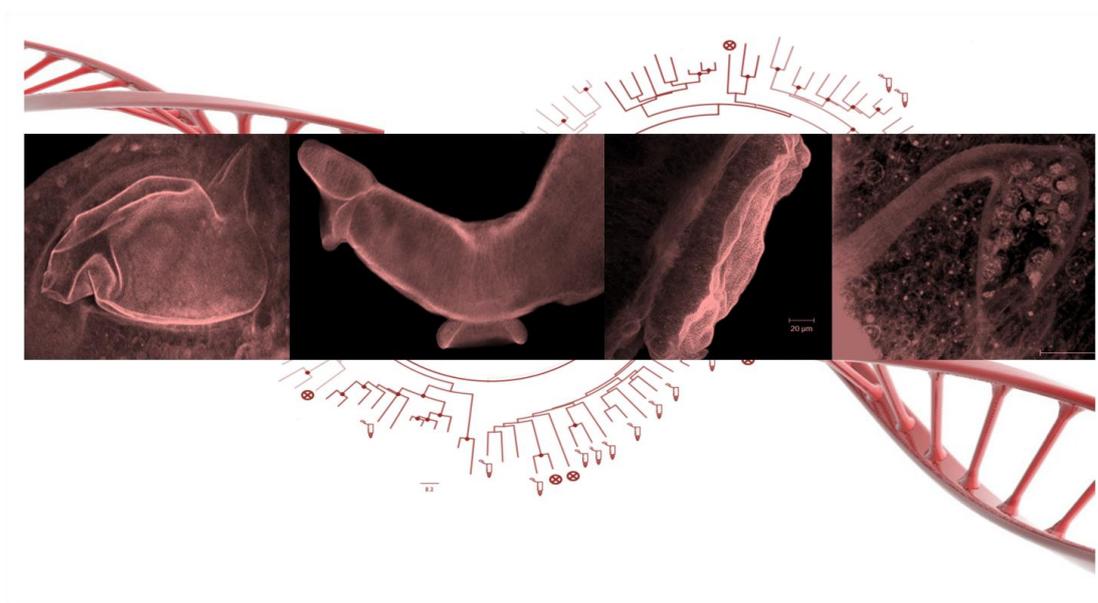


**Ministério da Saúde**  
**Fundação Oswaldo Cruz**  
**Centro de Pesquisas René Rachou**  
**Programa de Pós-graduação em Ciências da Saúde**

**Proteínas quinase eucarióticas (ePKs) de *Schistosoma*:  
identificação, anotação funcional e seleção de potenciais alvos  
terapêuticos**



por

**Luiza Freire de Andrade**

Orientador: Dr. Guilherme Oliveira

**Grupo de Genômica e Biologia Computacional**

Belo Horizonte, Minas Gerais – Brasil

08/2012

**TESE DBCM-CPqRR**

**L.F. ANDRADE**

**2012**

**Ministério da Saúde**  
**Fundação Oswaldo Cruz**  
**Centro de Pesquisas René Rachou**  
**Programa de Pós-graduação em Ciências da Saúde**

Proteínas quinase eucarióticas (ePKs) de *Schistosoma*: identificação,  
anotação funcional e seleção de potenciais alvos terapêuticos

por

**Luiza Freire de Andrade**

Orientador: Dr. Guilherme Oliveira

**Tese apresentada com vistas à obtenção  
do Título de Doutor em Ciências na área de  
concentração Biologia Celular e Molecular**

**Orientação: Guilherme Corrêa Oliveira**

Belo Horizonte, Minas Gerais – Brasil

08/2012

Catálogo-na-fonte  
Rede de Bibliotecas da FIOCRUZ  
Biblioteca do CPqRR  
Segemar Oliveira Magalhães CRB/6 1975

A553p

ANDRADE, LUIZA FREIRE DE

2012

Proteínas quinase eucarióticas (ePKs) de *Schistosoma*: identificação, anotação funcional e seleção de potenciais alvos terapêuticos / Luiza Freire de Andrade. – Belo Horizonte, 2012.

XVIII, 125 f: il.; 210 x 297mm.

Bibliografia: f. 120 - 143

Tese (doutorado) – Tese para obtenção do título de Doutor em Ciências pelo Programa de Pós-Graduação em Ciências da Saúde do Centro de Pesquisas René Rachou. Área de concentração: Biologia Celular e Molecular.

1. Esquistossomose/Quimioterapia 2. *Schistosoma*/Genética  
3. Proteínas Quinases/Análise I. Título. II. Oliveira, Guilherme (Orientação)

CDD – 22. ed. – 616.963

**Ministério da Saúde**  
**Fundação Oswaldo Cruz**  
**Centro de Pesquisas René Rachou**  
**Programa de Pós-graduação em Ciências da Saúde**

Proteínas quinase eucarióticas (ePKs) de *Schistosoma*: identificação,  
anotação funcional e seleção de potenciais alvos terapêuticos

por

**Luiza Freire de Andrade**

**Foi avaliada pela banca examinadora composta pelos seguintes membros:**

**Prof. Dr. Guilherme Corrêa Oliveira (Presidente)**  
**Prof. Dra. Diana Bahia**  
**Prof. Dra. Jaqueline Germano de Oliveira**  
**Prof. Dra. Cristiana Alves Brito**  
**Prof. Dr. Carlos Graeff**  
**Suplente: Aristóteles Goés Neto**

**Tese defendida e aprovada em: 17/08/2012**

## **AGRADECIMENTOS**

Os meus sinceros agradecimentos,

Ao Guilherme, meu orientador, por todos os ensinamentos, incentivo, pela confiança e pelas diversas oportunidades de crescimento. Sinto-me preparada para dar continuidade com meus projetos pessoais/profissionais e serei eternamente grata por tudo isso. Espero poder contribuir sempre para o crescimento e sucesso do grupo.

A Ângela por todo carinho, dedicação, incentivo e admiração. Pode contar comigo sempre!

A Marina pelas horas e mais horas de convivência, pelos ensinamentos e por todo incentivo para a conclusão desse trabalho. Muito obrigada pela orientação durante toda a parte experimental.

Laila, Jerônimo, Adhemar, Larissa, Lívia e Eric pela valiosa e essencial contribuição e pelas horas dedicadas para me ajudar na conclusão desse e de outros projetos.

À Marina e a Larissa agradeço também pelas sugestões e correções durante a redação dessa tese.

Juliana e Fernanda, minhas primeiras alunas, agradeço pela ajuda, pela força e pela amizade. Espero ter contribuído para o amadurecimento profissional de vocês. Para mim foi um aprendizado incrível.

A todos do LPCM: as secretárias, a Kenia (pela ajuda na organização e limpeza), as meninas da qualidade (pela difícil tarefa atribuída a vocês), aos amigos que fiz (algumas praticamente minhas irmãs - Fê e Mari) e a todos que contribuíram de alguma forma para a conclusão desse trabalho.

Ao Cebio pela excelente infraestrutura computacional e seus colaboradores por todo o apoio e amizade.

Ao Centro de Pesquisas René Rachou pela estrutura disponibilizada, sempre visando o sucesso de seus colaboradores.

Agradeço as secretárias Carolina Lara e Michelle Silva pela atenção e disponibilidade.

Aos colegas, professores e secretárias do Programa de Pós-Graduação em Ciências da Saúde do Centro de Pesquisas René Rachou.

Agradeço as Plataformas de Sequenciamento, de PCR em tempo Real e Proteômica do Programa de Desenvolvimento Tecnológico em Insumos para a Saúde (PDTIS), em especial a Elis, Fernanda Barbosa, Rosiane e Ramon pelos inúmeros auxílios prestados.

Ao moluscário do CPqRR e sua equipe por todo apoio e disponibilidade em ajudar.

À Equipe do Biotério e Central de Esterilização pela ajuda e colaboração sempre que necessário. Ao Guilherme Silva (LICM) e a Dra. Neusa Araujo Pereira (LESQ) pela ajuda essencial durante a infecção e perfusão dos camundongos. Muito Obrigada!

À Biblioteca do CPqRR em prover acesso gratuito local e remoto à informação técnico-científica em saúde custeada com recursos públicos federais, integrante do rol de referências desta tese, também pela catalogação da mesma.

Ao Dr. Conor Caffrey por ter me recebido e me supervisionado no período que estive na UCSF. Ao Dr. Jim McKerrow por ter autorizado a minha ida ao *Sandler Center for Basic research in parasitic diseases*. E ao Brian Suzuki por toda ajuda.

Ao Dr. José Roberto, Dra. Renata Neves e Frederico (UERJ) pela ajuda com as análises morfológicas do *Schistosoma* e pelos ensinamentos fundamentais com relação às estruturas do parasito que garantiram a finalização desse trabalho.

Ao Dr. Paulo Pimenta por ter permitido o uso do aparelho confocal do seu laboratório e ao Rafael Pimenta pela ajuda na manipulação do equipamento e pelas belas imagens registradas.

A banca examinadora por ter aceitado o convite para avaliar e ajudar a finalizar esse trabalho

Às Agências Financiadoras que contribuíram para a realização deste trabalho: National Institutes of Health - NIH/Fogarty International Center, European Community, SEtTReND, CNPq, FAPEMIG e Centro de Pesquisas René Rachou pela bolsa de doutorado.

A minha família: minha mãe, minha irmã, meu avos, meus tios e ao Cacau, Aninha, Serginho e Rê por todo amor, carinho, torcida, admiração e apoio em todos os momentos.

## **APOIO FINANCEIRO**

- ✓ National Institutes of Health - NIH/Fogarty International Center (D43 TW007012);
- ✓ European Community (FP7/2007-2013 FP7/2007-2011);
- ✓ Secretaria de Ciência, Tecnologia e Ensino Superior de Minas Gerais, Fundação de Amparo à Pesquisa do Estado de Minas Gerais, SECTES/FAPEMIG, Brazil (CBB-1181/08 and PPM-00439-10);
- ✓ Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq (CNPq-Universal 476036/2010-0, CNPq-INCT 573839/2008-5).

## SÚMARIO

LISTA DE FIGURAS .....	XI
LISTA DE TABELAS .....	XIII
LISTA DE QUADROS .....	XIII
LISTA DE ABREVIATURAS.....	XIV
RESUMO .....	XVII
ABSTRACT .....	XVIII
1. INTRODUÇÃO.....	19
1.1- A esquistossomose.....	19
1.2- <i>Schistosoma</i> .....	21
1.3- O genoma de três espécies de <i>Schistosoma</i> .....	24
1.4- Era Pós-Genômica .....	27
1.4-1. Anotação Funcional.....	27
1.4-2. Inferência Filogenética para classificação funcional de genes.....	28
1.4-3. O <i>Schistosoma</i> na era Pós-Genômica .....	31
1.4-4. Interferência por RNA (RNAi).....	32
1.5- Sinalização celular e o seu possível papel no desenvolvimento do <i>Schistosoma</i> .....	34
1.6- Proteínas quinase .....	35
2. JUSTIFICATIVA .....	38
3. OBJETIVOS.....	40
3.1- Objetivo geral .....	40
3.2- Objetivos específicos.....	40
4. MATERIAL E MÉTODO .....	41
4.1- Identificação, classificação e anotação funcional das proteínas quinase do <i>S. mansoni</i> , <i>S. haematobium</i> e <i>S. japonicum</i> .....	41
4.1-1. Organismos e Sequências .....	41

4.1-2.	Identificação das ePKs e proteínas relacionadas .....	41
4.1-3.	Classificação Funcional .....	42
4.1-4.	Inferência Filogenética .....	43
4.2-	Interferência por RNA (RNAi).....	44
4.2-1.	Seleção dos genes alvos.....	44
4.2-2.	Desenho dos iniciadores .....	44
4.2-3.	Obtenção dos parasitos .....	45
4.2-4.	Extração de RNA de esquistossômulos .....	46
4.2-5.	Síntese de cDNA .....	47
4.2-6.	Amplificação e sequenciamento.....	47
4.2-7.	Síntese de dsRNA.....	49
4.2-8.	Exposição dos parasitos aos dsRNAs.....	49
4.2-9.	Análise do silenciamento no nível da transcrição – QPCR.....	49
4.2-10.	Experimentos <i>in vivo</i> : infecção de camundongos com esquistossômulos silenciados .....	51
4.2-11.	Estudos morfométricos e morfológicos de vermes adultos de <i>S. mansoni</i> recuperados após perfusão de camundongos infectados com esquistossômulos silenciados .....	52
5.	RESULTADOS E DISCUSSÃO .....	53
5.1-	Proteínas quinase eucarióticas de três espécies <i>Schistosoma</i> .....	53
5.2-	Classificação das ePKs identificadas em <i>S. mansoni</i> , <i>S. haematobium</i> e <i>S. japonicum</i> .....	57
5.3-	Proteínas Serina/Treonina Quinase .....	61
5.3-1.	Análise das ePKs do grupo AGC encontradas em <i>Schistosoma</i> .....	61
5.3-2.	Análise das ePKs do grupo CaMK encontradas em <i>Schistosoma</i> .....	63
5.3-3.	Análise das ePKs do grupo CMGC encontradas em <i>Schistosoma</i> .....	65
5.3-4.	Análise das ePKs do grupo STE encontradas em <i>Schistosoma</i> .....	71

5.3-5.	Análise das ePKs do grupo CK1 encontradas em <i>Schistosoma</i> .....	71
5.4-	Proteínas tirosina quinase .....	72
5.4-1.	Análise das ePKs do grupo TK encontradas em <i>Schistosoma</i> .....	72
5.4-2.	Análise das ePKs do grupo RGC encontradas em <i>Schistosoma</i> .....	77
5.5-	Proteínas quinase híbridas .....	78
5.5-1.	Análise das ePKs do grupo TKL encontradas em <i>Schistosoma</i> .....	78
5.5-2.	Análise das ePKs do grupo <i>Other</i> encontradas em <i>Schistosoma</i> .....	81
5.6-	Domínios acessórios encontrados nas ePKs de <i>Schistosoma</i> .....	82
5.7-	Investigação do papel biológico de proteínas envolvidas na via de sinalização MAPK.....	87
5.7-1.	Proteínas selecionadas .....	87
5.7-2.	Síntese de dsRNAs e análise fenotípica de esquistossômulos.....	89
5.7-3.	Estudo do nível de transcrito em esquistossômulos tratados com dsRNA.....	92
5.7-4.	Testes <i>in vivo</i> : infecção de camundongos com esquistossômulos silenciados .....	94
5.7-5.	Efeito do silenciamento na transcrição de genes alvos da via MAPK .....	102
6.	CONCLUSÕES .....	107
7.	ANEXOS .....	108
7.1-	Tabelas.....	108
7.2-	Figuras .....	111
7.3-	Artigo Publicado em 2009 – NATURE.....	119
7.4-	Revisão Publicada em 2009 - CURRENT OPINION IN MICROBIOLOGY .....	119
7.5-	Artigo Publicado em 2011 - BMC GENOMICS.....	119
7.6-	Revisão Publicada em 2011 – JOURNAL OF SIGNAL TRANSDUCTION .....	119
7.7-	Artigo Publicado em 2012 – CURRENT PHARMACEUTICAL DESIGN.....	119
8.	REFERÊNCIAS BIBLIOGRÁFICAS .....	120

## LISTA DE FIGURAS

Figura 1: Distribuição global da esquistossomose. ....	20
Figura 2 - O <i>Schistosoma</i> . ....	21
Figura 3 - Ciclo biológico do <i>Schistosoma</i> (15). ....	23
Figura 4: A árvore filogenética. ....	30
Figura 5: Ilustração do mecanismo de RNAi. ....	33
Figura 6: Fosforilação reversível de proteínas quinase ....	36
Figura 7: Porcentagem de ePKs identificadas em diferentes organismos. ....	54
Figura 8: Relação de ortologia das ePKs entre três espécies de <i>Schistosoma</i> . ....	56
Figura 9: Esquema de classificação para as ePKs de <i>S. mansoni</i> , <i>S. haematobium</i> e <i>S. japonicum</i> . ....	58
Figura 10: Distribuição das ePKs de <i>S. mansoni</i> , <i>S. haematobium</i> e <i>S. japonicum</i> em seus grupos. ....	59
Figura 11: Análise dos resíduos conservados no subdomínio VIIb no domínio catalítico de ePKs do grupo RGC de <i>S. mansoni</i> , <i>S. haematobium</i> e <i>S. japonicum</i> . ....	61
Figura 12: Via de sinalização MAPK caracterizada em mamíferos e predita em <i>S. mansoni</i> . ....	70
Figura 13: Análise filogenética dos grupos ePKs parálogos de <i>S. mansoni</i> . ....	79
Figura 14: Representação da organização dos domínios proteicos de algumas ePKs dos grupos AGC, CaMK, CK1, CMGC e STE de <i>S. mansoni</i> . ....	85
Figura 15: Representação da organização dos domínios proteicos de algumas ePKs dos grupos TKL, Other, TK e RGC de <i>S. mansoni</i> . ....	86
Figura 16: Análise da síntese de dsRNA para os alvos SmCaMK2, SmJNK, SmERK-1, SmERK-2, SmRas e o controle inespecífico GFP. ....	90
Figura 17: Imagens dos esquistossômulos tratados com dsRNA SmCaMK2 no sétimo dia de cultivo. ....	91
Figura 18: Níveis de transcritos dos genes alvos em relação ao controle após tratamento com dsRNA durante 2, 4 ou 7 dias. ....	94

Figura 19: Níveis de transcritos de SmCaMK2, SmJNK e SmERK-1 dos esquistossômulos silenciados utilizados na infecção de camundongos. ....	95
Figura 20: Comparação do número de vermes adultos obtidos após perfusão de camundongos infectados com esquistossômulos silenciados em relação aos controles GFP e branco ( <i>negative control</i> ). ....	96
Figura 21: Imagens microscopia confocal a laser de vermes adultos macho de <i>S. mansoni</i> ....	97
Figura 22: Imagens microscopia confocal a laser de vermes adultos fêmea de <i>S. mansoni</i> . ...	98
Figura 23: Imagens de microscopia confocal a laser de vermes adultos macho (A, C e E) e fêmea (B, D e F) de <i>S. mansoni</i> tratados com dsRNA CaMK2. ....	99
Figura 24: Comparação do número de ovos no fígado obtidos após perfusão de camundongos infectados com esquistossômulos silenciados, em relação aos controles GFP e branco ( <i>negative control</i> ). ....	100
Figura 25: Imagens de microscopia confocal a laser de vermes adultos macho de <i>S. mansoni</i> . ....	101
Figura 26: Imagens de microscopia confocal a laser de vermes adultos fêmea de <i>S. mansoni</i> . ....	102
Figura 27: Níveis de transcritos dos genes SmCaMK2, SmJNK, SmRas, SmERK, SmSRF e SmC-fos em relação ao controle após tratamento com dsRNAs dsCaMK2, dsJNK em A e, dsRas e dsERK1 em B, durante 2, 4 ou 7 dias. ....	104
Figura 28: Via de sinalização hipotética MAPK/ERK ativada por EGF que leva a regulação do gene C-fos em <i>S. mansoni</i> . ....	106
Figura Anexo 1: Árvore filogenética baseada no método de distância para o grupo AGC. ...	111
Figura Anexo 2: Árvore filogenética baseada no método de distância do grupo CaMK.....	112
Figura Anexo 3: Árvore filogenética baseada no método de distância do grupo CMGC.....	113
Figura Anexo 4: Árvore filogenética baseada no método de distância do grupo STE.....	114
Figura Anexo 5: Árvore filogenética baseada no método de distância do grupo CK1. ....	115
Figura Anexo 6: Árvore filogenética baseada no método de distância do grupo TK. ....	116
Figura Anexo 7: Árvore filogenética baseada no método de distância do grupo RGC.....	117

Figura Anexo 8: Árvore filogenética baseada no método de distância do grupo TKL. .... 118

## LISTA DE TABELAS

Tabela 1: Comparação das características genômicas de três espécies de *Schistosoma* ..... 26

Tabela 2: Classificação das proteínas quinase (125) ..... 37

Tabela 3: Distribuição das ePKs em seus respectivos grupos para *S. mansoni*, *S. haematobium*, *S. japonicum* e organismos modelo ..... 60

Tabela 4: Relação de ortologia entre ePKs de *S. mansoni*, *S. haematobium*, *S. japonicum* e *C. elegans* e o fenótipo encontrado após silenciamento gênico em *C. elegans* (73). ..... 75

Tabela 5: ePKs de *S. mansoni* selecionadas como possíveis alvos para o desenvolvimento de droga baseado na importância funcional de seus genes e/ou dos ortólogos identificados nesse trabalho ..... 89

Tabela Anexo 1: Proteínas quinase eucarióticas do *Schistosoma mansoni* ..... 108

Tabela Anexo 2: Proteínas quinase eucarióticas ortólogas entre *S. mansoni*, *S. haematobium* e *S. japonicum* ..... 108

Tabela Anexo 3: Proteínas quinase eucarióticas do *Schistosoma haematobium* ..... 108

Tabela Anexo 4: Proteínas quinase eucarióticas do *Schistosoma japonicum* ..... 108

Tabela Anexo 5: Proteínas quinase eucarióticas sem ortólogos entre *S. mansoni* (Smp), *S. haematobium* (Sha) e *S. japonicum* (Sjp) ..... 109

Tabela Anexo 6: Sequências ePKs não preditas na versão atual (3.0) do genoma de *S. mansoni* ..... 110

Tabela Anexo 7: Sequências ePKs que tiveram o número de acesso modificado na versão atual (3.0) do genoma de *S. mansoni* ..... 110

## LISTA DE QUADROS

Quadro 1: Lista de iniciadores ..... 45

Quadro 2: Distribuição das principais famílias ePKs em *Schistosoma* e organismos modelo. 66

## LISTA DE ABREVIATURAS

µg	Microgramas
µL	Microlitro
aa	Aminoácido
Abl	Abelson Murine Leukemia Homolog
Ack	Activated Cdc42-Associated Tyrosine Kinase
AGC	Camp-Dependent Protein Kinase/Protein Kinase G/Protein Kinase C Extended
ATP	Adenosina Trifosfato
AUR	Aurora Kinase
Blast	Basic Local Alignment Search Tool
CAMK	Calcium/Calmodulin Regulated Kinase
CAMK1	CAMK Family 1
CAMK2	CAMK Family 2
CAMKL	Camk Like
CCK4	Colon Carcinoma Kinase 4
CD	Common Docking Domain
CDC7	Cell Division Control 7
CDD	Conserved Domain Database
CDK	Cyclin-Dependent Kinases
CDKL	Cyclin Dependent Kinase Like
cDNA	DNA Complementar.
CK1	Cell Kinase I
CK2	Cell Kinase 2
CLK	Cdc-Like Kinase
CMGC	Cyclin-Dependent- (CDK), Mitogen-Activated- (MAPK), Glycogen-Synthase- (GSK) And CDK-Like Kinases
DAPK	Death Associated Protein Kinase
DEPC	Dietilpirocarbonato
DMPK	Myotonic Dystrophy Protein Kinase
DNA	Ácido Desoxirribonucléico
DNase	Desoxirribonuclease.
dNTP	Desoxirribonucleosídeo Trifosfato
DTT	Dithioheitol
DYRK	Dual-Specificity Y (Tyrosine) Regulated Kinase
EDTA	Etileno Diamino Tetra
EGFR	Epidermal Growth Factor Receptor
Eph	Ephrin Receptors
ERK	Extracellular Signal-Regulated Protein Kinase
EST	Expressed Sequence Tags
Fak	Focal Adhesion Kinase
FGFR	Fibroblast Growth Factor Receptor
Fps	Sarcoma Aviário
GDP	Guanosina Difosfato
GO	Gene Ontology, Ontologia Gênica.
GRK	G-Protein Coupled Receptor Kinase
GTP	Guanosine Triphosphate

HMM	Hidden Markov Model
IL	Interleucina.
InsR	Insulin Receptor
IPTG	Isopropil $\beta$ -D-1-tiogalactopiranosideo
IRE	Inositol Requiring
JNK	C-Jun N-Terminal Kinase
Kb	Kilobase
kDa	Kilodalton
LB	Meio Luria Bertani
LISK	Family Containing Closely Related TESK and LIMK Sub-Families
LRRK	Leucine Rich Repeat Kinase
MAPK	Mitogen Activated Protein Kinase
MAPKAPK	Map Kinase Associated Protein Kinase
MAPKK	Mitogen Activated Protein Kinase Kinase
MAPKK ou MEK	Mitogen Activated Protein Kinase Kinase
MAPKKK	Mitogen Activated Protein Kinase Kinase Kinase
MAST	Microtubule Associated Serine/Threonine Kinase
Mb	Megabase
MKP	Map Kinase Phosphatase
mL	Mililitros
MLCK	Myosin Light Chain Kinases
MLK	Miked Lineage Kinases
mM	Milimolar
mRNA	RNA mensageiro.
Musk	Muscle-Specific Kinase
NAK	Numb-Associated Kinase
NDR	Nuclear Dbf2-Related Kinases
NEK	Mitotic Kinase Family
NKF2	New Kinase Family 2
OD	Densidade Ótica.
ORF	Open Read Frame
p38	Proteína Quinase P38
pb	Pares de Base.
PCR	Polimerase Chain Reaction
PDK1	Phosphatidylinositol 3,4,5-Triphosphate-Dependent Protein Kinase 1
pH	Potencial Hidrogniónico
PK	Protein Kinase
PKA	Protein Kinase A
PKC	Protein Kinase C
PKG	Protein Kinase G
PKN	Protein Kinase N
PLK	Polo Like Kinases
PM	Peso Molecular
RGC	Receptor Guanylate Cyclases
RNA	Ácido Ribonucleico.

RNase	Ribonuclease
Ror	Regeneron Orphan Receptors
rpm	Rotação por Minuto
RSK	Ribosomal S6 Kinase
RTK	Receptor Tyrosina kinase
SDS	Sódio Dodecil Sulfato
SGK	Serum And Glucocorticoid Resposive Kinase
SRPK	SR Protein Kinase; Phosphorylates SR Splicing Factors
STE	MAP Kinase Cascade Kinases
STE11	MAP3K (MAP Kinase Kinase Kinase) Genes
STE20	MAP4K (MAP Kinase Kinase Kinase Kinase) Genes
STE7	MAP2K (MAP Kinase Kinase) Genes
STKR	Serine/Threonine Kinase Receptors
TBCK	TBC Domain-Containing Kinase
TBE	Tris Borato EDTA.
TGF-b	Transforming Growth Factor Beta
TK	Tyrosine Kinase
TKL	Tyrosine Kinase Like
Tris	Tris hidroximetil amino metano
TTBK	Tau Tubulin Kinase
ULK	Unc-51 Like Kinase
VRK	Vaccinia Related Kinase
Wee	Homologs Of Yeast Wee1

## RESUMO

Apesar do praziquantel ser uma droga eficiente para o tratamento da esquistossomose, a prevalência da doença não mostrou redução significativa nos últimos anos e, até o momento, não existe uma alternativa eficaz para o tratamento dessa doença. Dessa forma, optamos por utilizar as poderosas ferramentas genômicas para identificar potenciais alvos para o desenvolvimento de um medicamento alternativo. Já que proteínas quinase eucarióticas (ePKs) são consideradas alvos para o desenvolvimento de drogas do ponto de vista médico e químico, e um número crescente de inibidores ePKs foram desenvolvidos e aprovados para o tratamento de diferentes doenças humanas, estas se tornaram o foco de estudo desse trabalho. As ePKs de *S. mansoni*, *S. japonicum* e *S. haematobium* foram identificadas nos proteomas preditos e classificadas em seus devidos grupos, famílias e subfamílias a partir de abordagens filogenéticas. Utilizando as informações dos ortólogos identificados, foi possível selecionar um grupo de ePKs com função predita essencial nesse parasito. A anotação funcional mostrou ainda que grande parte das ePKs selecionadas são ativadoras/efetoras da via de sinalização MAPK. Dessa forma, proteínas chave da via MAPK (SmRas, SmERK1, SmERK2, SmJNK e SmCaMK2), foram as escolhidas para validação experimental. Após redução significativa no nível de transcrito dos genes selecionados, nenhuma alteração fenotípica visível foi relatada em cultura de esquistossomulos. Contudo, o efeito da diminuição transcricional dos genes no desenvolvimento dos vermes diante do sistema imune do hospedeiro foi avaliado. Evidenciamos que proteínas MAPK JNK quando silenciada causa efeitos devastadores no tegumento de vermes adultos de *S. mansoni* que leva a morte dos mesmos. E, ePKs da subfamília ERK1 estão relacionadas com a produção de ovos, já que fêmeas com baixos níveis de transcritos SmERK1 e SmERK2 apresentam ovários pouco desenvolvidos e produção de ovos significativamente baixa. Além disso, foi comprovado que o fator de transcrição c-fos está diferencialmente expresso em parasitos silenciados para as proteínas MAPK SmJNK, SmCaMK2 e SmERK1/2. Dessa forma concluímos que o dado genômico, acoplado a ferramentas computacionais preditoras e abordagem experimental, compõem uma metodologia poderosa para o estudo dessa espécie. As proteínas MAPK, SmERK e SmJNK, são alvos de interesse para o desenvolvimento de drogas para tratamento da esquistossomose já que um inibidor contra essas proteínas provavelmente irá interromper o ciclo de vida de *Schistosoma* e impedir o progresso da doença.

## ABSTRACT

Although praziquantel is an efficacious drug for schistosomiasis treatment, in recent years problems of resistance have arisen and alternatives don't exist so far. To contribute to a solution, the important information of the recently sequenced genomes of these parasites was used to identify potential targets for the development of an alternative drug. Since eukaryotic protein kinases (ePKs) are good chemical and medical targets for drug development and an increasing number of ePK inhibitors have been approved for the treatment of different human disease, they have become the focus of this study. The ePKs were identified in the predicted proteomes of *S. mansoni*, *S. japonicum* and *S. haematobium* by Hidden Markov Model searches, the genes were classified in their group, family and subfamily by phylogeny approaches, and some ePKs were selected since they have essential function in the parasite. The functional annotation showed that the most selected ePKs are activators/effectors of MAPK signaling pathway, and key proteins of this pathway (SmRas, SmERK1, SmERK2, SmJNK, and SmCaMK2), were chosen for experimental validation. Gene silencing by RNA interference was used to elucidate the functional role of MAPK signaling pathways proteins in *S. mansoni*. After a significant reduction in the transcription level of the selected genes, no visible phenotypic changes were reported in schistosomula in culture. Therefore, mice were infected with the silenced schistosomulas and the development of adult worms was observed. It was shown that SmJNK has an important role in transformation and survival of the parasites as low number of adult worms was recovered and the tegument of survived worms was damaged. Moreover, SmERK1/SmERK2 was related to egg production, as mice infected with silenced schistosomulas, displayed significantly lower egg production and the recovered female worms had underdeveloped ovaries. Furthermore, it was shown that the c-fos transcription factor was overexpressed in parasites with low expression of SmERK1, SmJNK and SmCaMK2. We conclude that the genome contains valuable information and the functional annotation performed had an important role to improve the annotation of *S. mansoni*, *S. haematobium* and *S. japonicum* ePKs. Furthermore, MAPKs proteins, SmERK and SmJNK, are excellent targets for drug development as an inhibitor against these proteins will probably disrupt the life cycle of *Schistosoma* and prevent disease progression.

# 1. INTRODUÇÃO

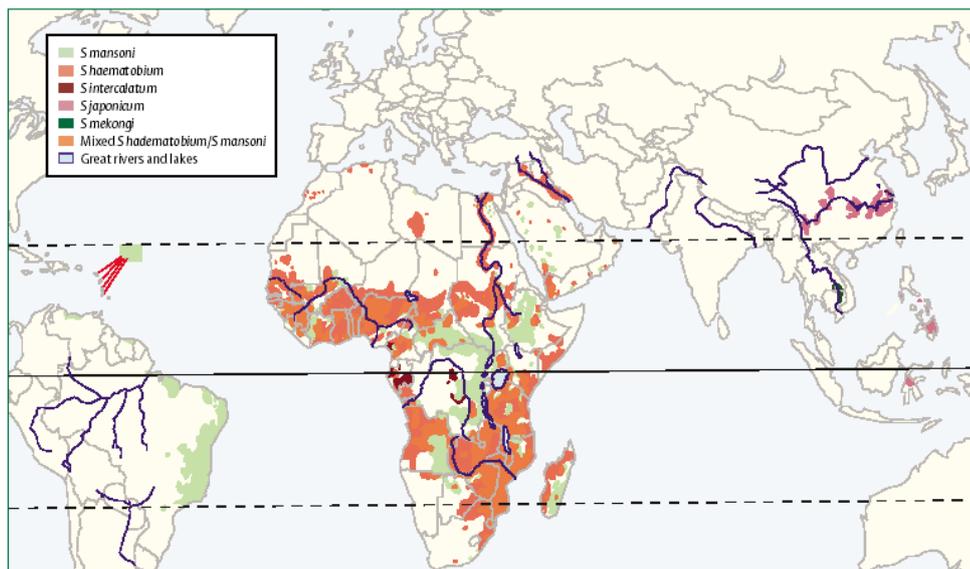
## 1.1- A esquistossomose

A esquistossomose é uma parasitose causada por vermes trematódeos do gênero *Schistosoma* e é considerada pela Organização Mundial de Saúde (OMS) a segunda doença parasitária mais socioeconomicamente prejudicial do mundo. As três principais espécies que parasitam o homem são: *Schistosoma mansoni*, *S. japonicum* e *S. haematobium*. Juntas estas espécies infectam cronicamente mais de 200 milhões de pessoas em 76 países e territórios localizados nas regiões tropicais e subtropicais (**Figura 1**) levando 200 mil pessoas a morte anualmente. A transmissão da esquistossomose é influenciada por fatores ambientais como as precárias instalações sanitárias (1;2), distribuição e presença do hospedeiro intermediário (moluscos do gênero *Biomphalaria*, *Oncomelania* e *Bulinus*, respectivamente para *S. mansoni*, *S. japonicum* e *S. haematobium*) e, além disso, por fatores intrínsecos ao hospedeiro humano, tais como, idade, contato com a água contaminada, genética, resposta imunológica e fisiológica (3). Novos focos da doença surgem principalmente em locais onde são realizados projetos de redirecionamento da água para a indústria e agricultura, ampliando as áreas alagadas e, portanto, o habitat dos hospedeiros intermediários. O posterior aumento da concentração de pessoas nestas áreas favorece a disseminação da doença, sendo que hoje em dia, 700 milhões de pessoas correm o risco de contrair a infecção (4).

A espécie *S. mansoni* ocorre principalmente na África, América Latina e Antilhas (**Figura 1**) (5). A infecção causada por esta espécie é denominada esquistossomose mansônica ou intestinal, pela localização dos parasitos nas vênulas da parede do intestino grosso, sigmóide e reto, com sintomas predominantemente intestinais. Nos casos mais graves, há envolvimento hepatoesplênico e hipertensão no sistema porta (6;7). Por sua vez, a espécie *S. japonicum* também é responsável por outra morbidade intestinal da doença, porém circunscrita ao extremo oriente e Pacífico ocidental, e é conhecida por esquistossomose japônica (6;7).

A espécie *S. haematobium* ocorre predominantemente na África, sua presença estende-se para a Bacia do Mediterrâneo, Oriente Próximo e Médio (**Figura 1**) (8). A doença causada por esta espécie é conhecida por esquistossomose hematóbica, vesical ou urinária e também

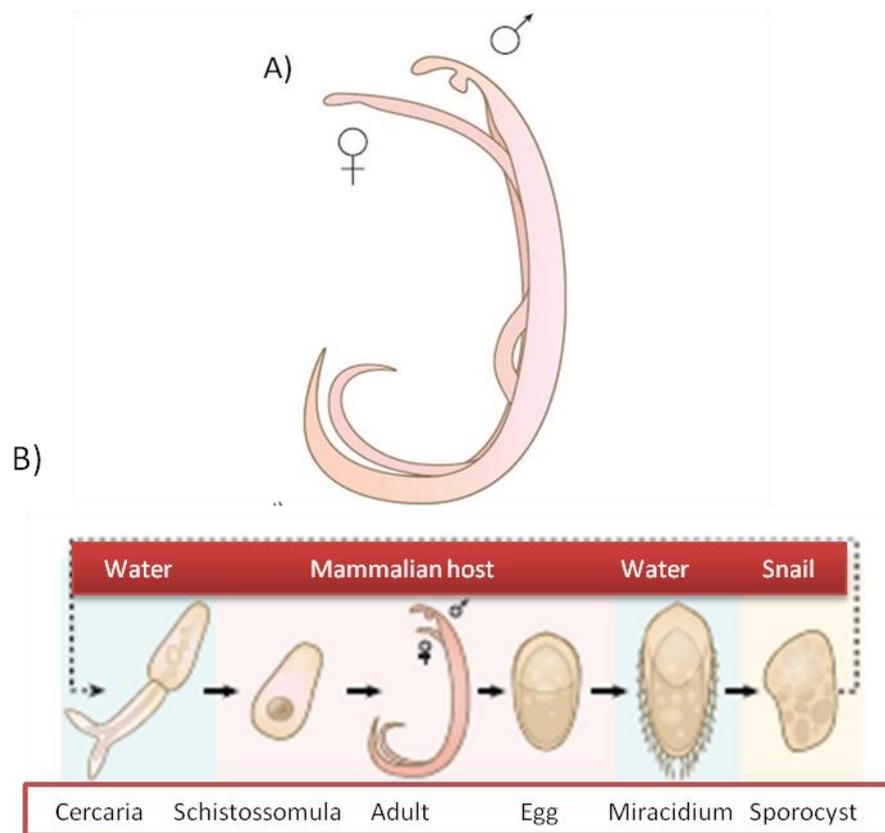
por genito-urinária. Os parasitos localizam-se preferencialmente no plexo vesical e a doença produz um quadro clínico com sintomas urinários (6;7).



**Figura 1: Distribuição global da esquistossomose.** *Schistosoma mansoni* (verde claro) é encontrado em grande quantidade na África Subsaariana, no Brasil, Suriname, Venezuela, Caribe, e em menor quantidade no Egito e península da Arábia; *S. haematobium* (laranja) tem maior prevalência na África Subsaariana, Egito, Sudão e na península da Arábia; *S. japonicum* (rosa) encontra-se ao longo dos lagos e rios centrais da China, em Mindanao, Leyte, em algumas outras ilhas das Filipinas e em pequenas regiões na Indonésia; *S. mekongi* (verde escuro) é encontrado na Bacia central Mekong no Laos e Camboja; *S. intercalatum* na África ocidental e central (Adaptado de (9)).

Além de infectar os humanos, *S. mansoni* é encontrado também em roedores e primatas, enquanto *S. japonicum* infecta várias espécies de mamíferos incluindo gatos, cães, porcos e roedores (10).

Há uma estimativa que 7,1 milhões de pessoas são infectadas por *S. mansoni* nas Américas, sendo 95% dessas no Brasil (11;12). Estudos de diferenciação gênica e isoenzimáticos entre parasitos do velho e novo mundo confirmaram que a esquistossomose mansônica foi introduzida no Brasil pelo tráfico de escravos (13).



**Figura 2 - *Schistosoma*.** A- Verme adulto macho (direita) e fêmea (esquerda); B- Estágios de desenvolvimento do ciclo de vida de *Schistosoma*. Fonte: (2)

### 1.2- *Schistosoma*

As espécies *S. mansoni*, *S. japonicum* e *S. haematobium* são parasitos digenéticos, pertencente ao filo Platelmino, classe Trematoda, ordem Digenea, da família Schistosomatidae e do gênero *Schistosoma*. As três espécies apresentam dimorfismo sexual, sendo que o macho aloja e fecunda a fêmea em um canal constituído por dobras longitudinais, chamado canal ginecóforo (**Figura 2**). O *Schistosoma* apresenta ciclo de vida bastante complexo, passando por seis diferentes formas (cercária, esquistossômulo, verme adulto, ovo, miracídio e esporocisto) em ambientes diversos: a água, o hospedeiro definitivo (mamíferos) onde se desenvolve a fase adulta sexuada e o hospedeiro intermediário (caramujo), onde desenvolve sua fase assexuada (**Figura 2**) (2).

As cercárias, larvas multicelulares, são liberadas pelo hospedeiro intermediário (caramujos) no meio aquático (14). Quando encontram um hospedeiro apropriado, penetram

pela sua pele ou mucosa liberando a cauda, dando início ao processo de transformação em esquistossômulos.

Após permanecerem na pele por aproximadamente 72 horas, os esquistossômulos iniciam a migração pelo do corpo do hospedeiro vertebrado. Através da circulação, eles chegam ao coração e aos pulmões em torno de quatro dias, tornando-se mais longos e delgados, o que facilita a sua migração através da rede vascular pulmonar. Do pulmão, os esquistossômulos voltam ao coração e são enviados pela circulação geral a todas as partes do corpo do hospedeiro. Somente quando alcançam o sistema porta intra-hepático completam seu desenvolvimento em vermes adultos (15).

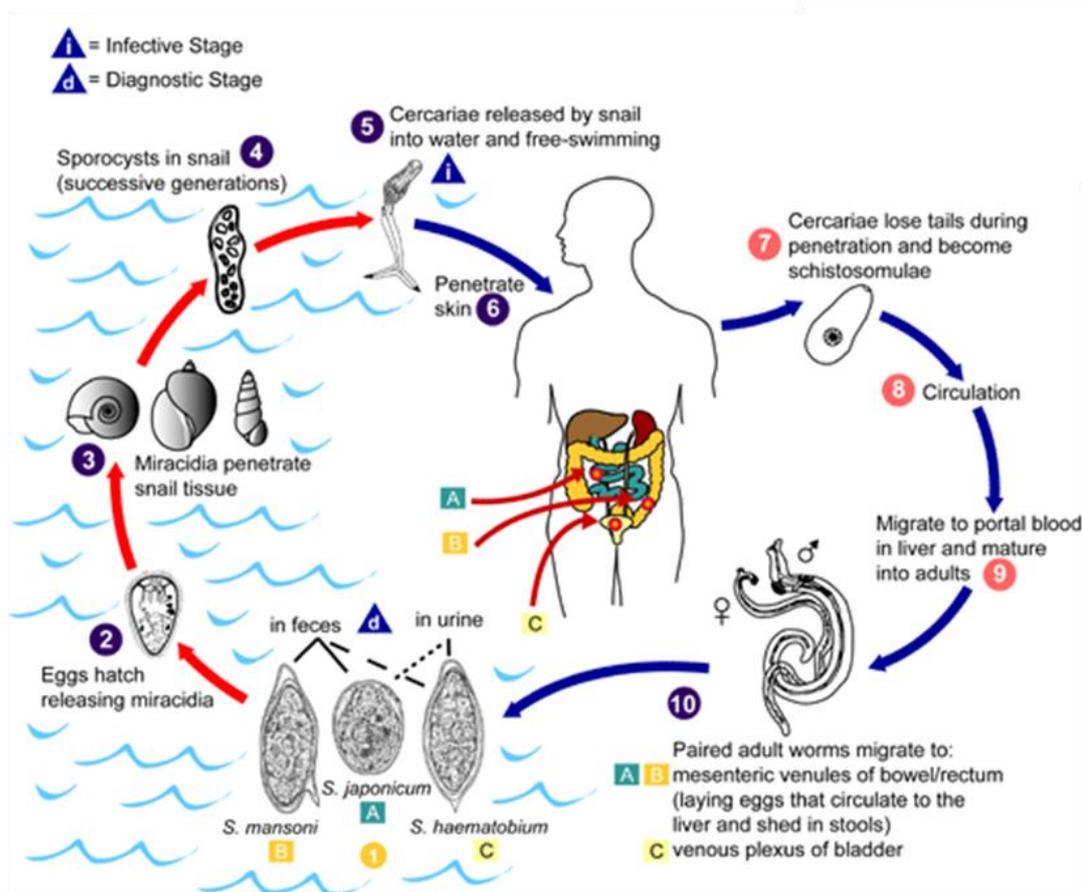
Quatro semanas após a infecção, os vermes encontram-se maduros e prontos para se acasalarem. Os vermes acasalados deslocam-se ativamente e migram para as veias mesentéricas (*S. mansoni* e *S. japonicum*) ou perivesicais (*S. haematobium*). Os vermes adultos são delgados, longos e possuem duas ventosas, uma oral e uma ventral (15). O par está em constante associação, encontrando-se a fêmea alojada no canal ginecóforo do macho. A maior parte da energia das fêmeas é gasta na produção e postura de ovos. Estas liberam entre 100 e 1000 ovos por casal e por dia (16). Estes ovos, por sua vez, migram através das paredes dos vasos em que foram depositados, atingindo os tecidos e, em seguida, o lúmen intestinal ou o lúmen da bexiga urinária. Finalmente, eles alcançam o meio externo através das fezes ou urina (15;17).

Os ovos eliminados eclodem ao encontrarem água fresca em condições adequadas, liberando o miracídio. Assim que a larva emerge começa a nadar ativamente na busca do hospedeiro intermediário (15). Os miracídios possuem seu corpo coberto por células ciliadas e organelas sensoriais que o auxiliam na localização do hospedeiro intermediário. Após a penetração o miracídio perde seu revestimento epitelial e seu órgão de penetração e sua musculatura atrofia (15). Por volta do oitavo dia o miracídio apresenta-se como um tubo enovelado, repleto de células germinativas em multiplicação, se transformando em esporocisto.

O esporocisto primário apresenta grandes células germinativas que se rompem pra liberar esporocistos filhos. Os esporocistos secundários apresentam células germinativas em constante multiplicação que formam aglomerados celulares e se diferenciam para formar

cercárias. A cercária é, então, liberada em água fresca, fechando o ciclo de vida do parasito (**Figura 3**) (15).

Ainda durante a infecção do hospedeiro definitivo, muitos dos ovos não ultrapassam a parede intestinal e são levados a órgãos e tecidos do hospedeiro pela corrente sanguínea. A presença destes ovos nos tecidos leva ao desenvolvimento de uma vigorosa reação granulomatosa, resultante da resposta do sistema imune do hospedeiro. A resposta de hipersensibilidade retém antígenos citotóxicos ao redor dos ovos, sendo, portanto, benéfica para o hospedeiro, mas, promove uma reação inflamatória que causa danos aos tecidos, principalmente ao fígado, podendo levar a uma grave fibrose hepática e hipertensão do sistema porta (18).



**Figura 3 - Ciclo biológico de *Schistosoma* (19).**

Vários esforços têm sido dedicados para o controle e tratamento esquistossomose. Contudo, apesar da aplicação de medidas sanitárias, supressão do hospedeiro intermediário e

tratamentos quimioterápicos, em algumas regiões, a prevalência da doença não mostrou redução significativa (20).

Algumas evidências sugerem a possibilidade de desenvolvimento de uma vacina contra a esquistossomose, tais como: o alto nível de proteção obtido pela vacinação com cercárias irradiadas (21), a existência de indivíduos não infectados em áreas endêmicas (22;23), o fato de indivíduos jovens serem mais propícios a infecção do que indivíduos adultos (24), a auto-cura relatada em macacos-rhesus (25) e a existência de vacinas recombinantes para helmintos, como *Taenia ovis* e *Echinococcus granulosus* (26;27). No entanto, embora alguns antígenos potencialmente candidatos a vacina para a esquistossomose tenham sido sugeridos, até o momento nenhum apresentou um nível de proteção efetivo (28). Um resultado promissor foi conseguido com o antígeno Sm14, uma proteína ligante de ácido graxo, que alcançou até 67,9% de proteção em camundongo (29).

Na ausência de uma vacina, de controle efetivo do vetor e saneamento básico adequado, o tratamento e o controle da esquistossomose são realizados através de uma única droga, o praziquantel (PZQ) (30). O PZQ atua em todas as espécies de *Schistosoma* e há 40 anos é utilizado. O PZQ é uma droga segura e eficiente, de baixo custo e ingerida em dose única (40 a 60 mg/kg). Ainda não foi relatado o desenvolvimento e/ou estabelecimento de resistência natural a essa droga (31;32). No entanto, o desenvolvimento de resistência induzida em laboratório ao PZQ (33), o fato desta ainda ser a única droga disponível no mercado e a relato de resistência a maioria dos anti-helmínticos veterinários, gerou o consenso de que o desenvolvimento de uma nova droga anti-schistosoma é necessária.

### **1.3- O genoma de três espécies de *Schistosoma***

Atualmente, a abordagem mais promissora para a identificação de novos alvos de drogas, vacinas e reagentes para diagnóstico, assim como para a compreensão da resistência a drogas, diversidade antigênica, infectibilidade e patologia, consiste em compreender e decifrar as informações nos genomas dos parasitos.

