

MARIANA SERPELONI

**Exportação de mRNAs do núcleo para o citoplasma:
análise do papel da proteína Sub2 em *Trypanosoma cruzi* e
*Toxoplasma gondii***

Tese apresentada como requisito à obtenção do grau de Doutor pelo Programa de Pós-Graduação em Biologia Celular e Molecular do Setor de Ciências Biológicas da Universidade Federal do Paraná e realizada no Instituto Carlos Chagas – Fiocruz Paraná.

Orientadora: Andréa Rodrigues Ávila

Co-orientador: Stenio Perdigão Fragoso

CURITIBA

ABRIL – 2014

UNIVERSIDADE FEDERAL DO PARANÁ
SETOR DE CIÊNCIAS BIOLÓGICAS
PROGRAMA DE PÓS-GRADUAÇÃO EM BIOLOGIA CELULAR E MOLECULAR
E INSTITUTO CARLOS CHAGAS

MARIANA SERPELONI

**Exportação de mRNAs do núcleo para o citoplasma:
análise do papel da proteína Sub2 em *Trypanosoma cruzi* e
*Toxoplasma gondii***

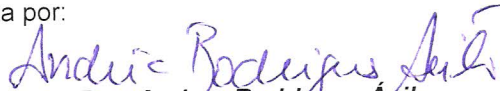
CURITIBA
ABRIL - 2014

PROGRAMA DE PÓS-GRADUAÇÃO EM BIOLOGIA CELULAR E MOLECULAR

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

PARECER

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



Prof. Dra. Andrea Rodrigues Ávila
Orientadora e presidente da banca
Instituto Carlos Chagas - ICC



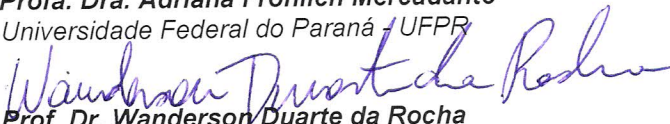
Prof. Dra. Carolina Borsoi Moraes Holanda de Freitas
Centro Nacional de Pesquisa em Energia e Materiais - CNPEM



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



Prof. Dra. Adriana Frohlich Mercadante
Universidade Federal do Paraná - UFPR



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

E tendo como suplentes,

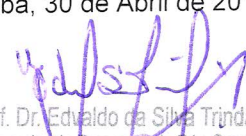
Prof. Dr. Alejandro Correa Dominguez
Instituto Carlos Chaves - ICC

Prof. Dr. Edvaldo da Silva Trindade
Universidade Federal do Paraná - UFPR

Após arguir a candidata **Mariana Serpeloni**, em relação ao seu trabalho intitulado: "Exportação de mRNAs do núcleo para o citoplasma: Análise do papel da proteína Sub2 em *Trypanosoma cruzi* e *Toxoplasma gondii*", são de parecer favorável à **APROVAÇÃO** da acadêmica, habilitando-a ao título de Doutora em Biologia Celular e Molecular, área de concentração em Biologia Celular e Molecular.

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

Curitiba, 30 de Abril de 2014



Prof. Dr. Edvaldo da Silva Trindade
Coordenador do Programa de Pós-Graduação
em Biologia Celular e Molecular - UFPR
Matr. 185795

À minha família: Walter, Marilena, Fernanda e André que me incentivam em todos os momentos e sempre me fizeram acreditar nos meus sonhos.

“Creativity is a challenge. It requires us to be fully human -- autonomous yet engaged, independent yet interdependent. Creativity bridges the conflict between our individualistic and our sociality. It celebrates the commonality of our species while simultaneously setting us apart as unique individuals.”

Dr. Greg Graffin

AGRADECIMENTOS

A Dra. Andréa Rodrigues Ávila pela orientação. Agradeço por todas as oportunidades que me foram proporcionadas, pela amizade, paciência, alegria contagiante, competência e confiança em meu trabalho.

Ao Dr. Stenio Perdigão Fragoso pela co-orientação, amizade e paciência em ensinar e discutir os resultados.

Ao Dr. Nilson Tonin Zanchin pela amizade e ajuda com as cromatografias e discussões científicas relevantes neste estudo.

Ao Dr. Edvaldo e Marlene, da UFPR, pela ajuda durante todos esses anos de Pós-Graduação.

Ao Dr. Newton de Medeiros Vidal pela ajuda da parte de bioinformática para obtenção dos resultados de filogenia. Agradeço pela amizade, ensinamentos e colaboração essencial nas discussões dos resultados.

Ao Dr. Henrique Preti pela ajuda e discussão dos resultados de duplo-híbrido.

Ao Dr. Paulo Costa Carvalho pela colaboração nas análises dos resultados de proteômica e agradeço a ajuda dos Drs. Michel Batista e Fabrício K. Marchini, da Plataforma de Proteômica do Instituto Carlos Chagas.

A todos do “Level 6” e do Laboratório do Markus Meissner pelo acolhimento e amizade num país tão gelado! Em especial aos amigos: Viviçoquis, Tiago, Fernando, Ellie, Nicole, Manu, Jo, Catarina, Saskia, Robyn, Jamie, Alli e Gurman, tanto pelos ensinamentos quanto pelas risadas!

A todos do Laboratório de Regulação da Expressão Gênica. Em especial Hálisson Joseph, Camaila, Pá, Haruo san e Sheila pela amizade e ajuda em todas as etapas. Agradeço aos amigos sinceros que fiz durante este tempo de ICC (Galera do mal), que sem eles todos esses anos de trabalho não teriam tanta graça e leveza, com certeza: Shacera Maria, Hálisson Joseph, Maycon, Ferrarini, Sharoline, Vane, Rafael, Monica, Alda Maria, Rô, Flavinha, Camaila, Pá, Axel, Bruno Dallagiovanna (bem fofo), Felipe e Sheila. O lema é: Dignidade já!

Ao Nilson, Vane, Tânia e Silvio por proporcionar a organização ideal, facilitando a realização deste estudo.

Aos meus pais Marilena e Walter e irmã, Fernanda, pelo apoio emocional, compreensão e carinho durante todo o desenvolvimento deste estudo. Mãe,

obrigada por ter lido a tese e entendido tudo heheheh. Agradeço ao cunhado Fred pela amizade e discussões bem nerds, científicas ou não.

Ao André pelo amor, apoio nas horas difíceis e pela alegria durante todos esses 15 anos de convivência.

Aos amigos que, mesmo com a distância, estão sempre presentes em todas as etapas da minha vida: Fer, Zan, Lina tchan, Thiaguinho, Junior Paião, Cynthia, Helena san, Cris, Melissa (Tirissa), Manuzes, Patti, Rafasguela, Vanessa (Pedrinhas) Nerso (empreendedorista), Paçoca, Poliguarro, Pri Cassola, Pri Henning e Pauleka. Vocês são demais!!!

SUMÁRIO

1	Introdução	1
1.1	<i>Trypanosoma cruzi</i> e a doença de Chagas	1
1.1.1	Expressão e regulação gênica de tripanossomatídeos	4
1.2	<i>Toxoplasma gondii</i> e a doença Toxoplasmose	8
1.2.1	Regulação da expressão gênica e abordagens de genética reversa em <i>T. gondii</i>	10
1.3	Exportação de mRNAs em eucariotos: do núcleo ao citoplasma	14
1.3.1	Monitoramento de transporte de mRNAs através do sistema MS2	24
1.3.2	Exportação de mRNAs em parasitas: abordagem do problema	25
2	Objetivos	30
2.1	Objetivo geral	30
2.1.1	Objetivos específicos	30
3	Capítulo 1. Comparative genomics of proteins involved in RNA nucleocytoplasmic export	32
4	Capítulo 2. An Essential Nuclear Protein in Trypanosomes Is a Component of mRNA Transcription/Export Pathway	35
5	Capítulo 3. TgUAP56: an essential nuclear protein related with mRNA export in <i>T. gondii</i> (artigo em preparação – Anexo I)	37
5.1	Materiais e Métodos	38
5.1.1	Análises de sequências de proteínas ortólogas de UAP56 e filogenia por ferramentas de bioinformática	38
5.1.2	Meios de cultura	40
5.1.3	Microrganismos	40
5.1.4	Soluções e tampões	42
5.1.5	Amplificação de DNA ou cDNA por PCR (Reação em cadeia da polimerase - <i>Polymerase Chain Reaction</i>) e subclonagem do produto em vetor pGEM-Teasy®	43
5.1.6	Reação de Ligação de DNA (fragmento e vetor, previamente digeridos com as mesmas enzimas de restrição)	45
5.1.7	Preparação e transformação de bactérias quimio-competentes	45
5.1.8	Confirmação de clonagens por PCR de colônia	46
5.1.9	Imunofluorescência para a localização subcelular de proteínas por microscopia de fluorescência e deconvolução	47
5.1.10	Ensaio de <i>Western blot</i> e revelação por quimioluminescência	47
5.1.11	Linhagem de células de mamíferos HFF (<i>human foreskin fibroblast</i>)	49

5.1.12	Cultivo e expansão de células HFF	49
5.1.13	Cultivo de <i>T. gondii</i> em linhagens de células de mamíferos	50
5.1.14	Congelamento de <i>T. gondii</i>	50
5.1.15	Transfecção de <i>T. gondii</i>	51
5.1.16	Obtenção de linhagem clonal de parasitas por diluição limitante	52
5.1.17	Isolamento de DNA genômico de <i>T. gondii</i>	52
5.1.18	Obtenção de cDNA de <i>T. gondii</i>	53
5.1.19	Ensaio de placa de lise (<i>plaque assay</i>).....	53
5.2	Caracterização funcional de TgUAP56 por abordagens de genética reversa .	54
5.2.1	Superexpressão de dd-GFP-TgUAP56 em <i>T. gondii</i> pelo sistema DD	54
5.2.1.1	Expressão e localização subcelular de proteína dd-GFP-TgUAP56 em <i>T. gondii</i> ...	56
5.2.2	Nocautado condicionado do gene <i>uap56</i> em <i>T. gondii</i> pelo sistema DiCre.....	57
5.2.2.1	Análise de integração do cassete para nocautado condicionado no genoma de parasitas da linhagem clonal cKOuap56.....	61
5.2.2.2	Análise de processamento por <i>splicing</i> de mRNAs após nocautado condicionado do gene <i>uap56</i>	63
5.2.3	Monitoramento de mRNAs em <i>T. gondii</i> pelo sistema MS2	64
5.2.3.1	Obtenção de parasitas transfectantes de MS2-Dip13-SL.....	64
5.2.4	Hibridação fluorescente <i>in situ</i> (FISH) para detecção de mRNAs em <i>T. gondii</i>	71
5.3	Resultados.....	73
5.3.1	Análise de conservação da proteína TgUAP56 na filogenia de eucariotos	73
5.3.2	Localização celular de TgUAP56 em <i>T. gondii</i> na linhagem DDGFPUAP56... ..	76
5.3.2.1	Obtenção de linhagem DDGFPUAP56 e estabilização da proteína dd-GFP-TgUAP56 após incubação com Shld-1	76
5.3.2.2	Análise de distribuição de mRNAs após indução da superexpressão de dd-GFP-TgUAP56... ..	81
5.3.3	Nocautado do gene causa acúmulo de mRNAs no núcleo	83
5.3.3.1	Obtenção da linhagem clonal cKOuap56 e indução do nocautado do gene <i>uap56</i> por rapamicina	83
5.3.3.2	Análise de acúmulo de mRNAs na linhagem cKOuap56 após nocautado do gene <i>uap56</i>	87
5.3.3.3	Análise do efeito do nocautado do gene <i>uap56</i> no processamento por <i>splicing</i> em alguns mRNAs.....	90
5.3.4	Monitoramento de mRNAs em <i>T. gondii</i> pelo sistema MS2	93
5.3.4.1	Análise de localização de proteína NLS-MS2-YFP-DD e mRNAs Dip13-SL.....	93
5.4	Discussão	101

6	Capítulo 4. Busca de evidências de proteínas associadas a TcSub2 bem como de proteínas relacionadas com a via de exportação de mRNAs em <i>T. cruzi</i>.	107
6.1	Materiais e Métodos	108
6.1.1	Meios de cultura	109
6.1.2	Microrganismos	109
6.1.3	Preparo dos extratos protéicos de <i>T. cruzi</i> para a imunoprecipitação de TcSub2	109
6.1.3.1	Obtenção do extrato protéico para a imunoprecipitação de TcSub2 em tampão de lise com Tris-EDTA e Noct	109
6.1.4	Ensaio de Imunoprecipitação de TcSub2 e preparo das frações eluídas para identificação por espectrometria de massas	112
6.1.4.1	Identificação de proteínas por espectrometria de massas	114
6.1.5	Cromatografia de gel-filtração (cromatografia de exclusão por tamanho)	117
6.2	Resultados e Discussão	118
6.2.1	Análise de interações de proteínas por ensaios de duplo-híbrido em leveduras	118
6.2.2	Imunoprecipitação para isolamento de complexos protéicos de <i>T. cruzi</i> contendo TcSub2	135
7	Conclusões e Perspectivas	145
8	Referências Bibliográficas	147
9	Anexo I. <i>TgUAP56: an essential nuclear protein related with mRNA export in T. gondii</i> (artigo em preparação)	175

LISTA DE FIGURAS

Figura 1. Esquema de processamento de mRNAs em mamíferos por <i>cis-splicing</i> e em tripanossomatídeos por <i>trans-splicing</i>	6
Figura 2. Sistema baseado na fusão de ddFKBP12 a uma proteína de interesse para controle de estabilidade.....	11
Figura 3. Expressão gênica condicional através de desestabilização de proteínas .	12
Figura 4. Ligação de Cre a sítios <i>LoxP</i>	13
Figura 5. Sistema DiCre para nocaute condicionado em <i>T. gondii</i>	14
Figura 6. Desenho esquemático do transporte de macromoléculas pelo ciclo de Ran em humanos	17
Figura 7. Diferentes vias de exportação de RNAs em eucariotos	18
Figura 8. Transcrição ou <i>splicing</i> acoplado com a exportação de mRNAs.....	21
Figura 9. Exportação de mRNPs	23
Figura 10. Sistema MS2 para detecção de mRNAs	25
Figura 11. Número de identificação das proteínas ortólogas de UAP56/Sub2 do GeneBank (GI) identificadas em 43 espécies representativas dos grupos de eucariotos, pintado em cores.....	39
Figura 12. Mapa do vetor pGEM® – Teasy utilizado para as etapas de subclonagens	44
Figura 13. Estratégia para a construção do vetor de superexpressão de dd-GFP-TgUAP56 em <i>T. gondii</i>	56
Figura 14. Estratégia de clonagem para a construção do vetor utilizado para nocaute condicionado do gene <i>uap56</i> em <i>T. gondii</i>	60
Figura 15. Distribuição de mRNAs esperada após o nocaute condicionado do gene <i>uap56</i> de <i>T. gondii</i>	65
Figura 16. Vetores utilizados para implementação do sistema MS2 (MS2-Dip13) em <i>T. gondii</i>	66
Figura 17. Estratégia de clonagem para a construção do vetor utilizado para obtenção de linhagem DiCre-MS2 em <i>T. gondii</i>	70
Figura 18. UAP56 é uma proteína conservada em eucariotos	75
Figura 19. TgUAP56: Proteína nuclear e essencial em <i>T. gondii</i>	78

Figura 20. Localização de dd-GFP-TgUAP56 e TgUAP56 por imunofluorescência na linhagem DDGFPUAP56	80
Figura 21. Análise de distribuição de mRNAs após superexpressão de dd-GFP-TgUAP56 na linhagem DDGFPUAP56	82
Figura 22. Análise por PCR de integração do cassete para nocaute condicionado do gene <i>uap56</i> no genoma de parasitas da linhagem clonal cKOuap56.....	85
Figura 23. Análise de distribuição de mRNAs após nocaute condicionado do gene <i>uap56</i> induzido por rapamicina em parasitas da linhagem clonal cKOuap56.....	89
Figura 24. Análise de processamento de alguns mRNAs de <i>T. gondii</i> após 24 e 48 horas da indução do nocaute do gene <i>uap56</i>	92
Figura 25. Análise do sistema MS2 em <i>T. gondii</i>	94
Figura 26. Análise da integração do cassete MS2-Dip13-SL no locus de <i>ku80</i> em parasitas da linhagem DiCre-MS2	96
Figura 27. Localização de NLS-MS2-YFP-DD após indução por 24 horas com diferentes concentrações de Shld-1.	98
Figura 28. Localização de mRNAs Dip13-SL e proteína NLS-MS2-YFP-DD na ausência ou presença de 1 µM de Shld-1 por 24 horas	99
Figura 29. Localização de proteína Ty-Dip13 na linhagem DiCre-MS2 por imunofluorescência.....	100
Figura 30. Cromatograma padrão de um teste funcional da coluna de gel-filtração Superdex 200 10/300 GL (GE Healthcare).	117
Figura 31. Busca por prováveis proteínas ortólogas de Mex67 em <i>T. cruzi</i>	120
Figura 32. Rede de interações obtida através de resultados de ensaio de duplo-híbrido para TcSub2 e proteínas candidatas a componentes da via de exportação de RNAs de <i>T. cruzi</i>	125
Figura 33. Interações identificadas entre TcSub2 e proteínas RBPs por ensaio de duplo-híbrido.....	127
Figura 34. Esquema representativo do domínio encontrado na proteína “Tclmp3”.	131
Figura 35. Rede de interação de Imp3 de <i>S. cerevisiae</i> pela ferramenta STRING	132
Figura 36. Interações identificadas para Tclmp3 por ensaio de duplo-híbrido.	134
Figura 37. Cromatograma e análise das frações coletadas do extrato protéico referente ao protocolo A	136

Figura 38. Cromatograma e análise das frações coletadas do extrato protéico referente ao protocolo C	137
Figura 39. Cromatograma e análise das frações coletadas do extrato protéico referente ao protocolo D	138
Figura 40. Imunoprecipitação da proteína TcSub2 em extrato protéico de <i>T. cruzi</i> obtido através de lise celular por TENoct (protocolo A).....	139
Figura 41. Lista de proteínas identificadas por espectrometria de massas a partir da imunoprecipitação de TcSub2 em diferentes extratos protéicos de <i>T. cruzi</i>	141
Figura 42. Localização de TcImp3 fusionada a GFP em <i>T.cruzi</i>	143

LISTA DE TABELAS

Tabela 1. Meios de cultura para bactérias, células HFF ou <i>T. gondii</i>	40
Tabela 2. Cepas e linhagens de <i>T. gondii</i> utilizadas e obtidas neste estudo	41
Tabela 3. Soluções e Tampões	43
Tabela 4. Anticorpos primários e as diluições utilizadas neste estudo	48
Tabela 5. Anticorpos secundários e as diluições utilizadas neste estudo	49
Tabela 6. Estratégias de seleção para obtenção de linhagens de parasitas transfectantes de <i>T. gondii</i>	52
Tabela 7. <i>Primers</i> utilizados para amplificar região de <i>uap56</i> a partir de cDNA para clonagem em vetor de superexpressão em <i>T. gondii</i>	54
Tabela 8. <i>Primers</i> utilizados para amplificar cDNA de <i>uap56</i> e regiões não-codificantes 5' e 3'UTRs para as clonagens em vetor de nocaute condicionado de <i>uap56</i>	58
Tabela 9. <i>Primers</i> utilizados para confirmação de integração do cassete para nocaute condicionado no locus do gene <i>uap56</i> no genoma de <i>T. gondii</i>	62
Tabela 10. <i>Primers</i> utilizados para análise de processamento de mRNAs após nocaute condicionado de <i>uap56</i>	64
Tabela 11. <i>Primers</i> utilizados para amplificar sequência de NLS-MS2-YFP-DD bem como regiões não codificantes 5' e 3' UTR relativas ao gene <i>ku80</i>	68
Tabela 12. <i>Primers</i> utilizados para confirmação de integração do cassete MS2-Dip13-SL no locus do gene <i>ku80</i> no genoma de <i>T. gondii</i>	71
Tabela 13. Soluções e Tampões	108
Tabela 14. Meio de cultura para <i>T. cruzi</i>	109
Tabela 15. Protocolos relacionados a obtenção de extratos protéicos de <i>T. cruzi</i> e lavagens referentes aos ensaios de imunoprecipitação de TcSub2	112
Tabela 16. Lista de proteínas de <i>T. cruzi</i> selecionadas para o ensaio de duplo-híbrido.....	123
Tabela 17. Lista de interações obtidas por ensaio de duplo-híbrido para TcSub2 e outras proteínas candidatas a componentes das vias de exportação de RNAs de <i>T. cruzi</i>	124
Tabela 18. Interações identificadas entre 2 alvos por ensaio de duplo-híbrido	130

LISTA DE ABREVIATURAS E SIGLAS

- ARE: *AU-rich element* (elemento rico em AU)
- ATP: Adenosina trifosfato
- BCIP: 5-bromo 4-cloro 3-indolil fosfato
- BLAST: Ferramenta de comparação de seqüências em banco de dados (*Basic Local Alignment Search Tool*)
- bp: Pares de bases (*base pair*)
- BSA: Albumina de soro bovino (*Bovine Serum Albumin*)
- cDNA: DNA complementar
- cm: Centímetro
- Da: Dalton
- DAPI: 4,6-diamidino-2-phenylindole
- DIC: Contraste interferencial (*Differential Interference Contrast*)
- Dm: *Didelphis marsupialis*
- DMSO: Dimetilsulfóxido
- DNA: Ácido desoxirribonucléico
- DNase: Desoxirribonuclease
- dNTP: Desoxirribonucleosídeo trifosfato
- E-64: L-transepoxisuccinil-leucilamido (4-guanidino)butano
- EDTA: Ácido etileno-diamino-tetracético
- FISH: Hibridação *in situ* por fluorescência (*Fluorescence in situ Hybridization*)
- g: Grama (unidade de medida de massa)
- GAP: Proteína ativadora de GTPase (*GTPase-activation protein*)
- GEF: Fator de substituição por GTP (*Guanine nucleotide exchange factor*)
- GFP: Proteína verde fluorescente (*Green Fluorescent Protein*)
- IgG: Imunoglobulina G
- IPTG: Isopropil- β -D-tiogalactopiranosídeo
- kb: Quilobase
- kDa: QuiloDalton
- kV: Quilovolt
- L: Litro
- LB: Meio Luria-Bertani
- LIT: Meio de infusão de fígado e triptose (*Liver Infusion Tryptose*)

LRR: *Leucine Rich Repeat*

M: Molar

mA: Miliamperagem

mg: Miligrama

mL: Mililitro

mm: Milímetro

mM: Milimolar

miRNA: microRNAs

mRNA: RNA mensageiro

mRNP: Complexo ribonucleoprotéico

ms: Milissegundo

mTorr: *millitorr*

NBT: *nitro blue tetrazolium chloride*

NCBI: *National Center for Biotechnology Information*

NES: Sinal de exportação nuclear (*Nuclear Export Signal*)

NLS: Sinal de Localização Nuclear (*Nuclear Localization Signal*)

ng: Nanograma

nm: Nanômetro

NPC: Complexo poro nuclear

OD: Densidade ótica

ORF: Região aberta de leitura (*Open Reading Frame*)

pb: Pares de base

PBS: Solução salina tamponada (*Phosphate Buffered Saline*)

PCR: Reação em cadeia da polimerase (*Polymerase Chain Reaction*)

pH: Potencial hidrogeniônico

pmol: Picomol

RBP: Proteína de ligação a RNA (*RNA Binding Protein*)

RNA: Ácido ribonucléico

RNase: Ribonuclease

RNP: Ribonucleoproteína

rpm: Rotações por minuto

rRNA: RNA ribossomal

SDS: Dodecil sulfato de sódio

SDS-PAGE: Eletroforese em gel de poliacrilamida com SDS

SFB: Soro fetal bovino

SL: RNA *spliced leader* (mini-éxon)

snRNA: Pequeno RNA nuclear

snoRNA: Pequeno RNA nucleolar

SSC: *Standard Saline-Citrate*

TAU: Meio Urina artificial de Triatomíneo (*Triatomine Artificial Urine*)

TFA: Ácido Trifluoroacético

TBE: Tampão Tris-Borato-EDTA

TRIS: Tris-hidroximetil aminometano

tRNA: RNA transportador

Tween 20: Monolaurato de polioxietileno (20) sorbitana

URE: Elemento de instabilidade rico em U

UTR: Região não traduzida (*Untranslated region*)

VSG: Glicoproteína variável de superfície (*Variable Surface Glycoprotein*)

X-gal: 5-bromo-4-cloro-3-indol- β -D-galactopiranosídeo

μ g: Micrograma

μ L: Microlitro

μ M: Micromolar

μ m: Micrômetro

μ s: Microsegundo

LISTA DE SÍMBOLOS

°C: Grau Celsius

%: Porcentagem

g: Aceleração de gravidade

U: Unidade

V: Volt

RESUMO

A doença de Chagas e toxoplasmose são problemas de saúde pública no Brasil e entender com detalhes a biologia dos agentes causais, *T. cruzi* e *T. gondii* respectivamente, pode ser um ponto crucial para gerar informações relevantes voltadas para o desenvolvimento de tratamento efetivo no combate a estas doenças. No entanto, a busca de novos alvos para quimioterapia de algumas doenças parasitárias esbarra, por exemplo, na ausência de ferramentas que possibilitem o estudo da função de genes essenciais para a sobrevivência do organismo. Dentro desta problemática, o objetivo central deste estudo é a identificação de proteínas envolvidas no transporte nucleocitoplasmático de RNAs, mais especificamente a exportação de mRNAs, com a utilização de ferramentas de genética reversa para análise de proteínas essenciais envolvidas nestas vias em *T. cruzi* e *T. gondii*. Além disso, *T. gondii* tem sido o modelo de estudo escolhido por diversos grupos, uma vez que dispõe de várias ferramentas de genética reversa consideradas inviáveis ou laboriosas em outros parasitas (como *Plasmodium* ou *Trypanosoma*). A escolha do tema foi baseada em dois fatos principais: a) dentro do contexto de eventos moleculares essenciais, o transporte de RNA mensageiro (mRNA) ao local de tradução é uma etapa essencial e crucial na seleção de genes a serem expressos em uma célula eucariótica; b) escassez de publicações envolvendo diversos grupos de parasitas, demonstrando ser um tema da pesquisa básica em parasitologia que necessita ser explorado. Os dados gerados neste estudo permitiram identificar que a maioria das proteínas relacionadas com a exportação de mRNAs não são tão conservadas ao longo da filogenia de eucariotos, principalmente em relação aos protozoários que divergiram na base da árvore filogenética. A exceção encontrada foi Sub2/UAP56, uma proteína relacionada com exportação de mRNAs em fungos e humanos, respectivamente, que se encontra altamente conservada ao longo da filogenia de eucariotos. Em *T. cruzi* foi observado que TcSub2 é uma proteína nuclear, relacionada com alguns sítios de transcrição por RNA Polimerase II e essencial, não sendo possível obter o nocaute duplo do respectivo gene. Em *T. brucei* e *T. gondii* foi observado que a diminuição gradual das proteínas em questão, TbSub2 e TgUAP56 respectivamente, resulta em acúmulo de mRNAs poliadenilados nos núcleos afetando a exportação dos mRNAs e resultando em morte celular. A partir da observação da participação de proteínas ortólogas de Sub2 na exportação de mRNAs em parasitas, a outra parte deste estudo foi identificar proteínas associadas a TcSub2 para iniciar o entendimento sobre esta via bem como verificar a presença de proteínas específicas e provavelmente essenciais de parasitas para futuros estudos funcionais.

ABSTRACT

Chagas disease and toxoplasmosis are public health issues in Brazil and understanding in detail the biology of its causative agents, *T. cruzi* and *T. gondii* respectively, could generate relevant information to develop effective treatment(s) against these diseases. The search for new chemotherapy targets against parasitic disease faces a number of challenges as the lack of tools that enable the study of essential gene function in parasites. Taking this into consideration, the aim of this study was to develop reverse genetic tools to identify and analyse essential proteins in *T. cruzi* and *T. gondii*, in particular those involved in nucleocytoplasmic transport of RNAs, specifically in mRNA export. *T. gondii* has served as model organism because it has an ample wide repertoire of reverse genetic tools. However, transferring these tools to other parasites, i.e. Plasmodium or Trypanosoma, has not been a straightforward task. The hypothesis of this study was based on two main facts: a) the transport of mRNA to the site of translation, within the context of key molecular events, is an essential step in the selection of genes to be expressed in an eukaryotic cell and b) the lack of publications in this topic in parasites proving to be a subject of basic research in parasitology that needs to be explored. The data obtained in this study showed that most proteins related to the export of mRNAs are not well conserved throughout eukaryotic phylogeny, especially eukaryotes that diverged at the base of the phylogenetic tree, the parasites. An exception to this observation was Sub2/UAP56, a protein directly related to export of mRNAs in yeast and human respectively, which is highly conserved throughout eukaryotic phylogeny. In *T. cruzi*, TcSub2 is an essential nuclear protein, associated with a number of sites of RNA polymerase II transcription and, a double knockout of the gene is lethal. It has been observed in *T. brucei* and *T. gondii* that the knockdown of TbSub2 and TgUAP56 respectively, results in nuclear accumulation of polyadenylated mRNAs, effectively reducing mRNA export and leading to cell death. Extending these observations that highlight the important role of Sub2 orthologues in mRNA export, we aim to identify proteins partners of TcSub2 to better understand this pathway in *T. cruzi* and verify the presence of specific and likely essential proteins for future studies in this parasite.

1 Introdução

1.1 *Trypanosoma cruzi* e a doença de Chagas

T. cruzi é um protozoário flagelado pertencente ao reino Protista, infra-reino Excavata, subreino Protozoa, incluído no filo Sarcomastigophora, subfilo Mastigophora, classe Zoomastigophora e ordem Kinetoplastida. Esta ordem abrange protozoários flagelados que possuem uma região especializada conhecida como cinetoplasto, onde reside o DNA mitocondrial, também denominado de kDNA (SHAPIRO e ENGLUND, 1995; TELLERIA *et al.*, 2006). É na ordem Kinetoplastida que está incluída a família Trypanosomatidae, na qual está presente o gênero *Trypanosoma* (LEVINE *et al.*, 1980). Este gênero constitui um dos mais importantes da família Trypanosomatidae, pois além de incluir *T. cruzi*, agente etiológico da doença de Chagas (CHAGAS, 1909), inclui outras duas subespécies de *Trypanosoma brucei* (*T. brucei*) que ocasionam doenças em humanos: *T. brucei rhodesiense* e *T. brucei gambiense*, ambos agentes da doença do sono. Adicionalmente, ainda há a subespécie *T. brucei brucei*, causadora da enfermidade em animais conhecida como nagana (HIDE, 1999).

A doença de Chagas, ou tripanossomíase americana, ocorre principalmente nas Américas, desde o sul dos Estados Unidos até o sul da Argentina e Chile. Segundo a Organização Mundial da Saúde, há aproximadamente 7 a 8 milhões de indivíduos infectados no mundo, sendo a maioria na América Latina (WHO - *Bulletin of the World Health Organization*, 2013). Nesta doença duas fases distintas podem ser observadas: aguda e crônica. A fase aguda inicia-se com a formação de um edema no local de entrada do parasita (chagoma de inoculação) que tende a desaparecer durante o curso da infecção. Este primeiro momento é seguido por um período pré-patente, de tempo/ variável, no qual os parasitas se multiplicam de forma exponencial (KROPF, AZEVEDO e FERREIRA, 2000). Quando da presença de sintomas, é caracterizada por alta parasitemia e intenso parasitismo tecidual, o que resulta em mal-estar, cefaléia, edema subcutâneo, disfunção cardíaca com miocardite focal, hepatomegalia, esplenomegalia e conjuntivite unilateral (sinal de Romana) (ANDRADE, 1999). A letalidade pode estar relacionada à meningoencefalite e insuficiência cardíaca (RASSI, RASSI JÚNIOR e RASSI, 2000). Quando a manifestação aguda diminui, segue um período de infecção latente onde

eventualmente pode ocorrer miocardite crônica acompanhada de cardiomegalia, arritmia, dispneia e edema periférico sem evidências ou com rápida retomada da multiplicação do parasita (ANSELMINI *et al.*, 1966; LARANJA e ANDRADE, 1980). Os casos agudos podem evoluir para a fase crônica com sintomas (forma determinada), onde os indivíduos doentes apresentam distúrbios cardíacos e digestivos como megaesôfago e megacólon, ou com ausência de sintomas (forma indeterminada) (DIAS, 2000).

T. cruzi é um microrganismo heteroxênico, com ciclo biológico alternado entre hospedeiros vertebrados (mamíferos) e invertebrados (triatomíneos). Existem diferentes formas de transmissão da doença, dentre as quais se destaca a via vetorial que consiste na transmissão do parasita por meio das excretas do inseto vetor hematófago pertencente à ordem Hemiptera, família Reduviidae, subfamília Triatominae. Há aproximadamente 123 espécies conhecidas de triatomíneos distribuídas em 15 gêneros, sendo que alguns dos gêneros - *Triatoma*, *Panstrongylus* e *Rhodnius* - estão relacionados com a transmissão natural da doença (DIAS *et al.*, 1956). A principal característica biológica dos triatomíneos é que são obrigatoriamente hematófagos e necessitam do repasto sanguíneo para completar o desenvolvimento (DIAS, 2000). O parasita pode ainda ser transmitido aos mamíferos por transfusão sanguínea ou via oral e, com menos frequência, por via congênita, acidentes de laboratório e transplantes de órgãos (BELTRAO HDE *et al.*, 2009; STEINDEL *et al.*, 2008; TANOWITZ *et al.*, 1992). Praticamente todo tipo de célula nucleada do hospedeiro mamífero pode ser parasitada considerando o tropismo das diferentes cepas (LENZI *et al.*, 1996). Outra característica deste parasita é apresentar diferentes morfologias durante o ciclo biológico. A classificação de suas formas baseia-se também na posição do cinetoplasto em relação ao núcleo e do local de onde emerge o flagelo (HOARE, 1971). Por exemplo, o flagelo emergindo na extremidade posterior que percorre todo o corpo formando uma membrana ondulante é característica da morfologia tripomastigota. Os tripomastigotas são alongados, possuem cinetoplasto posterior ao núcleo e são as formas infectivas extracelulares, não multiplicativas, encontradas no sangue de hospedeiros vertebrados (tripomastigotas sanguíneos) ou então na porção distal do intestino do inseto vetor (tripomastigotas metacíclicos). As formas amastigotas são arredondadas ou ovóides e apresentam flagelo que não se exterioriza. A

multiplicação destas formas ocorre por fissão binária no interior de células do hospedeiro vertebrado. Já no hospedeiro invertebrado, os epimastigotas são as formas multiplicativas extracelulares, com corpo alongado e cinetoplasto anterior bem próximo ao núcleo (DE SOUZA, 1999). A forma infectante para o hospedeiro mamífero no ciclo de transmissão natural é a tripomastigota metacíclica que está presente nas porções distais do intestino de triatomíneos infectados e pode ser eliminada durante o repasto sanguíneo junto às fezes e urina. A penetração dos tripomastigotas metacíclicos pode ocorrer através de mucosas ou descontinuidades da pele de mamíferos, seguido de internalização celular através de pseudópodos das células hospedeiras onde o parasita pode permanecer algumas horas no interior do vacúolo parasitóforo (DE SOUZA, 2002; GARCIA e AZAMBUJA, 1991; TEIXEIRA, NASCIMENTO e STURM, 2006). A fusão desse vacúolo com os lisossomos leva a um aumento da acidez no interior desta estrutura onde se encontra o parasita, permitindo o processo de diferenciação na forma amastigota. Paralelamente, ocorre a degradação gradual da membrana vacuolar pela ação de uma enzima secretada pelo parasita, a TcTOX, cuja molécula é semelhante à porina (ANDREWS, 1993). Uma vez no citoplasma das células, os amastigotas se dividem sucessivamente por fissão binária e iniciam o processo de diferenciação em tripomastigotas sanguíneos (MEYER e DE SOUZA, 1976). Depois de cinco dias, em média, as células hospedeiras são rompidas havendo liberação de tripomastigotas que invadem novas células ou se disseminam pela circulação sanguínea para outros órgãos. Os triatomíneos se infectam através da ingestão destas formas sanguíneas, dando continuidade ao ciclo de infecção (DE SOUZA, 2002). As formas sanguíneas ingeridas pelo triatomíneo se diferenciam em epimastigotas/esferomastigotas no estômago - porção anterior do trato digestivo e se dividem por fissão binária. Uma vez estabelecidas, as formas epimastigotas aderem às membranas perimicrovilares do intestino - porção distal do trato digestivo - diferenciando-se em formas tripomastigotas metacíclicas, que posteriormente são liberadas junto com fezes do triatomíneo (BONALDO *et al.*, 1988; GONZALEZ *et al.*, 1999). Esse processo de diferenciação celular de formas epimastigotas para tripomastigotas metacíclicas, denominado de metaciclogênese, é uma etapa crucial do ciclo biológico do parasita e pode ocorrer *in vitro* em meio LIT ou em outros meios de cultura (CAMARGO, 1964; CASTELLANI, RIBEIRO e FERNANDES, 1967; LANAR, 1979; RONDINELLI

et al., 1988). Nesses meios, a diferenciação ocasionalmente é acompanhada por multiplicação celular e por isso, Contreras e colaboradores (CONTRERAS *et al.*, 1985) desenvolveram condições para metaciclogênese *in vitro* na ausência de multiplicação celular. Por esse método, epimastigotas previamente cultivados em meio LIT são incubados em uma solução ligeiramente hipertônica que se assemelha às condições iônicas encontradas na urina do inseto vetor (TAU). Devido à carência de nutrientes do TAU, esta etapa promove um estresse nutricional e desequilíbrio metabólico que afetam o programa de expressão gênica do parasita (GOLDENBERG *et al.*, 1984). Posteriormente, esses protozoários são incubados em meio de diferenciação TAU3AAG (TAU suplementado com L-prolina, L-ácido glutâmico, L-ácido aspártico e glicose), permitindo a diferenciação em tripomastigotas metacíclicos que são liberados no sobrenadante da cultura (GOLDENBERG *et al.*, 1987). Há evidências de que a expressão de genes específicos de tripomastigotas precede as alterações morfológicas deste estágio e as primeiras modificações na expressão de genes ocorrem nas primeiras 24 horas. Com isso, a metaciclogênese *in vitro* pode ser considerada um modelo de estudo para caracterização de genes diferencialmente expressos (AVILA *et al.*, 2001; CONTRERAS, MOREL e GOLDENBERG, 1985).

