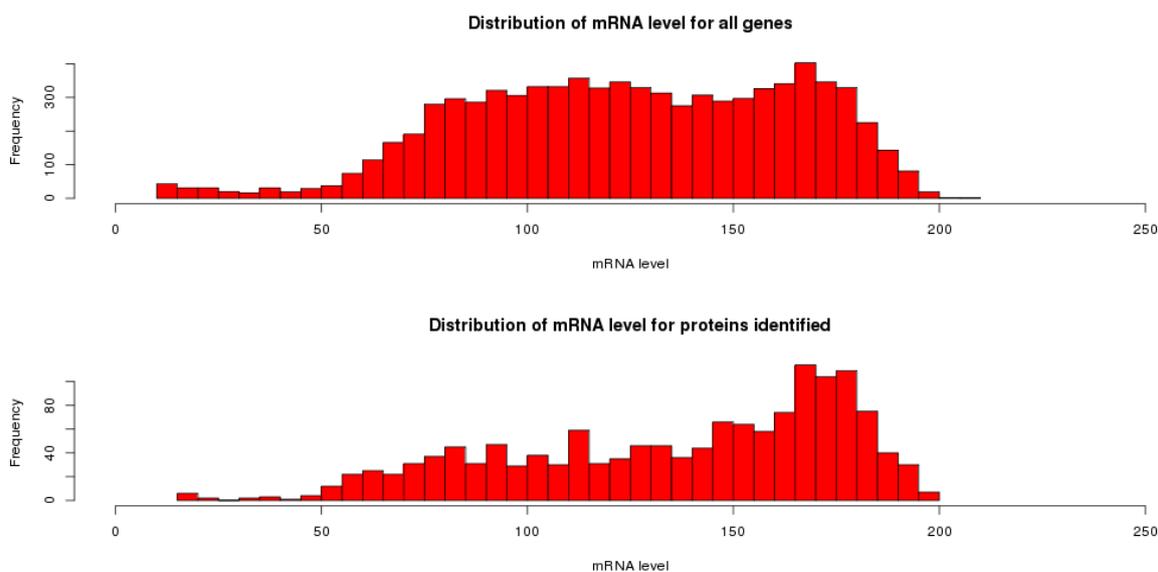


Figura 21: Correlação geral entre a proteômica e o RNAseq. No eixo X está o número de reads mapeados e no eixo Y o número de mRNAs.

5.6 Análise Fosfoproteômica

5.6.1 Perfil das análises

O extrato proteico obtido para a análise proteômica também foi utilizado para a análise fosfoproteômica. Após o fracionamento dos peptídeos, 90% da amostra foi direcionada para fosfoproteômica. A correlação de Pearson entre as réplicas biológicas foi de 0,75, 0,73 e 0,81, réplica 1 versus 2, 1 versus 3 e 2 versus 3, respectivamente (Figura 22).

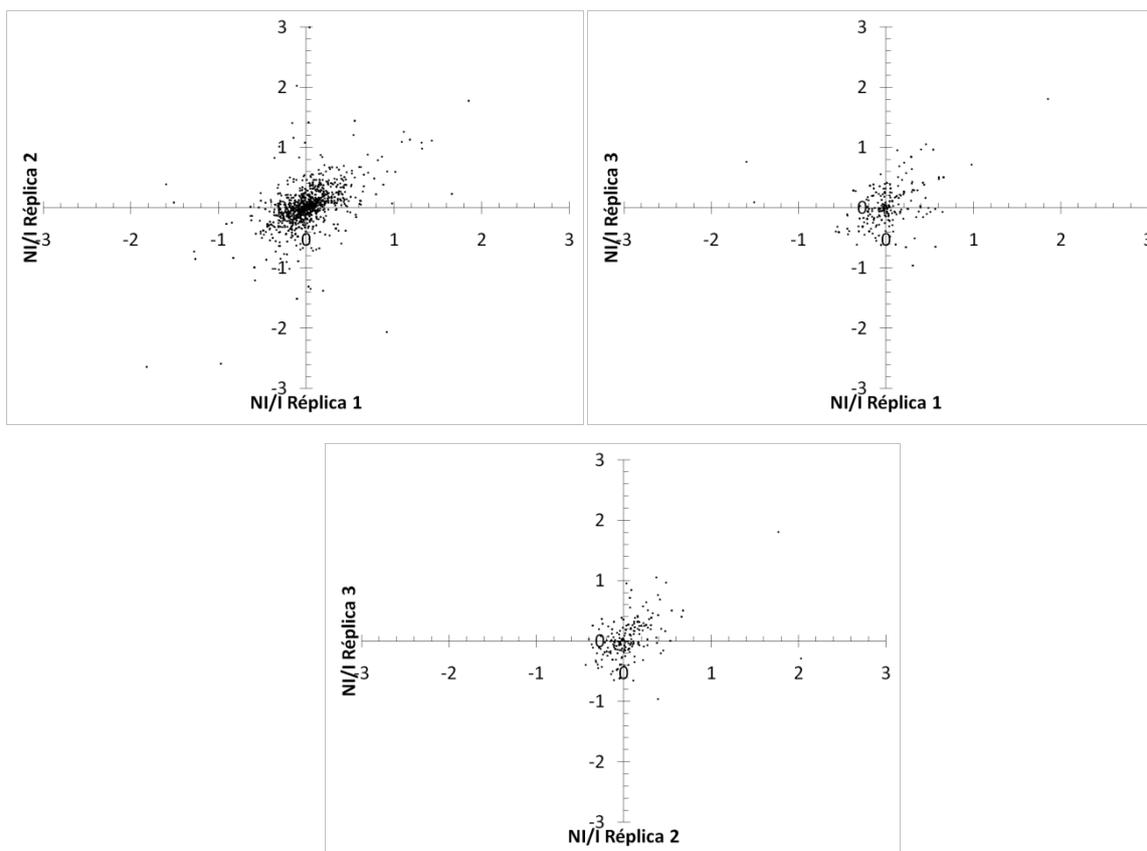


Figura 22: Correlação entre as intensidades NI/I dos fosfossítios nas 3 réplicas biológicas. Eixos x e y \log_2 , NI – não induzido, I – induzido.

No geral foram identificados 1.510 sítios de fosforilação (fosfossítios) em 1.153 peptídeos pertencentes a 648 proteínas. Os números totais de identificação por réplica biológica estão detalhados na Tabela 13.

Tabela 13: Resumo do número de proteínas, peptídeos e fosfossítios identificados.

	Proteínas	PO4	% PO4	Peptídeos	PO4	% PO4	Fosfossítios
Réplica 1	884	620	70.1%	1633	1088	66.6%	1439
Réplica 2	806	572	71.0%	1421	952	67.0%	1273
Réplica 3	384	221	57.6%	535	230	43.0%	293
Total	953	648	68.0%	1854	1153	62.2%	1510
Réplicas 1 e 2	758	551	72.7%	1282	896	69.9%	1215
Réplicas 1 e 3	352	212	60.2%	453	217	47.9%	276
Réplicas 2 e 3	342	205	59.9%	429	208	48.5%	266
Réplicas 1, 2 e 3	331	203	61.3%	410	204	49.8%	262

A fim de aumentar o número de fosfossítios identificados, foram realizados três enriquecimentos sequenciais dos fosfopeptídeos, ou seja, a fração não ligada após a primeira incubação com TiO₂, foi incubada com novas esferas de TiO₂ e isto foi repetido. A contribuição do terceiro enriquecimento sequencial por TiO₂ na identificação de fosfossítios únicos foi em média de 7%. É possível que a quantidade de material inicial não fosse suficiente para o terceiro enriquecimento sequencial. Outra observação é o baixo número de fosfossítios no passo de eluição ACN40, em concordância com os dados obtidos na proteômica (Tabela 14).

Tabela 14: Fosfossítios quantificados por fração e por passo de enriquecimento dos fosfopeptídeo.

	TiO ₂ -1			TiO ₂ -2		TiO ₂ -3	
	Réplica	Únicos	Total	Únicos	Total	Únicos	Total
ACN10%	1	37	198	68	318	10	268
	2	82	82	1	1	0	0
	3	28	30	0	2	0	0
	1;2;3	40	107	0	53	0	41
	Total	187	417	69	374	10	309
ACN14%	1	56	215	51	252	33	180
	2	17	17	0	0	0	0
	3	10	10	0	0	0	0
	1;2;3	33	65	3	34	5	32
	Total	116	307	54	286	38	212
ACN18%	1	67	147	23	111	3	48
	2	9	9	0	0	0	0
	3	2	2	0	0	0	0
	1;2;3	15	35	7	24	0	7
	Total	93	193	30	135	3	55
ACN26%	1	44	90	9	56	10	35
	2	8	8	0	0	0	0
	3	2	2	0	0	0	0
	1;2;3	14	21	1	9	0	4
	Total	68	121	10	65	10	39
ACN40%	1	16	22	2	14	0	4
	2	1	1	0	0	0	0
	3	0	0	0	0	0	0
	1;2;3	4	8	0	5	0	3
	Total	21	31	2	19	0	7

A fim de se avaliar a modulação no perfil de fosforilação do *knockdown*, os dados de proteômica foram utilizados na normalização das intensidades dos fosfossítios. Isto porque um determinado fosfossítio pode sofrer modulação em decorrência da modulação na expressão da proteína que contém este fosfossítio. Fazendo-se a normalização das intensidades dos fosfossítios através das intensidades das proteínas correspondentes obtidas na proteômica foi possível observar as flutuações específicas aos fosfossítios. Aproximadamente 50% dos fosfossítios não puderam ser normalizados pela proteômica, pois as proteínas correspondentes não foram identificadas (Tabela 15).

Tabela 15: Número de fosfossítios normalizados pela proteômica.

Réplica	Fosfossítios		
	Total	Normalizado pela Proteômica	
2 réplicas	974	513	52,7%
1;2	950	506	53,3%
1;3	203	94	46,3%
2;3	187	93	49,7%
1;2;3	187	90	48,1%

Diante da normalização parcial pela proteômica, optou-se pela tentativa de normalizar as razões das intensidades dos fosfossítios pelas intensidades dos peptídeos não fosforilados (identificados no fosfoproteoma) pertencentes à mesma proteína. Para isso, usou-se a mediana das razões das intensidades I/Ni. Desta maneira seria possível a análise dos fosfossítios, cujas proteínas não foram identificadas na proteômica. Primeiramente, verificou-se a correlação entre os peptídeos não fosforilados (dados de fosfoproteômica) e as respectivas proteínas (dados da proteômica). Verificou-se um coeficiente de variação médio de 16,6% nesta correlação, o que inviabilizou a normalização dos fosfossítios pelos peptídeos não fosforilados da mesma proteína. Na Figura 24 é possível observar a variação na quantificação dos fosfossítios modulados aplicando ou não as diferentes normalizações. Esta figura mostra a normalização por peptídeos não fosforilados, contém menos fosfossítios comparados, já que não foi possível tal normalização para todos mostrados na Figura 23. Foi evidenciada pouca variação entre os fosfossítios normalizados pela proteômica e os mesmos fosfossítios não normalizados (Figura

23). Diante deste fato a lista dos peptídeos modulados é composta por fosfósítios normalizados pela proteômica e não normalizados.

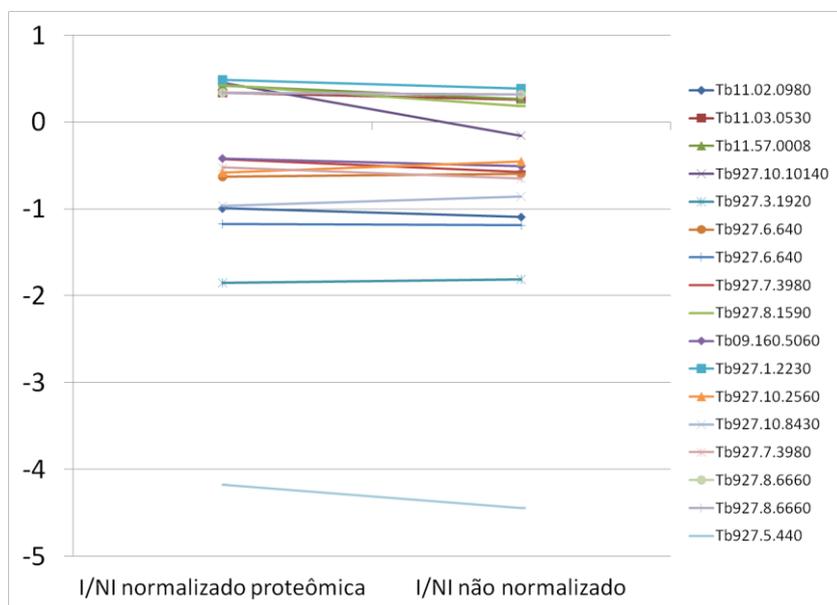


Figura 23: Fosfósítios modulados normalizados pela proteômica e não normalizados.

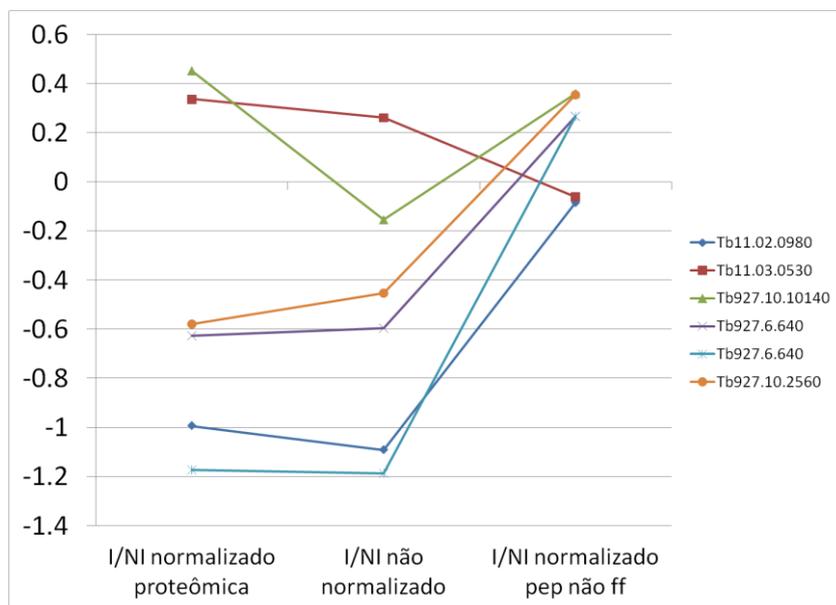


Figura 24: Fosfósítios modulados normalizados pela proteômica, não normalizados e normalizados pelos peptídeos não fosforilados.

5.6.2 Perfil geral dos fosfósítios

Na avaliação geral dos resíduos fosforilados foi possível observar predominância de fosforilação em serina, seguida por treonina e por tirosina. Dos

1.185 fosfosítios selecionados com probabilidade de localização do resíduo fosforilado maior ou igual a 0,75, ou seja, no mínimo 75% de probabilidade de estar correta a localização do resíduo fosforilado, 940 (79,3%) contêm fosfo-serinas, 215 (18,1%) fosfo-treoninas e 30 (2,5%) fosfo-tirosinas. No trabalho de Nett e colaboradores, foram identificados 906 resíduos de fosfo-serinas (75%), 259 fosfo-treoninas (21,5%) e 39 fosfo-tirosinas (3,5%) em *T. brucei* [9]. No trabalho de Marchini e colaboradores, foram encontrados 2.162 (84,1%) fosforilações em resíduos de serina, 384 (14,9%) em resíduos de treonina e 26 (1,0%) em resíduos de tirosina em *T. cruzi* [16].