*S. mansoni*, *S. haematobium* e *S. japonicum* possuem genoma diploide com sete pares de cromossomos autossômicos e um par de cromossomos sexuais (34-36). A fêmea é heterogamética, possuindo o par de cromossomos sexuais ZW e o macho o par ZZ (37;38). O sexo do parasito é determinado ainda nos ovos, mas apenas no hospedeiro definitivo as

diferenças morfológicas entre macho e fêmea se desenvolvem. Nos estágios larvais, fêmeas e machos são geneticamente diferentes, mas morfológicamente idênticos (39).

O sequenciamento completo dos genomas nucleares de *S. mansoni* e *S. japonicum* foi publicado em 2009 (40;41) e mais recentemente, em 2011, o de *S. haematobium* (42). O *Wellcome Trust Sanger Institute* juntamente com um grupo internacional de pesquisadores realizou o sequenciamento do genoma nuclear de *S. mansoni* e *S. japonicum* utilizando a técnica de *Whole Genome Shotgun* (WGS). As bibliotecas foram construídas após extração do DNA de milhares de vermes adultos e o sequenciamento foi realizado pela técnica convencional (*Sanger Capillary Sequencing*) o que resultou em genomas muito fragmentados (19.022 *scaffolds* para *S. mansoni* e 25.048 *scaffolds* para *S. japonicum*). Desde então as técnicas de sequenciamento foram modificadas e o genoma nuclear de *S. haematobium* foi sequenciado a partir de um único par de vermes adultos que foi submetido à técnica *Whole Genome amplification* (WGA). O DNA mais homogêneo permitiu uma montagem mais robusta que gerou um genoma menos fragmentado com um total de 7.475 *scaffolds*, menos da metade do que foi conseguido com os outros dois genomas (43).

Recentemente, uma nova versão do genoma de *S. mansoni* foi disponibilizada e todos os dados foram melhorados devido ao uso do dado de mapeamento genético e de tecnologia de sequenciamento em massa, além do transcriptoma (RNA-seq) de diferentes fases do parasito (44). Essa nova metodologia ajudou na melhor identificação de regiões transcricionalmente ativas e definiu com mais precisão a coordenada dos exons de cada gene. Além disso, a natureza quantitativa do RNA-seq permitiu identificar genes diferencialmente expressos de uma forma mais sensível e precisa do que os outros métodos anteriormente empregados (45;46). Na primeira versão do genoma, pela técnica de FISH (*fluorescence in situ hybridization*), foi possível localizar, aproximadamente, 43% dos fragmentos montados em um dos sete cromossomos autossômicos ou no cromossomo sexual de *S. mansoni* (47). Na nova versão já foi possível mapear 81% das sequências em seus cromossomos, sendo o cromossomo 1 o maior deles. Além disso, o número de *scaffolds* foi reduzido para 885 e o número de genes preditos passou de 11.807 para 10.852. Dos genes excluídos, 65% haviam sido anotados como proteínas hipotéticas (48).

Na Tabela 1 os dados comparativos relacionados ao genoma dos três parasitos foram demonstrados.

Tabela 1: Comparação das características genômicas de três espécies de *Schistosoma*

Genomic features	<i>S. mansoni</i>	<i>S. haematobium</i>	<i>S. japonicum</i>
Sequencing Methods	WGS	WGA	WGS
Genome size (Mb)	381	385	403
Repetitive elements (%)	45	47,2	40,1
GC content (%)	34,7	34,3	33,5
Proportion of genome that is coding (%)	4,72	4,43	4,32
Number of putative coding genes	10.852	13.073	13.469
Genes length (average)	13.397	11.952	10.003
Exon number per gene (average)	6,2	5,4	5,3
Gene exons length (average)	218	246	223
Gene intron length (average)	2.331	2.442	2.058

As análises comparativas mostraram ainda que há uma maior sintonia entre os genomas de *S. haematobium* e *S. mansoni* (89,4%) do que entre *S. haematobium* e *S. japonicum* (51,7%) ou entre *S. mansoni* e *S. japonicum* (67,0%). Além disso, existem quatro vezes mais rearranjos intracromossômicos entre *S. mansoni* e *S. japonicum* do que entre *S. mansoni* e *S. haematobium* (49). Foram identificadas 10.880 proteínas comuns entre as três espécies. Destas, 6.142 possuem homólogos em outros platelmintos como *Fasciola hepática*, *Fasciola gigantica*, *Clonorchis sinensis* e *Opisthorchis viverrini*. Além disso, 1.333 classes de proteínas são comuns entre *S. haematobium* e *S. japonicum*, sendo que a maior parte destas são enzimas (quinases, glicosil-transferases e peptidases), proteínas do citoesqueleto, de reparo do DNA, envolvidas em replicação, recombinação, proteínas do spliceossomo e elementos do sistema de ubiquitinação. Inferência das relações genéticas entre os trematódeos mostrou que *S. mansoni* e *S. haematobium* são evolutivamente mais próximos. A comparação com outros genomas mostrou que o *Schistosoma* possui mais ortólogos em vertebrados como *H. sapiens* do que com outros nematódeos como *C. elegans* (50;51).

Os dados de transcriptoma em *S. haematobium* mostraram que o número de genes transcritos é maior em vermes fêmea dos que em machos. Nas fêmeas as vias de sinalização

enriquecidas estão relacionadas a hematofagia e produção de ovos. Já em machos as vias enriquecidas estão relacionadas a mobilidade, alimentação e cópula (52).

Todos as informações dos genomas acima descritos podem ser visualizados no banco de dados (<http://www.schistodb.net>) (53). Com o SchistoDB é possível ter acesso a todas as sequências, anotações e qualquer outro tipo de dado integrado as informações genômicas, além de possuir diferentes ferramentas para analisar, recuperar e visualizar os dados.

## **1.4- Era Pós-Genômica**

### **1.4-1. Anotação Funcional**

A finalização do sequenciamento de um genoma é sucedida por processos que visam a montagem, predição gênica e anotação funcional. Embora o conhecimento da função gênica seja de extrema importância para a compreensão dos processos biológicos, cerca de 30-35% dos genes/proteínas, na maioria dos organismos, permanece sem caracterização experimental ou predita e, grande parte das demais sequências codificantes possui uma anotação funcional muito genérica (54). Nos últimos anos, o crescente número de genomas completamente sequenciados vem adicionando novas dimensões às ferramentas de análise de sequência e predição da função proteica (55).

O método computacional mais antigo e simples para atribuir função para a proteína é baseado na detecção de homólogos com função conhecida (previsão de função baseada em homologia) (56). Lembrando que homologia implica em ancestralidade comum, que é diferente de similaridade, que se refere ao grau de proximidade entre duas ou mais sequências. A anotação funcional para caracterizar relações de homologia se baseia no fato de que 70 a 90% dos genes em um genoma têm homólogos em outras espécies (57), e esse número cresce à medida que novos genomas são sequenciados. Entretanto, para aproximadamente 40% dos genes na maioria dos genomas, não existe predição clara de função, e para a outra grande maioria a predição de função é muito genérica (58). Soma-se a isso, o fato que sequências podem exibir alto grau de similaridade e não compartilharem um ancestral comum (evolução convergente) ou mesmo terem alto grau de divergência (baixa similaridade) e terem se originado de um único ancestral (evolução divergente). Além disso, já é sabido que proteínas com alto grau de similaridade entre suas sequências e até mesmo estrutura podem desempenhar diferentes funções bioquímicas (59;60).

Com o intuito de melhorar a predição da função das proteínas, novas formas de atribuir função a genes preditos estão surgindo, entre eles, métodos de detecção de ortologia em larga escala e métodos que exploram a relação entre genes e genomas, como fusão de genes, proximidade cromossomal, distribuição entre espécies ou coexpressão de genes conservados (61;62).

#### **1.4-2. Inferência Filogenética para classificação funcional de genes**

A inferência filogenética pode ser aplicada em diferentes contextos, como por exemplo, em estudos taxonômicos, de epidemiologia molecular, estudos de biodiversidade, análises forenses e predição funcional. Nesta última, a análise filogenética adiciona uma perspectiva evolutiva às análises biológicas comparativas tornando a predição de função proteica mais robusta e confiável (63). A abordagem filogenômica se baseia em múltiplos passos que incluem seleção dos genes/proteínas alvo, identificação de potenciais homólogos através de métodos baseados na busca por similaridade de sequências, alinhamento múltiplo, construção de árvores filogenéticas, anotação da topologia da árvore, discriminação entre ortólogos e parálogos e, finalmente, inferência da função da proteína tendo como base as informações disponíveis para os potenciais ortólogos, seja em bancos de dados e/ou literatura (64) (**Figura 4**). A inferência filogenética, quando aplicada em escala genômica, não é uma técnica simples e exige recursos computacionais mais pesados. No entanto, uma variedade de ferramentas e softwares foi desenvolvida para auxiliar em todas essas etapas (65-75).

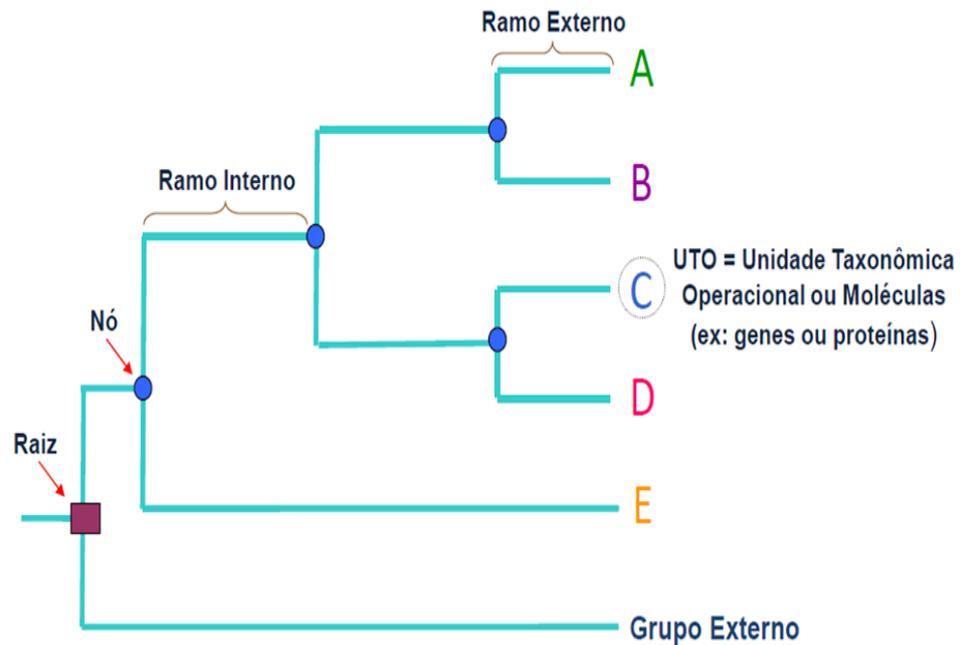
A reconstrução filogenética, utilizando dados moleculares, sequências de genes e/ou proteínas, tem como resultado uma árvore evolutiva que hipotetiza o relacionamento dessas sequências. A filogenia resultante tem numerosas aplicações como a detecção de eventos evolutivos e o estabelecimento acurado de relações de homologia, ortologia e paralogia. Ortólogos são proteínas ou genes homólogos derivados de um ancestral comum por especiação, enquanto parálogos são homólogos derivados por duplicação gênica. A detecção de relações de ortologia é idealmente realizada através da identificação de eventos de especiação e duplicação evidenciados pelas árvores filogenéticas (76). Uma árvore é um grafo utilizado para representar a história evolutiva de um grupo de moléculas ou organismos e consiste de nós conectados por ramos. As unidades taxonômicas operacionais (UTO) representam as sequências ou os organismos a partir dos quais os dados são provenientes. Os nós internos representam ancestrais hipotéticos e, o ancestral comum a todas as sequências ou

organismos que compõem a árvore, é a raiz da árvore filogenética, no caso de árvores enraizadas (**Figura 4**) (77).

Nas árvores filogenéticas que utilizam caracteres moleculares, o número de evidências é consideravelmente maior, uma vez que cada nucleotídeo ou aminoácido é interpretado como um caractere que evolui de forma independente (78). Diferentes proteínas ou segmentos de DNA não evoluem na mesma velocidade e isto deve ser considerado na escolha da metodologia científica aplicada à reconstrução filogenética. Atualmente, os métodos utilizados para converter a informação armazenada nas sequências nucleotídicas ou proteicas em árvores evolutivas podem ser divididos em dois grandes grupos: baseados em distância e baseados em caracteres (79;80).

Os métodos baseados em distância (81;82) comparam as diferenças entre todos os pares de sequências alinhadas e em seguida convertem as diferenças em uma matriz de distâncias que representa a distância evolutiva entre cada par. Desta forma, a matriz representa o número ou a proporção de diferenças observadas, ou o número esperado de substituições por sítio sob as premissas de um modelo de evolução molecular (79). As desvantagens dos métodos baseados em distância incluem a perda de informações quando caracteres discretos são convertidos em uma matriz de distância e o fato das estimativas confiáveis de distância não serem obtidas para sequências muito divergentes (79;83).

Os métodos baseados em caracteres como máxima verossimilhança (do inglês, *maximum likelihood*, ML) (Felsenstein, 1981) analisam cada caractere individualmente. Este método permite a incorporação de modelos evolutivos complexos na análise produzindo árvores filogenéticas mais acuradas (84). Entretanto, este método demanda mais tempo e capacidade de processamento computacional.



**Figura 4: A árvore filogenética.** Nesta árvore hipotética as letras A, B, C, D e E representam as seqüências proteicas a partir das quais a árvore foi reconstruída. A árvore representa a história evolutiva deste grupo de moléculas e consiste de nós conectados por ramos. Os nós internos representam ancestrais hipotéticos e o ancestral comum a todas as seqüências ou organismos que compõem a árvore é a raiz da árvore filogenética. O enraizamento da árvore é realizado a partir do grupo externo.

Dois pontos centrais têm de ser levados em consideração quando se utiliza a inferência filogenética para classificação funcional dos genes: a topologia da árvore tem que estar correta e o banco de dados utilizado para transferir a informação tem que ser confiável. Na prática, avaliar a precisão provável de uma árvore é normalmente realizado através de análise de *bootstrap* ou pela comparação das árvores construídas utilizando diferentes métodos de reconstrução filogenética (85). A transferência da anotação, seja no contexto genômico ou para casos isolados, idealmente, só deveria ser realizada quando os gene/proteína fonte da anotação possui comprovação experimental. No entanto, apenas 3% das proteínas depositadas em um banco de dados públicos (UniProt - <http://www.uniprot.org/>), por exemplo, têm suporte experimental. Até mesmo os organismos mais estudados como *E. coli*, levedura e *C. elegans*, contém um grande número de genes, cuja função ainda não foi experimentalmente comprovada (79). Com isso, a transferência da informação sem suporte experimental deve seguir padrões rigorosos, e a escolha de banco de dados que utilizam técnicas robustas de anotação, ou seja, aqueles que possuem curadoria manual e/ou que usam métodos de predição baseados não só em similaridade, devem ser priorizados.

### 1.4-3. O *Schistosoma* na era Pós-Genômica

Muitas abordagens distintas são usadas para explorar as propriedades funcionais dos genes (86). Nesse sentido, o recente sequenciamento dos genomas do *S. mansoni*, *S. haematobium* e *S. japonicum* criaram novas oportunidades para se pesquisar e entender a biologia desse complexo parasito.

Entre outros diversos avanços gerados na era pós sequenciamento de *Schistosoma*, podemos ressaltar o progresso conseguido na identificação das regiões gênicas e, a partir dos experimentos de microarranjos e, mais recente, a identificação do perfil transcricional das diferentes fases de vida do *S. mansoni* e *S. haematobium* por RNA-seq (87-91). Com o genoma e o transcriptoma ainda foi possível identificar o proteoma do *S. mansoni* (92) e com isso, os estudos que envolvem a caracterização de proteínas na superfície do parasito avançaram (93;94) melhorando os métodos de busca por candidatos a vacina e imunomoduladores (95).

No entanto, os estudos com *Schistosoma* têm suas limitações e, grandes dificuldades no progresso das pesquisas estão surgindo. Os métodos confiáveis de genética reversa e de manipulação gênica que são utilizados em larga escala no nematódeo de vida livre, *C. elegans* (96), por exemplo, não têm a mesma eficácia em *Schistosoma*. *Schistosoma* é um parasito com ciclo de vida complexo, com muitos estágios diferentes e ainda não existe uma técnica para a manutenção *in vitro* de todas as fases do ciclo. É necessário infectar e sacrificar animais para conseguir parasitos para experimentos laboratoriais, o que dificulta e limita muitos estudos. *Schistosoma* é um organismo eucariótico, multicelular que apresenta tecidos distintos cada um contendo um conjunto específico de células para as quais ainda não existe uma linhagem contínua que possa ser cultivada *in vitro*. Somado a isso, ainda não é possível a produção de organismos transgênicos, limitando muito as pesquisas que poderiam se beneficiar dessa técnica (97).

Apesar de todas as dificuldades, nos últimos anos foi possível aprimorar a técnica de genética reversa aplicada a esse parasita. Uma ferramenta promissora nesse contexto é a interferência por RNA (RNAi). Através dessa ferramenta está sendo possível aprofundar os estudos das funções gênicas para revelar o(s) suposto(s) papel(is) de cada proteína.

#### 1.4-4. Interferência por RNA (RNAi)

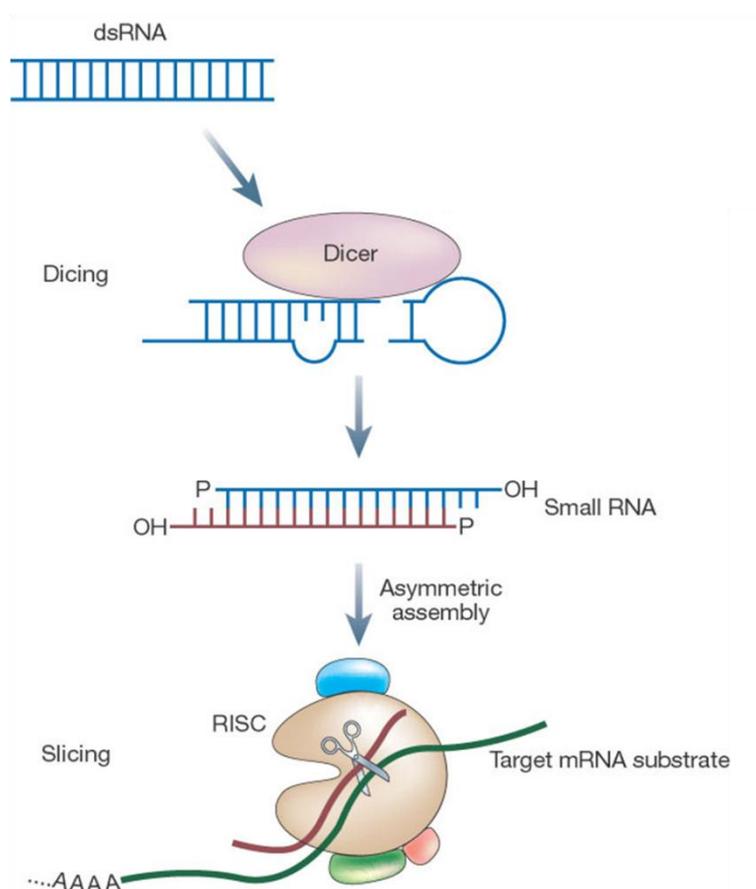
A técnica de RNAi foi primeiramente descrita em 1998, quando Fire e colaboradores (98) descobriram a habilidade de RNAs de dupla fita longos (dsRNA) silenciarem a expressão gênica em *C. elegans*. Logo em seguida, outro trabalho demonstrou que pequenos RNAs, conhecidos como *short interfering RNAs* (siRNAs), são capazes de bloquear sequências específicas em células de mamíferos e foram então usadas em experimento *in vivo* em camundongos (99).

Em células eucarióticas, RNAi é uma via fundamental na qual siRNAs específicos são capazes de ligar ao seu RNA mensageiros (mRNA) complementar. Esse mecanismo é dirigido por um complexo chamado RISC (*RNA-induced silencing complex*) que identifica e degrada, seletivamente, o mRNA que é complementar a fita antisenso do siRNA (100). Provavelmente, o RNAi faz parte de um mecanismo de defesa celular contra infecção por vírus RNA dupla fita e/ou elementos móveis, como retrotransposons, conservando a estabilidade genética. Em teoria, siRNAs podem ser utilizados para silenciar qualquer gene e, combinado ao conhecimento adquirido com o sequenciamento dos genomas, é uma poderosa ferramenta de caracterização funcional.

Resumidamente, o mecanismo de RNAi envolve um complexo proteico, chamado DICER, que possui atividade de ribonuclease III e cliva os dsRNAs que entram na célula, produzindo pequenos fragmentos de 21 a 23 pb de RNA dupla fita (que são os siRNAs). Os siRNAs contêm dois nucleotídeos não pareados de cada um dos lados que se ligam ao complexo RISC. Uma de suas fitas, em geral a fita antisenso, é selecionada para permanecer ligada ao complexo. O complexo é guiado ao mRNA de sequência complementar ao siRNA que está ligado à RISC. Uma endonuclease, chamada argonauta 2 (Ago2), presente no complexo, cliva o mRNA em um ponto interno central às fitas pareadas siRNA/mRNA. Os fragmentos liberados são degradados devido a perda de elementos essenciais à estabilidade dos mRNAs, tais como cauda poliA, suprimindo assim a tradução do mRNA alvo do silenciamento (101;102) (**Figura 5**).

Logo após a descoberta que os dsRNAs (*Double stranded RNA*) eram ativadores da via de RNAi, foi observada a existência de uma classe de moléculas de RNA endógeno chamadas microRNAs (miRNAs). Os miRNAs são precursores em forma de grampo de

cabelo e utilizam parte da maquinaria de RNAi para regular a expressão gênica através da degradação do mRNA ou supressão da tradução (103).



**Figura 5: Ilustração do mecanismo de RNAi.** Longos dsRNA são degradados pela enzima Dicer em pequenos RNAs (siRNA). Os siRNAs se ligam ao complexo RISC e são utilizados como molde para a ligação no mRNA complementar. Esse mRNA é então clivado, reprimindo assim, a tradução da proteína (104).

O silenciamento gênico por RNAi tem revolucionado a biologia experimental, sendo aplicada até mesmo em escala genômica para caracterização de genes (96;105). O sucesso da técnica está diretamente relacionada a estabilidade da proteína-alvo, ao organismo no qual a técnica está sendo aplicada, a forma de distribuição do RNA de dupla fita e a sequência utilizada para síntese do dsRNA/siRNA. Além disso, o silenciamento pode ser disparado através da inoculação de dsRNAs, de siRNAs ou de vetores transcrevendo *short hairpin* RNAs (shRNA) e a eficiência dos siRNAs dependerá da sua estabilidade termodinâmica ou acessibilidade para se ligar ao transcrito alvo (106).

Desde as primeiras publicações em 2003 (107;108), o RNAi têm mostrado ser uma ferramenta importante no estudo da perda ou diminuição da função para alguns genes de interesse do *S. mansoni*, incluindo proteases (109-114), quinases (112;114;115), receptores TGF-beta (116), enzimas antioxidantes (117), aquaporina (118), transportadores de glicose (119), tetraspaninas (120), entre outros. Em alguns estudos, nítidos fenótipos que afetaram a viabilidade e sobrevivência dos parasitos, foram obtidos. Essas evidências são ainda sustentadas por trabalhos que sugerem que a maquinaria molecular do RNAi está presente no *S. mansoni* (121-124).

### **1.5- Sinalização celular e o seu possível papel no desenvolvimento do *Schistosoma***

Como mencionado anteriormente, o *Schistosoma* tem um ciclo de vida complexo. Durante o seu desenvolvimento, sinais no ambiente estimulam mudanças fisiológicas, morfológicas e bioquímicas. A capacidade de infectar ativamente, assim como a migração pelo sistema de ambos os hospedeiros, demonstra uma sofisticada coordenação dos seus sistemas fisiológicos em diferentes fases de desenvolvimento (125).

Na relação parasito-hospedeiro, o verme é capaz de sobreviver por décadas no hospedeiro vertebrado utilizando-se de suas biomoléculas como fonte de alimento, energia e também como moléculas sinalizadoras. Já foi visto que o parasito utiliza de TNF- $\alpha$  do hospedeiro como sinal estimulatório para a produção de ovos (126), IL-7 já foi relacionada ao o crescimento e fecundação de vermes adultos (127) e a presença do receptor de TGF- $\beta$  expresso na superfície do tegumento sincicial indica a utilização deste fator pelo parasito (128). O Fator de crescimento epitelial humano (EGF) induz a meiose em ovócitos (129) e produtos de células T CD4+ modulam o desenvolvimento do verme (130). Em hospedeiros imunodeprimidos o desenvolvimento do parasito é retardado. Uma possível explicação para este fenômeno é que o parasito aguarda melhores condições de sobrevivência do hospedeiro para então finalizar o seu processo de crescimento, acasalamento, fecundidade e transmissão (131). Estes resultados indicam que o parasito está tão adaptado ao meio, que utiliza proteínas imunoreguladoras do hospedeiro como sinais para sua replicação (132).

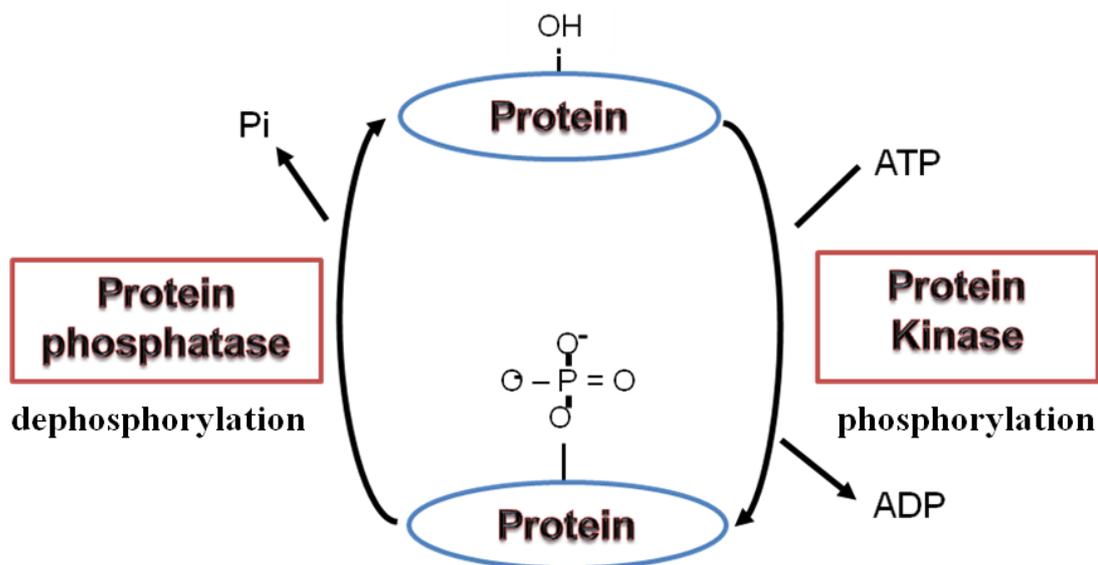
A fêmea necessita do contato com o macho para sua maturação, com possível transferência de moléculas sinalizadoras como hormônios esteroides, que são captadas por

receptores (133). Foi observado que a separação do casal implica na interrupção da transcrição de alguns genes responsáveis pela diferenciação dos vitelócitos (134). O desenvolvimento reprodutivo do macho não depende da interação com a fêmea. No entanto, o macho necessita da fêmea para regular o mecanismo de acúmulo de glutatona e lipídeo, na utilização de lipase e estimulando a incorporação de tirosina (135). Além disso, estudos de microarranjos mostraram a expressão diferencial de genes entre vermes machos e fêmeas (136-138).

Os fenômenos acima citados acontecem a partir de comunicações moleculares, sendo estas interações mediadas por processos de transdução de sinal (139). Transdução de sinal envolve redes integradas não lineares que interagem principalmente por mudança de status de atividade através de fosforilação (por proteínas quinase - PKs) e defosforilação (por proteínas fosfatase) de resíduos de aminoácidos, ou pela incorporação de GTP. Outros mensageiros celulares não-proteicos incluem AMP cíclico,  $Ca^{2+}$  e diacilglicerol. A resposta para fosforilação de PKs e consequente ativação das cascatas de sinalização é a transcrição/tradução de genes alvos-específicos ou o bloqueio destes.

### **1.6- Proteínas quinase**

As proteínas quinase (PKs) pertencem a uma grande família de enzimas, muitas das quais mediam respostas a estímulos externos em células eucarióticas (140). As PKs são proteínas de fosforilação reversível (**Figura 6**) que desempenham um papel central na regulação de funções básicas de todos os eucariotas, tais como: replicação, controle do ciclo celular, transcrição gênica, rearranjo do citoesqueleto, apoptose, metabolismo energético e outras. São também requeridas em funções mais finas, em eucariotas mais complexos, como: diferenciação celular de tecidos e órgãos, comunicação entre células, interação da célula com o substrato e mediação de interações complexas com o ambiente externo (141). As proteínas quinase agem na fosforilação do substrato, transferindo um grupo fosfato (P) do ATP para a cadeia de aminoácido de uma proteína substrato (**Figura 6**). O fosfato é carregado negativamente o que leva a mudança na conformação da proteína, causando assim modificações na sua estrutura e como consequência na sua atividade. Muitas PKs são moduladas por autofosforilação ou fosforiladas por outras proteínas quinase (142).



**Figura 6: Fosforilação reversível de proteínas quinase.** As proteínas quinase catalisam a transferência de um grupo fosfato (P) do ATP [ADP(P)] a uma proteína substrato específica. A proteína fosfatase remove o grupo fosfato, permitindo que a proteína substrato volte ao seu estado inativo.

As PKs possuem um domínio catalítico, que liga e fosforila proteínas alvo, e um domínio ou uma região regulatória, que controla a interação entre duas proteínas. O domínio catalítico da ePKs é composto por 250-300 aminoácidos, dividido em 12 subdomínios composto por aminoácidos individuais ou motivos altamente conservados (143).

PKs são classificadas em duas superfamílias: as proteínas quinase eucariotas ou convencionais (ePK), que compartilham um domínio catalítico conservado, e as proteínas quinase atípicas (aPKs) (Tabela 2). aPKs são proteínas que possuem atividade bioquímica quinase, mas não possuem um domínio catalítico conservado entre seus membros (144).

De acordo com os sítios de reconhecimento do substrato, ePKs são divididas em duas grandes classes: serina treonina quinases (STKs) e tirosina quinases (TKs). Proteínas quinase de dupla especificidade (Híbridas), também são encontradas. As ePKs foram classificadas em oito grandes grupos com base na similaridade da sequência de seus domínios catalítico, na presença e similaridade de seus domínios acessórios e com base no seu modo de regulação (145-148). De acordo com a classificação contida no KinBase (149), um banco de dados que contém informações de PKs codificadas no genoma humano e os seus homólogos de outros eucariotos, os oito grupos ePK são: AGC (proteína quinases dependente de cAMP/proteína

quinase G/proteína quinase C), CaMK (quinases reguladas por cálcio e calmodulina), CK1 (Cell Kinase I), CMGC (quinases dependentes de ciclina e outros parentes próximos), RGC (receptor Guanilato ciclase), STE (quinases da via de sinalização das MAP Quinase), TK (proteína tirosina quinase) e TKL (tirosina quinase *Like*). Um nono grupo, chamado *Other*, consiste de uma coleção mista de quinases, que não podem ser classificado facilmente nas famílias anteriores (150) (Tabela 1).

PKs são consideradas alvos para o desenvolvimento de drogas do ponto de vista médico e químico e um número crescente de inibidores PKs foram desenvolvidos e aprovados para o tratamento de diferentes doenças humanas (151). Um exemplo de um inibidor de PK bem-sucedida é o Gleevec®, que induz uma mudança conformacional na proteína alvo e impede a ativação das PKs subsequentes (152). Além disso, PKs têm ganhado interesse como estratégia de tratamento para combater muitos parasitos, incluindo *S. mansoni* (153-155). Recentemente um estudo demonstrou o efeito do inibidor para ePKs humanas Gleevac ® na morfologia e fisiologia de *S. mansoni* em experimentos *in vitro*. O inibidor causou alterações patológicas na gastroderme dos vermes adultos que finalmente causaram a morte dos parasitos (156).

Tabela 2: Classificação das proteínas quinase (157)

Superfamily	Class	Group
<b>Eukaryotic Protein Kinases (ePKs)</b>	Serine/Threonine Kinases	AGC
		CaMK
		CMGC
		STE
		CK1
	Tyrosine Kinases	TK
		RGC
	Hybrid	TK like
		Other
<b>Atypical Protein Kinases (aPKs)</b>	<b>Atypical</b>	

## 2. JUSTIFICATIVA

A esquistossomose ainda mata mais de 200 mil pessoas anualmente e, mesmo com todas as medidas preventivas e de tratamento adotadas, a prevalência da doença não mostrou redução significativa nos últimos anos em algumas regiões. Mais de 700 mil pessoas em todo mundo moram em áreas endêmicas e correm o risco constante de contrair a infecção (158;159). Um único medicamento é usado massivamente para o tratamento da doença e, apesar de ser uma droga eficiente, seguro e de baixo custo, ela não previne uma reinfecção e vem sendo utilizado há mais de 40 anos. Em áreas endêmicas os indivíduos estão constantemente se reinfectando e, alguns casos isolados de resistência ao praziquantel já foram relatados (160;161). Não existe uma alternativa eficiente para o tratamento e a OMS está incentivando os estudos que tentam entender a biologia do *Schistosoma* e/ou que tentam identificar novos alvos para drogas e vacinas. Recentemente, o genoma de *S. mansoni*, *S. haematobium* e *S. japonicum* foi sequenciado (162-164) criando novos meios para decifrar as informações codificadas pelos genes e compreender o papel de cada proteína no desenvolvimento e sucesso adaptativo desse parasito.

O *Schistosoma* é um parasito que possui um ciclo de vida complexo em que seis diferentes formas de vida são encontradas em ambientes distintos (2). Diante disto, foram escolhidas como alvo de estudo desse trabalho as proteínas quinase, enzimas envolvidas em diferentes vias de sinalização. Vias essas que levam a processos celulares críticos como replicação, controle do ciclo celular, transcrição gênica, rearranjo do citoesqueleto, apoptose e metabolismo energético (165), processos essenciais para qualquer organismo que passa constantemente por mudanças fisiológicas, estruturais e adaptativas como o *Schistosoma*. Além disso, proteínas quinase são consideradas bons alvos do ponto de vista médico e químico e diversos inibidores quinase já foram desenvolvidos para tratamento de doenças humanas como o câncer e também outras parasitoses (166).

As proteínas quinases são diversas e possuem uma classificação específica que as separam em nove diferentes grupos e em diversas famílias e subfamílias. Essa classificação separa aquelas proteínas com atividade funcional similar (167). Dessa forma, identificar e classificar as proteínas quinase em um organismo se torna um meio para se entender o papel funcional desses genes. Com o intuito de identificar e classificar as ePKs de *Schistosoma* de uma maneira confiável, optamos por adotar uma metodologia mais precisa para anotação funcional. Foi escolhida assim, a inferência filogenética que, dentro de um contexto evolutivo,

é mais eficiente do que métodos automáticos e mais utilizados para classificação de genes (168).

O mecanismo de RNA de interferência é um processo intrínseco à célula da maioria dos eucariotos, que entre outras funções, promove o silenciamento gênico através da degradação de mRNAs específicos. Como ferramenta molecular para analisar a função gênica, o RNAi é particularmente importante em organismos como o *Schistosoma* em que as técnicas clássicas de manipulação gênica ainda não são completamente válidas (96). Dessa forma, o RNAi se tornou uma ferramenta poderosa para validar experimentalmente a função de genes nesse parasito (169;170) e, foi a técnica escolhida nesse trabalho.

Todos os aspectos acima citados justificam o estudo com *Schistosoma*, a identificação de novos alvos para o desenvolvimento de drogas para o tratamento da esquistossomose, assim como as metodologias utilizadas para a conclusão do trabalho.

### **3. OBJETIVOS**

#### **3.1- Objetivo geral**

Caracterizar as proteínas quinase eucarióticas (ePKinoma) de *Schistosoma mansoni*, *Schistosoma haematobium* e *Schistosoma japonicum* visando a identificação de potenciais alvos terapêuticos.

#### **3.2- Objetivos específicos**

- ✓ Identificar todas as proteínas quinase eucarióticas (ePKinoma) no proteoma predito do *S. mansoni*, *S. haematobium* e *S. japonicum*.
- ✓ Padronizar a metodologia de inferência filogenética para classificar e anotar funcionalmente as ePKs identificadas.
- ✓ Selecionar ePKs que possam ser alvos para o desenvolvimento de drogas de acordo com os ortólogos identificados e literatura disponível.
- ✓ Investigar o papel biológico de proteínas envolvidas na via de sinalização das proteínas MAPK através do silenciamento gênico e frente ao sistema imune do hospedeiro.
- ✓ Verificar diferenças de expressão dos prováveis alvos regulados pela via MAPK após silenciamento dos genes selecionados.

## 4. MATERIAL E MÉTODO

### 4.1- Identificação, classificação e anotação funcional das proteínas quinase do *S. mansoni*, *S. haematobium* e *S. japonicum*

#### 4.1-1. Organismos e Sequências

*S. mansoni* [NCBI taxid: 6183], *S. haematobium* [NCBI taxid: 6185] e *S. japonicum* [NCBI taxid: 6182], e seis outros organismos foram selecionados para os processos de identificação e/ou classificação das proteínas quinase, incluindo o *Homo sapiens* [taxid: 9606], *Mus musculus* [taxid: 10090], *Drosophila melanogaster* [taxid: 7227], *Caenorhabditis elegans* [taxid: 6239], *Brugia malayi* [taxid: 6279] e *Saccharomyces cerevisiae* [taxid: 4932]. Os proteomas preditos do *S. mansoni*, *S. haematobium* e *S. japonicum* foram obtidos do SchistoDB, versão [3.0] (171). O conjunto de dados de PKs dos organismos *H. sapiens*, *M. musculus*, *D. melanogaster* e *C. elegans*, utilizados nas análises filogenéticas, foram transferidos do banco de dados KinBase (149). Já as sequências de aminoácidos das proteínas quinase de *B. malayi*, foram obtidas do KEGG (172). As sequências do KinBase e do KEGG foram extraídas no período de maio a outubro de 2009.

#### 4.1-2. Identificação das ePKs e proteínas relacionadas

A fim de identificar possíveis homólogos às ePKs de *S. mansoni*, sequências de aminoácidos de proteínas quinase conhecidas de cinco organismos modelo (*H. sapiens*, *M. musculus*, *D. melanogaster*, *C. elegans* e *S. cerevisiae*) foram selecionadas para a construção do HMM. As 68 sequências de aminoácidos que compõem o domínio catalítico de diferentes grupos e famílias ePKs que compartilhavam grau de identidade inferior a 50% (verificados pelo programa alstat) foram alinhadas utilizando o programa MAFFT (72) (sem modificar os parâmetros padrão do programa) e editadas manualmente utilizando o programa *BioEdit* (173). Na edição manual as regiões não alinhadas e *gaps* foram retiradas e as regiões mais conservadas foram mantidas. Os HMMs locais e globais foram construídos a partir de três módulos do pacote HMMER (<http://hmmer.wustl.edu>). Sendo que o alinhamento múltiplo de sequência editado foi utilizado para a construção do HMM a partir do módulo *hmmbuid*. O HMM gerado foi então calibrado utilizando o módulo *hmmcalibrate*. O HMM calibrado foi utilizado nas buscas por homólogos no proteoma predito do *S. mansoni*. Os programas do pacote HMMER foram executados em um servidor do Centro de Excelência em

Bioinformática (CEBio) com as seguintes características: Processador Intel 2X Xeon X5460 Quad-Core de 3.16 Ghz, 6MB cache memory, 24 GB RAM, 1,5TB HD.

Para a identificação de ePKs homólogas no proteoma predito de *S. haematobium* e *S. japonicum* foram utilizados HMMs construídos pelo banco de dados de famílias proteicas do Pfam (174), que foram integrados na versão atual do SchistoDB (3.0) (<http://qa.schistodb.net>). O Pfam possui um modelo para identificar proteína quinase serina/treonina (PF00069) e outro para identificar proteínas tirosina quinase (PF07477). Utilizamos os dois modelos nas nossas buscas. Em seguida, foram definidas as ePKs ortólogas nas três espécies (ver item 4.1-3).

Além das ePKs, utilizamos HMMs para identificar uma proteína GTPase da família Ras, homóloga a proteína let – 60 de *C. elegans*, assim como para identificar os fatores de transcrição SRF e C-fos ativados pela via MAPK. Para isso, um conjunto com 13 sequências de aminoácidos similares ao gene let-60 foi obtido no banco de dados de proteínas do NCBI e 41 sequências de aminoácido similares a SRF e 54 sequências similares a c-Fos foram obtidas do Uniprot (<http://www.uniprot.org>). Cada conjunto de sequência foi alinhado e editado manualmente. O HMM foi construído com o pacote HMMER, como descrito anteriormente, e os HMMs gerados foram utilizados nas buscas por homólogos no proteoma de *S. mansoni*.

#### **4.1-3. Classificação Funcional**

As proteínas quinase de *S. mansoni* foram caracterizadas por abordagens combinadas de similaridade de sequência e relações filogenéticas. A classificação funcional das proteínas quinase seguiu a hierarquia proposta e descrita em outros estudos (175-177) em que nove grupos, diversas famílias e subfamílias definem cada ePK. Dessa forma, em um primeiro momento, cada ePK identificada em *S. mansoni* foi classificada em seu respectivo grupo (AGC, CaMK, CMGC, CK1, STE, TK, RGC, TKL ou Other), utilizando dados do InterProScan (178) e PSI-BLAST (rpsblast) (66) integrandos ao programa Artemis (179). Em um segundo momento, as proteínas de um mesmo grupo ePK foram agrupadas para serem usadas no processo de inferência filogenética (ver item 4.1-4).

As ePKs de *S. haematobium* e *S. japonicum* não foram incluídas nas análises filogenéticas. Mas, foram classificadas a partir das sequências já anotadas de *S. mansoni*, utilizando os dados do orthoMCL (180), pré-computado no SchistoDB. Sendo assim, cada

ePK de *S. haematobium* e *S. japonicum* foi classificada a partir da transferência da anotação das ePKs ortólogas de *S. mansoni*.

O domínio catalítico e os domínios proteicos não catalíticos (domínios acessórios) de cada ePK de *S. mansoni*, *S. japonicum* e *S. haematobium* foram determinados com base em dados disponíveis no banco de dados de famílias de proteínas – Pfam (181).

#### 4.1-4. Inferência Filogenética

Sequências de aminoácidos correspondentes ao domínio catalítico conservado (Pfam: PF0069 ou PF07714) de cada grupo ePK, de *S. mansoni*, *H. sapiens*, *M. musculus*, *D. melanogaster*, *C. elegans* e *B. malayi*, foram alinhadas separadamente (um alinhamento por grupo) usando os parâmetros padrões do programa MAFFT (72). Os alinhamentos múltiplos gerados foram filtrados usando a ferramenta *Decrease Redundance* (182), para manter apenas sequências de proteínas que compartilhassem entre 50% a 90% de identidade, e editados manualmente para remover regiões ambíguas usando o *BioEdit* (173). Os alinhamentos finais foram utilizados nas reconstruções filogenéticas através de várias ferramentas disponíveis no *Phylogeny Inference Package* - PHYLIP, versão 3,69 (183), entre eles: *Seqboot*, *Protdist*, *Fitch* e *Consense*. Os programas do pacote PHYLIP foram executados em um servidor do Centro de Excelência em Bioinformática (CEBio) com as seguintes características: Processador Intel 2X Xeon X5460 Quad-Core de 3.16 Ghz, 6MB cache memory, 24 GB RAM, 1,5TB HD.

Inicialmente, 1.000 pseudo réplicas foram criadas para cada alinhamento usando *seqboot* com parâmetros padrão. O *Protdist* calculou, para cada conjunto de dados (1000 pseudo réplicas), uma matriz de distância baseado no modelo JTT e distribuição gama. Em seguida, filogenias foram estimadas a partir de dados da matriz de distância adotando o critério de *Fitch-Margoliash* através do programa *Fitch*. Finalmente, os resultados dos conjuntos de dados aleatórios foram resumidos no programa *Consense*, que calcula a árvore consenso pelo método *majority-rule consensus tree*. As árvores filogenéticas foram visualizadas e editadas usando *Tree Figure Drawing Tool*, *FigTree*, versão 1.3.1 (74). Nós com valores de apoio (*bootstrap*) de, no mínimo 80%, foram considerados significativos para apoiar as classificações e predições funcionais.

## 4.2- Interferência por RNA (RNAi)

### 4.2-1. Seleção dos genes alvos

A partir das análises filogenéticas, ePKs de *S. mansoni* foram selecionadas para os estudos experimentais. A escolha foi baseada no quão importante, em termos de função, é determinado gene para o ortólogo identificado na anotação e no potencial de especificidade para a transferência do sinal (proteínas únicas de uma família ou subfamília). Além disso, foram escolhidas ePKs que são prováveis sinalizadores da via MAPK, umas das vias mais bem estudadas em diversos organismos e que está envolvida em processos essenciais para a viabilidade celular em todos os eucariotos (184). Esses genes foram: SmCaMK2 (Smp\_011660.2), SmJNK (Smp\_172240), SmERK1 (Smp\_142050) e SmERK2 (Smp\_047900).

Por ser uma importante proteína na ativação da via MAPKs RTK/Ras/ERK, uma GTPase da família Ras - SmRas (Smp\_179910), similar à proteína let-60 de *C. elegans*, também foi escolhida como alvo para os estudos a seguir.

Em *H. sapiens*, proteínas MAPK ativadas fosforilam o fator de transcrição elk-1 que, por sua vez, forma um complexo com o fator de resposta ao soro – SRF. Ambos os fatores se ligam ao promotor do gene c-Fos iniciando a transcrição deste. Para entendermos como essa via é regulada em *S. mansoni*, os fatores de transcrição SRF SmSRF (Smp\_097730) e c-Fos SmC-Fos (Smp\_124600) foram selecionados.

### 4.2-2. Desenho dos iniciadores

Iniciadores específicos para síntese de dsRNAs e para reações de QPCR foram desenhados com base nas sequências preditas dos genes de interesse listados no item anterior. Utilizamos como controle negativo inespecífico do silenciamento, a proteína fluorescente verde (GFP - pCRII-GFP) e como controle normalizador das reações de QPCR o gene que codifica para a proteína C Oxidase (COX) de *S. mansoni* (GenBank AF216698). Os iniciadores foram desenhados pelo programa Primer 3 (<http://frodo.wi.mit.edu>) para amplificar em torno de 500pb para síntese de dsRNA e 100 a 150pb para reações de QPCR. O promotor da RNA polimerase T7 (taatacgactcactataggg) foi adicionado em cada extremidade 5' dos iniciadores utilizados na síntese de dsRNA. Os programas *Oligo analyzer*

(<http://www.idtdna.com/analyzer/Applications/OligoAnalyzer>),

*Amplifix*

(<http://ifrjr.nord.univ-mrs.fr/AmplifX-Home-page>)

e

*FastPCR*

(<http://primerdigital.com/fastpcr.html>) foram utilizados para verificar a possível formação de *primer-dimer*, grampos, estruturas secundária e para evitar a baixa especificidade na ligação na extremidade 3'. As seqüências dos iniciadores estão descritas no Quadro 1.