1.1.1 Expressão e regulação gênica de tripanossomatídeos

A maturação dos transcritos em parasitas da ordem Kinetoplastida, entre eles os tripanossomatídeos, difere da maioria dos eucariotos (CLAYTON, 2002). As regiões codificadoras estão ordenadas em grupos de uma mesma fita de DNA e na maioria dos casos não apresentam interrupções por íntrons, com exceção dos genes que codificam para a enzima poli-A polimerase, relatado em *T. cruzi* e *T. brucei*, e uma RNA helicase putativa de *T. brucei* (JAE *et al.*, 2010; MAIR *et al.*, 2000). A transcrição resulta na formação de mRNAs denominados policistrônicos que possuem a informação de diversos genes, os quais não possuem necessariamente funções relacionadas (JOHNSON, KOOTER e BORST, 1987; TEIXEIRA e DAROCHA, 2003). Uma peculiaridade neste processo é que genes de um mesmo transcrito policistrônico podem ser expressos em níveis diferentes, ressaltando a

importância de mecanismos pós-transcricionais no controle da expressão gênica (revisão por (CLAYTON e SHAPIRA, 2007; VANHAMME e PAYS, 1995)).

Outra peculiaridade inclui a diferença observada entre as RNA polimerases dos tripanossomatídeos em relação às de eucariotos superiores que são as alterações na extensão C-terminal da RNA Polimerase II (ausência das repetições em *tandem* de heptapeptídeos nos tripanossomatídeos). Há também a susceptibilidade das RNA Polimerases dos tripanossomatídeos a diferentes concentrações de inibidores como a α -amanitina, o que auxilia no estudo funcional destas enzimas *T. brucei* e *T. cruzi*, por exemplo (CAMPBELL, THOMAS e STURM, 2003; DOSSIN FDE e SCHENKMAN, 2005; LAUFER *et al.*, 1999). Anteriormente ao início da tradução, os transcritos policistrônicos requerem processamento para gerar RNAs mensageiros (mRNAs) maduros e individualizados referentes aos respectivos genes. Neste caso, o processamento é decorrente de um evento denominado *trans-splicing* que adiciona uma sequência conservada na extremidade 5' de cada mRNA e da adição de uma cauda de adeninas (poli-A) na extremidade 3' (LIANG *et al.*, 2003). O *trans-splicing* foi inicialmente caracterizado em tripanossomatídeos com estudos sobre glicoproteínas variáveis de superfície (VSGs) de *T. brucei*, quando foi observado que os mRNAs apresentavam aproximadamente 39 nucleotídeos na extremidade 5', denominada então de sequência líder - *spliced leader* (SL), ou mini-exon (BOOTHROYD e CROSS, 1982; NELSON *et al.*, 1983) Posteriormente, a presença de sequências homólogas de SL foram observadas em transcritos de outros tripanossomatídeos como *T. cruzi* e *Leptomonas collosoma* (DE LANGE *et al.*, 1984; NELSON *et al.*, 1983). O SL origina-se de um transcrito de aproximadamente 141 nucleotídeos e é adicionado na porção 5' da sequência codificante presente nos RNAs policistrônicos precursores, originando unidades monocistrônicas que serão traduzidas (VAN DER PLOEG *et al.*, 1982). Como pode ser observado na figura 1, uma porção rica em polipirimidinas, crucial para o processo de *trans-splicing*, está localizada na região intergênica que precede o sítio de adição de SL do pré-mRNA (HUG *et al.*, 1994). A adição de SL é direcionada pela região aceptora de SL-RNA que possui o dinucleotídeo AG na região 3' da região do RNA policistrônico primário que será removida (LEBOWITZ *et al.*, 1993; MATTHEWS, TSCHUDI e ULLU, 1994). Sabe-se também que a conservação da distância entre a região aceptora - 3' AG - e a região de poliadenilação é importante

para que o processamento ocorra corretamente, revelando assim o papel de regiões intergênicas na regulação do *trans-splicing* (revisto por (TEIXEIRA, 1998)) e consequentemente na expressão gênica nestes organismos.

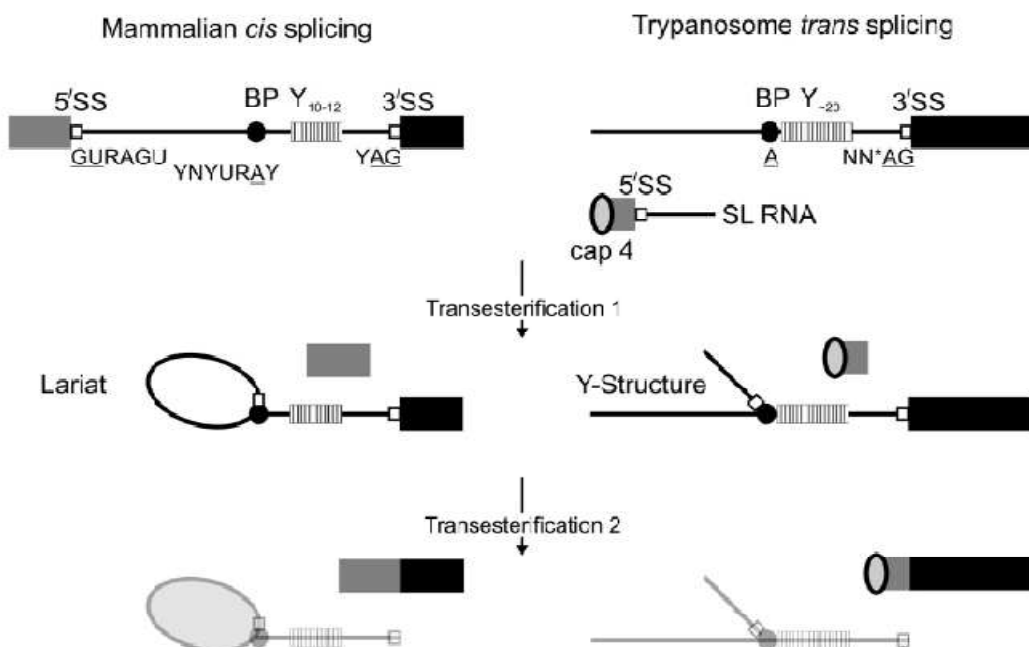


Figura 1. Esquema de processamento de mRNAs em mamíferos por cis-splicing e em tripanossomatídeos por trans-splicing. O éxon *upstream* e *spliced leader* (SL) estão representados em forma de retângulo cinza. O éxon *downstream* e o gene de tripanossomatídeo estão representados em forma de retângulos pretos. Sítios de *splicing* (SSs) estão representados por pequenos retângulos sem preenchimento, *Branch points* (BPs) por círculos pretos, trato de pirimidina pelo retângulo listrado e a estrutura *cap 4* de SL em forma elíptica. Sequências conservadas estão demonstradas abaixo do desenho, e os resíduos que não variam estão sublinhados. Enquanto em mamíferos 5'SSs, BPs e 3'SSs exibem conservação parcial de sequência (R, purina; Y, pirimidina; N, qualquer base), não há uma sequência óbvia em BPs de tripanossomatídeos (LUCKE *et al.*, 2005) e 3'SSs, embora tenham mostrado para este último a presença do dinucleotídeo AC (*) precedendo o dinucleotídeo AG reduz a eficiência de *splicing* ao menos que um dinucleotídeo AG esteja presente na região 5' UTR (SIEGEL, TAN e CROSS, 2005). Parece que o trato de polipirimidina torna-se mais importante quando a sequência consenso está faltando. Fungos possuem sequências SS e BPs altamente conservada e alguns introns não possuem trato de polipirimidina. As sequências parcialmente conservadas em mamíferos requerem um trato de polipirimidina pequeno, por volta de 10 a 12 resíduos (Y10-12), enquanto em tripanossomatídeos o trato é maior (Y20) e é uma sequência essencialmente importante para o *splicing* eficiente e tipicamente começa *downstream* ao BP (LUCKE *et al.*, 2005; SIEGEL, TAN e CROSS, 2005). Depois da primeira reação de transesterificação, estas estruturas intrônicas são desramificadas e rapidamente degradadas. FONTE: GUNZL, 2010.

As características peculiares na estrutura de genes e transcrição de mRNAs fazem com que a regulação da expressão gênica em tripanossomatídeos seja praticamente, se não exclusivamente pós-transcricional, diferente da maioria dos eucariotos superiores, onde é evidente a importância de um controle transcricional complementando os controles em nível pós-transcricional. Neste último caso,

proteínas de ligação a RNA (RBPs) possuem um papel principal (revisto por (DREYFUSS, KIM e KATAOKA, 2002)) e originam complexos ribonucleoprotéicos (mRNPs) que direcionam o mRNA ao citoplasma para tradução ou degradação (GLISOVIC *et al.*, 2008; KEENE, 2001). Além disso, as RBPs participam de diferentes eventos moleculares como *splicing*, poliadenilação, exportação, localização, tradução e estabilidade de mRNAs (revisto por (GLISOVIC *et al.*, 2008)). Os tripanossomatídeos também possuem proteínas de ligação a RNA como protagonistas na regulação da expressão gênica (CARO *et al.*, 2006; D'ORSO, DE GAUDENZI e FRASCH, 2003; DE GAUDENZI, FRASCH e CLAYTON, 2005), como por exemplo, PABP1, TcUBP1, TcSR, p34 e p37 (revisto por (D'ORSO, DE GAUDENZI e FRASCH, 2003)). De modo geral, estas proteínas reconhecem sequências presentes nas regiões não traduzidas localizadas na extremidade 3' de mRNAs denominadas de 3' UTR.

Ensaio com gene repórter luciferase mostraram que sequência 3' UTR de diferentes genes estágios-específicos, como gp72, gp85 e gp82, podem alterar os níveis de expressão de luciferase. Por exemplo, as inserções de 3' UTR dos genes estágios-específicos gp72 e gp85 resultaram em diminuição dos níveis de transcritos em epimastigotas, promovendo a instabilidade ou maturação ineficiente do mRNA (NOZAKI e CROSS, 1995). Em mamíferos, a estabilidade de mRNAs é regulada por um elemento desestabilizador rico em AU na região 3' UTR, denominada ARE – *AU rich element* (CHEN e SHYU, 1995). Da mesma forma, elementos estáveis ricos em U - UREs, estrutural e funcionalmente relacionados aos AREs, foram encontrados na região 3' UTR de diferentes transcritos de *T. brucei*, que são estáveis na forma procíclica mas instáveis na forma sanguínea (revisto por (CLAYTON e SHAPIRA, 2007)), bem como em alguns mRNAs de *T. cruzi* (D'ORSO, DE GAUDENZI e FRASCH, 2003) e de *Leishmania* (HAILE e PAPADOPOULOU, 2007).

Existem várias evidências para o papel de elementos presentes em moléculas de mRNA e de RBPs no controle da estabilidade e tradução de mRNA em tripanossomatídeos, contudo, pouco foi publicado até o momento quanto a maquinaria e mecanismos envolvidos com outros eventos pós-transcricionais, como por exemplo, a exportação de mRNAs do núcleo para o citoplasma.

1.2 *Toxoplasma gondii* e a doença Toxoplasmose

Toxoplasma gondii (*T. gondii*) é um protozoário pertencente ao reino protista, infra-reino Alveolata, subreino Protozoa, filo Apicomplexa, classe Conoidasida e ordem Eucoccidiorida. O Filo Apicomplexa consiste em protozoários de grande importância clínica, como por exemplo os gêneros *Plasmodium*, *Toxoplasma*, *Eimeria*, *Sarcocystis*, *Teileria* e *Babesia*. Os parasitas pertencentes a este Filo são intracelulares obrigatórios e possui uma estrutura de morfologia única, denominada complexo apical (HSU, CAMPBELL e LEVINE, 1973; LEVINE, 1988). Este complexo é requerido para adesão e invasão da célula do hospedeiro e é geralmente composto pelos anéis polares apicais, conóide, micronemas e roptrias (MORRISSETTE *et al.*, 1994; NICHOLS e CHIAPPINO, 1987). A grande maioria dos parasitas pertencentes a este Filo possui a organela denominada de apicoplasto, a qual foi adquirida por endossimbiose a partir de uma alga vermelha (HARPER, WAANDERS e KEELING, 2005; KOHLER *et al.*, 1997) e está relacionada com a síntese de heme, isoprenóides e lípidos neutros (MARECHAL e CESBRON-DELAUW, 2001).

T. gondii foi descrito primeiramente em roedores por Nicolle e Manceaux na Tunísia, África (Nicolle e Manceaux, 1908). O nome derivou da morfologia: *toxos* = arco, *plasma* = forma, e do hospedeiro: *gondii*, por ter sido descoberto em um roedor da região do Gundi. Em 1967 foi demonstrado que o *T. gondii* pode ser contraído pela ingestão de oocistos eliminados nas fezes dos felinos. A descoberta destes como hospedeiros definitivos, em 1970, levou ao entendimento do ciclo biológico deste parasita (DUBEY, 2008). *T. gondii* é um microrganismo heteroxênico em que a fase sexual ocorre no hospedeiro definitivo, o gato, e a reprodução assexuada ocorre em qualquer animal vertebrado de sangue quente, havendo uma enorme gama de possíveis hospedeiros (HILL, CHIRUKANDOTH e DUBEY, 2005). Existem duas formas distintas no hospedeiro: as formas taquizoítas, que replicam rapidamente e provocam infecção aguda em pacientes imunocomprometidos, e as formas bradizoítas que replicam lentamente, formando cistos que escapam do sistema imune e podem persistir ao longo da vida do hospedeiro, caracterizando a fase crônica da doença. É importante ressaltar que mesmo na fase crônica existe

possibilidade de transmissão para outro hospedeiro via ingestão oral, não necessariamente passando pelo ciclo sexual, bem como a possibilidade de reativação para formas taquizoítas que podem invadir novas células (DUBEY, 1998; DUBEY, 2009; SU *et al.*, 2003).

Toxoplasma gondi é um patógeno oportunista e agente etiológico da toxoplasmose, de distribuição mundial. Ele está relacionado com aproximadamente 30% de infecções humanas mais comuns, sendo a toxoplasmose a principal causa de morte que é atribuída a doenças transmitidas por alimentos (HILL, CHIRUKANDOTH e DUBEY, 2005). A infecção primária é geralmente subclínica, mas alguns casos de linfadenopatia cervical, toxoplasmose ocular (OT), encefalite ou sistêmicos podem ser observados em pacientes imunodeprimidos, especialmente aqueles que possuem HIV/AIDS. A infecção adquirida durante a gravidez também pode causar graves danos ao feto (MONTROYA e LIESENFELD, 2004; ROBERT-GANGNEUX e DARDE, 2012), embora dados epidemiológicos mostrem que muitos casos de OTs são adquiridos após o nascimento (GILBERT e STANFORD, 2000; HOLLAND, 1999; HOLLAND, 2003). *T.gondii* é um parasita versátil e a prevalência da infecção varia entre os países, áreas geográficas e grupos étnicos que vivem dentro de uma região específica (TENTER, HECKEROTH e WEISS, 2000). No Brasil, os altos níveis de doença ocular são associadas a toxoplasmose e os maiores valores de prevalência sorológica de infecção por *T. gondii* ocorre em estados do norte e do sul, incluindo 50-80% da população adulta (BAHIA-OLIVEIRA *et al.*, 2003; SILVEIRA *et al.*, 2001). Além disso, os sintomas da doença são mais agressivos para pacientes brasileiros, em comparação com os pacientes dos Estados Unidos e Europa, provavelmente devido a patogenicidade da cepa de infecção e/ou a resposta imunológica do hospedeiro (DUBEY *et al.*, 2012; MATOS *et al.*, 2013). Estudos tem mostrado que a cepa do Tipo I pode estar relacionada com OTs em pacientes brasileiros (VALLOCHI *et al.*, 2005), e a maioria das cepas brasileiras tem diferentes genótipos de linhagens clonais que predominam na América do Norte e na Europa (KHAN *et al.*, 2006), o que pode estar relacionado aos casos mais severos de OT entre crianças brasileiras com toxoplasmose congênita em comparação com crianças européias (GILBERT *et al.*, 2008). Felizmente, a infecção aguda podem ser tratada com pirimetamina e sulfadiazina

como agentes terapêuticos (MONTROYA e LIESENFELD, 2004), no entanto nenhum tratamento contra a infecção crônica (cistos teciduais) se encontra disponível no momento. Estudos de fatores essenciais e os mecanismos especializados para a sobrevivência do parasita podem levar a novos alvos para o tratamento eficaz de doenças parasitárias.

1.2.1 Regulação da expressão gênica e abordagens de genética reversa em *T. gondii*

O desenvolvimento é um processo altamente regulado em organismos eucarióticos e a diferenciação de *Toxoplasma* não é exceção, envolvendo mecanismos reguladores que operam na transcrição, bem como no nível pós-transcricional (SULLIVAN JR, SMITH e JOYCE, 2009). Em nível transcricional, os mecanismos de regulação gênica requerem fatores de transcrição, epigenética e modificações da cromatina (BEHNKE *et al.*, 2008; BOUGDOUR *et al.*, 2010; RADKE *et al.*, 2005). Em nível pós-transcricional, a maioria dos estudos envolve o controle de tradução com ênfase particular em resposta ao estresse celular encontrado durante o ciclo de vida do parasita (BOUGDOUR *et al.*, 2010; VONLAUFEN *et al.*, 2008; ZHANG *et al.*, 2013). Outros mecanismos moleculares podem estar envolvidos na regulação da expressão de genes durante o desenvolvimento de *Toxoplasma*, tais como *splicing* ou estabilidade do mRNA, mas os estudos são atualmente escassos. Recentemente tem surgido alguns trabalhos que mostram o papel dos fatores de *splicing* de mRNAs, grânulos de RNAs não-codificantes que regulam o ciclo celular, a viabilidade extracelular e diferenciação dos parasitas, mostrando a importância do nível pós-transcricional no controle de expressão gênica em *Toxoplasma* (LIRUSSI e MATRAJT, 2011; PATIL *et al.*, 2012; SUVOROVA *et al.*, 2013).

As metodologias para estudos de genes em *T. gondii* são bem estabelecidas e devido às vantagens experimentais este parasita tem se tornado um importante modelo para o estudo da biologia de espécies pertencentes ao Filo Apicomplexa (KIM e WEISS, 2004; MEISSNER *et al.*, 2007). Um sistema bem estabelecido é a superexpressão de proteínas através do sistema DD (*Destabilization Domain*). Este sistema induzido foi inicialmente desenvolvido para células de mamíferos e é baseado na fusão da proteína de interesse ao domínio de desestabilização

modificado de FKBP12 (ddFKBP12), resultando em degradação da proteína fusionada pela via de proteossomo em situações normais (BANASZYNSKI *et al.*, 2006). No entanto, existe uma molécula permeável, denominada de Shield-1 (Shld-1), que é análoga de rapamicina. Shld-1 se liga a este domínio possibilitando o enovelamento correto e estabilizado, permitindo assim a restauração dos níveis da proteína fusionada de modo condicionado. O esquema desse sistema baseado em ddFKBP12 pode ser observado na figura 2.

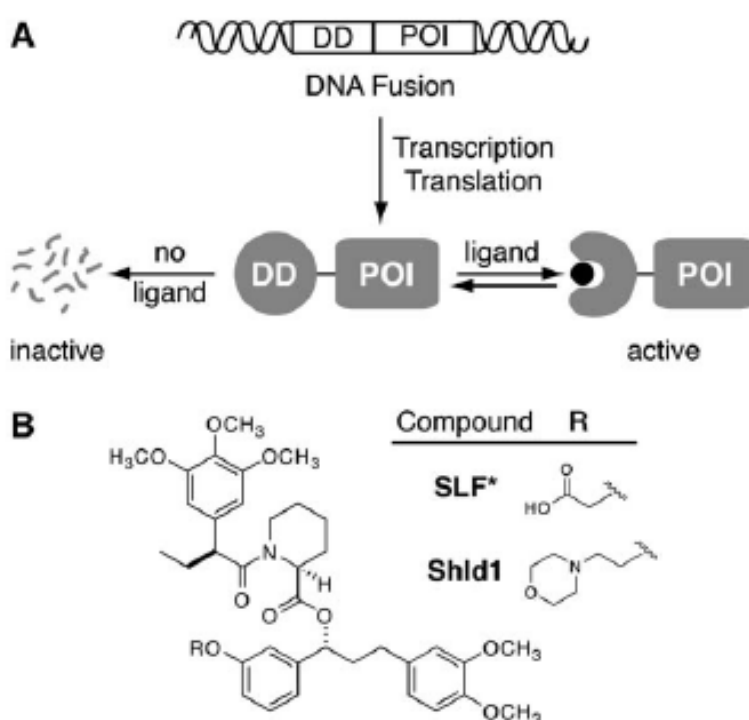


Figura 2. Sistema baseado na fusão de ddFKBP12 a uma proteína de interesse para controle de estabilidade. A. Fusão do domínio de desestabilização de FKBP12 (ddFKBP12 - DD) à proteína de interesse (POI – *protein of interest*) resulta em degradação da proteína fusionada quando na ausência do ligante. A adição do ligante protege a degradação da proteína fusionada. B. Ligantes sintéticos para FKBP12 mutada F36 V. SLF*: Primeiro composto a ser utilizado nos testes de estabilização de ddFKBP12. Shld-1: SLF* modificado e otimizado para estabilização de ddFKBP12 - Shield-1. FONTE: BANASZYNSKI *et al.*, 2006.

Esse sistema permite a regulação dos níveis de proteína, de modo rápido e reversível em parasitas do Filo Apicomplexa (ARMSTRONG e GOLDBERG, 2007; HERM-GOTZ *et al.*, 2007), como mostrado por exemplo na figura 3 em *T. gondii*. Este sistema é considerado vantajoso em relação aos sistemas já conhecidos principalmente por não depender de controle da expressão gênica de um elemento

promotor específico, como no caso do sistema de tetraciclina (MEISSNER, SCHLUTER e SOLDATI, 2002), evitando complicações resultantes de um promotor inapropriado, bem como pela relação diretamente proporcional entre concentração da droga utilizada e níveis de proteínas que podem ser obtidos. O sistema DD vem sendo recentemente empregado para estudo funcional de proteínas em *T. cruzi*, *Leishmania major* e *T. gondii* para melhor compreensão da biogênese de organelas, vias secretoras e fatores de virulência nestes parasitas (KREMER *et al.*, 2013; MA, WEISS e HUANG, 2011; MADEIRA DA SILVA *et al.*, 2009; PIEPERHOFF *et al.*, 2013).

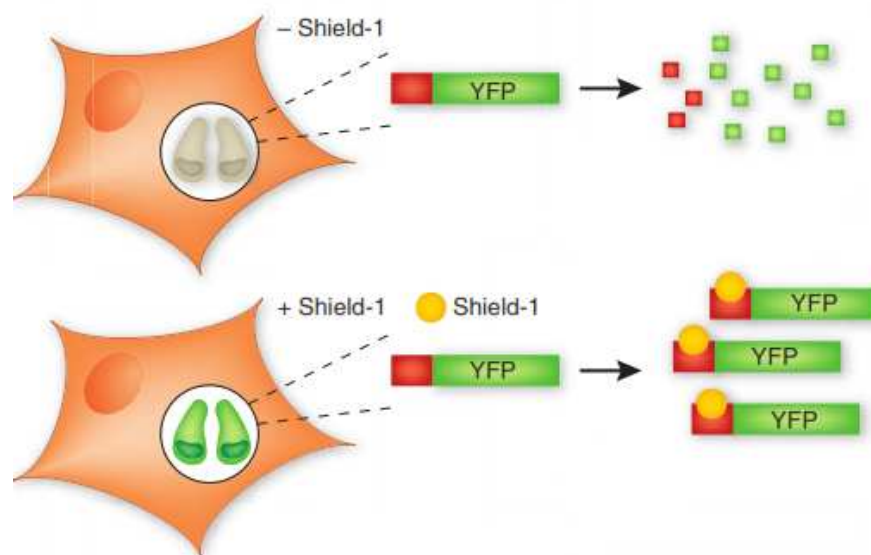


Figura 3. Expressão gênica condicional através de desestabilização de proteínas A célula hospedeira abriga dois parasitas (cinza claro) dentro do vacúolo parasitóforo (branco). Os parasitas são transfectados com uma construção de DNA contendo a sequência que codifica para a proteína alvo (YFP – *yellow fluorescent protein*) fusionada ao domínio de desestabilização de FKBP12 (ddFKBP12) na porção N-terminal. Na ausência do ligante Shield-1, YFP-ddFKBP12 é degradada. Após a adição de Shield-1, YFP-ddFKBP12 é estabilizada e facilmente detectável. FONTE: STRIEPEN, 2007.

Outro sistema que foi recentemente estabelecido em *T. gondii* é o sistema de nocaute condicionado para a inativação de genes de modo regulado que é baseado na sistema de recombinase Cre/LoxP, originário em bacteriófagos P1 (JULLIEN *et al.*, 2007; JULLIEN *et al.*, 2003). Cre é uma recombinase que catalisa a recombinação sítio-específica entre duas sequências *LoxP* através de sucessivas etapas de quebras e ligação de DNA (VAN DUYNE, 2001), resultando em excisão

do gene de interesse que esteja flanqueado por elas. Estas sequências *LoxP* possuem 34 pares de bases e são introduzidas no genoma de forma artificial. Também são compostas por duas regiões de ligação à recombinase com 14 pares de bases (RBE – *recombinase binding element*) e organizadas como repetições invertidas tendo como centro uma região que permite a recombinação sítio-específica por Cre, como ilustrado na figura 4.

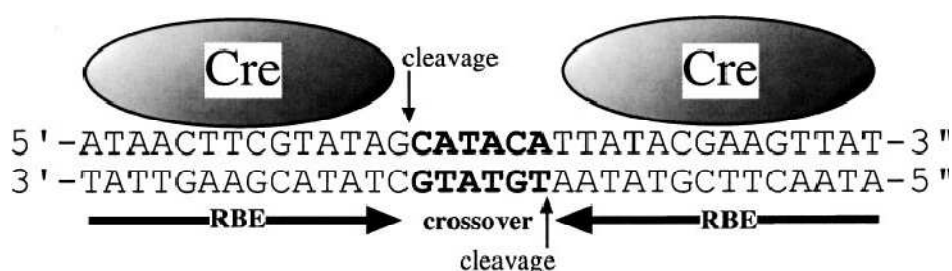


Figura 4. Ligação de Cre a sítios *LoxP*. Cada sequência de *LoxP* é composta por 2 elementos de 14 pares de bases que se ligam à recombinase (RBE – *recombinase binding element*), os quais são arranjados em forma de repetições invertidas em torno de uma região central de recombinação de 6 pares de bases, onde pode ocorrer a recombinação por recombinase Cre FONTE: HOESS e ABREMSKI, 1995.

Com o objetivo de diminuir o vazamento da atividade desta enzima, bem como tornar o sistema regulado, o sistema foi desenvolvido através da divisão da enzima em dois fragmentos inativos - DiCre (domínios N e C-terminal) fusionados a FKBP12 (FK506-binding protein) e FRB (*binding domain of the FKBP12 - rapamycin-associated protein*) respectivamente. Após a adição de rapamicina, e consequente associação de FKBP12 com o FRB, há a reconstituição da atividade enzimática de DiCre, tornando-a funcional e condicionada à presença desta droga (JULLIEN *et al.*, 2003). Esse sistema DiCre foi incorporado e funcional em *T. gondii* (ANDENMATTEN N *et al.*, 2013), sendo possível o *gene-swap*, ou seja, a retirada do gene de interesse flanqueado por *LoxP* no genoma haplóide do parasita após 24 horas da adição de rapamicina. Esta estratégia permite observar o fenótipo ao longo do tempo após a indução do nocaute, viabilizando o estudo da função de genes essenciais neste parasita (ANDENMATTEN *et al.*, 2013; BARGIERI *et al.*, 2013). O sistema DiCre em *T. gondii* está ilustrado na figura 5.

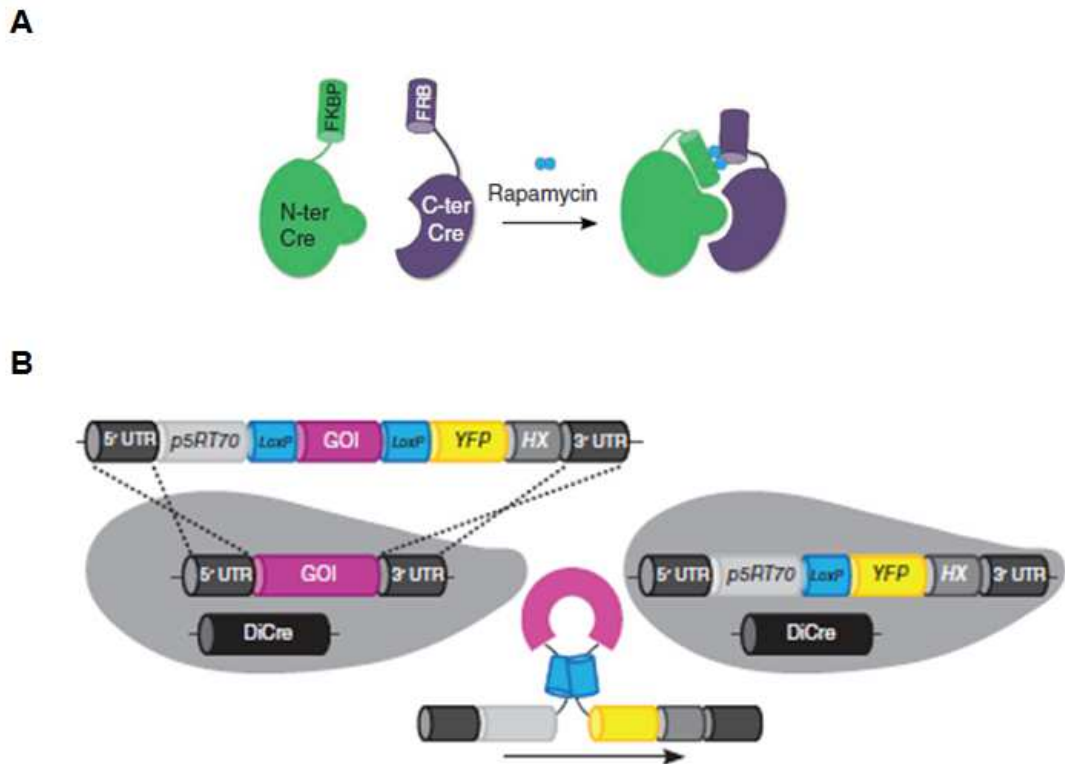


Figura 5. Sistema DiCre para nocaute condicionado em *T. gondii*. A. Neste sistema a porção N-terminal (N-ter: aminoácidos 1 a 59) de Cre recombinase está fusionada a FKBP12 e a porção C-terminal (C-ter: aminoácidos 60 a 343) está fusionada a FRB. Rapamicina dimeriza as subunidades de DiCre havendo a reconstituição da atividade DiCre recombinase. B. Na estratégia de *gene-swap* o gene de interesse, em rosa - GOI - *gene of interest* - é substituído pelo cassete indicado através de recombinação homóloga pelas regiões intergênicas 5'e 3'UTR (UTR - *untranslated region*). O cDNA de GOI, flanqueado por *LoxP*, que neste caso está sob regulação do promotor constitutivo *p5RT70* é seguido de gene-repórter *YFP* e marcador de seleção dos parasitas *HX*. Após a adição de rapamicina, há a excisão do cDNA de GOI, mediado por DiCre, e *YFP* passa a estar sob controle do promotor, produzindo parasitas nocaute fluorescentes. FONTE: ANDENMATTEN *et al.*, 2013.

1.3 Exportação de mRNAs em eucariotos: do núcleo ao citoplasma

Em células eucarióticas, a delimitação entre núcleo e citoplasma ocorre pela presença de uma dupla membrana lipídica que forma o envelope nuclear. Embora pareça uma estrutura estática, o envelope nuclear possui papéis importantes na organização da cromatina, regulação da expressão gênica e até mesmo na exportação de vírus em células de mamíferos ou complexos ribonucleoprotéicos em *D. melanogaster* através de um mecanismo não usual por brotamento (SPEESE *et al.*, 2012; VAN DE VOSSE *et al.*, 2011). No envelope nuclear estão presentes proteínas denominadas nucleoporinas, que formam os complexos de poro nuclear

(NPC) totalizando uma massa de aproximadamente 125 MDa, podendo ser compostos por até 100 proteínas diferentes (STOFFLER, FAHRENKROG e AEBI, 1999). A livre passagem é permitida a moléculas menores que 40 kDa, processo conhecido como difusão passiva, enquanto as macromoléculas atravessam com o auxílio de proteínas, caracterizando um processo seletivo com necessidade de sinais de transporte (YONEDA *et al.*, 2000).

Os mRNAs exportados comumente seguem para regiões de tradução no citoplasma, sendo um ponto de controle que garante o transporte dos transcritos corretamente processados e a síntese protéica em determinado local ou tempo dentro de uma célula ou organismo (DU, SCHMID e JANSEN, 2007). O modelo geral de transporte núcleo-citoplasma de RNAs tem a participação de proteínas da família das β -carioferinas (importinas e exportinas responsáveis pela importação e exportação, respectivamente) que utilizam energia fornecida por Ran-GTP para o transporte de cargas (STROM e WEIS, 2001).

A importina- β em fungos, denominada Kap95, é um fator de transporte que interage com o adaptador Kap60 - importina- α - pelo reconhecimento de sinais de localização nuclear (NLS – *nuclear localization signal*) presentes em proteínas que tem papel funcional no núcleo (POLLARD *et al.*, 1996). Outra representante da classe de importinas é a transportina 1 que reconhece a sequência denominada M9, e também está relacionada com a importação de macromoléculas para o núcleo (IZAURRALDE *et al.*, 1997). As exportinas CAS e CRM1 de fungos, proteínas ortólogas funcionais de CSE e *Exportin-1* (XPO1) de humanos respectivamente, são responsáveis principalmente pela exportação de proteínas e algumas espécies de RNAs. Geralmente, as exportinas interagem especificamente a algum tipo de RNA, apesar de ser observado também que a mesma exportina pode transportar tipos diferentes de RNAs. Em fungos, por exemplo, a exportação majoritária de tRNAs e miRNAs ocorre por exportinas diferentes: Los1 e Msn5, respectivamente, enquanto snRNAs, snoRNAs e rRNAs são exportados pela mesma exportina: CRM1 (KOHLENER e HURT, 2007; SHIBATA *et al.*, 2006; YAO *et al.*, 2007).

As proteínas que possuem afinidade a CRM1 geralmente possuem sinal de exportação nuclear (NES – *nuclear export signal*) rico em leucinas conhecido como motivo LRR, primeiramente descrito pelo grupo de Takahashi (TAKAHASHI, TAKAHASHI e PUTNAM, 1985). Este motivo apresenta um consenso de

aminoácidos com conservação na posição do aminoácido leucina: **LX₁₋₃**, **LX₂₋₃**, **LXL**, onde L é preferencialmente uma leucina e X um aminoácido qualquer (BOGERD *et al.*, 1996; IKUTA *et al.*, 1998). Esses motivos LRR também são encontrados em proteínas de exportação de mRNA, posicionados na porção N-terminal, concomitantemente a outro domínio de reconhecimento de RNA – RRM - de aproximadamente 90 aminoácidos (COOK *et al.*, 2007; DE GAUDENZI, D'ORSO e FRASCH, 2003).

Diferentes exportinas como CRM1, Los1 e Msn5 obedecem ao modelo geral de transporte de RNAs dependente da proteína Ran (revisto por (KOHLEER e HURT, 2007)). A proteína Ran é uma GTPase pequena e solúvel, presente tanto no núcleo quanto no citoplasma de todas as células eucarióticas e tem um papel crítico na exportação de RNAs: exportinas se ligam à carga junto com RanGTP no núcleo para formar o complexo ternário (RNA-exportina-RanGTP). Este complexo é então translocado para o citoplasma, onde se dissocia após hidrólise de GTP, ocorrendo a liberação do RNA. A proteína Rna1 de fungos (RanGAP em humanos) auxilia na hidrólise de GTP e liberação do RNA no citoplasma. O fator de transporte nuclear 2 (NTF2) carrega RanGDP para o núcleo, onde ocorre a mudança para RanGTP com auxílio da proteína Srm1 (RanGEF ou RCC1 em humanos) (COOK *et al.*, 2007; STEWART, 2000; TALCOTT e MOORE, 1999). O esquema desse transporte de macromoléculas mediado pelo ciclo de Ran pode ser observado na figura 6.