Neste trabalho foram identificadas 46 quinases com especificidade por proteínas e por outros tipos de moléculas, dentre as quais 30 são proteíno-quinases. Comparando-se com o dados do kinoma de *T. brucei* publicado por Parsons e colaboradores, foi obtida uma cobertura de 19% deste kinoma, porcentagem representante de vários grupos/famílias de proteíno-quinases. A classificação destas proteínas está representada na Tabela 16.

Tabela 16: Classificação das proteíno-quinases identificadas na fosfoproteômica.

Id	Proteína	Grupo/Família
Tb09.160.4560	arginine kinase	-
Tb09.211.2260	protein kinase, putative	Única
Tb09.211.2360	protein kinase A catalytic subunit isoform 2,protein kinase A catalytic subunit	CAMK
Tb927.9.11100	protein kinase A catalytic subunit isoform 1,protein kinase A catalytic subunit	AGC/PKA
Tb09.211.3590	glycerol kinase, glycosomal	-
Tb11.01.0330	aurora B kinase	Aurora
Tb11.01.3180	guanine nucleotide-binding protein beta subunit-like protein, activated protein kinase c receptor	-
Tb11.01.6650	serine/threonine-protein kinase, putative	NEK
Tb11.01.6980	phosphatidylinositol 3-kinase catalytic subunit, putative	-
Tb11.01.7800	nucleoside diphosphate kinase	-
Tb11.02.0640	protein kinase, putative,dual-specificity protein kinase, putative	CMGC/DYRK
Tb11.02.4150	pyruvate phosphate dikinase	-
Tb11.03.0090	ribokinase, putative	-

Tb927.1.1530	protein kinase, putative	STE
Tb927.1.1930	phosphatidylinositol 3-kinase, putative	-
Tb927.1.710	phosphoglycerate kinase	-
Tb927.10.12040	protein kinase, putative,mitogen-activated protein kinase 11, putative	CMGC/MAPK
Tb927.10.13010	protein kinase A catalytic subunit	AGC/PKA
Tb927.10.13780	glycogen synthase kinase 3	CMGC/GSK
Tb927.10.14770	protein kinase, putative	CAMK
Tb927.10.14800	protein kinase, putative,mitogen-activated protein kinase 9, putative	CMGC/RCK
Tb927.10.15020	protein kinase, putative	CMGC/DYRK
Tb927.10.15300	protein kinase, putative	CAMKK
Tb927.10.2010	hexokinase	-
Tb927.10.2020	hexokinase	-
Tb927.10.3230	protein kinase, putative	CMGC/MAPK
Tb927.10.460	NIMA-related protein kinase	NEK
Tb927.10.4940	Nuclear Dbf2-related kinase	AGC/NDR
Tb927.10.5140	protein kinase, putative,mitogen-activated protein kinase 2, putative	CMGC/MAPK
Tb927.10.5310	protein kinase, putative,SNF1-related protein kinases, putative	CAMK/CAMKL
Tb927.2.1820	protein kinase, putative	CAMK/CAMKL
Tb927.2.4210	glycosomal phosphoenolpyruvate carboxykinase	-
Tb927.2.4510	cdc2-related kinase, putative,protein kinase, putative	CMGC/CDK
Tb927.3.1570	protein kinase, putative	Única
Tb927.3.3270	ATP-dependent phosphofructokinase	-
Tb927.3.4560	serine/threonine protein kinase, putative,protein kinase, putative	CAMK/CAMKL
Tb927.3.690	protein kinase, putative	CMGC/RCK
Tb927.4.420	phosphatidylinositol 3-kinase (tor2)	-
Tb927.5.2820	protein kinase, putative	NEK
Tb927.6.4220	mitogen-activated protein kinase 5	CMGC/MAPK
Tb927.6.4970	serine/arginine-rich protein specific kinase SRPK, putative,protein kinase, putative	CMGC/SRPK
Tb927.7.3880	protein kinase, putative	CMGC/DYRK
Tb927.7.4400	inositol polyphosphate kinase-like protein, putative	-
Tb927.7.5770	Nuclear Dbf2-related kinase	AGC/NDR
Tb927.8.3770	mitogen-activated protein kinase, putative	CMGC/MAPK
Tb927.8.6210	phosphatidylinositol 3-kinase, putative	-

As proteíno-quinases em negrito foram encontradas fosforiladas.

Fazendo-se a comparação do presente trabalho com outros dois trabalhos que analisaram o fosfoproteoma de *T. brucei*, Nett e colaboradores e Urbaniak e colaboradores foi observada baixa sobreposição entre os sítios de fosforilação e as fosfoproteínas identificadas quando se compara o trabalho de Nett com o de Urbaniak ou com o presente trabalho. O trabalho de Nett identificou 1.204 fosfosítios em 491 fosfoproteínas nas formas sanguíneas de *T. brucei*. O dado utilizado nesta comparação foi extraído do material suplementar (Table S2 Nett.xls), onde constam 1.906 fosfosítios. A baixa sobreposição dos fosfosítios deste trabalho com os apresentados aqui pode (287), em parte, estar relacionada à análise de formas diferentes de *T. brucei* (procíclico vs sanguíneo) ou à menor cobertura do fosfoproteoma entre os dois trabalhos, quando comparada ao trabalho de Urbaniak (Tabela 17 e Figura 25). As sobreposições entre os dados de Nett ou Batista e Urbaniak representa quase metade dos fosfosítios identificados nos dois primeiros trabalhos. A sobreposição de 50% dos fosfosítios identificados neste trabalho com os identificados por Urbaniak demonstra a complexidade do fosfoproteoma, visto que somando-se os fosfosítios presentes nos três trabalhos, é provável que se tenha apenas uma sub-representação do fosfoproteoma de *T. brucei*.

Tabela 17: Sobreposição dos fosfoproteomas identificados em diferentes trabalhos.

Venn	Trabalho	Fosfoproteínas		Fosfosítios	
		Total	Únicos	Total	Únicos
a	Batista	648	76	1505	510
b	Urbaniak	2545	1668	10077	8314
c	Nett	594	56	1906	796
ab	Batista_Urbaniak	566	345	940	708
ac	Batista_Nett	227	6	287	55
bc	Urbaniak_Nett	532	311	1055	823
Abc	Batista_Urbaniak_Nett	221	221	232	232

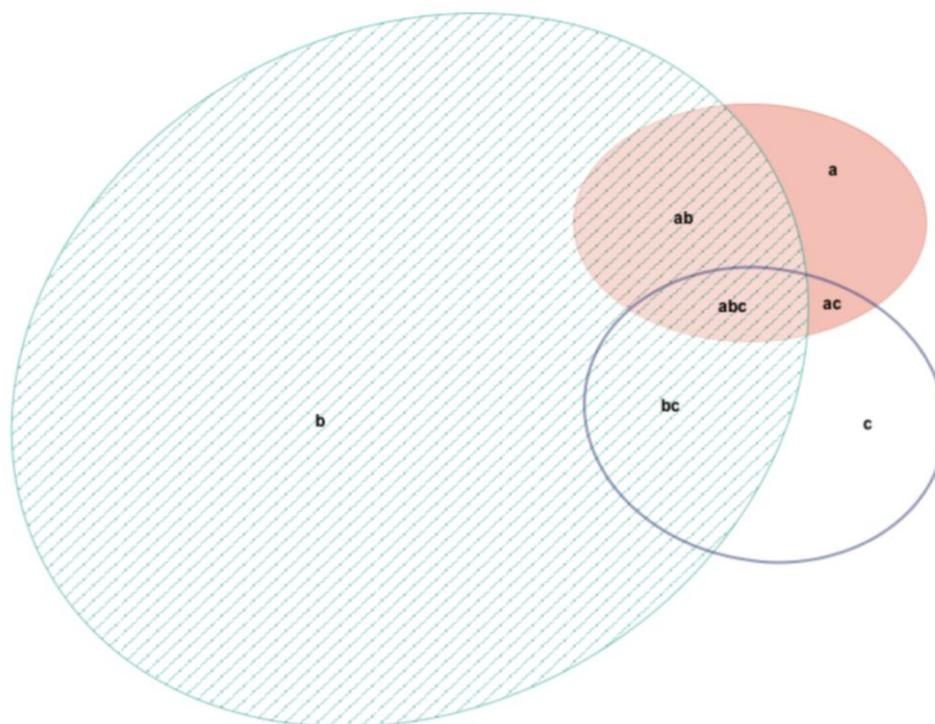


Figura 25: Sobreposição dos fosfosítios identificados entre diferentes trabalhos. a – Batista, b – Urbaniak, c – Nett.

Em relação às proteínas fosforiladas, houve maior sobreposição dos dados que a observada para os fosfosítios, com pouco mais de um terço das proteínas compartilhadas entre os trabalhos de Batista e Nett (Tabela 17 e Figura 26). Através dos dados gerados no presente trabalho foram identificados 510 fosfosítios e 76 proteínas fosforiladas não identificadas em dois trabalhos de grande impacto para a área de sinalização celular em *Trypanosomatídeos*.

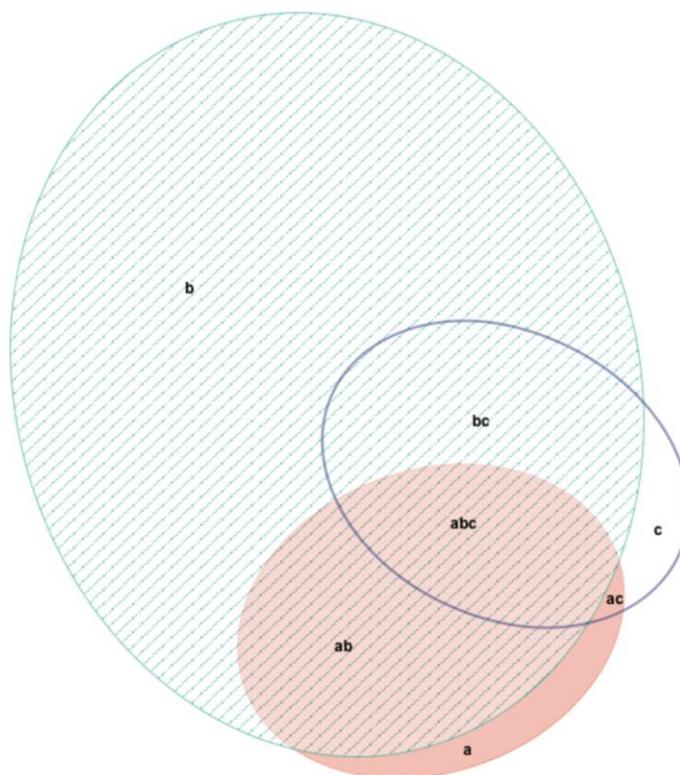


Figura 26: Sobreposição das fosfoproteínas identificadas entre diferentes trabalhos. a – Batista, b – Urbaniak, c – Nett.

5.6.3 Seleção dos fosfossítios modulados

Os fosfossítios foram normalizados pela respectiva réplica da proteômica. Após aplicação do teste-t, foram filtrados os fosfossítios com valor de p menor ou igual a 0,05 e com razão maior que 0,32 ou menor que -0,32 em \log_2 (*fold change* de 1,25). Os valores positivos refletem um aumento na intensidade do fosfossítio no *knockdown*, enquanto os valores negativos indicam diminuição. A lista dos fosfossítios modulados em pelo menos 25%, destacando os que estão modulados pelo menos 50% (+/- 0,58 em \log_2) está representada na

Tabela 18 e na Tabela 19.

Tabela 18: Fosfotios diminuídos no *knockdown* normalizados ou não normalizados com modulação mínima de 25%.

Id	Proteína	I/NI 1	I/NI 2	I/NI 3	AA	Posição	p value	Normalizado
Tb927.5.440	trans-sialidase, putative	-4.31	-4.04	-	Y	527	0.0207	S
Tb927.3.1920	NOT5 protein	-1.89	-1.75	-1.91	S	395	0.0007	S
Tb927.7.4900	5'-3' exonuclease XRNA, putative,exoribonuclease 1, putative	-1.18	-1.12	-	S	1282	0.0154	N
Tb927.6.640	kinetoplastid-specific phospho-protein phosphatase, putative	-1.15	-1.19	-	S	68	0.0098	S
Tb927.11.3490	hypothetical protein, conserved	-1.01	-0.98	-	S	377	0.0095	S
Tb927.10.8430	40S ribosomal protein S12, putative	-0.96	-0.97	-	T	14	0.0014	S
Tb927.9.11100	protein kinase A catalytic subunit isoform 1,protein kinase A catalytic subunit	-0.81	-0.79	-	S	315	0.0094	N
Tb927.6.640	kinetoplastid-specific phospho-protein phosphatase, putative	-0.64	-0.61	-	S	209	0.0196	S
Tb927.10.2560	mitochondrial malate dehydrogenase	-0.57	-0.58	-	T	85	0.0036	S
Tb927.7.3980	immunodominant antigen, putative,tc40 antigen-like	-0.56	-0.39	-0.32	S	372	0.0265	S
Tb927.7.3980	immunodominant antigen, putative,tc40 antigen-like	-0.52	-0.52	-	T	387	0.0021	S
Tb927.9.8180	hypothetical protein, conserved	-0.52	-0.57	-	T	438	0.0283	N
Tb927.11.1890	hypothetical protein, conserved	-0.44	-0.38	-	S	19	0.0441	N
Tb927.7.5940	Protein Associated with Differentiation	-0.44	-0.51	-	S	14	0.0465	N
Tb927.7.5940	Protein Associated with Differentiation	-0.43	-0.37	-	T	587	0.0455	N
Tb927.3.1200	hypothetical protein, conserved	-0.41	-0.48	-	T	737	0.0478	N
Tb927.10.3230	protein kinase, putative	-0.39	-0.40	-	S	369	0.0134	N
Tb927.8.1300	hypothetical protein, conserved	-0.36	-0.38	-	T	86	0.0212	N
Tb927.1.4040	hypothetical protein, conserved	-0.34	-0.34	-	S	165	0.0006	N
Tb927.9.5040	cAMP-specific phosphodiesterase	-0.34	-0.35	-	S	920	0.0038	N
Tb927.4.2070	antigenic protein, putative	-0.30	-0.40	-0.43	S	2261	0.0109	N
Tb927.9.6920	hypothetical protein, conserved	-0.26	-0.49	-0.51	T	2	0.0336	S

I – induzido, NI – não-induzido, AA – resíduo de aminoácido fosforilado, S – sim, N – não.

Tabela 19: Fosfotios aumentados no *knockdown* normalizados ou não normalizados com modulação mínima de 25%.