Quadro 1: Lista de iniciadores

Gene ID	Family	Extremidade 5'	Extremidade 3'
<b>DsRNA primers</b>			
Smp_142050	ERK1	taatacgactcactatagggTTGGTCAATTGGTTGTATTATGG	taatacgactcactatagggGGAACAATGGCACCAGGAAT
Smp_047900	ERK1	taatacgactcactatagggTCTGCCAGCGAACATATCG	taatacgactcactatagggGGATCACCAAGTCGTGAAGA
Smp_011660.2	CAMK2	taatacgactcactatagggGATGACATTCAGGACGAAGG	taatacgactcactatagggTCGCAGGACTGACTGTTAG
Smp_172240	JNK	taatacgactcactatagggACATGCAGCCGGTATAATCC	taatacgactcactatagggTACTTCAGAGTCTTCATACCATACG
Smp_179910	Ras	taatacgactcactatagggTGGCACCAGAACTTATCAGG	taatacgactcactatagggGATATAGAGCAGTCATTGCATTCC
pCRII-GFP	GFP	taatacgactcactataggggtcttcaagtcgceatg	taatacgactcactataggggtctcaggtagtggttgctc
<b>RT-qPCR primers</b>			
Smp_142050	ERK1	TGCAACATCTTGTGAATGC	GCACGATACCAACGTGTAGC
Smp_047900	ERK1	TTATCCTTCGGCGGATGC	AGCAACAGGCTCATCACTAGG
Smp_011660.2	CAMK2	ACGACTATGCTAGCCACACG	CAGACGATTCCTTAATACCATCG
Smp_172240	JNK	TCCTCCTGGGTATCATGTCCG	GCTACAACAAGCCCTGAGC
Smp_179910	Ras	GACTGAGTACAAGTTAGTTGTGTTGG	TTCTATAAGAGTCCCTATCGTTGG
Smp_124600	c-Fos	GAGGCTGCAAGAGAATGTCG	CAAAGTGCTTTAACTTTCTGAAGC
Smp_097730	SRE	GATACCTATTGAATTTATTTCTGATCG	CGGTTAATTCAGCCAATTCC
AF216698.1	COX	TACGGTTGGTGGTGTACACAG	ACGGCCATCACCATACTAGC

### 4.2-3. Obtenção dos parasitos

Os parasitos utilizados nesse trabalho são da espécie *S. mansoni* cepa LE, sendo que as cercárias foram provenientes de infecções do ciclo do parasito mantidos pelo Moluscário do CPqRR/FIOCRUZ e gentilmente cedidos para execução deste trabalho.

As cercárias obtidas foram distribuídas em alíquotas de 50mL e incubadas por 60 minutos no gelo. Em seguida, os tubos contendo as cercárias foram centrifugados a 1.000 rpm por 3 minutos a 4°C e o sobrenadante descartado. O sedimento foi ressuspensionado em 2mL de meio MEM contendo com 1% de penicilina/estreptomicina e o conteúdo de três tubos transferidos para um tubo de 15mLs (utilizando pipetas Pasteur descartáveis) e estes foram centrifugados à 1.000 rpm por 3 minutos.

Novamente, o sedimento foi ressuspensionado com 2mL de meio MEM contendo 1% de penicilina/estreptomicina. Os tubos foram submetidos à agitação máxima em vórtex por três vezes, durante 2 minutos, com intervalos de 1 minuto. O volume do tubo foi transferido para um frasco estéril de cultura de 50mL e o volume completado para 10mL com meio. A garrafa

de cultura foi incubada em estufa a 37°C, com atmosfera de 5% de CO<sup>2</sup> e 95% de umidade por 1 hora. Após a incubação, as amostras foram transferidas para tubos estéreis de 15mL dentro de fluxo laminar.

Os tubos contendo parasitos foram submetidos a sedimentação por 5 minutos. Em seguida, o sobrenadante foi retirado. O sedimento foi ressuspensionado em 5mL de meio suplementado e sedimentado por mais 5 minutos. Novamente, o sobrenadante foi removido e este passo repetido por mais cinco vezes diminuindo apenas o tempo de sedimentação (duas vezes 5 minutos, duas vezes 4 minutos, duas vezes 3 minutos). Após a última sedimentação, o sobrenadante foi ressuspensionado em 1mL de meio MEM (*Minimum Essential Medium Eagle* – *Sigma-Aldrich*) suplementado com meio *Schneider* (5%), HEPES (20mM), hidrolisado de lactalbumina (0,1%), hipoxantina ( $5 \times 10^{-7}$  M), serotonina ( $1 \times 10^{-6}$  M), triiodotironina ( $2 \times 10^{-7}$  M), hidrocortisona ( $1 \times 10^{-6}$  M), MEM vitaminas (0,5%) e soro fetal bovino - *Sigma* (1%). Após as lavagens, o sedimento foi observado com o auxílio de lupa para análise da presença de caudas e cercárias na suspensão.

Os parasitos foram incubados em garrafas de cultura de 50mLs, em B.O.D. a 37°C e cultivados durante 7 dias para posterior extração de RNA e síntese de cDNA utilizado para amplificação dos fragmentos, ou foram cultivados em placa 24 poços nos experimentos de silenciamento (ver item 4.2-8).

#### **4.2-4. Extração de RNA de esquistossômulos**

As extrações de RNA dos esquistossômulos foram realizadas utilizando o reagente TRIzol (*Invitrogen - Life Technologies*) e purificadas com o *RNeasy mini Kit* (QIAGEN). Resumidamente, os parasitos foram transferidos para tubos 1,5mL e centrifugados durante 1 minuto em baixa velocidade. O sobrenadante foi retirado, foram adicionados 50µL de TRIzol e os parasitos foram macerados com auxílio de um pistilo. Em seguida, foram adicionados 950µL de TRIzol e as amostras foram mantidas por 5 minutos a temperatura ambiente. Acrescentou-se 500µL de clorofórmio e as amostras foram agitas vigorosamente por 15 segundos e, logo em seguida, foram mantidos à temperatura ambiente por 3 minutos, com posterior centrifugação das amostras a 12.000 rpm por 15 minutos a 4°C. A fase aquosa, na qual se encontra o RNA, foi transferida para tubo novo e 280µL de etanol 70% foram adicionados e transferidos para a coluna do *RNeasy mini Kit* (QIAGEN). O tubo foi então centrifugado a 12.000 rpm a 4°C por 1 minuto. O sobrenadante foi removido e 700µL do

*Buffer RW1* foi acrescentado à coluna (fórmula proprietária do fabricante), que foi então levada a centrifugação por 15 segundos a 10.000 rpm. O eluato foi descartado e 500µL de solução RPE (fórmula proprietária do fabricante) foram adicionados à coluna. Mais um ciclo de centrifugação foi realizado por 15 segundos a 10.000 rpm. Em seguida, mais 500µL de solução RPE foram adicionados e a centrifugação dessa vez foi realizada por 2 minutos. As amostras foram centrifugadas por mais 1 minuto e a coluna transferida para tubo novo. Para eluição, 30µLs de água livre de RNAase foram acrescentados à coluna.

Objetivando a remoção de DNA genômico contaminante nas amostras de RNA, foi realizado o tratamento das amostras com DNase por 30 minutos a 37°C, utilizando o *kit TURBO DNA-free (Ambion - Life Technologies)*, segundo manual do fabricante. Posteriormente, a densidade óptica OD260/280 do RNA tratado foi medida utilizando o *Nanodrop Spectrometer ND-1000 (Thermo Fisher Scientific)*.

#### **4.2-5. Síntese de cDNA**

Para a síntese de cDNA foi utilizado o *kit SuperScript III System Rnase H (Invitrogen-Life Technologies)*. Ao RNA foram adicionados 1µL de oligo(dt) (500µg/mL) e 1µL de dNTP (10mM) em um volume final de 10µL. A amostra foi incubada a 65°C por 5 minutos e colocada imediatamente no gelo. Ao tubo foram acrescentados 4µL de tampão de síntese de primeira fita 5X (250mM Tris-HCl, pH=8,3; 375mM KCl; 15mM MgCl<sup>2</sup>), 2µL de DTT (0,1mM) e 1µL de *Rnase-Out* (40µg/mL, *Invitrogen- Life Technologies*). O tubo foi incubado a 42°C por 2 minutos e foi, em seguida, acrescentado 1µL de *SuperScript III* (200 U/mL). A reação foi então incubada a 42°C por 50 minutos e, em seguida, incubada a 70°C por 15 minutos. Para remoção de possíveis RNAs, que por ventura tenham permanecido no tubo, foi adicionado 1µL de RNase H e incubado 37°C por 20 minutos. As amostras foram armazenadas a -20°C para posterior utilização.

#### **4.2-6. Amplificação e sequenciamento**

Antes de iniciarmos a síntese dos dsRNAs os fragmentos dos genes selecionados foram amplificados e sequenciados para certificarmos que estávamos trabalhando com as sequências corretas e de qualidade desejável. Para amplificação foi utilizado o *kit Taq DNA polymerase (Invitrogen - Life Technologies)*, seguindo o protocolo sugerido pelo fabricante: MgCl<sub>2</sub> (1,5µL); dNTP (1µL); *Taq DNA polymerase* (0,2µL); 10X *PCR Buffer* (5µL); iniciador a 10 pmol (1µ de cada iniciador - 10pmol/µ); cDNA (1µL); água (quantidade

suficiente para 50 $\mu$ L). Utilizamos o termociclador *Gene Amp – PCR System 9700* (*Applied Biosystems*), com o seguinte ciclo de amplificação para todos os genes: Desnaturação inicial: 94°C por 2 minutos, desnaturação: 94°C por 40 segundos, anelamento: 55°C 40 segundos, extensão 72°C 2 minutos, extensão final: 72°C durante 10 minutos, o ciclo foi repetido 40 vezes.

Após amplificação, os produtos da reação de PCR foram aplicados em gel de agarose 1% (peso/volume) e a imagem foi digitalizada com o aparelho *ImageQuant LAS 4000* (*GE Healthcare Life Sciences*). As bandas foram extraídas e purificadas utilizando o *QIAquick Gel Extraction Kit* (QIAGEN), segundo orientações do fabricante.

Visando confirmar a autenticidade do produto amplificado, estes foram sequenciados. Para tal o kit de sequenciamento *DYEnamic ET Dye terminator* (*GE Healthcare Life Sciences*) foi utilizado, e as sequências foram obtidas através do sequenciador capilar de DNA MegaBACE 1000 (*GE Healthcare Life Sciences*). Em ambos os casos, os procedimentos foram executados segundo orientações do fabricante. Na placa de PCR para sequenciamento foi colocado o produto de PCR (concentração final 90ng/ $\mu$ L), 4 $\mu$ L pré-mix de sequenciamento *DYEnamic terminator*, 1 $\mu$ L de iniciador (10pmol/ $\mu$ L) para volume total de 10 $\mu$ L. O sequenciamento foi realizado com o uso dos iniciadores específicos para cada gene. Para a reação de sequenciamento foi utilizado o termociclador *MasterCycler* (Eppendorf) programado com os seguintes passos: 95°C por 20 segundos (desnaturação), 50°C por 15 segundos (anelamento) e 60°C por 1 minuto (extensão) por 25 ciclos consecutivos. A placa foi centrifugada em centrífuga 5804R (Eppendorf) e, em seguida, foi adicionado 1 $\mu$ L de 7,5M acetato de amônio e 27 $\mu$ L de etanol 100%. Após misturar os reagentes por inversão, a placa foi centrifugada a 800 rpm por 1 minuto, incubando-a, em seguida, por 10 minutos a temperatura ambiente e ao abrigo de luz. A placa foi então centrifugada a 3.700 rpm por 45 minutos e, em seguida, o sobrenadante foi descartado. 100 $\mu$ L de etanol 70% foram adicionados à placa que foi centrifugada a 3.700 rpm por 10 minutos. O sobrenadante foi descartado e a placa foi centrifugada invertida a 800rpm por 1 minuto para secagem do precipitado. Após 10 minutos, a reação de sequenciamento foi ressuspensa em 10 $\mu$ L de tampão de amostra do kit (70% formamida, 1mM EDTA) e a placa foi centrifugada a 3.700rpm por 2 minutos. A placa foi então sequenciada no aparelho de sequenciamento de DNA MegaBace1000 (*GE Healthcare*).

As sequências geradas pelo sequenciador foram comparadas às sequências proteicas preditas e depositadas no SchistoDB através de alinhamento múltiplo global realizado pelo programa ClustalX2 (<http://www.clustal.org>). Foi realizado também alinhamento múltiplo local através do BLASTn.

#### **4.2-7. Síntese de dsRNA**

Após confirmação dos fragmentos, os dsRNAs foram sintetizados utilizando o kit *T7 RiboMAX Express RNAi Systems* (Promega) seguindo o protocolo do fornecedor. Com o intuito de aumentar a concentração e a quantidade de dsRNAs produzidos, modificamos o protocolo para a produção final de 400µL de dsRNA. Para isso, na reação inicial, adicionamos 30µL do DNA linearizado, 35µL *RiboMAX™ Express T7 2X Buffer* e 5µL *Enzyme Mix T7 Express*. As amostras foram incubadas a 37°C por 16 horas, com o intuito de aumentar concentração dos dsRNAs e, em seguida, tratadas com DNase durante 30 minutos a 37°C. O produto foi precipitado com etanol e dissolvido em água livre de nucleases. A concentração das amostras foi medida utilizando o *Nanodrop Spectrometer ND-1000* (Nanodrop) e a integridade foi conferida em gel agarose 1% não desnaturante. Os dsRNAs foram armazenados a -70°C para posterior utilização.

#### **4.2-8. Exposição dos parasitos aos dsRNAs**

Após prévia padronização da cultura, os dsRNAs sintetizados, com tamanho aproximado de 500pb, foram inseridos na concentração de 100nM em culturas contendo aproximadamente 2 mil esquistossômulos e 2mL de meio MEM suplementado (ver item 4.2-3). Em cada uma das placas os tratamentos foram realizados em duplicata e tiveram a duração de dois, quatro e sete dias. As culturas foram observadas diariamente por microscopia de inversão (*ABO 100 - ZEISS*) para verificação de mudanças fenotípicas. Imagens dos parasitos foram registradas utilizando o software *Axion Vision REL 4*. O tratamento com dsRNA para cada um dos genes foi realizado em pelo menos 3 réplicas biológicas.

Após cada intervalo de tratamento, foi feito extração de RNA e síntese de cDNA, como descrito nos itens anteriores.

#### **4.2-9. Análise do silenciamento no nível da transcrição – QPCR**

Os níveis de transcritos dos genes silenciados em cultura de esquistossômulos foram avaliados por PCR em tempo real quantitativa (QPCR). Para a síntese de cDNA, as

mesmas quantidades de RNA total foram utilizadas para cada amostra dentro de um mesmo experimento. A síntese foi realizada utilizando a *SuperScript III reverse transcriptase* (Invitrogen – Life technologies) (ver item 4.2-5). Os iniciadores para os experimento de QPCR foram desenhados para amplificar fragmentos de 100 a 150 pb fora da região do dsRNA (Quadro 1).

Todos os iniciadores passaram, primeiramente, por uma padronização, para certificar a eficiência da amplificação e definição das concentrações. Para isso, três diferentes concentrações: 200 nM, 400 nM e 600 nM foram utilizadas para cada iniciador. O cDNA dos esquistossômulos foi utilizado em 5 diferentes concentrações (diluição seriada a partir de 100ng/μl de RNA total). O gene normalizador foi o Citocromo C Oxidase (COX) de *S. mansoni* (GenBank AF216698) (185;186). Esse gene é constitutivamente expresso em várias fases do ciclo do parasito (186;187). Os valores CT (*Cycle threshold*) foram plotados em função da diluição do DNA. A eficiência dos iniciadores foi medida através da análise das inclinações das curvas (entre -3,5 e -3,2) combinada aos cálculos dos coeficientes R<sup>2</sup> (foram considerados valores maiores que 0,98) indicando alta reprodutibilidade entre as triplicatas (188). Apenas os iniciadores que exibiram esses parâmetros foram considerados.

Para verificar o nível de transcrito dos genes silenciados, utilizamos as seguintes amostras: cDNA de esquistossômulos tratados com dsRNA dos genes selecionados e cDNA de vermes tratados com dsRNA-GFP (controle inespecífico). Na mesma placa, triplicatas de todas as amostras acima relacionadas foram amplificadas com iniciadores específicos e com iniciadores para o gene COX (controle para normalização). As reações foram realizadas no aparelho 7500 Real Time PCR System (*Applied Biosystems*) e em placas *96-Well Optical Reaction Plate* (*ABI PRISM, Applied Biosystems*) que continham, em cada poço, 12,5μl de *SYBR GREEN PCR Master Mix* (*Applied Biosystems*), 1,0μl de cDNA e 0,5μl de cada iniciador (200nM concentração final por reação), em um volume final de 25μl por reação. Em cada placa havia também um controle negativo (contendo todos os reagentes para síntese de cDNA e para amplificação, exceto o RNA) e controle de DNA genômico das amostras (onde todos os reagentes foram adicionados, com exceção da transcriptase reversa). A expressão relativa foi analisada através do método do  $\Delta\Delta CT$  (189) onde os valores Ct de cada amostra foram subtraídos do Ct do normalizador (COX) ( $\Delta Ct$ ). Em seguida, os valores correspondentes ao  $\Delta Ct$  encontrado para cada amostra foi subtraído do  $\Delta Ct$  normalizado. Os níveis de expressão de cada gene foram comparados com os níveis de expressão de transcritos

do mesmo gene em parasitos tratados com GFP. A significância dos resultados foi realizada pelo teste estatístico Wilcoxon-Sum of Ranks (Mann-Whitney, significância  $p > 0,05$ ) utilizando os  $\Delta\Delta CTs$  obtidos.

#### **4.2-10. Experimentos *in vivo*: infecção de camundongos com esquistossômulos silenciados**

Após silenciamento dos esquistossômulos durante dois dias, como descrito no item 4.2-8, 300 parasitos foram imersos em 0,2mL de meio de cultura MEM suplementado e inoculados em dose única, na região dorsal com aplicação subcutânea, em camundongos swiss (*Mus musculus*). Em cada experimento foram utilizados seis camundongos por gene silenciado, além de seis camundongos para o controle, onde parasitos não tratados com dsRNA foram utilizados. Três réplicas biológicas foram realizadas. Os animais foram mantidos no biotério por 35 dias e eutanasiados por deslocamento da cervical. Os possíveis vermes adultos foram então recuperados por perfusão de acordo com o descrito (190).

Após a perfusão, o fígado de cada camundongo foi retirado, pesado e acondicionado em tubos com KOH a 10% onde ficaram por 16 horas a 4°C e em seguida, incubados a 37°C, por uma hora. Os tubos foram centrifugados por cinco minutos a 2000 rpm e o sobrenadante foi descartado. O sedimento foi ressuspensionado com salina 0,85% para lavagem. Esta etapa foi repetida por mais duas vezes. Depois da última centrifugação, o sedimento foi ressuspensionado em 1mL de salina 0,85% e o número de ovos presentes em cada amostra foi contado por microscopia a partir de 3 alíquotas de 10 $\mu$ l.

Tanto o número de ovos quanto o número de vermes adultos recuperados para cada alvo silenciado nos três experimentos foram comparados com o número de ovos e vermes adultos recuperados nos camundongos infectados com esquistossômulos silenciados com o controle inespecífico GFP e com o controle negativo (esquistossômulos não tratados com nenhum dsRNA). Como as amostras não seguiam uma distribuição normal, a significância dos resultados foi realizada utilizando o teste de Mann-Whitney (*Wilcoxon-Sum of Ranks*,  $p < 0,05$ ,  $N \geq 3$ ).

#### **4.2-11. Estudos morfométricos e morfológicos de vermes adultos de *S. mansoni* recuperados após perfusão de camundongos infectados com esquistossômulos silenciados**

Amostras de vermes adultos recuperados após a perfusão de camundongos infectados com os esquistossômulos silenciados foram analisadas por microscopia confocal em busca de alterações fenotípicas. Para isso, os parasitos foram fixados em AFA (ácido acético 2%, formaldeído 10% e álcool 48%) e armazenados a temperatura ambiente para posterior utilização. Em seguida os vermes foram corados com carmim hidrocloreídrico 2,5%, desidratados por passagem em álcool 70%, 90% e 100%, clarificados com salicilato de metila e Bálsamo do Canadá (1:2) e montados individualmente em lâminas de vidro.

Análises morfométricas foram realizadas nos vermes macho e fêmea utilizando imagens computacionais (*Image Pro Plus - Media Cybernetics, EUA*) capturadas por uma câmera Sony acoplado (*640 x 480 pixels, RGB*) ao microscópio de luz (*Olympus BX50*). Os seguintes parâmetros foram determinados: número e área dos lobos testiculares, presença de tubérculos, a área do ovário, presença de ovos e glândulas vitelínicas e a integridade do tegumento. A significância dos resultados foi realizada utilizando o teste de Mann-Whitney (*Wilcoxon-Sum of Ranks, p<0,05*).

Imagens confocal (LSM-410, Zeiss), utilizando laser 488 nm He/Ne e filtro LP 585 no modo refletido foram usados para estudar o sistema reprodutivo de machos (lóbulo testicular) e fêmeas (ovário, útero, oótipo, e glândulas vitelínicas) além da integridade do tegumento e forma das ventosas oral e ventral de ambos sexos.

## 5. RESULTADOS E DISCUSSÃO

### 5.1- Proteínas quinase eucarióticas de três espécies *Schistosoma*

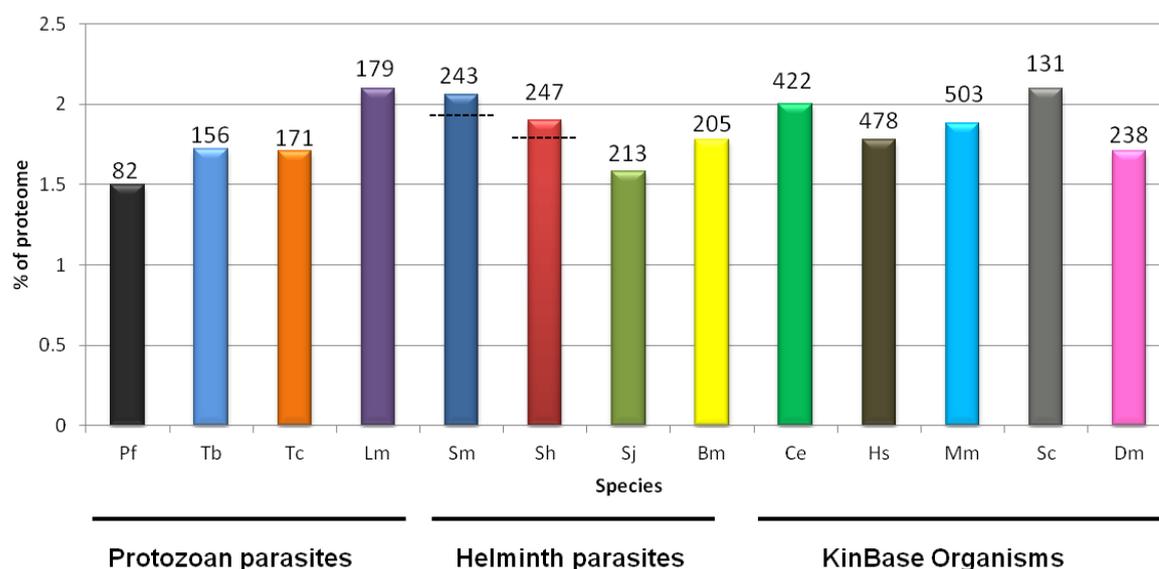
O conjunto de proteínas quinase eucarióticas do *S. mansoni*, definido por ePKinome, foi identificado através de buscas no proteoma predito do parasito utilizando HMMs que, por sua vez, foram construídos a partir do domínio catalítico de diferentes ePKs de cinco organismos selecionados no KinBase. HMMs são modelos estatísticos de alinhamentos múltiplos de sequência que capturam informações posição-específica sobre o quão conservado é cada coluna de um alinhamento e quais são os prováveis resíduos em cada coluna. O HMM faz um perfil da consulta que atribui um sistema de pontuação posição específico para substituições, inserções e deleções. Por causa da força dos seus modelos de probabilidade subjacentes, em comparação com BLAST e alinhamento de sequências e outras ferramentas de banco de dados de pesquisa, a utilização de HMMs é muito mais precisa e mais eficaz para detectar homólogos.

O ePKinome de *S. japonicum* e *S. haematobium* foi identificado a partir de HMMs integrados a versão atual do banco de dados SchistoDB, previamente construídos pelo PFAM e que definem proteínas que possuem o domínio quinase em suas sequências.

As análises revelaram 243 ePKs em *S. mansoni* (Tabela Anexo 1), 247 ePKs em *S. haematobium* (Tabela Anexo 3), e 213 ePKs em *S. japonicum* (Tabela Anexo 4). Esses valores correspondem a 2,06%, 1,9% e 1,58% do proteoma predito para os três parasitos, respectivamente. Algumas ePKs identificadas em *S. mansoni* (20 sequências) e *S. haematobium* (15 sequências) são produtos de *splicing* alternativo. Com isso, parte da diferença da proporção das ePKs identificadas entre as três espécies se deve ao fato de não existir predição de *splicing* alternativo para *S. japonicum*. O Gráfico na **Figura 7** mostra a porcentagem de ePKs encontradas em cada organismo e destaca a diferença ocasionada pela adição dos *splicing* alternativos em *S. mansoni* e *S. haematobium*.

A análise comparativa entre os ePKinomes de diferentes espécies de protozoários, helmintos e outros eucariotos cujas sequências estão depositadas do KinBase, mostrou que, apesar se haver uma grande variação no número de ePKs (de 82 a 503 sequências), a porcentagem em relação ao proteoma varia apenas entre 1,5 a 2,2% (**Figura 7**). O ePKinome das três espécies de *Schistosoma* analisadas neste trabalho segue este padrão. No entanto, entre *S. mansoni* (2,06%) e *S. japonicum* (1,58%) essa variação foi significativamente grande.

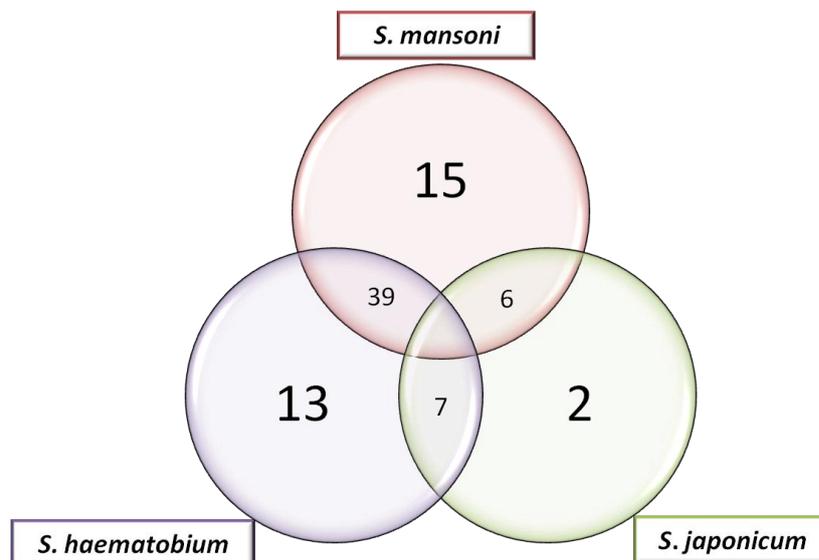
Parte disso foi ocasionada pela não inclusão dos produtos de *splicing* alternativo para *S. japonicum*, mas outra parte é provavelmente um efeito da fragmentação elevada desse último genoma. O genoma *S. japonicum* está dividido em 25.048 *scaffolds*, o de *S. haematobium* em 7.485 e o de *S. mansoni* em apenas 885. O grande número de *Scaffolds* propicia o maior número de erros durante a predição dos genes, pois exons que deveriam estar formando um único gene ficam separados em diferentes *scaffolds* e são preditos como parte de dois genes diferentes ou nem são preditos. Isso também explica parte do desaparecimento de 955 genes de *S. mansoni* entre a primeira versão do genoma, que estava dividida em 19.022 *scaffolds*, e a atual versão com apenas 885 *scaffolds*. Entre as ePKs de *S. mansoni* identificadas na primeira versão, 9 sequências não existem mais na nova versão. Dessas, 3 eram produtos de *splicing* alternativo e duas eram proteínas com classificação desconhecida (*unknown*) (Tabela Anexo 6).



**Figura 7: Porcentagem de ePKs identificadas em diferentes organismos.** Foram identificados um total de 243 ePKs em *S. mansoni*, 247 ePKs em *S. haematobium* e 213 ePKs em *S. japonicum*. Para comparação, a porcentagem (%) de proteínas que codificam para ePKs no proteoma predito e o número total de ePKs (mostrado no topo de cada barra) estão representadas para 4 parasitos protozoários: Pf - *Plasmodium falciparum* (191); Tc - *Trypanosoma cruzi*, Tb - *Trypanosoma brucei*, Lm - *Leishmania major* (192); dois helmintos parasitos: Bm - *Brugia malayi* (193) e *S. mansoni* (194); e organismos modelo do KinBase: Ce - *Caenorhabditis elegans* (195), Hm - *Homo sapiens* (196), Mm - *Mus musculus* (197), Dm - *Drosophila melanogaster* (198) e Sc - *Saccharomyces cerevisiae* (199). A linha tracejada (---) sob as barras destaca a porcentagem de ePKs quando não adicionados os produtos de *splicing* alternativo para *S. mansoni* (1,89%) e *S. haematobium* (1,78%).

Os possíveis ortólogos entre as ePKs das três espécies de *Schistosoma* foram preditas a partir de análises prévias realizadas pelo OrthoMCL (200;201) e integradas no banco de dados SchistoDB. O OrthoMCL é um algoritmo de escala genômica para o agrupamento de sequências de proteínas ortólogas. Ele fornece não apenas grupos compartilhados por duas ou mais espécies e genomas, mas também grupos de expansão de famílias gênicas espécie específicas. Assim, ele serve como uma importante ferramenta para anotação automatizada de genomas eucarióticos. O OrthoMCL primeiro computa os melhores hits recíprocos dentro de um genoma (pares parálogos) e depois relaciona os melhores hits recíprocos de qualquer par potencial entre dois genomas (pares ortólogos). A busca é feita a partir de ferramentas que se baseiam apenas na similaridade de sequências. Como estávamos trabalhando com espécies de um mesmo gênero, consideramos a escolha da metodologia para identificação de prováveis ortólogos e classificação das ePKs de *S. haematobium* e *S. japonicum*.

Para a maioria das ePKs um ou mais ortólogo entre *S. mansoni*, *S. haematobium* e *S. japonicum* foi identificado. No entanto, para 15 ePKs de *S. mansoni*, 13 de *S. haematobium* e 2 de *S. japonicum* nenhum ortólogo foi encontrado (**Figura 8**). Essas são, portanto, ePKs exclusivas e que definem cada uma das espécies. Observamos também que existem mais ePKs similares entre *S. mansoni* e *S. haematobium* (39 ePKs) do que entre *S. mansoni* e *S. japonicum* (6 ePKs) ou entre *S. haematobium* e *S. japonicum* (7 ePKs) (**Figura 8**). Tais diferenças podem diminuir à medida que novas e melhores versões dos genomas forem saindo ou podem ser fatos reais que caracterizam os ePKinomes de cada organismo. A análise comparativa de todo o genoma das três espécies realizado recentemente (202) mostrou que existe maior sintonia entre *S. haematobium* e *S. mansoni* (89.4%) do que entre *S. haematobium* e *S. japonicum* (51.7%) ou *S. mansoni* e *S. japonicum* (67.0%). Além disso, foi constatado pelos autores que *S. mansoni* está mais próximo filogeneticamente de *S. haematobium* do que de *S. japonicum* (203). Esses são dados similares ao observado aqui para as ePKs desses organismos.



**Figura 8: Relação de ortologia das ePKs entre três espécies de *Schistosoma*.** Cada círculo colorido representa uma espécie que está indicada nos retângulos de mesma cor. O número no centro de cada círculo indica a quantidade de ePKs que não possuem ortólogos nas outras duas espécies, sendo 15 em *S. mansoni*, 13 em *S. haematobium* e 2 em *S. japonicum*. Entre dois círculos está o número ePKs que são similares entre as duas espécies circunscritas. Algumas sequências têm mais de um ortólogo na mesma espécie, os parálogos, e estes foram incluídos na contagem. Sendo assim, entre *S. mansoni* e *S. haematobium* existem 39 ePKs que não possuem ortólogos em *S. japonicum*. Entre *S. mansoni* e *S. japonicum* encontramos 6 ePKs que não tem ortólogos em *S. haematobium*. Entre *S. haematobium* e *S. japonicum* existem 7 ePKs que não tem ortólogo em *S. mansoni*. Para o restante das ePKs foram encontrados ortólogos nas três espécies e esses dados não foram representadas nessa figura.

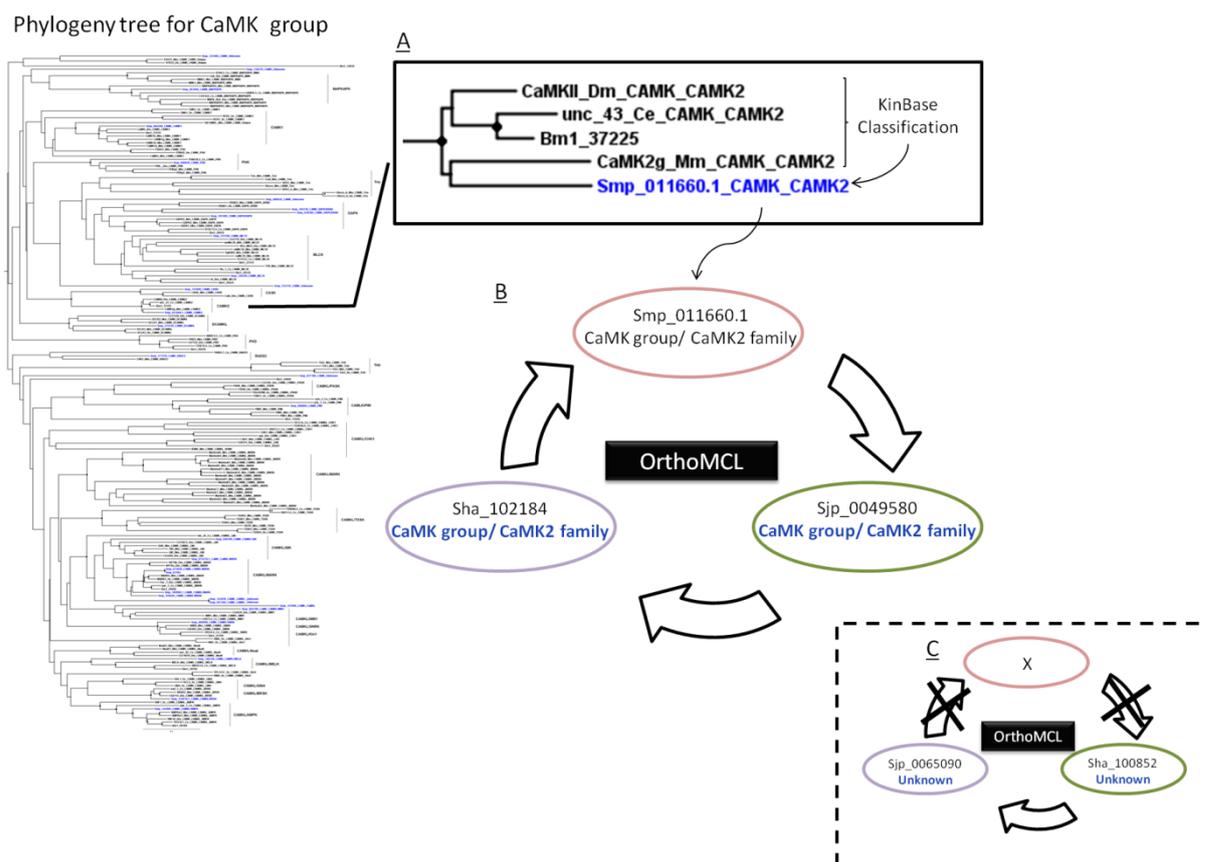
Baseado no mapeamento genético atualizado na última versão do genoma de *S. mansoni* (204;205), das 243 ePKs identificadas nesse organismo, 76,5% têm sua localização definida em um dos sete cromossomos autossômicos ou no cromossomo sexual (Tabela Anexo 1). A maioria das ePKs está localizadas no cromossomo 1 (19 %) ou nos cromossomos sexuais (18,5%). O restante está dividido nos cromossomos 2 (7,8 %), 3 (8,6%), 4 (7,4%), 6 (6.5%), 5 e 7 (2% cada). Como dito anteriormente, no *S. mansoni* o sexo é determinado pelos cromossomos sexuais, sendo a fêmea heterogamética (ZW) e o macho homogamético (ZZ). Criscione e colaboradores (2009) (206) mostraram que a maior parte dos cromossomos sexuais de *S. mansoni* sofrem recombinação, com exceção de uma pequena região de aproximadamente 18 Mb. Recentemente, Grunau e colaboradores (2012) (207) constataram que 70% das regiões não repetitivas do cromossomo sexual Z e W de *S. mansoni* são idênticas (pseudoautosomos). As sequências específicas do cromossomo W são na sua maioria representadas por regiões de repetição localizadas na heterocromatina e nenhum gene

específico foi identificado. No entanto, 15 *scaffolds* foram preditos como sendo específicos do cromossomo Z e neles 205 genes putativos foram encontrados (207). Destes, seis genes (Smp\_165470, Smp\_011700, Smp\_158720, Smp\_171610, Smp\_137610 e Smp\_134260) são ePKs identificadas nesse trabalho. Sendo assim, pode-se deduzir que essas ePKs têm alguma função relacionada a diferenciação sexual de *S. mansoni*, apesar de análises adicionais para esclarecer o real papel dessas proteínas serem necessárias.

Um trabalho analisou o perfil transcricional de vermes adultos de *S. mansoni* após tratamento com praziquantel (208). Dos genes diferencialmente expressos, 27 são ePKs identificadas nesse trabalho (Tabela Anexo 1) e destas, 81% estão super expressas em relação ao controle não tratado.

## 5.2- Classificação das ePKs identificadas em *S. mansoni*, *S. haematobium* e *S. japonicum*

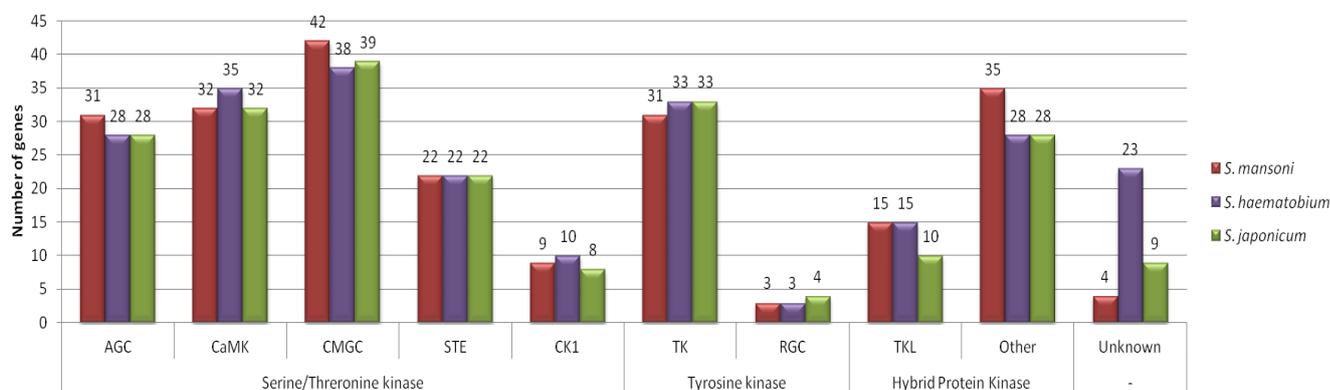
As sequências de aminoácidos correspondente ao domínio catalítico das ePKs de *S. mansoni* foram alinhadas utilizando o MAFFT (72) e as análises filogenéticas, baseadas no método de distância, foram implementadas (209). Para cada grupo de ePKs de *S. mansoni* foram também incluídas as ePKs homologas de seis outros organismos: *Homo sapiens*, *Mus musculus*, *Drosophila melanogaster*, *Caenorhabditis elegans*, *Saccharomyces cerevisiae* e *Brugia malayi* (**Figura 9**). Utilizando essa metodologia foi possível classificar cada ePK de *S. mansoni* em seu respectivo grupo, família e subfamília, baseado na hierarquia proposta por outros autores (149;210-212) e, em alguns casos, esclarecer o papel funcional e evolutivo dessas proteínas. A partir da classificação feita para as ePKs de *S. mansoni* e baseado nos dados de ortologia do OrthoMCL, as ePKs de *S. haematobium* e *S. japonicum* foram classificadas (**Figura 9**). As informações detalhadas estão disponíveis nas Tabelas 1, 2, 3 e 4, em anexo, que contêm, entre outros dados, todas as ePKs de *S. mansoni*, *S. haematobium* e *S. japonicum* com os identificadores correspondentes e ligados ao banco de dados SchistoDB (213). No SchistoDB é possível acessar as informações da anotação já disponível para os três genomas além de ferramentas para analisar e recuperar os dados.



**Figura 9: Esquema de classificação para as ePKs de *S. mansoni*, *S. haematobium* e *S. japonicum*.** As árvores filogenéticas foram construídas para cada grupo ePK e nelas foram incluídas sequências de aminoácidos correspondente ao domínio catalítico das ePKs de *S. mansoni* e outras seis espécies cujas ePKs já classificadas estão depositadas no KinBase. (A) Valores de Bootstrap (1000 pseudo réplicas) iguais ou maiores que 80% (•) foram considerados para transferir a anotação do KinBase para as ePKs de *S. mansoni* (Exemplo: Smp\_011660.1 que foi classificada como sendo do grupo CaMK e da família CaMK2). (B) Para cada ePK de *S. mansoni* buscamos pelos possíveis ortólogos em *S. haematobium* e *S. japonicum* com base nos dados e ortologia do OrthoMCL. As sequências ortólogas antes sem classificação foram então anotadas. (C) Em alguns casos as sequências de *S. haematobium* e *S. japonicum* possuíam ortólogos entre si, mas nenhum ortólogo em *S. mansoni*. Nesses casos as ePKs ficaram sem classificação.

Como mostra a **Figura 10**, as ePKs de *S. mansoni*, *S. haematobium* e *S. japonicum* têm representantes no principais grupos (AGC, CaMK, CK1, CMGC, RGC, STE, TK e TKL). As ePKs que não se encaixam nesses grupos são incluídas no grupo Other, que por sua vez, possui diversas famílias já definidas (143). O grupo STE é o único que possui a mesma quantidade de ePKs nas três espécies de *Schistosoma*. Para os demais grupos nenhuma diferença significativa foi observada. Entre as ePKs sem classificação - unknown (ePKs que possuem o domínio quinase característico mas que não se encaixam em nenhum grupo, nem

mesmo no grupo Other), *S. mansoni* se destaca por possuir o ePKinome mais bem classificado (1,6% do ePKinome dessa espécie continua sem classificação) e *S. haematobium* é aquele com maior número de ePKs sem classificação (9,2% do ePKinome dessa espécie). Isso porque as ePKs de *S. haematobium* e *S. japonicum* foram anotadas com base nas possíveis ePKs ortólogas de *S. mansoni*. Sendo assim, todas as ePKs dessas duas espécies que não possuem ortólogo em *S. mansoni* não puderam ser classificadas (**Figura 9C**). Temos como uma das perspectivas desse trabalho a inclusão das ePKs de *S. haematobium* e *S. japonicum* nas análises filogenéticas pois isso permitirá uma melhor anotação das mesmas. A porcentagem de ePKs sem classificação para *C. elegans*, *D. melanogaster* e *S. cerevisiae* está em torno de 19% a 38% do seus ePKinomes. Nesse contexto, os dados nos mostra que a metodologia utilizada nesse trabalho para classificar as proteínas quinase foi eficiente.



**Figura 10: Distribuição das ePKs de *S. mansoni*, *S. haematobium* e *S. japonicum* em seus grupos.** Os números acima de cada barra representam a contagem exata de genes (os produtos de *splicing* alternativo não foram incluídos nessa contagem) sendo as barras vermelhas as ePKs de *S. mansoni*, as barras azuis as ePKs de *S. haematobium* e as barras verdes as ePKs de *S. japonicum*. Os grupos ePKs incluem: AGC (*cAMP-dependent protein kinase/protein kinase G/protein kinase C extended*), CaMK (*Calcium/Calmodulin regulated kinases*), CMGC (*Cyclin-dependent Kinases and other close relatives*), STE (*MAP Kinase cascade kinases*), CK1 (*Cell Kinase I*), TK (*Protein Tyrosine Kinase*), RGC (*Receptor Guanylate Cyclases*), TKL (*Tyrosine Kinase Like*) e o grupo Other. Unknown são ePKs que não foram incluídas em nenhum dos grupos acima descrito.

As três espécies de *Schistosoma* têm mais representantes no grupo CMGC, uma característica única desse gênero e, tem como menor grupo o RGC, um característica comum a de vários outros organismos analisados na Tabela 3. Uma explicação para a grande quantidade de ePKs do grupo CMGC é que as proteínas desse grupo são necessárias no

controle da proliferação celular e para garantir a replicação e segregação correta de organelas. Juntos esses mecanismos são essenciais para parasitos que possuem um complexo ciclo de vida como é o caso de *Schistosoma*.

Tabela 3: Distribuição das ePKs em seus respectivos grupos para *S. mansoni*, *S. haematobium*, *S. japonicum* e organismos modelo

ePKs Groups	<i>Schistosoma mansoni</i>	<i>Schistosoma haematobium</i>	<i>Schistosoma japonicum</i>	<i>Brugia malayi</i>	<i>Caenorhabditis elegans</i>	<i>Homo sapiens</i>	<i>Drosophila melanogaster</i>	<i>Saccharomyces cerevisiae</i>
AGC	33	28	28	22	29	63	30	17
CaMK	37	40	32	41	41	83	34	22
CMGC	44	43	39	33	47	63	34	23
STE	27	22	22	27	24	46	19	14
CK1	9	11	8	31	87	12	10	4*
TK	31	37	33	35	85	90	34	-
RGC	3*	3*	4*	4*	27	5*	6*	-
TKL	18	16	10	12	15*	43	17	-
Other	37	28	28	-	67	81	44	37
Total	243	248	213	160	320	391	178	76

Os \* destacam o grupo menos representado para cada organismo e os números na cor vermelha o grupo com mais representantes em cada organismo

Das 243 ePKs identificadas em *S. mansoni* apenas 24 foram experimentalmente estudadas como mostra a Tabela Anexo 1. Para as ePKs de *S. mansoni* foi verificado ainda a presença dos 12 subdomínios conservados no domínio catalítico e também a presença da lisina no subdomínio II a dos resíduos de ácido aspártico no subdomínio VIb e VII, que têm papel essencial na atividade catalítica dessas enzimas (214-216). De acordo com essa análise, 12 ePKs de *S. mansoni* foram preditas como sendo cataliticamente inativas já que um ou mais desses três resíduos essenciais não estão conservados (Tabela Anexo 1). Isso inclui todos os membros do grupo RGC, em que o resíduo de ácido aspártico do subdomínio VIb foi substituído por um R (Arginina), N (Asparagina), ou W (Triptofano), todos são aminoácidos com características físico químicas diferentes do ácido aspártico (**Figura 11**). As proteínas do grupo RGC de *S. japonicum* e *S. haematobium* possuem exatamente as mesmas substituições nesse resíduo (**Figura 11**), com exceção da ePK de *S. japonicum* (Sjp\_0091120) em que não foi possível verificar essa substituição pois o gene tem o domínio catalítico truncado, provavelmente ocasionado por um erro de predição.

```

Sha_102682      VLRRDSIPLSWSLRIGFLTGLANGLAYLHHCQIVHG-RLNSSNCVPRLIMITFSLMSD---
Smp_085680      VLRRDSIPLSWSLRIGFLTGLANGLAYLHNYHIVHG-RLNSSNCVVSDTWTCKITDYGLD
Sjp_0082080     -----ILGFLTGLANGLAYLHNYPIVHG-RLNSSNCVVNDKWICKITDYGLD
Sha_106688      VLENEQIKLDWMFKFSLMQDICRGVMYLHQIFGPHG-NLKSSNCLVDSRFVLKITDFGLP
Smp_142620      VLENEQIKLDWMFKFSLMQDICRGVMYLHQIFGPHG-NLKSSNCLVDSRFVLKITDFGLP
Sjp_0006550     VLENEQIKLDWMFKFSLMQDICRGVMYLHQIFGPHG-NLKSSNCLVDSRFVLKITDFGLP
Sha_101830      I IKKESINLDWEFKLSLITDVVVKGMRYIHSSPIKKHGWLKSTNCCVDGRWVVKITDYGLT
Smp_125450      I IKKDSINLDWEFKLSLITDVVVKGMRYIHSCPIKKHGWLKSTNCCVDGRWVVKITDYGLP
Sjp_0091120     -----

```



DxxxxN - Subdomain VIb

**Figura 11: Análise dos resíduos conservados no subdomínio VIb no domínio catalítico de ePKs do grupo RGC de *S. mansoni*, *S. haematobium* e *S. japonicum*.** Parte do alinhamento múltiplo das ePKs do grupo RGC mostrando a substituição do resíduo de ácido aspártico (D) na região conservada DxxxxN, em que x é qualquer aminoácido. As ePKs de *S. mansoni* estão representadas pelos identificadores Smp\_085680, Smp\_142620 e Smp\_125450; as de *S. haematobium* pelos identificadores Sha\_102682, Sha\_106688 e Sha\_101830; e as de *S. japonicum* pelos identificadores Sjp\_0082080, Sjp\_0006550 e Sjp\_0091120, sendo que este último possui o domínio quinase truncado.