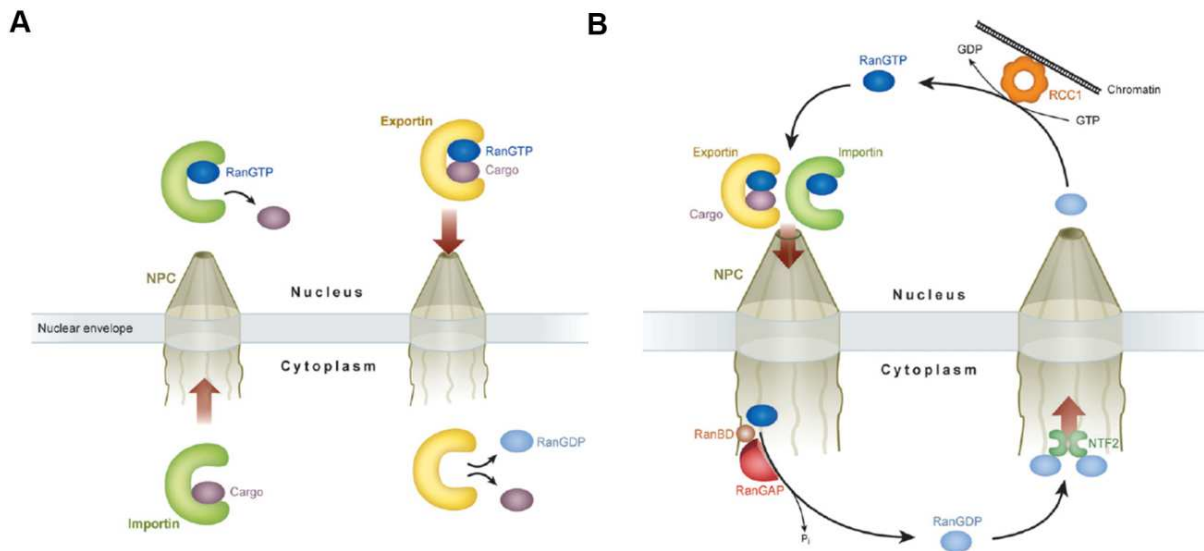


Figura 6. Desenho esquemático do transporte de macromoléculas pelo ciclo de Ran em humanos. Processos de exportação e importação mediados por β -carioferinas. O complexo de poro nuclear (NPC) mostra características de cesta nuclear e fibrilas citoplasmáticas. A. A importina se liga à carga no citoplasma e a libera através de ligação a RanGTP no núcleo (esquerda). A exportina se liga à carga e a RanGTP no núcleo e libera a carga no citoplasma após hidrólise de RanGTP em RanGDP (direita). B. A alta concentração de RanGDP no citoplasma é mantida por RanGAP (*GTPase-activating protein*), que está ligada a fibrilas citoplasmáticas do NPC. A alta concentração de RanGTP no núcleo é mantida pelo regulador de condensação de cromossomo 1 (RCC1), um fator de ligação à cromatina de troca de guaninas (RanGEF-*Guanine nucleotide Exchange Factor*) que age em RanGDP que é transportada para o núcleo através de NTF2. FONTE: COOK *et al.*, 2007.

Estudos em *Drosophila melanogaster*, humanos e fungos mostraram que CRM1 não é a via principal para a exportação de mRNAs, que na realidade essa via é uma exceção para o modelo geral dependente de RanGTP e carioferinas (HEROLD, TEIXEIRA e IZAURRALDE, 2003; KATAHIRA *et al.*, 1999; SCHUTZ *et al.*, 2006). Alguns trabalhos mostram que CRM1 pode estar envolvida no transporte especializado de um subconjunto de mRNAs, snRNAs ricos em uridina (UsnRNAs) (< 200 nucleotídeos) por recrutamento via proteína PHAX (OHNO *et al.*, 2000), rRNAs, RNAs que possuam sinal de reconhecimento (SRP), ou até mesmo na exportação de transcritos parcialmente processados em casos de mRNA viral (revisado por (CULLEN, 2003a; HUTTEN e KEHLENBACH, 2007)).

Citando leveduras como exemplo, a exportação da maioria dos mRNAs é dependente de um dímero que substitui a exportina, e considerado o receptor de exportação: Mex67-Mtr2 (TAP-p15, em humanos). Este dímero é responsável por receber o transcrito processado ainda no núcleo e carregá-lo para o citoplasma com o auxílio de diversas proteínas adaptadoras (KOHLER e HURT, 2007; SEGREF *et al.*,

1997; THAKURTA *et al.*, 2002). A comparação entre as diferentes vias de exportação de RNAs pode ser observada na figura 7: as vias dependentes de RanGTP e, no caso do transporte de mRNA, independente de RanGTP com atuação de Mex67-Mtr2/TAP-p15 ao invés de exportinas.

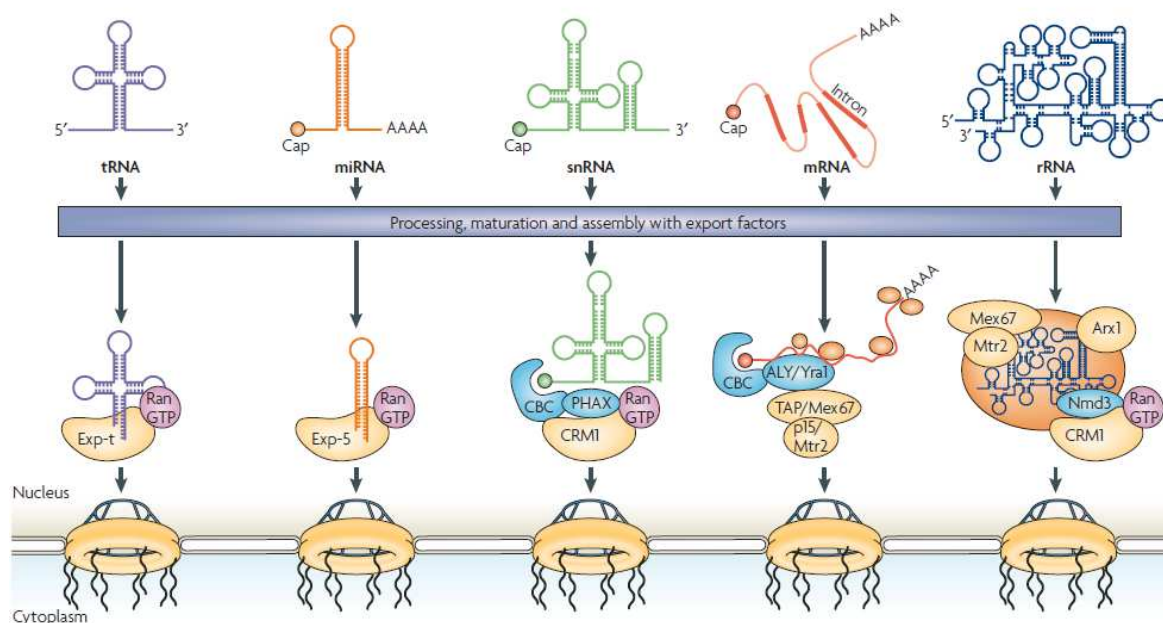


Figura 7. Diferentes vias de exportação de RNAs em eucariotos. As vias majoritárias de transporte estão ilustradas: RNA transportador (tRNA), microRNA (miRNA), pequenos RNAs nucleares (snRNA), RNA mensageiro (mRNA) e RNA Ribossomal (rRNA). Em cada caso, o transcrito primário está ilustrado, bem como a proteína transportadora correspondente após processamento dos respectivos RNA: Exp-t, Exp-5, CRM1, Mex67. Os adaptadores de exportação estão ilustrados em azul e receptores de exportação em amarelo. No caso de mRNA, ambos os nomes em humanos e fungos estão indicados, separados por “/”. Em humanos, os homólogos são: Aly, TAP e p15. Em fungos: Yra1, Mex67 e Mtr2, respectivamente. CBC, fator de ligação ao cap, exportina. FONTE: KÖHLER e HURT, 2007.

Enquanto em leveduras apenas um receptor de exportação de mRNAs (Mex67) é encontrado, em humanos há 4 proteínas distintas pertencentes à família NXF (NXF1-4), sendo NXF1 (ou Tap) a proteína ortóloga funcional de Mex67 e responsável pela exportação da maioria dos mRNAs. Já NXF3 parece mediar a exportação de mRNAs em parceria com a exportina-1 (CRM1) apenas em testículos (YANG *et al.*, 2001). Embora existam proteínas exclusivas de metazoários, a conservação encontrada entre as proteínas básicas da via de exportação em fungos, plantas e humanos indica que o desenvolvimento de sistemas mais complexos para esta via durante a evolução não alterou o modo fundamental para resolução de transporte de mRNAs nucleares. Uma diferença significativa na estrutura gênica entre fungos e metazoários é a complexidade em que 80% dos genes de *D.*

melanogaster possuem introns comparados com apenas 5% de *S. cerevisiae* (SAKHARKAR e KANGUEANE, 2004). Isso sugere que a maquinaria de *splicing* deve ter um papel crítico na regulação do processamento de mRNAs em metazoários, bem como a exportação, uma vez que esta deve ocorrer apenas para os mRNAs processados corretamente para o citoplasma (revisto por (WOLYNIAK e COLE, 2008)).

Em fungos, durante a transcrição do mRNA existe a participação do complexo THO, constituído pelas proteínas nucleares Hpr1, Tho2, Mft1 e Thp2 (CHAVEZ *et al.*, 2000), que interage com as proteínas Yra1, Tex1 e a RNA-helicase Sub2 originando o complexo TREX1 (*TR*anscription/*EX*port 1), relacionado ao processo de transcrição e exportação de mRNAs (JENSEN *et al.*, 2001; REED e CHENG, 2005; STRASSER *et al.*, 2002). Este processo também ocorre em metazoários, sendo o complexo THO, em humanos, constituído pelos homólogos THOC1 (Hpr1), THOC2 (Tho2) e outras proteínas não identificadas em fungos: THOC5, THOC6 e THOC7 (MASUDA *et al.*, 2005). Quanto ao complexo TREX1, há dois homólogos para Sub2: UAP56 e URH49 (KAPADIA *et al.*, 2006) e apenas um homólogo de Yra1, denominado de Aly (LUO *et al.*, 2001). Em humanos o complexo TREX1 é recrutado durante o *splicing* de mRNAs, em locais chamados *speckles*, onde estão presentes a RNA Polimerase II, fatores de remodelamento de cromatina e componentes de processamento de mRNAs (BROWN *et al.*, 2008; CUSTODIO *et al.*, 2004; DIAS *et al.*, 2010; KATAOKA *et al.*, 2000; KOTA *et al.*, 2008).

Durante a transcrição há o recrutamento de proteínas hnRNPs (*heterogeneous nuclear ribonucleoprotein*) e SRs. As hnRNPs são proteínas nucleares que interagem predominantemente com os introns de pré-mRNAs, empacotando-os em complexos hnRNPs (BEYER *et al.*, 1977; KUMAR e PEDERSON, 1975) de maneira similar ao empacotamento por histonas (WEIGHARDT, BIAMONTI e RIVA, 1996), gerando “nucleossomos de RNAs”. Aproximadamente 150 a 250 nucleotídeos do pré-mRNA se enrolam em torno de um tetrâmero de hnRNPC para então serem processados por *splicing* (HUANG *et al.*, 1994; MCAFEE *et al.*, 1996). Foi demonstrado recentemente em células de humanos que hnRNPCs podem funcionar como “régua molecular” e funcionam como proteína-chave para a exportação de RNAs via exportina-1 (CRM1) ou TAP, classificando o transcrito de RNA Polimerase II a ser exportado pelo tamanho

(MCCLOSKEY *et al.*, 2012). Há uma competição entre PHAX e hnRNPs pelo complexo de ligação ao *cap* (CBC). RNAs menores que 200 nucleotídeos (snRNAs) não conseguem ser estabilizados pelo tetrâmero de hnRNPs, permitindo assim o recrutamento de PHAX e conseqüentemente de CRM1. Já os mRNAs, maiores que 200 nucleotídeos, são estabilizados e impedidos de se ligar a PHAX, permitindo interação de Aly com CBC e conseqüentemente o dímero Tap-p15 para a exportação (MCCLOSKEY *et al.*, 2012).

As fosfoproteínas SRs possuem na região C-terminal um domínio característico rico em repetições de resíduos de serina (S) e arginina (R), sujeitas a fosforilação dinâmica (BEYER *et al.*, 1977; GRAVELEY, 2000; LONG e CACERES, 2009; REED e CHENG, 2005). São consideradas fatores de *splicing* que interagem com o pré-mRNA co-transcricionalmente quando hiperfosforiladas e recrutam a maquinaria de processamento. As proteínas SR mais estudadas são Npl3 e Gbp2, de fungos, e 9G8 e ASF/SF2, de humanos (HUANG e STEITZ, 2001). Durante o processamento de mRNAs, estas proteínas são defosforiladas e facilitam o trânsito pelo NPC através de interações com fatores de exportação pertencentes ao TREX1 (HUANG e STEITZ, 2005).

Para a exportação dos mRNAs processados, de modo simplificado em humanos e fungos, o complexo THO se associa primeiramente ao molde de DNA ativo e as proteínas Aly/Yra1 e UAP56/Sub2, formadoras do complexo TREX1, são recrutadas co-transcricionalmente junto com proteínas hnRNPs e SRs, interagindo com o transcrito nascente (ABRUZZI, LACADIE e ROSBASH, 2004; JURICA e MOORE, 2003). Após o *splicing* e poliadenilação do mRNA, UAP56/Sub2 é substituída por Aly/Yra1, que recruta o dímero TAP-p15 / Mex67-Mtr2 ao transcrito. Este complexo ribonucleoprotéico (mRNP) atravessa o poro nuclear através de interação com proteínas adaptadoras e nucleoporinas e libera o complexo mRNP para a tradução no citoplasma (KATAHIRA *et al.*, 1999; SANTOS-ROSA *et al.*, 1998; SEGREF *et al.*, 1997). A via de exportação de mRNAs de fungos e humanos pode ser visualizada na figura 8.

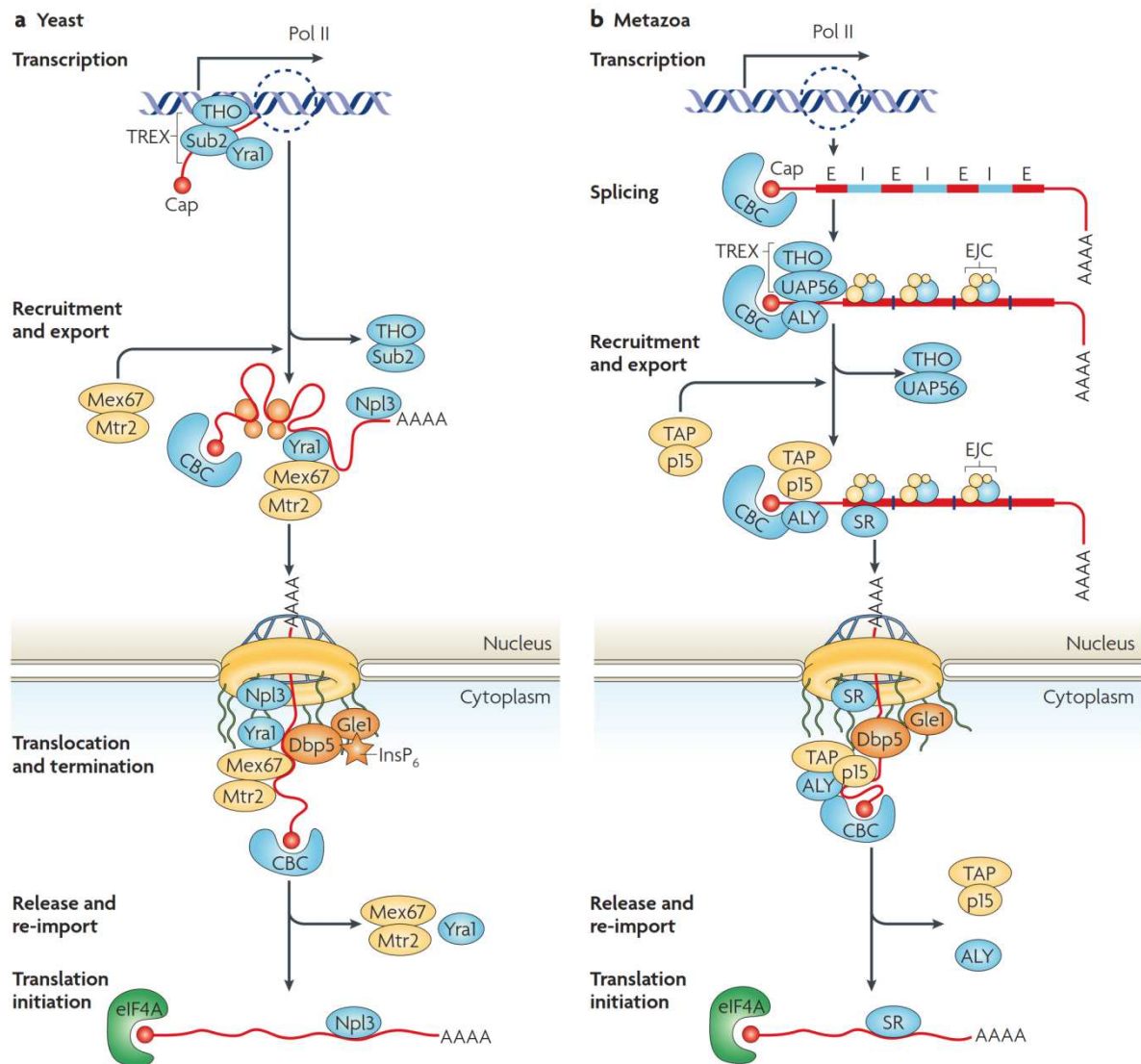


Figura 8. Transcrição ou *splicing* acoplado com a exportação de mRNAs. Em fungos (*yeast*), o transcrito nascente, gerado por RNA Polimerase II (Pol II), interage com o complexo THO/TREX (subcomplexo THO e proteínas Sub2 e Yra1) de modo co-transcricional. Após a associação com complexo de ligação ao *cap* (CBC) e proteínas ligadoras de RNAs (em laranja), o dímero receptor de exportação Mex67-Mtr2 é recrutado ao complexo mRNP através da proteína adaptadora Yra1. No complexo de poro nuclear (NPC), o mRNA interage com a RNA-helicase Dbp5, ativador Gle1 e molécula de sinalização *Insp*₆ (*inositol hexakisphosphate*). Após a dissociação do mRNA de Mex67-Mtr2 e outros fatores de exportação, este mRNA ainda pode conter proteínas que migram entre núcleo e citoplasma, como Npl3, que influenciam na tradução. O mRNA é recrutado para a maquinaria de iniciação da tradução via *cap*, que é reconhecido pelo fator de iniciação eIF4A e fatores adicionais. Os fatores de transporte são reimportados para o núcleo, o que em alguns casos requer fosforilações. B. Em metazoários (Metazoa), muitos modelos de exportação de mRNAs vem sendo descritos. Em humanos, a exportação de mRNAs é dependente de *cap*, recrutamento de TREX e *splicing* de mRNAs. Os eventos após o processamento correto de mRNA, que incluem o receptor de exportação TAP-p15, são similares em metazoários e fungos. Os receptores de exportação estão indicados em amarelo e as proteínas adaptadoras, em azul. Pra simplificar, os componentes de EJC (*exon-junction complex*) não estão sendo mostrados. E, éxon; I, intron; TREX, *transcription-export complex*. FONTE: KÖHLER e HURT, 2007.

Há também a participação de outro complexo em mamíferos e fungos que interage sinergisticamente com TREX1 denominado de TREX-2 ou THSC, formado

em fungos pelas proteínas THP1-SAC3-SUS1-CDC31 (PCID2- GANP- ENY2-centrinas, em mamíferos) que podem interagir também com o receptor de exportação Mex67-Mtr2 (TAP-p15) (JANI *et al.*, 2012; KOHLER e HURT, 2007). Este complexo está intimamente relacionado com a transcrição devido à associação com outro conjunto de proteínas denominado SAGA, que tem como função a promoção de acetilação de histonas e interação com cromatina ativa (FISCHER *et al.*, 2002; GONZALEZ-AGUILERA *et al.*, 2008; TIMMERS e TORA, 2005). A proteína Sus1 é um componente central para a interação com proteínas do poro nuclear e há evidências de que em conjunto com a proteína Sac3 e Thp1 há o recrutamento da maquinaria de transcrição de mRNAs para próximo do NPC. Este processo de interação reforça a hipótese de *gene gating*, que é a aproximação de genes ativos para a periferia do núcleo facilitando o direcionamento do transcrito para o citoplasma (BLOBEL, 1985; FISCHER *et al.*, 2002; IGLESIAS e STUTZ, 2008; RODRIGUEZ-NAVARRO *et al.*, 2004). Em contraste com fungos, a maior parte dos genes ativos nas células de mamíferos se encontram em “fábricas de transcrição” no nucleoplasma (DARZACQ, SINGER e SHAV-TAL, 2005; FRASER e BICKMORE, 2007; SHAV-TAL *et al.*, 2004), onde também são encontradas as proteínas de TREX-2, que ajudam na transferências de mRNAs para o poro nuclear, com o auxílio de GANP (*germinal-centre associated nuclear protein*) que parece atuar principalmente como chaperona, auxiliando os mRNPs maduros a atravessar o NPC (JANI *et al.*, 2012; WICKRAMASINGHE *et al.*, 2010).

Por fim, alguns mRNAs de células humanas, que codificam para proteínas mitocondriais, que serão secretadas ou irão se ligar a membranas, possuem elementos que ativam uma via alternativa de exportação de mRNAs denominada ALREX, independente de *splicing* ou *cap 5'* funcional (PALAZZO *et al.*, 2007). Estes mRNPs também estão presentes em *speckles* onde se associam com proteínas do complexo TREX1, UAP56 ou URH49, bem como com o dímero TAP-p15 (AKEF *et al.*, 2013). Um esquema geral e atual sobre as vias de exportação de mRNPs em células de mamíferos pode ser observada na figura 9.

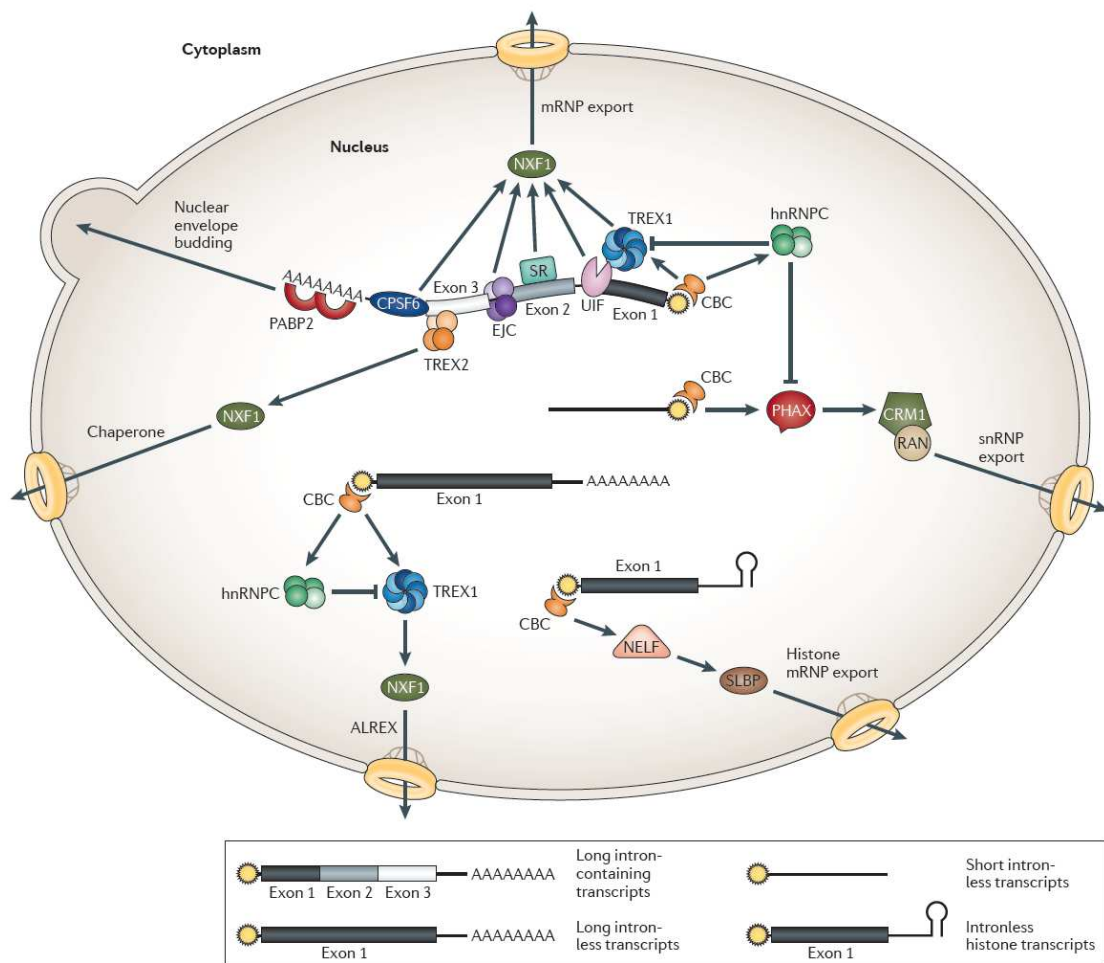


Figura 9. Exportação de mRNPs. Exemplos de diferentes mRNPs contendo quatro tipos diferentes de mRNAs transcritos por RNA polimerase II e as diferentes vias para exportação nuclear. A maioria dos mRNPs completamente processados são geralmente exportados pelo fator exportação RNA nuclear 1 (conhecida como TAP ou NXF1). Através da interação com diferentes proteínas adaptadoras mRNA – Complexo de transcrição e exportação 1 (TREX1), fator de interação a UAP56 (UIF), fator 6 de clivagem e poliadenilação (CPSF6) ou proteínas SR – TAP/NXF1 é recrutada para o transcrito longo que contém introns em uma maneira dependente de *splicing*. Quando os mRNPs são grandes demais para passar através dos poros nucleares, eles podem sair através de protrusões do envelope nuclear (*budding*). mRNAs de histonas que não possuem introns são exportados pela proteína ligadora de *stem-loops* (SLBP) que é recrutada através de interação com o complexo de ligação ao *cap* 5' (CBC) e fator de alongamento negativo (NELF). Chaperonas do complexo de transcrição e exportação 2 (TREX2) auxiliam desde o processamento até o transporte para o NPC. As quatro subunidades de TREX2 (GANP - *germinal centre-associated nuclear protein*, PCID2 - *PCI domain-containing protein 2*, DSS1 e ENY2 – participam na exportação de mRNA e provavelmente formam o complexo, embora apenas GANP1 seja a única que interaja diretamente com TAP/NXF1. O complexo TREX1 consiste no complexo THO (THOC1–7), bem como proteínas UAP56, TEX1 e CIP29. A exportação de transcritos longos sem introns ocorre através da via alternativa ALREX, envolvendo o recrutamento de Aly, THOC5 e UAP56 na região 5' do transcrito na região do CBC, independente de *splicing*, recrutando TAP/NXF1 aos mRNAs. PHAX (*phosphorylated adaptor for RNA export*), hnRNPC (*heterogeneous nuclear RNP C*) e Aly competem entre si para a interação com CBC. Transcritos pequenos e sem introns os quais hnRNPC não conseguem estabilizar (snRNAs), são ligados a PHAX e subsequentemente exportados via CRM1 (*chromosomal region maintenance protein 1*) – pela via dependente de RanGTP. EJC, *exon junction complex*; PABP2, *poly(A) binding protein 2* (homóloga nuclear de PABP). FONTE: MÜLLER-MCNICOLL e NEUGEBAUER, 2013.

1.3.1 Monitoramento de transporte de mRNAs através do sistema MS2

A proteína de capsídeo do bacteriófago do grupo I, fago 17 (MS2), reconhece especificamente uma estrutura secundária em forma de grampo (*hairpin / stem-loops*) do RNA genômico simples-fita. Esta estrutura é formada por 19 nucleotídeos, contendo a sequência Shine-Dalgarno e códon de iniciação do gene replicase. Esta proteína MS2 tem por função a repressão traducional de replicase após a interação com estas estruturas *stem-loops*, estabilizando-as e impedindo o acesso de ribossomos à sequência Shine-Dalgarno para o início da tradução (FOUTS, TRUE e CELANDER, 1997; STOCKLEY *et al.*, 1995). Esse mecanismo MS2-*stem-loops* permitiu o desenvolvimento de uma ferramenta para visualização de mRNAs *in vivo* nas células e vem sendo utilizada largamente em estudos de localização em eucariotos em diferentes contextos biológicos (BERTRAND *et al.*, 1998; FUSCO *et al.*, 2003; HOCINE *et al.*, 2013; LIONNET *et al.*, 2011; RODRIGUEZ *et al.*, 2007). O sistema se baseia na expressão simultânea da proteína MS2 fusionada a GFP (*green fluorescent protein*), bem como da transcrição do mRNA alvo contendo várias repetições de sequências *stem-loops* na região 3' não codificante (BERTRAND *et al.*, 1998). Foi demonstrado que 24 repetições é o suficiente para a detecção indireta das moléculas de mRNA por MS2-GFP, sendo facilmente visualizado por microscopia de fluorescência (FUSCO *et al.*, 2003). Esse sistema está esquematizado na figura 10.

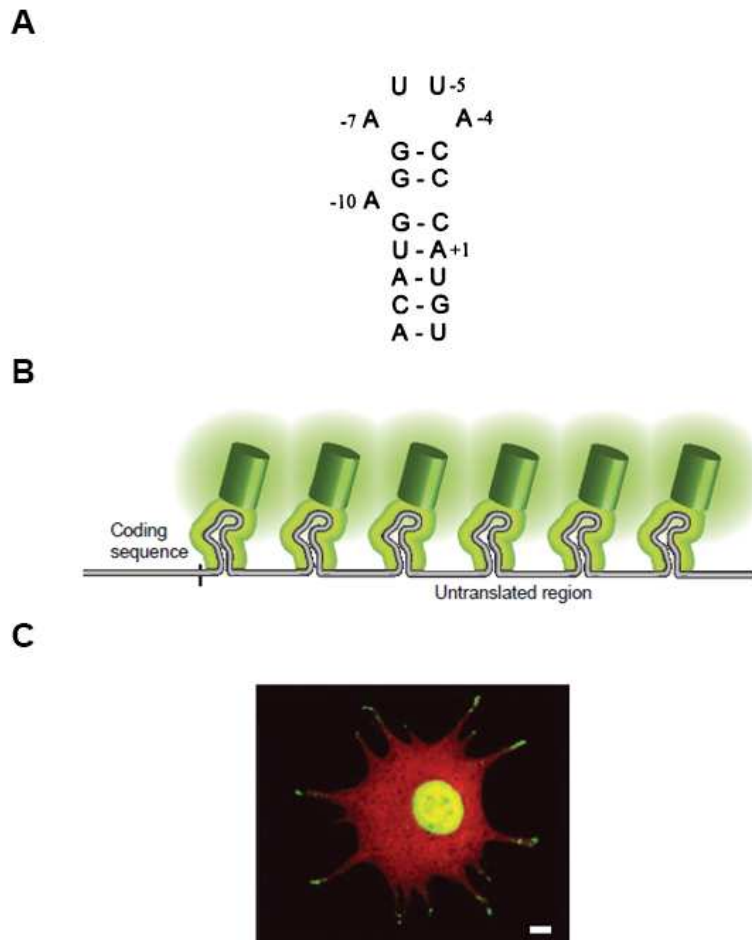


Figura 10. Sistema MS2 para detecção de mRNAs. A. Sequência do operador traducional MS2 formando a estrutura secundária do RNA: *stem-loop*. Os números são relativos ao códon inicial de replicase, sendo o A de AUG o número +1. Para a interação com a proteína MS2, além do reconhecimento da estrutura secundária, há a exigência de algumas purinas de localização conservada: posições 4, 7 e 10. Os traços indicam pareamento de base de Watson-Crick (STOCKLEY *et al.*, 1995). B. Representação esquemática do sistema MS2 onde a proteína MS2 reconhece as estruturas *stem-loops*. Para a utilização desta ferramenta, as repetições da sequência *stem-loop* são introduzidas na região não codificadora 3'UTR (UTR - *untranslated region*) do mRNA alvo. Na célula, após a transcrição do mRNA alvo contendo *stem-loops*, há o reconhecimento destas estruturas pela proteína MS2, fusionada a GFP (MS2-GFP), também presente na célula, sendo possível observar a localização indireta do mRNA alvo por microscopia de fluorescência no canal verde. C. Imagem de fluorescência de fibroblasto expressando a proteína MS2-GFP a qual reconhece o mRNA *RAB13*, que contém 24 cópias de sequências *stem-loops* na porção 3'UTR, nas protrusões da célula. Neste caso foi utilizada uma proteína de distribuição citossólica como controle, em vermelho (mRFP). Barra: 10 μ m (TYAGI, 2009).

1.3.2 Exportação de mRNAs em parasitas: abordagem do problema

A interação entre transcrição, processamento e transporte de RNAs reflete no sinergismo entre os eventos nucleares e citoplasmáticos no que diz respeito ao controle de qualidade do processamento do mRNA, assim como, a modulação da expressão gênica propriamente dita. A diferença entre os componentes envolvidos

nesses processos possivelmente reflete características específicas no metabolismo de RNA de cada organismo. Por exemplo, em leveduras, o processamento de mRNA por *cis-splicing* ocorre apenas para a minoria dos genes (WARNER, 1999), enquanto em metazoários o *cis-splicing* passa a ser regra. No grupo de organismos que divergiram cedo durante a evolução, as peculiaridades são evidentes, como é o caso dos tripanossomatídeos, onde um RNA policistrônico é processado por *trans-splicing*, com algumas diferenças quando comparado ao processamento por *cis-splicing*. Sendo assim, a caracterização de eventos pós-transcricionais como no caso da exportação de mRNAs nestes organismos, é sem dúvida relevante para enriquecer o que se sabe sobre a biologia molecular principalmente de parasitas de importância médica.

Dentre os eventos essencialmente pós-transcricionais para o controle de expressão gênica em eucariotos, a maquinaria da via de exportação nucleocitoplasmática de RNAs ainda é pouco entendida em parasitas.

Em *T. brucei* foram identificadas algumas nucleoporinas, como por exemplo, Nup1 que está envolvida com a organização de heterocromatina perinuclear, TbGle2 e TbSec13, proteínas possivelmente relacionadas ao transporte de macromoléculas através do NPC mas ainda não caracterizadas funcionalmente (DEGRASSE *et al.*, 2009; ROUT e FIELD, 2001).

Em *T. cruzi*, foi caracterizada a proteína ortóloga de Nup1 (TcNup1), uma nucleoporina que ancora cromossomos no envelope nuclear (PICCHI *et al.*, 2011), bem como uma proteína denominada TcUBP1 que possui o motivo consenso LRR e parece estar indiretamente relacionada com exportação de mRNAs (CASSOLA e FRASCH, 2009; D'ORSO e FRASCH, 2001). Neste mesmo trabalho, foi observado que o tratamento com leptomicina B (LMB), uma substância que inibe a função de CRM1 (KUDO *et al.*, 1999), resulta em acúmulo no núcleo de mRNAs, e proteínas TcPABP1 e TcUBP2/1 e com isso os autores propuseram que CRM1 estaria envolvida na exportação de mRNAs (CUEVAS, FRASCH e D'ORSO, 2005).

Os resultados de um trabalho prévio do nosso grupo, utilizando ferramentas de bioinformática para comparações de sequências de proteínas envolvidas nas diferentes vias de exportação de RNAs (ver Capítulo 1, (SERPELONI *et al.*, 2011)), mostraram que os componentes básicos da via dependente de RanGTP estão

conservados ao longo da filogenia de eucariotos, enquanto proteínas envolvidas na via independente de RanGTP são menos conservadas, principalmente em eucariotos que divergiram na base da árvore filogenética. CRM1, por exemplo, é uma exportina altamente conservada e, levando em consideração a evidência em *T. cruzi* com LMB, foi proposto um modelo evolutivo das vias de exportação de RNAs, onde o eucarioto ancestral poderia ter uma exportina ancestral, muito parecida com CRM1 e, um pouco antes da origem de LECA, foram surgindo diferentes exportinas, inclusive CRM1, cuja exportação ainda seriam dependentes da via de RanGTP. Com o surgimento de inovações na regulação da expressão gênica houve a necessidade de uma exportação específica para mRNAs, dependente de Mex67-Mtr2/TAP-p15. No entanto, a presença de proteínas menos conservadas não indica ausência de ortólogos em parasitas e pode refletir a presença de proteínas mais específicas e não tão similares, podendo apresentar alguns domínios diferentes, mas cuja função pode ser conservada. Esse foi o caso da proteína Mex67/TAP, o receptor de exportação como citado anteriormente, que pelas nossas análises, possui baixo grau de conservação ao longo da filogenia de eucariotos. Um trabalho em 2010 (KRAMER, KIMBLIN e CARRINGTON, 2010) observou que a proteína ortóloga de *T. brucei* possui um motivo *zinc-finger*, ausente nas proteínas ortólogas de fungos e mamíferos, mas conservado entre tripanossomatídeos e recentemente esse motivo foi identificado como relacionado diretamente com a exportação de mRNAs neste parasita (DOSTALOVA *et al.*, 2013).

Embora existam evidências sobre participação de CRM1 e Mex67 na via de exportação de mRNAs em tripanossomatídeos, ainda não foi feita uma caracterização mais detalhada da maquinaria para verificar se há redundância das vias ou se existem diferentes vias para exportação de mRNA nestes organismos.

Ainda dentre as proteínas identificadas no nosso trabalho, foi identificada uma proteína altamente conservada ao longo da filogenia de eucariotos que não depende da via de RanGTP, inclusive em eucariotos que divergiram cedo durante a evolução, como tripanossomatídeos e parasitas do filo Apicomplexa: UAP56/Sub2. Esta proteína é uma DEAD-box helicase relacionada com o complexo TREX1 na exportação de mRNAs. Em outro trabalho do nosso grupo (ver Capítulo 2, (SERPELONI *et al.*, 2011)), identificamos as proteínas ortólogas TcSub2 e TbSub2,

de *T. cruzi* e *T. brucei*, respectivamente as quais nomeamos genericamente de TrypSub2 por serem altamente similares. Foi observado que se trata de uma proteína nuclear, essencial e componente da via de transcrição/exportação de mRNAs em tripanossomatídeos. Então demos continuidade aos estudos funcionais de TcSub2 em *T. cruzi* focando na identificação de proteínas associadas a ela para melhor compreensão do contexto em que está inserida no metabolismo de mRNA. Como uma análise complementar, buscamos por interações diretas com TcSub2 por análises de duplo-híbrido, bem como de algumas proteínas identificadas em nosso trabalho anterior (SERPELONI *et al.*, 2011) utilizando-as como alvos. Estas proteínas, embora menos conservadas, podem indicar futuros possíveis candidatos para análises da via de exportação de mRNAs neste tripanossomatídeo.