Id	Proteína	I/NI 1	I/NI 2	I/NI 3	AA	Posição	p value	Normalizado
Tb927.11.14700	hypothetical protein, conserved	6.30	7.18	4.74	Y	215	0.0135	N
Tb927.4.4100	hypothetical protein, conserved	5.55	6.50	-	S	351	0.0497	N
Tb927.4.4100	hypothetical protein, conserved	5.55	6.50	-	Y	348	0.0497	N
Tb927.11.4380	ATP-dependent RNA helicase, putative,DEAD/DEAH box RNA helicase, putative	-	4.51	3.99	Y	862	0.0395	N
Tb927.9.12360	RNA-binding protein, putative	0.83	0.84	-	S	373	0.0032	N
Tb927.9.12360	RNA-binding protein, putative	0.83	0.84	-	S	376	0.0032	N
Tb927.10.10140	paraflagellar rod component, putative	0.62	0.46	0.27	S	320	0.0468	S
Tb927.5.2620	hypothetical protein, conserved	0.50	0.50	-	S	96	0.0035	N
Tb927.1.2230	calpain-like protein fragment, putative	0.50	0.48	-	T	118	0.0093	S
Tb927.10.6730	hypothetical protein, conserved	0.48	0.45	-	S	29	0.0224	N
Tb927.10.7930	2,3-bisphosphoglycerate-independent phosphoglycerate mutase	0.46	0.40	-	S	74	0.0440	N
Tb927.9.12360	RNA-binding protein, putative	0.44	0.32	0.35	S	373	0.0095	N
Tb927.9.2670	hypothetical protein, conserved	0.44	-	0.44	T	289	0.0035	N
Tb927.8.1590	ubiquitin-protein ligase, putative	0.43	0.43	-	S	2449	0.0043	S
Tb927.9.12360	RNA-binding protein, putative	0.42	0.30	0.39	S	372	0.0101	N
Tb927.11.1090	calpain-like protein, putative,cytoskeleton associated protein, putative	0.40	0.45	-	S	1227	0.0390	S
Tb927.11.14780	phosphomannose isomerase, putative	0.40	0.44	-	S	55	0.0344	N
Tb927.8.5010	69 kDa paraflagellar rod protein,PFR2	0.36	0.32	-	S	587	0.0326	N
Tb927.11.550	hypothetical protein SCD6.10	0.35	0.32	-	S	129	0.0283	S
Tb927.8.6660	paraflagellar rod component, putative	0.35	0.33	-	S	599	0.0244	S
Tb927.8.6660	paraflagellar rod component, putative	0.35	0.33	-	T	600	0.0244	S
Tb927.7.220	CDP-diacylglycerol synthetase, putative	0.34	0.37	-	S	18	0.0279	N
Tb927.11.12560	hypothetical protein	0.27	0.32	0.32	S	427	0.0030	N

I – induzido, NI – não-induzido, AA – resíduo de aminoácido fosforilado, S – sim, N – não.

Entre os fosfossítios modulados foram encontradas fosforilações em resíduos de serina, treonina e tirosina. Entre os 45 peptídeos fosforilados identificados como modulados, 42 sítios de fosforilação foram individualmente quantificados. Foram encontradas 11 fosfo-treoninas entre os 44 resíduos identificados, representando 25% dos fosfossítios. Esta porcentagem representa um leve enriquecimento de fosfo-treoninas em relação à representação deste sítio no fosfoproteoma identificado neste trabalho (18,1%). Isto pode refletir um direcionamento da MAPK-A ou de outras proteíno-quinases, reguladas por esta MAPK, para sítios contendo treoninas em comparação com a distribuição deste fosfossítios no fosfoproteoma.

Foi realizada busca no *Gene Ontology* para enriquecimento de GO usando-se as proteínas com fosfossítios modulados como entrada, no entanto o resultado foi de baixa confiabilidade, com e-values não significativos. O baixo número de genes anotados, 10 entre os 37 genes usados como entrada contribuiu para a baixa confiabilidade nos resultados.

Na tentativa de identificar motivos presentes nos sítios de fosforilação, fez-se a busca pelo algoritmo Motif-X. Usando-se os 46 sítios identificados e selecionando-se janelas de 15, 25 e 35 aminoácidos flanqueando o resíduo fosforilado não foram encontrados motivos enriquecidos.

5.6.4 Fosfossítios diminuídos após depleção da MAPK-A

Fazendo-se uma análise dos sítios modulados pelo menos 25%, entre os 36 sítios quantificados individualmente, 22 se apresentaram diminuídos, enquanto 20 se apresentaram aumentados no *knockdown*. Quinze fosfossítios modulados pertencem à proteínas hipotéticas, uma porcentagem relativamente baixa, visto que os genomas dos Tripanossomatídeos possuem um grande percentual de proteínas sem atribuição de função, chegando a mais de 50% dos genes com função desconhecida [113]. Fazendo-se a busca por domínios presentes nas proteínas hipotéticas aqui identificadas, dentre os fosfossítios diminuídos, foram encontrados domínios com significância para apenas uma delas, Tb927.11.1890, com 3 domínios WD40 (e-values 3.00E-06, 4.00E-05 e 0.0004), envolvidos com interação proteína-proteína.

A fim de se avaliar os dados obtidos experimentalmente, buscou-se na literatura informações sobre as proteínas contendo modulação de no mínimo 50% no

seu fosfotio. Como exemplo, foram encontrados vários fosfotios para a proteína NOT5, mas apenas um deles foi selecionado pelo critério estatístico (S395), tendo uma diminuição com *fold change* de 3,6. A quantificação dos demais fosfotios e os respectivos resíduos estão demonstrados na Figura 27. No trabalho de Nett e colaboradores, também foram encontrados alguns fosfotios em NOT5 na forma sanguínea de *T. brucei*, sendo que somente a fosforilação no resíduo S322 se sobrepõe com os resultados aqui apresentados [9].

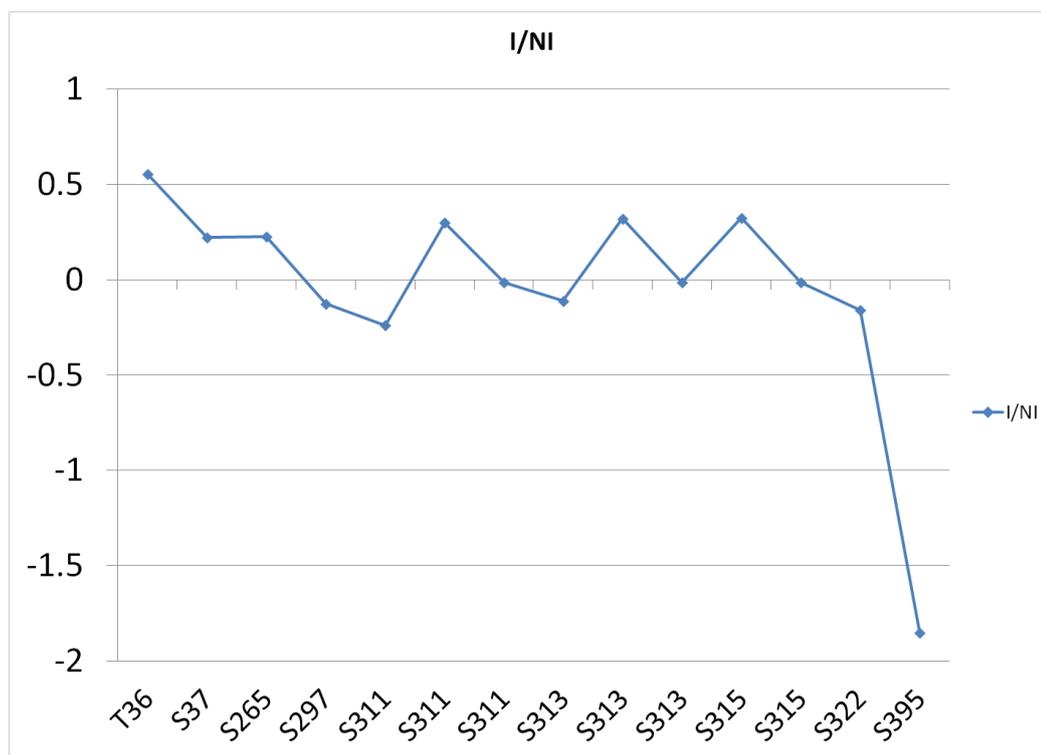


Figura 27: Modulação nos fosfotios de NOT5 quantificados. Eixo y em \log_2 .

A proteína NOT5 participa do complexo Ccr4/Caf1/Not, o qual está envolvido na deadenilação de mRNAs em eucariotos, regulando assim a estabilidade dos mesmos. A depleção de proteínas deste complexo leva a inibição da proliferação celular e diminuição na deadenilação de mRNAs nas formas sanguíneas de *T. brucei* [114]. Outro fosfotio encontrado diminuído pertence à 5'-3' exonuclease XRNA, putative, exoribonuclease 1, putative. Esta proteína está envolvida com a rápida degradação de mRNAs exclusivos da forma procíclica de *T. brucei* [115].

Entre as modulações observadas, uma delas foi detectada em um sítio contendo um resíduo de fosfo-tirosina de uma trans-sialidase. Proteínas tirosino-

quinases convencionais não estão presentes no genoma de *T.brucei*, *T. cruzi* e *L. major*), embora fosfo-tirosinas tem sido frequentemente encontrada nos trabalhos de fosfoproteômica aplicados a Tripanossomatídeos [9, 16, 85]. Em *T. cruzi*, tem sido demonstrada a associação entre fosforilação em trans-sialidases como produto da atividade de PKA, sendo sugerido a importância deste evento no endereçamento das trans-sialidases para o retículo endoplasmático [116]. Aqui foi observada a diminuição de um fosfósíto no domínio catalítico de uma trans-sialidase, acompanhada de uma diminuição de uma fosfo-serina na subunidade catalítica da proteína quinase A. No entanto, o fosfósíto encontrado na trans-sialidase foi uma fosfo-tirosina, sendo que a PKA é uma serino/treonino quinase.

Foram encontrados dois fosfósítos diminuídos pertencentes a uma fosfatase específica da Ordem Kinetoplastidae. Esta fosfatase pertence a um grupo com substituições de importantes resíduos catalíticos, não tendo atividade de defosforilação, no entanto podendo executar diferentes funções em Tripanossomatídeos[10].

A proteína com Id Tb927.7.3980, um antígeno imunodominante descrito em *T. cruzi* [117] apresentou diminuição em dois fosfósítos. Esta proteína possui um domínio WD40, o qual está envolvido em interações proteína-proteína.

Também foi identificada diminuição em um fosfósíto presente na proteína ribossomal S12. Foram evidenciadas fosforilações em resíduos de serina e treonina em diferentes proteínas ribossomais, incluindo S12 na bactéria filamentosa *Streptomyces collinus*. Sugere-se a possível influência destas fosforilações na interação códon-anticódon e na ligação de fatores de iniciação da tradução [118].

Além da PKA, outra proteína-quinase contendo fosfósíto diminuído, embora com uma baixa modulação (*fold change* de 1,3), está uma fosfo-serina S369 pertencente a uma MAPK (identificada como B neste trabalho).

5.6.5 Fosfósítos aumentados após depleção da MAPK-A

Entre as proteínas que apresentaram aumento nos fosfósítos quantificados em pelo menos 50%, está uma proteína pouco estudada anotada como *RNA-binding protein, putative* (Tb927.9.12360). Através de busca por domínios conservados, esta proteína apresentou dois domínios Pfam com alinhamentos não significativos. Um

dos domínios é da família RRM (*RNA Recognition Motif*) e o outro pertence à família PAM-2, associado em eucariotos com a proteínas ligadoras de poli-A. Estas tem sido associadas com regulação pós-transcricional, regulando processos como exportação em meia vida de mRNAs.

Avaliando-se as proteínas classificadas como hipotéticas que contém fosfotios aumentados no *knockdown*, apenas duas delas apresentam domínios Pfam significativos. Dentre estas, está a proteína Tb927.11.550, que possui um domínio Sm Scd6-like (e-value 6.00E-29). Este domínio pertence a uma família de proteínas que se ligam a RNA [119]. Membros desta família foram associados com *P-bodies* em mamíferos e fungos, estando também envolvidos no controle pós transcricional em oocitos de *Xenopus* [120]. A outra proteína hipotética (Tb927.10.6730) possui um domínio denominado *plasma-membrane choline transporter* (e-value 3.10E-83) presente em *C. elegans*.

Dentre os fosfotios aumentados com maior valor de *fold change* (19,0) está uma fosfo-tirosina (Y862) pertencente a uma *ATP-dependent RNA helicase, putative, DEAD/DEAH box RNA helicase, putative*. Esta proteína tem sido associada com processo de *splicing* em eucariotos. Tem sido demonstrado o envolvimento de RNA helicases na edição de RNAs em *T. brucei* e *Leishmania* [121][122].

Diante das modulações observadas (aumento ou diminuição) é possível fazer algumas considerações. Entre os fosfotios encontrados modulados, alguns estão relacionados a proteínas que participam de processos de regulação de mRNA, como é caso da proteína hipotética com domínio Sm Scd6-like, da NOT5, da RNA helicase e da 5'-3' exonuclease XRNA. A MAPK aqui analisada pode ter uma influência em processos de regulação de degradação/estabilidade de mRNAs, os quais são bastante evidenciados em Tripanossomatídeos. Este dado, aparentemente não se relaciona com os dados obtidos na proteômica, que mostraram modulação de algumas proteínas mitocondriais. É importante salientar o desafio do entendimento das modulações na expressão gênica, abordadas aqui por RNAseq e proteômica, combinadas à modificações pós-traducionais por fosforilação. Tem sido demonstrada baixa correlação entre níveis de proteína e de mRNA [77]. Além disso, a dinâmica dos eventos de fosforilação insere uma dificuldade adicional no entendimento dos mecanismos de sinalização celular.

5.7 Avaliação da Interação entre proteínas da via das MAPKs

As MAPKs participam de cascatas de sinalização nas células em resposta a estímulos do ambiente, de maneira que um determinado grupo, por exemplo, as MAP2Ks, especificamente ativam por fosforilação em sítios de tirosina e treonina as MAPKs. Neste sentido, existe uma interação quinase substrato entre os membros desta família, sendo que estas interações são desconhecidas em Tripanossomatídeos, com raras exceções provenientes de ensaios *in vitro* [32]. Aqui neste trabalho, dentre as metodologias para detecção de interação proteína-proteína, optou-se pela abordagem de duplo-híbrido (Y2H), por ser uma abordagem passível de uso em larga escala e por estar implantada e ser de domínio do grupo de trabalho. Dentre os 40 membros da via das MAPKs, incluindo MAPKs, MAP2Ks (*Ste7-like*), MAP3Ks (*Ste11*) e MAP4Ks (*Ste20*). Foram testadas interações entre 34 membros da família das MAPKs, em um total de 1.024 combinações para cada uma das duas cepas de *S. cerevisiae* empregadas. Os resultados mostraram uma taxa de auto-ativação de aproximadamente 30% (Tabela 21), com maior ocorrência na cepa MAV, no meio sem histidina com 100 mM de AT. Uma proteína auto-ativadora, geralmente quando está ligada ao domínio de ligação ao DNA, é capaz de ativar a transcrição do gene repórter, mesmo na ausência do domínio de ativação. A justificativa para o uso de duas cepas está baseada nas possíveis diferenças na detecção de interações e na questão de auto-ativação, pois algumas proteínas podem ser auto-ativadoras em uma cepa, mas não em outra. O uso de diferentes meios auxilia no balanceamento entre ruído ou auto-ativação e interações reais. Foram identificadas 22 interações entre as 34 proteínas avaliadas (Figura 28).