### 5.3- Proteínas Serina/Treonina Quinase

#### 5.3-1. Análise das ePKs do grupo AGC encontradas em *Schistosoma*

Aproximadamente 13 famílias foram classificadas como membros do grupo AGC de eucariotos (143). Nas espécies de *Schistosoma* analisadas nesse trabalho, 12 famílias estão presentes, sendo a maioria pertencente às famílias PKA (*Protein Kinase A*) (5 ePKs para cada espécie de *Schistosoma*), PKG (*Protein Kinase G*) (4 ePKs em *S. mansoni* e *S. japonicum* e, 3 ePKs em *S. haematobium*), DMPK (*Myotonic Dystrophy Protein Kinase*) (4 ePKs e 1 produto de *splicing* alternativo para *S. mansoni* e *S. haematobium*), e PKC (*Protein Kinase C*) (4 ePKs em *S. mansoni* e *S. haematobium*, e 5 em *S. japonicum*) (Quadro 2). A família MAST que possui um representante em *S. mansoni* não possui ortólogos nas outras duas espécies de *Schistosoma*. Já para a família AKT, que possui um representante em *S. mansoni* e um em *S. haematobium*, foram encontradas três ePKs em *S. japonicum* (Tabela Anexo 2).

Nas proteínas da família PKA as atividades regulatórias e catalíticas são realizadas por genes diferentes, conhecidos por PKA-R e PKA-C, respectivamente. Já nas proteínas PKG, o domínio *cNMP-binding* (PF00027), responsável pela atividade regulatória, e o domínio catalítico estão presentes no mesmo polipeptídeo. Sendo assim, a conformação inativa de uma PKA é um heterotetrâmero formado por duas subunidades PKA-R e duas PKA-C, enquanto proteínas PKG são homodímeros (217).

A subunidade PKA-C possui o domínio catalítico quinase e cinco homólogos PKA-C foram preditos nesse trabalho como ePKs em *S. mansoni*, *S. haematobium* e *S. japonicum*. Swierczewski e colegas (2012) (218) ainda relataram que as sequências dos ortólogos entre as três espécies são altamente conservadas com mais de 99% de identidade entre suas sequências de aminoácidos. Além disso, seis subunidades PKA-R estão no proteoma predito *S. mansoni* (Smp\_131050, Smp\_147320, Smp\_079010, Smp\_030400, Smp\_019280, Smp\_022100), cinco em *S. haematobium* (Sha\_104102, Sha\_100985, Sha\_107430, Sha\_104580 e Sha\_103301) e cinco em *S. japonicum* (Sjp\_00418990, Sjp\_0121360, Sjp\_0056650, Sjp\_0116860 e Sjp\_0079660) o que possibilita a formação de várias e diferentes haloenzimas nesses parasitos. Todas as subunidades PKA-R preditas nas três espécies de *Schistosoma* possuem o domínio *cNMP-binding* (PF00027) responsável pela atividade regulatória dessas proteínas.

Alguns estudos demonstraram que proteínas PKG de *Toxoplasma* (219) e *Eimeria* (220) e, proteínas PKG e PKA de *Plasmodium* (221;222) são essenciais para a sobrevivência desses organismos. Além disso, foi demonstrado recentemente que a inibição da subunidade catalítica de uma PKA (SmPKA-C/Smp\_152330), expressa em vermes adultos de *S. mansoni*, resulta na morte dos parasitos (115). Esse resultado, assim como a variedade de haloenzimas que podem ser formadas pelas PKAs de *Schistosoma*, indica que os genes dessa família podem ser essenciais para o desenvolvimento do parasito e assim representar bons alvos para o desenvolvimento de droga.

Uma ePK da família PKG foi recentemente estudada em *S. mansoni* e, utilizando inibidores específicos para essa família para tratar vermes adultos, os autores observaram vários fenótipos, menos drásticos do que os conseguidos com a inibição de proteínas PKA, mas que incluem pouca movimentação e produção reduzida de ovos (223).

Proteínas PKC pertencem a uma grande família que é classificada em quatro importantes subfamílias: PKC alpha, que contém as PKCs convencionais ( $\gamma$ ,  $\beta$ I,  $\beta$ II e  $\alpha$ ) e são sensitivas ao diacilglicerol (DAG) e  $Ca^{2+}$ ; as subfamílias PKC Eta e Delta que contém as PKCs *novel* ( $\epsilon$ ,  $\delta$ ,  $\eta$  e  $\theta$ ) que são reguladas apenas por DAG; e a subfamília PKC Iota, que contém as PKCs atípicas ( $\zeta$  and  $\iota$ ) e são insensíveis a DAG e ao  $Ca^{2+}$  (224;225). As subfamílias PKC são consideradas reguladoras do desenvolvimento em vertebrados, pois têm um papel importante no crescimento e diferenciação (226-228). Em *S. mansoni* (**Figura**

**Anexo 1)** assim como em *S. haematobium* e *S. japonicum* encontramos representantes em três das quatro subfamílias mencionadas acima (Iota, Eta e Alpha) mas não encontramos homólogos da subfamília delta, que está presente em *C. elegans*, *D. melanogaster*, *M. musculus* e *H. sapiens*. Em *C. elegans*, a PKC da subfamília delta parece participar de vias de sinalização envolvidas no processo de alimentação e crescimento (229). As duas PKC da subfamília Alpha que foram encontradas em *S. mansoni* (Smp\_128480 e Smp\_176360), pertencem a isoforma PKC $\beta$ I e foram caracterizadas experimentalmente (230-232). As duas foram encontradas na massa neural, vesícula excretora, tegumento e células germinativas em esquistossômulos e miracídio, sugerindo um possível papel na transformação dos parasitos (233-235).

Uma ePK de *S. mansoni* do grupo AGC, Smp\_157370, continua sem classificação com relação a família. Na árvore filogenética do grupo, essa ePK aparece próxima ao agrupamento da família GRK (**Figura Anexo 1**) mas, apesar da boa conservação do domínio catalítico, o domínio acessório característico da família GRK não foi encontrado. Smp\_157370 possui ainda um ortólogo em *S. haematobium* (Sha\_101172) e um em *S. japonicum* (Sjp\_0049550).

### 5.3-2. Análise das ePKs do grupo CaMK encontradas em *Schistosoma*

O cátion de cálcio bivalente (Ca<sup>2+</sup>) é um dos íons mais amplamente utilizados como mensageiro secundário na sinalização celular, sendo que grande parte desse processo é controlado pelas kinases ligadoras de calmodulina (CaMK). Por isso, alguns membros do grupo CaMK são dependentes da ligação ao Ca<sup>2+</sup>/CaM (236). No ePKinome de *S. mansoni*, 32 genes foram classificados no grupo CaMK sendo a maioria destes (14 genes que codificam 17 proteínas) pertencentes à família CaMKL (*Calcium/Calmodulin Regulated Kinase - like*). Um número similar foi encontrado em *S. haematobium*, *S. japonicum* e nos os outros organismos analisados no Quadro 2. *Schistosoma* ainda possui representantes nas famílias DAPK (*death associated protein kinase*), MAPKAPK (*MAPK associated protein kinase*), MLCK (*myosin light chain kinase*) e PHK (*phosphorylase kinase*) (Tabela Anexo 2 e **Figura Anexo 2**).

MLCK é uma família dependente de Ca<sup>2+</sup>/CaM cujo único substrato conhecido é a miosina II reguladora da cadeia leve (RLC- *myosin II regulatory light chain*) (237). A função primária das proteínas MLCK é estimular a contração muscular através da fosforilação de

RLC, uma proteína motora encontrada em eucariotos que interage com a actina nos processos de contração muscular. Apesar de MLCK possuir apenas um substrato conhecido, essas proteínas participam de vários processos celulares devido às diversas funções de RLC (238). O gene MLCK é encontrado em quatro diferentes isoformas em eucariotos: MYLK (musculatura lisa), MYLK2 (musculatura esquelética), MYLK3 (musculatura cardíaca) e MYLK4 (*novel*) (239). Uma proteína MLCK específica de musculatura lisa foi identificada no ePKinome de *S. mansoni* (Smp\_121780). Esta proteína tem seus ortólogos em *S. haematobium* e *S. japonicum* as ePKs Sjp\_0015670 e Sha\_100927, respectivamente. Um trabalho do nosso grupo identificou o perfil transcricional de vermes adultos de *S. mansoni* após tratamento com praziquantel (208) e Smp\_121780 se mostrou muito expresso em relação ao controle não tratado (Tabela Anexo 1). Já se sabe que o praziquantel atua nos canais de cálcio aumentando a permeabilidade desse íon e induz a contração muscular dos parasitos resultando em paralisia (240;241). Diante dessas evidências, possivelmente Smp\_121780 é uma ePK envolvida no processo de morte dos parasitos após tratamento com praziquantel.

Outra MLCK foi reconhecida em *S. mansoni* - Smp\_126240 e esta é a maior ePK identificada nesse parasito. Sua sequência com 6.010 aminoácidos possui diversos domínios proteicos I-set (*Immunoglobulin I-set*) e fn3 (*Fibronectin type III*) que caracterizam as proteínas MLCK denominadas titina. Proteínas titina são geralmente as maiores proteínas encontradas em eucariotos. Elas funcionam como uma mola molecular que é responsável pela elasticidade passiva do músculo. É composta por diversos domínios ligados por sequências peptídicas não estruturadas. Estes domínios desdobram quando a proteína é esticada e redobram quando a tensão acaba sendo importantes no processo de contração da musculatura estriada (242). Em *S. haematobium* duas sequências (Sha\_100748 e Sha\_101294) ortólogos a Smp\_126240 foram identificadas, mas nenhum ortólogo foi encontrado em *S. japonicum*.

A família DCAMLK (*Doublecortin and CaMK-like*) está envolvida na regulação do citoesqueleto dos microtúbulos em mamíferos e em galinhas está especificamente expressa nos cérebros em desenvolvimento (243;244). Um único homólogo á essa família foi encontrado em *S. mansoni*, *S. haematobium* e *S. japonicum*.

Proteínas da família CASK participam dos processos de adesão celular (245) e, de acordo com as análises filogenéticas, um único homólogo a essa família foi identificado em *S. mansoni*. Identificamos também um ortólogos em *S. haematobium* e três em *S. japonicum* (Tabela Anexo 2).

Enquanto a família CaMK2 (*CaMK family 2*) é representada por 4 genes em humanos, um único gene CaMK2, com dois produtos de splicing alternativo, foi encontrado em *S. mansoni*. A CaMK2 de *S. mansoni* foi destacada como bom alvo para o desenvolvimento de drogas através de uma abordagem comparativa entre o proteoma de *S. mansoni* e dois organismos modelo, *C. elegans* e *D. melanogaster* (246). A função dessa proteína em *S. mansoni* ainda é desconhecida. No ouriço do mar, CaMK2 é necessária para a ruptura do envelope nuclear seguido de fertilização (247) e em *C. elegans* ela está envolvida na ativação da via de sinalização MAPK responsável pelo desenvolvimento dos vermes (248). Outra ePK, também envolvida da ativação de proteínas MAPK, é a Smp\_053560 que é único membro da família MAPKAPK em *S. mansoni*. O ortólogo a esse gene em *C. elegans* produziu um fenótipo de letalidade em embriões após silenciamento do gene (Tabela 4).

### 5.3-3. Análise das ePKs do grupo CMGC encontradas em *Schistosoma*

O grupo CMGC é o que possui a maior quantidade de ePKs em *Schistosoma* (Tabela 3). As principais famílias do grupo estão conservadas entre *Schistosoma*, *S. cerevisiae*, *C. elegans*, *M. musculus*, *H. sapiens* e *D. melanogaster*, incluindo CDK (*Cyclin Dependent Kinase*), MAPK (*Mitogen Activated Protein Kinase*), GSK (*Glycogen Synthase 3 Kinase*), CLK (*CDC-Like Kinase*), SRPK (*SR Protein Kinase*), CK2 (*Cell Kinase 2*), DYRK (*Dual-specificity Tyrosine Regulated Kinase*) e RCK (Quadro 2 e **Figura Anexo 3**).

*S. mansoni*, *S. haematobium* e *S. japonicum* possuem 14 ePKs da família CDK (Quadro1) incluindo representantes em todas as subfamílias atualmente conhecidas (CDK7, CDK4, CDK8, CRK7, CDK9, PITSLRE, CDK10, PCTAIRE, PFTAIRE, VDK5 e CDC2), sendo esta a maior família do grupo CMGC. Em *C. elegans* número semelhante de proteínas CDK (13 ePKs) foi identificado comparado com apenas 7 em *S. cerevisiae*. Por outro lado, para outras quatro famílias (RCK, SRPK, CK2 e CLK) um único representante foi identificado em *Schistosoma*.

Quadro 2: Distribuição das principais famílias ePKs em *Schistosoma* e organismos modelo

	AGC Group					CaMK Group				CK1 Group					CMGC Group			
	PKA	DMPK	PKC	RSK	PKG	CaMKL	CAMK2	MLCK	TSSK	CK1	TTBK1	Worm6	CK1-unique	Dual	CDK	MAPK	DYRK	GSK
<i>S. mansoni</i>	5	4	4	3	4	14	1	3	0	6	0	0	0	0	14	9	8	3
<i>S. haematobium</i>	5	4	4	3	3	16	1	3	0	7	0	0	0	0	14	6	9	3
<i>S. japonicum</i>	5	4	5	2	4	13	2	1	0	5	0	0	0	0	14	8	6	2
<i>C. elegans</i>	2	2	4	4	2	16	1	1	4	3	31	28	6	3	13	14	6	7
<i>H. sapiens</i>	5	7	9	8	2	20	4	2	5	7	0	0	0	0	20	14	10	2
<i>M. musculus</i>	3	7	9	8	2	44	4	5	6	6	0	0	0	0	19	14	10	2
<i>D. melanogaster</i>	4	3	5	3	3	13	1	5	5	7	0	0	0	0	12	7	5	3
<i>S. cerevisiae</i>	3	0	1	3	0	12	0	0	0	4	0	0	0	0	7	6	1	4

	RGC Group	STE Group			TK Group							TKL Group		Other Group				
	RGC	STE20	STE11	STE7	Src	EGFR	InsR	Fer	TK-unique	KIN16	KIN6	STKR	MLK	Nek	PEK	SCY1	Haspin	Other-unique
<i>S. mansoni</i>	3	13	2	7	5	3	2	1	0	0	0	6	4	5	3	5	0	0
<i>S. haematobium</i>	3	12	3	7	5	3	2	1	0	0	0	5	4	4	3	2	0	0
<i>S. japonicum</i>	3	13	2	7	5	3	2	2	0	0	0	4	3	3	3	2	0	0
<i>C. elegans</i>	27	12	2	10	3	1	1	37	7	15	5	3	7	4	3	2	13	11
<i>H. sapiens</i>	5	30	8	7	11	4	3	2	1	0	0	12	10	11	4	3	1	3
<i>M. musculus</i>	7	30	8	7	11	4	3	2	1	0	0	12	10	11	4	3	1	3
<i>D. melanogaster</i>	6	12	2	4	2	1	1	1	5	0	0	5	6	2	2	4	1	1
<i>S. cerevisiae</i>	0	5	5	4	0	0	0	0	0	0	0	0	0	1	1	1	2	2

As proteínas RCK são semelhantes à proteína MAK (*male germ cell-associated kinase*) que foi primeiramente identificada em mamíferos e estão envolvidas na meiose espermatogênica e em vias de tradução para visão e cheiro (249). Não se sabe a real função dessa ePK em *Schistosoma*. No entanto, a RCK de *S. mansoni* (Smp\_132890) foi identificada no cromossomo sexual W (Tabela Anexo 1) reforçando a hipótese que essa proteína esteja envolvida nos processos de reprodução sexuada dos parasitos.

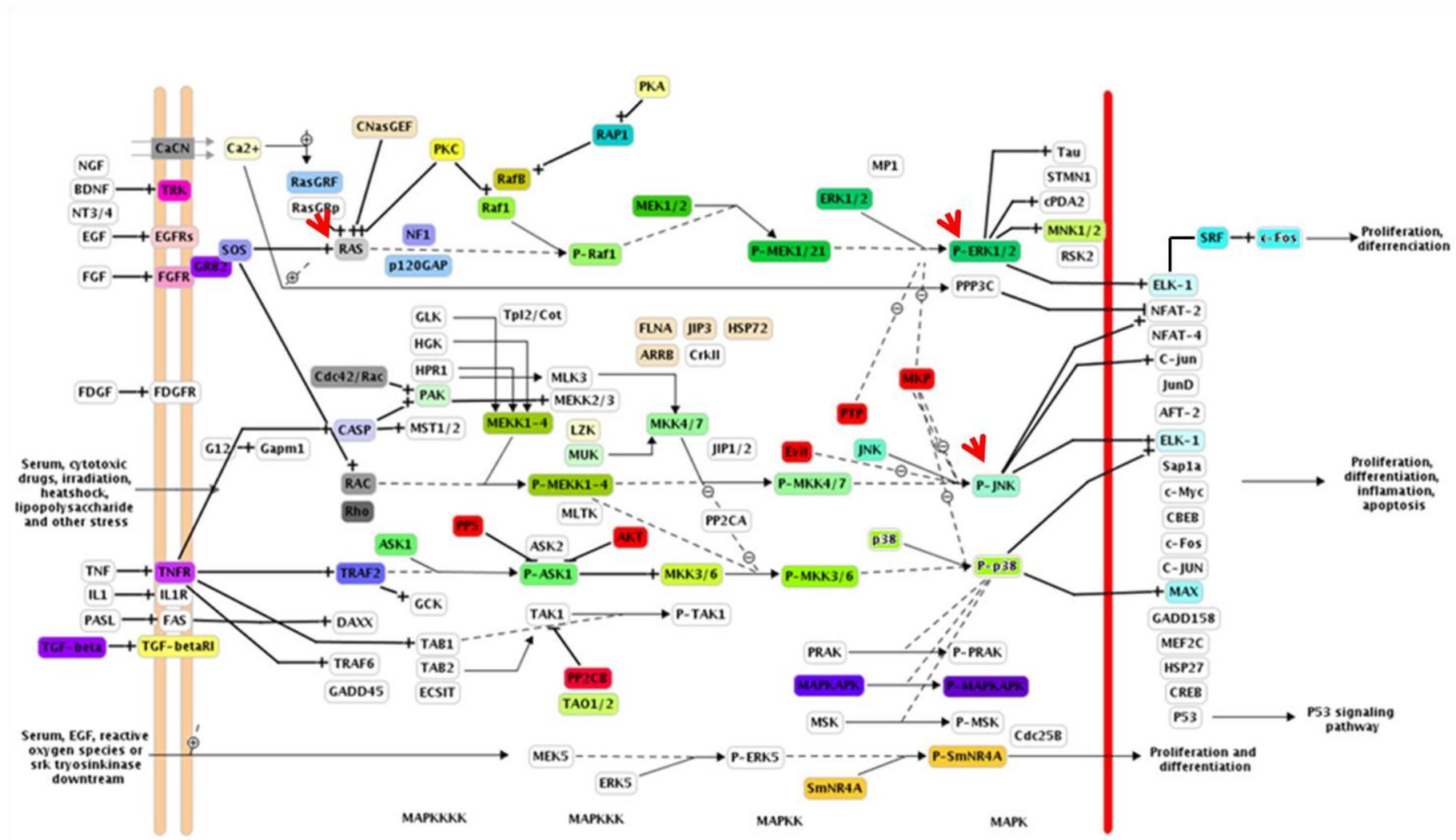
Proteínas da família GSK estão representadas por três ePKs em *S. mansoni* e *S. haematobium* e duas em *S. japonicum*. Uma GSK em especial (Smp\_008260.1) foi selecionada como possível alvo para o desenvolvimento de drogas através de uma abordagem comparativa entre os proteomas de *S. mansoni* e outras diferentes espécies de organismos modelo (246). Não é conhecida a função dessa proteína em *Schistosoma*, mas em mamíferos GSKs estão envolvidas no desenvolvimento e proliferação celular, estão altamente expressas em carcinoma de cólon e, em *C. elegans*, regulam positivamente a via de sinalização Wnt durante o desenvolvimento embrionário e durante a transição ovócito/embrião (250).

A via de sinalização MAPK é o mais bem caracterizado sistema de sinalização. Em *S. mansoni* foram encontradas nove ePKs da família MAPK, comparado com 6 em *S. haematobium*, 8 em *S. japonicum*, sete em *D. melanogaster* e 14 em humanos e *C. elegans* (Quadro 2).

Como mostra a **Figura 12**, mamíferos possuem pelo menos cinco cascatas de sinalização descritas na via MAPK que incluem: a cascata das proteínas ERK (*extracellular signal-regulated kinase*), que regula o crescimento e diferenciação celular; a cascata das JNK/SAPK (*c-Jun N-terminal kinase/stress-activated protein kinase*); e a cascata das proteínas p38, que funcionam essencialmente na resposta ao estresse, tais como inflamação e apoptose (251). Em *D. melanogaster* e *C. elegans*, as vias MAPK estão envolvidas em processos celulares críticos (252;253). *S. cerevisiae* tem quatro vias de sinalização MAPK distintas que são mediadoras prováveis de respostas ao ferormônio, a fome nutricional e ao estresse celular ou osmótica (254). As vias de sinalização MAPK parecem estar bem conservadas em *Schistosoma* (**Figura Anexo 3**) já que encontramos representantes das subfamílias ERK, p38, JNK e NLK. Outras ePKs de diferentes grupos como STE, PKA, PKC, Raf os receptores tirosina quinase (RTKs), também participam da via de sinalização das

MAPK. RTKs estão ancorados a membrana e têm papel importante na transmissão do sinal do meio extracelular para o citoplasma (**Figura 12**) (248). Membros de todas essas famílias foram encontrados em *S. mansoni*, *S. haematobium* e *S. japonicum* (Quadro 2). ePKs da subfamília ERK5 não foram identificados em *Schistosoma*. No entanto, a via de sinalização dessa MAPK é encontrada especialmente em mamíferos (**Figura Anexo 3**) e não era esperado que homólogos fossem encontrados nesses parasitos.

Em geral, na via das MAPK, cada subfamília é ativada por estímulos diferentes que geram diferentes respostas biológicas (255-257). Em *S. mansoni* apenas uma ePK foi identificada nas subfamílias JNK (Smp\_172240) e p38 (Smp\_133020) e estas subfamílias por sua vez desempenham papéis importantes na função da célula humana (258) e no desenvolvimento de *C. elegans* incluindo fenótipo de letalidade embrionária e esterilidade em vermes silenciados (248). Em *S. mansoni*, p38 está envolvida na transformação de miracídio para esporocisto sendo que o papel dessa ePK em vermes adultos ainda não foi validado (259;260). Sendo assim, acreditamos que as proteínas MAPK podem ter um papel importante na sobrevivência de *Schistosoma* e representam bons alvos para abordagens experimentais, principalmente àquelas que são membros únicos de uma subfamília e cujo sinal deverá ser interrompido caso a proteína não seja ativada.



**Figura 12: Via de sinalização MAPK caracterizada em mamíferos e predita em *S. mansoni*.** A via de sinalização MAPK de mamíferos está representada na figura. Os blocos coloridos correspondem as proteínas identificadas no proteoma predito de *S. mansoni*. Os blocos na cor branca são as proteínas de mamíferos sem homólogos no parasito. O sinal (+) representa ativação da proteína e o sinal (-) inibição da proteína. Nessa via de sinalização, além das proteína MAPK (ERK, p38, JNK), também participam outras ePKs incluindo proteínas do grupo STE (MEK1 / 2, MKK) AGC (PKC e PKA), CaMK (MAPKAPK), TK (EGFR, FGFR), e TKL (RAF e TGFbeta-RI). As proteínas MAPK são expressas em todas as células eucarióticas e são ativadas por estímulos diversos que variam de citocinas, fatores de crescimento, neurotransmissores, hormônios, estresse celular e adesão celular. O conjunto base de ativação da via MAPK é um módulo com três componentes onde participam três ePKs que estabelecem uma via de ativação sequencial. Nela estão uma MAPK quinase quinase (MKKK, MEKK, STE11, ou STE13), MAPK quinase (MKK, MEK ou STE7), e MAPK. A MAPK ativada pode translocar para o núcleo e se ligar a fatores de transcrição como elk-1 e MAX. As setas vermelhas indicam as proteínas selecionadas para validação experimental nesse trabalho.

#### 5.3-4. Análise das ePKs do grupo STE encontradas em *Schistosoma*

Proteínas do grupo STE têm com principal função transmitir o sinal nas vias de sinalização MAPK (**Figura 12**). Em *S. mansoni*, *S. haematobium* e *S. japonicum*, o grupo STE inclui sete ePKs da família STE7 (MEK ou MAPKK), duas STE11 (MEKK ou MAPKKK), e 13 STE20 (MEKKK) (Quadro 2). O grande número de membros da família STE em *Schistosoma* pode ser traduzido em um enorme potencial para a especificidade e diversidade do sinal traduzido para as vias MAPK (**Figura 12**).

A SmSLK (Smp\_150260) é uma proteína da família Ste20, recentemente caracterizada em *S. mansoni*, que foi capaz de ativar proteínas MAPK/JNK em células HEK (*Human Embryonic Kidney*), bem como em oócitos de *Xenopus*. Além disso, por imunofluorescência observou-se que SmSLK está abundantemente distribuída no tegumento de verme adultos de *S. mansoni* (261). Estas descobertas são evidências de que os sinais detectados no ambiente podem ser reconhecidos por proteínas do parasito que levarão a ativação da cascata de sinalização das MAPK. Curiosamente, as proteínas que codificam moléculas ativadoras das vias MAPK, como hormônios EGF, FGF, FDGF e TNF, não foram encontradas no proteoma predito de *S. mansoni* (**Figura 12**). Já foi demonstrado, no entanto, que *S. mansoni* utiliza as proteínas do hospedeiro para o seu crescimento e desenvolvimento (262;263).

Em *C. elegans* um estudo selecionou um grupo de proteínas essenciais para a sobrevivência do organismo a partir de experimentos de RNA de interferência e *knockout* de genes alvo (96). Através das análises filogenéticas foi possível identificar as ePKs de *S. mansoni* que são homólogas às ePKs de *C. elegans* (**Figura Anexo 4**) e que possuem fenótipo de letalidade e esterilidade, ou seja, que são essenciais (Tabela 4). Interessantemente, a maioria das ePKs destacadas pertencem aos grupos CMGC e STE, sugerindo o envolvimento destas proteínas em vias de sinalização que culminam em processos celulares essenciais, entre elas a via das MAPK.

#### 5.3-5. Análise das ePKs do grupo CK1 encontradas em *Schistosoma*

Em *Schistosoma* os dois menores grupos são CK1 e RGC (Tabela 3). Em contraste, em *C. elegans* CK1 é o maior grupo e o grupo RGC está muito expandido. Entretanto, a

comparação com os outros eucariotos analisados nesse trabalho, mostra que essas expansões são uma característica única de *C. elegans* (Tabela 4).

O grupo CK1 possui três famílias ePKs principais: CK1, VRK (*Vaccinia Related Kinase*) e TTBK (*Tau Tubulin Kinase*) que formam três agrupamentos bem definidos na árvore filogenética do grupo (**Figura Anexo 5**). *S. mansoni*, *S. haematobium* e *S. japonicum* possuem representantes em cada uma dessas famílias (Quadro 2 e Tabela Anexo 2) que também incluem ePKs de *C. elegans*, *D. melanogaster*, *M. musculus*, *H. sapiens*, *S. cerevisiae* e *B. malayi*. Os nematoides, *C. elegans* e *B. malayi*, ainda possuem duas famílias que parecem ser específicas desse grupo taxonômico, TTBKL e Worm6 (**Figura Anexo 5**). Além disso, outra família denominada Worm8, foi identificada apenas em *Caenorhabditis* até o momento. A diversificação do grupo CK1 em *C. elegans* foi relacionada por seus autores como uma adaptação que permite a reparação melhorada do DNA em resposta à exposição excessiva à mutagênicos ambientais (264). Um gene que codifica CK1 (SPE-6) tem funções relacionada a espermatogênese, e pelo menos metade das proteínas deste grupo são seletivamente expressas no esperma de *C. elegans* como mostrado em análises de microarranjo (265;266). O papel destas proteínas em parasitos *Schistosoma* ainda não é conhecido.

#### 5.4- Proteínas tirosina quinase

##### 5.4-1. Análise das ePKs do grupo TK encontradas em *Schistosoma*

PTKs podem ser classificadas, com base na presença ou ausência de domínios transmembrana, em tirosina quinase receptoras (RTKs) e tirosina quinase citoplasmáticas (CTKs). RTKs são capazes de transmitir sinais do meio extracelular ao intracelular (267). Dos RTKs encontrados no ePKinome de *S. mansoni* identificamos representantes nas subfamílias InsRs (*Insulin Receptors*), EGFRs (*Epidermal Growth Factor Receptor*), VKRs (*Venus Flytrap Kinase Receptors*), Ephs (*Ephrin receptors*), Ror, CCK4 (*Colon Carcinoma Kinase 4*) e MUSK (*Muscle\_Specific kinase*). Para a família Ror não encontramos ortólogos em *S. haematobium* e *S. japonicum*. Para todas as demais, pelo menos um provável ortólogo foi identificado.

Dois InsRs em *S. mansoni*, SmIR-1 (Smp\_009990) e SmIR-2 (Smp\_074030), foram experimentalmente validados e apresentaram funções distintas durante o desenvolvimento do parasita (268). Em nossas análises filogenéticas, SmIR-1 e SmIR-2 estão agrupados junto aos

outros membros da família InsR de outros organismos. No entanto, verificamos que os InsRs de *S. mansoni* são mais divergentes do que as proteínas de mamíferos e *D. melanogaster* (Figura 6 - Anexo). Um ortólogo para cada um dos InRs de *S. mansoni* foi encontrado em *S. haematobium* e *S. japonicum*. Em *S. japonicum* os receptores de insulina foram caracterizados (SjIR-1 e SjIR-2) e apresentaram alta identidade com SmIR-1 e SmIR-2, respectivamente (269;270).

SmIR-1 foi localizada nos músculos, epitélio intestinal, e membrana basal de adultos machos e fêmeas e na periferia dos esquistossômulos, principalmente no tegumento (271). Além disso, SmIR-1 foi co-localizada, no tegumento de esquistossômulos, com transportadores de glicose, sugerindo um papel na regulação da captação de glicose, um nutriente essencial para a fase adulta de *S. mansoni*. SmIR-2, em contraste, está distribuído no parênquima de machos e fêmeas adultos, indicando um possível envolvimento do receptor no crescimento do parasito. Em *S. japonicum*, SjIR-1 foi localizada na membrana basal do tegumento e no epitélio interno de vermes adultos e, assim como SmIR-2, SjIR-2 foi localizada no parênquima de machos e também no tecido vitelínico das fêmeas (272;273). *S. mansoni* e *S. japonicum* são os primeiros invertebrados com dois receptores de insulina caracterizados e que parecem ter funções distintas, como em vertebrados (274-276). Mamíferos têm dois membros InsR: IGFR (*Insulin-like growth factor receptor*) que têm um papel no controle do crescimento, e InR, que está relacionado a regulação metabólica (277) e, ao que parece, os vermes adultos de *Schistosoma* são capazes de utilizar a insulina produzida pelo hospedeiro mamífero para seu próprio desenvolvimento e crescimento (278).

Em *C. elegans* a via de sinalização dos receptores EGFR induzem à latência comportamental (175). Um homólogo ao receptor EGFR em *S. mansoni* (Smp\_093930.2) foi localizado no músculo do parasito e talvez esteja relacionada com a função ou com o desenvolvimento muscular (279). Curiosamente, esse receptor EGFR de *S. mansoni* foi localizada por Lepesant e colaboradores (2012) (207) em um *Scaffold* encontrado especificamente no cromossomo sexual Z (Tabela Anexo 1). Já foi visto que o hormônio EGF de vertebrados ativa EGFR de *S. mansoni* e leva a ativação da via de sinalização das MAPK-ERK (**Figura 12**), indicando a conservação da função do EGFR em *S. mansoni* (280). Além disso, foi demonstrado que o hormônio EGF humano aumenta a síntese de proteínas e DNA bem como a fosforilação da proteína em vermes adultos de *S. mansoni*, suportando a hipótese de que o EGF do hospedeiro poderia regular o desenvolvimento desse parasito (281). A

semelhança das proteínas de *S. mansoni* aos receptores hormonais sexuais dos mamíferos fornece um bom exemplo de relação hospedeiro-parasito, em que o verme adulto depende da síntese da hormonal do hospedeiro para a sua maturação e reprodução (2).

Tabela 4: Relação de ortologia entre ePKs de *S. mansoni*, *S. haematobium*, *S. japonicum* e *C. elegans* e o fenótipo encontrado após silenciamento gênico em *C. elegans* (96).

Group	<i>S. mansoni</i>	<i>S. haematobium</i>	<i>S. japonicum</i>	<i>C. elegans</i>	<i>C. elegans RNAi phenotype</i>
<b>CaMK</b>	Smp_053560	Sha_104673	Sjp_0052020	C44C8.6_Ce	Embryonic Lethal
<b>AGC</b>	Smp_123640	Sha_107453	Sjp_0030520	F19C6.1_Ce	Embryonic lethal, maternal sterile, organism morphology variant.
	Smp_136750	Sha_100632	Sjp_0048140	Y47D3A.16_Ce	Larval lethal, fertility reduced
	Smp_158560	Sha_105127	Sjp_0089570	K08B12.5_Ce	Body morphology defect, slow growth
	Smp_096310	Sha_101385	Sjp_0054180	pkc-3_Ce_A	Embryonic lethal, sterile, body morphology defect
<b>TK</b>	Smp_139480	Sha_103222	Sjp_0106390	vab-1_Ce_T	Embryonic lethal, body morphology defect
<b>STE</b>	Smp_026510	Sha_100164	Sjp_0093560	mek-2_Ce	Embryonic lethal, sterile, reduced brood size, exploted through vulva, slow growth
	Smp_151670	Sha_105501	Sjp_0095630		
	Smp_096640	Sha_106372	Sjp_0074710	gck_1_Ce	Cortical dynamics defective early embryonic, maternal sterile
	Smp_163420	Sha_100496	Sjp_0083800	mig_15_Ce	Abnormal cell migration, protruding vulva, developmental delay
	Smp_068060	Sha_103298	Sjp_0069910	kin_18_Ce	Embryonic lethal, body morphology defect, slow growth, sterile, larval lethal, dumpy
	Smp_131800	Sha_107737	Sjp_0051030	Y59A8B.23_CE	More depolarized oocytes
	Smp_150260	Sha_103379	Sjp_0065920	C04A11.3_Ce	Slow growth, larval lethal late L3/L4
<b>TKL</b>	Smp_079760	Sha_103781	Sjp_0072270 Sjp_0132890	pat4_Ce_TK	Embryonic lethal, body morphology defect, larval lethal, sterile, paralyzed, uncoordinated movement, paralyzed
	Smp_176990	Sha_101467	Sjp_0100500	lin45_Ce_T	Sterile, osmotic integrity problems
<b>CMGC</b>	Smp_068960	Sha_105780	Sjp_0066800	F22D6.5_Ce	Embryonic lethal, slow growth, sterile progeny, uncoordinated movement
	Smp_041770	Sha_107226	Sjp_0000700	spk_1_Ce	Embryonic lethal, larval lethal, maternal sterile
	Smp_172700	Sha_104207	Sjp_0031950	cdk_4_Ce	Locomotion variant, larval lethal, sterile progeny
	Smp_133020	-	Sjp_0108760	pmk-1_Ce, pmk-2_Ce	Apoptosis increased, sterile progeny, reduced brood size, embryonic lethal
	Smp_134260	Sha_101970	Sjp_0049420		
	Smp_133490	Sha_106849	Sjp_0001270	C05D10.2	Reduced brood size, embryonic lethal
	Smp_133500		Sjp_0001260		
	Smp_080730	-	-	cdk_1_Ce	embryonic lethal, sterile progeny
	Smp_155720	Sha_103078	Sjp_0083510	Y18D10A.5_Ce	Larval lethal, slow growth, embryonic lethal
	Smp_125310	Sha_102724	Sjp_0026020	C44H4.6_Ce	
	Smp_008260	Sha_104108			
	Smp_156990	Sha_200050	Sjp_0034250	F39H11.3_Ce	Sterile
	Smp_176620	Sha_100367	Sjp_0082800	B0495.2_Ce	Embryonic lethal
	Smp_003000	Sha_105528	Sjp_0098610	H25P06.2_Ce	Embryonic lethal, slow growth, uncoordinated movement
	Smp_150040	Sha_106204	-	Cdk_7_Ce	Embryonic lethal
	Smp_140700	Sha_101439	Sjp_0066980	lit_1_Ce	Embryonic lethal, larval lethal
	Smp_074080	Sha_104362	Sjp_0035070		
	Smp_141230	Sha_102259	Sjp_0051010	mbk_2_Ce	Embryonic lethal, sterile progeny

Cinco ePK de *S. mansoni* não se agruparam com nenhuma família RTK como mostra a análise filogenética (**Figura Anexo 6**). Duas delas (Smp\_175590 e Smp\_157300) possuem domínio catalítico truncado e outra (Smp\_173590) não foi predita na versão atual do genoma. As outras duas são RTK específicas com o domínio *venus flytrap* (família VKR). VKR é uma família de receptores encontrados em invertebrados, especialmente em insetos. Uma VKR de *S. mansoni*, Smp\_153500 (SmVKR1), foi experimentalmente caracterizada e os autores indicam que a provável função desse gene esteja relacionada ao desenvolvimento e reprodução (282). SmVKR1 é composta por um domínio extracelular *Venus flytrap module* (VFT) ligado por um domínio transmembrana ao domínio catalítico (domínio tirosina quinase) intracelular que é similar ao dos receptores de insulina (IRs) (282). Um ortólogo a essa ePK foi encontrado em *S. haematobium* (Sha\_104501) e outro em *S. japonicum* (Sjp\_0109220). Nós identificamos outra ePK de *S. mansoni* (Smp\_019790) agrupada junto com SmVRK (**Figura Anexo 6**) e com alta similaridade entre suas sequências. Recentemente essa ePK foi clonada (Smp\_019790 – SmVKR2) e sua atividade quinase foi comprovada. Os autores indicam que SmVKR1 e SmVKR2 representam parálogos com diferentes funções no parasito (283). Em *S. haematobium* e *S. japonicum* foram identificados dois prováveis ortólogos a Smp\_019790. Apesar da semelhança do domínio catalítico das proteínas VKR com os InsRs, as duas proteínas VKRs de *S. mansoni* não estão agrupadas com a família InsR (**Figura Anexo 6**). Curiosamente proteínas da família VKR não são encontradas em mamíferos e podem assim representar bons alvos para o desenvolvimento de drogas já que um inibidor específico para esta família, provavelmente, não afetará qualquer proteína do hospedeiro (282).

As CTKs em *S. mansoni* são representadas por 11 famílias diferentes (Tabela Anexo 1). Dentre as seis ePKs localizadas em regiões específicas do cromossomo sexual Z, duas (Smp\_137610 e Smp\_011700) são CTKs. As duas são representantes únicas da sua família (Fak e Sev, respectivamente) sendo que nenhum ortólogo a ePK da família Fak foi identificado em *S. japonicum* e *S. haematobium* (Tabela Anexo 2).

SmTK3 (Smp\_054500) e SmTK5 (Smp\_136300) - membros da família src, e SmTK4 (Smp\_149460) - família syk, são CTKs presentes nos órgãos reprodutores e possivelmente envolvidas no desenvolvimento de gônadas e multiplicação de células germinais e vitelínicas (282;284-286). Proteínas da família Abl de *S. mansoni* (Smp\_128790 e Smp\_169230) foram estudadas recentemente usando um inibidor comercial específico para essa família (*Imatinib*,

Gleevec<sup>®</sup>). Os resultados mostraram uma importante alteração morfológica em vermes adultos de *S. mansoni* que levou à morte dos parasitos (287). Em vermes tratados com praziquantel, essas duas ePKs da família Abl estão super expressas em relação ao controle não tratado (Tabela Anexo 1).

*C. elegans* contém 42 membros da família Fer, enquanto apenas um único membro, SmFes, foi encontrado em *S. mansoni*. A ePK Fer de *S. mansoni* (SmFes, Smp\_164810) apresenta os traços característicos da subfamília Fes/Fps/Fer (*fes*, *feline sarcoma*; *fps*, *Fujinami poultry sarcoma*; *fer*, *fes related*). Por ensaios de imunolocalização foi mostrado que SmFes é particularmente expressa no terebratorium de miracídios e tegumento de cercária e esquistossômulos. Estas descobertas sugerem que SmFes pode desempenhar um papel nas vias de transdução de sinal envolvidas na transformação dos parasitos após a penetração nos hospedeiros intermediário e definitivo (288;289). Por ser o único membro da sua família encontrada em *S. mansoni* e pelas evidências experimentais sugerirem um importante papel durante o ciclo de vida do parasito, SmFes, é possivelmente um bom alvo para o desenvolvimento de droga.

#### 5.4-2. Análise das ePKs do grupo RGC encontradas em *Schistosoma*

Proteínas do grupo RGC compartilham sequência similar no domínio catalítico (Pfam: PF07714) às do grupo TK (229) e são filogeneticamente próximas a esse grupo (**Figura 13**). O grupo RGC é sub-representado na maioria das espécies, exceto em *C. elegans* que tem uma grande expansão destas proteínas e, *S. cerevisiae*, que não tem nenhuma proteína com domínio catalítico semelhante às ePKs do grupo TK (Quadro 2). Outros parasitos como *L. major*, *T. brucei*, *T. cruzi* e *P. falciparum* também não possuem representantes no grupo RGC (192;290).

Apenas três membros RGC foram identificados no ePKinome de *S. mansoni*, *S. haematobium* e *S. japonicum*. As RGC de *S. mansoni* estão mais próximas filogeneticamente às RGC de mamíferos e insetos do que as RGC de *B. malayi* e *C. elegans* como sugere a árvore filogenética (**Figura Anexo 7**) em que as ePKs de *C. elegans* e *B. malayi* formam pelo menos duas famílias visivelmente divergentes das RGC de *S. mansoni*, *D. melanogaster*, *M. musculus* e *H. sapiens*.

A maioria das ePKs desse grupo ainda não foram caracterizadas e com isso a verdadeira função desses genes ainda é um mistério. Em *C. elegans*, várias ePKs RGC estão altamente expressas em conjuntos restritos de neurônios envolvidos nos sentidos químicos e, uma está especialmente relacionada ao desenvolvimento de uma fase específica do verme (*Dauer stage formation*) (291).

As três proteínas RGC de *Schistosoma* têm uma substituição no aminoácido ácido aspártico (D) no subdomínio VIb do domínio catalítico, que as torna cataliticamente inativas (**Figura 11**). Embora o domínio catalítico de uma enzima seja geralmente conservado, alguns relatos mostram a existência de proteínas, como as ePKs do grupo de RGC, com substituições nas posições essenciais e que convertem a enzima numa forma cataliticamente inativa. Um estudo recente mostrou que as enzimas inativas são encontradas numa ampla variedade de famílias conservadas entre espécies de metazoários e que apesar de terem perdido a sua atividade catalítica, adotaram novas funções, e estão envolvidos em processos de regulação (292;293).

## 5.5- Proteínas quinase híbridas

### 5.5-1. Análise das ePKs do grupo TKL encontradas em *Schistosoma*

TKL consiste de um grupo que é filogeneticamente próximo ao grupo TK (**Figura 13**). No entanto, as proteínas TKL têm um domínio catalítico incomum que é um híbrido entre os domínios serina/treonina e tirosina-quinase (294). Sendo assim, no grupo TKL algumas ePKs podem exibir um domínio catalítico semelhante ao domínio catalítico das tirosina quinase (Pfam: PF07714) ou ao domínio a serina/treonina quinase (Pfam: PF00069) (143;295). Em alguns casos, para as ePKs de *S. mansoni*, encontramos os dois domínios em uma mesma sequência (Smp\_047870.1, Smp\_124450 e Smp\_144390) (Tabela Anexo 1).



Em *Schistosoma*, o grupo TKL inclui representantes nas famílias MLK (*Mixed Lineage Kinases*), LISK (*Family containing closely related LIMK and TESK sub-families*), Raf, RIPK (*Receptor Interacting Protein Kinase*), STKR (*Serine/threonine kinase Receptors for activin and TGF $\beta$  ligands*) e LRRK (*Leucine Rich Repeat Kinase*), sendo esta última família não encontrada em *S. haematobium* e *S. japonicum* (Quadro 2 e Tabela Anexo 2). *S. mansoni* não tem proteínas homólogas à família IRAK (*interleukin-1 (IL-1) receptor-associated kinase*) que, por sua vez, está presente em outros vermes como *C. elegans* e *B. malayi* além de outras espécies como *D. melanogaster*, *Homo sapiens*, e *M. musculus* (**Figura Anexo 8**). Embora *S. cerevisiae* não possua qualquer homólogo às proteínas TKL, outras espécies de fungos contêm (296).

Raf (também conhecido como MAPKKK) é uma ePK da família TKL que desempenha um papel importante na ativação de proteínas STE na cascata de sinalização que culmina na ativação da MAPK ERK1/2 (**Figura 12**) (297). Um estudo recente mostrou que o bloqueio da expressão do homólogo da proteína Raf de *S. mansoni* (Smp\_176990) em *C. elegans* por RNAi, gera um fenótipo de esterilidade, que suporta a hipótese do envolvimento de proteína Raf no desenvolvimento da linhagem germinativa, ovogênese, espermatogênese, ovulação ou fertilização (Tabela 4). Proteínas Raf como Smp\_176990 são únicos membros da sua subfamília, possuem um ortólogo em *S. haematobium* e *S. japonicum* e podem com isso representar bons alvos para o desenvolvimento de drogas para combate da esquistossomose.

Os membros da família STKR que se ligam a TGF $\beta$  (*Transforming growth factor*) são receptores de membrana que podem ser divididos em duas subfamílias (Tipo I e Tipo II). O receptor tipo II liga TGF $\beta$  e, em seguida, recruta o receptor tipo I. O receptor TGF $\beta$  tipo I foi clonado em *S. mansoni* (SmT $\beta$ RI - Smp\_173400.2) e verificou-se a sua localização na superfície do parasito. (298;299). Apesar de não ter sido relatado pelos autores, outro receptor STKR tipo I (Smp\_049760) também foi identificado em *S. mansoni* (Tabela Anexo 1 e **Figura Anexo 8**). Além disso, foi observado dois receptores STKR tipo II (receptores TGF $\beta$  do tipo II) em *S. mansoni*, sendo um deles produto de *splicing* alternativo. Forrester e colaboradores (2004) revelaram a presença de dois transcritos que são traduzidos em duas isoformas diferentes de receptor tipo II (300). Estes transcritos são produzidos por *splicing* alternativo dos últimos dois exons. Os mesmos autores indicaram que os receptores do tipo II podem fazer parte de vias de sinalização que atuam em diferentes células ou estágios de desenvolvimento. Além disso, este estudo demonstrou que, na presença de TGF $\beta$  humano,

SmTBR2 (Smp\_165310) ativa SmTBR1. Os resultados também fornecem evidência que a via de sinalização de TGF- $\beta$  do macho está induzindo o desenvolvimento reprodutor da fêmea (301;302). Provavelmente essas vias de sinalização também estão conservadas em *S. haematobium* e *S. japonicum* já que identificamos ortólogos para os dois tipos de receptores citados acima.

### 5.5-2. Análise das ePKs do grupo *Other* encontradas em *Schistosoma*

Nesse trabalho a classificação das ePKs de *Schistosoma* seguiu a classificação estabelecida e descrita no banco de dados KinBase. Sendo assim, verificamos que 15% das ePKs de *S. mansoni* não se encaixam em nenhum dos 8 grupos já detalhados (AGC, CaMK, CK1, CMGC, STE, TK, TKL e RGC), mas formam 20 pequenas e conservadas famílias pertencentes ao grupo *Other*. Todas essas 20 famílias possuem ortólogos em *S. haematobium* e *S. japonicum* (Tabela Anexo 2).