Em *T. gondii* existe um estudo de predição de proteínas ortólogas, principalmente para as proteínas da via RanGTP mostrando que, assim como nossos dados, estas proteínas estão bem conservadas neste parasita (FRANKEL e KNOLL, 2009). Existem alguns estudos funcionais sobre transporte nuclear que mostram a interação entre TgGCN5 (histona acetiltransferase) com TgIMP α (importina- α) (BHATTI e SULLIVAN, 2005), bem como entre TgRCC1 e TgRan, sendo esta última de distribuição por toda a célula, que difere da localização nuclear encontrada em outros organismos (FRANKEL e KNOLL, 2008). Em *Plasmodium* também foram preditas algumas proteínas ortólogas (TUTEJA e MEHTA, 2010), e o mesmo grupo realizou análises bioquímicas com as proteínas ortólogas de DDX19 e UAP56 (Dbp5/Sub2 de fungos) em *Plasmodium falciparum*. A proteína ortóloga de DDX19/Dbp5 neste parasita, PfD66, é uma DNA/RNA helicase e ATPase dependente de ácido nucléico e a proteína ortóloga de UAP56/Sub2, PfU52, é uma ATPase dependente de ácido nucléico, envolvida em *splicing in vitro* (SHANKAR, PRADHAN e TUTEJA, 2008). Apesar destes estudos em transporte nuclear em Apicomplexas, ainda pouco se sabe a respeito das vias de exportação de RNAs nestes organismos. Por Sub2 se tratar de uma proteína altamente conservada em eucariotos, iniciamos os estudos funcionais da via de exportação de mRNAs em *T. gondii* com a proteína ortóloga, nomeada TgUAP56. Foram utilizadas as ferramentas de genética reversa já bem estabelecidas neste parasita como o sistema DD (HERM-GOTZ *et al.*, 2007) e o nocaute condicionado (ANDENMATTEN *et al.*, 2013)

do gene *uap56* que abriam perspectivas de estudos do envolvimento de TgUAP56 com metabolismo de mRNAs, bem como proteínas envolvidas na via de exportação em *T. gondii*. Com isso pretendemos futuramente compreender o quão conservada se encontra a via de exportação de mRNAs, bem como avaliar a participação de proteínas essenciais e possivelmente específicas de parasitas nesse processo.

Pensando em uma aplicação em larga escala para análise fenotípica, foi estabelecido o sistema MS2 para monitoramento de distribuição celular de mRNAs em *T. gondii*. Neste caso foi desenvolvida uma ferramenta de detecção de exportação de mRNA através do rastreamento de mRNA de Dip13 contendo *stem-loops* que são reconhecidos pela proteína MS2. O gene *dip13* (TGME49_295450) codifica para uma proteína citoplasmática que não é essencial ao parasita e a função ainda não foi estudada (dados não publicados). Esta proteína foi fusionada ao NLS de *T. gondii* (*nuclear localization signal* - sinal de localização nuclear) (BHATTI e SULLIVAN, 2005), YFP (proteína amarela fluorescente) e ao domínio de desestabilização ddfKBP12. O sistema DD, como já citado anteriormente, permite a expressão de modo induzido quando da presença de Shld-1, havendo estabilização da proteína fusionada. Logo, esperamos que em uma situação normal a proteína NLS-MS2-YFP-DD esteja ligada ao mRNA de Dip13 no citoplasma e, após interromper a via de exportação de mRNA, o mRNA de Dip13 não saia do núcleo, havendo conseqüente acúmulo nuclear da proteína modificando a relação entre núcleo e citoplasma, facilmente detectável por microscopia de fluorescência. Logo, o nocaute condicionado do gene *uap56* de *T. gondii*, resultando em diminuição de TgUAP56, poderia ser usado como prova de conceito deste sistema. Além disso, acreditamos que este sistema poderá ser utilizado em estudos de larga-escala para varredura de drogas que inibam a exportação de mRNAs, visto que se trata de uma via essencial ao parasita.

2 Objetivos

2.1 Objetivo geral

Identificar proteínas de parasitas envolvidas na exportação de mRNAs do núcleo para o citoplasma utilizando ferramentas de bioinformática, genética reversa, biologia molecular e celular.

2.1.1 Objetivos específicos

Objetivo 1 - Identificar as proteínas envolvidas na exportação de RNAs em eucariotos por análises *in silico*, incluindo espécies de parasitas.

Meta 1.1. Listar candidatos para as análises de genômica comparativa através de buscas na literatura de proteínas envolvidas nas vias de exportação do núcleo para o citoplasma dos diferentes tipos de RNAs e com evidências funcionais em *S. cerevisiae* e *H. sapiens* (organismos-modelo).

Meta 1.2. Executar análises comparativas das proteínas encontradas através de alinhamentos de estruturas primárias entre espécies de diferentes grupos de eucariotos visando cobrir a árvore filogenética dos eucariotos.

Meta 1.3. Identificar proteínas mais conservadas ao longo da filogenia de eucariotos e seguir com as análises funcionais em parasitas.

Objetivo 2 - Comprovar a homologia da proteína mais conservada, neste caso a candidata à proteína ortóloga de Sub2, na exportação de mRNAs em tripanossomatídeos por abordagens de genética reversa.

Meta 2.1. Avaliar o fenótipo causado por nocaute gênico ou silenciamento por interferência de RNA (RNAi) de proteínas ortólogas de Sub2 de *T. cruzi* e *T. brucei*, respectivamente.

Objetivo 3 - Verificar a função da proteína ortóloga de Sub2 na exportação de mRNAs por abordagens de genética reversa em espécie de parasita pertencente a um grupo diferente dos tripanossomatídeos.

Meta 3.1. Avaliar o fenótipo causado por nocaute gênico ou superexpressão de proteína ortóloga de Sub2 de *T. gondii*, espécie do grupo Apicomplexa.

Objetivo 4 - Ampliar a identificação de componentes da via de exportação de mRNAs de *T. cruzi*.

Meta 4.1. Identificar proteínas candidatas através de análises de interação por ensaio de duplo-híbrido.

Meta 4.2 Identificar complexos de proteínas por ensaios de imunoprecipitação seguida de espectrometria de massas usando como isca a proteína de *T. cruzi* ortóloga de Sub2 (TcSub2).

3 Capítulo 1.

Comparative genomics of proteins involved in RNA nucleocytoplasmic export

Autores: Mariana Serpeloni, Newton M Vidal, Samuel Goldenberg, Andréa R Ávila, Federico G Hoffmann

Publicado no periódico BMC Evolutionary Biology. 11:7 LMS2.
doi:10.1186/1471-2148-11-7

Muitos mecanismos de exportação de mRNAs têm sido estudados em organismos-modelo como leveduras e humanos e pouco se sabe a respeito deste tipo de transporte em protozoários de importância médica, os quais divergiram na base da árvore filogenética de eucariotos. Por esse motivo, foi realizado um trabalho pelo nosso grupo sobre análise de genômica comparativa para verificar a conservação das proteínas envolvidas com o transporte de RNAs ao longo da evolução (artigo em anexo neste capítulo (SERPELONI *et al.*, 2011)), incluindo as linhagens eucarióticas Chromalveolata e Excavata, nas quais estão inseridas várias espécies de parasitas de cada grupo, como por exemplo *T. gondii* e *T. cruzi*, respectivamente.

Foi feita uma busca inicial na literatura nestes organismos-modelo para a identificação de proteínas envolvidas em transporte de diferentes tipos de RNAs, entre eles mRNAs. Foram identificadas 64 proteínas de *S. cerevisiae* funcionalmente caracterizadas e, a partir destas, foram verificadas quais apresentam caracterização funcional em humanos, bem como indicadas como proteínas ortólogas pelo banco de dados do *Ensembl*. As sequências das proteínas de *S. cerevisiae* e *H. sapiens* identificadas foram utilizadas como sementes de busca (*query*) em bancos de dados de diferentes eucariotos para análises de conservação de estrutura primária através de alinhamento local por algoritmos de BLASTP, PSI-BLAST e TBLASTN (ALTSCHUL *et al.*, 1997). As sequências encontradas através das *queries* são denominadas *subjects* e as regiões de maior pontuação para o alinhamento local foram denominadas HSP (*High-scoring Segment Pair*), que é o comprimento da região alinhada entre *query* e *subject*. Os bancos de proteínas utilizados para os alinhamentos foram de representantes dos 5 supergrupos, cobrindo a árvore filogenética dos eucariotos: Opisthokonta, Amoebozoa, Plantae, Chromalveolata e

Excavata. Foram analisadas 65 espécies pertencentes a estes supergrupos e foi considerado como resultado final o melhor *hit* encontrado em cada espécie. Através dessa busca foi possível inferir a conservação de cada proteína ao longo da evolução, em busca de prováveis ortólogos em diferentes organismos.

Os critérios utilizados para determinação da conservação foram exigentes quanto à similaridade e cobertura de alinhamento entre as proteínas em virtude da distância filogenética entre as sequências incluídas na análise. Esta exigência diminuiu o repertório de identificação mas garantiu a confiabilidade, principalmente pela questão de envolver organismos que divergiram há mais de 1 bilhão de anos (ROGER e HUG, 2006). A classificação da conservação de estrutura primária das proteínas ao longo da filogenia de eucariotos foi realizada por números de 1 a 5 em escala decrescente de conservação, sendo os números 1 e 2 considerados os resultados de proteínas altamente conservadas.

Pelas análises de conservação de proteínas foi possível verificar que algumas poucas são altamente conservadas em todas as linhagens de eucariotos, inclusive em protozoários. Podemos então inferir que LECA (*Last Eukaryotic Ancestor*) possuía os componentes básicos de exportação de RNAs, principalmente das vias dependentes de RanGTP. A hipótese atual é que a complexidade aumentou de acordo com a evolução eucariótica, com o surgimento de inovações de proteínas com funções menos generalizadas durante a evolução, principalmente em relação à via de exportação de mRNAs, que é independente do gradiente de RanGTP e possui componentes específicos em leveduras e metazoários. É importante ressaltar que isso não implica em ausência de uma via para exportação para mRNAs em protozoários, seja através de atuação de exportinas relacionadas à via de RanGTP (CUEVAS, FRASCH e D'ORSO, 2005) ou até mesmo proteínas mais específicas, não necessariamente muito similares às existentes em organismos-modelo (DOSTALOVA *et al.*, 2013; SCHWEDE *et al.*, 2009).

Os resultados mostraram que no grupo de proteínas altamente conservadas, classificadas como 1 ou 2 ao longo da filogenia, estão incluídas principalmente proteínas pertencentes às vias dependentes de RanGTP, como Ran e exportina CRM1, relacionadas com o transporte nucleocitoplasmático de proteínas (HUTTEN e KEHLENBACH, 2007; JEKELY, 2003; YONEDA *et al.*, 1999), bem como com as vias de exportação de rRNAs, tRNAs e snRNAs (CULLEN, 2003b; HUTTEN e

KEHLENBACH, 2007). Outra proteína altamente conservada é Nmd3, parceira de CRM1 na exportação de rRNAs em leveduras e já caracterizada em *T. brucei* (HO, KALLSTROM e JOHNSON, 2000; PROHASKA e WILLIAMS, 2009). Quanto às proteínas específicas da via de exportação de mRNAs, ou seja, do grupo independente de RanGTP, foi observado que a maioria das proteínas não é altamente conservada ao longo da filogenia de eucariotos, principalmente em protozoários. Uma exceção foi a proteína UAP56/Sub2 de humanos e fungos, respectivamente, uma proteína diretamente relacionada com a exportação e processamento de mRNAs em organismos-modelo (KAMMEL *et al.*, 2013; REED e CHENG, 2005; WAHL, WILL e LUHRMANN, 2009) que se destacou pela alta conservação inclusive em diferentes protozoários.

Em virtude do foco inicial do nosso trabalho ser a análise de componentes da via de exportação de mRNAs em tripanossomatídeos, mais especificamente *T. cruzi* e *T. brucei*, optamos por realizar as análises funcionais da proteína ortóloga de UAP56/Sub2 nestes parasitas e os resultados obtidos serão discutidos no próximo capítulo. Algumas possíveis proteínas ortólogas de *T. cruzi* relacionadas com a exportação de mRNAs foram analisadas também neste trabalho e serão discutidas no Capítulo 4.

RESEARCH ARTICLE

Open Access

Comparative genomics of proteins involved in RNA nucleocytoplasmic export

Mariana Serpeloni^{1,2†}, Newton M Vidal^{1,2†}, Samuel Goldenberg^{1,2}, Andréa R Ávila^{1,2*}, Federico G Hoffmann^{2,3,4*}

Abstract

Background: The establishment of the nuclear membrane resulted in the physical separation of transcription and translation, and presented early eukaryotes with a formidable challenge: how to shuttle RNA from the nucleus to the locus of protein synthesis. In prokaryotes, mRNA is translated as it is being synthesized, whereas in eukaryotes mRNA is synthesized and processed in the nucleus, and it is then exported to the cytoplasm. In metazoa and fungi, the different RNA species are exported from the nucleus by specialized pathways. For example, tRNA is exported by exportin-t in a RanGTP-dependent fashion. By contrast, mRNAs are associated to ribonucleoproteins (RNPs) and exported by an essential shuttling complex (TAP-p15 in human, Mex67-mtr2 in yeast) that transports them through the nuclear pore. The different RNA export pathways appear to be well conserved among members of Opisthokonta, the eukaryotic supergroup that includes Fungi and Metazoa. However, it is not known whether RNA export in the other eukaryotic supergroups follows the same export routes as in opisthokonts.

Methods: Our objective was to reconstruct the evolutionary history of the different RNA export pathways across eukaryotes. To do so, we screened an array of eukaryotic genomes for the presence of homologs of the proteins involved in RNA export in Metazoa and Fungi, using human and yeast proteins as queries.

Results: Our genomic comparisons indicate that the basic components of the RanGTP-dependent RNA pathways are conserved across eukaryotes, and thus we infer that these are traceable to the last eukaryotic common ancestor (LECA). On the other hand, several of the proteins involved in RanGTP-independent mRNA export pathways are less conserved, which would suggest that they represent innovations that appeared later in the evolution of eukaryotes.

Conclusions: Our analyses suggest that the LECA possessed the basic components of the different RNA export mechanisms found today in opisthokonts, and that these mechanisms became more specialized throughout eukaryotic evolution.

Background

Protein synthesis in all living cells involves the transcription of DNA into messenger RNA (mRNA) and its subsequent translation into polypeptides. In prokaryotes, transcription and translation are physically and temporally linked, and each mRNA molecule is translated by the ribosomes as it is transcribed. By contrast, in eukaryotes transcription and mRNA processing are physically and temporally separated from translation by the nuclear

membrane. This separation is hypothesized to have been a major factor in the emergence of the nuclear membrane [1]. As a result of the establishment of the nuclear membrane, the different RNA species involved in protein synthesis such as mRNAs, ribosomal RNAs (rRNAs), and transfer RNAs (tRNAs), need to be shuttled from the nucleus to the cytoplasm. The general model of RNA export involves exportins as transport receptors that carry RNA through the nuclear pore complex (NPC) in a RanGTP-dependent manner [2]. In Metazoa and Fungi, the nuclear export of most RNA species, such as microRNAs (miRNAs), rRNAs, small nuclear RNAs (snRNAs), and tRNAs, has been shown to follow the RanGTP-exportin model of transport, and specific exportins are involved with the different export pathways. However,

* Correspondence: aravila@fiocruz.br; federico.g.hoffmann@gmail.com

† Contributed equally

¹Departamento de Biologia Celular e Molecular, Universidade Federal do Paraná - UFPR, Curitiba, Brazil

²Instituto Carlos Chagas - ICC, Curitiba, Brazil

Full list of author information is available at the end of the article

nuclear export of most mRNAs does not follow the RanGTP-exportin pathway [3-5]. The mRNA export machinery is highly integrated with mRNA processing, and it includes a different set of nuclear transport adaptors plus other mRNA binding proteins, RNA helicases, and NPC associated proteins [4-6].

It is not known whether the proteins involved in the different RNA export pathways are conserved among the different eukaryotic supergroups, and which components of the nucleocytoplasmic export of RNA are traceable to the last eukaryotic common ancestor (LECA). The goal of this study was to explore the evolutionary history of the different RNA export pathways across a diverse array of eukaryotes, with special emphasis on members of the supergroups Excavata and Chromalveolata, in order to identify lineage-specific innovations and make inferences regarding RNA export evolution. To do so, we screened the genome of 65 species of eukaryotes to explore the presence of homologs of proteins involved in RNA export in Metazoa and Fungi. Results from our bioinformatic comparisons suggest that the basic components of the RanGTP-dependent RNA pathways are conserved across eukaryotes, whereas proteins involved in RanGTP-independent mRNA export are less conserved.

Methods

Data Sources

We identified sixty four proteins from baker's yeast (*S. cerevisiae*) experimentally shown to be involved in RNA export and their putative human orthologs from Ensembl (release 54), plus PHAX, a human protein involved in RNA export that apparently lacks a yeast ortholog. For simplicity, proteins have been named according to their human ortholog when present, the full list of proteins used is provided in Table S1 (Additional File 1). Using all the proteins previously identified to seed bioinformatic searches, we interrogated the genomes of 65 eukaryotic species to identify their putative homologs. Our initial set of species included representatives of five of the different eukaryotic supergroups (Opisthokonta, Amoebozoa, Chromalveolata, Plantae, and Excavata). The genomes of most species were downloaded from the NCBI Reference Sequence (RefSeq) collection except where noted (see Table 1). The initial set included 2 or more representatives of the genera *Caenorhabditis*, *Entamoeba*, *Cryptosporidium*, *Leishmania*, *Plasmodium*, and *Trypanosoma* which provided taxonomic redundancy and served as consistency controls for our protocols.

Bioinformatic searches

Bioinformatic searches were locally performed using the BLASTP, PSI-BLAST and TBLASTN algorithms [7],

part of the NCBI C++ toolkit [8]. In the case of BLASTP, we compared results obtained under three different matrices (BLOSUM62, BLOSUM45 and PAM250) and in the case of PSI-BLAST we compared the results obtained for 2, 3 and 5 iterations. There were no major differences between the different search strategies implemented. Thus, we report results for BLASTP with the BLOSUM62 matrix in the manuscript. All results, including Reciprocal Best Hit (RBH), also evaluated using BLASTP with the BLOSUM62 matrix, are presented on Tables S2 and S3 (Additional File 1). Results were then ranked into 5 categories (1-5; with 1 being the most conserved) as described below. Hits with E-values better than 10^{-5} were classified according to the following criteria: category 1 [strong similarity], a candidate homolog sequence (CHS) of similar length to the query protein sequence (QS), showing a similarity of 60% or greater, and having a match that covers at least 80% of the QS; category 2 [similar], a CHS of similar length to the QS, similarity higher than 50%, and covering at least 60% of QS; category 3 [weak similarity], a CHS of similar length to QS, with similarity higher than 40%, and covering at least 45% of QS; category 4 [partial similarity], a CHS having 30% similarity or greater, and covering at least 30% of QS; and category 5 [very weak similarity], a CHS with E-value better than 10^{-5} that did not match any of the criteria above; candidates with E-value worse than 10^{-5} were not classified (NC). In cases where human and yeast orthologs were present, results from the two queries were assessed together to ensure consistency. Proteins in categories 1 and 2 were considered as strong homology candidates to the query sequences, proteins in category 3 were considered as potential homologs of the query sequences, proteins in categories 4 and 5 were considered as showing some local homologies due to sharing conserved domains. Exemplary alignments for each category are provided in Additional File 2. TBLASTN only improved results by uncovering category 5 CHS in cases where no CHS had been identified, and the results are presented in Table S3 (Additional File 1) but are not discussed further. It is important to note that a negative result does not prove the absence of a particular protein homolog, it is just an indication that our protocols, based on sequence similarity, were not able to identify a homolog candidate.

Reconstruction of RNA export pathways in LECA

Our reconstruction of the different RNA export pathways in the LECA was based on a tree adapted from four recent reports that group the six eukaryotic supergroups into three separate megagroups [9-12]. The first megagroup includes Opisthokonta and Amoebozoa, the second one includes Chromalveolata, Plantae, and Rhizaria, and the third one corresponds to Excavata (Figure 1).

Table 1 List of species included in this study

Species	Code	Lineage	Supergroup	Data Source
<i>Homo sapiens</i>	Hsa	Metazoa	Opisthokonta	NCBI RefSeq
<i>Mus musculus</i> ^a	Mmu	Metazoa	Opisthokonta	NCBI RefSeq
<i>Rattus norvegicus</i> ^a	Rno	Metazoa	Opisthokonta	NCBI RefSeq
<i>Gallus gallus</i> ^a	Gga	Metazoa	Opisthokonta	NCBI RefSeq
<i>Xenopus tropicalis</i> ^a	Xtr	Metazoa	Opisthokonta	NCBI RefSeq
<i>Danio rerio</i> ^a	Dre	Metazoa	Opisthokonta	NCBI RefSeq
<i>Takifugu rubripes</i>	Fru	Metazoa	Opisthokonta	Fugu Genome Project
<i>Ciona intestinalis</i>	Cin	Metazoa	Opisthokonta	NCBI RefSeq
<i>Anopheles gambiae</i> ^a	Aga	Metazoa	Opisthokonta	NCBI RefSeq
<i>Drosophila melanogaster</i>	Dme	Metazoa	Opisthokonta	NCBI RefSeq
<i>Apis mellifera</i> ^a	Ame	Metazoa	Opisthokonta	NCBI RefSeq
<i>Caenorhabditis briggsae</i> ^a	Cbr	Metazoa	Opisthokonta	NCBI RefSeq
<i>Caenorhabditis elegans</i>	Cel	Metazoa	Opisthokonta	NCBI RefSeq
<i>Monosiga brevicollis</i>	Mbr	Choanoflagellate	Opisthokonta	NCBI RefSeq
<i>Saccharomyces cerevisiae</i> ^a	Scs	Fungi	Opisthokonta	NCBI RefSeq
<i>Candida glabrata</i>	Cgl	Fungi	Opisthokonta	NCBI RefSeq
<i>Ashbya gossypii</i> ^a	Ago	Fungi	Opisthokonta	NCBI RefSeq
<i>Kluyveromyces lactis</i>	Kla	Fungi	Opisthokonta	NCBI RefSeq
<i>Debaryomyces hansenii</i>	Dha	Fungi	Opisthokonta	NCBI RefSeq
<i>Yarrowia lipolytica</i>	Yli	Fungi	Opisthokonta	NCBI RefSeq
<i>Magnaporthe oryzae</i> ^a	Mor	Fungi	Opisthokonta	NCBI RefSeq
<i>Neurospora crassa</i> ^a	Ncr	Fungi	Opisthokonta	NCBI RefSeq
<i>Fusarium oxysporum</i> ^a	Fox	Fungi	Opisthokonta	NCBI RefSeq
<i>Aspergillus fumigatus</i>	Afu	Fungi	Opisthokonta	NCBI RefSeq
<i>Schizosaccharomyces pombe</i>	Spo	Fungi	Opisthokonta	NCBI RefSeq
<i>Cryptococcus neoformans</i>	Cne	Fungi	Opisthokonta	NCBI RefSeq
<i>Ustilago maydis</i> ^a	Uma	Fungi	Opisthokonta	NCBI RefSeq
<i>Encephalitozoon cuniculi</i>	Ecu	Fungi	Opisthokonta	NCBI RefSeq
<i>Dictyostelium discoideum</i>	Ddi	Mycetozoa	Amoebozoa	NCBI RefSeq
<i>Entamoeba dispar</i> ^a	Edi	Archamoebae	Amoebozoa	NCBI RefSeq
<i>Entamoeba histolytica</i>	Ehi	Archamoebae	Amoebozoa	NCBI RefSeq
<i>Arabidopsis thaliana</i>	Ath	Streptophyta	Plantae	NCBI RefSeq
<i>Oryza sativa</i>	Osa	Streptophyta	Plantae	NCBI RefSeq
<i>Populus trichocarpa</i>	Pop	Streptophyta	Plantae	NCBI RefSeq
<i>Physcomitrella patens</i>	Ppa	Streptophyta	Plantae	NCBI RefSeq
<i>Chlamydomonas reinhardtii</i>	Cre	Chlorophyta	Plantae	NCBI RefSeq
<i>Volvox carteri</i>	Vca	Chlorophyta	Plantae	NCBI RefSeq
<i>Ostreococcus lucimarinus</i> ^a	Olu	Chlorophyta	Plantae	NCBI RefSeq
<i>Cyanidioschyzon merolae</i>	Cme	Rhodophyta	Plantae	U. of Tokyo
<i>Guillardia theta</i> ^b	Gth	Cryptophyta	Chromalveolata	NCBI RefSeq
<i>Bigeloviella natans</i> ^b	Bna	Cercozoa	Rhizaria	NCBI RefSeq
<i>Emiliania huxleyi</i> ^c	Ehu	Haptophyta	?	Joint Genome Institute
<i>Theileria parva</i>	Tpa	Alveolata	Chromalveolata	NCBI RefSeq
<i>Plasmodium berghei</i> ^a	Pbe	Alveolata	Chromalveolata	NCBI RefSeq
<i>Plasmodium falciparum</i>	Pfa	Alveolata	Chromalveolata	NCBI RefSeq
<i>Plasmodium knowlesi</i> ^a	Pkn	Alveolata	Chromalveolata	NCBI RefSeq
<i>Plasmodium vivax</i> ^a	Pvi	Alveolata	Chromalveolata	NCBI RefSeq
<i>Plasmodium yoelii</i> ^a	Pyo	Alveolata	Chromalveolata	NCBI RefSeq
<i>Toxoplasma gondii</i>	Tgo	Alveolata	Chromalveolata	ToxoDB
<i>Cryptosporidium hominis</i> ^a	Cho	Alveolata	Chromalveolata	NCBI RefSeq
<i>Cryptosporidium parvum</i>	Cpa	Alveolata	Chromalveolata	NCBI RefSeq

Table 1 List of species included in this study (Continued)

<i>Paramecium tetraurelia</i>	Pte	Alveolata	Chromalveolata	NCBI RefSeq
<i>Tetrahymena thermophila</i>	Tth	Alveolata	Chromalveolata	NCBI RefSeq
<i>Phaeodactylum tricornutum</i>	Ptr	Stramenopila	Chromalveolata	NCBI RefSeq
<i>Thalassiosira pseudonana</i>	Tps	Stramenopila	Chromalveolata	NCBI RefSeq
<i>Phytophthora infestans</i>	Pin	Stramenopila	Chromalveolata	NCBI RefSeq
<i>Leishmania braziliensis</i> ^a	Lbr	Kinetoplastida	Excavata	NCBI RefSeq
<i>Leishmania infantum</i> ^a	Lin	Kinetoplastida	Excavata	NCBI RefSeq
<i>Leishmania major</i>	Lma	Kinetoplastida	Excavata	NCBI RefSeq
<i>Trypanosoma brucei</i>	Tbr	Kinetoplastida	Excavata	NCBI RefSeq
<i>Trypanosoma cruzi</i> ^a	Tcr	Kinetoplastida	Excavata	NCBI RefSeq
<i>Trypanosoma vivax</i> ^a	Tvi	Kinetoplastida	Excavata	GeneDB
<i>Naegleria gruberi</i>	Ngr	Heterolobosea	Excavata	NCBI RefSeq
<i>Giardia lamblia</i>	Gla	Diplomonadida	Excavata	NCBI RefSeq
<i>Trichomonas vaginalis</i>	Tva	Parabasalia	Excavata	NCBI RefSeq

^a Results for these species are available in the Supplementary Online Material.

^b The nucleomorph genomes of these species were included in the searches. Despite the phylogenetic position of the hosts, they were members of placed close to Plantae in the tables because of the algal origin of the nucleomorphs.

^c The phylogenetic position of *Emiliana* is still unclear (see [11,12]).

Despite significant advances in our understanding of eukaryotic phylogeny, no consensus has emerged regarding the placement of the root of the eukaryotic tree (see [12-14] for a discussion). To err on the side of caution, we have based our inferences on an unrooted tree (Figure 1), and set relatively stringent criteria in order to reconstruct the different RNA export pathways in the LECA. Proteins inferred as likely to be present in the LECA had to have matches in categories 1 or 2 in members of the three megagroups. For example, a protein with strong homology matches in Opisthokonta or Amoebozoa, strong homology matches in Chromalveolata or Plantae, and strong homology matches in Excavata would be considered as likely to have already been present in the LECA. In turn, proteins with matches in the top 3 categories in members of the three megagroups were inferred as probably present in the LECA.

Three additional sets of proteins were included for validation purposes. Firstly, we included members of the SAGA transcription complex (Table S2 and S3, Additional File 1), which is specific of fungi lineages [15,16]. As an additional test, we included the proteins that make the U5snRNP spliceosomal complex as reported by Collins and Penny [17] and checked that our results were consistent with those reported previously (Table S4, Additional File 1). Finally, we compared our inferences with those of two previous studies that focused on the components of the NPC [18,19], and found that results are similar for all overlapping proteins, suggesting that the search protocols are comparable (Table S5, Additional File 1). Importantly, our analysis broadened the number of proteins analyzed and cover a much wider portion of eukaryotic diversity.

Results and discussion

In this study, we examined conservation of the proteins involved in RNA export across eukaryotes. The different RNA export pathways can be divided into two different groups. In opisthokonts, miRNAs, rRNAs, snRNAs, and tRNAs are exported through RanGTP-dependent pathways (Figure 2), whereas mRNA export follows a RanGTP-independent pathway (Figure 3). We examined conservation of the proteins involved in the different RNA export pathways across eukaryotes, with special focus on distinguishing components that are traceable to the LECA from those that derive from lineage-specific innovations. To facilitate the discussion, the different proteins were grouped according to the pathways they are involved with. In the case of proteins that are involved in more than one pathway, they were discussed in the context where their role appears to be better understood. Thus, results for the TAP-p15 heterodimer are discussed in the context of mRNA export, although TAP-p15 also plays a role in rRNA export. We present results for the most representative proteins in Tables 2 and 3, and complete results for all species are available in Tables S2 and S3 (Additional File 1).

Ran-GTP dependent RNA export (Table 2; and Tables S2 and S3 from Additional File 1)

Ran is a small, soluble GTPase present in both the nucleus and cytoplasm of all eukaryotic cells that plays a critical role in RNA export. Exportins, the nuclear export transport receptors, bind nuclear cargo together with RanGTP in the nucleus to form a ternary complex (RNA-exportin-RanGTP) that is translocated to the cytoplasm, where the complex subsequently dissociates

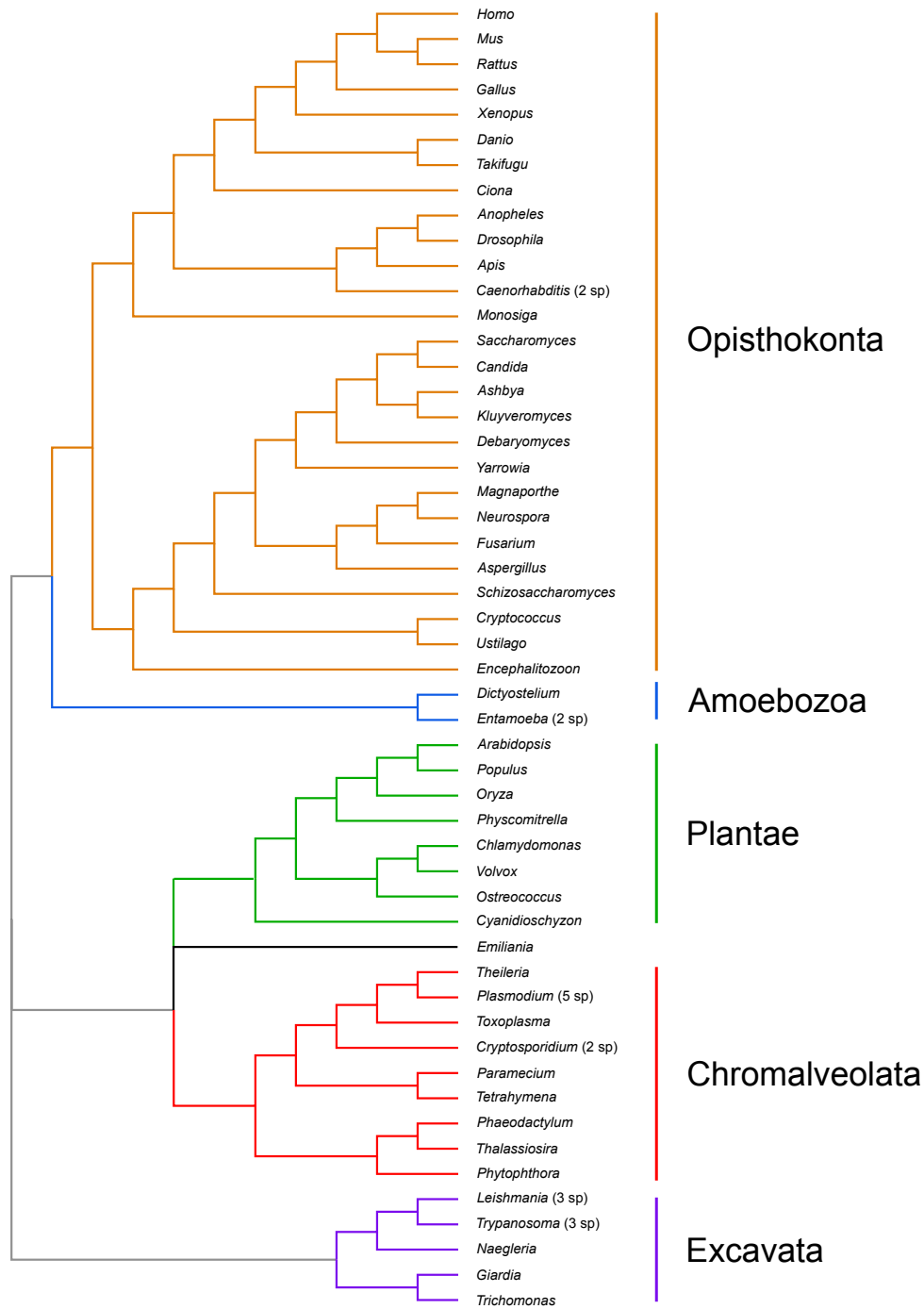


Figure 1 Phylogenetic relationships among the species included in our study, based on results from four recent studies that group the six supergroups into three different megagroups [9-12]. Relationships within Fungi follow Fitzpatrick et al. [73]. The genomes of the nucleomorphs of *Guillardia theta* and *Bigeloviella natans* were included in the bioinformatic searches but not in tree because its phylogenetic position does not match the position of the host. The different eukaryotic supergroups are highlighted. The phylogenetic position of *Emiliana* is still unclear (see [11,12]), and thus, it was left in black.

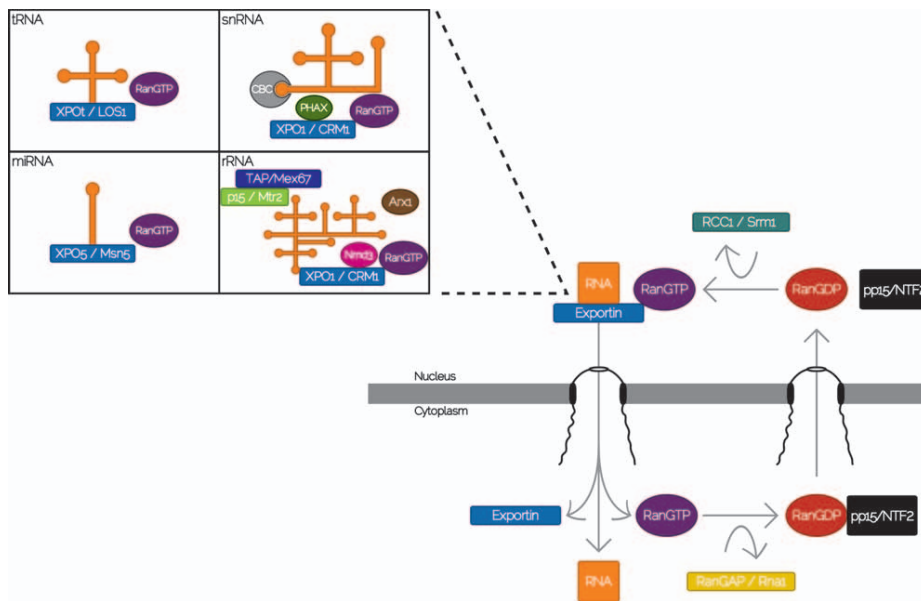


Figure 2 Schematic representation of the RanGTP-dependent RNA export pathways in yeast (see [4] for a full description). Nuclear export of miRNA, rRNA, snRNA, and tRNA has been shown to follow the RanGTP-dependent exportin model. The nuclear export receptors, exportins, bind nuclear cargo together with RanGTP in the nucleus to form a ternary complex (RNA-exportin-RanGTP) that is translocated to the cytoplasm. The complex dissociates in the cytoplasm upon hydrolysis of RanGTP to release the cargo molecule. Different exportins are preferentially involved with the different RNA export pathways. Exportin t (XPOt) is involved with tRNA export, Exportin 5 (XPO5) in miRNA export whereas Exportin 1 (XPO1) is in charge of snRNA and rRNA export, as shown in the corresponding box. In the case of snRNA and rRNA export, additional adaptors are needed. After GTP hydrolysis in the cytoplasm, the import receptor, NTF2, carries RanGDP into the nucleus, where nucleotide exchange occurs by RCC1 to generate RanGTP.