Tabela 20: Interações detectadas e auto-ativação por duplo-híbrido.

AD	Nome	BD		MAV		Y		
		Id	Nome	-URA	100 mM 3AT	1 mM 3AT	3 mM 3AT	-ADE
Tb927.10.10870	A	Tb927.10.14780	STE11	+				
Tb927.8.3770	C	Tb927.10.5140	I					+++
Tb927.8.3550	D	Tb927.8.5730	STE20	+++	+++			
Tb927.10.12040	H	Tb927.10.5140	I					+
Tb927.10.5140	I	Tb927.10.5140	I					+
Tb927.10.7780	J	Tb927.10.10350	STE11					+/-
Tb927.11.16790	L	Tb927.8.3550	D	+				
Tb927.11.16790	L	Tb927.11.8170	K	+				
Tb927.11.16790	L	Tb927.10.14300	STE11	+				
Tb927.11.16790	L	Tb927.10.10350	STE11	+/-				
Tb927.10.14800	M	Tb927.10.10350	STE11			+	+++	
Tb927.3.690	N	Tb927.11.8170	K	+				
Tb927.3.690	N	Tb927.10.7780	J	+				
Tb927.3.690	N	Tb927.10.14300	STE11	+				
Tb927.3.690	N	Tb927.10.10350	STE11	+				
Tb927.3.690	N	Tb927.10.10870	A	+				
Tb927.3.690	N	Tb927.8.3550	D	+				
Tb927.3.690	N	Tb927.3.4860	STE7-like	+				
Tb927.3.690	N	Tb927.11.14270	STE7-like	+				
Tb927.3.690	N	Tb927.10.5140	I	+				
Tb927.3.690	N	Tb927.10.14800	M	+				
Tb927.3.4860	STE7-like	Tb927.10.10350	STE11					+++
Tb927.11.14270	STE7-like	Tb927.3.690	N	+++				
Tb927.8.5950	STE7-like	Tb927.10.10350	STE11	+	+			
Tb927.10.14780	STE11	Tb927.10.10350	STE11	++				
Tb927.6.2030	STE11	Tb927.10.10350	STE11					+
Tb927.9.3120	STE11	Tb927.10.10350	STE11	++				
Tb927.10.2040	STE(?)	Tb927.9.3120	STE11				+++	
Tb927.3.2060	STE(?)	Tb927.11.14270	STE7-like			+++		
Tb927.3.2060	STE(?)	Tb927.10.5140	I			+		
Tb927.3.2060	STE(?)	Tb927.9.3120	STE11				+	
Tb927.8.5730	STE20	Tb927.10.14780	STE11	+				

Autoativadoras

MAV	Y	MAV e Y
-----	---	---------

Tabela 21: Autoativações obtidas no ensaio de duplo híbrido.

Autoativação								
Id	Nome	AD	BD	MAV		Y		
				-URA	100 mM 3AT	1 mM 3AT	3 mM 3AT	-ADE
Tb927.10.10870	A		X		+	+++	++	
Tb927.6.4220	G		X	+/-				
Tb927.10.5140	I		X		+		+++	
Tb927.11.8170	K		X		+	+++	+++	++
Tb927.11.16790	L		X				++	
Tb927.10.14800	M		X				++	
Tb927.3.690	N	x		+				
Tb927.7.6680	STE11		X		+	+++	+++	++
Tb927.7.3650	STE11		X				++	
Tb927.10.14300	STE11		X		+	+++	+++	
Tb927.9.12400	STE11		X		+			
Tb927.10.2040	STE(?)							
Tb927.3.2060	STE(?)		X	+++	+++	+++	+++	+++

AD = domínio de ativação, BD = domínio de ligação a DNA.

Os genes listados na Tabela 22: Proteínas não avaliadas por Y2H. não foram inseridos na plataforma para os ensaios de Y2H.

Tabela 22: Proteínas não avaliadas por Y2H.

Id	Nome
Tb927.10.3230	B
Tb927.11.14070	STE11
Tb927.7.5220	STE11
Tb927.10.1910	STE11
Tb927.3.5650	STE11-like
Tb927.1.1530	STE(?)

Avaliando-se a rede de interações obtidas foram evidenciadas várias interações convergindo para uma MAP3K classificada como Ste11 (Tb927.10.10350). Tais interações ocorrem entre esta MAP3K e outras quinases classificadas nos três níveis da cascata das MAPKs (MAP1K, MAP2K e MAP3K). Esta quinase pode exercer um papel central na regulação desta via. Obviamente

que a comprovação de tais resultados necessita de ensaios adicionais de interação através de outras metodologias como *pull down* e imunoprecipitação. Entre as outras interações observadas, foram detectadas duas interações entre quinases de níveis adjacentes na cascata de sinalização. Uma MAPK (Tb927.3.690), interagiu com uma MAP2K (Tb927.11.142.70) e uma MAP3K (Tb927.10.14780) interagiu com uma MAP4K (Tb927.8.5730). As duas MAPKs, cujo nível na cascata não está predito (MAP?K) interagiram com proteínas MAPK, MAP2K e MAP3K. Uma das MAPKs (Tb927.10.5140) teve uma auto-interação (Figura 28).

Foi evidenciada a fosforilação *in vitro* de uma MAPK (LinJ.19.1480) por uma MAP2K de *L. infantum* (LinJ36.1770) [32]. Os ortólogos em *T. brucei* (MAP2K - Tb927.10.5270 e MAPK – Tb927.10.16030), não apresentaram interação. Eventos de fosforilação podem resultar de interações dependentes de outras proteínas, fato ausente no Y2H. Além disso, usando a mesma justificativa, ensaios de quinase-substrato por fosforilação *in vitro* podem não refletir os eventos biológicos no contexto celular.

Além de se pensar nas interações entre as MAPKs tendo como consequência a fosforilação direta e ativação em cascata. Outra maneira de interação entre as vias de MAPKs é observada com a participação de STEs como proteínas *scaffolds* [123] [124]. Como exemplo pode-se citar a proteína Ste5, que em levedura atua como *scaffold*. Esta interação resulta na formação do complexo de levedura que participa na resposta a ferormônio [125].

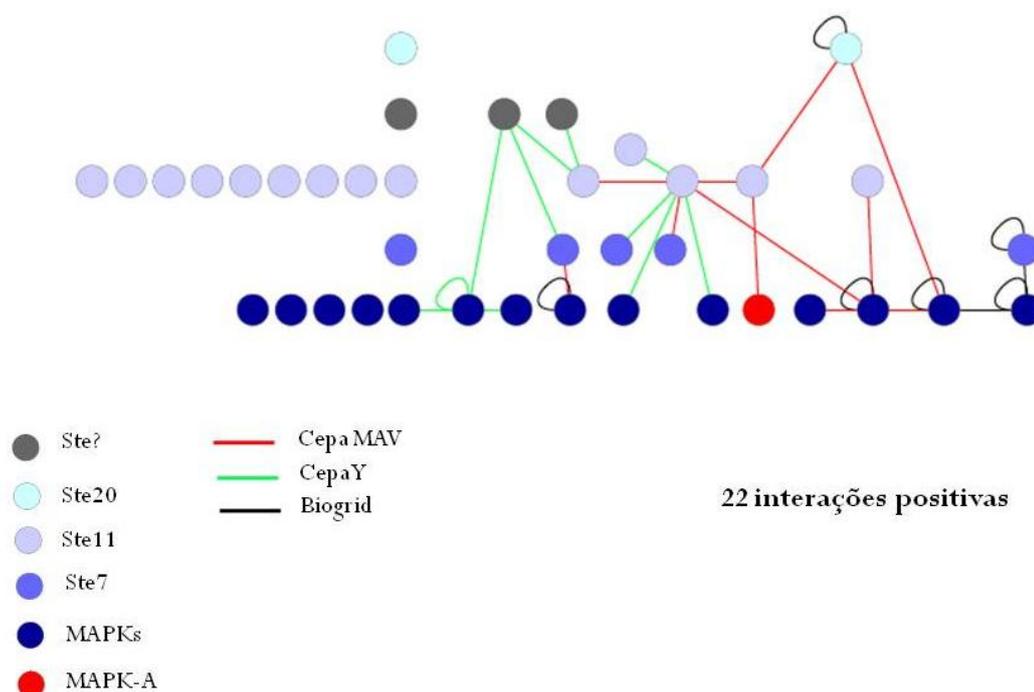


Figura 28: Interações entre MAPKs detectadas por Y2H.

As interações obtidas foram comparadas com os bancos de dados de interação Intact e Biogrid. Buscou-se por interações entre os ortólogos das proteínas avaliadas aqui e outros parceiros com ortólogos em *T. brucei*. O banco de dados Intact possui 314.020 interações enquanto o Biogrid possui 696.238. A Figura 29 mostra a união das redes obtidas por Y2H neste trabalho com as interações encontradas nos bancos de dados explorados. Não foi observada sobreposição das interações aqui descritas com as interações encontradas nos bancos de dados, embora as MAPKs D, E, I, L e N e a STE Tb927.10.5270 possuem interações nestes bancos de dados.

Interações entre as proteínas das vias das MAPKs adicionando-se outros interatores.

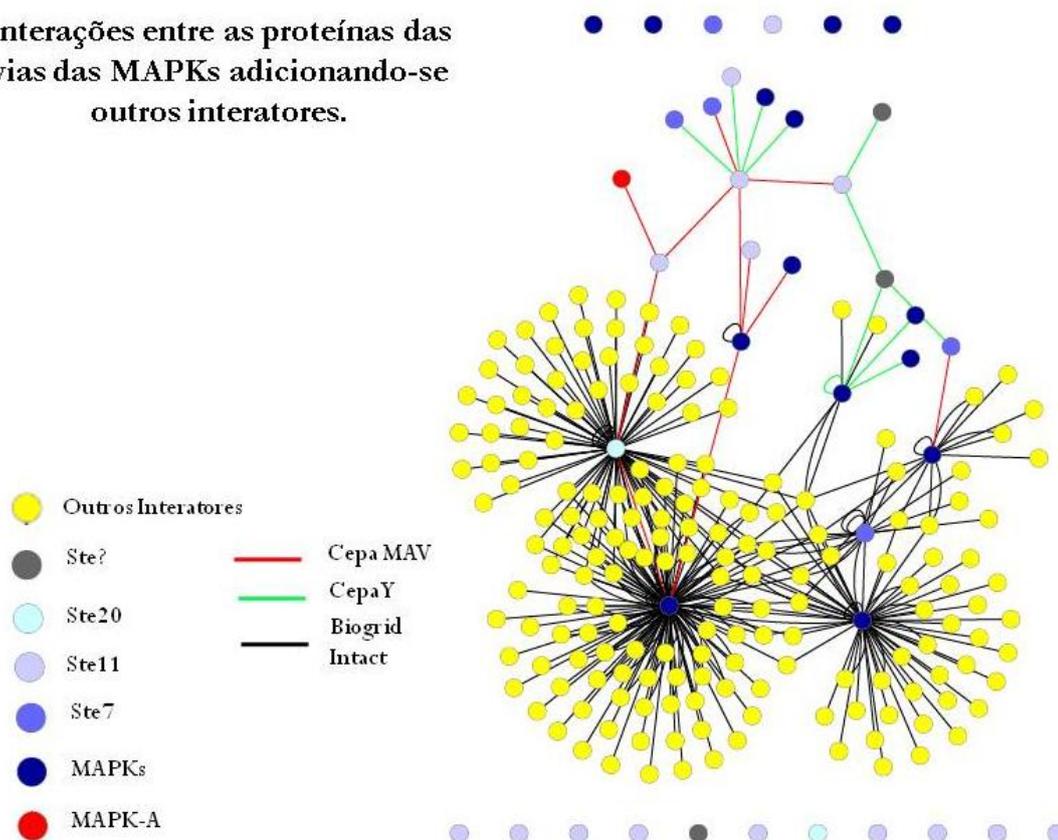


Figura 29: União das redes de interação envolvendo ortólogos de proteínas das vias das MAPKs. Azul – Interações testadas no presente trabalho. Amarelo - Interações não testadas. Verde – Interações encontradas nos bancos Biogrid e Intact.

6. CONCLUSÃO

O conjunto de dados apresentados aqui revelou a influência de cinco MAPKs sobre a proliferação celular de *T. brucei*. As análises do transcriptoma, do proteoma e do fosfoproteoma, realizadas para uma destas quinases (Tb927.10.10870, MAPK-A) demonstraram pouca modulação no *knockdown*, diante da significativa influência desta MAPK na proliferação celular. No entanto, dentre os fosfossítios modulados, uma parte pertence a proteínas envolvidas em processos de regulação da estabilidade de mRNAs na célula. Isto demonstra a possível relação da MAPK-A com estas proteínas e com processos de regulação da estabilidade de mRNAs na célula. A possível influência da MAPK-A avaliada sobre estes processos merece uma avaliação aprofundada. Uma das formas de se compreender a influência de cada um dos principais fosfossítios modulados pela MAPK-A seria através de mutações dos resíduos presentes nestes sítios e avaliação dos efeitos sobre a célula. Adicionalmente, um terço dos fosfossítios identificados não havia sido publicado em pelo menos dois importantes trabalhos abordando o fosfoproteoma de *T. brucei*. Dentre as interações entre proteínas da via das MAPKs testadas por Y2H foram encontradas 22 interações positivas, com uma MAP3K interagindo MAPKs de diferentes níveis dentro da cascata de fosforilação. Estes dados, em conjunto com a crescente elucidação da função das proteínas classificadas como hipotéticas e também da confirmação da função das proteínas anotadas por inferência podem fornecer pistas importantes sobre a biologia dos Tripanossomatídeos.

7. ANEXOS

Tabela 23: Oligonucleotídeos usados no RNAi e na QPCR

Gene	Oligonucleotídeo	
	RNAi	QPCR
MAPK-A	F	TTTGGTTCGTTTGCATGGTA
	R	ATCAATCCATTTCTGCTGCC
MAPK-B	F	GAGTGCCGACTTCTTTTTCG
	R	TTTGGATCCTGCGTATCCTC
MAPK-C	F	ATCAAGAGTTGTGCGAGCCT
	R	GGATAGTGACGCTACGGGAA
MAPK-D	F	ATCAAGAGTTGTGCGAGCCT
	R	GGATAGTGACGCTACGGGAA
MAPK-E	F	CCTAGGGCTCATCTTTGCTG
	R	TGCCAAAATAATACGCCACA
MAPK-F	F	GCGTTCATTACCCAAGAGGA
	R	CGAAGCACTGAAATCTGCAA
MAPK-G	F	TCCTCCCGAGCTATTGCTTA
	R	GCATGCAAGTCAAACGAGAA
MAPK-H	F	ACCAGTTGGACGACCGTTAC
	R	ACACATCAATTTGTGCGGAA
MAPK-I	F	GCTTGAAAAGGAGCAAGTGG
	R	TACTCTTGCATGGTGAAGCG
MAPK-J	F	AAATTGTCTATGCCGTTGCC
	R	GAGCAAAGTAAGGGTGCTGC
MAPK-K	F	TTCTCCCTTTATTGCAACCG
	R	TTTCTTACGGGGCTATGGTG
MAPK-L	F	ACGCCACCCTAACGTCATAG
	R	CAGCGTCTTCCACATCAA
MAPK-M	F	GAAGGTTACCAGCTCTTGCG
	R	ATCGTCAAGTCCAGGTTTGG
MAPK-N	F	GGCAGAGGCACTGAAGAAAC
	R	GTCGCATTGCTTTGAAGTGA
MAPK-O	F	GCGATCAAGTGCATGAGAAA
	R	CCAACGGGTCGAGACATACT

Tabela 24: Protocolo de preparo do meio SDM-79.