O grupo *Other* é formado por uma diversidade de ePKs com representantes em todos os eucariotos. Várias famílias já foram identificadas dentro do grupo *Other* como: SCY1, NEK (*Mitotic Kinase family*, também conhecida como *NRK*), PEK, Haspin, WEE, NAK (*Numb-Associated Kinase*), ULK (*Unc-51 Like Kinase*), IRE (*Inositol Requiring*), PLK (*Polo Like Kinases*), AUR (*Aurora Kinase*), e CDC7 (*Cell Division Control 7*). Essas famílias não têm qualquer semelhança entre si e poderiam formar novos grupos ePKs. Para isso, um trabalho de caracterização detalhado precisa ser feito. Tentamos construir uma árvore filogenética para classificar as ePKs do grupo *Other*, mas devido a alta divergência entre as sequências das diferentes famílias essa árvore não pode ser criada.

Recentemente, Dissou e colaboradores (2011) (303) clonaram e caracterizaram duas ePKs do grupo *Other* e família PLK (*Polo-like kinases*) em *S. mansoni* e, através de análises estruturais e funcionais, demonstraram possível conservação da sua função em relação às proteínas de mamíferos que estão envolvidas na regulação do ciclo celular durante a transição das fases G2/M. Em mamíferos PLK regula positivamente a mitose celular e está super expressa em vários tipos de câncer. Um inibidor comercial para as proteínas dessa família foi utilizado em *S. mansoni* e os resultados indicam que PLK está envolvida especificamente no processo de gametogênese do parasito (304).

## 5.6- Domínios acessórios encontrados nas ePKs de *Schistosoma*

A estrutura do domínio catalítico das ePKs é altamente conservada entre diferentes organismos pelo fato de todos eles reconhecerem e se ligarem ao ATP em um sítio comum. Entretanto, o domínio catalítico é suficientemente divergente para, por si só, discriminar grupos, famílias e subfamílias (**Figura 13**). A maioria das ePKs possuem também um domínio envolvido na interação proteína-proteína que regula alostericamente a atividade do domínio catalítico (305). Nesse trabalho, apenas a sequência correspondente ao domínio catalítico das ePKs de *S. mansoni* foi utilizada nas análises filogenética. Curiosamente, quando os domínios acessórios foram integrados a anotação, observamos uma correlação entre a diversidade da arquitetura proteica e os pares filogenéticos.

Já é sabido que diferentes arquiteturas de domínios proteicos refletem diferenças funcionais nas proteínas (306). Assim, compreender os mecanismos que geram tais diversidades é essencial para a compreensão da função biológica de cada gene. Além disso, arquiteturas de domínios proteicos incomuns podem gerar especificidade no reconhecimento de substratos e reguladores e são geralmente os alvos de ligação para os inibidores utilizados no tratamento de diferentes doenças (307).

As análises dos domínios acessórios, identificados pelo banco de dados PFAM (308), mostraram que aproximadamente 36% das ePKs de *S. mansoni* são proteínas que possuem vários domínios regulatórios e de sinalização juntos ao domínio quinase catalítico (**Figura 14** e **Figura 15**). Em *S. haematobium* essa proporção cai para 28% e em *S. japonicum* para 19%. Mais uma vez, os domínios acessórios que muitas vezes são característicos e essenciais em certas famílias ePKs, não foram identificados nessas duas últimas espécies por erros na predição do gene provavelmente ocasionados pela alta fragmentação do genoma.

Os domínios acessórios mais comuns encontrados pelo Pfam nas ePKs de *Schistosoma* foram Pkinase\_C (Pfam: PF00433) todos eles no grupo AGC; C1\_1 (Pfam: PF00130) encontrado nos grupos AGC e TKL; SH2 (PF00017) todos encontrados no grupo TK; e SH3 (Pfam: PF00018) encontrado em TKs e TKLs (Tabela Anexo 2). Estes domínios são também comumente encontrados em famílias ePKs em outras espécies como observamos no KinBase (<http://www.kinase.com/kinbase>) (309).

O domínio Pkinase\_C (C1\_1) é comumente encontrado em ePKs da família PKC e está conservado em todas as ePKs dessa família em *S. mansoni*. O domínio C1\_1 se liga a ésteres de forbol (PE) e DAG (diacilglicerol). DAG é um importante mensageiro secundário e ésteres de forbol são análogos a DAG [34,97]. O domínio C1\_1 está presente em uma ou duas cópias, dependendo da isoenzima do qual a PKC pertence (**Figura 14**).

cNMP (PF00027) é um domínio N-terminal das proteínas PKG que se liga a proteínas G receptoras acopladas para aliviar a inibição do domínio catalítico (217).

A longa ePK MLCK (**Figura 14**) possui um grande número de repetições de Ig (I-set, V-set e Ig) que, em outras espécies, estão envolvidos em uma variedade de funções, incluindo o reconhecimento célula-célula, receptores na superfície celular, estrutura muscular e com o sistema imune (310) e, fn3 é um domínio de aproximadamente 100 aminoácidos comumente encontrados em vários organismos.

Os grupos CMGC e CK1 têm nenhum ou alguns poucos domínios acessórios em *Schistosoma* (Tabela Anexo 2) e também em grande parte dos eucariotos. No entanto, sabe-se que pequenas regiões nestas ePKs desempenham um papel importante no reconhecimento e ligação ao substrato (311;312). Por exemplo, o domínio de CD (domínio de ancoragem comum) é uma região C-terminal das proteínas MAPK compostos por um conjunto de aminoácidos carregados negativamente que é usado para ancorar ativadores de proteína (tal como proteínas do grupo STE), substratos (tais como MAPKAPK) e proteínas inativadoras (tais como as MAPK fosfatases)(313). Assim, esta região regula uma série de cascatas de transdução de sinal das reações das MAPKs. Outras regiões, incluindo o sítio ED (*ERK docking*), trabalha com o domínio de CD e garante a especificidade e a força de interação (314) (**Figura 14**).

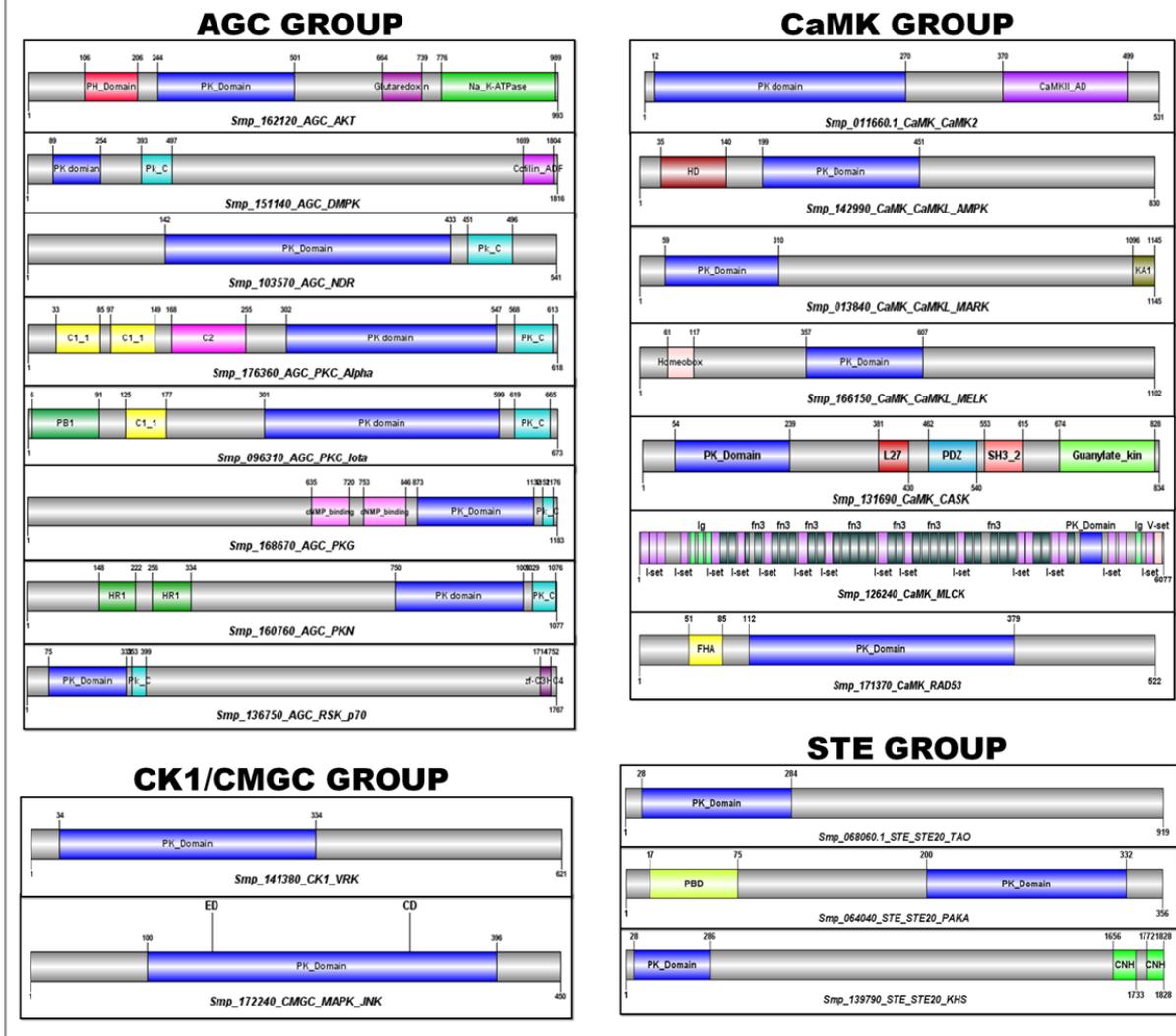
O domínio PDB (PF00786 - *p21-Rho-binding*) e o domínio C-terminal CNH (PF00780) são normalmente encontrados nas ePKs da família STE20 (**Figura 14**). PDB liga-se a Cdc42 e GTPases ativando a cascata de sinalização das MAPK. O domínio CNH interage com o pequenas GTPases para atuar na regulação do citoesqueleto de actina (315).

Os domínios SH3 e SH2 (Pfam: PF00018 e PF00017, respectivamente) são comumente encontrados nas proteínas CTKs. O SH2 funciona como módulo regulador da

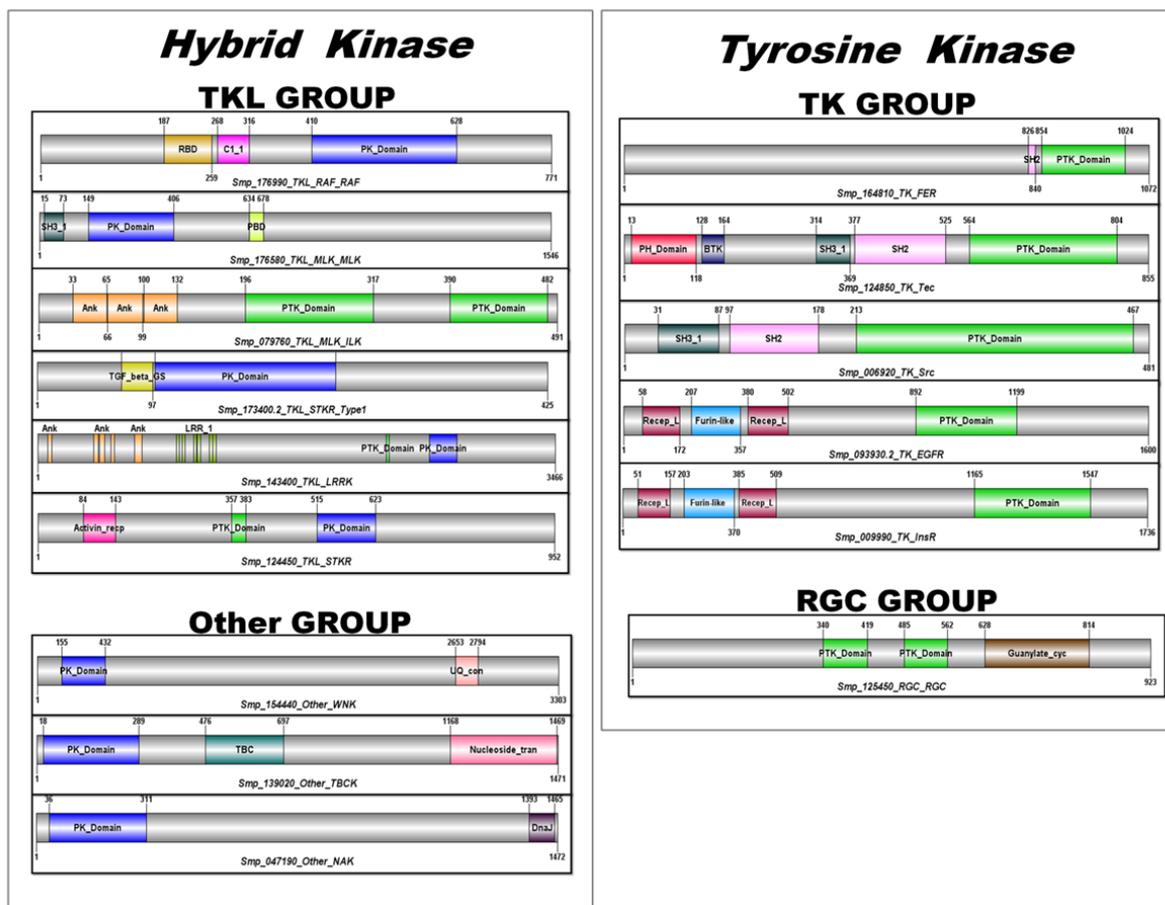
casca de sinalização intracelular e foi encontrado na maioria das CTK das famílias Src, SYK e Tec em *Schistosoma*. As CTK da família Fer são geralmente compostas por três domínios: domínio FCH (PF00611), SH2 e o domínio quinase C-terminal e essa arquitetura é encontrada em proteínas Fer de *H. sapiens*, *M. musculus*, e *D. melanogaster*. No entanto, a ePK da família Fer de *Schistosoma* (SmFes - Smp\_164810 e Sha\_103100) (316) e as 42 proteínas Fer de *C. elegans* parecem ter perdido o domínio FCH N-terminal (**Figura 15**).

RTKs são caracterizados por um domínio extracelular e um domínio quinase intracelular (317). A região de ligação extracelular da proteínas EGFR e das proteínas InsR são compostas por dois domínios receptor\_L entre um domínio *Furin\_like* (**Figura 15**). Outros domínios extracelulares encontrados em *S. mansoni* são Ephrin\_Ibd (Pfam: PF01404) nos receptores Efrina (Eph) e domínios Ig (Pfam: PF00047) em proteínas CCK4 (Tabela Anexo 1).

## Serine/Threonine Kinase



**Figura 14: Representação da organização dos domínios proteicos de algumas ePKs dos grupos AGC, CaMK, CK1, CMGC e STE de *S. mansoni*.** Os identificadores e a classificação de cada ePK estão mostrados abaixo de cada estrutura. As abreviações dos domínios são: PK\_Domain (Domínio *Protein kinase*), PH\_Domain (Domínio *Pleckstrin Homology*), Na\_K-ATPase (*Sodium/potassium ATPase beta chain*), PK\_C- (Domínio *Protein Kinase C terminal*), Zf-C3HC4 (*zinc-finger, C3HC4 type RING finger*), HR1 (*Hr1 repeat*), PB1 (*Phox and Bem1p domain*), C1\_1 (Domínio *Phorbol esters/diacylglycerol binding*), C2 (Domínio  $Ca^{2+}$ -dependent), CNH (Domínio *citron homology*), cNMP\_binding (Domínio *cyclic nucleotide-binding*), Cofilin\_ADF (*Cofilin/tropomyosin-type actin-binding protein*), CaMKII\_AD (*Calcium/calmodulin dependent protein kinase II Association*), HD (*HD homeobox domain*), L27 (Domínio L27), PDZ (Domínio PDZ, também conhecido por *DHR* ou *GLGF*), SH3\_2 (Domínio *Variant SH3*), Guanylate\_kin (*Guanylate kinase*), Ig (Domínio *Immunoglobulin*), fn3 (Domínio *Fibronectin type III*), V-set (Domínio *Immunoglobulin V-set*), FHA (Domínio *Forkhead-associated*) e PBD (Domínio *P21-Rho-binding*). A arquitetura dos domínios foi gerada por DOG 1.0 (318) baseado nos limites dos domínios do Pfam (319).



**Figura 15: Representação da organização dos domínios proteicos de algumas ePKs dos grupos TKL, Other, TK e RGC de *S. mansoni*.** Os identificadores e a classificação de cada ePK estão mostrados abaixo de cada estrutura. As abreviações são: **PK\_domain** (Domínio *Protein kinase*), **RBD** (Domínio *Raf-like Ras-binding*), **C1\_1** (Domínio *Phorbol esters/diacylglycerol binding*), **SH3\_1** (*Src homology 3*), **PBD** (Domínio *P21-Rho-binding*), **Ank** (*Ankyrin repeat*), **PTK\_domain** (Domínio *Protein tyrosine kinase*), **TGF\_beta\_GS** (*Transforming growth factor beta type I GS-motif*), **LRR\_1** (*Leucine Rich Repeat*), **Activin\_recep** (Domínio *Activin types I and II receptor*), **UQ\_con** (*Ubiquitin-conjugating enzyme*), **TBC** (Domínio *TBC*), **Nucleoside\_tran** (*Nucleoside transporter*), **DnaJ** (*DnaJ domains*), **SH2** (Domínio *Src homology 2*) **PH\_domain** (Domínio *Pleckstrin Homology*), **BTK** (*Bruton's tyrosine kinase motif*), **Recep\_L\_domain** (Domínio *Receptor L*), **furin-like** (*Furin-like cysteine rich region*) e **Guanylate\_cyc** (*Receptor family ligand binding region*). A arquitetura dos domínios foi gerada por DOG 1.0 (320) baseado nos limites dos domínios do Pfam (321).

## 5.7- Investigação do papel biológico de proteínas envolvidas na via de sinalização MAPK.

### 5.7-1. Proteínas selecionadas

A anotação funcional realizada para as ePKs das espécies de *Schistosoma* aqui estudadas permitiu o melhor entendimento do papel dessas proteínas. Baseado na importância funcional dos ortólogos, foram selecionadas algumas ePKs com função predita essencial para o desenvolvimento e sobrevivência desse organismo (Tabela 5). As ePKs selecionadas são, na maioria das vezes, representantes únicas de uma família/subfamília evitando assim a substituição de sua atividade nas vias das quais participam. Além disso, grande parte delas (Tabela 5) são ativadoras/efetoras da via de sinalização MAPK.

O mecanismo de RNA de interferência é um processo intrínseco à célula da maioria dos eucariotos, que entre outras funções, promove o silenciamento gênico através da degradação de mRNAs específicos. Esse mecanismo pode ser desencadeado *in vitro* pela presença de um RNA de dupla fita (dsRNA). Como ferramenta molecular para analisar a função gênica, o RNAi é particularmente importante em organismos como *Schistosoma* em que as técnicas clássicas de manipulação gênica não estão plenamente desenvolvidas. Através da análise do genoma já foi possível identificar todas as moléculas envolvidas na maquinaria de RNAi em *S. mansoni* (322). Sendo assim, essa metodologia foi escolhida para avaliar a função de algumas ePKs envolvidas na via de sinalização MAPK e ter uma ideia da real função e de como essa via é regulada nesse parasito.

As proteínas selecionadas para os estudos experimentais são: SmCaMK2 (Smp\_011660.2), SmJNK (Smp\_172240), SmRas (Smp\_179910), SmERK1 (Smp\_142050) e SmERK2 (Smp\_047900). Na via de sinalização das MAPK representadas na **Figura 12** essas proteínas estão destacadas por uma seta vermelha. SmCaMK2 não aparece nessa figura pois ela está presente apenas na via de sinalização MAPK/JNK do nematódeo *C. elegans* (170). SmCaMK2 é uma proteína quinase do grupo CaMK e da família CaMK2. As proteínas do grupo CaMK, em sua grande maioria, são ativadas pelo mensageiro secundário  $Ca^{2+}$  (323). Apesar da existência de quatro membros dessa família em humanos, apenas uma proteína foi encontrada em *C. elegans* (unc-43) e em *S. mansoni* (SmCaMK2). SmCaMK2 possui ainda dois produtos de *splicing* alternativos preditos. Recentemente, Caffrey e colaboradores (2009) (246), selecionaram SmCaMK2 como potencial alvo terapêutico devido ao fenótipo de

letalidade verificado em experimentos de RNAi nos ortólogos em *C. elegans* e *D. melanogaster*.

SmJNK é uma proteína quinase do grupo CMGC, da família MAPK e subfamília JNK. Curiosamente, três proteínas da família JNK são encontradas em humanos, cinco são encontradas em *C. elegans* e, apenas uma, em *S. mansoni*. Em *C. elegans*, a JNK (*jnk-1*) é ativada pela proteína CaMK2 (*unc-43*) em resposta ao estresse, regulando a apoptose celular (248;324).

As proteínas Ras não são ePKs, mas fazem parte de uma superfamília de pequenas moléculas ativadas por GTP (*small GTPases*) e estão envolvidas em processos de sinalização celular, entre elas, na ativação da via de sinalização MAPK. Os genes Ras foram primeiramente identificados como oncogenes por estarem mutados em diversos tipos de cânceres. Existem mais de 30 proteínas da superfamília Ras em humanos ativando diversas vias de sinalização (325). Em *C. elegans* já foram identificados 56 membros da superfamília Ras. No entanto, a única função conhecida para proteína Ras de nome *let-60*, é ativar a via de sinalização RTK/Ras/MAPK (248). A fim de identificar possíveis homólogos às proteínas Ras em *S. mansoni*, sequências de aminoácidos de proteínas Ras conhecidas e depositadas no banco de dados de proteínas do NCBI, foram selecionados. Entre as proteínas identificadas, aquela com maior similaridade à proteína *let-60* foi escolhida, a SmRas (*Smp\_179910*). SmRas já havia sido caracterizada em *S. mansoni* e é expressa em todos os estágio de desenvolvimento do parasito mas, com expressão aumentada em vermes fêmeas quando comparada ao macho (326).

As proteínas ERK são proteínas quinase do grupo CMGC e família MAPK. Existem 5 subfamílias para as proteínas MAPK/ERK (ERK, ERK1, ERK3, ERK5 e ERK7). No entanto, a subfamília MAPK envolvida na via de sinalização das proteínas RTK/Ras/MAPK, é a ERK1. Em *S. mansoni* identificamos dois ortólogos a essa família, SmERK1 e SmERK2. Em geral, proteínas ERK1, quando ativadas, se tornam capazes de migrar para o núcleo e ativar fatores de transcrição como *elk1* e SRF que, por sua vez, iniciam a transcrição de genes específicos como *c-Fos*, que estão envolvidos em processos de proliferação e diferenciação. SRF e *c-Fos* também foram identificadas no proteoma predito de *S. mansoni* através de buscas por similaridade de sequência utilizando HMMs. Esses genes são: SmSRF (*Smp\_097730*) e SmC-Fos (*Smp\_124600*).

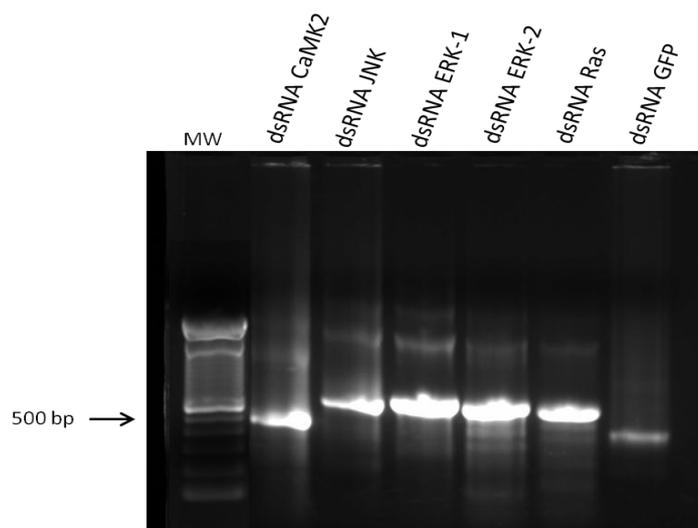
Tabela 5: ePKs de *S. mansoni* selecionadas como possíveis alvos para o desenvolvimento de droga baseado na importância funcional de seus genes e/ou dos ortólogos identificados nesse trabalho

ePK group and family	ePK ID	Important Characteristics
AGC_PKA	Smp_152330	Inhibition causes parasite death (88)
*CaMK_CAMk2	Smp_011660	Single family member and important phenotype in other organisms after inhibition (162)
*CaMK_MAPKAPK	Smp_053560	Single family member, lethal phenotype in <i>C. elegans</i> after inhibition (71).
CMGC_GSK	Smp_008260	Important phenotype in other organisms after inhibition. Probably role in embryonic development (162)
*CMGC_MAPK_JNK	Smp_172240	Single subfamily member. Probably role in parasite reproduction (164)
*CMGC_MAPK_p38	Smp_133020	Single subfamily member. Lethal phenotype in <i>C. elegans</i> after inhibition (164)
TK_VKR	Smp_153500	Not found in mammals. Probably role in the parasite development and reproduction (186)
TK_VKR	Smp_019790	Not found in mammals. Probably role in the parasite development and reproduction (178)
TK_Fer	Smp_164810	Single family member. Probably role in the parasite transformation (191)
*TKL_Raf_Raf	Smp_176990	Single subfamily member. Lethal phenotype in <i>C. elegans</i> after inhibition (71)

\*ePKs envolvidas na via de sinalização MAPK

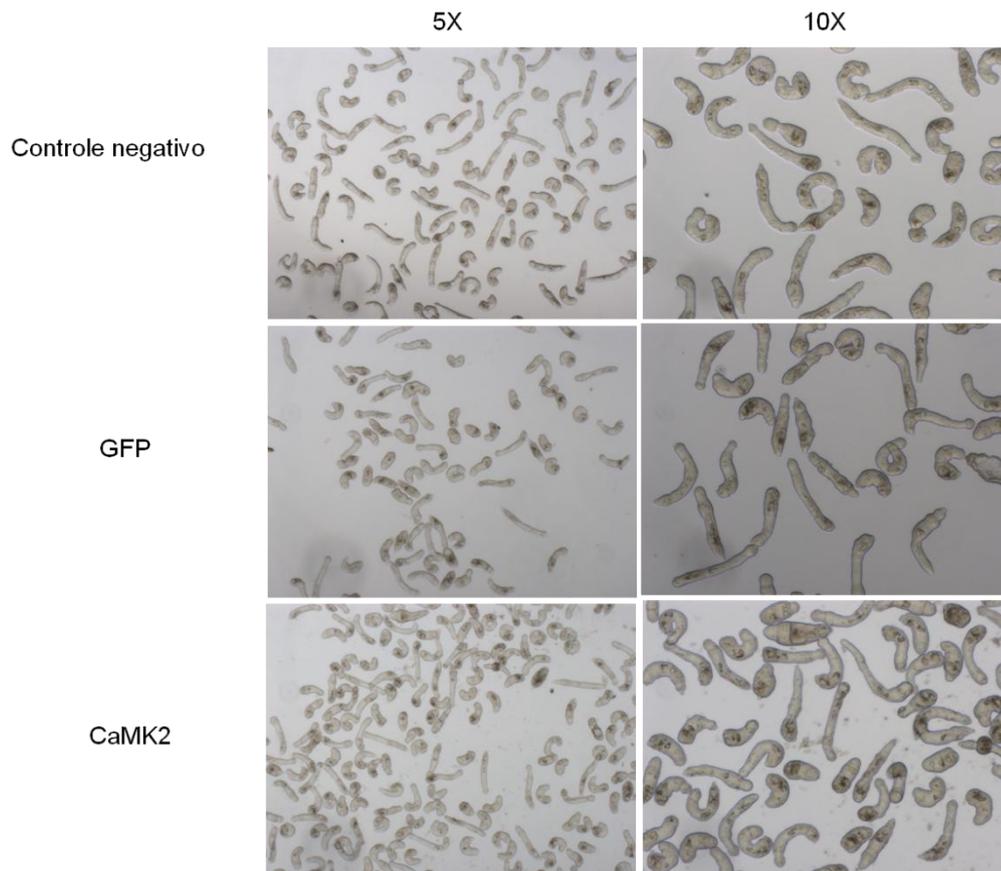
### 5.7-2. Síntese de dsRNAs e análise fenotípica de esquistossômulos

A região dos genes utilizados na síntese de dsRNA foi primeiramente amplificada e sequenciada, e as sequências geradas foram comparadas às sequências preditas. Nenhum erro de predição foi encontrado após análise do sequenciamento. Dessa forma, seguimos com a síntese *in vitro* dos dsRNAs. Os dsRNAs dos genes selecionados, assim como o dsRNA do controle inespecífico (vetor pCRII-GFP) foram sintetizados, purificados e analisados em gel de agarose 1% não desnaturante (**Figura 16**). A eficiência da síntese variou entre os alvos escolhidos, fato que pode estar relacionado à concentração dos produtos recuperados após purificação da PCR ou ao próprio processo de síntese do dsRNA. No entanto, após quantificação, a mesma concentração de dsRNA (100nM) foi utilizada nos experimentos subsequentes.



**Figura 16: Análise da síntese de dsRNA para os alvos SmCaMK2, SmJNK, SmERK-1, SmERK-2, SmRas e o controle inespecífico GFP.** Gel agarose 1% não desnaturante representando os dsRNAs sintetizados. Caneleta 1 (MW) padrão peso molecular 100 pares de base (TrackIt™ DNA Ladders – Life Technologies). A seta indica a banda referência de 500 pares de base. O tamanho esperado para cada dsRNA é: CaMK2 (504 pb); JNK (570pb); ERK-1 (566 pb); ERK-2 (538 pb); Ras (506 pb); GFP (360 pb).

Os dsRNA foram inseridos na cultura de esquistossômulos e observados durante todos os dias por até 7 dias. Foi feita uma busca por alterações fenotípicas como morte, perda ou lenta movimentação, ruptura do tegumento ou outra alteração clara no comportamento do parasito. As análises foram feitas através de comparação aos grupos de parasitos não tratados ou tratados com dsRNA GFP. Imagens foram obtidas após sete dias e nenhum fenótipo relacionado aos tratamentos com dsRNAs foi observado nas culturas de esquistossômulos (**Figura 17**).



**Figura 17: Imagens dos esquistossômulos tratados com dsRNA SmCaMK2 no sétimo dia de cultivo.** As imagens foram obtidas em microscópio invertido sob lente de aumento de 5x e 10X. As duas primeiras imagens são esquistossômulos sem tratamento, as imagens do meio são esquistossômulos tratados com dsRNA GFP e as últimas são esquistossômulos tratados com dsRNA CaMK2.

A técnica de RNAi não produz o *knockout* do gene e sim a diminuição da expressão do gene alvo e esse fato pode estar relacionado a ausência de alterações fenotípicas observados nesse ponto do trabalho. Além disso, dependendo da função desempenhada por cada produto gênico, uma diminuição modesta na expressão pode ser prejudicial, assim como uma grande supressão da expressão gênica não pode ocasionar mudanças aparentes. Ainda deve ser considerado que o gene alvo pode estar envolvido com funções essenciais em outras fases do desenvolvimento do parasito e que as alterações fenotípicas ocasionadas pela supressão do gene só venham causar algum efeito em outro momento do ciclo de vida do parasito, como por exemplo, na transformação dos esquistossômulos em vermes adultos. Somado a isso, existe o fato de que a absorção do dsRNA numa população de

esquistossômulos não é uniforme (112;327) e que órgãos específicos podem ou não estar sendo atingidos, não favorecendo a observação de fenótipos.

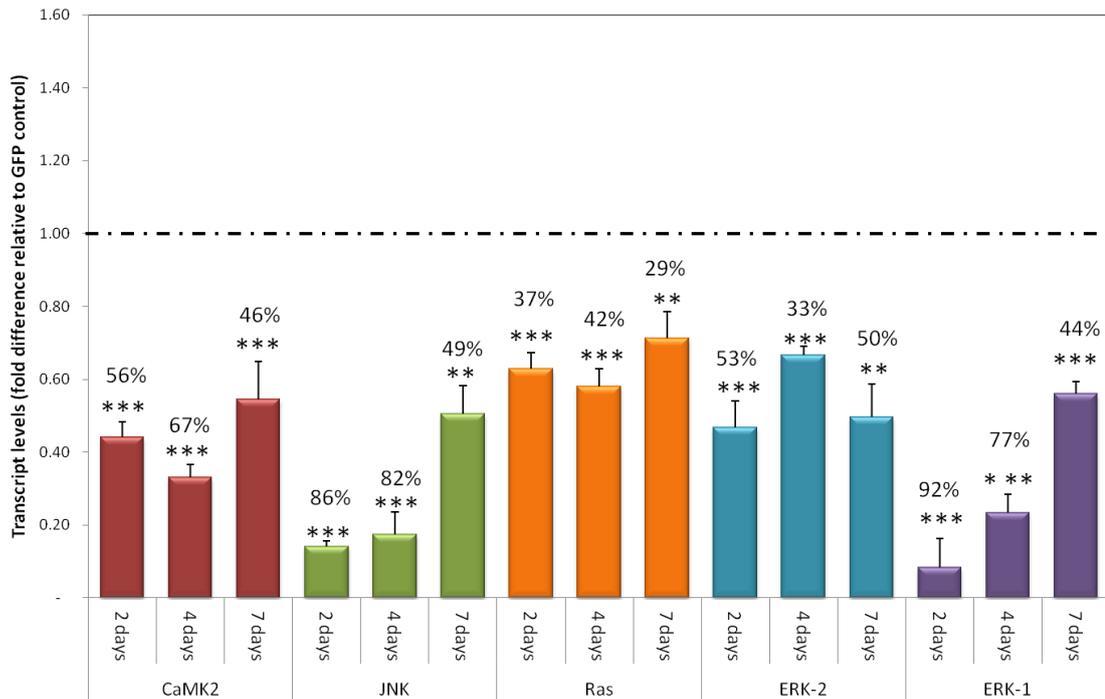
### 5.7-3. Estudo do nível de transcrito em esquistossômulos tratados com dsRNA

A PCR em tempo real quantitativa (QPCR) foi utilizada para verificar os níveis de transcritos dos genes silenciados nos parasitos tratados com dsRNAs. Os níveis de transcritos de cada alvo foi comparado com o nível dos mesmos transcritos em esquistossômulos tratados com dsRNA GFP (controle).

É preciso ressaltar que todos os iniciadores utilizados na QPCR foram desenhados para amplificar uma pequena região do gene fora da região utilizada para síntese do dsRNA, evitando assim a detecção de transcritos de dsRNAs durante a PCR. Foi constatado nesse trabalho que SmCaMK2 (Smp\_011660.1) possui ainda dois produtos de *splicing* alternativo (Smp\_011660.2 e Smp\_011660.3). Assim, uma região 100% idêntica entre os três transcritos preditos foi selecionada para a síntese do dsRNA para esse gene. O iniciador usado para avaliar os níveis de transcritos SmCaMK2 na RT\_qPCR foi desenhado fora da região utilizada para a síntese de dsRNA mas, em uma parte também conservada entre as isoformas.

Foi observado por Mourão e colaboradores (112;328) que a redução no nível de transcrito por RNAi em *Schistosoma* não é constante e que cada transcrito apresenta um *turnover* diferente. Dessa forma, foi avaliado nesse trabalho o tempo ideal para o silenciamento de cada um dos genes escolhidos. Para isso, a extração do RNA total da cultura de esquistossômulos foi realizada após dois, quatro e sete dias da adição dos dsRNA. Como mostra a **Figura 18**, foi constatado que a eficiência do silenciamento variou de gene para gene e também durante os diferentes dias, sendo que a maior foi a redução de 92% observada para SmERK-1 após dois dias e a menor foi a redução de 29% para SmRas após 7 dias. No entanto, foi observado uma diminuição significativa nos níveis de transcritos para todos os genes silenciados, sendo mais evidente para SmJNK (86%) e SmERK-1 (92%) após dois dias de tratamento (**Figura 18**). Ainda para esses dois genes, um aumento gradativo da expressão no quarto e sétimo dias de tratamento foi observado, chegando a aproximadamente 50% ao final das análises. Diferente dos demais genes, o silenciamento de SmCaMK2 e SmRas foi melhor no quarto dia, sendo este último, o que apresentou menor eficiência no silenciamento (no máximo 42% de inibição).

Algumas hipóteses podem explicar a pouca eficiência de silenciamento alcançada para alguns genes, entre elas, a possibilidade do gene alvo ser expresso em células que são restritivas para a maquinaria do RNAi ou devido a estrutura secundária do mRNA ser menos acessível ao complexo RISC (329). Além disso, foi utilizado nesse trabalho longos dsRNAs (~500pb) que após processamento pelo complexo DICER é degradado em pequenos siRNAs (~22 pb) com sequências diversas. Já foi descrito que a eficiência do silenciamento é dependente das sequências utilizadas para a síntese dos dsRNAs (330). Dessa forma, para evitar degradação inespecífica de outros transcritos foi necessário escolher, para a síntese dos dsRNAs, parte das sequências que são pouco conservadas entre diferentes ePKs, ou seja, regiões que não caracterizam um domínio proteico. No entanto, ePKs do grupo CMGC (ERK e JNK) (**Figura 14**) possuem um domínio catalítico que cobre quase toda a extensão do gene. Essa característica pode ter gerado pouca especificidade na ligação dos siRNAs ao gene alvo levando a resultados não muito satisfatórios como os observados para SmRas e SmERK-2. Existe ainda a possibilidade do dsRNA de SmERK-1 estar interferindo na transcrição de SmERK-2, já que esses genes pertencem a mesma subfamília e possuem sequências altamente similares. Para testar essa hipótese, os níveis de transcritos de SmERK-2, foram analisadas nas amostras tratadas durante dois dias com dsRNA ERK-1. Foi constatada uma diminuição de 56% nos níveis de SmERK2 (nível de transcrito em relação a GFP = 0,44 +/- 0,06), resultado mais satisfatório do que aquele conseguido após o tratamento com dsRNA específico para gene SmERK-2 (**Figura 18**), confirmando assim o *off-target*. Sendo assim, o dsRNA SmERK-1 silencia os dois genes da subfamília ERK1 de *S. mansoni* (SmERK-1 e SmERK-2) e com isso, qualquer alteração fenotípica ocasionada nos parasitos tratados com dsRNA SmERK1, é, na verdade, um efeito do silenciamento dos dois genes da subfamília ERK.

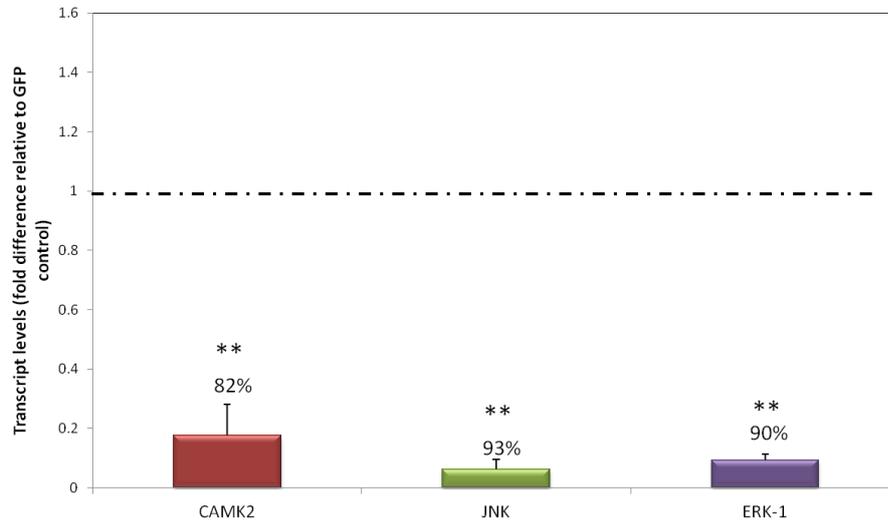


**Figura 18: Níveis de transcritos dos genes alvos em relação ao controle após tratamento dos esquistossômulos com dsRNA durante 2, 4 ou 7 dias.** Os níveis dos transcritos foram comparados com aqueles dos esquistossômulos tratados com dsRNA GFP (controle). Os resultados foram analisados pelo método de  $\Delta\Delta CT$  seguido do teste de Mann-Whitney (*Wilcoxon signed-rank*);  $P < 0.006$ ;  $N \geq 3$ . Todos os genes alvos tiveram diminuição significativa no nível de transcrito (\*). A linha pontilhada representa o nível normalizado dos transcritos na população controle. A porcentagem de inibição para cada gene está descrito acima das barras.

#### 5.7-4. Testes *in vivo*: infecção de camundongos com esquistossômulos silenciados

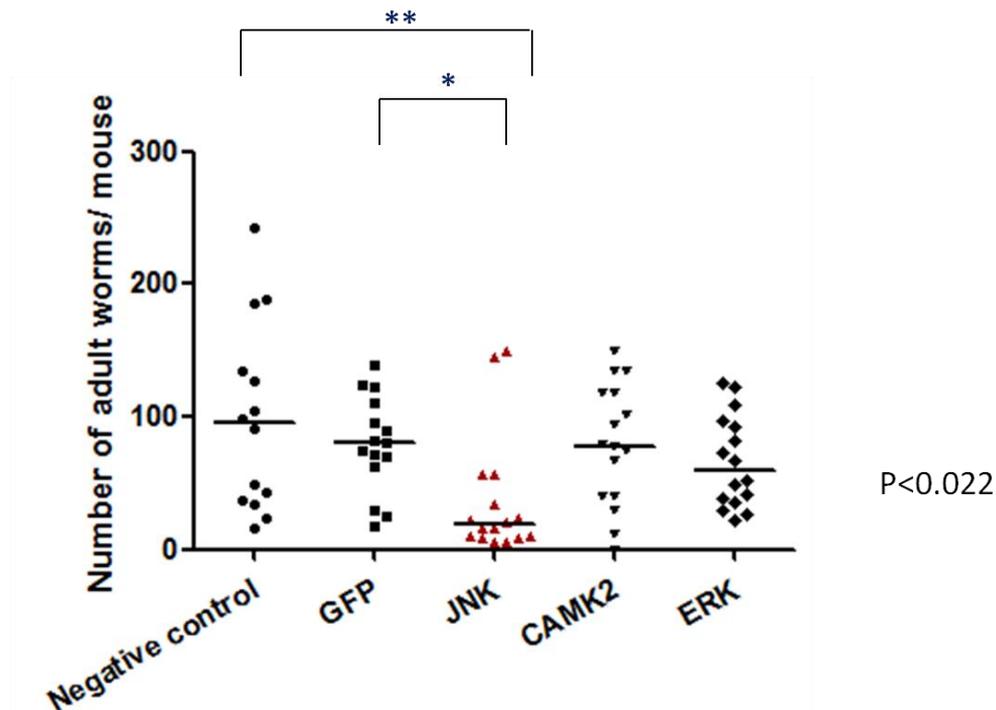
Para investigar o papel biológico de proteínas envolvidas na via de sinalização das proteínas MAPK frente ao sistema imune do hospedeiro, camundongos foram infectados com esquistossômulos silenciados. Assim, foi possível observar a transformação dos esquistossômulos em vermes adultos e o desenvolvimento do parasito de forma mais natural ao encontrado por *Schistosoma* durante o seu ciclo de vida. Para isso, dsRNAs SmCaMK2, SmJNK, SmERK1, que tiveram maior redução da expressão após silenciamento, e dsRNA do controle GFP, foram novamente utilizados em cultura de esquistossômulos. Após dois dias, 300 parasitos foram injetados em camundongos. Para garantir que o silenciamento foi eficaz nessas amostras, outra parte dos esquistossômulos foi separada para análise por QPCR.

A análise por QPCR confirmou uma diminuição significativa dos transcritos alvos (**Figura 19**), como esperado, o que garante que qualquer fenótipo observado nos testes *in vivo* possa ser um efeito do silenciamento. Sendo assim, após 35 dias da infecção foi feita a perfusão dos camundongos, os vermes foram contados, o fígado foi retirado e os ovos contidos nesse órgão foram extraídos e contados.



**Figura 19: Níveis de transcritos de SmCaMK2, SmJNK e SmERK-1 dos esquistossômulos silenciados utilizados na infecção de camundongos.** Os níveis dos transcritos foram comparados com aqueles dos parasitos tratados com dsRNA GFP (controle). Os resultados foram analisados pelo método de  $\Delta\Delta\text{CT}$  seguido do teste de Mann-Whitney (Wilcoxon signed-rank);  $P < 0.002$ ;  $N = 3$ . A linha pontilhada representa o nível normalizado dos transcritos na população controle. A porcentagem de inibição para cada gene está descrito acima das barras. Todos os genes alvos tiveram diminuição significativa no nível de transcrito (\*).

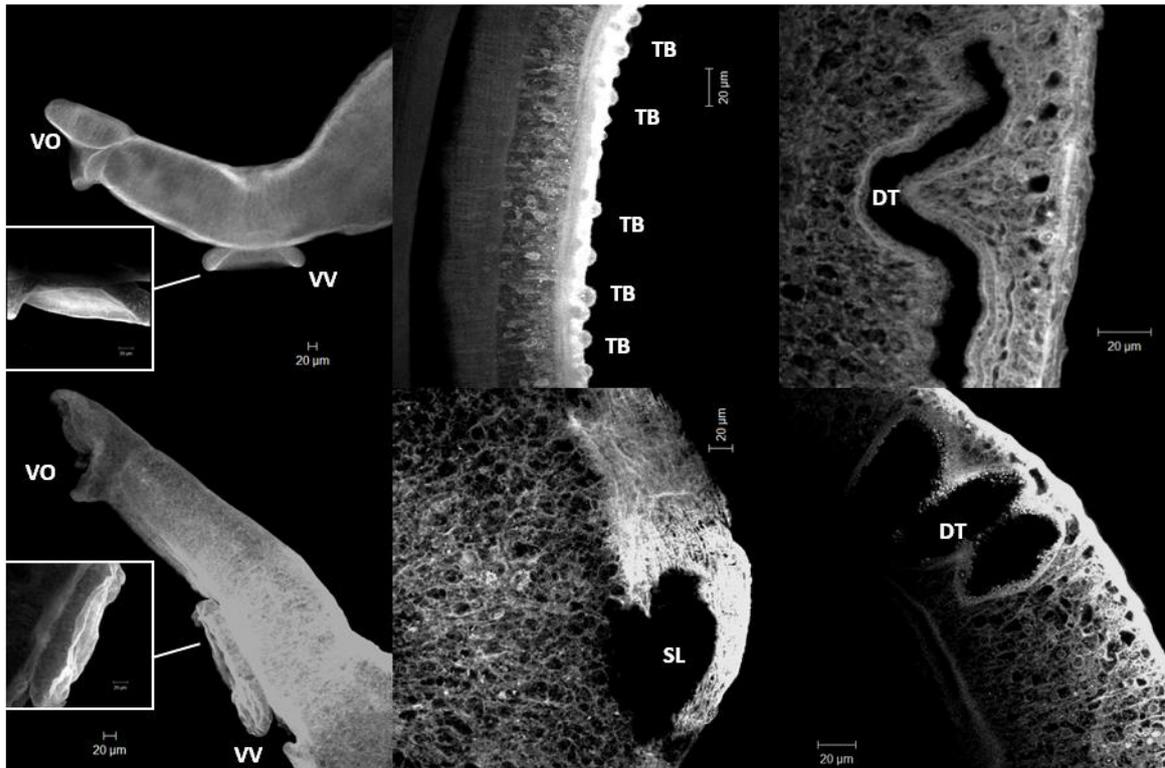
Com relação a contagem de vermes adultos, foi observada uma diminuição significativa no número de parasitos naqueles camundongos infectados com esquistossômulos silenciados para SmJNK (**Figura 20**). Esse dado sugere que essa ePK é essencial para a transformação e/ou sobrevivência de *S. mansoni* no organismo dos hospedeiros. Esse fato é ainda reforçado pela função previamente descrita para a JNK em outras espécies como em *D. melanogaster*, em que a via JNK é responsável por processos críticos do desenvolvimento (331) e, em planária (*Dugesia japonica*), em que via JNK coordena a formação do blastema (massa de células capazes de crescer e regenerar dentro dos órgãos ou partes do corpo) através da regulação da transição G2/M do ciclo celular durante o processo de regeneração desses animais (332).