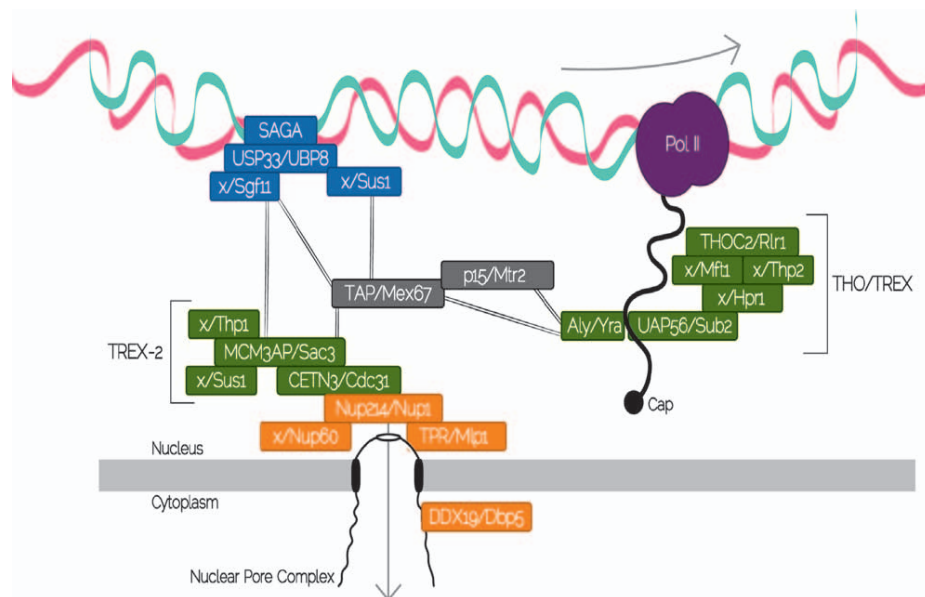


Figure 3 Schematic representation of the mRNA export pathway in yeast (see [4] for a full description). mRNAs associate with protein factors into messenger RNPs (mRNPs) which are then exported through the NPC by the TAP/p15 heterodimer (Mex67/Mtr2 in yeast), that functions as a general mRNA export receptor. TAP/p15 operates in association with TREX (Transcription/Export), a complex that connects transcription with mRNA export and the THO complex. In yeast there is an additional complex named TREX-2, that is also capable of mediating the nuclear export of mRNPs. Each box represents a human/yeast protein and the absence of human predicted orthologs, is represented by "X". Proteins are colored according to the stage in which they act and lines represent interactions between the complexes. TREX: transcription-coupled to export complex. Pol II: RNA Polymerase II.

Table 3 Bioinformatic results for proteins involved the RanGTP-independent mRNA export pathway

Queries	Opisthokonta									Amo				Plantae							Chromalveolata							Excavata									
	Hsa	Cin	Dme	Cel	Mbr	Cgl	Afu	Spo	Cne	Ecu	Ddi	Ehi	Ath	Osa	Pop	Ppa	Cre	Vca	Cme	Ehu	Tpa	Pfa	Tgo	Cpa	Pte	Tth	Ptr	Tps	Pin	Lma	Tbr	Ngr	Gla	Tva			
Hsa_CETN3	1	1	4	1	1	1	4	1	1	2	1	1	1	1	1	1	1	1	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
Scce_Cdc31p	1	1	4	1	1	1	4	1	1	2	1	2	1	1	1	1	1	1	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
Hsa_MCM3AP	1	3	5	5	5	5	5	5	5	5	5	5	4	5	5	5	5	5	5	5	NC	NC	5	5	5	5	5	5	5	5	5	5	5	5	5	5	
Scce_Sac3p	4	5	5	5	5	1	5	4	5	5	5	NC	5	5	5	5	5	5	5	5	NC	NC	NC	NC	NC	NC	5	5	5	5	5	5	5	5	5	5	
Hsa-no ortholog																																					
Scce_Thp1p	4	4	3	3	3	1	NC	3	NC	NC	3	NC	3	4	NC	4	NC	NC	3	NC	NC	NC	NC	NC	3	NC	4	4	4	NC	NC	4	NC	NC	NC	NC	
Hsa-no ortholog																																					
Scce_Sus1p	2	1	2	NC	2	NC	NC	2	2	NC	2	NC	NC	NC	2	2	3	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	
Hsa_THOC2	1	1	1	2	4	3	4	2	4	NC	4	5	2	2	2	2	NC	4	5	4	4	5	4	NC	3	3	3	3	2	NC	NC	3	NC	NC	NC		
Scce_Rlr1p	3	3	3	3	4	1	4	3	3	NC	4	5	3	3	4	3	NC	4	NC	NC	NC	5	NC	NC	4	4	4	4	4	4	NC	NC	4	NC	NC	4	
Hsa-no ortholog																																					
Scce_Tho1p	NC	NC	NC	NC	NC	2	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	5	NC	NC	NC	NC	NC	NC	NC	
Hsa-no ortholog																																					
Scce_Hpr1p	NC	NC	NC	NC	NC	2	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	
Hsa-no ortholog																																					
Scce_Mft1p	NC	NC	NC	NC	NC	2	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	
Hsa-no ortholog																																					
Scce_Thp2p	NC	NC	NC	NC	NC	1	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	
Hsa_UAP56	1	1	1	1	1	1	1	1	1	2	1	1	1	1	1	1	1	1	1	4	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2	1
Scce_Sub2	1	1	1	1	1	1	1	1	1	2	1	1	1	1	1	1	1	1	1	4	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2	1
Hsa-no ortholog																																					
TREX																																					
Scce_Yra1p	NC	NC	3	2	NC	1	3	3	2	NC	NC	4	NC	4	2	2	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	3	NC	NC	NC	NC	NC	NC	4	
Hsa_THOC3	1	1	1	2	2	3	2	2	4	3	4	3	1	1	1	1	2	2	4	1	3	4	4	4	2	2	2	4	2	3	4	4	3	3	3		
Scce_Tex1p	3	5	3	3	3	1	NC	4	4	NC	3	5	3	3	3	3	4	NC	4	NC	3	NC	NC	3	3	NC	NC	3	NC	NC	NC	NC	NC	NC	5		
Hsa_TAP	1	2	2	2	3	3	2	2	3	NC	4	4	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	3	NC	NC	
Scce_Mex67p	3	3	3	NC	NC	1	2	3	3	NC	4	4	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	5	5	NC	NC	NC	NC	5	5	4	NC	NC	NC	NC		
Hsa_p15	1	2	1	2	NC	3	1	4	2	NC	3	3	2	3	2	3	NC	NC	3	NC	3	NC	NC	3	NC	NC	NC	3	2	2	3	NC	NC	3	3		
Scce_Mtr2p	NC	NC	NC	NC	NC	1	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	
Hsa_DDX19	1	1	1	1	1	1	1	1	1	2	1	2	1	1	1	1	1	1	1	1	1	4	1	1	1	1	1	2	1	1	2	1	1	2	1	2	
Scce_Dbp5p	1	1	1	1	1	1	1	1	1	2	1	2	2	2	2	1	1	1	1	2	1	4	1	1	1	2	2	2	1	2	2	1	2	2	1	2	
Hsa-no ortholog																																					
Scce_Gle1p	4	3	3	NC	NC	1	NC	2	3	NC	NC	NC	NC	3	NC	NC	5	NC	4	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	4	NC	NC
Hsa_RAE1	1	1	1	1	2	2	2	1	1	2	1	3	1	1	1	1	1	1	2	2	2	4	2	2	4	2	1	1	1	3	3	2	3	3	3		
Scce_Gle2p	2	2	1	2	2	1	1	1	1	2	1	3	2	1	1	1	1	1	2	2	2	NC	2	1	4	2	1	1	1	3	2	2	3	2	2	2	
Hsa_TPR	1	5	3	4	NC	4	3	5	5	NC	4	NC	3	5	3	5	NC	4	NC	NC	NC	NC	NC	NC	NC	NC	3	3	3	NC	NC	5	NC	NC	NC		
Scce_Mlp1p	2	4	3	NC	NC	1	2	3	NC	NC	NC	NC	4	NC	4	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	5	4	5	NC	NC	5	NC	NC	NC		
Hsa-no ortholog																																					
Scce_Nab2p	4	NC	NC	NC	NC	1	3	4	3	NC	NC	NC	NC	NC	NC	NC	NC	3	NC	NC	NC	4	5	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	NC	
Hsa_HNRNPM	1	3	4	4	4	4	4	4	4	4	NC	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	
Scce_Gbp2p	4	2	4	4	3	2	3	3	4	4	4	4	4	4	4	4	4	4	4	3	4	4	4	4	4	4	4	4	3	4	2	4	4	4	4		

Results were ranked into 5 categories (1-5; with 1 being the most conserved). NC (Not Classified). We present results for a representative subset of species in here, and full results are presented in Table S2 (Additional File 1)

upon hydrolysis of RanGTP to release the cargo molecule. Nuclear export of miRNA, rRNA, snRNA, and tRNA has been shown to follow the RanGTP-exportin model. Bioinformatic searches showed that Ran was well conserved in all species included in this study, with a Ran CHS in category 1 (see categories definition in the Methods section) in almost all genomes surveyed, and high levels of sequence conservation. The Ran orthologs from human and *Trypanosoma brucei* are 72.2% identical/79% similar, whereas the *S. cerevisiae* and *T. brucei* orthologs are 70.51% identical/79% similar. The high conservation scores and phyletic distribution of Ran CHS indicate that the LECA was highly likely to have possessed an ortholog of Ran, and extends previous work that included fewer eukaryotic species in their analyses [18,19].

Ran Binding Protein 1 (RanBP1, Yrb1p in yeast), another protein of the Ran cycle, had a high level of conservation as well. RanBP1 stimulates RanGTP hydrolysis by RanGAP [20,21]. Because we found strong CHS for RanBP1 in representatives of the three eukaryotic megagroups (Table 2), it is likely that RanBP1 was present in the LECA. Experimental evidence showing that RanBP1 is involved in the Ran cycle in the Excavate *Leishmania major* [22] would suggest its function has been conserved throughout eukaryotic evolution.

RCC1 (Srm1 in yeast) and RanGAP (Rna1 in yeast) are responsible for the establishment of a RanGTP-RanGDP gradient across the nuclear envelope that drives RanGTP dependent transport. These two proteins were less conserved than either Ran or RanBP1 in our searches (Table 2). RCC1 acts in the nucleus as the guanine nucleotide exchange factor, and RanGAP acts in the cytoplasm and regulates the GTPase activity of Ran. We only found CHS for the human or yeast RCC1 and RanGAP orthologs in categories 1 and 2 in opisthokonts (Table 2). Because we found some CHS in category 3 for RCC1 and RanGAP among plants, chromalveolates, and excavates, we infer that these two proteins were probably present in the LECA (Table 2).

The proteins involved in the Ran cycle discussed above play role in intracellular transport of proteins and nucleic acids. Their strong conservation across all species examined is in line with previous reports [18-23]. Further, these results are consistent with the predictions of the evolutionary scenario presented by Jékely [24], where RanGTP-dependent transport is strongly related to the origin of the nucleus. In addition to the proteins involved in the Ran cycle, other export factors are required to provide specificity for the different RNA export mechanisms. In the following sections we discuss conservation for the proteins that are specifically involved in the different RNA export pathways.

MicroRNA and tRNA export pathways

MicroRNAs are short non-coding RNAs involved in the regulation of gene expression in eukaryotic cells [25], and tRNAs are non-coding RNAs that transfer specific amino acids to a growing polypeptide chain. In the case of miRNA and tRNA export, the exportins involved in the export of miRNA precursors in human and yeast is exportin-5 (XPO5, Msn5 in yeast), whereas XPOt (Los1 in yeast) is the one involved in tRNA export [26].

There were no conservation scores better than 3 for yeast XPO5 other than its CHS in yeast (Table 2). Conversely, conservation scores outside members of Metazoa were 3 or worse in searches involving human XPO5. These results agree with previous studies that had failed to find XPO5 homologs in members of Euglenozoa and Apicomplexa [27,28]. Interestingly, several of the species that lack XPO5 homologs, such as *T. brucei* or *Trichomonas vaginalis*, possess a functional miRNA machinery [29,30], suggesting that there are additional pathways for the export of miRNA other than the one used in humans. A simple explanation for the complex phyletic distribution of XPO5 CHS is difficult to reconcile with current eukaryotic phylogenies. There are known orthologs of XPO5 in Plantae, and Opisthokonta, and there are CHS of score 3 in the three different megagroups, which would suggest that an XPO5 ortholog was probably present in the LECA and would have been secondarily lost in several lineages. Under alternative eukaryotic phylogenies [9-14] the presence of orthologs in plants and opisthokonts would suggest that XPO5 was probably present in the LECA, as previously suggested [28].

In the case of XPOt, CHS for the yeast ortholog in category 1 were only found in Fungi, whereas CHS for the human ortholog in categories 1 and 2 were found among members of Metazoa and Amoebozoa, respectively (Table 2). In addition, we found CHS in category 3 or higher for the human XPOt ortholog in members of Opisthokonta, Plantae, Chromalveolata and Excavata. Although some of the CHS in *Leishmania* and *Trypanosoma* (Excavata) and *Plasmodium* (Chromalveolata) are not strongly conserved, phylogenies show that they are true orthologs [28]. Orthologs of XPOt are involved in tRNA export in yeast and plants, but they are not an essential factor in some species. For example, nucleocytoplasmic export of tRNA is not fully blocked when the XPOt gene is deleted in *S. cerevisiae*, *S. pombe* and *A. thaliana* [31-33]. Similar to XPO5, orthologs of XPOt are present in most of the major eukaryotic supergroups in this study, which would imply that the LECA was likely to possess an ortholog of this gene.

XPO5 has also been postulated to play a minor role in tRNA export for some organisms. In this case, XPO5

mediates the tRNA transport in association to a ribosomal elongation factor named eEF1A, TEF2 in yeast [34,35]. This ribosomal elongation factor is also highly conserved, with CHS in category 1 for most species in this study (Table 2), once again, suggesting that it was likely present in the LECA.

Small nuclear RNAs export pathway

Small nuclear RNAs are a group of non-coding transcripts involved in RNA splicing, transcription factor regulation and telomere maintenance. The different snRNAs are synthesized in the nucleus, assembled into snRNPs in the cytoplasm and re-imported into the nucleus. Export of snRNAs presents some departures from the simpler model of exportin-RanGTP-cargo seen in miRNAs and tRNAs. In this case the exportin does not bind RNA directly, and additional adaptor proteins are needed. Exportin-1 (XPO1, CRM1 in yeast) is the transport receptor in charge of carrying snRNAs to the cytoplasm, in association to CBP20 (CBC2 in yeast), which binds to the snRNA; CBP80 (Sto1 in yeast), which ensures high affinity binding; and PHAX, which provides the nuclear export signal [36,37].

Our analyses indicate that XPO1 is the most conserved exportin among the three included in our study (Table 2). There were CHS for human and yeast orthologs in categories 1 and 2 in most of eukaryotic lineages, and sequence conservation was also high. The human and *T. brucei* XPO1 orthologs were 32% identical/52% similar. In the case of CBP20 we found CHS in categories 1 or 2 in all eukaryotic supergroups, whereas we failed to find CHS of CBP80 in several species (Table 2). By contrast, conservation scores for PHAX are either low, or below our detection threshold for all eukaryotes in this study other than members of Metazoa, in agreement with a previous study [36]. These results indicate that the LECA was likely to have XPO1 and CBP20 orthologs, but that PHAX appears to be a lineage-specific innovation. The case for CBP80 is more complex, there are some matches in category 3 in all eukaryotic megagroups, but no CHS were identified in several species. We infer that CBP80 could have been present in the LECA, but that it was either lost or diverged beyond our ability to detect it in several species.

rRNA export pathway

Ribosomes are made up of a combination of different rRNAs and a large number of ribosomal proteins that are organized into the large (60S) and small (40S) pre-ribosomal subunits. In yeast, nuclear export of the pre-ribosomal subunits depends on XPO1 and RanGTP [38,39], and also involves additional proteins. In yeast, the nuclear export of the pre-60S ribosomal subunit requires the proteins Nmd3, Arx1 and the Mex67/Mtr2 heterodimer, but export of 40S is poorly understood [40-42]. The role of Nmd3, Arx1 and the Mtr2/Mex67

has not been completely elucidated yet. On the one hand, deletion of the Arx1 gene leads to pre-60S accumulation in the nucleus [43]. On the other hand, a subsequent study from the same group shows that in the absence of Arx1, the addition of a nuclear export signal to the pre-60S subunit would be enough to restore its export [44], which suggests that the export of pre-60S is relatively flexible.

XPO1 conservation scores have already been discussed with the other snRNA export proteins. Results for the two subunits of the general transport adapter, (TAP/p15 in human, Mex67/Mtr2 in yeast) are low (Table 3) and will be discussed in the next section because of their major role in mRNA export. The phyletic distribution of XPO1, XPO5, and XPO5 suggests that these exportins are involved with RNA export in all the different eukaryotic supergroups analyzed, and also indicate that each of them is traceable to the LECA. We observed that XPO1, which is mostly involved in rRNA export in opisthokonts, is the most conserved one, while XPO1 and XPO5, involved in miRNA, tRNA, snRNA export, are not as conserved. Even though orthologs of these exportins are found among excavates and chromalveolates, their role in these lineages is not fully understood. The little functional data available suggest that their roles are not entirely the same as in opisthokonts. For example, XPO1 has been shown to be involved with mRNA export in the parasites of the genus *Trypanosoma* [45], but it is mostly involved with rRNA export in opisthokonts. Based on the functional data and their phyletic distribution, we speculate that there is some level redundancy in the functional role of the different exportins and propose that in the LECA, XPO1 could have acted as the protein responsible for transport of most of the RNAs through the NPC.

For the additional rRNA export proteins, our analyses show that Nmd3 is well conserved in the vast majority of species included in this study (Table 2), with 58% similarity/38% identity in comparisons between human and *T. brucei* orthologs. This would indicate that Nmd3 is highly likely to have been present in the LECA. By contrast some of the additional adaptors, such as Arx1, Alb1p, and PHAX have few matches in category 3, and seem to be lineage specific innovations. Strong CHS for Arx1 are only found among Fungi, strong CHS for Alb1p are restricted to yeasts in the subphylum Saccharomycotina, whereas strong CHS for PHAX are only found among animals.

As in the case with the exportins, the phyletic distribution of the transport adaptors discussed above suggests that these proteins are involved with export of tRNA, miRNAs, snRNAs and rRNAs in all the different eukaryotic supergroups, and that they are traceable to the LECA, such as Nmd3, which also appear to be

conserved among the different eukaryotic supergroups. The specific combination of exportins and adaptors varies in a lineage-specific manner through a combination of gene gains and losses. For example, XPO5 has been secondarily lost in fruit fly, whereas XPOt has been secondarily lost in *C. elegans*. On the other hand, we also found that Arx1, Alb1p and PHAX are restricted to different groups within Opisthokonta, suggesting that they are innovations specific to this group. These results would suggest that the specific export mechanism for the different RNAs varies in a lineage-specific manner.

RanGTP-independent transport: mRNA export pathway (Table 3; and Tables S2 and S3 from Additional File 1)

In Metazoa and Fungi, nucleocytoplasmic export of the majority of the mRNAs does not follow the RanGTP-dependent exportin pathway. From yeast to humans, mRNAs associate with protein factors into messenger RNPs (mRNPs) which are then exported through the NPC by an essential shuttling heterodimer, TAP/p15 in human and Mex67/Mtr2 in yeast, which functions as a general mRNA export receptor to transport mRNPs through the NPCs [46-48]. The dimeric export receptor operates in association with TREX (Transcription/Export), a multiprotein complex that connects transcription with mRNA export [49,50]. Human TREX consists of the RNA helicase UAP56 (Sub2 in yeast), the RNA-binding adaptor protein Aly (Yra1 in yeast) and the THO complex [51-53]. In yeast, an additional complex named TREX-2, consisting of Thp1, Sac3, Sus1, and Cdc31 is also capable of mediating the nuclear export of mRNPs in concert with Mex67/Mtr2 [54-56].

Our analyses indicate that few of the proteins involved in the mRNA export pathway are conserved across all eukaryotes. This is perhaps not unexpected given the differences in mRNA processing among eukaryotes. In the case of the TAP or Mex67 subunit of the heterodimer, conservation scores outside Fungi and Metazoa were 4 or 5 with the exception of some Excavata members. Whereas bioinformatic searches for the yeast Mtr2 subunit only identified strong CHS in fungi. This is not a surprising result, as similarities between Mtr2 and its human functional analog, p15, are restricted to three-dimensional structure [57], which cannot be detected by sequence similarity analyses. In the case of TREX, the UAP56 subunit (Table 3) is well conserved across all eukaryotes: we have strong CHS for the yeast ortholog in all species in this study except for *Emiliana huxleyi*. The Thoc3 subunit, Tex1 in yeast (Table 3), has good CHS among opisthokonts and plants, and moderate candidates in Amoebozoa, Chromalveolata, and Excavata. By contrast, we could not identify CHS for the yeast Yra1 (Table 3) in several Chromalveolata and Excavata. A similar situation is observed for the TREX-2 complex,

where only CETN3, CDC31 in yeast (Table 3), is well conserved across all eukaryotes. CETN3 is a centrin involved in the duplication and segregation of the centrosome during cytokinesis that is also involved in mRNA export in yeast [55]. However, evidence for a role of CETN3 in mRNA export in some members of Excavata is lacking, as it is only required for the initiation of cytokinesis in *Leishmania* and *T. brucei* [58].

The cellular fate of the different mRNAs depends mainly on the shuttle proteins TAP/p15, Yra1, Rae1 and Gle1, DDX19, which are also required for proper translation [4,59]. Rae1 acts in delivering TAP to the NPC [60] and DDX19 is responsible for triggering the mRNPs remodeling at the NPC cytoplasmic-face scaffold, delivering the mRNA to the cytoplasm [61]. Both Rae1 and DDX19 are well conserved, with CHS in categories 1 and 2 for all the supergroups. However, the adaptor protein Gle1, that plays an important role in the control of mRNA export in yeast [62], appears to be missing from Amoebozoa, Plantae, Chromalveolata, and Excavata, suggesting it is an innovation specific to Opisthokonta.

The proteins involved in the RanGTP-independent export of mRNA are generally less conserved than those involved in RanGTP-dependent export of all other RNA species. Some of the major components of the mRNA export pathway are conserved across the different eukaryotic supergroups examined. This is the case for UAP56, DDX19, and Rae1. However, there are many proteins that play essential roles in mRNA export in opisthokonts that lack strong CHS outside this group, and appear to represent opisthokont-specific innovations. This group includes Mlp1, Mlp2, and the shuttling proteins TAP, Yra1, Nab2, and Gbp2. These results suggest that mRNA export in members of the other supergroups might follow slightly different mechanisms. One possibility is that mRNA export might also follow a RanGTP-dependent pathway. Interestingly, the fact that in the excavate *Trypanosoma* the XPO1 ortholog is involved in mRNA export [45] is consistent with our inference. We speculate that mRNA might have been exported by XPO1 in a RanGTP-dependent manner early in eukaryote evolution, and that it later followed more specialized pathways.

Conclusions

Results from our genomic comparisons indicate that several of the key proteins involved in the different RanGTP-dependent RNA export pathways are conserved across most eukaryotic lineages, and thus we infer that orthologs of them were highly likely to have been already present in the LECA. Examples of these are the exportins XPO1 and XPOt, Nmd3 and most of the proteins involved in the Ran cycle. Despite the relatively strong level of conservation, we also document how these export mechanisms vary in a lineage-specific manner as a consequence of the

differential gene gains and losses, as documented by the secondary loss of XPO1 in *C. elegans*, and the emergence of Arx1 in Fungi and of PHAX in animals. In agreement with inferences drawn from studies of components of the cytoskeleton, endomembrane, NPC, and spliceosome [17-19,62-67] a number of key RNA nucleocytoplasmic transport factors can be traced to the LECA. This would suggest that the different RanGTP-dependent RNA export mechanisms were already present in the LECA.

By contrast, our analyses indicate that several of the proteins involved in the RanGTP-independent export pathway from opisthokonts are lineage-specific innovations. The mRNA export pathway is the most complex and the least conserved among those examined in this study. We found CHS for few of the yeast proteins involved in mRNA export in most eukaryotic lineages, which suggests that mRNA export among them is different from what is observed in yeast. We document the acquisition of several lineage-specific innovations in the

mRNA export mechanisms of opisthokonts relative to the other supergroups included in this study. It seems plausible that the evolution of a RanGTP-independent mRNA export pathway in opisthokonts might be related to the observed differences in the regulation of gene expression. In Amoebozoa, Chromalveolata and Excavata, most regulation is post-transcriptional [68], whereas in human and yeast the presence of both transcriptional and post-transcriptional regulation is linked to a more refined control of gene expression. One possibility is that having an mRNA export pathway fully separated from the other RNA export pathways increases the range possibilities of gene-specific control. This would be particularly important to allow both coordination and versatility in gene expression control [69,70], as well as to open additional opportunities to connect active transcriptional sites to the NPC, allowing a fine control of gene expression in yeast and human [71,72]. Functional data regarding mRNA export in

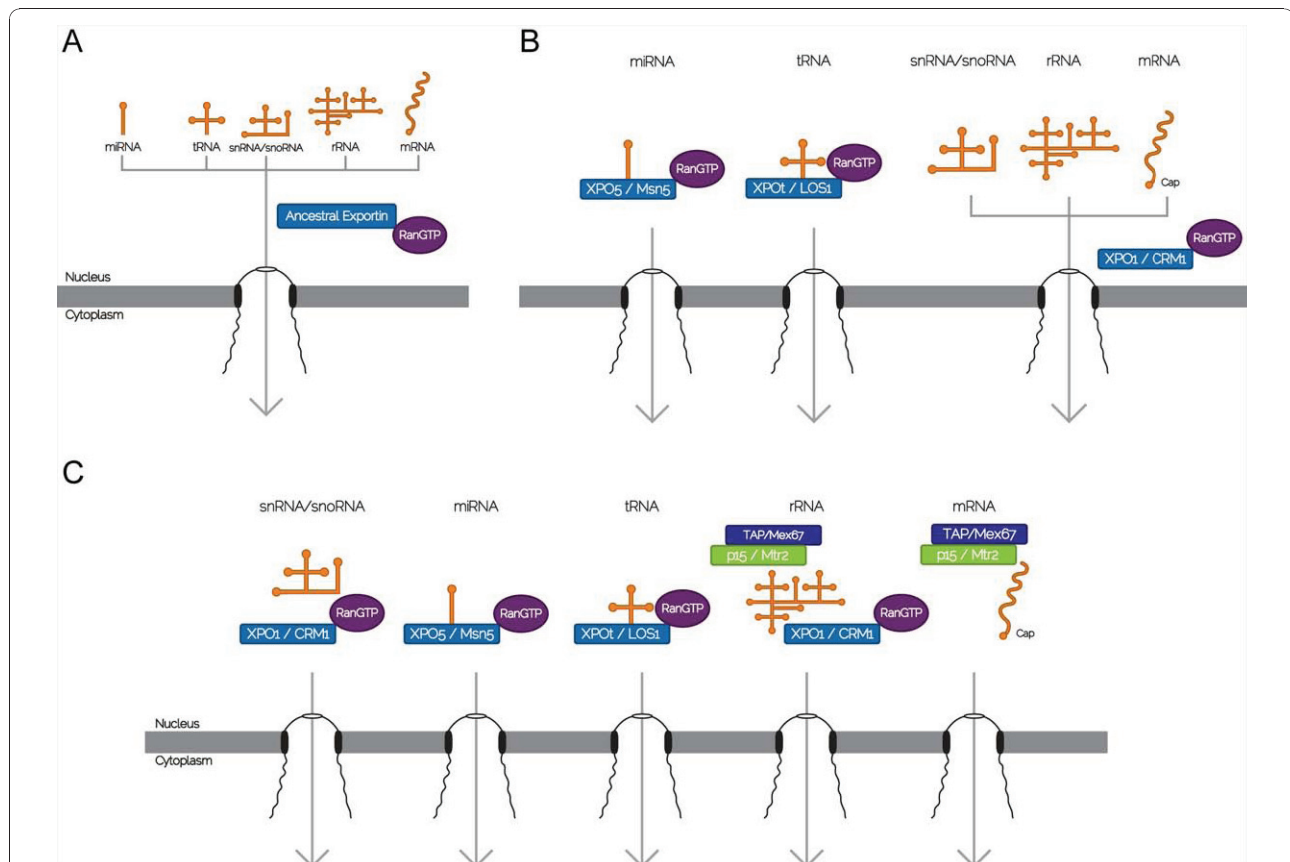


Figure 4 Model proposed of the evolution of the different RNA export pathways throughout eukaryotic history. (A) Early in the history of eukaryotes a single ancestral exportin similar to XPO1 was in charge of nucleocytoplasmic export of all RNAs in a RanGTP-dependent manner. (B) Prior to the emergence of the LECA, different exportins originated and specialized in nucleocytoplasmic export of the different RNA species. (C) In opisthokonts, a further innovation was the emergence of a RanGTP-independent pathway of mRNA export. This pathway relies on the TAP-p15 heterodimer as the transport receptor instead of an exportin, and includes several other adaptors which are restricted to opisthokonts. Each box represents a human/yeast protein.

excavates and chromalveolates is still limited, and despite the fact that some studies suggest it follows less complex routes [45], the possibility that alternative specializations have arisen in these groups cannot be discarded.

Taken together, our analyses suggest that the LECA possessed most of the basic components of the different RanGTP-dependent RNA export mechanisms, which are also well conserved among the different eukaryote supergroups included in this study. In addition, we also show that some of the major components of the RanGTP-independent mRNA export pathways can also be traced to the LECA, but that several of the proteins that play key roles in opisthokonts derive from lineage-specific innovations. It is likely that early in eukaryote evolution a single generalized ancestral exportin was probably responsible for nucleocytoplasmic transport of all RNA species (Figure 4A). Prior to the emergence of the LECA, orthologs of XPO1, XPO5, and XPOt would have already emerged and specialized in transporting a subset of the RNAs (Figure 4B). We speculate that in the LECA mRNA might have been exported by XPO1 in RanGTP-dependent manner, as in excavates in the genus *Trypanosoma*. The emergence of the RanGTP-independent mRNA export pathway in opisthokonts coincides with the origin of several lineage-specific innovations, and might be related to refinements in the regulation of gene expression in this supergroup (Figure 4C).

Additional material

Additional File 1: Spreadsheet file containing tables S1-S5. Table S1. Full list of human and yeast proteins used as queries. ^a- yeast orthologs in human, according to Ensembl. ^b- Not described as a yeast ortholog in Ensembl, used as negative control. **Table S2.** Results for all proteins in all species with BLASTP (BLOSUM62) and PSI-BLAST (2 iterations) ranked in 1-5 code according to Methods. Code: Category of similarity, from 1 (most conserved) to 5. RBH: Reciprocal Best Hit, 1:True; 0: False. NC: Not Classified. NA: Not Analyzed. **Table S3.** Individual results for each protein, using different algorithms for ranking 1-5 code. I: % Identity, S: %Similarity. QC: Query Coverage. **Table S4.** Comparative analysis of U5snRNP spliceosomal complex using 1-5 ranking according to Methods in all species. **Table S5.** Comparison of the protein repertoire in relation to further evolutionary studies. In blue, common proteins analyzed in at least one study. In red, proteins that were not analyzed in previous works.

Additional File 2: Exemplary alignments of proteins from the different categories.

Acknowledgements

This work was funded by grants from the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) to FGH (CNPq 476739/2008-0) and ARA, grants to ARA from Fundação Araucária, and grants to SG from Fiocruz and CNPq. MS was funded by a studentship from CNPq, NMV was funded by a studentship from Fundação Araucária, FGH was funded by a postdoctoral fellowship from CNPq, and SG is a research fellow from CNPq. We thank J. Peter Gogarten and three anonymous reviewers for helpful suggestions on the manuscript, Christian M. Probst for bioinformatic assistance, Paula Emilia Kranz for help with the figures, and Fernando

Alvarez-Valín, Alejandro Correa, Augusto Ramos, Florencia Meyer, and Bruno Dallagiovana for valuable comments.

Author details

¹Departamento de Biologia Celular e Molecular, Universidade Federal do Paraná - UFPR, Curitiba, Brazil. ²Instituto Carlos Chagas - ICC, Curitiba, Brazil. ³School of Biological Sciences, University of Nebraska, Lincoln, USA. ⁴Department of Biochemistry and Molecular Biology, Mississippi State University, Mississippi State, USA.

Authors' contributions

MS, NMV, ARA and FGH designed the research. MS and ARA identified the yeast and human queries. NMV and MS did the bioinformatic searches. MS, NMV, SG, ARA and FGH discussed the results. MS, ARA and FGH wrote the manuscript. All authors read and approved the final manuscript.

Received: 15 December 2009 Accepted: 11 January 2011

Published: 11 January 2011

References

1. Martin W, Koonin EV: Introns and the origin of nucleus-cytoplasmic compartmentalization. *Nature* 2006, **440**:41-45.
2. Fried H, Kutay U: Nucleocytoplasmic transport: taking an inventory. *Cell Mol Life Sci* 2003, **60**:1659-1688.
3. Cullen BR: Nuclear RNA export. *J Cell Sci* 2003, **116**:587-597.
4. Köhler A, Hurt E: Exporting RNA from the nucleus to the cytoplasm. *Nat Rev Mol Cell Biol* 2007, **8**:761-773.
5. Rodríguez-Navarro S, Fischer T, Luo M, Antúnez O, Brettschneider S, Lechner J, Pérez-Ortín JE, Reed R, Hurt E: Sus1, a functional component of the SAGA histone acetylase complex and the nuclear pore-associated mRNA export machinery. *Cell* 2004, **116**:75-86.
6. Wolyniak MJ, Cole CN: Harnessing genomics to explore the processes and evolution of mRNA export. *RNA Biol* 2008, **5**:68-72.
7. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ: Basic local alignment search tool. *J Mol Biol* 1990, **215**:403-410.
8. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL: BLAST+: architecture and applications. *BMC Bioinformatics* 2009, **10**:421.
9. Burki F, Shalchian-Tabrizi K, Pawlowski J: Phylogenomics reveals a new 'megagroup' including most photosynthetic eukaryotes. *Biol Lett* 2008, **4**:366-369.
10. Hampl V, Hug L, Leigh JW, Dacks JB, Lang BF, Simpson AGB, Roger AJ: Phylogenomic analyses support the monophyly of Excavata and resolve relationships among eukaryotic "supergroups". *Proc Natl Acad Sci USA* 2009, **106**:3859-3864.
11. Burki F, Inagaki Y, Bråte J, Archibald JM, Keeling PJ, Cavalier-Smith T, Sakaguchi M, Hashimoto T, Horak A, Kumar S, Klaveness D, Jakobsen KS, Pawlowski J, Shalchian-Tabrizi K: Large-scale phylogenomic analyses reveal that two enigmatic protist lineages, Telonemia and Centroheliozoa, are related to photosynthetic chromalveolates. *Genome Biol Evol* 2009, **1**:231-238.
12. Parfrey LW, Grant J, Tekle YI, Lasek-Nesselquist E, Morrison HG, Sogin ML, Patterson DJ, Katz LA: Broadly sampled multigene analyses yield a well-resolved eukaryotic tree of life. *Syst Biol* 2010, **59**:518-533.
13. Yoon HS, Grant J, Tekle YI, Wu M, Chaon BC, Cole JC, Logsdon JMJ, Patterson DJ, Bhattacharya D, Katz LA: Broadly sampled multigene trees of eukaryotes. *BMC Evol Biol* 2008, **8**:14.
14. Rogozin IB, Basu MK, Csürös M, Koonin EV: Analysis of rare genomic changes does not support the unikont-bikont phylogeny and suggests cyanobacterial symbiosis as the point of primary radiation of eukaryotes. *Genome Biol Evol* 2009, **2009**:99-113.
15. Cabal GG, Genovesio A, Rodríguez-Navarro S, Zimmer C, Gadal O, Lesne A, Buc H, Feuerbach-Fournier F, Olivo-Marin J, Hurt EC, Nehrbass U: SAGA interacting factors confine sub-diffusion of transcribed genes to the nuclear envelope. *Nature* 2006, **441**:770-773.
16. Kurshakova MM, Krasnov AN, Kopytova DV, Shidlovskii YV, Nikolenko JV, Nabirochkina EN, Spehner D, Schultz P, Tora L, Georgieva SG: SAGA and a novel *Drosophila* export complex anchor efficient transcription and mRNA export to NPC. *EMBO J* 2007, **26**:4956-4965.
17. Collins L, Penny D: Complex spliceosomal organization ancestral to extant eukaryotes. *Mol Biol Evol* 2005, **22**:1053-1066.