Componente	Concentração estoque (mg/mL)	Volume solução estoque (mL)	Solvente	Concentração final mg/L	uso (mL/mg)
Sais					
CaCl ₂	NA			180	180
NaH ₂ PO ₄ .H ₂ O	NA			126	126
acetate.Na	NA			10	10
Pyruvate.Na	NA			100	100
MgSO ₄	NA			87.9	21528.9
NaCl				6120	
KCl				360	
Glucose				1900	
Glucosamine.HCl				50	
Phenol Red				11	
HEPES (sodium salt)				8000	
MOPS (free acid)				5000	
Fe(NO ₃) ₃ .9H ₂ O	1.4	20		0.14	0.1
D-ribose	1			0.1	
deoxy-D-ribose	1			0.1	
L-aminoácidos					
Cysteine	2	10		0.02	0.01
L-glutathione (reduced)	1			0.01	
Alanine	27.225	20		210	7.713
Asparagine	5.94	20		8	1.347
Aspartic acid	31.02	20		14	0.451
Cystine	37.18	5		29	0.780
Glutamic acid	58.795	20		24	0.408
Glutamine	23.1	20		320	13.853
Glycine	78.72549	5.1		15	0.191
Histidine.HCl.H ₂ O	21.85333	15		38	1.739
Hydroxyproline	19.06667	3		2	0.105
Isoleucine	24.53	20		46	1.875
Leucine	44.77	20		54	1.206
Methionine	28.6	10		85	2.972
Phenylalanine	21.384	25		111	5.191
Proline	38.3625	40		615	16.031
Serine	26.10667	15		71	2.720
Taurine	35.2	5		160	4.545
Threonine	36.52	20		394	10.789
Tryptophan	8.25	20		10	1.212
Tyrosine	56.1	10		150	2.674

Valine	38.64667	15		42	1.087
Vitaminas e outros					
p-aminobenzoic acid	2.01	10		2.01	1
myo-Inositol	1.41			1.41	
Folic Acid	4.7	10	1 M NaOH (50 mg/ml)	4.7	1
Biotin	0.02	1000		0.2	10
D-Ca pantothenate	0.7	100		0.7	1
Choline chloride	0.8			0.8	
Niacinamide	0.705			0.705	
Thiamine.HCl	0.702			0.702	
Pyridoxal.HCl (-20 °C)	0.705			0.705	
Ascorbic acid	0.033	300		0.01	0.3
Retinyl acetate	0.093			0.028	
Riboflavin	0.072	1000		0.072	1
(+) alpha tocopherol succinate	0.08			0.001914	0.024
Menadione	0.09	9		0.002	0.022
Ergocalciferol	2	10		0.02	0.01
Niacin	0.2	50		0.005	0.025
Pyridoxine.HCl	0.2	50		0.005	0.025
Purinas e outros					
Adenosine	5	200	H ₂ O (7 mg/ml) ou 1 M NH ₄ OH (50 mg/ml)	10	2
Adenine SO ₄ . 2H ₂ O	2	30		2	1
Guanosine	20	10		10	0.5
AMP.H ₂ O (Adenosine-5'-Monophosphate)	0.24	50		0.04	0.167
ATP.Na ₂ .3H ₂ O	2	10		0.2	0.1
Hypoxanthine	0.12	100		0.06	0.5
Guanine.HCl	1.5	50	1 N NaOH (50 mg/ml)	0.05	0.033
Xanthine	1.5		1 M NaOH (50 mg/ml)	0.05	
Thymine	0.72	50	1 N NaOH (50 mg/ml)	0.06	0.083
Uracil	0.72		1 M NaOH (50 mg/ml)	0.06	
Lipídios e soro					
Cholesterol	4	6	Solúvel 1,29% em etanol	0.04	0.01
Tween 80				4	0.004
Elevar para 750 mL					
NaHCO ₃				3980	3980
Elevar para 900 mL					
Ajustar o pH para 7,4					
Antibióticos					
Penicilina				59	59
Streptomycin				133	133

Solvente informado quando diferente de água.

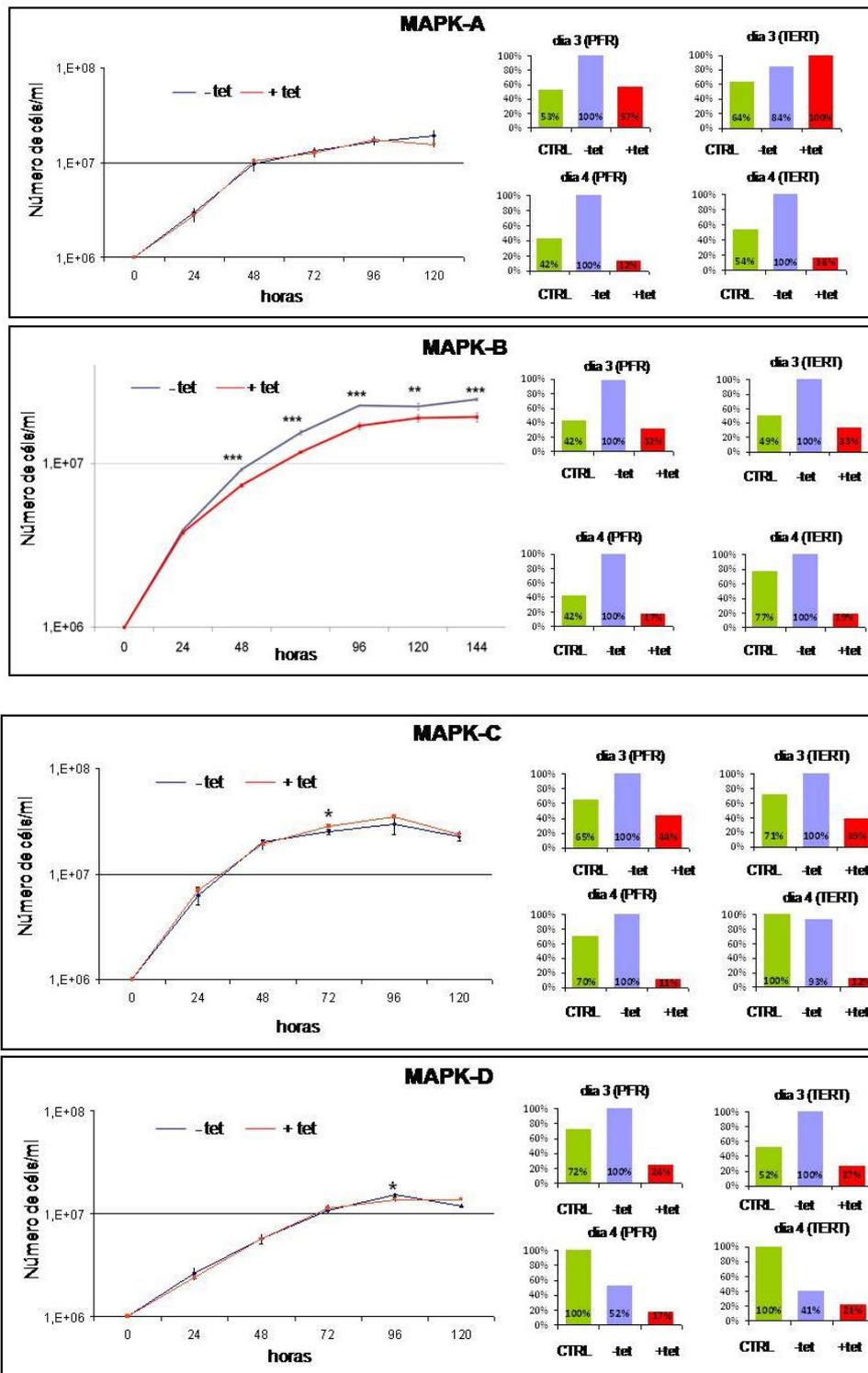
Tabela 25: Oligonucleotídeos utilizados nas amplificações dos alvos para duplo híbrido

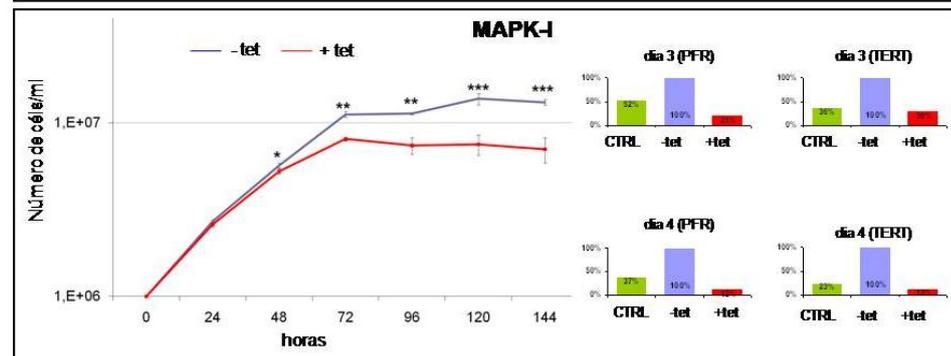
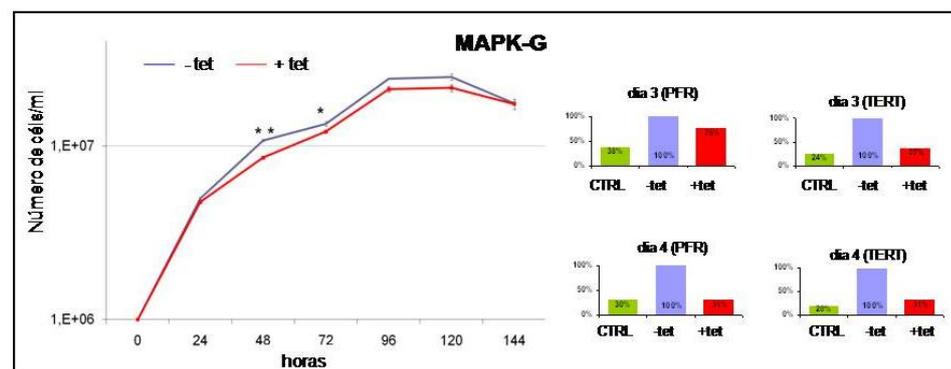
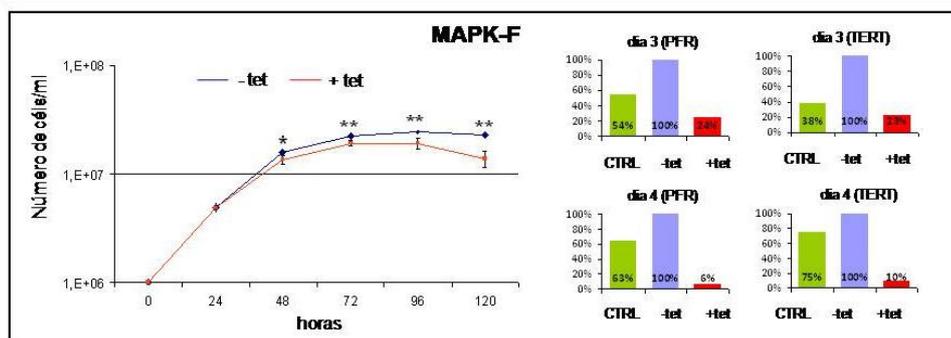
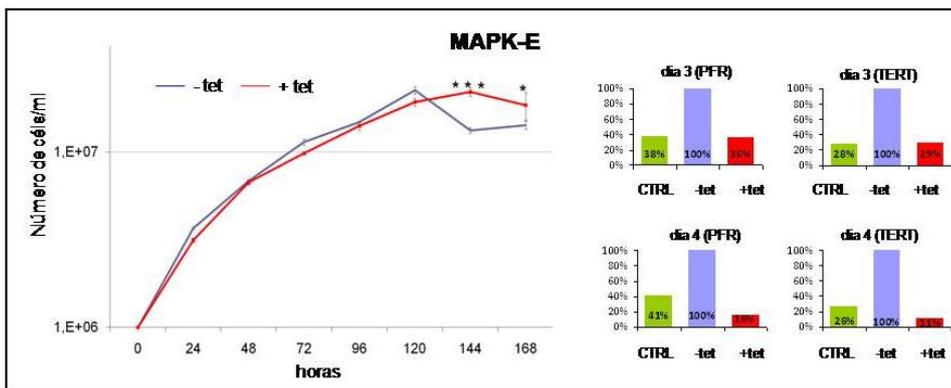
Gene	Oligonucleotídeo
Tb927.3.4860_F	5'GGCTCCACCATGAAAAGGAAGGGGATTGCGTTAG3'
Tb927.3.4860_R	5'TGGGTGGATYCACAGAACACCTAGCATTTC AATTTTATGT3'
Tb927.11.14270_F	5'GGCTCCACCATGGTCTCTCTCCGGATCAGGAAAAC3'
Tb927.11.14270_R	5'TGGGTGGATYCAACACAAGTGGGGAAATTTTTCATC3'
Tb927.11.14070_F	5'GGCTCCACCATGACGAAAGAGGATCACGGTGAGG3'
Tb927.11.14070_R	5'TGGGTGGATYCACAAACAAAATGCATATGAAAGTAAACTGG3'
Tb927.8.6810_F	5'GGCTCCACCATGAGTTGCACACCTGCAACGTTAG3'
Tb927.8.6810_R	5'TGGGTGGATYCATTCGTTTCCCCAGTATCCAC3'
Tb927.8.1100_F	5'GGCTCCACCATGAGTGAGGACAGCGCAGCG3'
Tb927.8.1100_R	5'TGGGTGGATYCACGGGTGACCAAGAAAGGATG3'
Tb927.7.6680_F	5'GGCTCCACCATGCCTGTCGTGAAGGGCG3'
Tb927.7.6680_R	5'TGGGTGGATYCACTCATCGTCTGGTGGCACC3'
Tb927.7.5220_F	5'GGCTCCACCATGACACAGCCCTCGGCAGATG3'
Tb927.7.5220_R	5'TGGGTGGATYCAATTACCTTCCACAGCAGCGTTAC3'
Tb927.8.1780_F	5'GGCTCCACCATGGACCATAAACTGCTACGGTAGATG3'
Tb927.8.1780_R	5'TGGGTGGATYCACGCTTTTGAAGGTTTGAGC3'
Tb927.7.3650_F	5'GGCTCCACCATGGGAAGTGAAGGAAGTGTAGTAGCA3'
Tb927.7.3650_R	5'TGGGTGGATYCAAATATTGGGGAAAGAAACACCCA3'
Tb927.10.2040_F	5'GGCTCCACCATGCGGGCTAACGACATTGTGAA3'
Tb927.10.2040_R	5'TGGGTGGATYCATCCACATTTGTCTGTGCCTG3'
Tb927.10.14300_F	5'GGCTCCACCATGAGTTGATGGAGGAATTACACGC3'
Tb927.10.14300_R	5'TGGGTGGATYCACACATCCCCTGAATGCGTGG3'
Tb927.10.1910_F	5'GGCTCCACCATGAAGCAAGAGTTTGTTCGCCATG3'
Tb927.10.1910_R	5'TGGGTGGATYCAAAGGGAAAAAACGGGAAGTGCC3'
Tb927.10.14780_F	5'GGCTCCACCATGCCCGCAGGCATGTGTTCC3'
Tb927.10.14780_R	5'TGGGTGGATYCACACCTTTAAGAAGGGATGTTGCAAGAG3'
Tb927.6.2030_F	5'GGCTCCACCATGTCCCTTTGGAAAATTGCCAC3'
Tb927.6.2030_R	5'TGGGTGGATYCACGGCTGACAATGGGC3'
Tb927.10.10350_F	5'GGCTCCACCATGAGTACACAAGAATCTCATGGTTCTGTG3'
Tb927.10.10350_R	5'TGGGTGGATYCAAAACGTAAGAAATGGATGTTTCAAAG3'
Tb927.1.1530_F	5'GGCTCCACCATGGTTTCTAAAGAGGAGAATGCGAAC3'
Tb927.1.1530_R	5'TGGGTGGATYCACCCACTCAATTTTGCAGTTCC3'
Tb927.11.2040_F	5'GGCTCCACCATGAAGCTGAGGGCTGACTCCC3'
Tb927.11.2040_R	5'TGGGTGGATYCATTTGATCAACCACTTGATCACCG3'
Tb927.8.5950_F	5'GGCTCCACCATGCCACCCAAACGACCCAAG3'
Tb927.8.5950_R	5'TGGGTGGATYCACGATACTGTGGTGAACCATCTTCC3'
Tb927.8.5730_F	5'GGCTCCACCATGAGCGACGAGGGGTTTGAAG3'
Tb927.8.5730_R	5'TGGGTGGATYCACAGCGGTAGACACTTCAACAC3'
Tb927.10.5270_F	5'GGCTCCACCATGACCGGGTTGAAGATTGACATC3'