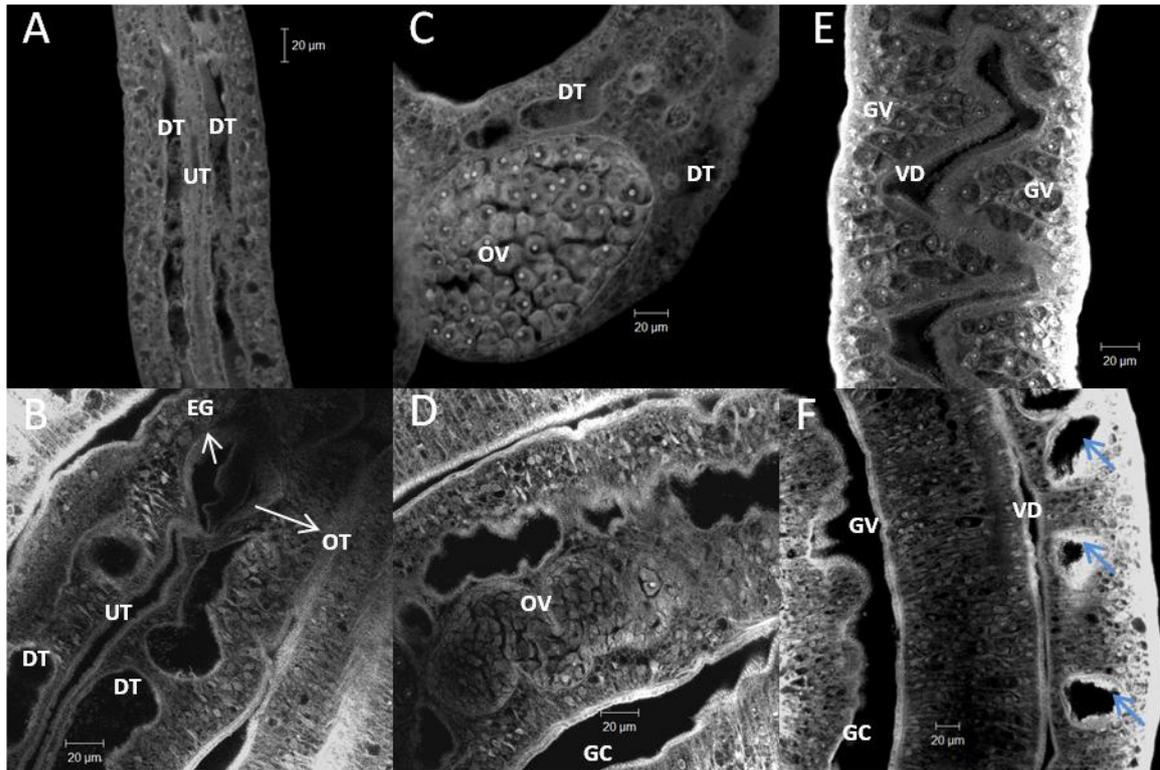
**Figura 20: Comparação do número de vermes adultos obtidos após perfusão de camundongos infectados com esquistossômulos silenciados em relação aos controles GFP e branco (*negative control*). Os (\*) representam aquelas amostras que tiveram diminuição significativa ( $P < 0.022$ ) no número de vermes perfundidos por camundongo, quando comparados aos controles. As linhas pretas horizontais representam a mediana do número de vermes total para cada gene alvo.**

As análises morfométricas comparando os vermes silenciados para o gene SmJNK e os controles, não mostraram qualquer alteração na área ou no número dos lóbulos testiculares dos machos e na área dos ovários das fêmeas. No entanto, as avaliações das estruturas dos parasitos por microscopia confocal mostraram que os vermes machos e fêmeas analisados apresentavam dilatações incomuns no tegumento e/ou no tubo digestivo (**Figura 21 e Figura 22**). Além disso, apesar de termos encontrado ovos, os ovários se apresentaram imaturos na maioria das fêmeas (**Figura 22 D**). Sendo assim, é provável que o pequeno número de vermes adultos recuperados nos camundongos infectados com esquistossômulos silenciados para o gene JNK seja um efeito da má formação do tegumento que levou à morte a maioria dos vermes. Além disso, como dito anteriormente, a absorção do dsRNA numa população de esquistossômulos não é uniforme e, com isso, existe ainda a possibilidade dos poucos vermes sobreviventes no grupo tratado com dsRNA SmJNK serem aqueles cujo silenciamento não foi eficaz suficiente para ocasionar a morte dos parasitos. O completo *knockout* do gene seria

necessário para avaliar se a falta de SmJNK seria suficientemente letal no processo de crescimento de *S. mansoni*.

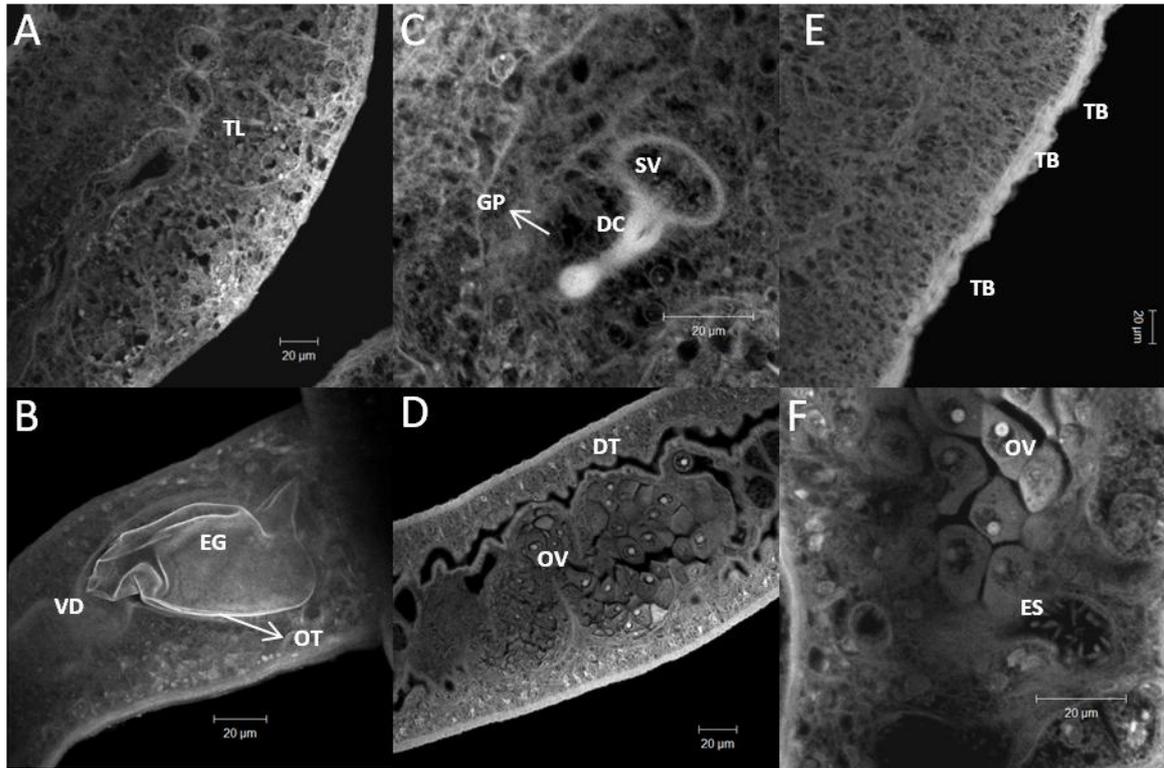


**Figura 21: Imagens microscopia confocal a laser de vermes adultos macho de *S. mansoni*.** A, C e E são vermes adultos tratados com dsRNA GFP e B, D e F são vermes adultos tratados com dsRNA JNK. As estruturas estão indicadas pelas seguintes letras VO: ventosa oral; VV: ventosa ventral; TB: tubérculos; SL: lesão subtegumentar; DT: tubo digestivo. Em A e B podemos observar a ventosa ventral destacada e irregularidade no tegumento da ventosa em B quando comparado ao controle. Em D, além de não observamos os tubérculos na superfície do macho, dilatações anormais também são encontradas (SL). Em F o tubo digestivo encontra-se anormalmente dilatado. As barras indicam 20 µm.



**Figura 22: Imagens microscopia confocal a laser de vermes adultos fêmea de *S. mansoni*.** A, C e E são vermes adultos tratados com dsRNA GFP e B, D e F são vermes adultos tratados com dsRNA JNK. As estruturas estão indicadas pelas seguintes letras: DT: tubo digestivo; UT: útero; EG: ovo; OT: oótipo; OV: ovário; GC: canal ginecóforo; GV: glândulas vitelínicas; VD: viteloducto. Em B e F (destacados pelas setas azuis) podemos observar dilatações no intestino das fêmeas. Em D observamos o ovário pequeno e com células não diferenciadas. As barras indicam 20 µm.

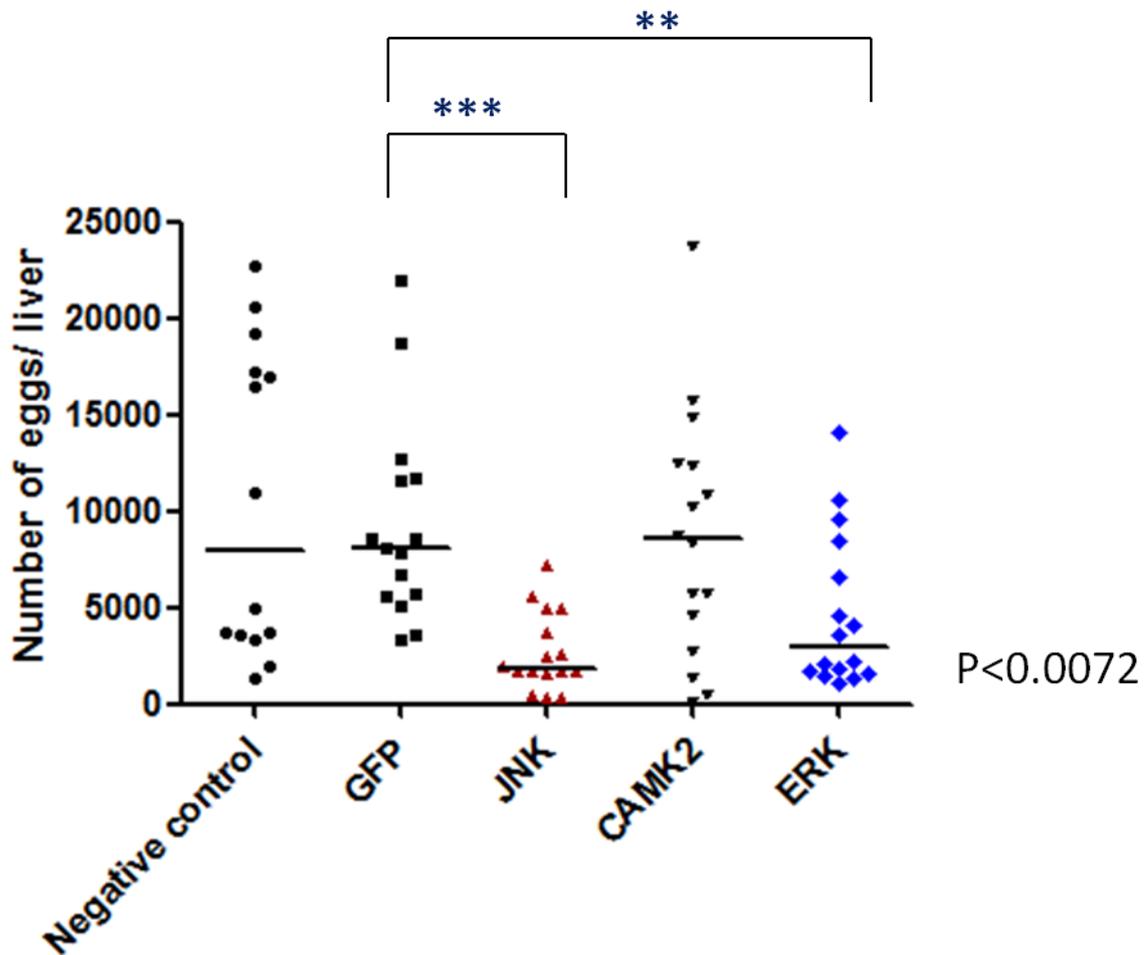
Em uma das vias MAPK de *C. elegans*, CaMK2 ativa JNK (333). É possível deduzir pelo gráfico da (Figura 20) e pelas análises por microscopia, que CaMK2 não ativa ou não é o único ativador da via JNK em *S. mansoni*, já que o efeito do silenciamento do gene ativador, no caso SmCaMK2, teria que ser o mesmo do gene ativado (SmJNK). O que foi visto é que não houve qualquer alteração significativa no número de vermes naqueles camundongos silenciados para o gene SmCaMK2 quando comparado aos controles e nem nas características estruturais dos órgãos dos vermes adultos recuperados após perfusão. Vermes adultos recuperados nos camundongos infectados com esquistossômulos silenciados para o gene SmCaMK2, apresentaram todas as estruturas normais e visivelmente viáveis (Figura 23).



**Figura 23: Imagens de microscopia confocal a laser de vermes adultos macho (A, C e E) e fêmea (B, D e F) de *S. mansoni* tratados com dsRNA CaMK2.** As estruturas estão indicadas pelas seguintes letras TL: lóbulos testiculares; GP: poro genital; SV: vesícula seminal; DC: ducto para o poro genital; TB: tubérculos; EG: ovo; OT: oótipo; VD: viteloducto; DT: tubo digestivo; OV: ovário; ES: espermateca. Todos os órgãos encontram-se normais. Observamos os lóbulos testiculares normais (A) além da presença da vesícula seminal com espermatozoides. O tegumento está íntegro e com tubérculos (E). O ovário está diferenciado (D) e na espermateca podemos observar a presença de espermatozoides (F). O ovo se apresenta em formato normal e fecundado (B). As barras indicam 20 µm.

Ao contar o número de ovos retirados dos fígados dos camundongos (**Figura 24**) foi observado uma diminuição significativa naqueles infectados com esquistossômulos silenciados para SmJNK. Isso era esperado já que o número de vermes adultos também estava reduzido nessas amostras. No entanto, para aqueles camundongos cuja infecção foi feita com esquistossômulos silenciados para o gene SmERK-1, foi também visualizado uma redução significativa no número de ovos. Isso sugere que a diminuição nos níveis de transcritos SmERK-1 e SmERK-2 não têm nenhum efeito na transformação e sobrevivência de vermes adultos de *S. mansoni*, mas provavelmente está relacionado com o desenvolvimento dos órgãos reprodutores e/ou no acasalamento dos parasitos. Em *Xenopus laevis* proteínas ERK ajudam a coordenar a maturação dos oócitos e progressão do ciclo celular com ovulação e

fertilização (334) e em *C. elegans knockout* para o gene ERK-1 não há desenvolvimento da vulva e dos oócitos e conseqüentemente também não há postura de ovos (335).



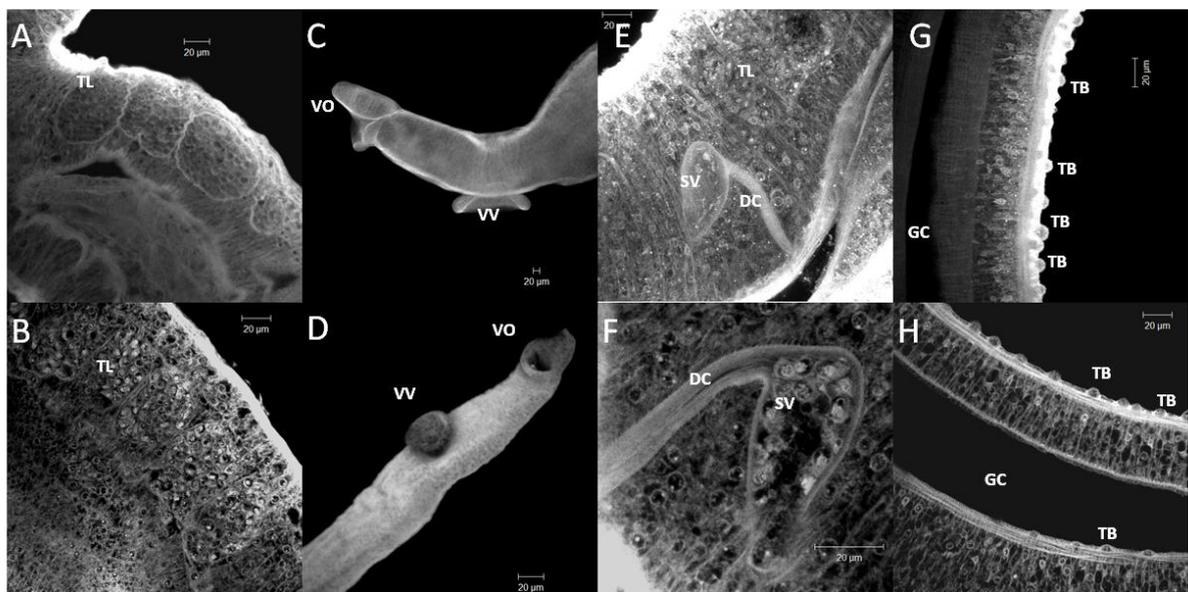
**Figura 24:** Comparação do número de ovos no fígado obtidos após perfusão de camundongos infectados com esquistossômulos silenciados, em relação aos controles GFP e branco (*negative control*). Os (\*) representam aquelas amostras que tiveram diminuição significativa ( $P < 0.0072$ ) no número ovos quando comparados aos controles. As linhas pretas horizontais representam a mediana do número de ovos total para cada gene alvo.

As análises por microscopia confocal mostraram que os machos recuperados dos camundongos infectados com esquistossômulos silenciados para o gene SmERK1 possuem lóbulos testiculares normais, vesícula seminal funcional com espermatozoides visíveis e tegumento íntegro (**Figura 25**). No entanto, a maioria das fêmeas apresenta ovário com células não diferenciadas e as análises morfométricas mostraram uma redução significativa na área dos ovários das fêmeas quando comparadas com o controle GFP (valor de  $P = 0.03$ ;

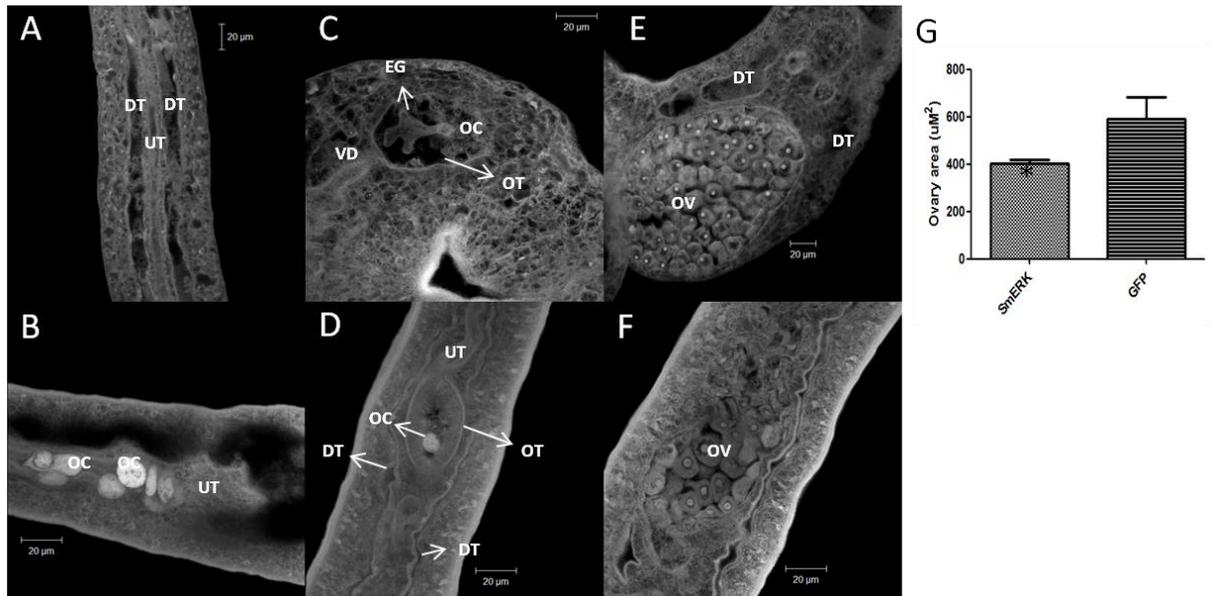
N=5). Uma das fêmeas ainda apresentava oócitos acumulados no útero e oócito não fecundado no oótipo (**Figura 26**).

Sendo assim, é possível que SmERK e SmJNK possam ser interessantes alvos para o desenvolvimento de drogas e que um inibidor contra essas proteínas provavelmente irá interromper o ciclo de *S. mansoni* impedindo o progresso da doença.

Apesar de SmCaMK2 ser o único gene da sua família identificado em *S. mansoni* e de ter sido destacado como potente alvo para o desenvolvimento de drogas, a diminuição desse transcrito não foi suficiente para causar qualquer alteração fenotípica nesse parasito. E este, provavelmente, não está relacionado ou não é o único ativador da via MAPK ERK ou JNK. Outra hipótese é de que o *turnover* desse gene seja mais rápido do que os de SmERK ou SmJNK sendo necessário o completo *knockout* do gene para a não ativação da via por SmCaMK2.



**Figura 25: Imagens de microscopia confocal a laser de vermes adultos macho de *S. mansoni*.** A, C, E e G são vermes adultos tratados com dsRNA GFP e B, D, F e H são vermes adultos tratados com dsRNA ERK1. As estruturas estão indicadas pelas seguintes letras: TL: lóbulos testiculares; VO: ventosa oral; VV: ventosa ventral; SV: vesícula seminal; DC: ducto para o poro genital; TB: tubérculos; GC: canal ginecóforo. Observamos os lóbulos testiculares normais (B) além da presença da vesícula seminal com espermatozoides (F). O tegumento está íntegro e com tubérculos (D e H). As barras indicam 20 µm.



**Figura 26: Imagens de microscopia confocal a laser de vermes adultos fêmea de *S. mansoni*.** A, C e E são vermes adultos tratados com dsRNA GFP e B, D e F são vermes adultos tratados com dsRNA ERK-1. As estruturas estão indicadas pelas seguintes letras DT: tubo digestivo; UT: útero; EG: ovo; OT: oótipo; OC: oócito; OV: ovário. Em B observamos um aglomerado anormal de oócitos no útero da fêmea e em D um oócito (não fecundado) no oótipo. Em F observamos o ovário pequeno e com células não diferenciadas. As barras indicam 20 µm. Em G está representado um gráfico comparando a área dos ovários das fêmeas tratadas com dsRNA ERK1 e controle GFP. O (\*) representa diminuição significativa na área (Teste de Mann-Whitney - Wilcoxon signed-rank -  $P < 0.03$ ;  $N = 5$ ).

### 5.7-5. Efeito do silenciamento na transcrição de genes alvos da via MAPK

Existem pelo menos dois modos principais de regulação gênica. No primeiro modo os genes exibem uma taxa variável de transcrição e necessitam que um fator específico se ligue ao gene promotor para que se inicie a tradução do gene e o sinal seja então transmitido, sendo esta uma via mais lenta. No segundo, o gene está presente numa taxa fixa de transcrição e em uma conformação estável nos modos ativo ou inativo e estão ligados diretamente a vias de sinalização rápidas que não necessitam que novos produtos sejam transcritos para a indução do sinal, como é o caso da maioria das vias de sinalização onde ePKs estão envolvidas, incluindo a via MAPK (336).

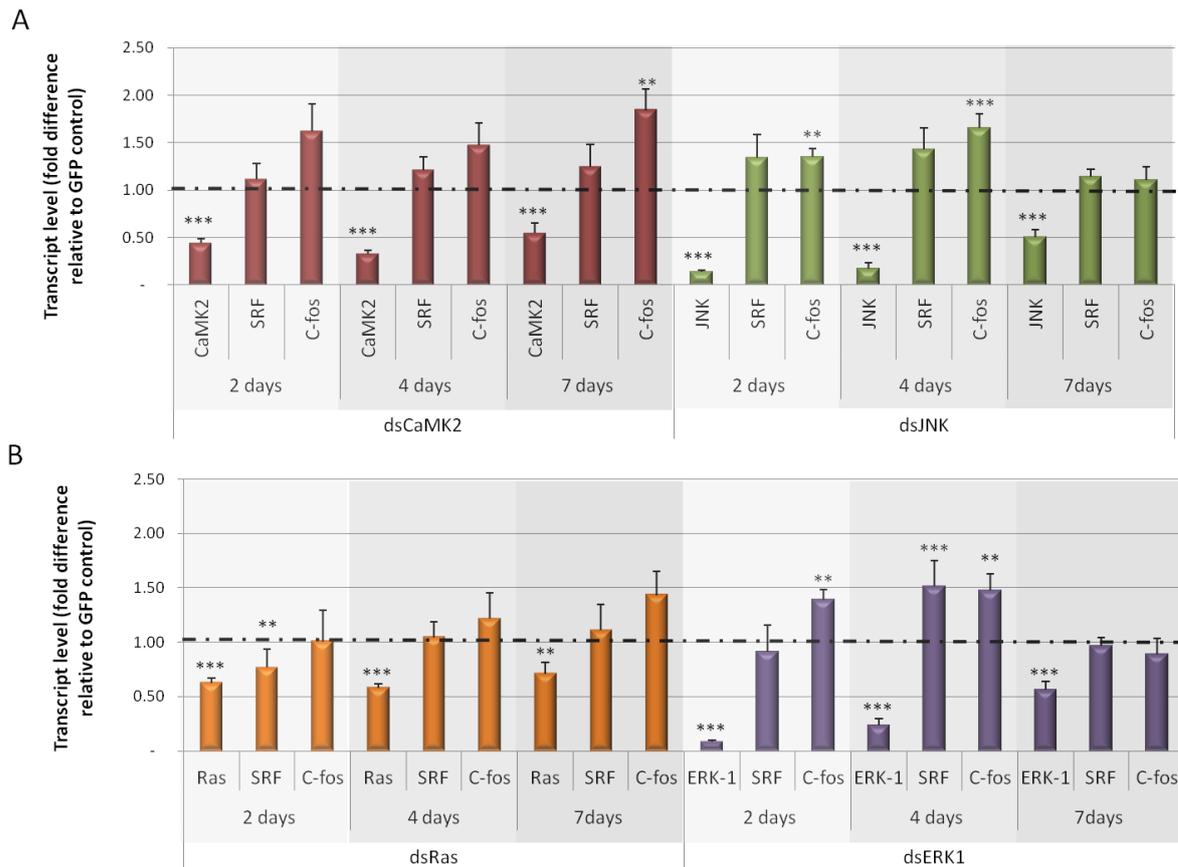
Na via MAPK, uma proteína ancorada a membrana transmite o sinal para o meio intracelular que rapidamente passa esse sinal ativando e/ou desativando uma série de proteínas ali presentes até chegar as MAPK propriamente ditas (337). JNK e ERK ativadas podem se translocar para o núcleo e ativar fatores de transcrição como elk-1 que forma um

complexo com outro fator de transcrição, o SRF, que induz, entre outras, a transcrição de genes específicos como c-Fos (338). Sendo assim, c-fos tem uma taxa de transcrição variável que é dependente da ligação do complexo elk-1/SRF, que por sua vez, tem uma taxa de transcrição fixa. O fator de transcrição c-Fos regula uma ampla gama de processos celulares, incluindo a proliferação celular, morte, sobrevivência e diferenciação. Estudos *in vivo* em camundongos apontam a importância de c-Fos para a fertilidade dos animais (339).

Com o intuito de verificar se a via MAPK que induz c-Fos em mamíferos está conservada em *S. mansoni*, foi avaliado o nível dos transcritos c-Fos e SRF após silenciamento dos genes SmRas, SmERK-1, SmJNK e SmCaMK2. O material (RNA total) utilizado nesse experimento foi o mesmo utilizado para quantificar os transcritos que são os alvos diretos do silenciamento (**Figura 18**). Sendo assim, já era sabido que o nível de transcrito SmRas, SmERK-1, SmJNK e SmCaMK2 estava significativamente reduzido nessas amostras.

O resultado mostra (**Figura 27**) que a quantidade de transcrito SmSRF se manteve constante na maioria das vezes, independente do alvo silenciado. Já Smc-Fos, está altamente transcrito em relação ao controle GFP nas amostras em que observamos uma baixa transcrição de SmERK-1, SmCaMK2 e SmJNK. Isso indica que há uma relação entre os genes silenciados e a transcrição de SmC\_fos, como acontece em mamíferos. No entanto, era esperado que essa relação fosse positiva, ou seja, baixos níveis de SmERK-1 e SmJNK iriam impedir a ativação de SRF que por sua vez não seria capaz de se ligar ao promotor de c-Fos impedindo a transcrição deste. O que foi observado foi uma relação negativa, no qual baixos níveis de transcrito SmERK-1, SmCaMK2 e SmJNK induzem a transcrição de c-Fos. Além disso, o gráfico da **Figura 27** nos mostra que à medida que a quantidade de transcrito SmJNK e SmERK aumenta o nível de transcrito SmC-fos diminui.

Apesar do aumento de transcritos SmC-fos após silenciamento de SmRas, esse aumento não foi significativo (**Figura 27**) muito provavelmente pelo fato do silenciamento de SmRas não ter sido tão significativo quanto os demais (SmJNK e SmERK).

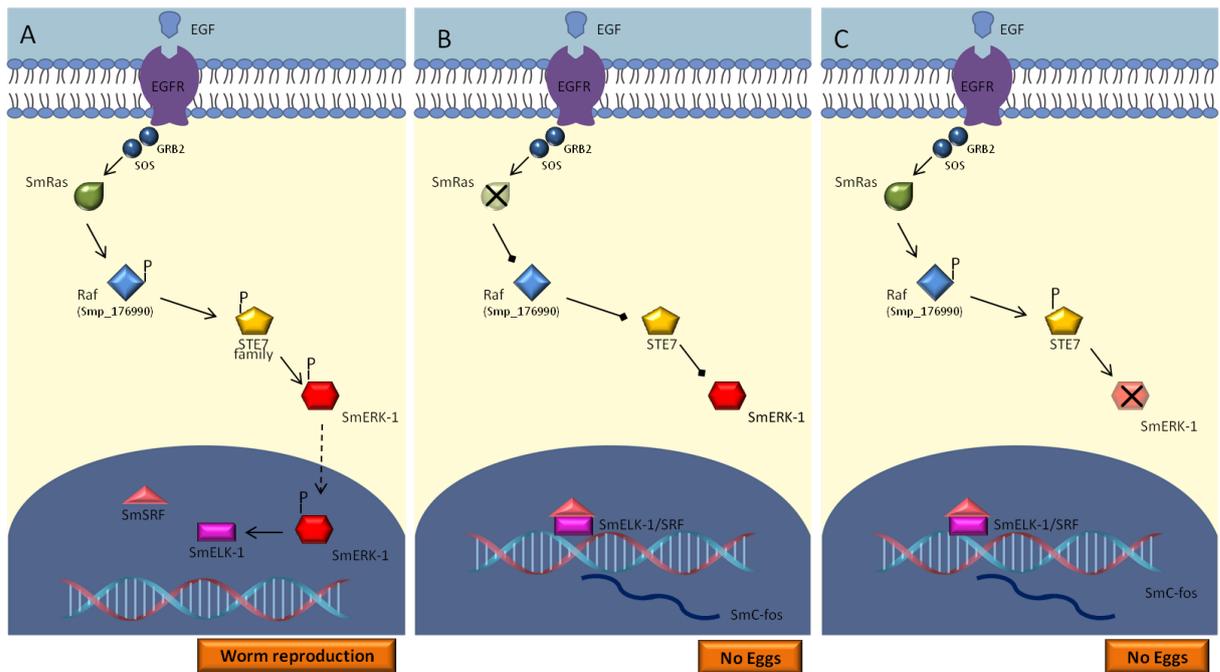


**Figura 27: Níveis de transcritos dos genes SmCaMK2, SmJNK, SmRas, SmERK, SmSRF e SmC-fos em relação ao controle após tratamento com dsRNAs dsCaMK2, dsJNK em A e, dsRas e dsERK1 em B, durante 2, 4 ou 7 dias. Os níveis dos transcritos foram comparados com aqueles dos parasitos tratados com dsRNA GFP (controle). Os resultados foram analisados pelo método de  $\Delta\Delta\text{CT}$  seguido do teste de Mann-Whitney (*Wilcoxon signed-rank*);  $P < 0.005$ ;  $N \geq 3$ . Os genes que tiveram diminuição significativa no nível de transcrito estão destacados por (\*). A linha pontilhada representa o nível normalizado dos transcritos na população controle.**

Possivelmente as vias de sinalização que ativam c-fos em *S. mansoni* são mais similares às de *C. elegans*, em que elk-1 (LIN-31 em *C. elegans*) e SRF (LIN-1 em *C. elegans*) formam um complexo quando MAPK não é fosforilada e este complexo ativa c-Fos que inibe o desenvolvimento da vulva (órgão necessário para a postura de ovos e para a cópula com machos). Quando MAPK é fosforilada este complexo é rompido e LIN-31 promove o desenvolvimento da vulva. Em *C. elegans* essa via é ativada pelo fator de crescimento epitelial (EGF) (340). Além disso, recentemente foi relatado que essa mesma via pode ser ativada pelo influxo de cálcio através da via de sinalização CaMK que induz c-Fos

pela fosforilação de CaMK2 que, por sua vez, fosforila SRF, sem uma relação direta com proteínas MAPK ERK ou JNK (341).

Sendo assim, é possível que o alto nível de transcrito c-Fos, verificado após silenciamento de SmERK, SmJNK e SmCaMK2, tenha uma relação com a não ativação da via de sinalização MAPK e isso levou a formação do complexo elk-1/SRF e indução da transcrição de C-Fos (**Figura 28**). Nesse caso SmCaMK2 e SmJNK agiram independente das vias Ras/ERK. Como não foi observado nenhum fenótipo relacionado à baixa produção de ovos em parasitos tratados com dsRNA CaMK2, é possível deduzir que SmCaMK2 atue em uma via diferente ou em células específicas que não estão relacionadas ao desenvolvimento dos órgão reprodutores. Já SmRas e SmERK atuariam em uma mesma via, como o que acontece em *C. elegans*, Mamíferos e *D. melanogaster*. EGF estaria ativando SmRas que então fosforilaria proteínas STE e MAPK ERK. SmERK se translocaria para o núcleo, fosforilando o fator de transcrição ELK-1 que estaria envolvido no desenvolvimento de órgão reprodutores relacionados a produção de ovos e/ou acasalamento (**Figura 28 A**). As fêmeas de *S. mansoni* não tem vulva, como em *C. elegans*, mas possuem uma abertura (poro genital) onde os ovos saem e onde ocorre a entrada dos espermatozóides. Não foi possível visualizar essa estrutura nas fêmeas que foram tratadas com dsRNA ERK1 e analisadas por microscopia confocal, mas, ovários pequenos e oócitos não fecundados e aglomerados no útero foram observados (**Figura 26**), o que comprova o envolvimento da via MAPK na reprodução de *Schistosoma*.



**Figura 28: Via de sinalização hipotética MAPK/ERK ativada por EGF que leva a regulação do gene C-fos em *S. mansoni*.** Em (A) EGF ativa a via de sinalização Ras/ERK. EGFR, uma proteína receptora localizada na membrana celular, transmite o sinal para meio intracelular através da ativação de Ras e fosforilação sequencial de SmRAF (ePK do grupo TKL e família raf), SmSTE7 (ePK do grupo STE e família STE7) e SmERK (ePK do grupo CMGC, família MAPK e subfamília ERK1/2). Ao ser fosforilada ERK se transloca para o núcleo e impede a formação do complexo ELK1/SRF e a postura de ovos se mantém constante. Em (B e C) SmRas ou SmERK são silenciadas e o sinal não é transmitido. ELK1 forma um complexo com SRF que se liga ao promotor c-Fos e inicia transcrição do gene que, de alguma forma, impede a postura de ovos.

## 6. CONCLUSÕES

Os resultados apresentados nesse trabalho mostram que as informações contidas nos genomas, quando bem trabalhadas, se tornam ferramentas de grande importância para melhorar o conhecimento em relação a funcionalidade de proteínas preditas e para compreender os mecanismos relacionados ao desenvolvimento dos organismos. Utilizando uma combinação de metodologias baseadas em similaridade de sequências e abordagens filogenéticas, foi possível identificar 243 ePKs em *S. mansoni*, 247 ePKs em *S. haematobium* e 213 ePKs em *S. japonicum* além de classificar mais de 95% desses genes. Em *S. japonicum* não existia qualquer informação relacionada às ePKs identificadas e, para *S. haematobium* e *S. mansoni* a anotação era baseada em métodos pouco confiáveis. Em *S. mansoni* apenas 12% das ePKs possuíam alguma validação experimental relacionada. Dessa forma foi possível melhorar a anotação para mais de 40% das ePKs de *S. haematobium* e *S. mansoni*. O ganho para a anotação das ePKs de *S. japonicum* foi inquestionável. A inferência filogenética se mostrou uma ferramenta eficaz para anotação do proteoma e também para seleção de alvos promissores. Foi possível entender melhor o papel das ePKs em *Schistosoma* e, os genes selecionadas para validação experimental, se mostraram promissores como alvos para tratamento da esquistossomose, o que reforça a eficácia da metodologia.

O silenciamento gênico por RNA de interferência é uma importante ferramenta para validação funcional de alvos. Uma diminuição significativa nos níveis de transcritos para todos os genes trabalhados foi conseguida. Apesar da não visualização de diferenças fenotípicas advindas do silenciamento em esquistossômulos cultivados *in vitro*, esse efeito foi visível em vermes adultos recuperados de camundongos infectados. Foi possível comprovar que fatores de transcrição relacionados a via MAPK em outros organismos também participam dessa via em *S. mansoni* e que existe uma conservação funcional da via em vermes. Finalmente, temos evidências de que as proteínas MAPK JNK e ERK de *S. mansoni*, são alvos importantes que devem ser melhor trabalhados para confecção de inibidores específicos já que esses, possivelmente, irão interromper o ciclo de vida do parasito e consequentemente a infecção.

## **7. ANEXOS**

### **7.1- Tabelas**

Tabela Anexo 1: Proteínas quinase eucarióticas do *Schistosoma mansoni*

Material digital, disponível em CD

Tabela Anexo 2: Proteínas quinase eucarióticas ortólogas entre *S. mansoni*, *S. haematobium* e *S. japonicum*

Material digital, disponível em CD

Tabela Anexo 3: Proteínas quinase eucarióticas do *Schistosoma haematobium*

Material digital, disponível em CD

Tabela Anexo 4: Proteínas quinase eucarióticas do *Schistosoma japonicum*

Material digital, disponível em CD

Tabela Anexo 5: Proteínas quinase eucarióticas sem ortólogos entre *S. mansoni* (Smp), *S. haematobium* (Sha) e *S. japonicum* (Sjp)

Protein accession	Kinase type	Group	Family	Subfamily
Smp_174820	Serine/Threonine Kinase	AGC	MAST	MASTL
Smp_157890	Serine/Threonine Kinase	AGC	NDR	
Smp_168120	Serine/Threonine Kinase	CK1	CK1	
Smp_047900	Serine/Threonine Kinase	CMGC	MAPK	ERK
Smp_133500	Serine/Threonine Kinase	CMGC	MAPK	ERK7
Smp_177920	Hybrid Protein Kinase	Other	NEK	
Smp_156890	Hybrid Protein Kinase	Other	SCY1	
Smp_094090	Serine/Threonine Kinase	STE	STE20	MST
Smp_137610	Tyrosine Kinase	TK	Fak	
Smp_118100	Tyrosine Kinase Receptor	TK	Ror	
Smp_054500	Tyrosine Kinase	TK	Src	
Smp_172200	Tyrosine Kinase	TK	Src	
Smp_143400	Hybrid Protein Kinase	TKL	LRRK	
Smp_158950	Protein Kinase	Unknown	Unknown	
Sjp_0072900	Serine/Threonine Kinase	Unknown	Unknown	
Sjp_0065080	Serine/Threonine Kinase	Unknown	Unknown	
Sha_301280	Serine/Threonine Kinase	Unknown	Unknown	
Sha_200410	Serine/Threonine Kinase	Unknown	Unknown	
Sha_104532	Serine/Threonine Kinase	Unknown	Unknown	
Sha_100538	Serine/Threonine Kinase	Unknown	Unknown	
Sha_102105	Serine/Threonine Kinase	Unknown	Unknown	
Sha_107692	Serine/Threonine Kinase	Unknown	Unknown	
Sha_105753	Serine/Threonine Kinase	Unknown	Unknown	
Sha_103671	Serine/Threonine Kinase	Unknown	Unknown	
Sha_104973	Serine/Threonine Kinase	Unknown	Unknown	
Sha_106629	Serine/Threonine Kinase	Unknown	Unknown	
Sha_109265	Serine/Threonine Kinase	Unknown	Unknown	
Sha_103569	Tyrosine kinase	Unknown	Unknown	
Sha_101889	Tyrosine kinase	Unknown	Unknown	

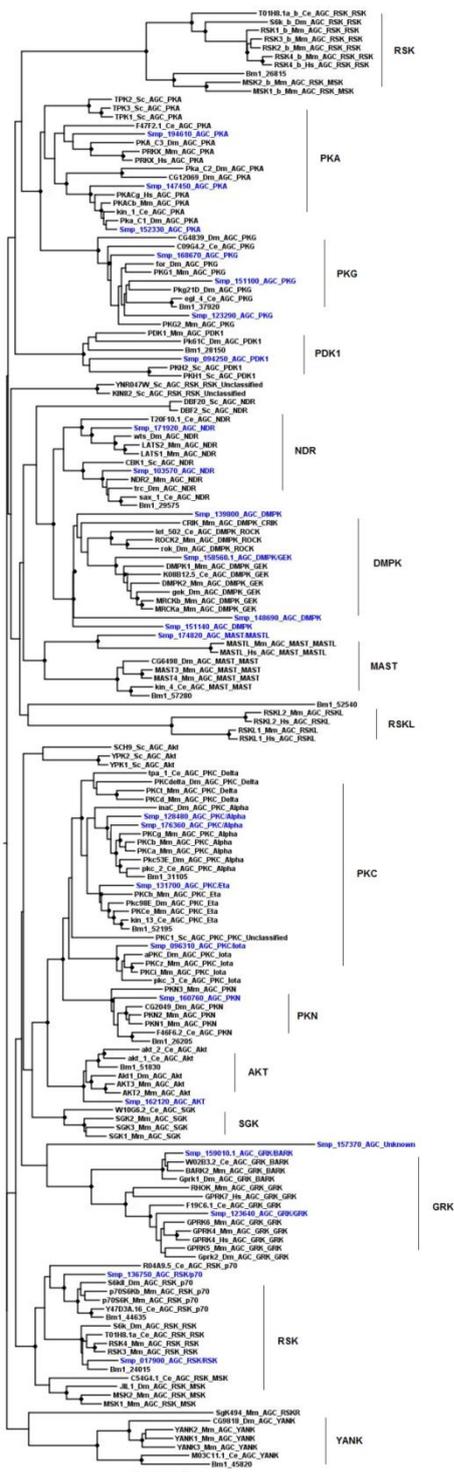
Tabela Anexo 6: Sequências ePKs não preditas na versão atual (3.0) do genoma de *S. mansoni*

Protein accession	Kinase type	Group	Family	Subfamily
Smp_152930	Serine/Threonine Kinase	CAMK	CAMKL	Unknown
Smp_159010.2	Serine/Threonine Kinase	AGC	GRK	BARK
Smp_126280.1	Hybrid Protein Kinase	Other	Haspin	
Smp_126280.2	Hybrid Protein Kinase	Other	Haspin	
	Tyrosine Kinase			
Smp_093500	Receptor	TK	Unknown	
	Tyrosine Kinase			
Smp_173590	Receptor	TK	EGFR	
Smp_118490	Tyrosine Kinase	TK	Src	
Smp_165310	Hybrid Protein Kinase	TKL	STKR	Type2

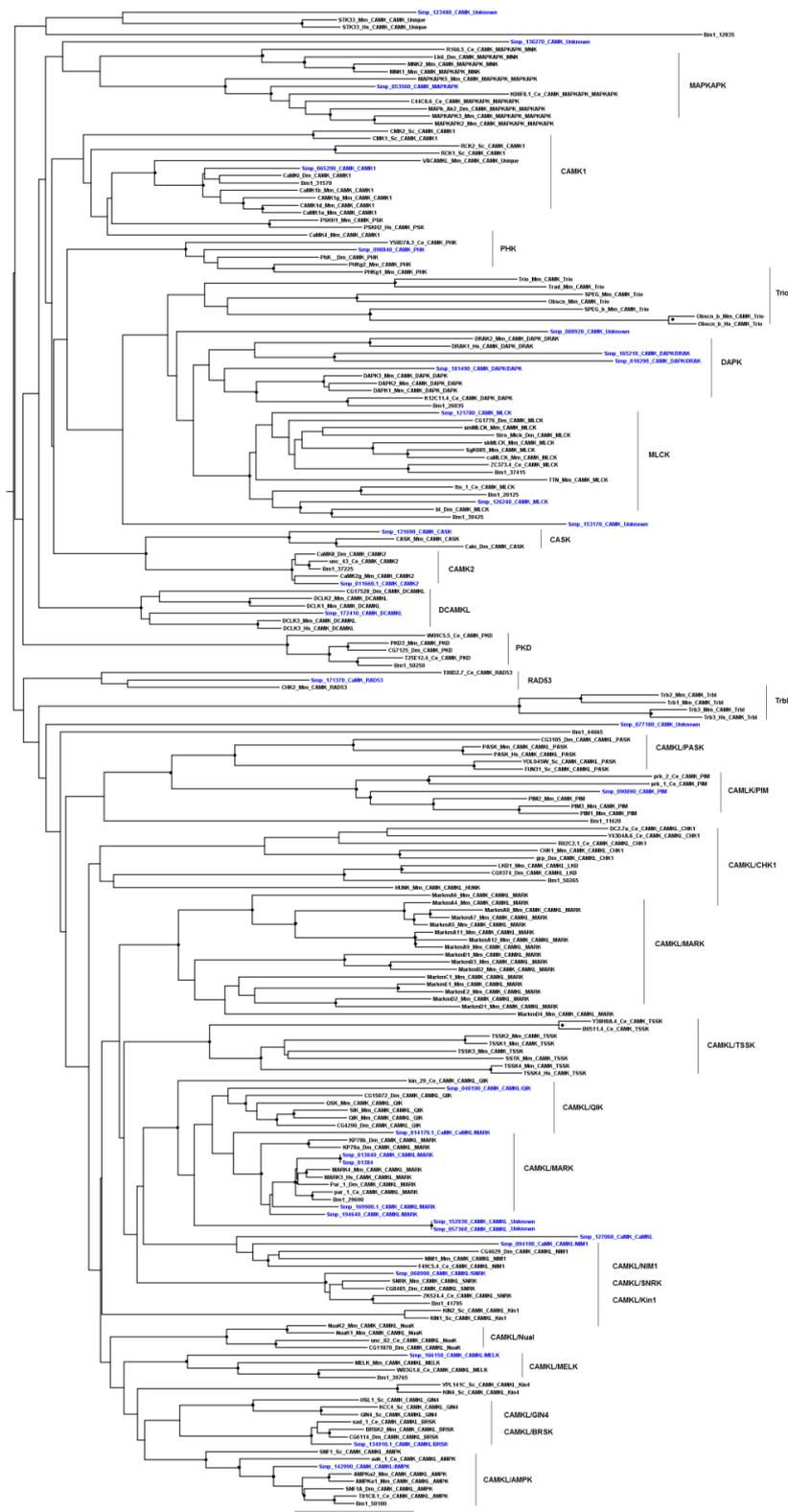
Tabela Anexo 7: Sequências ePKs que tiveram o número de acesso modificado na versão atual (3.0) do genoma de *S. mansoni*

Protein accession (Version 2.0)	Protein accession (Version 3.0)
Smp_162120	Smp_073930.1 e Smp_073930.2
Smp_080770	Smp_164380
Smp_131930	Smp_204050
Smp_173400.2	Smp_093540.3
Smp_008260.1	Smp_008260