18. Mans BJ, Anantharaman V, Aravind L, Koonin EV: **Comparative genomics, evolution and origins of the nuclear envelope and nuclear pore complex.** *Cell Cycle* 2004, **3**:1612-1637.
19. Bapteste E, Charlebois RL, MacLeod D, Brochier C: **The two tempos of nuclear pore complex evolution: highly adapting proteins in an ancient frozen structure.** *Genome Biol* 2005, **6**:R85.
20. Bischoff FR, Görlich D: **RanBP1 is crucial for the release of RanGTP from importin beta-related nuclear transport factors.** *FEBS Lett* 1997, **419**:249-254.
21. Lounsbury KM, Macara IG: **Ran-binding protein 1 (RanBP1) forms a ternary complex with Ran and karyopherin beta and reduces Ran GTPase-activating protein (RanGAP) inhibition by karyopherin beta.** *J Biol Chem* 1997, **272**:551-555.
22. Casanova M, Portalès P, Blaineau C, Crobu L, Bastien P, Pagès M: **Inhibition of active nuclear transport is an intrinsic trigger of programmed cell death in trypanosomatids.** *Cell Death Differ* 2008, **15**:1910-1920.
23. Jékely G: **Small GTPases and the evolution of the eukaryotic cell.** *Bioessays* 2003, **25**:1129-1138.
24. Jékely G: **Origin of the nucleus and Ran-dependent transport to safeguard ribosome biogenesis in a chimeric cell.** *Biol Direct* 2008, **3**:31.
25. Bartel DP: **MicroRNAs: genomics, biogenesis, mechanism, and function.** *Cell* 2004, **116**:281-297.
26. Kim VN, Han J, Siomi MC: **Biogenesis of small RNAs in animals.** *Nat Rev Mol Cell Biol* 2009, **10**:126-139.
27. Frankel MB, Knoll LJ: **The ins and outs of nuclear trafficking: unusual aspects in apicomplexan parasites.** *DNA Cell Biol* 2009, **28**:277-284.
28. Murphy D, Dancis B, Brown JR: **The evolution of core proteins involved in microRNA biogenesis.** *BMC Evol Biol* 2008, **8**:92.
29. Djikeng A, Shi H, Tschudi C, Ullu E: **RNA interference in *Trypanosoma brucei*: cloning of small interfering RNAs provides evidence for retroposon-derived 24-26-nucleotide RNAs.** *RNA* 2001, **7**:1522-1530.
30. Lin W, Li S, Lin W, Shin J, Hu S, Yu X, Huang T, Chen S, Chen H, Chen S, Huang P, Gan RR, Chiu C, Tang P: **Identification of microRNA in the protist *Trichomonas vaginalis*.** *Genomics* 2009, **93**:487-493.
31. Hurt DJ, Wang SS, Lin YH, Hopper AK: **Cloning and characterization of LOS1, a *Saccharomyces cerevisiae* gene that affects tRNA splicing.** *Mol Cell Biol* 1987, **7**:1208-1216.
32. Li J, Chen X: **PAUSED, a putative exportin-t, acts pleiotropically in ranked into 5 categories (1-5; with 1 being the most conserved).NC (Not Classified) development but is dispensable for viability.** *Plant Physiol* 2003, **132**:1913-1924.
33. Steiner-Mosonyi M, Mangroo D: **The nuclear tRNA aminoacylation-dependent pathway may be the principal route used to export tRNA from the nucleus in *Saccharomyces cerevisiae*.** *Biochem J* 2004, **378**:809-816.
34. Grosshans H, Hurt E, Simos G: **An aminoacylation-dependent nuclear tRNA export pathway in yeast.** *Genes Dev* 2000, **14**:830-840.
35. Takano A, Endo T, Yoshihisa T: **tRNA actively shuttles between the nucleus and cytosol in yeast.** *Science* 2005, **309**:140-142.
36. Ohno M, Segref A, Bachi A, Wilim M, Mattaj JW: **PHAX, a mediator of U snRNA nuclear export whose activity is regulated by phosphorylation.** *Cell* 2000, **101**:187-198.
37. Segref A, Mattaj JW, Ohno M: **The evolutionarily conserved region of the U snRNA export mediator PHAX is a novel RNA-binding domain that is essential for U snRNA export.** *RNA* 2001, **7**:351-360.
38. Hurt E, Hannus S, Schmelzl B, Lau D, Tollervey D, Simos G: **A novel in vivo assay reveals inhibition of ribosomal nuclear export in Ran-cycle and nucleoporin mutants.** *J Cell Biol* 1999, **144**:389-401.
39. Moy TI, Silver PA: **Nuclear export of the small ribosomal subunit requires the Ran-GTPase cycle and certain nucleoporins.** *Genes Dev* 1999, **13**:2118-2133.
40. Gadal O, Strauss D, Kessl J, Trumpower B, Tollervey D, Hurt E: **Nuclear export of 60 s ribosomal subunits depends on Xpo1p and requires a nuclear export sequence-containing factor, Nmd3p, that associates with the large subunit protein Rpl10p.** *Mol Cell Biol* 2001, **21**:3405-3415.
41. Ho JH, Kallstrom G, Johnson AW: **Nmd3p is a Crm1p-dependent adapter protein for nuclear export of the large ribosomal subunit.** *J Cell Biol* 2000, **151**:1057-1066.
42. Yao W, Roser D, Köhler A, Bradatsch B, Bassler J, Hurt E: **Nuclear export of ribosomal 60S subunits by the general mRNA export receptor Mex67-Mtr2.** *Mol Cell* 2007, **26**:51-62.
43. Hung N, Lo K, Patel SS, Helmke K, Johnson AW: **Arx1 is a nuclear export receptor for the 60S ribosomal subunit in yeast.** *Mol Biol Cell* 2008, **19**:735-744.
44. Lo K, Johnson AW: **Reengineering ribosome export.** *Mol Biol Cell* 2009, **20**:1545-1554.
45. Cuevas IC, Frasch ACC, D'Orso I: **Insights into a CRM1-mediated RNA-nuclear export pathway in *Trypanosoma cruzi*.** *Mol Biochem Parasitol* 2005, **139**:15-24.
46. Grüter P, Taberner C, von Kobbe C, Schmitt C, Saavedra C, Bachi A, Wilim M, Felber BK, Izaurralde E: **TAP, the human homolog of Mex67p, mediates CTE-dependent RNA export from the nucleus.** *Mol Cell* 1998, **1**:649-659.
47. Santos-Rosa H, Moreno H, Simos G, Segref A, Fahrenkrog B, Panté N, Hurt E: **Nuclear mRNA export requires complex formation between Mex67p and Mtr2p at the nuclear pores.** *Mol Cell Biol* 1998, **18**:6826-6838.
48. Segref A, Sharma K, Doye V, Hellwig A, Huber J, Lührmann R, Hurt E: **Mex67p, a novel factor for nuclear mRNA export, binds to both poly(A)+ RNA and nuclear pores.** *EMBO J* 1997, **16**:3256-3271.
49. Chávez S, Beilharz T, Rondón AG, Erdjument-Bromage H, Tempst P, Svejstrup JQ, Lithgow T, Aguilera A: **A protein complex containing Tho2, Hpr1, Mft1 and a novel protein, Thp2, connects transcription elongation with mitotic recombination in *Saccharomyces cerevisiae*.** *EMBO J* 2000, **19**:5824-5834.
50. Strässer K, Masuda S, Mason P, Pfannstiel J, Oppizzi M, Rodríguez-Navarro S, Rondón AG, Aguilera A, Struhl K, Reed R, Hurt E: **TREX is a conserved complex coupling transcription with messenger RNA export.** *Nature* 2002, **417**:304-308.
51. Aguilera A: **Cotranscriptional mRNP assembly: from the DNA to the nuclear pore.** *Curr Opin Cell Biol* 2005, **17**:242-250.
52. Hautbergue GM, Hung M, Golovanov AP, Lian L, Wilson SA: **Mutually exclusive interactions drive handover of mRNA from export adaptors to TAP.** *Proc Natl Acad Sci USA* 2008, **105**:5154-5159.
53. Rougemaille M, Gudipati RK, Olesen JR, Thomsen R, Seraphin B, Libri D, Jensen TH: **Dissecting mechanisms of nuclear mRNA surveillance in THO/Sub2 complex mutants.** *EMBO J* 2007, **26**:2317-2326.
54. Fischer T, Strässer K, Rácz A, Rodríguez-Navarro S, Oppizzi M, Ihrig P, Lechner J, Hurt E: **The mRNA export machinery requires the novel Sac3p-Thp1p complex to dock at the nucleoplasmic entrance of the nuclear pores.** *EMBO J* 2002, **21**:5843-5852.
55. Fischer T, Rodríguez-Navarro S, Pereira G, Rácz A, Schiebel E, Hurt E: **Yeast centrin Cdc31 is linked to the nuclear mRNA export machinery.** *Nat Cell Biol* 2004, **6**:840-848.
56. Gallardo M, Luna R, Erdjument-Bromage H, Tempst P, Aguilera A: **Nab2p and the Thp1p-Sac3p complex functionally interact at the interface between transcription and mRNA metabolism.** *J Biol Chem* 2003, **278**:24225-24232.
57. Fribourg S, Conti E: **Structural similarity in the absence of sequence homology of the messenger RNA export factors Mtr2 and p15.** *EMBO Rep* 2003, **4**:699-703.
58. Selvapandian A, Kumar P, Morris JC, Salisbury JL, Wang CC, Nakhasi HL: **Centrin1 is required for organelle segregation and cytokinesis in *Trypanosoma brucei*.** *Mol Biol Cell* 2007, **18**:3290-3301.
59. Alcázar-Román AR, Tran EJ, Guo S, Wente SR: **Inositol hexakisphosphate and Gle1 activate the DEAD-box protein Dbp5 for nuclear mRNA export.** *Nat Cell Biol* 2006, **8**:711-716.
60. Blevins MB, Smith AM, Phillips EM, Powers MA: **Complex formation among the RNA export proteins Nup98, Rae1/Gle2, and TAP.** *J Biol Chem* 2003, **278**:20979-20988.
61. Alcázar-Román AR, Bolger TA, Wente SR: **Control of mRNA export and translation termination by inositol hexakisphosphate requires specific interaction with Gle1.** *J Biol Chem* 2010, **285**:16683-16692.
62. Brighthouse A, Dacks JB, Field MC: **Rab protein evolution and the history of the eukaryotic endomembrane system.** *Cell Mol Life Sci* 2010, **67**:3449-3465.
63. DeGrasse JA, DuBois KN, Devos D, Siegel TN, Sali A, Field MC, Rout MP, Chait BT: **Evidence for a shared nuclear pore complex architecture that is conserved from the last common eukaryotic ancestor.** *Mol Cell Proteomics* 2009, **8**:2119-2130.
64. Wickstead B, Gull K, Richards TA: **Patterns of kinesin evolution reveal a complex ancestral eukaryote with a multifunctional cytoskeleton.** *BMC Evol Biol* 2010, **10**:110.
65. Anantharaman V, Koonin EV, Aravind L: **Comparative genomics and evolution of proteins involved in RNA metabolism.** *Nucleic Acids Res* 2002, **30**:1427-1464.

66. Aravind L, Iyer LM, Koonin EV: **Comparative genomics and structural biology of the molecular innovations of eukaryotes.** *Curr Opin Struct Biol* 2006, **16**:409-419.
67. Neumann N, Lundin D, Poole AM: **Comparative genomic evidence for a complete nuclear pore complex in the last eukaryotic common ancestor.** *PLoS ONE* 2010, **5**:e13241.
68. Gomez C, Esther Ramirez M, Calixto-Galvez M, Medel O, Rodríguez MA: **Regulation of gene expression in protozoa parasites.** *J Biomed Biotechnol* 2010, **2010**:726045.
69. Casolari JM, Brown CR, Drubin DA, Rando OJ, Silver PA: **Developmentally induced changes in transcriptional program alter spatial organization across chromosomes.** *Genes Dev* 2005, **19**:1188-1198.
70. Colón-Ramos DA, Salisbury JL, Sanders MA, Shenoy SM, Singer RH, García-Blanco MA: **Asymmetric distribution of nuclear pore complexes and the cytoplasmic localization of beta2-tubulin mRNA in *Chlamydomonas reinhardtii*.** *Dev Cell* 2003, **4**:941-952.
71. Dieppois G, Stutz F: **Connecting the transcription site to the nuclear pore: a multi-tether process that regulates gene expression.** *J Cell Sci* 2010, **123**:1989-1999.
72. Capelson M, Liang Y, Schulte R, Mair W, Wagner U, Hetzer MW: **Chromatin-bound nuclear pore components regulate gene expression in higher eukaryotes.** *Cell* 2010, **140**:372-383.
73. Fitzpatrick DA, Logue ME, Stajich JE, Butler G: **A fungal phylogeny based on 42 complete genomes derived from supertree and combined gene analysis.** *BMC Evol Biol* 2006, **6**:99.

doi:10.1186/1471-2148-11-7

Cite this article as: Serpeloni *et al.*: Comparative genomics of proteins involved in RNA nucleocytoplasmic export. *BMC Evolutionary Biology* 2011 **11**:7.

**Submit your next manuscript to BioMed Central
and take full advantage of:**

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at
www.biomedcentral.com/submit



Categories	Criteria
1	S ≥ 60%, HSP/QC ≥ 80%, QL ~ SL
2	S ≥ 50%, HSP/QC ≥ 60%, QL ~ SL
3	S ≥ 40%, HSP/QC ≥ 45%, QL ~ SL
4	S ≥ 30%, HSP/QC ≥ 30%
5	S ≥ 30% or HSP/QC ≥ 30%
NC	E-value > E-05

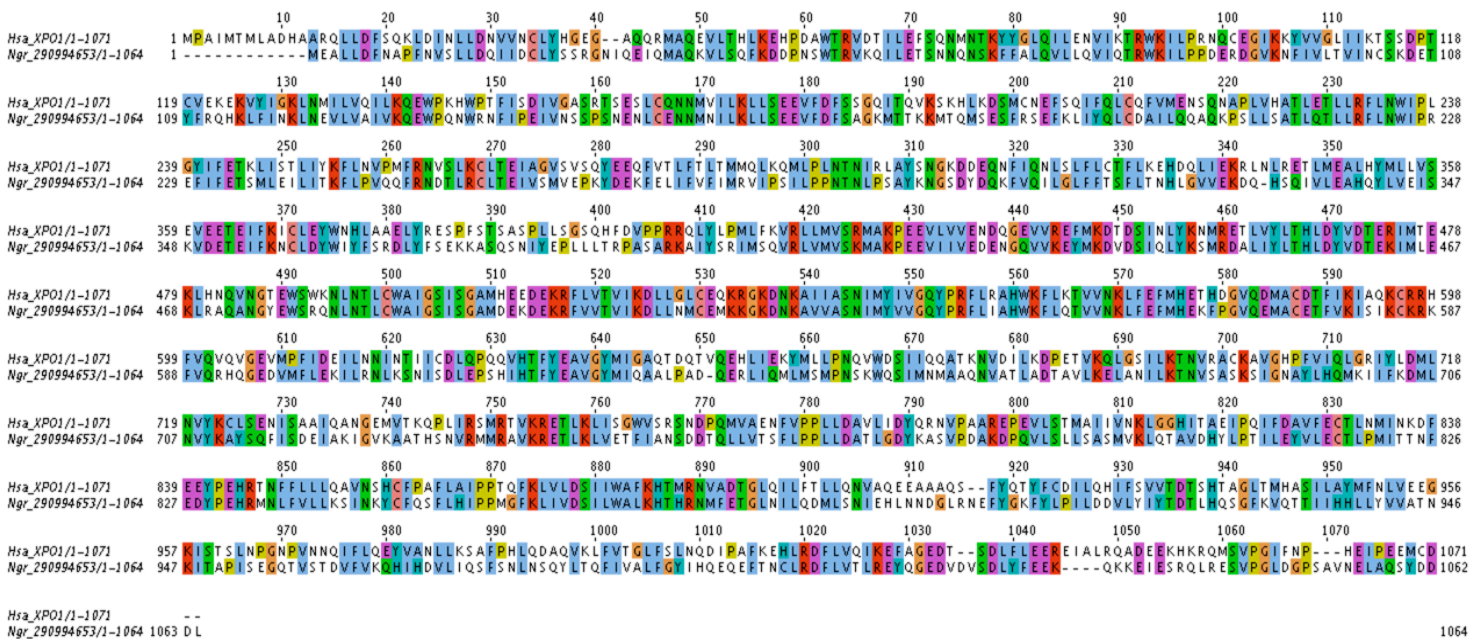
Figure S1. Categories are in 5-color degree-scale, matching the 1-5 similarity categories, with the respective criteria, defined in Methods. S = Similarity, HSP = High-scoring Segment Pair, QC = Query Coverage, QL = Query Length, SL = Subject Length, ~ = similar length.

Figure S2. Examples of alignments for human Exportin-1 (query Hsa_XPO1) in different eukaryotes (A-D). A: Category 1, B: Category 2, C: Category 3, D: Category 4, E: Category 5.

A

Hsa_XPO1 x Ngr (*Naegleria gruberi*)

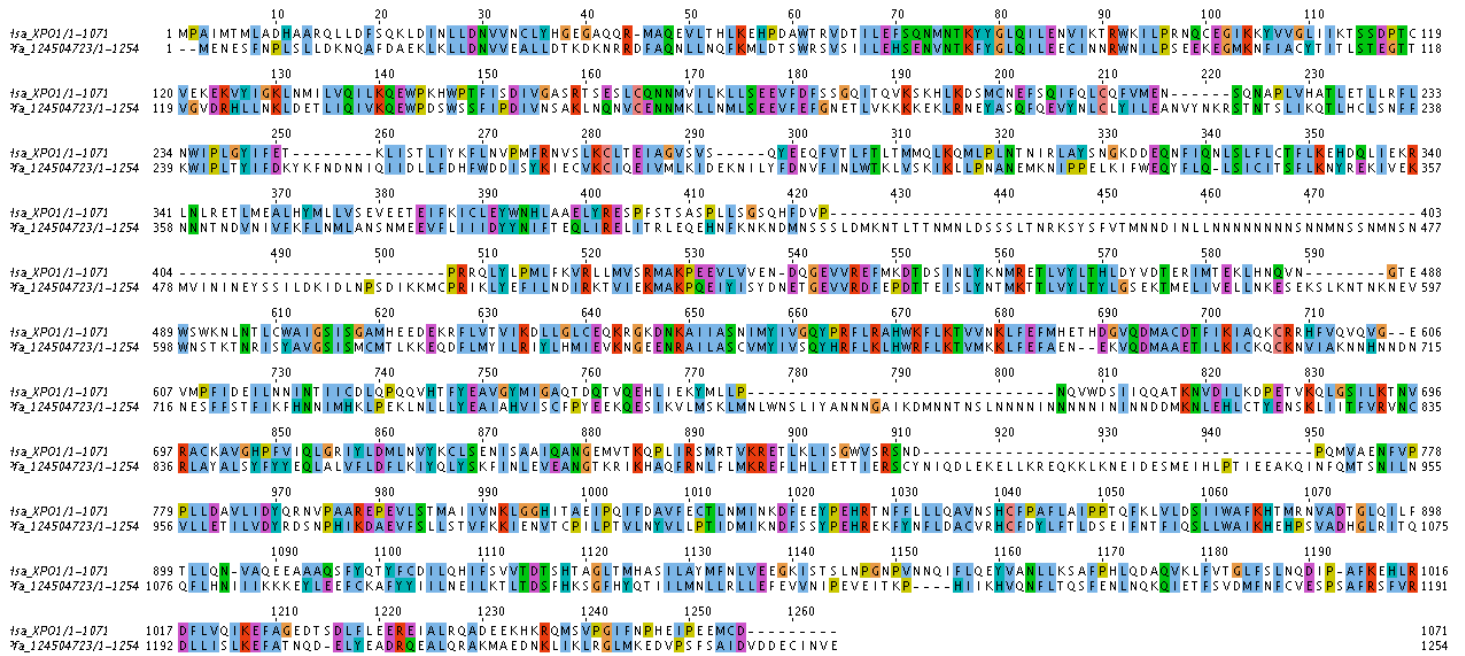
S = 68.7%, QL = 1071, SL = 1064, HSP/QC = 98% E-value 0.0



B

Hsa_XPO1 x Pfa (*Plasmodium falciparum*)

S = 56.3%, QL = 1071, SL = 1254, HSP/QC = 96%, E-value 2.21 E-82



C

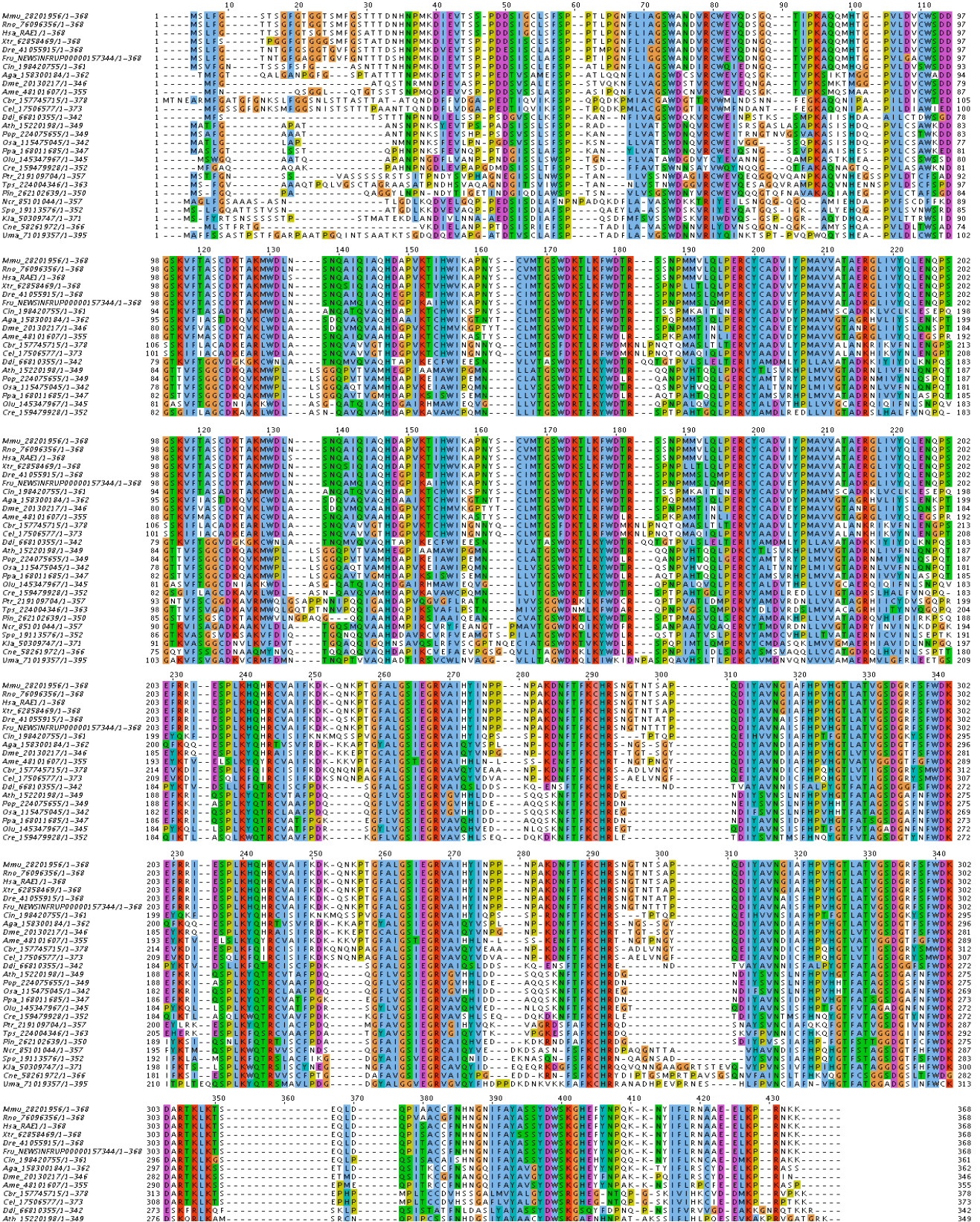
Hsa_XPO1 x Tva (*Trichomonas vaginalis*)

S = 46.9%, QL = 1071, SL = 1003, HSP/QC = 95%, E-value 2.03 E-70

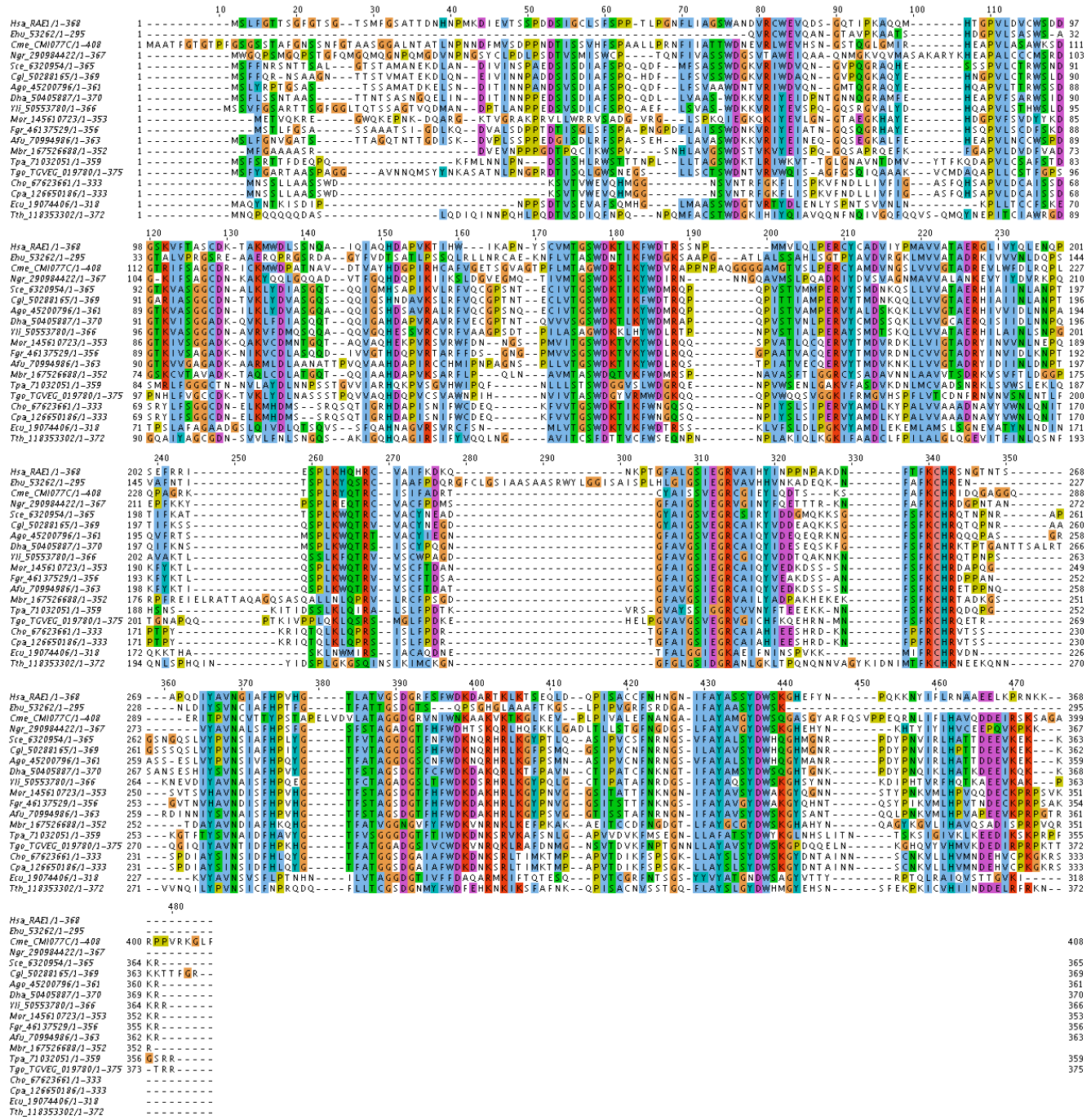


Figure S3. Examples of multiple alignments for human Rael1 (query Hsa_Rael1) in different eukaryotes (A-D). A: Category 1, B: Category 2, C: Category 3, D: Category 4, E: Category 5.

A



B



C

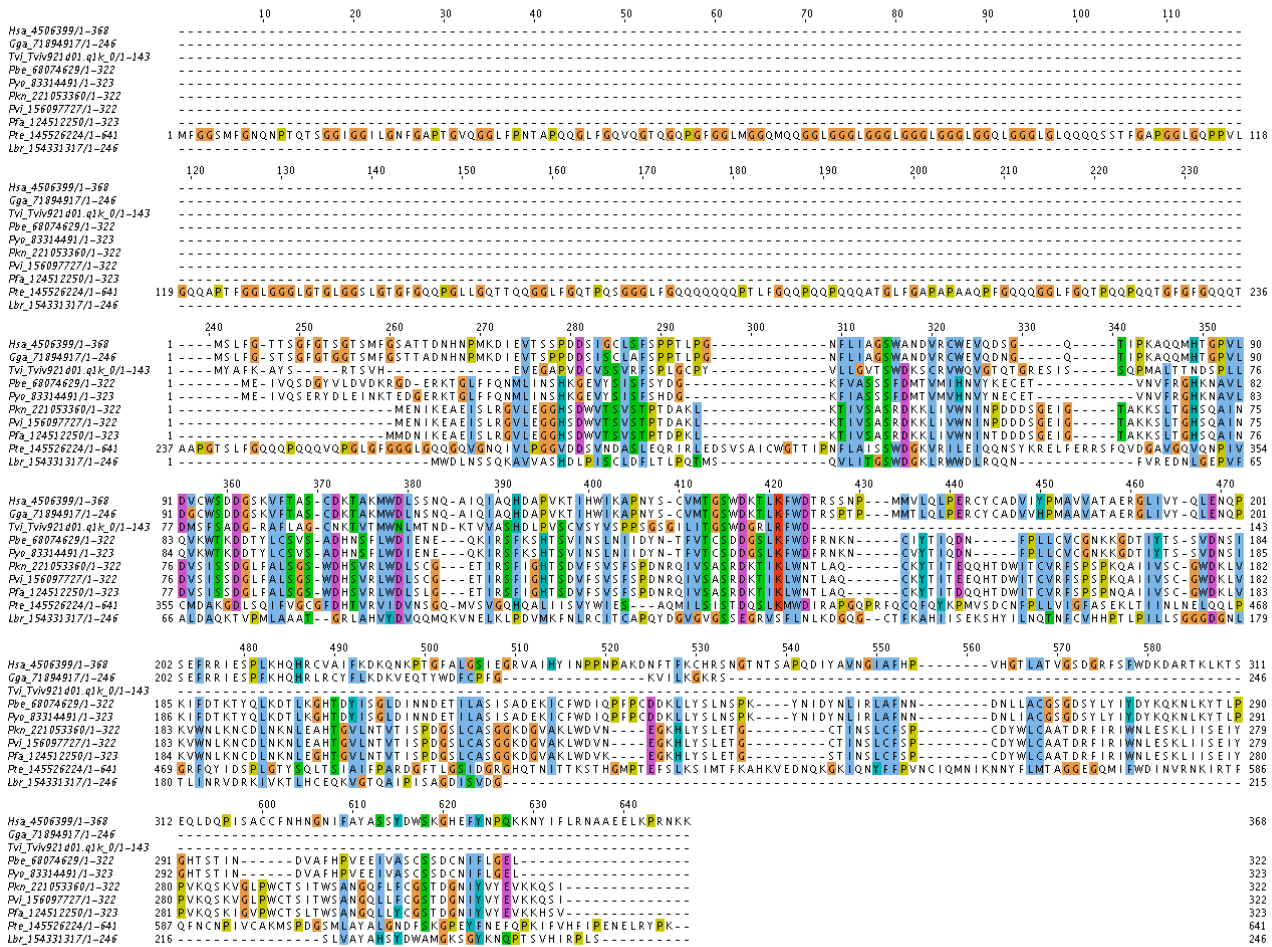
```
10      20      30      40      50      60      70      80      90      100     110
Edi_167389836/1-338 1  -----MSKQQSNSPVFPLGEGKTFQDTISGIEGHSQNGILVGVSWDKCKAFLVNSPTQTQIVGSSVNCNSPLTICKVH---ESQFLVGS-LNGGVYLWNVQ 95
Ehi_67477229/1-341 1  -----MSKQQSNSPVFPLGEGKTFQDTISGIEGYSQNGILVGVSWDKCKAFLVNSPTQTQIVGSSVNCNSPLTICKFKFCBESQFLVGS-LNGGVYLWNVQ 98
Lin_146074949/1-307 1  -----MLMGVAVWDDGCSIQWVARNEAGAVIQPMTTTHDSPLLTMSLSADGRVFFGGCSKAVMWDLNLN 66
Lma_15787825/1-307 1  -----MLMGVAVWDDGCSIQWVARNEAGAVIQPMTTTHDSPLLTMSLSADGRVFFGGCSKAVMWDLNLN 66
Tbr_71744326/1-349 1  -----MFFYSQDQKKAHEVQSPGGDTIISIRFSAGCRLLVGATWDKCRVWQVQDMSRSAAISKPLSLAEAGAPILDMSFSEDGRVFFGGCSKATMWNLT 102
Ter_7165359/1-349 1  -----MNTFNYIRSPGTQHEVQSPGGDQVTVRFSFGCGRYLLGAAWDKCRVWQVQIHQNAINIRSQPMTFAVDAPILDMSFSEDGRVFFGGCSKATMWNLT 102
Hsa_4506399/1-368 1MSLFGTTSGFGTSGTSMFGSATTDMHNMKDI E V T S S A D D S I G C L S F S P T L P G N F L I A G W A N D V R C W E V Q D S G ---QIPKAQQMHTGPVLDVCSDDGSKVFTASQDKLAKMWDLS 116
Tva_123420309/1-303 1  -----MFSNNFIELRIEETDGISGIRFCFKTNS--DYLAFSWDGLHLYDIKTKRLLKTRKFDQPLACEWAEHFCVSGGADGATISANG----- 83
Gla_159119294/1-363 1  -----METRRLLSVTPKGGITRMRFAADKQ---LLAVSWDENVYVYNAEKD---PDRFIFASLHLNKRITESVDVWSDATIVSGGADENLVLSVGG 87
Bna_161899407/1-380 1  -----MGNHKNKNFVHQKINENKFIEMKNSRQSITIEKLSLKKNKCTSNNDIKRKLLENVIRKQYHDKTKKEMAIRFKWKIINTIFALGWVRSIDENYSNSLEFSGSDRILKIWNFF 117

130     140     150     160     170     180     190     200     210     220     230     240
Edi_167389836/1-338 96QLSLOQVGG-FNDCISMKPLLET---NIVISLDGSLGLLDLNL-----SQN-ISKMF-VNNKVICMDAK---PYITFHTP-----NKWEVFNISGKLMTDLIGGQQ 187
Ehi_67477229/1-341 99QNSLQVGG-FNDCISMKPFFET---NIVISLDGSLGLLDLNL-----SQN-ISKMF-VNNKVICMDAK---PYITFHTP-----NKWEVFNISGKLMTDLIGGQQ 190
Lin_146074949/1-307 67SQQKAVVAS-HDLPISCLDFLSLPQTMQMLITGWDQKLRVWDLR-----QQSYVREEN-LGEPVFDALDAQKTPMMAAATG-----RLAHYVDVQMQKVNELKLL--P 162
Lma_15787825/1-307 67SQQKAVVAS-HDLPISCLDFLSLPQTMQMLITGWDQKLRVWDLR-----QQSYVREEN-LGEPVFDALDAQKTPMMAAATG-----RLAHYVDVQMQKVNELKLL--P 162
Tbr_71744326/1-349 103TQKTI VGS-HDLPISCLSVLSL-TGGDMLITGWDQKLRVWDMK-----QRRPVKEEL-LGEPVFDALDAQRSFMAAATG-----RKVHVFNMQFMSKVMLELDP--P 197
Ter_7165359/1-349 103TQKTI VGS-HDLPISCLSVLSN-TTEMLVITGWDQKLRVWDLR-----QRRPVKEEL-LGEPVFDALDAQRSFMAAATG-----RKVHVFNMQFMSKVMLELDP--P 197
Hsa_4506399/1-368 117SQAQIAQ-ADA P K T I H W I K A P ---NYSCEVMGWDKTLKFWDR-----SNPMVQLQ-LPERGCAVDIYRMAVVAER---GLIVRQLENQSEERRIEE--- 209
Tva_123420309/1-303 84---KQIGS-HNDAYCLALVLS---SCIISSFKTKYVWDLR-----SNPISLS-LODKVVSILDESVICGGGD---QNFIFDRRREKGVK--- 170
Gla_159119294/1-363 88KRATSLVGR-HAADVRECVLBSN---RNIVISGWDSHAAWQRTAAEILSSTHEIARIGSPPPLSKIA-AGGKIFCMDVRSRVLGLLGLRTRVQMWDCLPNDLGSBLFKQVLRV--E 201
Bna_161899407/1-380 118KGRLLSLPGHIDQILDLDQDQ---DLLYSALDKTKCWDLK-----QNRVINYYGKSGVYIINLFLKILFSGRRD---CVIKYVWDRSNKEVMNSMRG--- 211

250     260     270     280     290     300     310     320     330     340     350     360
Edi_167389836/1-338 188SITINSMOCLAVFPDRK---GIVYGGADGRANVLHFN-KSNQYTFAFRAQKKK---ISTTQIYYVNVNIFIFNET-----IFLAGSDGVNINWDRKNSLWNL---LTPIQNCR 290
Ehi_67477229/1-341 191SIVNNSMOCIAVFPDRK---GIVYGGADGRANVLHFN-KSNQYTFAFRAQKKK---ISTTQIYYVNVNIFIFNET-----IFLAGSDGVNINWDRKNSLWNL---LTPIQNCR 293
Lin_146074949/1-307 163DVMKFNRCITLCAQDYD---GVVGGSEGRVFSIMKD-APQ---CFKAHITTT---KSHYILSQTDFCVHHTLP-----LLLGGDQNLTVINRADKVKIKLQCEQKVGTAIR 266
Lma_15787825/1-307 163DVMKFNRCITLCAQDYD---GVVGGSEGRVFSIMKD-APQ---CFKAHITTT---KSHYILSQTDFCVHHTLP-----LLLGGDQNLTVINRADKVKIKLQCEQKVGTAIR 266
Tbr_71744326/1-349 198KMKMSELRCAVCSPOHD---GVAVSSEGRVSLPLQA-ESQ---CFKAHVVE---ENNVLVHQTDFVCSIDKTKG-----RMIFGGDGRILAVWQYKKNVCNYC-ENDRKLNRNNS 300
Ter_7165359/1-349 198PLKFNLCVAVCSPOKD---GVVVGSEGRVSLPLQA-EVQ---CFKAHVVE---DNVLMHQTDFVCSISPRV---HMIAGGQGRIGCWQYKKSIVGYGESETRLENRNS 301
Hsa_4506399/1-368 210-PLKHQRCAVIFKDKQKRTFGALGIEGRVAIHYINPNAKDNFTFKCHRSGTNTSAPQDIYAVNGIATFHPVH---LTAIVGSDGRFSFWDKDAITKLLK-----EQLDQR 317
Tva_123420309/1-303 171-PLHNISCVAAKDLF---AIGFEGRVGVDDNN---NTFCKAHVQV---DDSKLLYSINSMCFNPQTR-----DLVTAGSDGKIIIVMDIEMKQRYE-----LGYETS 264
Gla_159119294/1-363 202SPLRQHRVAIVSDIG---NQFCLGLEGRIAMDNVAG---PDRFAFCHRQDD-ANMVFVYPIATIKWCWRN-----ALFVGGSDGLYLWNLOSRRKFRLCESSGIASIDIS 308
Bna_161899407/1-380 212-ENENIMKILIDRKNPK---IIGTIDRTIKEDLRTAKCTGINKENNALRDLKLSNDGFIIGVSDLEKMKNSENNKIVTFLFNLKKSFTCLDLSKGGILIG-TEDGGIFVWHK 325

370     380     390     400     410
Edi_167389836/1-338 291ITADFCENKFLAVSVGVDWHQGVGGLLQGVSVVPIKVVDDPLVRL----- 338
Ehi_67477229/1-341 294ITADFCENKFLAVSVGVDWHQGVGGLLQGVSVVPIKVVDDPLVRL----- 341
Lin_146074949/1-307 267ISAGDISADGSLVAIAHSDWAMGKSGVR-NQPTSVHIRPLL----- 307
Lma_15787825/1-307 267ISAGDISADGSLVAIAHSDWAMGKSGVR-NQPTSVHIRPLS----- 307
Tbr_71744326/1-349 301ISAGDISADCSLLAIAHSDWAMKTRAITNEPHTIIRISANITQLKAR----- 349
Ter_7165359/1-349 302ISAGDISADCSLLAIAHSDWAMKTRAITNEPHTIIRISANITQLKAR----- 349
Hsa_4506399/1-368 318IACCFNHNENIFAAASDWSKHEFNHQKNYIFLNAAEELKRNNK----- 368
Tva_123420309/1-303 265ISIDFSANGNLLAIAISYGENGNISHKDRNLLFIQI----- 303
Gla_159119294/1-363 309QEHDTIALGLSYWENGWDFSRKNATIALHDKIVCSVYEFKDDKAQEQHQ 363
Bna_161899407/1-380 326FKQKLLNKLQVNMKRTILDEKNSFIIDKFDRLGSRFISDTSRRIKLVHQRSDY 380
```

D



4 Capítulo 2.

An Essential Nuclear Protein in Trypanosomes Is a Component of mRNA Transcription/Export Pathway

Autores: Mariana Serpeloni, Carolina Borsoi Moraes, João Renato Carvalho Muniz, Maria Cristina Machado Motta, Augusto Savio Peixoto Ramos, Rafael Luis Kessler, Alexandre Haruo Inoue, Wanderson Duarte daRocha, Sueli Fumie Yamada-Ogatta, Stenio Perdigão Fragoso, Samuel Goldenberg, Lucio H. Freitas-Junior, Andréa Rodrigues Ávila.