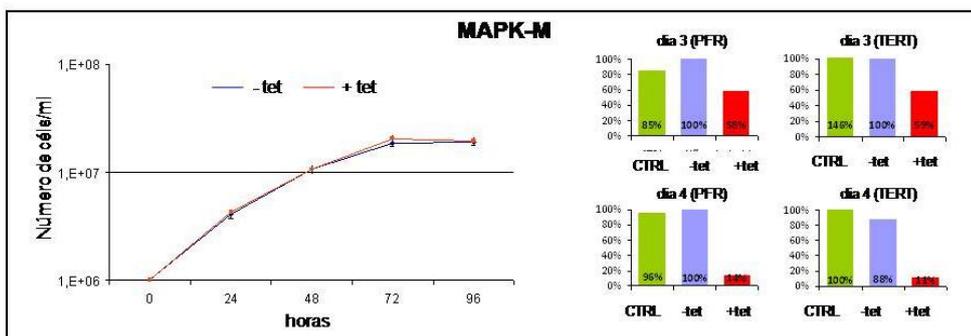
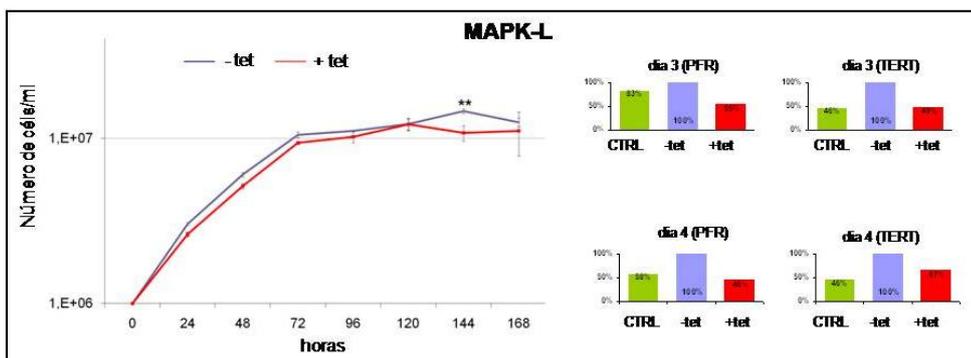
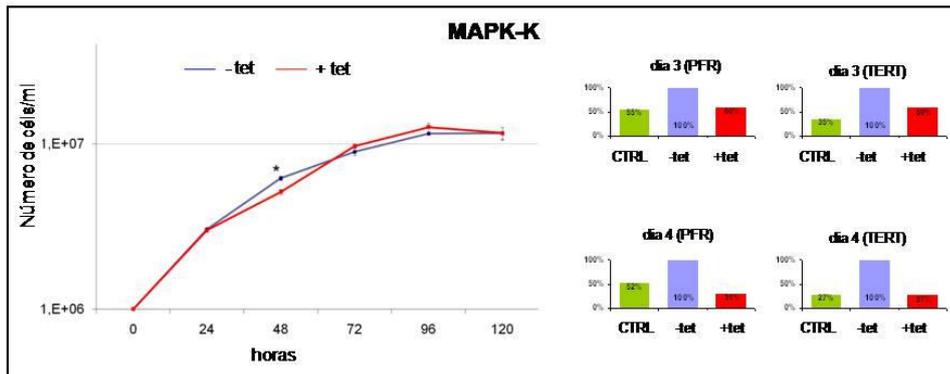
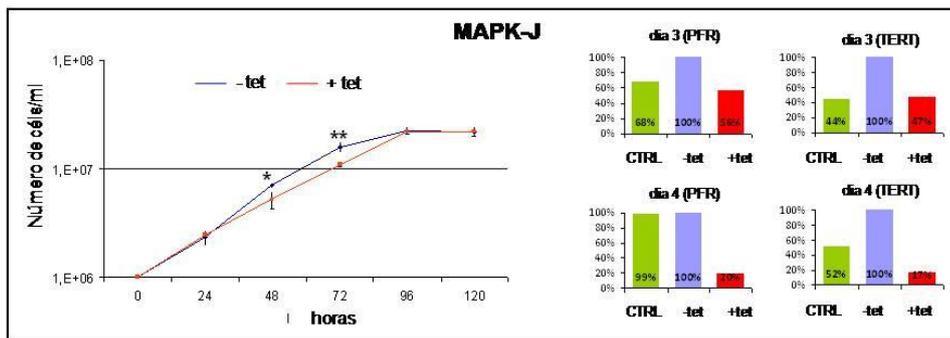
Tb927.10.5270_R	5'TGGGTGGATYCACTGCACCAACTATTTAGTTCCTCG3'
Tb927.3.5650_F	5'GGCTCCACCATGATAGAATATGCGTGTGGTTGGCTAC3'
Tb927.3.5650_R	5'TGGGTGGATYCACAAATGAAATTCGCTTCTCGAGATATTC3'
Tb927.9.12880_F	5'GGCTCCACCATGGATGACGGAAGTGGACACCTC3'
Tb927.9.12880_R	5'TGGGTGGATYCATATGGGACCCATAGCCCTTG3'
Tb927.9.3120_F	5'GGCTCCACCATGGGACCAACATGTGATCGAATAGC3'
Tb927.9.3120_R	5'TGGGTGGATYCACGAATGGGGTTGTGTGGTGATG3'
Tb927.9.12400_F	5'GGCTCCACCATGTGTGCCTTAAGGGATATCGCG3'
Tb927.9.12400_R	5'TGGGTGGATYCAAGTTTACCATAAAGTCATGGAACAAAAGCTC3'
Tb927.3.2060_F	5'GGCTCCACCATGGAGAACATTCGTAGGACGAATCG3'
Tb927.3.2060_R	5'TGGGTGGATYCACTCTGGGTCTTTACTCGACCGGG3'
Tb927.10.10870_F	5'GGCTCCACCATGCACAATGTCCCGTTACTCGC3'
Tb927.10.10870_R	5'TGGGTGGATYCAACGAGGTGATGAAAAGTGCCTCAG3'
Tb927.10.3230_F	5'GGCTCCACCATGAACAGTAATGCTCCGCTTATCTTTG3'
Tb927.10.3230_R	5'TGGGTGGATYCAATCCTCATCATCGCTAACTAGTAGGTC3'
Tb927.8.3770_F	5'GGCTCCACCATGGGTTCTAAAGGCGAGGAGCTG3'
Tb927.8.3770_R	5'TGGGTGGATYCAAGGCATAGAGGTGGGAAGATCTG3'
Tb927.8.3550_F	5'GGCTCCACCATGCACAAAAGTAATCAAGAGTTGTGCG3'
Tb927.8.3550_R	5'TGGGTGGATYCACGCGCCTTTCGGTATTTGG3'
Tb927.10.16030_F	5'GGCTCCACCATGGACATACCACAAGAACTGGCTTG3'
Tb927.10.16030_R	5'TGGGTGGATYCAAGTACCCTTTGGACCGGGAG3'
Tb927.6.1780_F	5'GGCTCCACCATGTTGACGCATGTGTCCAGTCCC3'
Tb927.6.1780_R	5'TGGGTGGATYCACCGTCTCGAAGCACTGAAATCTG3'
Tb927.6.4220_F	5'GGCTCCACCATGGTAACAGCAAATGGCTCCTCAAAC3'
Tb927.6.4220_R	5'TGGGTGGATYCACTGATAGTGTGCTGAATCTTTCCATATATCC3'
Tb927.10.12040_F	5'GGCTCCACCATGAACCGTAGTTCGAGGAACCTTGC3'
Tb927.10.12040_R	5'TGGGTGGATYCACTCAACGGAACCCACATATTTGTC3'
Tb927.10.5140_F	5'GGCTCCACCATGTCATCAGAAATAGAGCCACATATCCTC3'
Tb927.10.5140_R	5'TGGGTGGATYCATTTGTGCAACACACGAGAGGC3'
Tb927.10.7780_F	5'GGCTCCACCATGGTGTGCTTCAGCATCGATGG3'
Tb927.10.7780_R	5'TGGGTGGATYCACAAAGGCTAACTTTTGCTTCGACG3'
Tb927.11.8170_F	5'GGCTCCACCATGGAAGACTACGAAACCGTGCC3'
Tb927.11.8170_R	5'TGGGTGGATYCACCGCTTCGAGCCATTGTAATG3'
Tb927.11.16790_F	5'GGCTCCACCATGGATGCGTACGAAACACTTGGTATG3'
Tb927.11.16790_R	5'TGGGTGGATYCATCGTTTTGTTGCACCATCTCG3'
Tb927.10.14800_F	5'GGCTCCACCATGAATCGGTATACGATATTGGGTCAGC3'
Tb927.10.14800_R	5'TGGGTGGATYCAAAAGTTGAACGCTTCATCGTCAAGTC3'
Tb927.3.690_F	5'GGCTCCACCATGGAGAAGTACCAACTTCTCAATCAGTTAGG3'
Tb927.3.690_R	5'TGGGTGGATYCAACTGTTGGAACCTTGGTTGTGCTC3'
Tb927.9.9320_F	5'GGCTCCACCATGCAAAAGTACGCAATATTAGGAAAGAAAG3'
Tb927.9.9320_R	5'TGGGTGGATYCACCTTGAAGTTTGGGCAGAA3'
attB1	5'GGGACAAGTTTGTACAAAAAAGCAGGCTTC3'

attB2

5'GGGGACCACTTTGTACAAGAAAGCTGGGTC3'







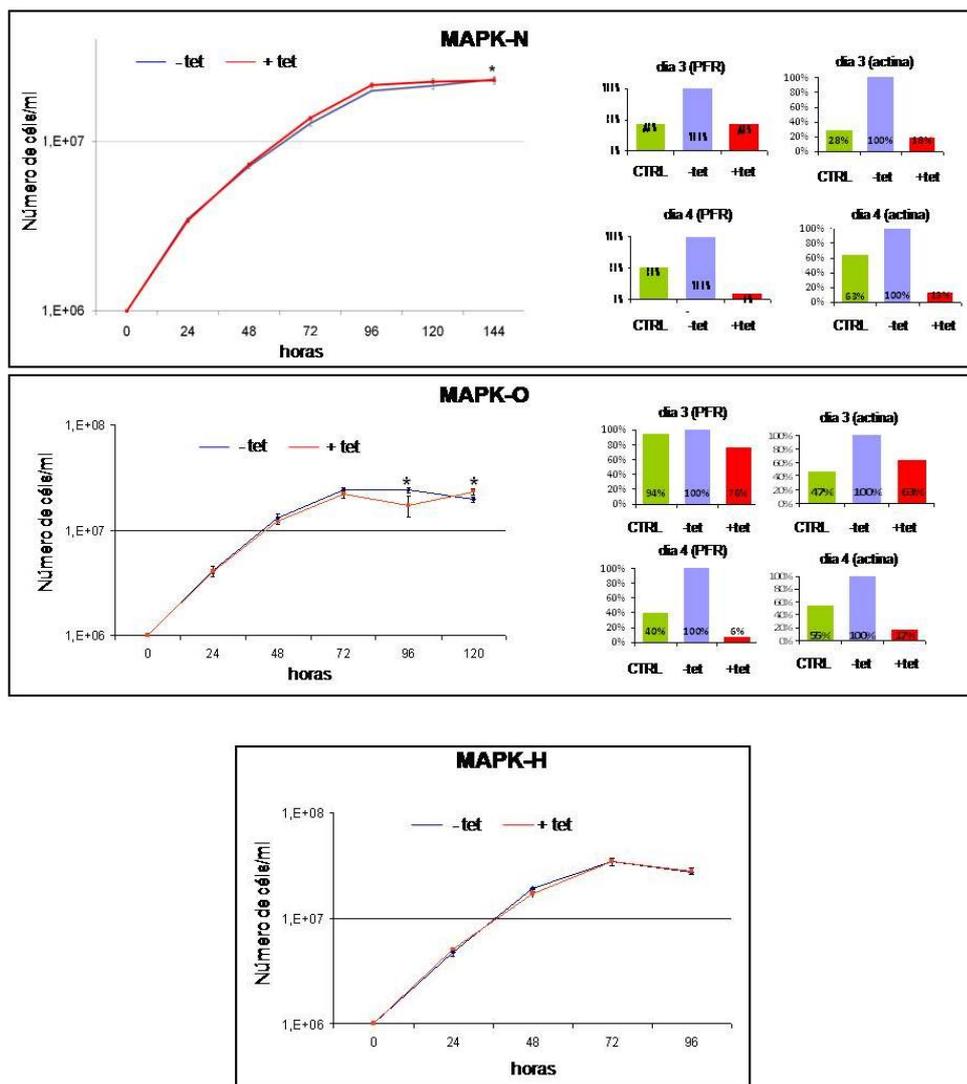


Figura 30: Avaliação dos níveis de mRNA e do efeito da depleção das MAPKs na proliferação celular de *T. brucei*. Foram avaliadas culturas induzidas (+ tet) e não-induzidas (-tet) para RNAi. * = 0,01 < pvalue < 0,05, ** = 0,001 < pvalue < 0,01, *** = pvalue < 0,001. Nos gráficos de crescimento celular (à esquerda), o eixo das ordenadas está em escala logarítmica. Nos gráficos que representam os níveis de mRNA (à direita), as barras verdes representam a relação controle/não-induzido, as barras azuis representam a relação não-induzido/não-induzido e as barras vermelhas representam a relação induzido/não-induzido.