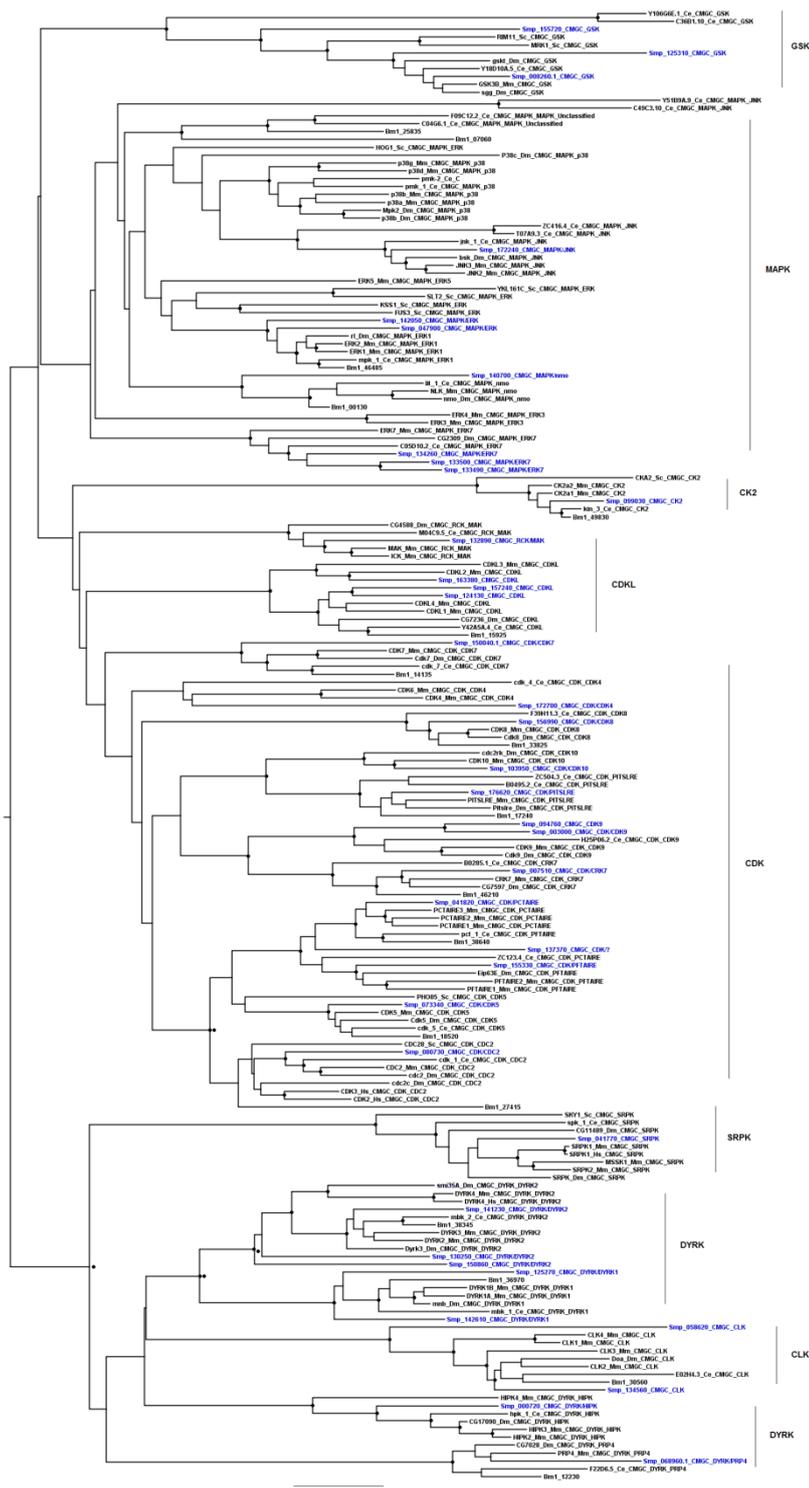
## 7.2- Figuras



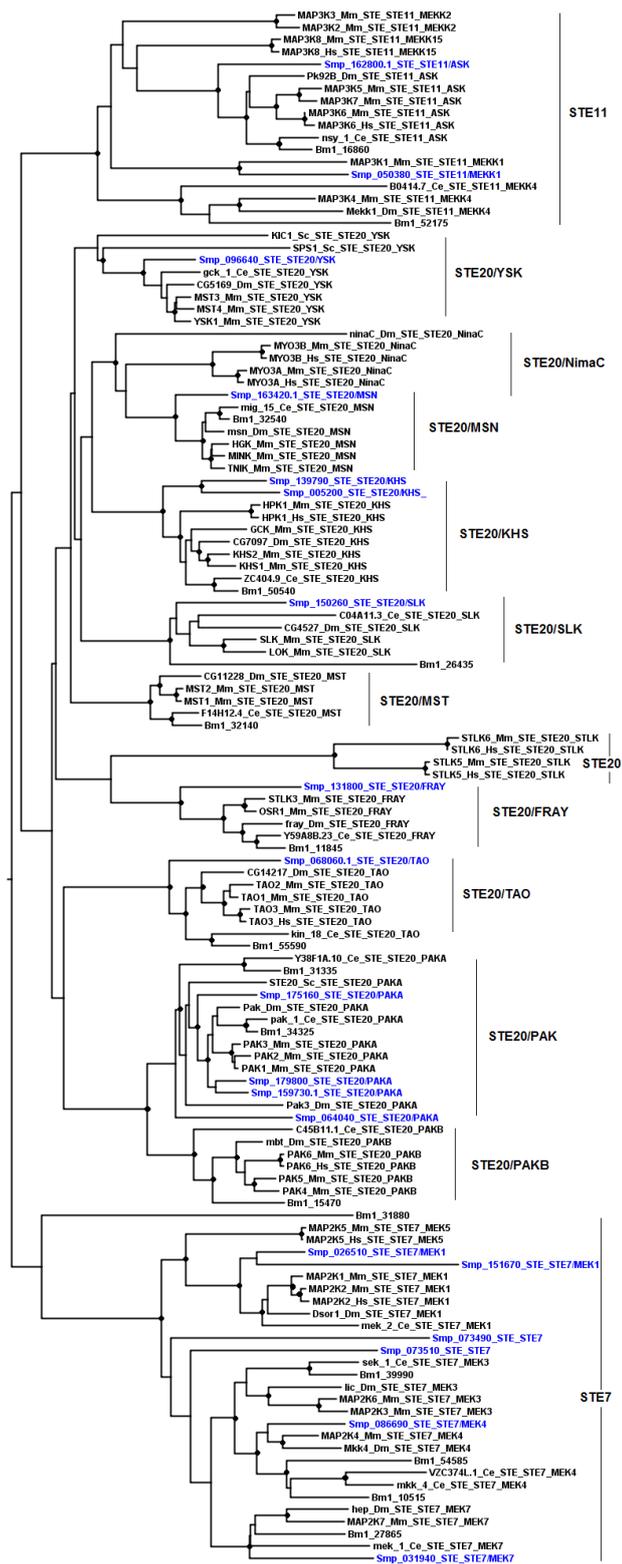
**Figura Anexo 1: Árvore filogenética baseada no método de distância para o grupo AGC. As ePKs de *S. mansoni* estão indicadas pelos labels na cor azul. Valores de Bootstrap (1000pseudo réplicas) iguais ou maiores que 80% (●) foram considerados para transferir a anotação de famílias/subfamílias (indicado individualmente em cada label). As famílias identificadas estão destacadas pelas barras verticais.**



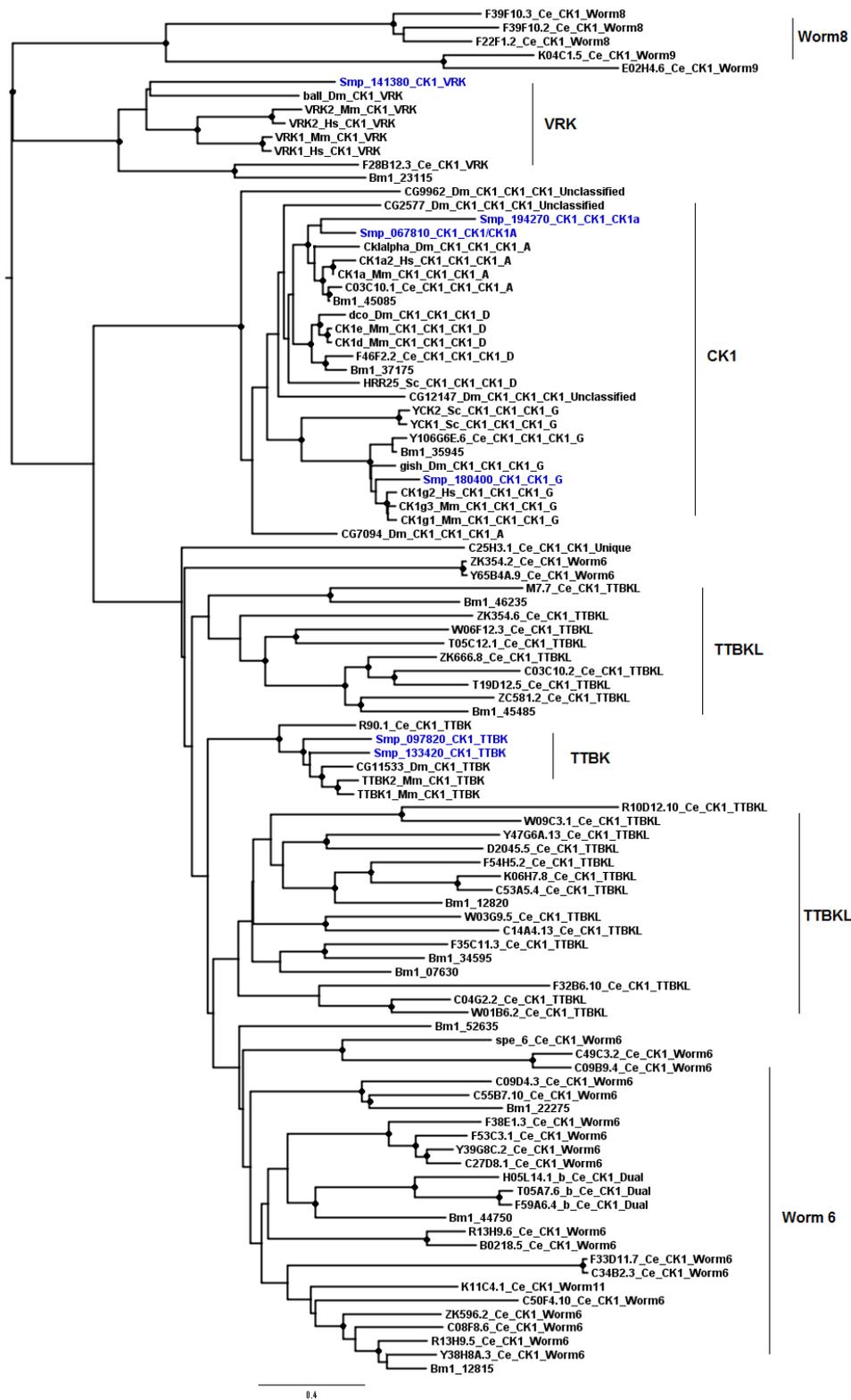
**Figura Anexo 2: Árvore filogenética baseada no método de distância do grupo CaMK.** As ePKs de *S. mansoni* estão indicadas pelos labels na cor azul. Valores de Bootstrap (1000pseudo réplicas) iguais ou maiores que 80% (●) foram considerados para transferir a anotação de famílias/subfamílias (indicado individualmente em cada label). As famílias identificadas estão destacadas pelas barras verticais.



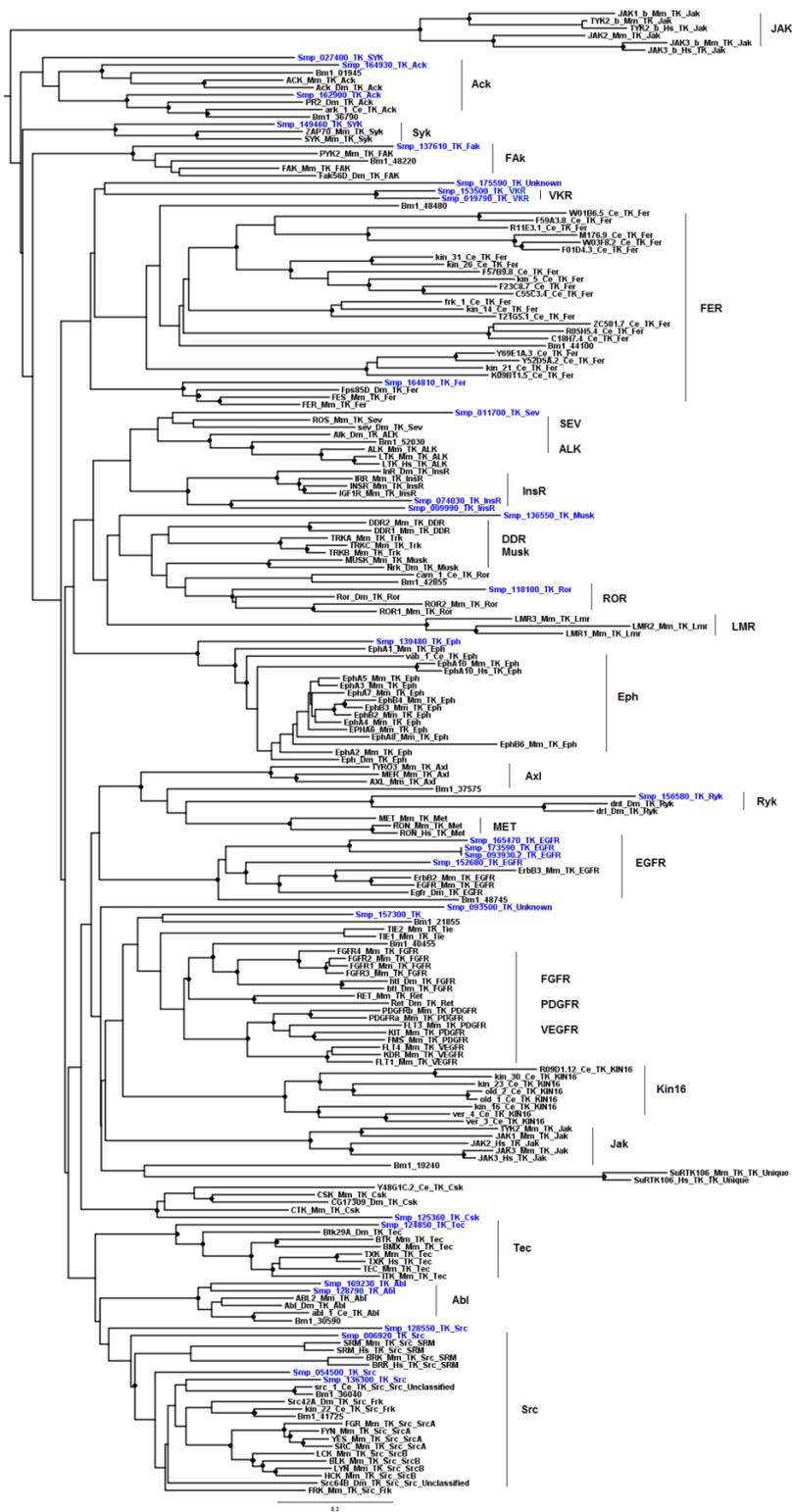
**Figura Anexo 3: Árvore filogenética baseada no método de distância do grupo CMGC.** As ePKs de *S. mansoni* estão indicadas pelos labels na cor azul. Valores de Bootstrap (1000pseudo réplicas) iguais ou maiores que 80% (●) foram considerados para transferir a anotação de famílias/subfamílias (indicado individualmente em cada label). As famílias identificadas estão destacadas pelas barras verticais.



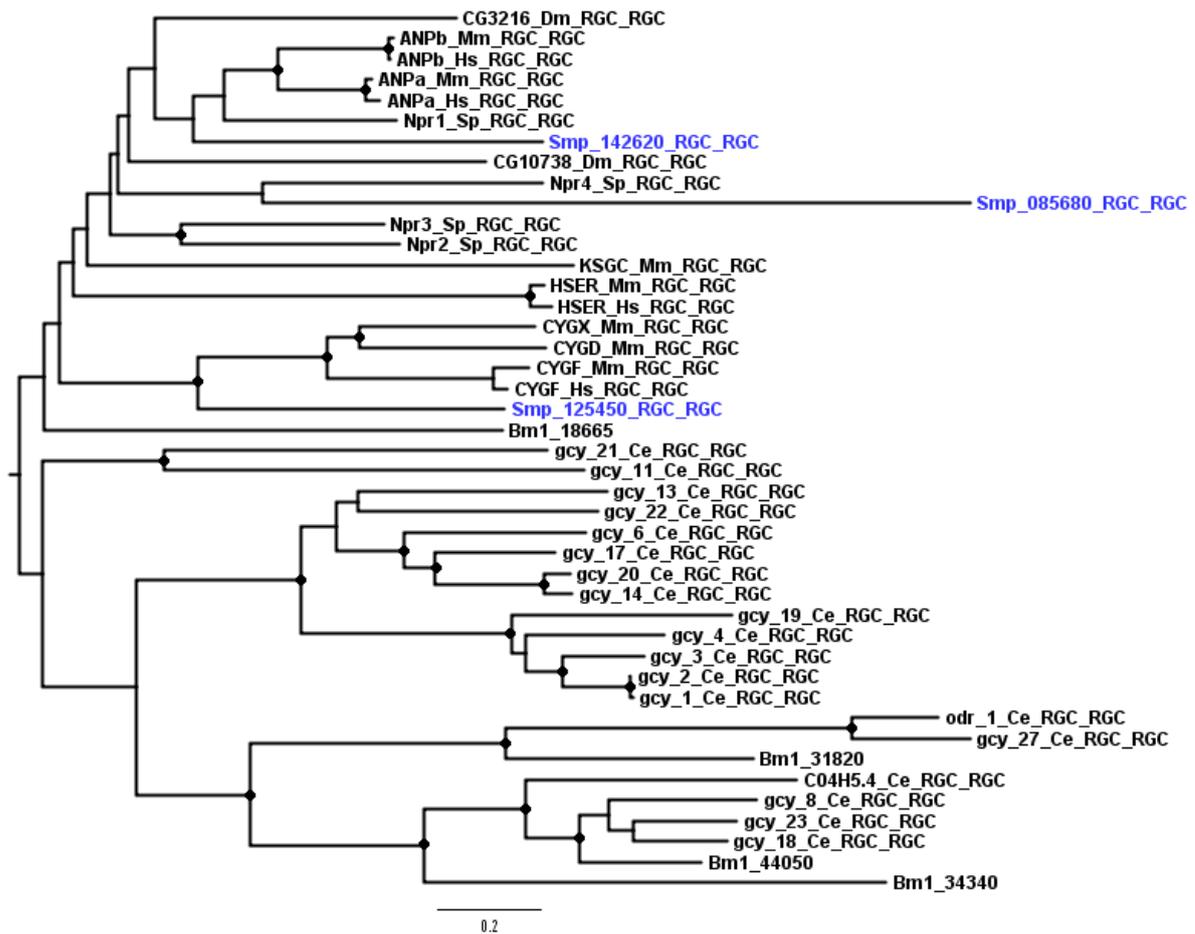
**Figura Anexo 4: Árvore filogenética baseada no método de distância do grupo STE.** As ePKs de *S. mansoni* estão indicadas pelos labels na cor azul. Valores de Bootstrap (1000pseudo réplicas) iguais ou maiores que 80% (●) foram considerados para transferir a anotação de famílias/subfamílias (indicado individualmente em cada label). As famílias identificadas estão destacadas pelas barras verticais.



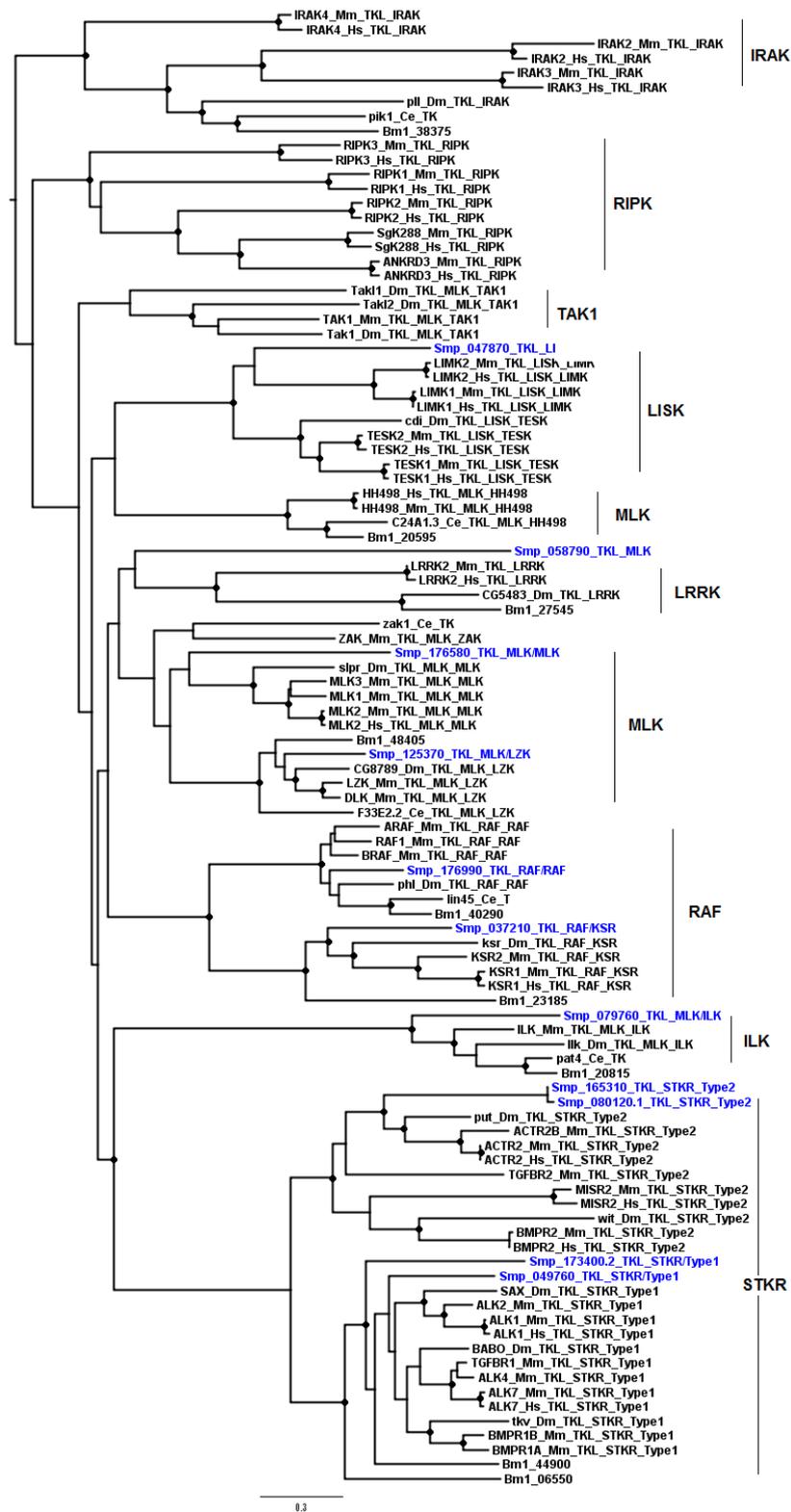
**Figura Anexo 5: Árvore filogenética baseada no método de distância do grupo CK1.** As ePKs de *S. mansoni* estão indicadas pelos labels na cor azul. Valores de Bootstrap (1000pseudo réplicas) iguais ou maiores que 80% (●) foram considerados para transferir a anotação de famílias/subfamílias (indicado individualmente em cada label). As famílias identificadas estão destacadas pelas barras verticais.



**Figura Anexo 6: Árvore filogenética baseada no método de distância do grupo TK.** As ePKs de *S. mansoni* estão indicadas pelos labels na cor azul. Valores de Bootstrap (1000pseudo réplicas) iguais ou maiores que 80% (●) foram considerados para transferir a anotação de famílias/subfamílias (indicado individualmente em cada label). As famílias identificadas estão destacadas pelas barras verticais.



**Figura Anexo 7: Árvore filogenética baseada no método de distância do grupo RGC.** As ePKs de *S. mansoni* estão indicadas pelos labels na cor azul. Valores de Bootstrap (1000 pseudo réplicas) iguais ou maiores que 80% (●) foram considerados para transferir a anotação de famílias/subfamílias (indicado individualmente em cada label). As famílias identificadas estão destacadas pelas barras verticais.



**Figura Anexo 8: Árvore filogenética baseada no método de distância do grupo TKL.** As ePKs de *S. mansoni* estão indicadas pelos labels na cor azul. Valores de Bootstrap (1000 pseudo réplicas) iguais ou maiores que 80% (●) foram considerados para transferir a anotação de famílias/subfamílias (indicado individualmente em cada label). As famílias identificadas estão destacadas pelas barras verticais.

### **7.3- Artigo Publicado em 2009 – NATURE**

Berriman, M., Haas, B. J., Loverde, P. T., Wilson, R. A., Dillon, G. P., Cerqueira, G. C., Mashiyama, S. T., Al-Lazikani, B., **ANDRADE, L. F.**, Ashton, P. D., Aslett, M. A., Bartholomeu, D. C., Blandin, G., Coffrey, C. R., Coghlan, A., Coulson, R., Day, T. A., Delcher, A., Demarco, R., Djikeng, A., Eyre, T., Gamble, J. A., Ghedin, E., Gu, Y., Hertz-Fowler, C., Hirai, H., Hirai, Y., Houston, R., Ivens, A., Johnston, D. A., Lacerda, D., Macedo, C. D., Mcveigh, P., Ning, Z., Oliveira, Guilherme, Overington, J. P., Parkhill, J., Perlea, M., Pierce, R. J., Protasio, A. V., Quail, M. A., Rajandream, M., Rogers, J., Sajid, M., Salzberg, S. L., Stanke, M., Tivey, A. R., White, O., Willians, D. L., Wortman, J., Wu, W., Zamanian, M., Zerlotini, A., Fraser-Liggett, C. M., Barrell, B. G., El-Sayed, N. M. The Genome of the Blood Fluke *Schistosoma mansoni*. Nature (London). 2009, v.460, p.352/08160 – 358.

### **7.4- Revisão Publicada em 2009 - CURRENT OPINION IN MICROBIOLOGY**

LoVerde, Philip T, **ANDRADE L.F**, OLIVEIRA, Guilherme. Signal transduction regulates schistosome reproductive biology. Current Opinion in Microbiology. 2009, v.12, p.422 - 428.

### **7.5- Artigo Publicado em 2011 - BMC GENOMICS**

**ANDRADE LF**, Nahum LA, Avelar LG, Silva LL, Zerlotini A, Ruiz JC, Oliveira G. Eukaryotic protein kinases (ePKs) of the helminth parasite *Schistosoma mansoni*. BMC Genomics. 2011 May 6;12:215.

### **7.6- Revisão Publicada em 2011 – JOURNAL OF SIGNAL TRANSDUCTION**

Avelar LG, Nahum LA, **ANDRADE LF**, Oliveira G. **Functional Diversity of the *Schistosoma mansoni* Tyrosine Kinases.** J Signal Transduct. 2011;2011:603290

### **7.7- Artigo Publicado em 2012 – CURRENT PHARMACEUTICAL DESIGN**

Pierce RJ, Dubois-Abdeselem F, Lancelot J, **ANDRADE L**, Oliveira G. Targeting Schistosome Histone Modifying Enzymes for Drug Development. Curr Pharm Des. 2012 May 11.

## 8. REFERÊNCIAS BIBLIOGRÁFICAS

- (1) Engels D, Chitsulo L, Montresor A, Savioli L. The global epidemiological situation of schistosomiasis and new approaches to control and research. *Acta Trop* 2002; 82(2):139-146.
- (2) Han ZG, Brindley PJ, Wang SY, Chen Z. Schistosoma genomics: new perspectives on schistosome biology and host-parasite interaction. *Annu Rev Genomics Hum Genet* 2009; 10:211-240.
- (3) Pereira WR, Kloos H, Crawford SB, Velasquez-Melendez JG, Matoso LF, Fujiwara RT et al. Schistosoma mansoni infection in a rural area of the Jequitinhonha Valley, Minas Gerais, Brazil: analysis of exposure risk. *Acta Trop* 2010; 113(1):34-41.
- (4) Sturrock RF. Schistosomiasis epidemiology and control: how did we get here and where should we go? *Mem Inst Oswaldo Cruz* 2001; 96 Suppl:17-27.
- (5) Gryseels B, Polman K, Clerinx J, Kestens L. Human schistosomiasis. *Lancet* 2006; 368(9541):1106-1118.
- (6) Rey L. Parasitologia: parasitas e doenças parasitárias do homem nas Américas e na África. 2nd ed. Rio de Janeiro: 1991.
- (7) TDR. WHO. <http://www.who.int/tdr/diseases/schisto/diseaseinfo.htm> . 2005.

Ref Type: Internet Communication

- (8) Gryseels B, Polman K, Clerinx J, Kestens L. Human schistosomiasis. *Lancet* 2006; 368(9541):1106-1118.
- (9) Gryseels B, Polman K, Clerinx J, Kestens L. Human schistosomiasis. *Lancet* 2006; 368(9541):1106-1118.
- (10) Gryseels B, Polman K, Clerinx J, Kestens L. Human schistosomiasis. *Lancet* 2006; 368(9541):1106-1118.
- (11) Katz N, Peixoto SV. [Critical analysis of the estimated number of Schistosomiasis mansoni carriers in Brazil]. *Rev Soc Bras Med Trop* 2000; 33(3):303-308.
- (12) Mohammad Bagher Rokni. Schistosomiasis. First ed. Janeza Trdine 9, 51000 Rijeka, Croatia, 2012.
- (13) Despres L, Imbert-Establet D, Monnerot M. Molecular characterization of mitochondrial DNA provides evidence for the recent introduction of Schistosoma mansoni into America. *Mol Biochem Parasitol* 1993; 60(2):221-229.
- (14) Davis AH. Schistosomiasis. *Epidemiology and Community Control of Disease in Warm Climate Countries*. London: Churchill Livingstone, 1985: 389-412.
- (15) Rollinson D, Simpson AJ. *The Biology of Schistosomes - From genes to Latrines*. London: Harcourt Brace Jovanovich, 1987.

- (16) Cheever AW, Macedonia JG, Mosimann JE, Cheever EA. Kinetics of egg production and egg excretion by *Schistosoma mansoni* and *S. japonicum* in mice infected with a single pair of worms. *Am J Trop Med Hyg* 1994; 50(3):281-295.
- (17) Mohammad Bagher Rokni. *Schistosomiasis*. First ed. Janeza Trdine 9, 51000 Rijeka, Croatia, 2012.
- (18) Katz N, Faria I, Reis FA. *Modernos Conhecimentos sobre Esquistossomose Mansônica*. Academia Mineira de Medicina 1986.
- (19) CDC - Centers for Disease Control and Prevention.  
<http://www.cdc.gov/parasites/schistosomiasis/biology.html> . 2010.

Ref Type: Electronic Citation

- (20) Engels D, Chitsulo L, Montresor A, Savioli L. The global epidemiological situation of schistosomiasis and new approaches to control and research. *Acta Trop* 2002; 82(2):139-146.
- (21) Wilson RA, Coulson PS. Strategies for a schistosome vaccine: can we manipulate the immune response effectively? *Microbes Infect* 1999; 1(7):535-543.
- (22) Correa-Oliveira R, Pearce EJ, Oliveira GC, Golgher DB, Katz N, Bahia LG et al. The human immune response to defined immunogens of *Schistosoma mansoni*: elevated antibody levels to paramyosin in stool-negative individuals from two endemic areas in Brazil. *Trans R Soc Trop Med Hyg* 1989; 83(6):798-804.
- (23) Gazzinelli G, Viana IR, Bahia-Oliveira LM, Silveira AM, Queiroz CC, Carvalho OS et al. Immunological profiles of patients from endemic areas infected with *Schistosoma mansoni*. *Mem Inst Oswaldo Cruz* 1992; 87 Suppl 4:139-142.
- (24) Wilson RA, Coulson PS. Immune effector mechanisms against schistosomiasis: looking for a chink in the parasite's armour. *Trends Parasitol* 2009; 25(9):423-431.
- (25) Wilson RA, Langermans JA, van Dam GJ, Vervenne RA, Hall SL, Borges WC et al. Elimination of *Schistosoma mansoni* Adult Worms by Rhesus Macaques: Basis for a Therapeutic Vaccine? *PLoS Negl Trop Dis* 2008; 2(9):e290.
- (26) Dalton JP, Mulcahy G. Parasite vaccines--a reality? *Vet Parasitol* 2001; 98(1-3):149-167.
- (27) Johnson KS, Harrison GB, Lightowers MW, O'Hoy KL, Cogle WG, Dempster RP et al. Vaccination against ovine cysticercosis using a defined recombinant antigen. *Nature* 1989; 338(6216):585-587.
- (28) McManus DP, Loukas A. Current status of vaccines for schistosomiasis. *Clinical Microbiology Reviews* 2008; 21(1):225-+.
- (29) Oliveira S.C, Fonseca C.T, Cardoso F.C, Farias L.P, Leite L.C. Recent advances in vaccine research against schistosomiasis in Brazil. *Acta Tropica* 2008; 108:256-262.

- (30) Doenhoff MJ, Pica-Mattoccia L. Praziquantel for the treatment of schistosomiasis: its use for control in areas with endemic disease and prospects for drug resistance. *Expert Rev Anti Infect Ther* 2006; 4(2):199-210.
- (31) Caffrey CR. Chemotherapy of schistosomiasis: present and future. *Curr Opin Chem Biol* 2007; 11(4):433-439.
- (32) Gryseels B, Polman K, Clerinx J, Kestens L. Human schistosomiasis. *Lancet* 2006; 368(9541):1106-1118.
- (33) Wang W, Wang L, Liang YS. Susceptibility or resistance of praziquantel in human schistosomiasis: a review. *Parasitology Research* 2012; 111(5):1871-1877.
- (34) Short RB, MENZEL MY. Chromosomes of nine species of schistosomes. *J Parasitol* 1960; 46:273-287.
- (35) Short RB, MENZEL MY, Pathak S. Somatic chromosomes of *Schistosoma mansoni*. *J Parasitol* 1979; 65(3):471-473.
- (36) Berriman M, Haas BJ, LoVerde PT, Wilson RA, Dillon GP, Cerqueira GC et al. The genome of the blood fluke *Schistosoma mansoni*. *Nature* 2009; 460(7253):352-358.
- (37) Grossman AI, Short RB, Kuntz RE. Somatic chromosomes of *Schistosoma rodhaini*, *S. mattheei*, and *S. intercalatum*. *J Parasitol* 1981; 67(1):41-44.
- (38) Short RB, MENZEL MY. Chromosomes of nine species of schistosomes. *J Parasitol* 1960; 46:273-287.
- (39) Lepesant JMJ, Roquis D, Emans R, Cosseau C, Arancibia N, Mitta G et al. Combination of de novo assembly of massive sequencing reads with classical repeat prediction improves identification of repetitive sequences in *Schistosoma mansoni*. *Experimental Parasitology* 2012; 130(4):470-474.
- (40) The *Schistosoma japonicum* Genome and Functional Analysis Consortium. The *Schistosoma japonicum* genome reveals features of host-parasite interplay. *Nature* 2009; 460(7253):345-351.
- (41) Berriman M, Haas BJ, LoVerde PT, Wilson RA, Dillon GP, Cerqueira GC et al. The genome of the blood fluke *Schistosoma mansoni*. *Nature* 2009; 460(7253):352-358.
- (42) Young ND, Jex AR, Li B, Liu SP, Yang LF, Xiong ZJ et al. Whole-genome sequence of *Schistosoma haematobium*. *Nature Genetics* 2012; 44(2):221-225.
- (43) Young ND, Jex AR, Li B, Liu SP, Yang LF, Xiong ZJ et al. Whole-genome sequence of *Schistosoma haematobium*. *Nature Genetics* 2012; 44(2):221-225.
- (44) Protasio AV, Tsai IJ, Babbage A, Nichol S, Hunt M, Aslett MA et al. A Systematically Improved High Quality Genome and Transcriptome of the Human Blood Fluke *Schistosoma mansoni*. *Plos Neglected Tropical Diseases* 2012; 6(1).

- (45) Marioni JC, Mason CE, Mane SM, Stephens M, Gilad Y. RNA-seq: An assessment of technical reproducibility and comparison with gene expression arrays. *Genome Research* 2008; 18(9):1509-1517.
- (46) Shendure J. The beginning of the end for microarrays? *Nature Methods* 2008; 5(7):585-587.
- (47) Berriman M, Haas BJ, LoVerde PT, Wilson RA, Dillon GP, Cerqueira GC et al. The genome of the blood fluke *Schistosoma mansoni*. *Nature* 2009; 460(7253):352-358.
- (48) Protasio AV, Tsai IJ, Babbage A, Nichol S, Hunt M, Aslett MA et al. A Systematically Improved High Quality Genome and Transcriptome of the Human Blood Fluke *Schistosoma mansoni*. *Plos Neglected Tropical Diseases* 2012; 6(1).
- (49) Young ND, Jex AR, Li B, Liu SP, Yang LF, Xiong ZJ et al. Whole-genome sequence of *Schistosoma haematobium*. *Nature Genetics* 2012; 44(2):221-225.
- (50) The *Schistosoma japonicum* Genome and Functional Analysis Consortium. The *Schistosoma japonicum* genome reveals features of host-parasite interplay. *Nature* 2009; 460(7253):345-351.
- (51) Young ND, Jex AR, Li B, Liu SP, Yang LF, Xiong ZJ et al. Whole-genome sequence of *Schistosoma haematobium*. *Nature Genetics* 2012; 44(2):221-225.
- (52) Young ND, Jex AR, Li B, Liu SP, Yang LF, Xiong ZJ et al. Whole-genome sequence of *Schistosoma haematobium*. *Nature Genetics* 2012; 44(2):221-225.
- (53) Zerlotini A, Heiges M, Wang H, Moraes RL, Dominitini AJ, Ruiz JC et al. SchistoDB: a *Schistosoma mansoni* genome resource. *Nucleic Acids Res* 2009; 37(Database issue):D579-D582.
- (54) Galperin MY, Koonin EV. Who's your neighbor? New computational approaches for functional genomics. *Nat Biotechnol* 2000; 18(6):609-613.
- (55) Gabaldon T, Huynen MA. Prediction of protein function and pathways in the genome era. *Cell Mol Life Sci* 2004; 61(7-8):930-944.
- (56) Galperin MY, Koonin EV. Who's your neighbor? New computational approaches for functional genomics. *Nat Biotechnol* 2000; 18(6):609-613.
- (57) Galperin MY, Koonin EV. Who's your neighbor? New computational approaches for functional genomics. *Nat Biotechnol* 2000; 18(6):609-613.
- (58) Gabaldon T, Huynen MA. Prediction of protein function and pathways in the genome era. *Cell Mol Life Sci* 2004; 61(7-8):930-944.
- (59) Nahum LA, Riley M. Divergence of function in sequence-related groups of *Escherichia coli* proteins. *Genome Res* 2001; 11(8):1375-1381.

- (60) Serres MH, Riley M. Gene fusions and gene duplications: relevance to genomic annotation and functional analysis. *BMC Genomics* 2005; 6(1):33.
- (61) Huynen M, Snel B, Lathe W, III, Bork P. Predicting protein function by genomic context: quantitative evaluation and qualitative inferences. *Genome Res* 2000; 10(8):1204-1210.
- (62) Marcotte EM, Pellegrini M, Ng HL, Rice DW, Yeates TO, Eisenberg D. Detecting protein function and protein-protein interactions from genome sequences. *Science* 1999; 285(5428):751-753.
- (63) Eisen JA. Phylogenomics: improving functional predictions for uncharacterized genes by evolutionary analysis. *Genome Res* 1998; 8(3):163-167.
- (64) Brown D, Sjolander K. Functional classification using phylogenomic inference. *PLoS Comput Biol* 2006; 2(6):e77.
- (65) KEGG: Kyoto Encyclopedia of Genes and Genomes. KEGG: Kyoto Encyclopedia of Genes and Genomes. <http://www.genome.jp/kegg/> . 5-17-2000. 10-6-2009.

Ref Type: Electronic Citation

- (66) PSI-BLAST. PSI-BLAST. <http://www.ncbi.nlm.nih.gov/Education/BLASTinfo/psi1.html> . 2000. 10-9-2008.

Ref Type: Electronic Citation

- (67) Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *J Mol Biol* 1990; 215(3):403-410.
- (68) Felsenstein J. Mathematics vs. Evolution: Mathematical Evolutionary Theory. *Science* 1989; 246(4932):941-942.
- (69) Finn RD, Mistry J, Tate J, Coggill P, Heger A, Pollington JE et al. The Pfam protein families database. *Nucleic Acids Res* 2010; 38(Database issue):D211-D222.
- (70) Higgins DG, Thompson JD, Gibson TJ. Using CLUSTAL for multiple sequence alignments. *Methods Enzymol* 1996; 266:383-402.
- (71) Hunter S, Apweiler R, Attwood TK, Bairoch A, Bateman A, Binns D et al. InterPro: the integrative protein signature database. *Nucleic Acids Res* 2009; 37(Database issue):D211-D215.
- (72) Katoh K, Asimenos G, Toh H. Multiple alignment of DNA sequences with MAFFT. *Methods Mol Biol* 2009; 537:39-64.
- (73) Quevillon E, Silventoinen V, Pillai S, Harte N, Mulder N, Apweiler R et al. InterProScan: protein domains identifier. *Nucleic Acids Res* 2005; 33(Web Server issue):W116-W120.

- (74) Rambaut A. Tree Figure Drawing Tool, Version 1.3.1.  
<http://tree.bio.ed.ac.uk/software/figtree> . 2010. 3-1-2010.

Ref Type: Electronic Citation

- (75) Rutherford K, Parkhill J, Crook J, Horsnell T, Rice P, Rajandream MA et al. Artemis: sequence visualization and annotation. *Bioinformatics* 2000; 16(10):944-945.
- (76) Delsuc F, Brinkmann H, Philippe H. Phylogenomics and the reconstruction of the tree of life. *Nat Rev Genet* 2005; 6(5):361-375.
- (77) Page R, Holmes E. *Molecular Evolution: A phylogenetic approach*. Blackwell Science, 1998.
- (78) Ridley M. *Evolução*. 3ed. ed. Porto Alegre: Artmed, 2006.
- (79) Nahum LA, Pereira SL. *Phylogenomics, Protein Family Evolution, and the Tree of life: An Integrated Approach between Molecular Evolution and Computational Intelligence*. Studies in Computational Intelligence. Springer-Verlag Berlin Heidelberg, 2008: 259-279.
- (80) Schneider H. *Métodos de Análise Filogenética*. Ribeirão Preto: Holod Editora, 2007.
- (81) Saitou N, Nei M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* 1987; 4(4):406-425.
- (82) Studier JA, Keppler KJ. A note on the neighbor-joining algorithm of Saitou and Nei. *Mol Biol Evol* 1988; 5(6):729-731.
- (83) Holder M, Lewis PO. Phylogeny estimation: traditional and Bayesian approaches. *Nat Rev Genet* 2003; 4(4):275-284.
- (84) Holder M, Lewis PO. Phylogeny estimation: traditional and Bayesian approaches. *Nat Rev Genet* 2003; 4(4):275-284.
- (85) Brown D, Sjolander K. Functional classification using phylogenomic inference. *PLoS Comput Biol* 2006; 2(6):e77.
- (86) Wilson RA, Ashton PD, Braschi S, Dillon GP, Berriman M, Ivens A. 'Oming in on schistosomes: prospects and limitations for post-genomics. *Trends Parasitol* 2007; 23(1):14-20.
- (87) Dillon GP, Feltwell T, Skelton JP, Ashton PD, Coulson PS, Quail MA et al. Microarray analysis identifies genes preferentially expressed in the lung schistosomulum of *Schistosoma mansoni*. *Int J Parasitol* 2006; 36(1):1-8.
- (88) Fitzpatrick JM, Johnston DA, Williams GW, Williams DJ, Freeman TC, Dunne DW et al. An oligonucleotide microarray for transcriptome analysis of *Schistosoma mansoni* and its application/use to investigate gender-associated gene expression. *Mol Biochem Parasitol* 2005; 141(1):1-13.

- (89) Protasio AV, Tsai IJ, Babbage A, Nichol S, Hunt M, Aslett MA et al. A Systematically Improved High Quality Genome and Transcriptome of the Human Blood Fluke *Schistosoma mansoni*. *Plos Neglected Tropical Diseases* 2012; 6(1).
- (90) Verjovski-Almeida S, DeMarco R, Martins EA, Guimaraes PE, Ojopi EP, Paquola AC et al. Transcriptome analysis of the acoelomate human parasite *Schistosoma mansoni*. *Nat Genet* 2003; 35(2):148-157.
- (91) Young ND, Jex AR, Li B, Liu SP, Yang LF, Xiong ZJ et al. Whole-genome sequence of *Schistosoma haematobium*. *Nature Genetics* 2012; 44(2):221-225.
- (92) Curwen RS, Ashton PD, Johnston DA, Wilson RA. The *Schistosoma mansoni* soluble proteome: a comparison across four life-cycle stages. *Mol Biochem Parasitol* 2004; 138(1):57-66.
- (93) Braschi S, Borges WC, Wilson RA. Proteomic analysis of the schistosome tegument and its surface membranes. *Mem Inst Oswaldo Cruz* 2006; 101 Suppl 1:205-212.
- (94) van Balkom BW, van Gestel RA, Brouwers JF, Krijgsveld J, Tielens AG, Heck AJ et al. Mass spectrometric analysis of the *Schistosoma mansoni* tegumental sub-proteome. *J Proteome Res* 2005; 4(3):958-966.
- (95) Wilson RA, Curwen RS, Braschi S, Hall SL, Coulson PS, Ashton PD. From genomes to vaccines via the proteome. *Mem Inst Oswaldo Cruz* 2004; 99(5 Suppl 1):45-50.
- (96) Kemphues K. Essential genes. *WormBook* 2005;1-7.
- (97) Wilson RA, Ashton PD, Braschi S, Dillon GP, Berriman M, Ivens A. 'Oming in on schistosomes: prospects and limitations for post-genomics. *Trends Parasitol* 2007; 23(1):14-20.
- (98) Fire A, Xu S, Montgomery MK, Kostas SA, Driver SE, Mello CC. Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* 1998; 391(6669):806-811.
- (99) McCaffrey AP, Meuse L, Pham TT, Conklin DS, Hannon GJ, Kay MA. RNA interference in adult mice. *Nature* 2002; 418(6893):38-39.
- (100) Ameres SL, Martinez J, Schroeder R. Molecular basis for target RNA recognition and cleavage by human RISC. *Cell* 2007; 130(1):101-112.
- (101) Vidalin O, Muslmani M, Estienne C, Echchakir H, Abina AM. In vivo target validation using gene invalidation, RNA interference and protein functional knockout models: it is the time to combine. *Curr Opin Pharmacol* 2009; 9(5):669-676.
- (102) Zamore PD, Tuschl T, Sharp PA, Bartel DP. RNAi: double-stranded RNA directs the ATP-dependent cleavage of mRNA at 21 to 23 nucleotide intervals. *Cell* 2000; 101(1):25-33.

- (103) Murchison EP, Hannon GJ. miRNAs on the move: miRNA biogenesis and the RNAi machinery. *Curr Opin Cell Biol* 2004; 16(3):223-229.
- (104) Hannon GJ, Rossi J. <http://www.ncbi.nlm.nih.gov/pubmed/15372045>. *Nature* 2004; 431(7006):371-378.
- (105) Lenz G. The RNA interference revolution. *Braz J Med Biol Res* 2005; 38(12):1749-1757.
- (106) Kurreck J. RNA interference: from basic research to therapeutic applications. *Angew Chem Int Ed Engl* 2009; 48(8):1378-1398.
- (107) Boyle JP, Wu XJ, Shoemaker CB, Yoshino TP. Using RNA interference to manipulate endogenous gene expression in *Schistosoma mansoni* sporocysts. *Mol Biochem Parasitol* 2003; 128(2):205-215.
- (108) Skelly PJ, Da'dara A, Harn DA. Suppression of cathepsin B expression in *Schistosoma mansoni* by RNA interference. *Int J Parasitol* 2003; 33(4):363-369.
- (109) Correnti JM, Brindley PJ, Pearce EJ. Long-term suppression of cathepsin B levels by RNA interference retards schistosome growth. *Mol Biochem Parasitol* 2005; 143(2):209-215.
- (110) Delcroix M, Sajid M, Caffrey CR, Lim KC, Dvorak J, Hsieh I et al. A multienzyme network functions in intestinal protein digestion by a platyhelminth parasite. *J Biol Chem* 2006; 281(51):39316-39329.
- (111) Morales ME, Rinaldi G, Gobert GN, Kines KJ, Tort JF, Brindley PJ. RNA interference of *Schistosoma mansoni* cathepsin D, the apical enzyme of the hemoglobin proteolysis cascade. *Mol Biochem Parasitol* 2008; 157(2):160-168.
- (112) Mourao MM, Dinguirard N, Franco GR, Yoshino TP. Phenotypic screen of early-developing larvae of the blood fluke, *Schistosoma mansoni*, using RNA interference. *PLoS Negl Trop Dis* 2009; 3(8):e502.
- (113) Rinaldi G, Morales ME, Alrefaei YN, Cancela M, Castillo E, Dalton JP et al. RNA interference targeting leucine aminopeptidase blocks hatching of *Schistosoma mansoni* eggs. *Mol Biochem Parasitol* 2009; 167(2):118-126.
- (114) Stefanic S, Dvorak J, Horn M, Braschi S, Sojka D, Ruelas DS et al. RNA interference in *Schistosoma mansoni* schistosomula: selectivity, sensitivity and operation for larger-scale screening. *PLoS Negl Trop Dis* 2010; 4(10):e850.
- (115) Swierczewski BE, Davies SJ. A schistosome cAMP-dependent protein kinase catalytic subunit is essential for parasite viability. *PLoS Negl Trop Dis* 2009; 3(8):e505.
- (116) Osman A, Niles EG, Verjovski-Almeida S, LoVerde PT. *Schistosoma mansoni* TGF-beta receptor II: role in host ligand-induced regulation of a schistosome target gene. *PLoS Pathog* 2006; 2(6):e54.