Publicado no periódico PLoS One. 6(6):e20730 (2011)
doi: 10.1371/journal.pone.0020730

Tendo em vista a alta conservação de UAP56/Sub2 ao longo da filogenia de eucariotos, inclusive em tripanossomatídeos, foi iniciado o estudo funcional das proteínas ortólogas de *T. cruzi* e *T. brucei*, sendo denominadas de TcSub2 e TbSub2 respectivamente, ou mais genericamente de Tryp-Sub2, por se tratar de proteínas altamente similares (89% de identidade e 96% de similaridade).

Foi possível prever a homologia estrutural de TcSub2 através da comparação com a proteína ortóloga de humanos, UAP56, devido a alta similaridade entre elas (70%), assim como dos domínios típicos de ATPases/helicases e os tamanhos similares, o que é uma evidência forte de possível conservação de função desta proteína. Foi verificado que TcSub2, assim como em outros eucariotos, é uma proteína nuclear, com distribuição preferencial em regiões entre cromatina mais frouxa e cromatina mais densa, ausente em nucléolo, e que existe uma associação de TcSub2 a alguns sítios de transcrição de mRNAs no núcleo de *T. cruzi*. Esta localização e associação com transcrição por RNA Polimerase II também ocorre com as proteínas homólogas de organismos-modelo como *C. tentans*, *C. elegans*, *D. melanogaster*, *S. cerevisiae* e *H. sapiens*. Neste caso, elas participam de modo co-transcricional em *loci* ativos e no transporte de mRNAs até o poro nuclear, sendo liberadas antes mesmo do complexo mRNA alcançar o citoplasma, permanecendo no núcleo (AGUILERA, 2005; EBERL *et al.*, 1997; GATFIELD *et al.*, 2001; KIESLER, MIRALLES e VISA, 2002; MACMORRIS, BROCKER e BLUMENTHAL, 2003).

Não foi possível obter o nocaute duplo do gene *sub2* de *T. cruzi* para as análises funcionais, provavelmente por se tratar de uma proteína essencial neste parasita. Levando em consideração a alta similaridade com a proteína de *T. brucei*,

TbSub2, as análises funcionais foram feitas então em *T. brucei* através de RNA interferência (RNAi). Assim foi possível observar que a distribuição de mRNAs após a indução do RNA de interferência ao longo do tempo. Após o silenciamento da expressão do gene *sub2* de *T. brucei* foi observado que TbSub2 é também essencial no parasita, resultando em acúmulo de mRNAs no núcleo, diminuição drástica da tradução de modo geral e consequente morte celular, semelhante ao que ocorre em outros eucariotos.(GATFIELD *et al.*, 2001; HAUTBERGUE *et al.*, 2009; KAPADIA *et al.*, 2006; MACMORRIS, BROCKER e BLUMENTHAL, 2003; STRASSER e HURT, 2001). Outro resultado interessante observado é que a diminuição desta proteína prejudica a divisão celular resultando em parasitas majoritariamente com dois núcleos, nos quais também foram encontrados mRNAs acumulados. Essa relação de Sub2 com a citocinese já foi descrita em células humanas (YAMAZAKI *et al.*, 2010) e os indícios obtidos corroboram com a conservação da função na exportação de mRNAs em tripanossomatídeos. Para avaliar a função nuclear de TcSub2 no metabolismo de mRNAs e entender o contexto celular em que esta proteína está inserida, o próximo passo foi a tentativa de identificação de proteínas associadas e os resultados serão discutidos no Capítulo 4.

An Essential Nuclear Protein in Trypanosomes Is a Component of mRNA Transcription/Export Pathway

Mariana Serpeloni^{1,2}, Carolina Borsoi Moraes³, João Renato Carvalho Muniz⁴, Maria Cristina Machado Motta⁵, Augusto Savio Peixoto Ramos², Rafael Luis Kessler^{1,2}, Alexandre Haruo Inoue^{1,2}, Wanderson Duarte da Rocha⁶, Sueli Fumie Yamada-Ogatta⁷, Stenio Perdigão Fragoso², Samuel Goldenberg², Lucio H. Freitas-Junior³, Andréa Rodrigues Ávila^{2*}

1 Departamento de Biologia Celular e Molecular, Universidade Federal do Paraná (UFPR), Curitiba, Brazil, **2** Laboratório de Regulação da Expressão gênica, Instituto Carlos Chagas (ICC), Curitiba, Brazil, **3** Center for Neglected Diseases Drug Discovery (CND3), Institut Pasteur Korea (IPK), Gyeonggi-do, South Korea, **4** University of Oxford, Oxford, United Kingdom, **5** Departamento de Biologia Celular e Parasitologia, Instituto de Biofísica Carlos Chagas Filho, Universidade Federal do Rio de Janeiro (UFRJ), Rio de Janeiro, Brazil, **6** Departamento de Bioquímica, Universidade Federal do Paraná (UFPR), Curitiba, Brazil, **7** Departamento de Microbiologia, Centro de Ciências Biológicas, Universidade Estadual de Londrina (UEL), Londrina, Brazil

Abstract

In eukaryotic cells, different RNA species are exported from the nucleus via specialized pathways. The mRNA export machinery is highly integrated with mRNA processing, and includes a different set of nuclear transport adaptors as well as other mRNA binding proteins, RNA helicases, and NPC-associated proteins. The protozoan parasite *Trypanosoma cruzi* is the causative agent of Chagas disease, a widespread and neglected human disease which is endemic to Latin America. Gene expression in *Trypanosoma* has unique characteristics, such as constitutive polycistronic transcription of protein-encoding genes and mRNA processing by trans-splicing. In general, post-transcriptional events are the major points for regulation of gene expression in these parasites. However, the export pathway of mRNA from the nucleus is poorly understood. The present study investigated the function of TcSub2, which is a highly conserved protein ortholog to Sub2/ UAP56, a component of the Transcription/Export (TREX) multiprotein complex connecting transcription with mRNA export in yeast/human. Similar to its orthologs, TcSub2 is a nuclear protein, localized in dispersed foci all over the nuclei—except the fibrillar center of nucleolus—and at the interface between dense and non-dense chromatin areas, proposing the association of TcSub2 with transcription/processing sites. These findings were analyzed further by BrUTP incorporation assays and confirmed that TcSub2 is physically associated with active RNA polymerase II (RNA pol II), but not RNA polymerase I (RNA pol I) or Spliced Leader (SL) transcription, demonstrating participation particularly in nuclear mRNA metabolism in *T. cruzi*. The double knockout of the TcSub2 gene is lethal in *T. cruzi*, suggesting it has an essential function. Alternatively, RNA interference assays were performed in *Trypanosoma brucei*. It allowed demonstrating that besides being an essential protein, its knockdown causes mRNA accumulation in the nucleus and decrease of translation levels, reinforcing that *Trypanosoma*-Sub2 (Tryp-Sub2) is a component of mRNA transcription/export pathway in trypanosomes.

Citation: Serpeloni M, Moraes CB, Muniz JRC, Motta MCM, Ramos ASP, et al. (2011) An Essential Nuclear Protein in Trypanosomes Is a Component of mRNA Transcription/Export Pathway. PLoS ONE 6(6): e20730. doi:10.1371/journal.pone.0020730

Editor: Fabio T. M. Costa, State University of Campinas, Brazil

Received: November 26, 2010; **Accepted:** May 11, 2011; **Published:** June 8, 2011

Copyright: © 2011 Serpeloni et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This work was funded by grants from the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) and Fundação Araucária to ARA, grants to SG from Fiocruz and CNPq. MS was funded by a studentship from CNPq and AHI by a studentship from CAPES and CBM and LHFJ were funded by the Ministry of Education, Science and Technology (MEST), the Gyeonggi government, and the Korean Institute of Science and Technology Information from South Korea. SG, SPF and MCM are research fellows from CNPq. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: aravila@fiocruz.br

Introduction

In eukaryotic cells, transcription and translation are physically and temporally separated by the nuclear membrane. Control of mRNA transport to regions of translation defines when and where proteins are expressed. This transport is initiated in the nucleus, where association with protein complexes determines the fate of mRNAs in the cytoplasm [1]. Different events in the nuclear metabolism of mRNA have crucial roles in controlling gene expression. The mRNA has to be correctly processed before being shuttled from the nucleus to the cytoplasm *via* a nuclear pore complex (NPC). The general model of RNA export involves exportins as transport receptors that carry RNA through the NPC

in a RanGTP-dependent manner; specific exportins are involved with the different RNA types [2]. In contrast, nucleocytoplasmic export of most mRNAs does not follow the RanGTP-exportin pathway. In yeast and humans, mRNAs associate with protein factors as messenger ribonucleoprotein complexes (mRNPs) which are then exported through the NPC by an essential general receptor-shuttling heterodimer: Mex67/Mtr2 in yeast and TAP/p15 in humans [3,4]. The dimeric export receptor operates in association with TREX (Transcription/Export), a multiprotein complex that connects transcription with mRNA export. Yeast TREX consists of the RNA helicase **Sub2** (UAP56 in human), the RNA-binding adaptor protein Yra1 (ALY in human), and the THO complex [5–7].

Excluding model eukaryotic organisms, the export machinery of other eukaryotes has yet to be determined. Using comparative genomics, we recently showed that the mRNA export pathway is the least conserved among early divergent eukaryotes, especially in excavates, a major kingdom of unicellular eukaryotes also known as Excavata [8]. In this lineage, we have suggested that mRNA export is quite different for several members. The phylogenetic category Excavata contains a variety of free-living and symbiotic forms, and also includes some major parasites affecting humans [9,10].

The protozoan parasite *Trypanosoma cruzi* is a member of Excavata kingdom and the causative agent of the Chagas disease, a widespread and neglected disease in humans which is endemic to Latin America. Over 16 million people worldwide have been estimated to be infected with this parasite, and more than 100 million live in endemic areas [11–13]. *T. cruzi* gene expression displays some unique characteristics such as constitutive polycistronic transcription of protein-coding genes and mRNA processing by trans-splicing. Because genes in the same polycistronic unit display different levels of processed mRNA, the consensus notion is that post-transcriptional events are the major points of gene expression regulation and play important roles in the adaptation of the parasite throughout its life cycle [14,15]. However, the machinery of the mRNA export pathway is not functionally studied in *T. cruzi*. We have previously shown that Sub2/UAP56, a component of the TREX complex, seems to be the most conserved protein among the factors involved in the mRNA export pathway [8]. In the present study we demonstrate, employing comparative modeling, that *T. cruzi* Sub2 (TcSub2) has a similar 3D structure when compared with the crystallized human ortholog. TcSub2 is a nuclear protein localized in foci dispersed all over the nucleus apart from the nucleolus. Furthermore, electron scanning microscopy revealed that TcSub2 is localized at the interface between dense and non-dense chromatin areas, a pattern of distribution typically associated with mRNA transcription/processing sites. BrUTP incorporation assays showed that TcSub2 colocalizes with RNA pol II transcription sites. These findings strongly suggest that TcSub2 is associated with transcription sites and may participate in nuclear mRNA metabolism. The double knockout of the TcSub2 gene is lethal in *T. cruzi*, indicating that, similarly to other eukaryotes, TcSub2 is an essential protein for trypanosomes. Alternatively, RNA interference (RNAi) assays in *T. brucei* confirmed it has an essential function. Besides, its knockdown causes mRNA accumulation in the nucleus and decreasing of translation levels, confirming that this protein is a component of mRNA transcription/export pathway in trypanosomes.

Materials and Methods

Trypanosoma cultures

Epimastigotes of *T. cruzi* Dm28c [16] were cultured at 28°C in LIT medium supplemented with 10% bovine fetal serum [17]. The culture was seeded at 1×10^6 cells ml^{-1} and the parasites were harvested when cultures reached a cell density of 1×10^7 cells ml^{-1} (during log phase of growth). Different forms of parasite during metacyclogenesis were obtained *in vitro* as previously described [18,19]. Amastigotes were obtained in LLC-MK2 (ATCC® – number: CCL-7TM) [20] infected cells as previously described [21]. The culture of *T. brucei* procyclic forms of Lister 427 29-13 [22] were performed as previously described [23].

Cloning and expression of TcSub2 recombinant protein for polyclonal antiserum production

The TcSub2 coding region (GenBank identification number 3538341) was amplified by PCR using the follow oligonucleotides:

5'GGGGACAAGTTTGTACAAAAAAGCAGGCTTCATGAG-CAGTGGACTTGCCGAC3' (forward) and 5'GGGGACCAC-TTTGTACAAGAAAGCTGGGTCTTACTGATTTCATG-TACTGGCTCTG3' (reverse). Genomic DNA of Dm28c was used as a template. Both oligonucleotides contain recombination sites for cloning the TcSub2 ORF into the entrance pDONR vector (Invitrogen – Gateway® technology). This ORF was recombined into the expression vector pDESTTM 17, in accordance with the manufacturer's protocol for production of His₆-tagged recombinant protein in the *Escherichia coli* BL21 (DE3) strain. The recombinant protein expression was induced with 1 mM IPTG for 3 h at 37°C; the protein was separated by SDS-PAGE, excised from the gel and eluted for inoculation in rabbit for polyclonal antiserum production.

Immunoblotting assay

Cell lysates from 5×10^6 parasites were loaded into SDS-PAGE and the proteins were transferred onto nitrocellulose membrane (Hybond C, Amersham Biosciences, England) using standard protocols. The membrane was incubated with the anti-TcSub2 polyclonal antiserum diluted 1:1000 in blocking solution (5% of nonfat powdered milk, 0.1% Tween-20 in phosphate buffered saline, pH 7.4) for 1 h. After washing in PBS, the membrane was incubated with polyclonal anti-GAPDH antiserum diluted 1:500 in blocking solution. This was then incubated for 45 min at room temperature (R.T.) with anti-rabbit horseradish peroxidase-conjugated IgG (Amersham Biosciences, England) diluted in PBS/Tween 0.5%. Detection of protein was performed according to the manufacturer's instructions.

Comparative analysis by protein sequence alignments and molecular modeling

BLAST searches were carried out for TcSub2 *T. cruzi* protein—homologues at the GenBank databases (<http://www.ncbi.nlm.nih.gov>) using the sequences of the human and *Saccharomyces cerevisiae* proteins as queries. Sequences were aligned using Clustal W (<http://www.cmbi.kun.nl/bioinf/tools/clustalw.shtml>) with homolog proteins of different organisms; including putative proteins of trypanosomatids and occasionally manual refinement of the alignments was performed. For the molecular modeling of TcSub2, the GenTHREADER software [24] was used to perform structural alignments based on the secondary structure of available proteins with known structure. The best match was the human HsaUAP56 (1XTK) protein [25]. The atomic coordinates from this structure were submitted with the alignment results to the program MODELLER [26] to produce the atomic-resolution models, which were then validated using the software PRO-CHECK [27], Verify 3D [28] and WHATIF [29].

Ultrastructural immunocytochemistry

Parasites were fixed in 0.3% glutaraldehyde, 4% formaldehyde, and 1% picric acid diluted in 0.1 M cacodylate buffer (pH 7.2) and then dehydrated at -20°C in a graded series of ethanol solutions. The material was progressively infiltrated with Unicryl at lower temperatures and resin polymerization was carried out in BEEM capsules at 20°C for 5 days under UV light. Ultrathin sections were obtained with a Leica ultramicrotome (Reichert, Ultracuts) and grids containing the sections were incubated with 50 mM NH₄Cl for 30 min. They were then incubated with blocking solution (3% BSA, 0.5% teleostean gelatin diluted in PBS, pH 8.0) for 30 min. This was followed by a 1 h incubation with anti-TcSub2 serum diluted 1:200 in blocking solution. The grids were then incubated for 30 min with gold-labeled goat anti-

rabbit IgG (Sigma) diluted 1:200 in blocking solution. Grids were washed and stained with uranyl acetate and lead citrate for further observation in a Zeiss 900 transmission electron microscope. For control assays, incubation with the primary antiserum was omitted.

In situ labeling of nascent RNAs

The assay was performed as previously described by Dossin and Schenkman (2005) [30]. The parasites were washed three times and resuspended to a cell density of 1×10^7 ml⁻¹ in transcription buffer containing 80 U µg ml⁻¹ of RNasin (Promega). These cells were then permeabilized with 2.5 mg ml⁻¹ of lysolecithin (Sigma) for 1 min on ice, centrifuged, washed in transcription buffer, and suspended to the same cell density in transcription buffer containing 2 mM ATP, 1 mM CTP, 1 mM GTP, 0.5 mM bromo-UTP (BrUTP) (Sigma), 200 mg ml⁻¹ of creatine kinase (Roche), and 50 mM creatine phosphate (Roche). Cells were then incubated for 15 min at 28°C. As a control, the parasites were incubated in transcription buffer for 15 min with different concentrations of α-amanitin (Sigma) prior to the addition of nucleotides. Based on Campbell *et al.* (2003) [31], incubation of the parasites with 75 µg ml⁻¹ of α-amanitin inhibits only RNA pol II while 200 µg ml⁻¹ inhibits both RNA pol II and III.

Fluorescent microscopy analysis of TcSub2

For N-terminal fusion with GFP the TcSub2 ORF was cloned into the pTcGFPN vector [32] and the construct was transfected in *T. cruzi* epimastigotes as previously described [33]. For analysis of the cells expressing TcSub2-GFP, aliquots of cells were harvested, washed in PBS, resuspended in PBS to a density of 10^7 cells ml⁻¹, fixed with 4% paraformaldehyde in PBS for 20 min. The fixed cells were then placed onto poly-l-lysine-coated slides for 20 min at room temperature and were analyzed by fluorescence microscopy. For the indirect immunofluorescence (IF) assays, the cells were fixed as described above, permeabilized with 0.1% Triton X-100 in PBS for 5 min and blocked in 4% BSA in PBS for 60 min. For TcSub2 localization analysis, the cells were incubated for 45 min with anti-TcSub2 serum diluted 1:1000. This was followed by a 45 min incubation with Alexa Fluor 488-conjugated donkey anti-rabbit IgG diluted 1:1600 (Invitrogen) and 5 µg ml⁻¹ DAPI. For transcription colocalization analysis, cells were incubated for 45 min with anti-TcSub2 serum diluted 1:1000 and mouse anti-BrdU (Invitrogen) diluted 1:300. This was followed by incubation for 45 min with Alexa Fluor 594-conjugated donkey anti-rabbit IgG diluted 1:1600 (Invitrogen), Alexa Fluor 488-conjugated goat anti-mouse IgG diluted 1:400

(Invitrogen), and 5 µg ml⁻¹ DAPI. All of the slides were mounted in VectaShield (Vector Laboratories). 2D images were collected with a Nikon E90i epifluorescence microscope using a 100×/1.4 PlanApoVC objective and a Nikon DS-QiMc camera. 3D images (z series) were collected with confocal Leica TCS SP2 AOBs using a 63x/HCX 1.4 PL Apo lbdBL oil immersion objective.

TcSub2 knockout assay

Regions located upstream (UPS) and downstream (DWN) to the TcSub2 coding sequence were cloned into the pKS/NEO and pKS/HYG vectors which respectively contain the neomycin (NEO) and hygromycin (HYG) resistance genes [34]. The UPS and DWN were cloned in flanking regions of NEO or HYG genes. The oligonucleotides used for amplification of UPS and DWN are listed in the supplementary material (Table 1).

The UPS/NEO/DWN construction (approximately 2100 bp) was PCR-amplified and purified by electroelution. A total of 10 µg of DNA was used to transfect *T. cruzi*, based on the method described by [33]. After 24 h, neomycin (G418 - Sigma, St. Louis, MO, USA) was added at a concentration of 300 µg ml⁻¹ to select resistant parasites. The selected parasites were cloned by serial dilution in 24-well plates. Single allele knockout of TcSub2 was confirmed by Pulse-Field electrophoresis followed by Southern blot, where the TcSub2 and NEO genes were labeled with α-[³²P]-dCTP (10 µCi/µl; 3000 Ci/mmol) and used as probes. Knockout was also confirmed by Western blot assays where protein extracts from knockout and wild-type parasites were used for detection of TcSub2, using GAPDH level as loading control. Comparative analysis for both proteins was performed through signal quantification using Scion Image program [35]. A single allele knockout parasites were then transfected with the UPS/HYG/DWN construction (approximately 2300 bp), which was PCR-amplified using the oligonucleotides UPSF e DWNr. The transfection was performed as described for the UPS/NEO/DWN construction, and 300 µg ml⁻¹ G418 and 300 µg ml⁻¹ hygromycin B (Sigma) were added to the LIT medium 24 h after transfection.

RNAi assay in *T. brucei*

According to previous data of comparative genomics [8], *Trypanosoma brucei* has an ortholog protein (TbSub2) of TcSub2 with the accession number Tb10.70.7730 in GenBank database. For TbSub2 RNAi approach, the target sequence and primers were chosen using RNAi software [36]. The target sequence was PCR-amplified and cloned into the p2T7-177 vector [37]. The vector was linearized by *NotI* and transfected into procyclic forms

Table 1. Oligonucleotides used for amplification of UPS and DWN regions of the TcSub2 gene.

Oligonucleotides	Sequence (5'- 3')	Enzyme	bp
UPS-NEO F	GGGG TCGAC GGTGTATTTCGCTCACGATGTGC	<i>SaI</i>	700
UPS-NEO R	GGGG TCGAC TTGCTTATGTCTATCTGTTTCTTA	<i>SaI</i>	
UPS HIGRO F	GGG CTAG AGGTGTATTTCGCTCACGATGTGC	<i>XbaI</i>	
UPS HIGRO R	GGG CTAG ATTGCTTATGTCTATCTGTTTCTTA	<i>XbaI</i>	
DOWN-NEO F	GGGG AATTC GATTTAAGGCCTCTGAAAATAAGAATGAGG	<i>EcoRI</i>	450
DOWN-NEO R	GGGG AATTC AACACAAAAATAAGAAGAAGTATTGGAATGGCAATTGC	<i>EcoRI</i>	
DOWN-HIGRO F	GGGG GATCC GATTTAAGGCCTCTGAAAATAAGAATGAGG	<i>BamHI</i>	
DOWN-HIGRO R	GGG AAGCTT AACACAAAAATAAGAAGAAGTATTGGAATGGCAATTGC	<i>HindIII</i>	

NEO: neomycin; HIGRO: hygromycin. bp base pair. The enzyme cleavage sites are in bold.
doi:10.1371/journal.pone.0020730.t001

of *T. brucei* 29-13 strain [22]. The stable DNA integration was selected using phleomycin ($5 \mu\text{g ml}^{-1}$) and RNAi induction was performed after addition of $1 \mu\text{g ml}^{-1}$ of tetracycline. Cultures grown to mid-log phase (10^6 – 10^7 cells ml^{-1}) were harvested for production of the total protein extracts. The knockdown of TbSub2 expression was confirmed by immunoblotting assay, using GAPDH level as loading control. Comparative analysis for both proteins was performed through signal quantification using Scion Image program [35].

Fluorescent *in situ* hybridization (FISH)

For TcSub2- telomere colocalization analysis, log phase epimastigotes were subjected to indirect immunofluorescence followed by telomere detection using Telomere PNA FISH Kit/FITC (DakoCytomation) in accordance with the manufacturer's instructions. To observe colocalization of TcSub2 with SL transcription sites in *T. cruzi*, log phase epimastigotes were subjected to IF followed by Fluorescent *in situ* hybridization (FISH), using the SL-RNA as a probe. The SL-RNA probe was prepared by PCR using as template the *T. cruzi* (Y strain) SL RNA gene cloned in vector pGEM-T Easy. The SL probe was labeled with digoxigenin by PCR using High prime DNA labeling kit (Roche Diagnostics), purified and used in hybridization assays as described by Dossin et al., 2005 [30]. For detection of polyA-containing mRNA, FISH was performed based on previous works with modifications [38,39]. Aliquots of cells (*T. cruzi* epimastigotes and *T. brucei* procyclic forms) were harvested, washed in PBS, resuspended in PBS to a density of 10^7 cells ml^{-1} and fixed with 4% paraformaldehyde in PBS for 20 min. Fixed cells were washed three times in PBS, resuspended in $200 \mu\text{l}$ PBS and then placed onto poly-l-lysine-coated slides for 20 min at R.T. The cells were then permeabilized with 0.2% Triton X-100 in PBS for 30 min and washed three times with PBS. Pre-hybridization was performed in $10\times$ Denhardt's solution, $4\times$ SSC, 1 mM EDTA, 35% deionized formamide, 0.5 mg ml^{-1} tRNA, 2 mU ml^{-1} RNase OUT for 30 min at room temperature. For control, cells were previously treated with $100 \mu\text{g ml}^{-1}$ of boiled RNase A in PBS for 60 min at 37°C . Six $\text{ng}/\mu\text{l}$ of digoxigenin (DIG) conjugated-5' oligo d(T)₃₀ was added to pre-hybridization buffer. The cells were heated to 65°C for 3 min before hybridization at R.T. overnight. The cells were washed twice with $2\times$ SSC for 15 min, twice with $1\times$ SSC for 15 min, twice with PBS for 5 min and then incubated for 45 min at R.T. with anti-digoxigenin monoclonal antibody (Sigma) diluted in PBS/ BSA (0.1 mg ml^{-1}). This was followed by incubation with Alexa Fluor 488-conjugated goat anti-mouse IgG diluted 1:400 (Invitrogen) in PBS/BSA 0.1 mg ml^{-1} and 100 ng ml^{-1} DAPI for 45 min at R.T. The slides were mounted in N-propyl gallate and examined with a Nikon E600 microscope. Images were acquired with the Image Pro program (Media Cybernetics, Bethesda, MD, USA).

Metabolic Labeling

The metabolic labeling of procyclic forms of *T. brucei* was performed according to [40] with modifications. A total 5×10^6 cells from 48 h-induced and non-induced cultures was harvested from medium, washed in methionine-free RPMI medium (Sigma) supplemented with fetal bovine serum 0.5% and labeled by incubation, in the same medium containing $50 \mu\text{Ci ml}^{-1}$ [³⁵S]-methionine (Amersham Bioscience), for one hour at 28°C . For control, parasites were treated with $100 \mu\text{g ml}^{-1}$ of cycloheximide for 5 min before addition of [³⁵S]-methionine. Aliquots of $10 \mu\text{l}$ were spotted in 3MM filters and boiled in 10% of trichloroacetic acid (TCA). After consecutive washes in 5% TCA, acetone an

ethanol, the filters were dried at R.T. and the radioactivity was counted by liquid scintillation.

Flow cytometry assays

All experiments were performed in a FACSCalibur flow cytometer (Becton-Dickinson, San-Jose, USA). A total of 20,000 events were acquired and data analysis was performed using FlowJo software (Treestar software). All experiments were done at least in duplicate.

Cell viability assays were performed using two parameters: propidium iodide staining and mitochondrial transmembrane potential [adapted from 41]. For the first assay, 2×10^6 cells were harvested, washed with PBS and stained with $5 \text{ ng } \mu\text{l}^{-1}$ propidium iodide in PBS at 28°C for 15 min. The cells were analyzed by flow cytometry using FL2-H detector. For the second assay, 2×10^6 cells were harvested, washed in PBS and incubated with $10 \mu\text{g ml}^{-1}$ of rhodamine 123 in PBS for 15 min at 28°C . The cells were then resuspended in PBS for flow cytometry analyses using FL1-H detector. Data analysis was executed considering only viable cells gated by forward and side scatter. Relative mitochondrial membrane potential (in comparison to control cells) was determined considering the ratio (treated/control cells) of geometric mean of rhodamine 123 fluorescence.

The cell cycle analysis was performed by DNA staining with propidium iodide. 2×10^5 cells were harvested, washed in PBS and resuspended in $500 \mu\text{l}$ of DNA staining solution containing 0.425 mM Tris-HCl, $9.375 \mu\text{M}$ Propidium Iodide, 0.0125% NP40 detergent, 1.25 mM NaCl and 8.75 U/L of RNase A. The cells were kept on ice until immediate flow cytometry quantification. Data analysis was done in FL2-Widht \times FL2-Area gated cells to exclude debris and doublets. The Dean-Jett-Fox algorithm of FlowJo software was used to estimate percentage of cells in G0/G1, S and G2/M phases of cell cycle.

Apoptosis was detected by phosphatidylserine exposure using BD PharMingen™ Annexin V-FITC Apoptosis Detection Kit according to manufacturer's instructions. Briefly, 2×10^6 cells were harvested, washed and staining with annexin-V-FITC (AV) for 30 min at 28°C followed by staining with propidium iodide (PI) for 15 min at 28°C . Unstained and single-stained cells were used to set up compensation of signals from the detectors. To observe morphological alterations, cell samples from each assay were treated using the staining method "Panótico Rápido" (Laborclin, Pinhais, PR, Brazil) as described previously [34]. The stained parasites were examined in bright field using Leica SP5-AOBS confocal microscopy.

Results

Cloning of TcSub2 and comparative analysis of ortholog proteins by molecular modeling

Previous studies have shown that only a small group of proteins involved in mRNA export is conserved throughout eukaryotic phylogeny. Among the components of the mRNA export pathway, Sub2 is the most conserved protein [8]. Its ortholog in *T. cruzi* (TcSub2) is annotated as a putative RNA helicase (Tc00.1047053508319.40) and is highly similar to SceSub2 (68%) and HsaUAP56 (69%). Multiple alignments of TcSub2 with similar proteins from different eukaryotes (Figure 1A) show a highly conserved amino acid sequence through eukaryotic lineages. Sequence analysis revealed that TcSub2 exhibits nine canonical helicase motifs, typical of DEADX box-RNA helicase proteins, which are associated with nearly all processes involving RNA, from transcription to RNA degradation [42]. We obtained the molecular model of TcSub2 based on the crystal structure of

human HsaUAP56 (Figure 1B). Comparative analysis between the molecular model of TcSub2 and HsaUAP56 shows a very conservative folding of RNA helicase motifs and two major helicase domains. Despite some amino acid differences between the sequences (Figure 1C) it does not seem to affect quality and overall aspect of the modeled structure of the TcSub2 RNA helicase (Figure 1D and E). Modifications in the N-terminal (regions Ia and II) are non-conservative (figure 1D) whereas all modifications of C-terminal regions (regions IV, V and VI) are conservative (figure 1E). In the N-terminal, the V116C substitution is located in a hydrophobic core and both A118T and Y123F are exposed to the solvent. These substitutions are therefore unlikely to cause any conformational changes to the structure as a whole. The substitution F194C, located in the region II, is part of a hydrophobic core and a few conservative substitutions surrounding the volumetrically bigger side chain of the phenylalanine is enough to accommodate this residue at that position. At C-terminal domain, the region V presents only one conservative substitution the V347M. This mutation is located next to one of the IPA binding sites of human structure but it is not directly related on its coordination. The substitution D350E is located between regions V and VI and it is structurally exposed, minimizing its impact on protein folding. The conservative variable residues for regions IV and VI (mutations I287V and G372A, respectively) are buried and could destabilize the TcSub2 structure. However, they are limited to the C-terminal domain and are therefore unlikely to cause disruption in the protein folding. Some of these amino acid substitutions (Y123F, I287V, G372A) were also observed in the Sub2 ortholog of *Plasmodium falciparum* (pFU52) and did not affect the activity of the protein domains (Figures 1C–E). pFU52 also has

RNA binding activity and is involved in ATPase and mRNA processing, similar to UAP56 in humans [43]. Since the TcSub2 structure is highly conserved, we assayed the functional ability of the TcSub2 protein to rescue the phenotype in *S. cerevisiae sub2* null mutants [44]. *SUB2* is an essential gene in yeast and the viability of the mutant cells is maintained by the episomal presence of the gene in a URA-3 plasmid. Functional complementation of heterologous genes in such strains can then be accessed by growth in FOA-containing plates. Since Sub2 levels are critical in yeast [45], we tested the expression of TcSub2 in low- and high-copy plasmids (as a control the *S. cerevisiae SUB2* gene was cloned into the same vectors). However, no growth was observed in any case even after 5 days of culture at 30°C (data not shown).