8. REFERÊNCIAS BIBLIOGRÁFICAS

1. Englund PT, Hajduk SL, Marini JC: **The molecular biology of trypanosomes**. *Annu Rev Biochem* 1982, **51**:695-726.
2. Malvy D, Chappuis F: **Sleeping sickness**. *Clin Microbiol Infect* 2011, **17**(7):986-995.
3. Steverding D: **The history of African trypanosomiasis**. *Parasit Vectors* 2008, **1**(1):3.
4. Vickerman K: **Developmental cycles and biology of pathogenic trypanosomes**. *Br Med Bull* 1985, **41**(2):105-114.
5. Matthews KR: **The developmental cell biology of *Trypanosoma brucei***. *J Cell Sci* 2005, **118**(Pt 2):283-290.
6. Manning G, Whyte DB, Martinez R, Hunter T, Sudarsanam S: **The protein kinase complement of the human genome**. *Science* 2002, **298**(5600):1912-1934.
7. Miranda-Saavedra D, Barton GJ: **Classification and functional annotation of eukaryotic protein kinases**. *Proteins* 2007, **68**(4):893-914.
8. Parsons M, Worthey EA, Ward PN, Mottram JC: **Comparative analysis of the kinomes of three pathogenic trypanosomatids: *Leishmania major*, *Trypanosoma brucei* and *Trypanosoma cruzi***. *BMC Genomics* 2005, **6**:127.
9. Nett IR, Martin DM, Miranda-Saavedra D, Lamont D, Barber JD, Mehlert A, Ferguson MA: **The phosphoproteome of bloodstream form *Trypanosoma brucei*, causative agent of African sleeping sickness**. *Mol Cell Proteomics* 2009, **8**(7):1527-1538.
10. Brenchley R, Tariq H, McElhinney H, Szoor B, Huxley-Jones J, Stevens R, Matthews K, Taberner L: **The TriTryp phosphatome: analysis of the protein phosphatase catalytic domains**. *BMC Genomics* 2007, **8**:434.
11. Wilkes JM, Doerig C: **The protein-phosphatome of the human malaria parasite *Plasmodium falciparum***. *BMC Genomics* 2008, **9**:412.
12. Alonso A, Rojas A, Godzik A, Mustelin T: **The dual-specific protein tyrosine phosphatase family**. In: *Topics in Current Genetics: Protein Phosphatases*. Edited by Springer, vol. 5; 2004: 26.
13. Cohen PTW: **Overview of protein serine/threonine phosphatases**. In: *Topics in Current Genetics: Protein Phosphatases*. vol. 5: Springer; 2004: 20.
14. Alonso A, Sasin J, Bottini N, Friedberg I, Osterman A, Godzik A, Hunter T, Dixon J, Mustelin T: **Protein tyrosine phosphatases in the human genome**. *Cell* 2004, **117**(6):699-711.
15. Stern A, Privman E, Rasis M, Lavi S, Pupko T: **Evolution of the metazoan protein phosphatase 2C superfamily**. *J Mol Evol* 2007, **64**(1):61-70.
16. Marchini FK, de Godoy LM, Rampazzo RC, Pavoni DP, Probst CM, Gnad F, Mann M, Krieger MA: **Profiling the *Trypanosoma cruzi* phosphoproteome**. *PLoS One* 2011, **6**(9):e25381.
17. Bossemeyer D: **Protein kinases--structure and function**. *FEBS Lett* 1995, **369**(1):57-61.

18. Tokmakov AA, Sato KI, Fukami Y: **Deregulation of mitogen-activated protein kinase at low pH due to a structural rearrangement of activation segment.** *Biochim Biophys Acta* 2000, **1476**(1):66-74.
19. Boulton TG, Nye SH, Robbins DJ, Ip NY, Radziejewska E, Morgenbesser SD, DePinho RA, Panayotatos N, Cobb MH, Yancopoulos GD: **ERKs: a family of protein-serine/threonine kinases that are activated and tyrosine phosphorylated in response to insulin and NGF.** *Cell* 1991, **65**(4):663-675.
20. Boulton TG, Yancopoulos GD, Gregory JS, Slaughter C, Moomaw C, Hsu J, Cobb MH: **An insulin-stimulated protein kinase similar to yeast kinases involved in cell cycle control.** *Science* 1990, **249**(4964):64-67.
21. Raman M, Chen W, Cobb MH: **Differential regulation and properties of MAPKs.** *Oncogene* 2007, **26**(22):3100-3112.
22. Cargnello M, Roux PP: **Activation and function of the MAPKs and their substrates, the MAPK-activated protein kinases.** *Microbiol Mol Biol Rev* 2011, **75**(1):50-83.
23. Chen RE, Thorner J: **Function and regulation in MAPK signaling pathways: lessons learned from the yeast *Saccharomyces cerevisiae*.** *Biochim Biophys Acta* 2007, **1773**(8):1311-1340.
24. Hua SB, Wang CC: **Interferon-gamma activation of a mitogen-activated protein kinase, KFR1, in the bloodstream form of *Trypanosoma brucei*.** *J Biol Chem* 1997, **272**(16):10797-10803.
25. Muller IB, Domenicali-Pfister D, Roditi I, Vassella E: **Stage-specific requirement of a mitogen-activated protein kinase by *Trypanosoma brucei*.** *Mol Biol Cell* 2002, **13**(11):3787-3799.
26. Ellis J, Sarkar M, Hendriks E, Matthews K: **A novel ERK-like, CRK-like protein kinase that modulates growth in *Trypanosoma brucei* via an autoregulatory C-terminal extension.** *Mol Microbiol* 2004, **53**(5):1487-1499.
27. Domenicali Pfister D, Burkard G, Morand S, Renggli CK, Roditi I, Vassella E: **A Mitogen-activated protein kinase controls differentiation of bloodstream forms of *Trypanosoma brucei*.** *Eukaryot Cell* 2006, **5**(7):1126-1135.
28. Wiese M: **A mitogen-activated protein (MAP) kinase homologue of *Leishmania mexicana* is essential for parasite survival in the infected host.** *EMBO J* 1998, **17**(9):2619-2628.
29. Wiese M, Wang Q, Gorcke I: **Identification of mitogen-activated protein kinase homologues from *Leishmania mexicana*.** *Int J Parasitol* 2003, **33**(14):1577-1587.
30. Bengs F, Scholz A, Kuhn D, Wiese M: **LmxMPK9, a mitogen-activated protein kinase homologue affects flagellar length in *Leishmania mexicana*.** *Mol Microbiol* 2005, **55**(5):1606-1615.
31. Wang Q, Melzer IM, Kruse M, Sander-Juelch C, Wiese M: **LmxMPK4, a mitogen-activated protein (MAP) kinase homologue essential for promastigotes and amastigotes of *Leishmania mexicana*.** *Kinetoplastid Biol Dis* 2005, **4**:6.
32. John von Freyend S, Rosenqvist H, Fink A, Melzer IM, Clos J, Jensen ON, Wiese M: **LmxMPK4, an essential mitogen-activated protein kinase of *Leishmania mexicana* is phosphorylated and activated by the STE7-like protein kinase LmxMKK5.** *Int J Parasitol* 2010, **40**(8):969-978.

33. Fire AZ: **Gene silencing by double-stranded RNA.** *Cell Death Differ* 2007, **14**(12):1998-2012.
34. Ngo H, Tschudi C, Gull K, Ullu E: **Double-stranded RNA induces mRNA degradation in Trypanosoma brucei.** *Proc Natl Acad Sci U S A* 1998, **95**(25):14687-14692.
35. Balana-Fouce R, Reguera RM: **RNA interference in Trypanosoma brucei: a high-throughput engine for functional genomics in trypanosomatids?** *Trends Parasitol* 2007, **23**(8):348-351.
36. Hammarton TC, Kramer S, Tetley L, Boshart M, Mottram JC: **Trypanosoma brucei Polo-like kinase is essential for basal body duplication, kDNA segregation and cytokinesis.** *Mol Microbiol* 2007, **65**(5):1229-1248.
37. Hellman KM, Ciganda M, Brown SV, Li J, Ruyechan W, Williams N: **Two trypanosome-specific proteins are essential factors for 5S rRNA abundance and ribosomal assembly in Trypanosoma brucei.** *Eukaryot Cell* 2007, **6**(10):1766-1772.
38. Taylor MC, Kaur H, Blessington B, Kelly JM, Wilkinson SR: **Validation of spermidine synthase as a drug target in African trypanosomes.** *Biochem J* 2008, **409**(2):563-569.
39. Aeby E, Palioura S, Pusnik M, Marazzi J, Lieberman A, Ullu E, Soll D, Schneider A: **The canonical pathway for selenocysteine insertion is dispensable in Trypanosomes.** *Proc Natl Acad Sci U S A* 2009, **106**(13):5088-5092.
40. Beverley SM: **Protozomics: trypanosomatid parasite genetics comes of age.** *Nat Rev Genet* 2003, **4**(1):11-19.
41. LaCount DJ, Barrett B, Donelson JE: **Trypanosoma brucei FLA1 is required for flagellum attachment and cytokinesis.** *J Biol Chem* 2002, **277**(20):17580-17588.
42. Wirtz E, Leal S, Ochatt C, Cross GA: **A tightly regulated inducible expression system for conditional gene knock-outs and dominant-negative genetics in Trypanosoma brucei.** *Mol Biochem Parasitol* 1999, **99**(1):89-101.
43. Redmond S, Vadivelu J, Field MC: **RNAit: an automated web-based tool for the selection of RNAi targets in Trypanosoma brucei.** *Mol Biochem Parasitol* 2003, **128**(1):115-118.
44. Broadhead R, Dawe HR, Farr H, Griffiths S, Hart SR, Portman N, Shaw MK, Ginger ML, Gaskell SJ, McKean PG *et al*: **Flagellar motility is required for the viability of the bloodstream trypanosome.** *Nature* 2006, **440**(7081):224-227.
45. Subramanian R, Gondi CS, Lakka SS, Jutla A, Rao JS: **siRNA-mediated simultaneous downregulation of uPA and its receptor inhibits angiogenesis and invasiveness triggering apoptosis in breast cancer cells.** *Int J Oncol* 2006, **28**(4):831-839.
46. de Godoy LM, Olsen JV, Cox J, Nielsen ML, Hubner NC, Frohlich F, Walther TC, Mann M: **Comprehensive mass-spectrometry-based proteome quantification of haploid versus diploid yeast.** *Nature* 2008, **455**(7217):1251-1254.
47. Olsen JV, Vermeulen M, Santamaria A, Kumar C, Miller ML, Jensen LJ, Gnad F, Cox J, Jensen TS, Nigg EA *et al*: **Quantitative phosphoproteomics**

- reveals widespread full phosphorylation site occupancy during mitosis.** *Sci Signal* 2010, **3**(104):ra3.
48. Lundby A, Lage K, Weinert BT, Bekker-Jensen DB, Secher A, Skovgaard T, Kelstrup CD, Dmytriyev A, Choudhary C, Lundby C *et al*: **Proteomic analysis of lysine acetylation sites in rat tissues reveals organ specificity and subcellular patterns.** *Cell Rep* 2012, **2**(2):419-431.
 49. Wu C, Tran JC, Zamdborg L, Durbin KR, Li M, Ahlf DR, Early BP, Thomas PM, Sweedler JV, Kelleher NL: **A protease for 'middle-down' proteomics.** *Nat Methods* 2012, **9**(8):822-824.
 50. Biemann K: **Contributions of mass spectrometry to peptide and protein structure.** *Biomed Environ Mass Spectrom* 1988, **16**(1-12):99-111.
 51. Roepstorff P, Fohlman J: **Proposal for a common nomenclature for sequence ions in mass spectra of peptides.** *Biomed Mass Spectrom* 1984, **11**(11):601.
 52. Pappin DJ, Hojrup P, Bleasby AJ: **Rapid identification of proteins by peptide-mass fingerprinting.** *Curr Biol* 1993, **3**(6):327-332.
 53. Henzel WJ, Billeci TM, Stults JT, Wong SC, Grimley C, Watanabe C: **Identifying proteins from two-dimensional gels by molecular mass searching of peptide fragments in protein sequence databases.** *Proc Natl Acad Sci U S A* 1993, **90**(11):5011-5015.
 54. Mann M, Hojrup P, Roepstorff P: **Use of mass spectrometric molecular weight information to identify proteins in sequence databases.** *Biol Mass Spectrom* 1993, **22**(6):338-345.
 55. James P, Quadroni M, Carafoli E, Gonnet G: **Protein identification by mass profile fingerprinting.** *Biochem Biophys Res Commun* 1993, **195**(1):58-64.
 56. Yates JR, 3rd, Speicher S, Griffin PR, Hunkapiller T: **Peptide mass maps: a highly informative approach to protein identification.** *Anal Biochem* 1993, **214**(2):397-408.
 57. Eng JK, McCormack AL, Yates JR: **An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database.** *J Am Soc Mass Spectrom* 1994, **5**(11):976-989.
 58. Peng J, Elias JE, Thoreen CC, Licklider LJ, Gygi SP: **Evaluation of multidimensional chromatography coupled with tandem mass spectrometry (LC/LC-MS/MS) for large-scale protein analysis: the yeast proteome.** *J Proteome Res* 2003, **2**(1):43-50.
 59. Gygi SP, Rist B, Gerber SA, Turecek F, Gelb MH, Aebersold R: **Quantitative analysis of complex protein mixtures using isotope-coded affinity tags.** *Nat Biotechnol* 1999, **17**(10):994-999.
 60. Ross PL, Huang YN, Marchese JN, Williamson B, Parker K, Hattan S, Khainovski N, Pillai S, Dey S, Daniels S *et al*: **Multiplexed protein quantitation in *Saccharomyces cerevisiae* using amine-reactive isobaric tagging reagents.** *Mol Cell Proteomics* 2004, **3**(12):1154-1169.
 61. Ong SE, Blagoev B, Kratchmarova I, Kristensen DB, Steen H, Pandey A, Mann M: **Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics.** *Mol Cell Proteomics* 2002, **1**(5):376-386.