- (117) Mourao MM, Dinguirard N, Franco GR, Yoshino TP. Role of the endogenous antioxidant system in the protection of *Schistosoma mansoni* primary sporocysts against exogenous oxidative stress. *PLoS Negl Trop Dis* 2009; 3(11):e550.
- (118) Faghiri Z, Skelly PJ. The role of tegumental aquaporin from the human parasitic worm, *Schistosoma mansoni*, in osmoregulation and drug uptake. *FASEB J* 2009; 23(8):2780-2789.
- (119) Krautz-Peterson G, Simoes M, Faghiri Z, Ndegwa D, Oliveira G, Shoemaker CB et al. Suppressing glucose transporter gene expression in schistosomes impairs parasite feeding and decreases survival in the mammalian host. *PLoS Pathog* 2010; 6(6):e1000932.
- (120) Tran MH, Freitas TC, Cooper L, Gaze S, Gatton ML, Jones MK et al. Suppression of mRNAs encoding tegument tetraspanins from *Schistosoma mansoni* results in impaired tegument turnover. *PLoS Pathog* 2010; 6(4):e1000840.
- (121) Gomes MS, Cabral FJ, Jannotti-Passos LK, Carvalho O, Rodrigues V, Baba EH et al. Preliminary analysis of miRNA pathway in *Schistosoma mansoni*. *Parasitol Int* 2009; 58(1):61-68.
- (122) Krautz-Peterson G, Skelly PJ. *Schistosoma mansoni*: the dicer gene and its expression. *Exp Parasitol* 2008; 118(1):122-128.
- (123) Krautz-Peterson G, Bhardwaj R, Faghiri Z, Tararam CA, Skelly PJ. RNA interference in schistosomes: machinery and methodology. *Parasitology* 2010; 137(3):485-495.
- (124) Simoes MC, Lee J, Djikeng A, Cerqueira GC, Zerlotini A, da Silva-Pereira RA et al. Identification of *Schistosoma mansoni* microRNAs. *BMC Genomics* 2011; 12(1):47.
- (125) LoVerde PT, Hirai H, Merrick JM, Lee NH, El-Sayed N. *Schistosoma mansoni* genome project: an update. *Parasitol Int* 2004; 53(2):183-192.
- (126) Amiri P, Locksley RM, Parslow TG, Sadick M, Rector E, Ritter D et al. Tumour necrosis factor alpha restores granulomas and induces parasite egg-laying in schistosome-infected SCID mice. *Nature* 1992; 356(6370):604-607.
- (127) Wolowczuk I, Nutten S, Roye O, Delacre M, Capron M, Murray RM et al. Infection of mice lacking interleukin-7 (IL-7) reveals an unexpected role for IL-7 in the development of the parasite *Schistosoma mansoni*. *Infect Immun* 1999; 67(8):4183-4190.
- (128) Davies SJ, Shoemaker CB, Pearce EJ. A divergent member of the transforming growth factor beta receptor family from *Schistosoma mansoni* is expressed on the parasite surface membrane. *J Biol Chem* 1998; 273(18):11234-11240.
- (129) Vicogne J, Cailliau K, Tulasne D, Browaeys E, Yan YT, Fafeur V et al. Conservation of epidermal growth factor receptor function in the human parasitic helminth *Schistosoma mansoni*. *J Biol Chem* 2004; 279(36):37407-37414.

- (130) Davies SJ, Grogan JL, Blank RB, Lim KC, Locksley RM, McKerrow JH. Modulation of blood fluke development in the liver by hepatic CD4+ lymphocytes. *Science* 2001; 294(5545):1358-1361.
- (131) Davies SJ, Grogan JL, Blank RB, Lim KC, Locksley RM, McKerrow JH. Modulation of blood fluke development in the liver by hepatic CD4+ lymphocytes. *Science* 2001; 294(5545):1358-1361.
- (132) Amiri P, Locksley RM, Parslow TG, Sadick M, Rector E, Ritter D et al. Tumour necrosis factor alpha restores granulomas and induces parasite egg-laying in schistosome-infected SCID mice. *Nature* 1992; 356(6370):604-607.
- (133) LoVerde PT, Andrade LF, Oliveira G. Signal transduction regulates schistosome reproductive biology. *Curr Opin Microbiol* 2009; 12(4):422-428.
- (134) Kunz W. Schistosome male-female interaction: induction of germ-cell differentiation. *Trends Parasitol* 2001; 17(5):227-231.
- (135) LoVerde PT, Hirai H, Merrick JM, Lee NH, El-Sayed N. *Schistosoma mansoni* genome project: an update. *Parasitol Int* 2004; 53(2):183-192.
- (136) Fitzpatrick JM, Johnston DA, Williams GW, Williams DJ, Freeman TC, Dunne DW et al. An oligonucleotide microarray for transcriptome analysis of *Schistosoma mansoni* and its application/use to investigate gender-associated gene expression. *Mol Biochem Parasitol* 2005; 141(1):1-13.
- (137) Waisberg M, Lobo FP, Cerqueira GC, Passos LK, Carvalho OS, Franco GR et al. Microarray analysis of gene expression induced by sexual contact in *Schistosoma mansoni*. *BMC Genomics* 2007; 8:181.
- (138) Waisberg M, Lobo FP, Cerqueira GC, Passos LK, Carvalho OS, El-Sayed NM et al. *Schistosoma mansoni*: Microarray analysis of gene expression induced by host sex. *Exp Parasitol* 2008; 120(4):357-363.
- (139) LoVerde PT, Hirai H, Merrick JM, Lee NH, El-Sayed N. *Schistosoma mansoni* genome project: an update. *Parasitol Int* 2004; 53(2):183-192.
- (140) Hanks SK, Quinn AM, Hunter T. The protein kinase family: conserved features and deduced phylogeny of the catalytic domains. *Science* 1988; 241(4861):42-52.
- (141) Plowman GD, Sudarsanam S, Bingham J, Whyte D, Hunter T. The protein kinases of *Caenorhabditis elegans*: a model for signal transduction in multicellular organisms. *Proc Natl Acad Sci U S A* 1999; 96(24):13603-13610.
- (142) Manning G, Plowman GD, Hunter T, Sudarsanam S. Evolution of protein kinase signaling from yeast to man. *Trends Biochem Sci* 2002; 27(10):514-520.
- (143) Hanks SK. Genomic analysis of the eukaryotic protein kinase superfamily: a perspective. *Genome Biol* 2003; 4(5):111.

- (144) Hanks SK, Hunter T. Protein kinases 6. The eukaryotic protein kinase superfamily: kinase (catalytic) domain structure and classification. *FASEB J* 1995; 9(8):576-596.
- (145) Hanks SK, Quinn AM, Hunter T. The protein kinase family: conserved features and deduced phylogeny of the catalytic domains. *Science* 1988; 241(4861):42-52.
- (146) Hanks SK, Hunter T. Protein kinases 6. The eukaryotic protein kinase superfamily: kinase (catalytic) domain structure and classification. *FASEB J* 1995; 9(8):576-596.
- (147) Manning G, Whyte DB, Martinez R, Hunter T, Sudarsanam S. The protein kinase complement of the human genome. *Science* 2002; 298(5600):1912-1934.
- (148) Miranda-Saavedra D, Barton GJ. Classification and functional annotation of eukaryotic protein kinases. *Proteins* 2007; 68(4):893-914.
- (149) KinBase. Kinbase - The kinase database at Sugen/Salk. <http://www.kinase.com/kinbase> . 2010. 8-8-2009.

Ref Type: Electronic Citation

- (150) Miranda-Saavedra D, Barton GJ. Classification and functional annotation of eukaryotic protein kinases. *Proteins* 2007; 68(4):893-914.
- (151) Boyle SN, Koleske AJ. Dissecting kinase signaling pathways. *Drug Discov Today* 2007; 12(17-18):717-724.
- (152) Eglen RM, Reisine T. The current status of drug discovery against the human kinome. *Assay Drug Dev Technol* 2009; 7(1):22-43.
- (153) Dissous C, Ahier A, Khayath N. Protein tyrosine kinases as new potential targets against human schistosomiasis. *Bioessays* 2007; 29(12):1281-1288.
- (154) Naula C, Parsons M, Mottram JC. Protein kinases as drug targets in trypanosomes and *Leishmania*. *Biochim Biophys Acta* 2005; 1754(1-2):151-159.
- (155) Ward P, Equinet L, Packer J, Doerig C. Protein kinases of the human malaria parasite *Plasmodium falciparum*: the kinome of a divergent eukaryote. *BMC Genomics* 2004; 5(1):79.
- (156) Beckmann S, Grevelding CG. Imatinib has a fatal impact on morphology, pairing stability and survival of adult *Schistosoma mansoni* in vitro. *Int J Parasitol* 2010; 40(5):521-526.
- (157) Andrade LF, Nahum LA, Avelar LG, Silva LL, Zerlotini A, Ruiz JC et al. Eukaryotic protein kinases (ePKs) of the helminth parasite *Schistosoma mansoni*. *BMC Genomics* 2011; 12:215.
- (158) Engels D, Chitsulo L, Montresor A, Savioli L. The global epidemiological situation of schistosomiasis and new approaches to control and research. *Acta Trop* 2002; 82(2):139-146.

- (159) Mohammad Bagher Rokni. Schistosomiasis. First ed. Janeza Trdine 9, 51000 Rijeka, Croatia, 2012.
- (160) Doenhoff MJ, Cioli D, Utzinger J. Praziquantel: mechanisms of action, resistance and new derivatives for schistosomiasis. *Curr Opin Infect Dis* 2008; 21(6):659-667.
- (161) Mohammad Bagher Rokni. Schistosomiasis. First ed. Janeza Trdine 9, 51000 Rijeka, Croatia, 2012.
- (162) Berriman M, Haas BJ, LoVerde PT, Wilson RA, Dillon GP, Cerqueira GC et al. The genome of the blood fluke *Schistosoma mansoni*. *Nature* 2009; 460(7253):352-358.
- (163) The *Schistosoma japonicum* Genome and Functional Analysis Consortium. The *Schistosoma japonicum* genome reveals features of host-parasite interplay. *Nature* 2009; 460(7253):345-351.
- (164) Young ND, Jex AR, Li B, Liu SP, Yang LF, Xiong ZJ et al. Whole-genome sequence of *Schistosoma haematobium*. *Nature Genetics* 2012; 44(2):221-225.
- (165) Hanks SK, Quinn AM, Hunter T. The protein kinase family: conserved features and deduced phylogeny of the catalytic domains. *Science* 1988; 241(4861):42-52.
- (166) Boyle SN, Koleske AJ. Dissecting kinase signaling pathways. *Drug Discov Today* 2007; 12(17-18):717-724.
- (167) Miranda-Saavedra D, Barton GJ. Classification and functional annotation of eukaryotic protein kinases. *Proteins* 2007; 68(4):893-914.
- (168) Eisen JA. Phylogenomics: improving functional predictions for uncharacterized genes by evolutionary analysis. *Genome Res* 1998; 8(3):163-167.
- (169) Boyle JP, Wu XJ, Shoemaker CB, Yoshino TP. Using RNA interference to manipulate endogenous gene expression in *Schistosoma mansoni* sporocysts. *Mol Biochem Parasitol* 2003; 128(2):205-215.
- (170) Skelly PJ, Da'dara A, Harn DA. Suppression of cathepsin B expression in *Schistosoma mansoni* by RNA interference. *Int J Parasitol* 2003; 33(4):363-369.
- (171) Zerlotini A, Heiges M, Wang H, Moraes RL, Dominitini AJ, Ruiz JC et al. SchistoDB: a *Schistosoma mansoni* genome resource. *Nucleic Acids Res* 2009; 37(Database issue):D579-D582.
- (172) KEGG: Kyoto Encyclopedia of Genes and Genomes. KEGG: Kyoto Encyclopedia of Genes and Genomes. <http://www.genome.jp/kegg/>. 5-17-2000. 10-6-2009.

Ref Type: Electronic Citation

- (173) Hall T. Bioedit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucl Acids Symp Ser* 1999; 41:95-98.

- (174) Finn RD, Mistry J, Tate J, Coggill P, Heger A, Pollington JE et al. The Pfam protein families database. *Nucleic Acids Res* 2010; 38(Database issue):D211-D222.
- (175) Deshmukh K, Anamika K, Srinivasan N. Evolution of domain combinations in protein kinases and its implications for functional diversity. *Prog Biophys Mol Biol* 2010; 102(1):1-15.
- (176) Hanks SK, Hunter T. Protein kinases 6. The eukaryotic protein kinase superfamily: kinase (catalytic) domain structure and classification. *FASEB J* 1995; 9(8):576-596.
- (177) Manning G, Whyte DB, Martinez R, Hunter T, Sudarsanam S. The protein kinase complement of the human genome. *Science* 2002; 298(5600):1912-1934.
- (178) Hunter S, Apweiler R, Attwood TK, Bairoch A, Bateman A, Binns D et al. InterPro: the integrative protein signature database. *Nucleic Acids Res* 2009; 37(Database issue):D211-D215.
- (179) Rutherford K, Parkhill J, Crook J, Horsnell T, Rice P, Rajandream MA et al. Artemis: sequence visualization and annotation. *Bioinformatics* 2000; 16(10):944-945.
- (180) Li L, Stoeckert CJ, Roos DS. OrthoMCL: Identification of ortholog groups for eukaryotic genomes. *Genome Research* 2003; 13(9):2178-2189.
- (181) Finn RD, Mistry J, Tate J, Coggill P, Heger A, Pollington JE et al. The Pfam protein families database. *Nucleic Acids Res* 2010; 38(Database issue):D211-D222.
- (182) Notredame C. Decrease redundancy tool. <http://www.expasy.ch/tools/redundancy> . 2009. 11-20-2009.

Ref Type: Electronic Citation

- (183) Felsenstein J. Mathematics vs. Evolution: Mathematical Evolutionary Theory. *Science* 1989; 246(4932):941-942.
- (184) Wilkinson MG, Millar JB. Control of the eukaryotic cell cycle by MAP kinase signaling pathways. *FASEB J* 2000; 14(14):2147-2157.
- (185) Le TH, Blair D, Agatsuma T, Humair PF, Campbell NJ, Iwagami M et al. Phylogenies inferred from mitochondrial gene orders-a cautionary tale from the parasitic flatworms. *Mol Biol Evol* 2000; 17(7):1123-1125.
- (186) Oliveira G, Busek S, Correa-Oliveira R. Transcription levels of two actin genes (SmAct and SmAct2), cytochrome C oxidase subunit II (SmCOXII), triosephosphate isomerase (TPI), and a putative translation regulatory protein EIF-5 during the first seven days of in vitro development of *Schistosoma mansoni* schistosomula. *Mem Inst Oswaldo Cruz* 1998; 93 Suppl 1:215-217.
- (187) Le TH, Blair D, Agatsuma T, Humair PF, Campbell NJ, Iwagami M et al. Phylogenies inferred from mitochondrial gene orders-a cautionary tale from the parasitic flatworms. *Mol Biol Evol* 2000; 17(7):1123-1125.

- (188) Shaner NC, Campbell RE, Steinbach PA, Giepmans BN, Palmer AE, Tsien RY. Improved monomeric red, orange and yellow fluorescent proteins derived from *Discosoma* sp. red fluorescent protein. *Nat Biotechnol* 2004; 22(12):1567-1572.
- (189) Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* 2001; 25(4):402-408.
- (190) Pellegrino J, Katz N. Experimental chemotherapy of *Schistosomiasis mansoni*. *Adv Parasitol* 1968; 6:233-290.
- (191) Anamika K, Srinivasan N. Comparative kinomics of *Plasmodium* organisms: unity in diversity. *Protein Pept Lett* 2007; 14(6):509-517.
- (192) 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.
- (193) Ghedin E, Wang S, Spiro D, Caler E, Zhao Q, Crabtree J et al. Draft genome of the filarial nematode parasite *Brugia malayi*. *Science* 2007; 317(5845):1756-1760.
- (194) Berriman M, Haas BJ, LoVerde PT, Wilson RA, Dillon GP, Cerqueira GC et al. The genome of the blood fluke *Schistosoma mansoni*. *Nature* 2009; 460(7253):352-358.
- (195) Plowman GD, Sudarsanam S, Bingham J, Whyte D, Hunter T. The protein kinases of *Caenorhabditis elegans*: a model for signal transduction in multicellular organisms. *Proc Natl Acad Sci U S A* 1999; 96(24):13603-13610.
- (196) Manning G, Whyte DB, Martinez R, Hunter T, Sudarsanam S. The protein kinase complement of the human genome. *Science* 2002; 298(5600):1912-1934.
- (197) Caenepeel S, Charydczak G, Sudarsanam S, Hunter T, Manning G. The mouse kinome: discovery and comparative genomics of all mouse protein kinases. *Proc Natl Acad Sci U S A* 2004; 101(32):11707-11712.
- (198) Zdobnov EM, von MC, Letunic I, Torrents D, Suyama M, Copley RR et al. Comparative genome and proteome analysis of *Anopheles gambiae* and *Drosophila melanogaster*. *Science* 2002; 298(5591):149-159.
- (199) Manning G, Plowman GD, Hunter T, Sudarsanam S. Evolution of protein kinase signaling from yeast to man. *Trends Biochem Sci* 2002; 27(10):514-520.
- (200) Chen F, Mackey AJ, Stoeckert CJ, Roos DS. OrthoMCL-DB: querying a comprehensive multi-species collection of ortholog groups. *Nucleic Acids Research* 2006; 34:D363-D368.
- (201) Li L, Stoeckert CJ, Roos DS. OrthoMCL: Identification of ortholog groups for eukaryotic genomes. *Genome Research* 2003; 13(9):2178-2189.

- (202) Young ND, Jex AR, Li B, Liu SP, Yang LF, Xiong ZJ et al. Whole-genome sequence of *Schistosoma haematobium*. *Nature Genetics* 2012; 44(2):221-225.
- (203) Young ND, Jex AR, Li B, Liu SP, Yang LF, Xiong ZJ et al. Whole-genome sequence of *Schistosoma haematobium*. *Nature Genetics* 2012; 44(2):221-225.
- (204) Berriman M, Haas BJ, LoVerde PT, Wilson RA, Dillon GP, Cerqueira GC et al. The genome of the blood fluke *Schistosoma mansoni*. *Nature* 2009; 460(7253):352-358.
- (205) Protasio AV, Tsai IJ, Babbage A, Nichol S, Hunt M, Aslett MA et al. A Systematically Improved High Quality Genome and Transcriptome of the Human Blood Fluke *Schistosoma mansoni*. *Plos Neglected Tropical Diseases* 2012; 6(1).
- (206) Criscione CD, Valentim CLL, Hirai H, LoVerde PT, Anderson TJC. Genomic linkage map of the human blood fluke *Schistosoma mansoni*. *Genome Biology* 2009; 10(6).
- (207) Lepesant J, Cosseau C, Boissier J, Freitag M, Portela J, Climent D et al. Chromatin structure changes around satellite repeats on the *Schistosoma mansoni* female sex chromosome suggest a possible mechanism for sex chromosome emergence. *Genome Biology* 2012; 14.
- (208) Regina Coeli Lage. **PZQ afeta de forma sistêmica a expressão de genes e proteínas de *Schistosoma mansoni***. 2011.
- (209) Felsenstein J. Mathematics vs. Evolution: Mathematical Evolutionary Theory. *Science* 1989; 246(4932):941-942.
- (210) Deshmukh K, Anamika K, Srinivasan N. Evolution of domain combinations in protein kinases and its implications for functional diversity. *Prog Biophys Mol Biol* 2010; 102(1):1-15.
- (211) Hanks SK, Hunter T. Protein kinases 6. The eukaryotic protein kinase superfamily: kinase (catalytic) domain structure and classification. *FASEB J* 1995; 9(8):576-596.
- (212) Manning G, Whyte DB, Martinez R, Hunter T, Sudarsanam S. The protein kinase complement of the human genome. *Science* 2002; 298(5600):1912-1934.
- (213) Zerlotini A, Heiges M, Wang H, Moraes RL, Dominitini AJ, Ruiz JC et al. SchistoDB: a *Schistosoma mansoni* genome resource. *Nucleic Acids Res* 2009; 37(Database issue):D579-D582.
- (214) Boudeau J, Miranda-Saavedra D, Barton GJ, Alessi DR. Emerging roles of pseudokinases. *Trends Cell Biol* 2006; 16(9):443-452.
- (215) Hanks SK, Quinn AM, Hunter T. The protein kinase family: conserved features and deduced phylogeny of the catalytic domains. *Science* 1988; 241(4861):42-52.
- (216) Hanks SK, Hunter T. Protein kinases 6. The eukaryotic protein kinase superfamily: kinase (catalytic) domain structure and classification. *FASEB J* 1995; 9(8):576-596.

- (217) Taylor SS, Buechler JA, Yonemoto W. cAMP-dependent protein kinase: framework for a diverse family of regulatory enzymes. *Annu Rev Biochem* 1990; 59:971-1005.
- (218) Swierczewski BE, Davies SJ. Conservation of protein kinase a catalytic subunit sequences in the schistosome pathogens of humans. *Experimental Parasitology* 2010; 125(2):156-160.
- (219) Donald RG, Allocco J, Singh SB, Nare B, Salowe SP, Wiltsie J et al. *Toxoplasma gondii* cyclic GMP-dependent kinase: chemotherapeutic targeting of an essential parasite protein kinase. *Eukaryot Cell* 2002; 1(3):317-328.
- (220) Gurnett AM, Liberator PA, Dulski PM, Salowe SP, Donald RG, Anderson JW et al. Purification and molecular characterization of cGMP-dependent protein kinase from Apicomplexan parasites. A novel chemotherapeutic target. *J Biol Chem* 2002; 277(18):15913-15922.
- (221) Diaz CA, Allocco J, Powles MA, Yeung L, Donald RG, Anderson JW et al. Characterization of *Plasmodium falciparum* cGMP-dependent protein kinase (PfPKG): antiparasitic activity of a PKG inhibitor. *Mol Biochem Parasitol* 2006; 146(1):78-88.
- (222) Knockaert M, Gray N, Damiens E, Chang YT, Grellier P, Grant K et al. Intracellular targets of cyclin-dependent kinase inhibitors: identification by affinity chromatography using immobilised inhibitors. *Chem Biol* 2000; 7(6):411-422.
- (223) Leutner S, Beckmann S, Grevelding CG. Characterization of the cGMP-dependent protein kinase SmcGK1 of *Schistosoma mansoni*. *Anais da Academia Brasileira de Ciencias* 2011; 83(2):637-648.
- (224) Mellor H, Parker PJ. The extended protein kinase C superfamily. *Biochem J* 1998; 332 ( Pt 2):281-292.
- (225) Parker PJ, Murray-Rust J. PKC at a glance. *J Cell Sci* 2004; 117(Pt 2):131-132.
- (226) Eckert JJ, McCallum A, Mears A, Rumsby MG, Cameron IT, Fleming TP. Specific PKC isoforms regulate blastocoel formation during mouse preimplantation development. *Dev Biol* 2004; 274(2):384-401.
- (227) Gallicano GI, Yousef MC, Capco DG. PKC--a pivotal regulator of early development. *Bioessays* 1997; 19(1):29-36.
- (228) Steinhart R, Kazimirsky G, Okhrimenko H, Ben-Hur T, Brodie C. PKCepsilon induces astrocytic differentiation of multipotential neural precursor cells. *Glia* 2007; 55(2):224-232.
- (229) Manning G. Genomic overview of protein kinases. *WormBook* 2005;1-19.
- (230) Bahia D, Avelar LG, Vigorosi F, Cioli D, Oliveira GC, Mortara RA. The distribution of motor proteins in the muscles and flame cells of the *Schistosoma mansoni* miracidium and primary sporocyst. *Parasitology* 2006; 133(Pt 3):321-329.

- (231) Bahia D, Avelar L, Mortara RA, Khayath N, Yan Y, Noel C et al. SmPKC1, a new protein kinase C identified in the platyhelminth parasite *Schistosoma mansoni*. *Biochem Biophys Res Commun* 2006; 345(3):1138-1148.
- (232) Ludtmann MH, Rollinson D, Emery AM, Walker AJ. Protein kinase C signalling during miracidium to mother sporocyst development in the helminth parasite, *Schistosoma mansoni*. *Int J Parasitol* 2009; 39(11):1223-1233.
- (233) Bahia D, Avelar LG, Vigorosi F, Cioli D, Oliveira GC, Mortara RA. The distribution of motor proteins in the muscles and flame cells of the *Schistosoma mansoni* miracidium and primary sporocyst. *Parasitology* 2006; 133(Pt 3):321-329.
- (234) Bahia D, Avelar L, Mortara RA, Khayath N, Yan Y, Noel C et al. SmPKC1, a new protein kinase C identified in the platyhelminth parasite *Schistosoma mansoni*. *Biochem Biophys Res Commun* 2006; 345(3):1138-1148.
- (235) Ludtmann MH, Rollinson D, Emery AM, Walker AJ. Protein kinase C signalling during miracidium to mother sporocyst development in the helminth parasite, *Schistosoma mansoni*. *Int J Parasitol* 2009; 39(11):1223-1233.
- (236) Swulius MT, Waxham MN. Ca(2+)/calmodulin-dependent protein kinases. *Cell Mol Life Sci* 2008; 65(17):2637-2657.
- (237) Plowman GD, Sudarsanam S, Bingham J, Whyte D, Hunter T. The protein kinases of *Caenorhabditis elegans*: a model for signal transduction in multicellular organisms. *Proc Natl Acad Sci U S A* 1999; 96(24):13603-13610.
- (238) Swulius MT, Waxham MN. Ca(2+)/calmodulin-dependent protein kinases. *Cell Mol Life Sci* 2008; 65(17):2637-2657.
- (239) Manning G, Whyte DB, Martinez R, Hunter T, Sudarsanam S. The protein kinase complement of the human genome. *Science* 2002; 298(5600):1912-1934.
- (240) Doenhoff MJ, Pica-Mattoccia L. Praziquantel for the treatment of schistosomiasis: its use for control in areas with endemic disease and prospects for drug resistance. *Expert Rev Anti Infect Ther* 2006; 4(2):199-210.
- (241) Doenhoff MJ, Cioli D, Utzinger J. Praziquantel: mechanisms of action, resistance and new derivatives for schistosomiasis. *Curr Opin Infect Dis* 2008; 21(6):659-667.
- (242) Labeit S, Kolmerer B. Titins - Giant Proteins in Charge of Muscle Ultrastructure and Elasticity. *Science* 1995; 270(5234):293-296.
- (243) Capes-Davis A, Tolhurst O, Dunn JM, Jeffrey PL. Expression of doublecortin (DCX) and doublecortin-like kinase (DCLK) within the developing chick brain. *Dev Dyn* 2005; 232(2):457-467.
- (244) Lin PT, Gleeson JG, Corbo JC, Flanagan L, Walsh CA. DCAMKL1 encodes a protein kinase with homology to doublecortin that regulates microtubule polymerization. *J Neurosci* 2000; 20(24):9152-9161.

- (245) Lehtonen S, Ryan JJ, Kudlicka K, Iino N, Zhou H, Farquhar MG. Cell junction-associated proteins IQGAP1, MAGI-2, CASK, spectrins, and alpha-actinin are components of the nephrin multiprotein complex. *Proc Natl Acad Sci U S A* 2005; 102(28):9814-9819.
- (246) Caffrey CR, Rohwer A, Oellien F, Marhofer RJ, Braschi S, Oliveira G et al. A comparative chemogenomics strategy to predict potential drug targets in the metazoan pathogen, *Schistosoma mansoni*. *PLoS One* 2009; 4(2):e4413.
- (247) Bradham CA, Foltz KR, Beane WS, Arnone MI, Rizzo F, Coffman JA et al. The sea urchin kinome: a first look. *Dev Biol* 2006; 300(1):180-193.
- (248) Sundaram MV. RTK/Ras/MAPK signaling. *WormBook* 2006;1-19.
- (249) Plowman GD, Sudarsanam S, Bingham J, Whyte D, Hunter T. The protein kinases of *Caenorhabditis elegans*: a model for signal transduction in multicellular organisms. *Proc Natl Acad Sci U S A* 1999; 96(24):13603-13610.
- (250) Kalis AK, Kroetz MB, Larson KM, Zarkower D. Functional genomic identification of genes required for male gonadal differentiation in *Caenorhabditis elegans*. *Genetics* 2010; 185(2):523-535.
- (251) Dohrmann PR, Sclafani RA. Novel role for checkpoint Rad53 protein kinase in the initiation of chromosomal DNA replication in *Saccharomyces cerevisiae*. *Genetics* 2006; 174(1):87-99.
- (252) Noselli S. JNK signaling and morphogenesis in *Drosophila*. *Trends Genet* 1998; 14(1):33-38.
- (253) Stronach B, Perrimon N. Activation of the JNK pathway during dorsal closure in *Drosophila* requires the mixed lineage kinase, slipper. *Genes Dev* 2002; 16(3):377-387.
- (254) Plowman GD, Sudarsanam S, Bingham J, Whyte D, Hunter T. The protein kinases of *Caenorhabditis elegans*: a model for signal transduction in multicellular organisms. *Proc Natl Acad Sci U S A* 1999; 96(24):13603-13610.
- (255) Kondoh K, Nishida E. Regulation of MAP kinases by MAP kinase phosphatases. *Biochim Biophys Acta* 2007; 1773(8):1227-1237.
- (256) Nishimoto S, Nishida E. MAPK signalling: ERK5 versus ERK1/2. *EMBO Rep* 2006; 7(8):782-786.
- (257) Wilkinson MG, Millar JB. Control of the eukaryotic cell cycle by MAP kinase signaling pathways. *FASEB J* 2000; 14(14):2147-2157.
- (258) Nishimoto S, Nishida E. MAPK signalling: ERK5 versus ERK1/2. *EMBO Rep* 2006; 7(8):782-786.

- (259) Ressurreicao M, Rollinson D, Emery AM, Walker AJ. A role for p38 mitogen-activated protein kinase in early post-embryonic development of *Schistosoma mansoni*. *Molecular and Biochemical Parasitology* 2011; 180(1):51-55.
- (260) Ressurreicao M, Rollinson D, Emery AM, Walker AJ. A role for p38 MAPK in the regulation of ciliary motion in a eukaryote. *Bmc Cell Biology* 2011; 12.
- (261) Yan Y, Tulasne D, Browaeys E, Cailliau K, Khayath N, Pierce RJ et al. Molecular cloning and characterisation of SmSLK, a novel Ste20-like kinase in *Schistosoma mansoni*. *Int J Parasitol* 2007; 37(14):1539-1550.
- (262) Hu W, Brindley PJ, McManus DP, Feng Z, Han ZG. Schistosome transcriptomes: new insights into the parasite and schistosomiasis. *Trends Mol Med* 2004; 10(5):217-225.
- (263) Salzet M, Capron A, Stefano GB. Molecular crosstalk in host-parasite relationships: schistosome- and leech-host interactions. *Parasitol Today* 2000; 16(12):536-540.
- (264) Plowman GD, Sudarsanam S, Bingham J, Whyte D, Hunter T. The protein kinases of *Caenorhabditis elegans*: a model for signal transduction in multicellular organisms. *Proc Natl Acad Sci U S A* 1999; 96(24):13603-13610.
- (265) Muhlrud PJ, Ward S. Spermiogenesis initiation in *Caenorhabditis elegans* involves a casein kinase 1 encoded by the *spe-6* gene. *Genetics* 2002; 161(1):143-155.
- (266) Reinke V, Smith HE, Nance J, Wang J, Van DC, Begley R et al. A global profile of germline gene expression in *C. elegans*. *Mol Cell* 2000; 6(3):605-616.
- (267) Bahia D, Andrade LF, Ludolf F, Mortara RA, Oliveira G. Protein tyrosine kinases in *Schistosoma mansoni*. *Mem Inst Oswaldo Cruz* 2006; 101 Suppl 1:137-143.
- (268) Ahier A, Khayath N, Vicogne J, Dissous C. Insulin receptors and glucose uptake in the human parasite *Schistosoma mansoni*. *Parasite* 2008; 15(4):573-579.
- (269) You H, Zhang WB, Moertel L, McManus DP, Gobert GN. Transcriptional profiles of adult male and female *Schistosoma japonicum* in response to insulin reveal increased expression of genes involved in growth and development. *International Journal for Parasitology* 2009; 39(14):1551-1559.
- (270) You H, Zhang WB, Jones MK, Gobert GN, Mulvenna J, Rees G et al. Cloning and Characterisation of *Schistosoma japonicum* Insulin Receptors. *Plos One* 2010; 5(3).
- (271) Ahier A, Khayath N, Vicogne J, Dissous C. Insulin receptors and glucose uptake in the human parasite *Schistosoma mansoni*. *Parasite* 2008; 15(4):573-579.
- (272) You H, Zhang WB, Moertel L, McManus DP, Gobert GN. Transcriptional profiles of adult male and female *Schistosoma japonicum* in response to insulin reveal increased expression of genes involved in growth and development. *International Journal for Parasitology* 2009; 39(14):1551-1559.

- (273) You H, Zhang WB, Jones MK, Gobert GN, Mulvenna J, Rees G et al. Cloning and Characterisation of *Schistosoma japonicum* Insulin Receptors. *Plos One* 2010; 5(3).
- (274) Ahier A, Khayath N, Vicogne J, Dissous C. Insulin receptors and glucose uptake in the human parasite *Schistosoma mansoni*. *Parasite* 2008; 15(4):573-579.
- (275) Dissous C, Ahier A, Khayath N. Protein tyrosine kinases as new potential targets against human schistosomiasis. *Bioessays* 2007; 29(12):1281-1288.
- (276) Khayath N, Vicogne J, Ahier A, BenYounes A, Konrad C, Trolet J et al. Diversification of the insulin receptor family in the helminth parasite *Schistosoma mansoni*. *FEBS J* 2007; 274(3):659-676.
- (277) Kim JJ, Accili D. Signalling through IGF-I and insulin receptors: where is the specificity? *Growth Horm IGF Res* 2002; 12(2):84-90.
- (278) Ahier A, Khayath N, Vicogne J, Dissous C. Insulin receptors and glucose uptake in the human parasite *Schistosoma mansoni*. *Parasite* 2008; 15(4):573-579.
- (279) Ramachandran H, Skelly PJ, Shoemaker CB. The *Schistosoma mansoni* epidermal growth factor receptor homologue, SER, has tyrosine kinase activity and is localized in adult muscle. *Mol Biochem Parasitol* 1996; 83(1):1-10.
- (280) Ramachandran H, Skelly PJ, Shoemaker CB. The *Schistosoma mansoni* epidermal growth factor receptor homologue, SER, has tyrosine kinase activity and is localized in adult muscle. *Mol Biochem Parasitol* 1996; 83(1):1-10.
- (281) Vicogne J, Cailliau K, Tulasne D, Browaeys E, Yan YT, Fafeur V et al. Conservation of epidermal growth factor receptor function in the human parasitic helminth *Schistosoma mansoni*. *J Biol Chem* 2004; 279(36):37407-37414.
- (282) Ahier A, Rondard P, Gougnard N, Khayath N, Huang S, Trolet J et al. A new family of receptor tyrosine kinases with a venus flytrap binding domain in insects and other invertebrates activated by aminoacids. *PLoS One* 2009; 4(5):e5651.
- (283) Gougnard N, Vanderstraete M, Cailliau K, Lescuyer A, Browaeys E, Dissous C. *Schistosoma mansoni*: Structural and biochemical characterization of two distinct Venus Kinase Receptors. *Experimental Parasitology*. In press.
- (284) Kapp K, Schussler P, Kunz W, Grevelding CG. Identification, isolation and characterization of a Fyn-like tyrosine kinase from *Schistosoma mansoni*. *Parasitology* 2001; 122(Pt 3):317-327.
- (285) Kapp K, Knobloch J, Schussler P, Sroka S, Lammers R, Kunz W et al. The *Schistosoma mansoni* Src kinase TK3 is expressed in the gonads and likely involved in cytoskeletal organization. *Mol Biochem Parasitol* 2004; 138(2):171-182.
- (286) Knobloch J, Winnen R, Quack M, Kunz W, Grevelding CG. A novel Syk-family tyrosine kinase from *Schistosoma mansoni* which is preferentially transcribed in reproductive organs. *Gene* 2002; 294(1-2):87-97.

- (287) Beckmann S, Grevelding CG. Imatinib has a fatal impact on morphology, pairing stability and survival of adult *Schistosoma mansoni* in vitro. *Int J Parasitol* 2010; 40(5):521-526.
- (288) Bahia D, Mortara RA, Kusel JR, Andrade LF, Ludolf F, Kuser PR et al. *Schistosoma mansoni*: expression of Fes-like tyrosine kinase SmFes in the tegument and terebratorium suggests its involvement in host penetration. *Exp Parasitol* 2007; 116(3):225-232.
- (289) Ludolf F, Bahia D, Andrade LF, Cousin A, Capron M, Dissous C et al. Molecular analysis of SmFes, a tyrosine kinase of *Schistosoma mansoni* orthologous to the members of the Fes/Fps/Fer family. *Biochem Biophys Res Commun* 2007; 360(1):163-172.
- (290) Ward P, Equinet L, Packer J, Doerig C. Protein kinases of the human malaria parasite *Plasmodium falciparum*: the kinome of a divergent eukaryote. *BMC Genomics* 2004; 5(1):79.
- (291) Vowels JJ, Thomas JH. Genetic analysis of chemosensory control of dauer formation in *Caenorhabditis elegans*. *Genetics* 1992; 130(1):105-123.
- (292) Pils B, Schultz J. Inactive enzyme-homologues find new function in regulatory processes. *J Mol Biol* 2004; 340(3):399-404.
- (293) Boudeau J, Miranda-Saavedra D, Barton GJ, Alessi DR. Emerging roles of pseudokinases. *Trends Cell Biol* 2006; 16(9):443-452.
- (294) Deshmukh K, Anamika K, Srinivasan N. Evolution of domain combinations in protein kinases and its implications for functional diversity. *Prog Biophys Mol Biol* 2010; 102(1):1-15.
- (295) Manning G, Plowman GD, Hunter T, Sudarsanam S. Evolution of protein kinase signaling from yeast to man. *Trends Biochem Sci* 2002; 27(10):514-520.
- (296) Miranda-Saavedra D, Barton GJ. Classification and functional annotation of eukaryotic protein kinases. *Proteins* 2007; 68(4):893-914.
- (297) Wilkinson MG, Millar JB. Control of the eukaryotic cell cycle by MAP kinase signaling pathways. *FASEB J* 2000; 14(14):2147-2157.
- (298) Beall MJ, Pearce EJ. Human transforming growth factor-beta activates a receptor serine/threonine kinase from the intravascular parasite *Schistosoma mansoni*. *J Biol Chem* 2001; 276(34):31613-31619.
- (299) Davies SJ, Shoemaker CB, Pearce EJ. A divergent member of the transforming growth factor beta receptor family from *Schistosoma mansoni* is expressed on the parasite surface membrane. *J Biol Chem* 1998; 273(18):11234-11240.

- (300) Forrester SG, Warfel PW, Pearce EJ. Tegumental expression of a novel type II receptor serine/threonine kinase (SmRK2) in *Schistosoma mansoni*. *Mol Biochem Parasitol* 2004; 136(2):149-156.
- (301) LoVerde PT, Osman A, Hinck A. *Schistosoma mansoni*: TGF-beta signaling pathways. *Exp Parasitol* 2007; 117(3):304-317.
- (302) Osman A, Niles EG, Verjovski-Almeida S, LoVerde PT. *Schistosoma mansoni* TGF-beta receptor II: role in host ligand-induced regulation of a schistosome target gene. *PLoS Pathog* 2006; 2(6):e54.
- (303) Dissous C, Grevelding CG, Long T. *Schistosoma mansoni* Polo-like kinases and their function in control of mitosis and parasite reproduction. *Anais da Academia Brasileira de Ciencias* 2011; 83(2):627-635.
- (304) Dissous C, Grevelding CG, Long T. *Schistosoma mansoni* Polo-like kinases and their function in control of mitosis and parasite reproduction. *Anais da Academia Brasileira de Ciencias* 2011; 83(2):627-635.
- (305) Eglen RM, Reisine T. The current status of drug discovery against the human kinome. *Assay Drug Dev Technol* 2009; 7(1):22-43.
- (306) Deshmukh K, Anamika K, Srinivasan N. Evolution of domain combinations in protein kinases and its implications for functional diversity. *Prog Biophys Mol Biol* 2010; 102(1):1-15.
- (307) Eglen RM, Reisine T. The current status of drug discovery against the human kinome. *Assay Drug Dev Technol* 2009; 7(1):22-43.
- (308) Finn RD, Mistry J, Tate J, Coghill P, Heger A, Pollington JE et al. The Pfam protein families database. *Nucleic Acids Res* 2010; 38(Database issue):D211-D222.
- (309) Kinbase - The kinase database at Sugen/Salk. <http://www.kinase.com/kinbase> . 2010. 8-8-2009.

Ref Type: Electronic Citation

- (310) Finn RD, Mistry J, Tate J, Coghill P, Heger A, Pollington JE et al. The Pfam protein families database. *Nucleic Acids Res* 2010; 38(Database issue):D211-D222.
- (311) Tanoue T, Yamamoto T, Nishida E. Modular structure of a docking surface on MAPK phosphatases. *J Biol Chem* 2002; 277(25):22942-22949.
- (312) Zhou B, Wu L, Shen K, Zhang J, Lawrence DS, Zhang ZY. Multiple regions of MAP kinase phosphatase 3 are involved in its recognition and activation by ERK2. *J Biol Chem* 2001; 276(9):6506-6515.
- (313) Tanoue T, Adachi M, Moriguchi T, Nishida E. A conserved docking motif in MAP kinases common to substrates, activators and regulators. *Nat Cell Biol* 2000; 2(2):110-116.

- (314) Tanoue T, Adachi M, Moriguchi T, Nishida E. A conserved docking motif in MAP kinases common to substrates, activators and regulators. *Nat Cell Biol* 2000; 2(2):110-116.
- (315) Finn RD, Mistry J, Tate J, Coggill P, Heger A, Pollington JE et al. The Pfam protein families database. *Nucleic Acids Res* 2010; 38(Database issue):D211-D222.
- (316) Ludolf F, Bahia D, Andrade LF, Cousin A, Capron M, Dissous C et al. Molecular analysis of SmFes, a tyrosine kinase of *Schistosoma mansoni* orthologous to the members of the Fes/Fps/Fer family. *Biochem Biophys Res Commun* 2007; 360(1):163-172.
- (317) Deshmukh K, Anamika K, Srinivasan N. Evolution of domain combinations in protein kinases and its implications for functional diversity. *Prog Biophys Mol Biol* 2010; 102(1):1-15.
- (318) Ren J, Wen L, Gao X, Jin C, Xue Y, Yao X. DOG 1.0: illustrator of protein domain structures. *Cell Res* 2009; 19(2):271-273.
- (319) Finn RD, Mistry J, Tate J, Coggill P, Heger A, Pollington JE et al. The Pfam protein families database. *Nucleic Acids Res* 2010; 38(Database issue):D211-D222.
- (320) Ren J, Wen L, Gao X, Jin C, Xue Y, Yao X. DOG 1.0: illustrator of protein domain structures. *Cell Res* 2009; 19(2):271-273.
- (321) Finn RD, Mistry J, Tate J, Coggill P, Heger A, Pollington JE et al. The Pfam protein families database. *Nucleic Acids Res* 2010; 38(Database issue):D211-D222.
- (322) Krautz-Peterson G, Skelly PJ. *Schistosoma mansoni*: the dicer gene and its expression. *Exp Parasitol* 2008; 118(1):122-128.
- (323) Swulius MT, Waxham MN. Ca(2+)/calmodulin-dependent protein kinases. *Cell Mol Life Sci* 2008; 65(17):2637-2657.
- (324) Sagasti A, Hisamoto N, Hyodo J, Tanaka-Hino M, Matsumoto K, Bargmann CI. The CaMKII UNC-43 activates the MAPKKK NSY-1 to execute a lateral signaling decision required for asymmetric olfactory neuron fates. *Cell* 2001; 105(2):221-232.
- (325) Malumbres M, Barbacid M. RAS oncogenes: the first 30 years. *Nat Rev Cancer* 2003; 3(6):459-465.
- (326) Osman A, Niles EG, LoVerde PT. Characterization of the Ras homologue of *Schistosoma mansoni*. *Mol Biochem Parasitol* 1999; 100(1):27-41.
- (327) Krautz-Peterson G, Bhardwaj R, Faghiri Z, Tararam CA, Skelly PJ. RNA interference in schistosomes: machinery and methodology. *Parasitology* 2010; 137(3):485-495.

- (328) Krautz-Peterson G, Bhardwaj R, Faghiri Z, Tararam CA, Skelly PJ. RNA interference in schistosomes: machinery and methodology. *Parasitology* 2010; 137(3):485-495.
- (329) Rao MK, Wilkinson MF. Tissue-specific and cell type-specific RNA interference in vivo. *Nature Protocols* 2006; 1(3):1494-1501.
- (330) Reynolds A, Leake D, Boese Q, Scaringe S, Marshall WS, Knvorova A. Rational siRNA design for RNA interference. *Nature Biotechnology* 2004;(22):326-330.
- (331) Noselli S. JNK signaling and morphogenesis in *Drosophila*. *Trends Genet* 1998; 14(1):33-38.
- (332) Tasaki J, Shibata N, Sakurai T, Agata K, Umesono Y. Role of c-Jun N-terminal kinase activation in blastema formation during planarian regeneration. *Development Growth & Differentiation* 2011; 53(3):389-400.
- (333) Ely HA, Mellon PL, Coss D. GnRH Induces the c-Fos Gene via Phosphorylation of SRF by the Calcium/Calmodulin Kinase II Pathway. *Molecular Endocrinology* 2011; 25(4):669-680.
- (334) Nebreda AR, Ferby I. Regulation of the meiotic cell cycle in oocytes. *Curr Opin Cell Biol* 2000; 12(6):666-675.
- (335) Hajnal A, Berset T. The *C.elegans* MAPK phosphatase LIP-1 is required for the G(2)/M meiotic arrest of developing oocytes. *EMBO J* 2002; 21(16):4317-4326.
- (336) Hazzalin CA, Mahadevan LC. MAPK-regulated transcription: A continuously variable gene switch? *Nature Reviews Molecular Cell Biology* 2002; 3(1):30-40.
- (337) Kyriakis JM, Avruch J. Mammalian mitogen-activated protein kinase signal transduction pathways activated by stress and inflammation. *Physiol Rev* 2001; 81(2):807-869.
- (338) Hazzalin CA, Mahadevan LC. MAPK-regulated transcription: A continuously variable gene switch? *Nature Reviews Molecular Cell Biology* 2002; 3(1):30-40.
- (339) Johnson RS, Spiegelman BM, Papaioannou V. Pleiotropic Effects of A Null Mutation in the C-Fos Protooncogene. *Cell* 1992; 71(4):577-586.
- (340) Tan PB, Lackner MR, Kim SK. MAP kinase signaling specificity mediated by the LIN-1 Ets/LIN31 WH transcription factor complex during *C.elegans* vulval induction. *Cell* 1998; 93(4):569-580.
- (341) Ely HA, Mellon PL, Coss D. GnRH Induces the c-Fos Gene via Phosphorylation of SRF by the Calcium/Calmodulin Kinase II Pathway. *Molecular Endocrinology* 2011; 25(4):669-680.