62. Cox J, Mann M: **MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification.** *Nat Biotechnol* 2008, **26**(12):1367-1372.
63. Aebersold R, Goodlett DR: **Mass spectrometry in proteomics.** *Chem Rev* 2001, **101**(2):269-295.
64. Reinders J, Sickmann A: **State-of-the-art in phosphoproteomics.** *Proteomics* 2005, **5**(16):4052-4061.
65. Zhou H, Watts JD, Aebersold R: **A systematic approach to the analysis of protein phosphorylation.** *Nat Biotechnol* 2001, **19**(4):375-378.
66. Andersson L, Porath J: **Isolation of phosphoproteins by immobilized metal (Fe³⁺) affinity chromatography.** *Anal Biochem* 1986, **154**(1):250-254.
67. Posewitz MC, Tempst P: **Immobilized gallium(III) affinity chromatography of phosphopeptides.** *Anal Chem* 1999, **71**(14):2883-2892.
68. Pinkse MW, Uitto PM, Hilhorst MJ, Ooms B, Heck AJ: **Selective isolation at the femtomole level of phosphopeptides from proteolytic digests using 2D-NanoLC-ESI-MS/MS and titanium oxide precolumns.** *Anal Chem* 2004, **76**(14):3935-3943.
69. Larsen MR, Thingholm TE, Jensen ON, Roepstorff P, Jorgensen TJ: **Highly selective enrichment of phosphorylated peptides from peptide mixtures using titanium dioxide microcolumns.** *Mol Cell Proteomics* 2005, **4**(7):873-886.
70. Bodenmiller B, Malmstrom J, Gerrits B, Campbell D, Lam H, Schmidt A, Rinner O, Mueller LN, Shannon PT, Pedrioli PG *et al*: **PhosphoPep--a phosphoproteome resource for systems biology research in Drosophila Kc167 cells.** *Mol Syst Biol* 2007, **3**:139.
71. **Abstracts of the 50th ASMS Conference on Mass Spectrometry and Allied Topics. June 2-6, 2002. Orlando, Florida, USA.** *J Am Soc Mass Spectrom* 2002, **13**(5 Suppl):1S-126S.
72. Schroeder MJ, Shabanowitz J, Schwartz JC, Hunt DF, Coon JJ: **A neutral loss activation method for improved phosphopeptide sequence analysis by quadrupole ion trap mass spectrometry.** *Anal Chem* 2004, **76**(13):3590-3598.
73. Olsen JV, Mann M: **Improved peptide identification in proteomics by two consecutive stages of mass spectrometric fragmentation.** *Proc Natl Acad Sci U S A* 2004, **101**(37):13417-13422.
74. Bodenmiller B, Wanka S, Kraft C, Urban J, Campbell D, Pedrioli PG, Gerrits B, Picotti P, Lam H, Vitek O *et al*: **Phosphoproteomic analysis reveals interconnected system-wide responses to perturbations of kinases and phosphatases in yeast.** *Sci Signal* 2010, **3**(153):rs4.
75. Bonaldi T, Straub T, Cox J, Kumar C, Becker PB, Mann M: **Combined use of RNAi and quantitative proteomics to study gene function in Drosophila.** *Mol Cell* 2008, **31**(5):762-772.
76. Corona DF, Siriaco G, Armstrong JA, Snarskaya N, McClymont SA, Scott MP, Tamkun JW: **ISWI regulates higher-order chromatin structure and histone H1 assembly in vivo.** *PLoS Biol* 2007, **5**(9):e232.
77. Hilger M, Bonaldi T, Gnad F, Mann M: **Systems-wide analysis of a phosphatase knock-down by quantitative proteomics and phosphoproteomics.** *Mol Cell Proteomics* 2009, **8**(8):1908-1920.

78. Leifso K, Cohen-Freue G, Dogra N, Murray A, McMaster WR: **Genomic and proteomic expression analysis of Leishmania promastigote and amastigote life stages: the Leishmania genome is constitutively expressed.** *Mol Biochem Parasitol* 2007, **152**(1):35-46.
79. Silverman JM, Chan SK, Robinson DP, Dwyer DM, Nandan D, Foster LJ, Reiner NE: **Proteomic analysis of the secretome of Leishmania donovani.** *Genome Biol* 2008, **9**(2):R35.
80. Silverman JM, Clos J, de'Oliveira CC, Shirvani O, Fang Y, Wang C, Foster LJ, Reiner NE: **An exosome-based secretion pathway is responsible for protein export from Leishmania and communication with macrophages.** *J Cell Sci* 2010, **123**(Pt 6):842-852.
81. Chawla B, Jhingran A, Panigrahi A, Stuart KD, Madhubala R: **Paromomycin affects translation and vesicle-mediated trafficking as revealed by proteomics of paromomycin -susceptible -resistant Leishmania donovani.** *PLoS One* 2011, **6**(10):e26660.
82. Paba J, Santana JM, Teixeira AR, Fontes W, Sousa MV, Ricart CA: **Proteomic analysis of the human pathogen Trypanosoma cruzi.** *Proteomics* 2004, **4**(4):1052-1059.
83. de Godoy LM, Marchini FK, Pavoni DP, Rampazzo Rde C, Probst CM, Goldenberg S, Krieger MA: **Quantitative proteomics of Trypanosoma cruzi during metacyclogenesis.** *Proteomics* 2012, **12**(17):2694-2703.
84. Urbaniak MD, Mathieson T, Bantscheff M, Eberhard D, Grimaldi R, Miranda-Saavedra D, Wyatt P, Ferguson MA, Frearson J, Drewes G: **Chemical proteomic analysis reveals the drugability of the kinome of Trypanosoma brucei.** *ACS Chem Biol* 2012, **7**(11):1858-1865.
85. Urbaniak MD, Martin DM, Ferguson MA: **Global quantitative SILAC phosphoproteomics reveals differential phosphorylation is widespread between the procyclic and bloodstream form lifecycle stages of Trypanosoma brucei.** *J Proteome Res* 2013, **12**(5):2233-2244.
86. Wickstead B, Ersfeld K, Gull K: **Targeting of a tetracycline-inducible expression system to the transcriptionally silent minichromosomes of Trypanosoma brucei.** *Mol Biochem Parasitol* 2002, **125**(1-2):211-216.
87. Brenndorfer M, Boshart M: **Selection of reference genes for mRNA quantification in Trypanosoma brucei.** *Mol Biochem Parasitol* 2010, **172**(1):52-55.
88. Wisniewski JR, Zougman A, Mann M: **Combination of FASP and StageTip-based fractionation allows in-depth analysis of the hippocampal membrane proteome.** *J Proteome Res* 2009, **8**(12):5674-5678.
89. Bodenmiller B, Mueller LN, Mueller M, Domon B, Aebersold R: **Reproducible isolation of distinct, overlapping segments of the phosphoproteome.** *Nat Methods* 2007, **4**(3):231-237.
90. Lundby A, Secher A, Lage K, Nordsborg NB, Dmytriyev A, Lundby C, Olsen JV: **Quantitative maps of protein phosphorylation sites across 14 different rat organs and tissues.** *Nat Commun* 2012, **3**:876.
91. Olsen JV, de Godoy LM, Li G, Macek B, Mortensen P, Pesch R, Makarov A, Lange O, Horning S, Mann M: **Parts per million mass accuracy on an Orbitrap mass spectrometer via lock mass injection into a C-trap.** *Mol Cell Proteomics* 2005, **4**(12):2010-2021.

92. Brun R, Schonenberger: **Cultivation and in vitro cloning or procyclic culture forms of Trypanosoma brucei in a semi-defined medium. Short communication.** *Acta Trop* 1979, **36**(3):289-292.
93. Montalvetti A, Fernandez A, Sanders JM, Ghosh S, Van Brussel E, Oldfield E, Docampo R: **Farnesyl pyrophosphate synthase is an essential enzyme in Trypanosoma brucei. In vitro RNA interference and in vivo inhibition studies.** *J Biol Chem* 2003, **278**(19):17075-17083.
94. Lai DH, Bontempi EJ, Lukes J: **Trypanosoma brucei solanesyl-diphosphate synthase localizes to the mitochondrion.** *Mol Biochem Parasitol* 2012, **183**(2):189-192.
95. Ferella M, Montalvetti A, Rohloff P, Miranda K, Fang J, Reina S, Kawamukai M, Bua J, Nilsson D, Pravia C *et al.*: **A solanesyl-diphosphate synthase localizes in glycosomes of Trypanosoma cruzi.** *J Biol Chem* 2006, **281**(51):39339-39348.
96. Li B, Espinal A, Cross GA: **Trypanosome telomeres are protected by a homologue of mammalian TRF2.** *Mol Cell Biol* 2005, **25**(12):5011-5021.
97. Downey N, Hines JC, Sinha KM, Ray DS: **Mitochondrial DNA ligases of Trypanosoma brucei.** *Eukaryot Cell* 2005, **4**(4):765-774.
98. Papadopoulou B, Roy G, Breton M, Kundig C, Dumas C, Fillion I, Singh AK, Olivier M, Ouellette M: **Reduced infectivity of a Leishmania donovani biopterin transporter genetic mutant and its use as an attenuated strain for vaccination.** *Infect Immun* 2002, **70**(1):62-68.
99. Vertommen D, Van Roy J, Szikora JP, Rider MH, Michels PA, Opperdoes FR: **Differential expression of glycosomal and mitochondrial proteins in the two major life-cycle stages of Trypanosoma brucei.** *Mol Biochem Parasitol* 2008, **158**(2):189-201.
100. Wisniewski JR, Dus K, Mann M: **Proteomic workflow for analysis of archival formalin-fixed and paraffin-embedded clinical samples to a depth of 10 000 proteins.** *Proteomics Clin Appl* 2013, **7**(3-4):225-233.
101. Gilar M, Olivova P, Daly AE, Gebler JC: **Two-dimensional separation of peptides using RP-RP-HPLC system with different pH in first and second separation dimensions.** *J Sep Sci* 2005, **28**(14):1694-1703.
102. Yang F, Shen Y, Camp DG, 2nd, Smith RD: **High-pH reversed-phase chromatography with fraction concatenation for 2D proteomic analysis.** *Expert Rev Proteomics* 2012, **9**(2):129-134.
103. Saas J, Ziegelbauer K, von Haeseler A, Fast B, Boshart M: **A developmentally regulated aconitase related to iron-regulatory protein-1 is localized in the cytoplasm and in the mitochondrion of Trypanosoma brucei.** *J Biol Chem* 2000, **275**(4):2745-2755.
104. Zhang X, Cui J, Nilsson D, Gunasekera K, Chanfon A, Song X, Wang H, Xu Y, Ochsenreiter T: **The Trypanosoma brucei MitoCarta and its regulation and splicing pattern during development.** *Nucleic Acids Res* 2010, **38**(21):7378-7387.
105. Sienkiewicz N, Jaroslowski S, Wyllie S, Fairlamb AH: **Chemical and genetic validation of dihydrofolate reductase-thymidylate synthase as a drug target in African trypanosomes.** *Mol Microbiol* 2008, **69**(2):520-533.
106. Bringaud F, Biteau N, Melville SE, Hez S, El-Sayed NM, Leech V, Berriman M, Hall N, Donelson JE, Baltz T: **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.
107. Rindisbacher L, Hemphill A, Seebeck T: **A repetitive protein from *Trypanosoma brucei* which caps the microtubules at the posterior end of the cytoskeleton.** *Mol Biochem Parasitol* 1993, **58**(1):83-96.
 108. Pao SS, Paulsen IT, Saier MH, Jr.: **Major facilitator superfamily.** *Microbiol Mol Biol Rev* 1998, **62**(1):1-34.
 109. Panigrahi AK, Zikova A, Dalley RA, Acestor N, Ogata Y, Anupama A, Myler PJ, Stuart KD: **Mitochondrial complexes in *Trypanosoma brucei*: a novel complex and a unique oxidoreductase complex.** *Mol Cell Proteomics* 2008, **7**(3):534-545.
 110. Panigrahi AK, Ogata Y, Zikova A, Anupama A, Dalley RA, Acestor N, Myler PJ, Stuart KD: **A comprehensive analysis of *Trypanosoma brucei* mitochondrial proteome.** *Proteomics* 2009, **9**(2):434-450.
 111. Horbinski C, Chu CT: **Kinase signaling cascades in the mitochondrion: a matter of life or death.** *Free Radic Biol Med* 2005, **38**(1):2-11.
 112. Alavi A, Hood JD, Frausto R, Stupack DG, Cheresch DA: **Role of Raf in vascular protection from distinct apoptotic stimuli.** *Science* 2003, **301**(5629):94-96.
 113. Salavati R, Najafabadi HS: **Sequence-based functional annotation: what if most of the genes are unique to a genome?** *Trends Parasitol* 2010, **26**(5):225-229.
 114. Schwede A, Ellis L, Luther J, Carrington M, Stoecklin G, Clayton C: **A role for Caf1 in mRNA deadenylation and decay in trypanosomes and human cells.** *Nucleic Acids Res* 2008, **36**(10):3374-3388.
 115. Li CH, Irmer H, Gudjonsdottir-Planck D, Freese S, Salm H, Haile S, Estevez AM, Clayton C: **Roles of a *Trypanosoma brucei* 5'->3' exoribonuclease homolog in mRNA degradation.** *RNA* 2006, **12**(12):2171-2186.
 116. Bao Y, Weiss LM, Ma YF, Kahn S, Huang H: **Protein kinase A catalytic subunit interacts and phosphorylates members of trans-sialidase superfamily in *Trypanosoma cruzi*.** *Microbes Infect* 2010, **12**(10):716-726.
 117. Lesenechal M, Duret L, Cano MI, Mortara RA, Jolivet M, Camargo ME, da Silveira JF, Paranhos-Baccala G: **Cloning and characterization of a gene encoding a novel immunodominant antigen of *Trypanosoma cruzi*.** *Mol Biochem Parasitol* 1997, **87**(2):193-204.
 118. Mikulik K, Janda I: **Protein kinase associated with ribosomes phosphorylates ribosomal proteins of *Streptomyces collinus*.** *Biochem Biophys Res Commun* 1997, **238**(2):370-376.
 119. Anantharaman V, Aravind L: **Novel conserved domains in proteins with predicted roles in eukaryotic cell-cycle regulation, decapping and RNA stability.** *BMC Genomics* 2004, **5**(1):45.
 120. Tanaka KJ, Ogawa K, Takagi M, Imamoto N, Matsumoto K, Tsujimoto M: **RAP55, a cytoplasmic mRNP component, represses translation in *Xenopus oocytes*.** *J Biol Chem* 2006, **281**(52):40096-40106.
 121. Hashimi H, Cicova Z, Novotna L, Wen YZ, Lukes J: **Kinetoplastid guide RNA biogenesis is dependent on subunits of the mitochondrial RNA binding complex 1 and mitochondrial RNA polymerase.** *RNA* 2009, **15**(4):588-599.

122. Li F, Herrera J, Zhou S, Maslov DA, Simpson L: **Trypanosome REH1 is an RNA helicase involved with the 3'-5' polarity of multiple gRNA-guided uridine insertion/deletion RNA editing.** *Proc Natl Acad Sci U S A* 2011, **108**(9):3542-3547.
123. Tanoue T, Nishida E: **Molecular recognitions in the MAP kinase cascades.** *Cell Signal* 2003, **15**(5):455-462.
124. Bardwell L: **Mechanisms of MAPK signalling specificity.** *Biochem Soc Trans* 2006, **34**(Pt 5):837-841.
125. Choi KY, Satterberg B, Lyons DM, Elion EA: **Ste5 tethers multiple protein kinases in the MAP kinase cascade required for mating in *S. cerevisiae*.** *Cell* 1994, **78**(3):499-512.
126. Kugeratski, FG. **Identificação dos Alvos e Vias Reguladas pelas Proteínas Ste7 de *Trypanosoma brucei*.** Dissertação de mestrado, Programa de Pós Graduação em Biociências e Biotecnologia do Instituto Carlos Chagas - FIOCRUZ/PR, 2012.
127. Vermelho AB, Melo ACN, Soares RA, Alviano DS, Souza EP, Souto-Padrón T, Rodrigues GC, Aguiar AP, Pereira MC, Ferreira-Pereira A, Rosa MSS, Meirelles MNL, Alviano CS. **Trypanosoma cruzi Peptidases: An Overview.** *The Open Parasitology Journal* 2010, 4, 120-131.