TcSub2 is a nuclear protein localized in mRNA transcription sites

The subcellular localization of TcSub2 was analyzed first by GFP fusion, where an exclusively nuclear distribution was observed by fluorescence microscopy in epimastigote forms (Figure 2, left). We observed that the protein is not evenly distributed in the nucleus, with some regions showing a more intense signal typical of a compartmentalized distribution. This distribution pattern was further confirmed by indirect immunofluorescence, using anti-TcSub2 polyclonal specific antibodies (Figure 2, right). This is a constitutive protein and the same pattern of protein localization occurs in different forms of the life cycle of the parasite and also in *Trypanosoma brucei* and *Leishmania major* (data not shown). To further investigate the nuclear distribution of TcSub2, ultrastructural immunocytochemical analyses were performed in *T. cruzi* epimastigotes (Figure 3). The results showed

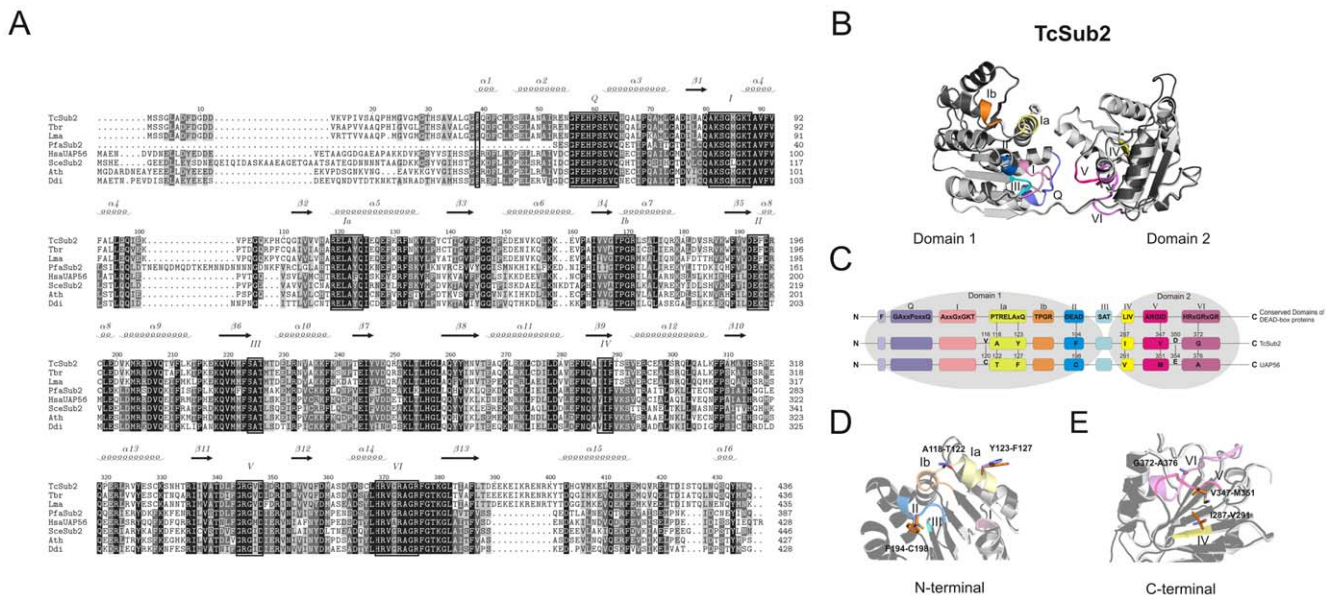


Figure 1. Comparison of amino acid sequences of TcSub2 with homolog proteins. (A) Amino acids are shaded according to their similarity: black for identical, dark grey if present in more than 60% of the sequences, and light grey if the similarity follows the BLOSUM 62 Matrix. When necessary, gaps were inserted within the various sequences (dashes) to allow alignment optimization. The nine motifs which are typical of DEAD-box RNA helicases are highlighted and predicted secondary structures are shown above the alignment. The GenBank accession numbers for the entries are listed as follows: TcSub2, (*T. cruzi* NP_808002.1); Tbr, (*T. brucei* XP_822312.1); Lma, (*Leishmania major* XP_001683110.1); PfaSub2, (*Plasmodium falciparum* XP_001349607.1); HsaUAP56, (Homo sapiens NP_004631.1); SceSub2, (*Saccharomyces cerevisiae* NP_010199.1); Ath, (*Arabidopsis thaliana* NP_568245.1); Ddi, (*Dictyostelium discoideum* XP_646415.1). (B) Molecular modeling of TcSub2 (light grey), based on the human homologue, HsaUAP56 (dark grey). Motifs are highlighted. Representations of the three-dimensional models were created with the software PyMol (<http://www.pymol.org>). (C) Analysis of TcSub2 motifs, compared with the general DEAD-box consensus and HsaUAP56 amino acid disposition. (D–E). The arrangement of different amino acids that differ between TcSub2 (light grey) and HsaUAP56 (dark grey) are evidenced in N-terminal (D) and C-terminal (E) regions.

doi:10.1371/journal.pone.0020730.g001

that this protein is localized mainly in less-compact chromatin regions (Figure 3A) and is absent from fibrillar center of the nucleolus (Figure 3B). Moreover, we observed the presence of grouped gold particles at the edge of electron-dense chromatin regions (Figure 3A). This distribution is similar to that of transcription sites [46], indicating that TcSub2 is located in active transcription regions. Analyses using indirect immunofluorescence combined with telomere FISH reinforced this notion because most of the protein does not colocalize with telomeres as indicated by the absence of TcSub2 signal over most of the telomeric repeats (Supplementary Figure S1). To test this hypothesis, *in situ* labeling of nascent RNAs followed by the immunofluorescence of TcSub2 was analyzed by confocal microscopy. This approach allowed the detection of BrUTP-labeled nascent RNAs (BrRNA), which corresponded to active transcription sites, and TcSub2 (Figure 4). The results showed that TcSub2 colocalizes with RNA pol II active transcription sites (Figure 4A). Based on previous work [31] incubation of the parasites with $75 \mu\text{g ml}^{-1}$ of α -amanitin inhibits only RNA pol II whereas $200 \mu\text{g ml}^{-1}$ inhibits both RNA pol II and III. We treated the cells using only $75 \mu\text{g ml}^{-1}$ of α -amanitin, but a very similar profile of signal detection was

observed compared with the $200 \mu\text{g ml}^{-1}$ (data not shown). We therefore decided to use $200 \mu\text{g ml}^{-1}$ of the drug as the inhibition concentration for RNA pol II. When BrUTP incorporation was carried out in the presence of the RNA pol II inhibitor α -amanitin, there was no colocalization of TcSub2 to the RNA pol I transcription site. This confirmed our previous finding that TcSub2 is excluded from the fibrillar center of nucleolus. The distribution of TcSub2 in relation to RNA pol II and RNA pol I transcription sites is better observed by comparison of the fluorescence intensity profiles of confocal pictures of the BrUTP-labeled nuclei. In the absence of α -amanitin treatment, fluorescence intensity peaks of nascent RNAs and TcSub2 often correlated (Figure 4B). However, after α -amanitin treatment, poor correlation was observed between the BrRNA and TcSub2 fluorescent peaks (Figure 4C). This colocalization pattern was observed in different focal planes (Supplementary Figure S2). In addition, no colocalization between TcSub2 and Spliced Leader RNA was observed by FISH and immunolocalization analysis (Figure 5). Taking together, these results strongly suggest that TcSub2 is related to mRNA transcription and/or processing in the nucleus of *T. cruzi*. We also observed uncorrelated peaks of

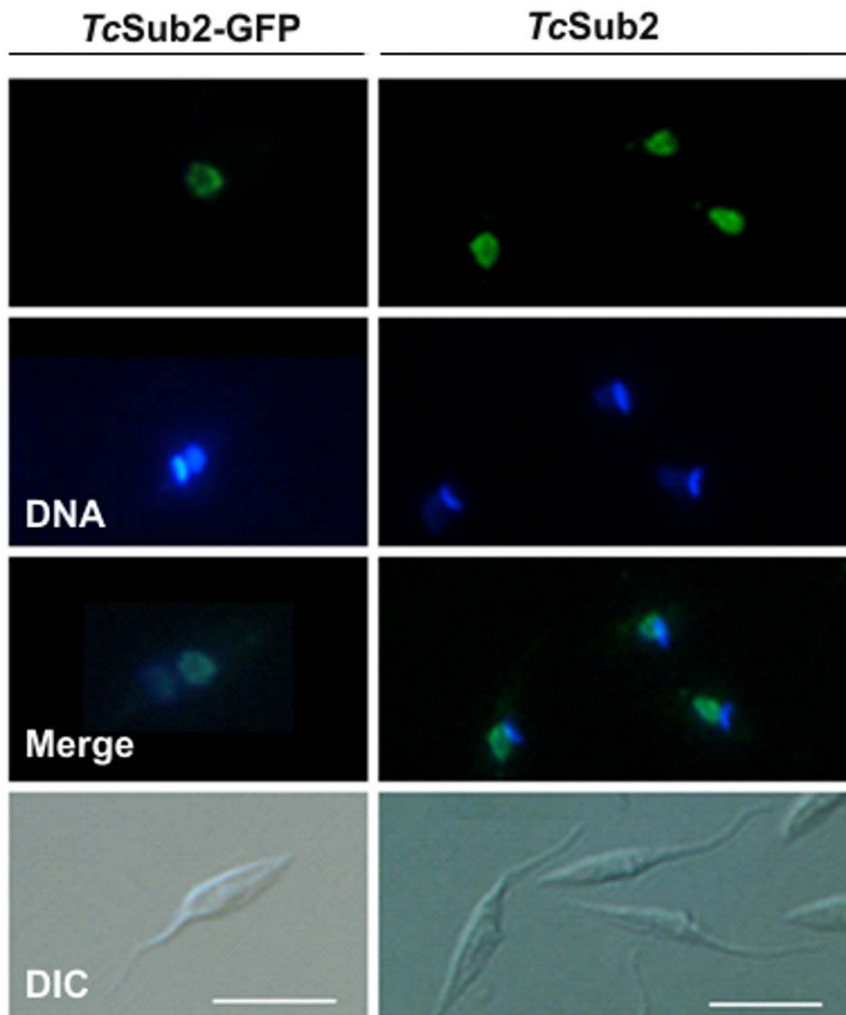


Figure 2. Localization of TcSub2 in *T. cruzi* Dm28c epimastigotes. The GFP fusion protein (TcSub2-GFP) and native protein (TcSub2) were localized by fluorescence microscopy. Specific polyclonal antibodies raised against TcSub2 were used to detect the native protein. Secondary antibodies were conjugated to Alexa-488. Cells were stained with DAPI to locate the nuclear and kinetoplast DNA. Bars = 10 μm . doi:10.1371/journal.pone.0020730.g002

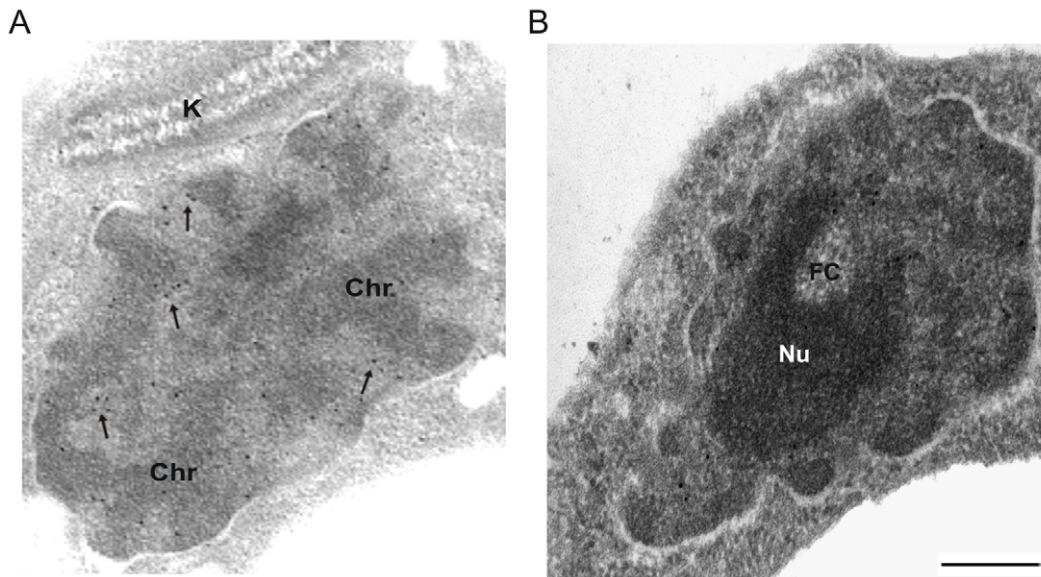


Figure 3. Immunoelectron microscopy localization of TcSub2 in *T. cruzi* Dm28c epimastigotes. (A) nuclear distribution of TcSub2: gold particles showing TcSub2 are seen at the edge of electron-dense chromatin (arrows). (B) Gold particles are present in the perinucleolar regions (arrows) and absent in the fibrillar center (FC) of nucleolus (Nu). Bars = 1 μ m. k = kinetoplast, Chr = dense chromatin. doi:10.1371/journal.pone.0020730.g003

TcSub2 and BrRNA, suggesting that RNA pol II transcription sites are not the only nuclear domain associated with TcSub2.

Trypanosome Sub2 ortholog proteins are essential for parasite survival

To analyze the function of TcSub2 we used a gene-knockout approach to abolish TcSub2 expression and observed the subsequent phenotypic alterations. A single-allele TcSub2 knockout of *T. cruzi* was successfully obtained (Figure 6). However, cell growth was not affected by the reduction in protein (Figure 6A). Western analysis using protein extracts from wild-type and single-allele knockout parasites, confirmed around 50% reduction on TcSub2 expression (Figure 6B). This reduction in protein expression is due to replacement of the TcSub2 allele by the neomycin selective marker as observed in the pulse field electrophoresis and Southern blot analysis (Figure 6C). We failed to achieve a null knockout because the cells were not viable during selection for hygromycin resistance. Three unsuccessful attempts were made to obtain the null knockout and we believe that this result is a consequence of the essential function of TcSub2 in *T. cruzi*. *Trypanosoma brucei* is a relevant model for analysis of gene function using an inducible system based on RNA interference [47]. Moreover, Sub2 is a highly conserved protein along eukaryotic evolution [8] and *T. brucei* has an ortholog protein which is 97% similar to TcSub2 (see alignment at Figure 1A). This encouraged us to use *T. brucei* as a suitable model to investigate the function of Sub2 in trypanosomes. First data showed that TbSub2 has an essential function since its knockdown in procyclic forms resulted in life cycle arrest in G2 phase (Figure 7A) with morphological changes (Figure 7B). Life cycle analysis by flow cytometry showed the presence of cells arrested in G2 phase (containing two nuclei and/or kinetoplasts) only after 48 h of RNAi induction, with an increasing of number of G2 arrested - cells after 72 h (Figure 7A). After 24 h of RNAi induction, the level of TbSub2 decreases around 45% and reaches 70% after 72 h of induction (Supplementary Figure S3A). The knockdown inhibits the cellular division after 24 h (Supplementary Figure S3B) but it

does not trigger death by apoptosis (assessed by phosphatidylserine exposure, Figure S4). In addition, the analysis of cell viability by propidium iodide internalization (Supplementary Figure S3C and S5A) and mitochondrial transmembrane potential (Supplementary Figure S3D and S5B) shows that a strong increase in the percentage of non-viable cells occurs only after 96 hours of RNAi induction. Therefore, the results observed in the present study can not be explained by cell death patterns, since necrotic cell death should be a late secondary event due to knockdown of TbSub2.

Knockdown of TbSub2 causes mRNA accumulation in the nuclei and translation inhibition

FISH assays to detect poliA⁺ RNA were performed to infer the distribution pattern of mRNA and investigate changes after decreasing Tryp-Sub2 protein levels in the parasite. As shown in supplementary figure S6, the distribution patterns of mRNA are very similar between wild type and single knockout cells of *T. cruzi*. It is observed a major signal in the cytoplasm and around the nucleus, but a very weak signal inside the nucleus. The analysis was done at different time-points of growth but no evident changes were observed.

By contrast, the knockdown of TbSub2 affected the mRNA distribution pattern in *T. brucei*. Differently from wild type cells, after 48 h of RNAi induction it is possible to observe a slight signal increase inside the nucleus (Figure 8A, on the top; Supplementary Figure S7); however the signal is still visible in the cytoplasm. This might correspond to mRNA accumulation in the nucleus that is clearly visible after 72 h of induction. (Figure 8A, below; Supplementary Figure S7). By analysis using metabolic labeling with [³⁵S]-methionine, it was demonstrated that the silencing of TbSub2 expression also caused translation inhibition after 48 h of induction (Figure 8B). It was possible to observe a deep decrease of radioactive counting in induced cells when compared to non-induced cells, similar to the result observed for non-induced cells treated with the translation inhibitor cycloheximide. We believe that the translation inhibition and mRNA accumulation is not due to cell death since knockdown of TbSub2 strongly affects the cell

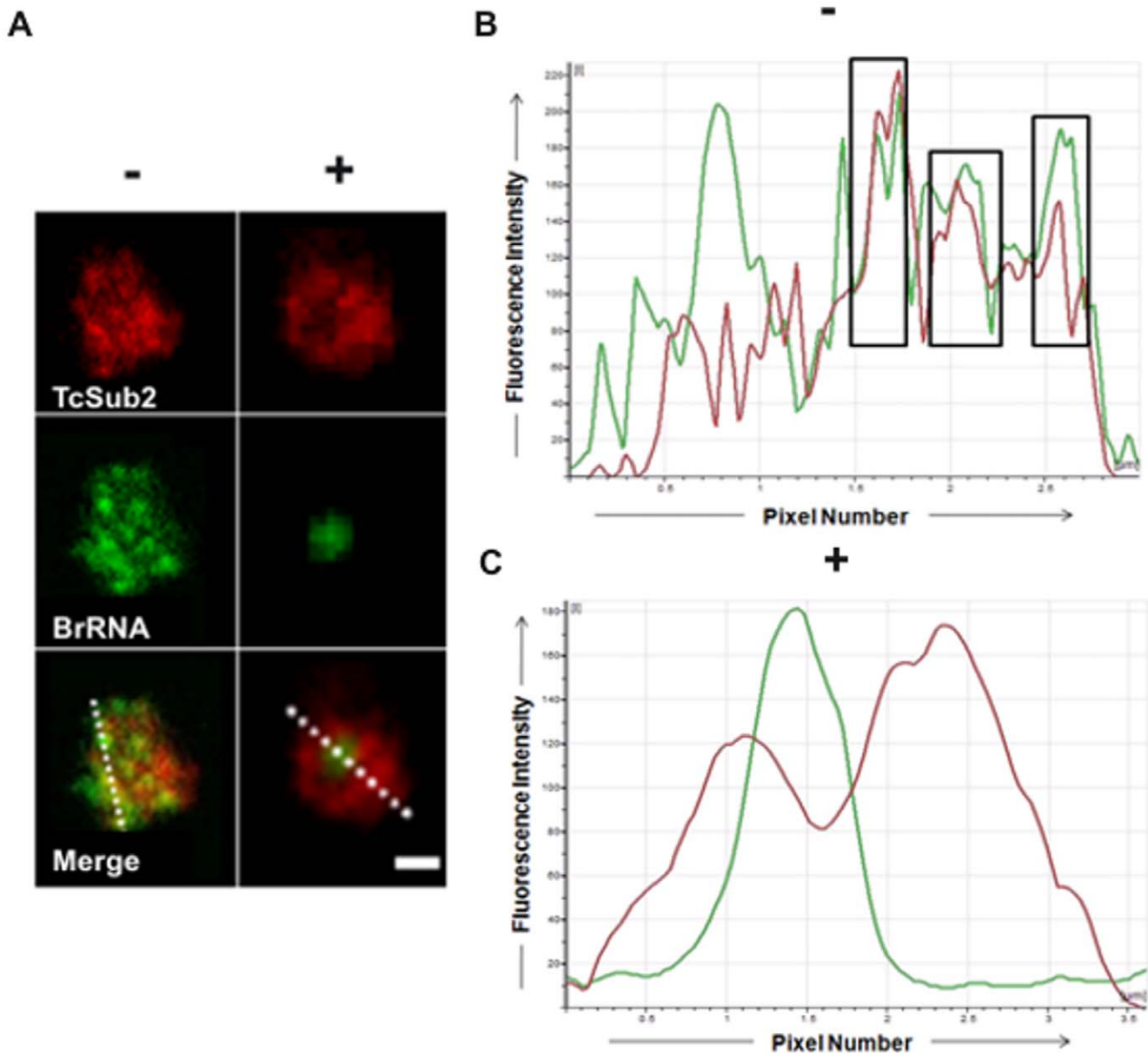


Figure 4. Nuclear colocalization of TcSub2 with active sites of transcription by detection of nascent RNAs in *T. cruzi*. (A) Immunolocalization of TcSub2 and active transcription sites (BrRNA) by confocal microscopy, without (–) and with (+) 200 µg ml⁻¹ of α-amanitin. (B) Analysis of colocalization of TcSub2 (red curves) and transcription sites (green curves) without α-amanitin. Squares show colocalization of TcSub2 and transcription sites. (C) Analysis of colocalization of TcSub2 (red curves) and transcription sites (green curves) with α-amanitin. The fluorescence intensity profiles of TcSub2 and transcription sites show the distribution of fluorescence across the dotted line (x-axis). The fluorescence intensities are plotted along the y-axis and the squares indicate colocalization. Bars = 1 µm.
doi:10.1371/journal.pone.0020730.g004

viability only after 96 h of induction, as shown before (Supplementary Figure S3C,D; Supplementary Figures S4 and S5).

Discussion

Post-transcriptional events are crucial for regulation of gene expression in trypanosomatids because of the absence of specific control mechanisms during transcription. Unlike most eukaryotic organisms, in which each gene transcribed by RNA Pol II has its own promoter, transcription in trypanosomatids is polycistronic without traditional promoter elements and genes in individual clusters do not necessarily code for functionally related proteins

[15]. Mature mRNAs are generated from primary transcripts by trans-splicing and polyadenylation and are then moved to the cytoplasm to be translated [48]. In the context of post-transcriptional events, the machinery of mRNA export is poorly understood in trypanosomes and this pathway might be an important step in regulation of gene expression in these parasites. We were therefore interested in investigating the factors that could be involved in this pathway.

The export of a few mRNAs in *T. cruzi* can be mediated by CRM1, a component of the RanGTP-exportin pathway. This pathway is commonly responsible for protein export but has no major role in mRNA export in higher eukaryotes [49]. Most

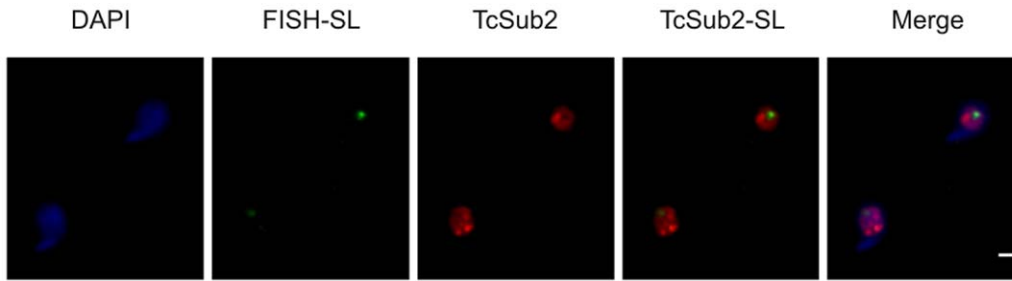


Figure 5. Nuclear colocalization of TcSub2 with SL transcription sites in *T. cruzi*. Immunofluorescence of TcSub2 (green) followed by FISH of SL transcription sites (red) was performed in epimastigote forms of *T. cruzi*. Cells were stained with DAPI to locate the nuclear and kinetoplast DNA. Bars = 1 μ m.

doi:10.1371/journal.pone.0020730.g005

reports related to this topic come from model organisms, especially *S. cerevisiae*, and the bulk of mRNA is exported in a RanGTP independent pathway involving the THO/TREX complex [revised by 2]. Based on our previous investigations, we started biological analysis of the most conserved component of the eukaryotic mRNA export pathway, the yeast DEAD-box RNA helicase Sub2 (UAP56 in humans). In the present study we cloned the gene encoding the *T. cruzi* protein that is highly similar to Sub2/UAP56, and has been named TcSub2. Sub2/UAP56 is a component of the TREX multiprotein complex that links transcription with mRNA export [revised by 50]. DEAD-box proteins are involved in the ATP-dependent unwinding of double-stranded RNA (dsRNA), displacement, RNA remodeling, or RNA/protein complexes. They are characterized by nine

conserved motifs distributed in two domains. In general, motifs I and II are implicated in ATP binding and hydrolysis, with contributions from motif VI. Motif III is believed to couple ATP hydrolysis with RNA unwinding, whereas motifs IV, V, and VI contribute to RNA binding [42,51]. We noted some changes of amino acid composition in the TcSub2 sequence when compared to the human sequence (UAP56). However, these changes did not affect the molecular model of TcSub2 that is very similar to the UAP56 crystal structure.

Although TcSub2 is highly conserved, it is unable to function as a substitute for SUB2. We speculate that this could be due to the specific functions of TcSub2 in *T. cruzi* because most transcripts in this parasite are processed by *trans*-splicing rather than *cis*-splicing as in other eukaryotic cells [52]. Specific point mutations could be

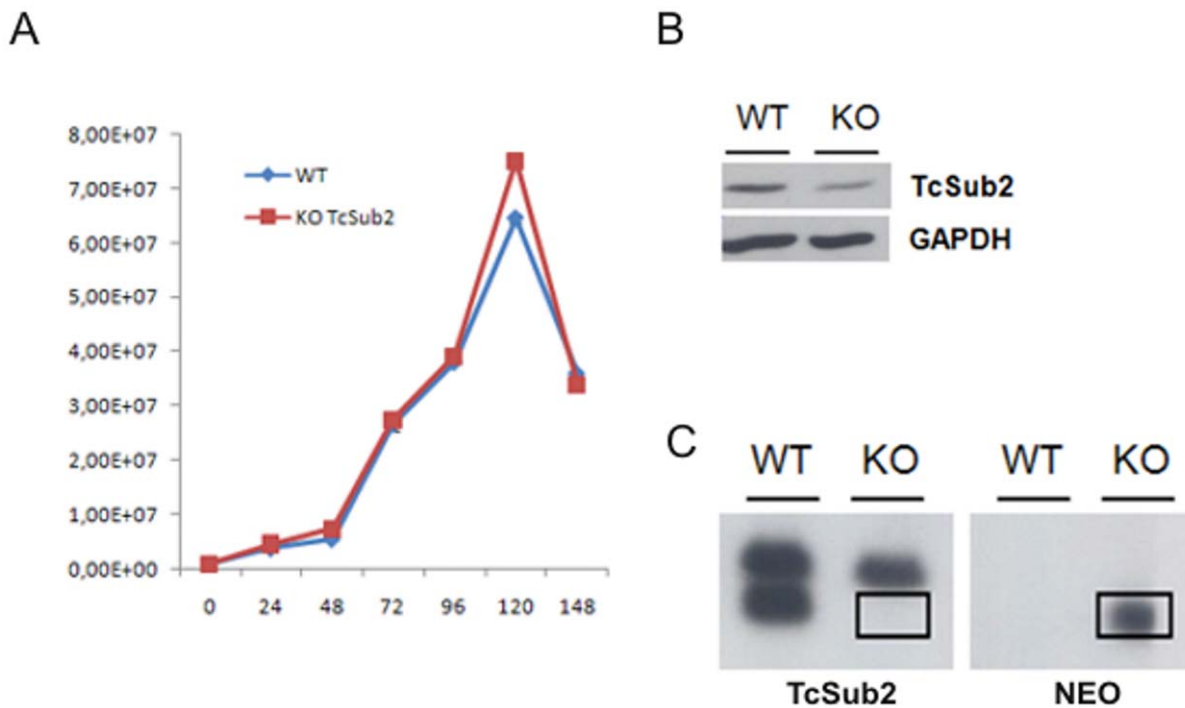


Figure 6. Cell proliferation analysis of single allele TcSub2 knockout of *T. cruzi*. (A) Growth curve of wild-type *T. cruzi* epimastigotes (blue squares) and single-allele knockout mutants (red squares). (B) Western-blot analysis comparing the expression of TcSub2 in wild-type (WT) and mutant parasites (KO). The membrane was incubated with 1:1000 polyclonal anti-TcSub2 and anti-GAPDH serum. The level of TcSub2 protein was decreased in single-allele TcSub2-knockout parasites. (C) Southern blot analysis confirming the knockout of one allele of TcSub2. The membrane containing chromosomes of wild-type (WT) and single-allele knockout mutant was hybridized with TcSub2 (TcSub2) and Neomycin genes (NEO) labeled with 32 P.

doi:10.1371/journal.pone.0020730.g006

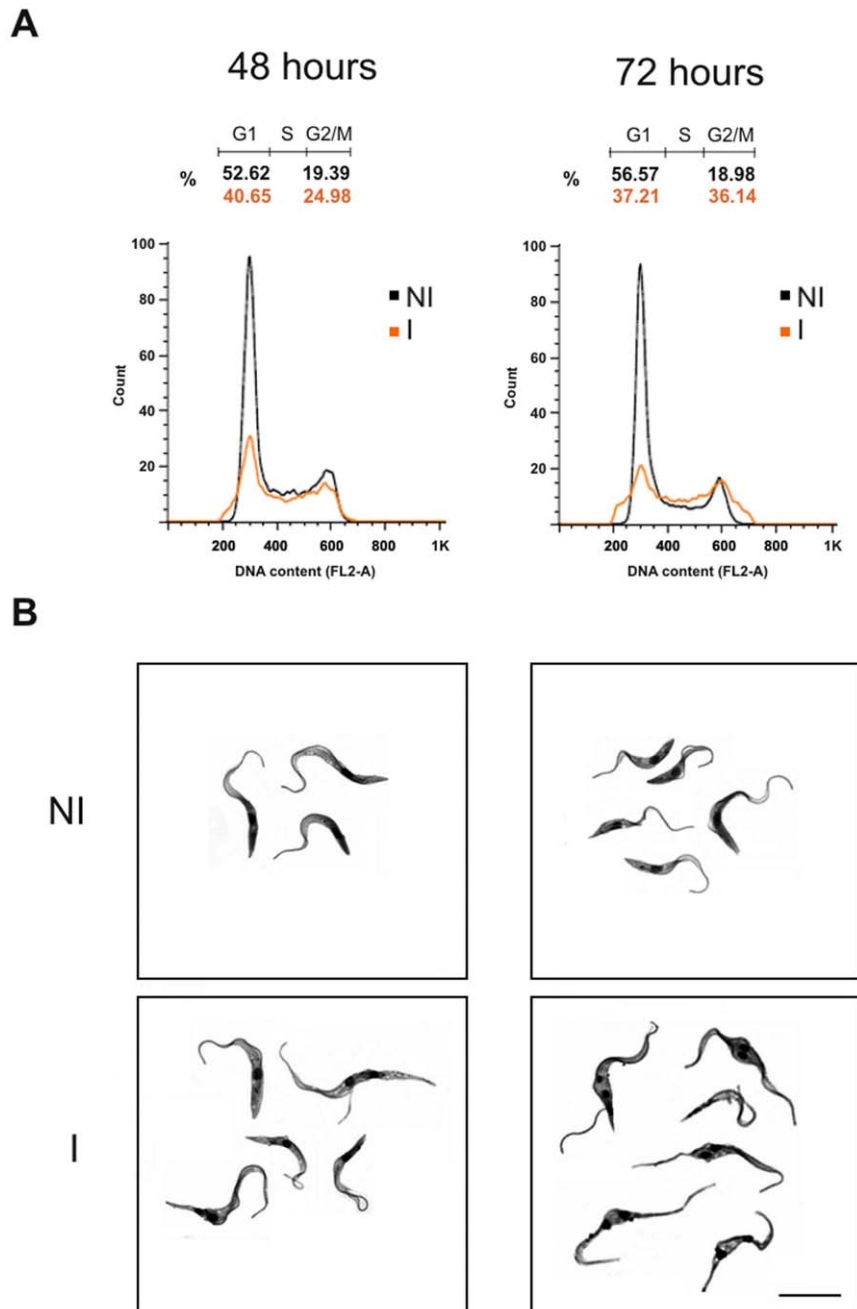


Figure 7. Analysis of cell cycle and morphology. (A) The cell cycle was analyzed by flow cytometry and the graphs correspond to data from RNAi assays after 48 and 72 hours of induction. Non-induced cells (NI) are indicated by black lines, induced cells (I) are indicated by orange lines. The cell cycle phases are indicated in each graph: G1, S and G2/M. (B) Morphology images from stained cells after 48 and 72 hours of RNAi induction obtained by confocal microscopy. Bars = 10 μ m.
doi:10.1371/journal.pone.0020730.g007

enough to create specific characteristics in the TcSub2 protein, and could, for example, prevent the interaction of this protein with components of TREX from *S.cerevisiae*. Such an example has been recently reported by Vázquez and colleagues (2009) [53]. TcU₂AF35, which is involved in the initial steps of *trans*-splicing in *T. cruzi*, is able to functionally complement yeast cells when only two aminoacid residues are modified. This demonstrates the importance of conserved residues for functional substitution of ortholog proteins between different organisms. Another possibility might be the development of novel components in the *T. cruzi* pathway, where specific factors would be necessary for functional activity of TcSub2.

In the export of mRNA, there is one revealing example involving the yeast Mex67-Mtr2 protein, an adaptor which is responsible for mRNA export, and its human counterpart, TAP-p15. TAP alone is unable to complement mex67- mutant cells neither does p15 complement mtr2⁻ mutants. However, the TAP-p15 complex partially restores the growth defect of mex67⁻/mtr2⁻ double mutants showing that TAP and p15 must interact to be functional in yeast whereas the TAP-Mtr2 and Mex67-p15 complexes may not form [54]. Like Sub2 and UAP56, TcSub2 is exclusively nuclear and is dispersed in loci all over the nuclei, and it is also present at the periphery of nucleolus, excluding the fibrillar center of nucleolus.

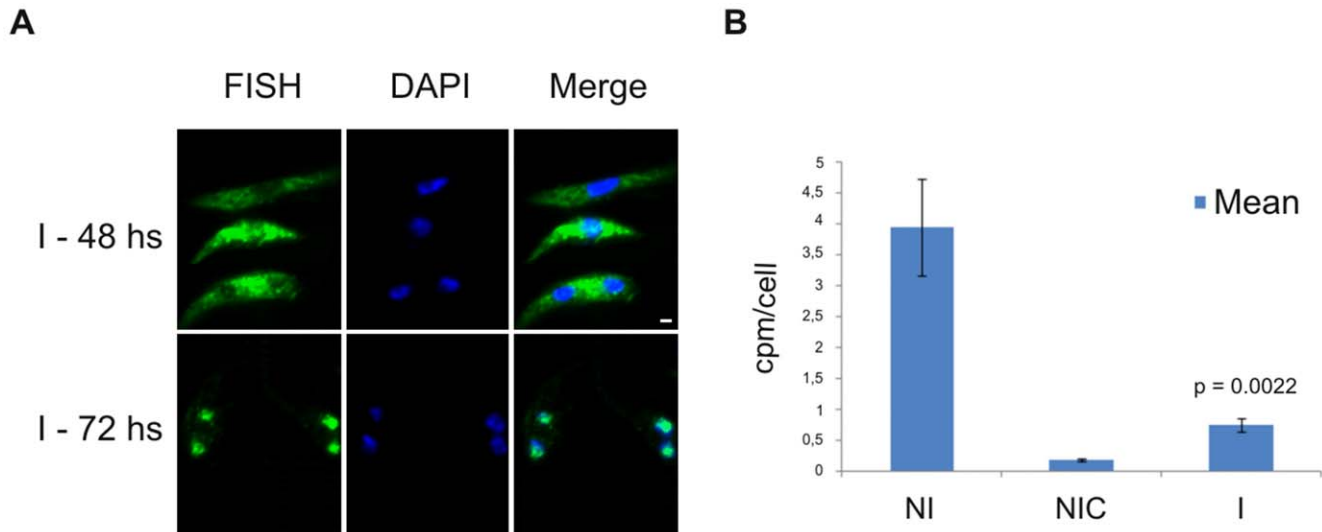


Figure 8. FISH of *poliA*⁺mRNA and translation level after *TbSub2* knockdown. (A) Fluorescence microscopy of *poliA*⁺ mRNAs by FISH using DIG-labeled oligo d(T)₃₀ after 48 and 72 hours of RNAi induction. Cells were stained with DAPI to localize the nuclear and kinetoplast DNA. Bars = 1 μm. (B) Metabolic labeling after 48 hours of RNAi induction. The graph shows the average of radioactive counting from triplicate experiments of non-induced cells (NI), non-induced cells treated with 100 μg ml⁻¹ of cycloheximide (NIC), and induced cells (I). The p value, calculated by unpaired t-test, is indicated in the graph. doi:10.1371/journal.pone.0020730.g008

The ultrastructural immunocytochemical assays showed that TcSub2 is concentrated in non-dense chromatin areas and grouped mainly at the interface between dense and non-dense chromatin. This speckled pattern of distribution is usually associated with mRNA transcription, processing, and nucleocytoplasmic export [45,55,56]. This distribution pattern is similar to human and arthropod (*Chironomus tentans*) TcSub2 homologues which are localized on the periphery of dense chromatin domains, termed interchromatin granule clusters (IGCs) [57]. These clusters contain mainly proteins related to mRNA processing, especially SC-35, and are closely related with perichromatin fibrils, where mRNA transcription occurs [revised by 58,59]. Many studies have demonstrated that nascent mRNAs are deposited in these interchromatin spaces [46,60–62]. A similar distribution pattern has been observed in trypanosomatids, at the interface between dense and non-dense chromatin for bromodomain factors (BDFs) and acetylated histones [63–65]. Recent investigations show that these histone modifications can serve as indicators of regions for initiation of RNA Pol II transcription [64,66,67]. Based on our findings for TcSub2 localization, we decided to investigate the relationship of TcSub2 and active transcription sites using BrUTP incorporation followed by immunocolocalization. This approach has been successfully used to observe the association of RNA Pol II transcription sites with proteins such as Hrp59 and Hrp65, and hnRNPs in *C. tentans*. It has also been used to define the localization of RNA Pol II transcription sites in *T. cruzi* [30,68,69]. Our results demonstrated that TcSub2 also colocalizes with nascent RNAs—specifically with those transcribed by RNA pol II—responsible for mRNA, snRNA and spliced-leader (SLRNA) transcription. Blocking of RNA pol II activity with α -amanitin abolished the transcription of nascent RNAs, resulting in RNAs transcribed only by RNA pol I. Under these conditions, we observed the absence of colocalization with TcSub2, indicating that TcSub2 might be associated with RNA pol II, but not with RNA pol I transcription. Besides, the protein does not localize with SL-RNA, reinforcing that the protein is related to mRNA transcription sites. The association of TcSub2 with RNA pol II transcription sites strongly suggests that

TcSub2 is functionally similar to homolog proteins in other eukaryotes. However we are still unable to confirm where TcSub2 is involved in the pathway of nuclear mRNA metabolism in this parasite. We also observed uncorrelated peaks of TcSub2 and BrRNA, suggesting that RNA pol II transcription sites are not the only nuclear domain associated with TcSub2. We speculate that the partial association with nascent mRNAs can be explained by the dynamics of interaction between different events in the cell, where DNA regions could be at different stages of transcription. Another explanation could be that TcSub2 is related to transcription of specific mRNAs as consequence of its role in regulation of gene expression. To further investigate the function of TcSub2, we attempted to knockout the TcSub2 gene to obtain phenotypic alterations related to mRNA export. However, the null knockout lineages of *T. cruzi* were not viable. Studies show that it is an essential protein in other eukaryotes because the knockdown of Sub2 homologs inhibits cell growth [45,55,70,71]. Unfortunately, knockout of a single TcSub2 allele was not enough to affect morphology and cell growth. Although the single-allele TcSub2 knockouts had around 50% less TcSub2 protein, this did not alter mRNA export. It is therefore necessary to apply a regulated system to completely silence TcSub2 expression to get more information regarding to its function in *T. cruzi*. The absence of approaches of inducible silencing, such as RNA interference, in *T. cruzi* preclude suitable studies for this. On the other hand, *T. brucei* is a suitable model to study essential genes using inducible system based on RNAi in trypanosomes. Taking in account that Sub2 is highly conserved in trypanosomes [8], we decided to use *T. brucei* to study the function of this protein. The data obtained demonstrated that it is really an essential protein since its knockdown inhibits the cellular growth. Similar to other eukaryotes, the silencing of TcSub2 expression caused translation inhibition and robust accumulation of bulk polyadenylated mRNAs in the nucleoplasm. These phenotypes are also observed when the expression of homologs is silenced in metazoan [45,55,71–76]. Interestingly, we also observed an arrest of cell cycle in G2 phase, mainly after 72 hours of RNAi induction when the level of the proteins was reduced by 70%. In the same

way, knockdown of UAP56, the homolog in humans, resulted in mitotic defects. Taking together, the trypanosome protein has similar characteristics to its orthologs in higher eukaryotes, possibly reflecting a similar function in these parasites.

In the context of nuclear mRNA metabolism in trypanosomads, nucleocytoplasmic transport is not well understood and data in this area will lend further insights into regulation of gene expression in these parasites. Herein, we identified a novel trypanosome protein that it is related to nuclear mRNA metabolism and proposed that the Tryp-Sub2 is a component of mRNA transcription/export pathway in these parasites.

Supporting Information

Figure S1 Nuclear colocalization of TcSub2 with telomere repeats. FISH of telomeric repeat regions followed by immunofluorescence of TcSub2 was performed in epimastigote forms. Telomeric repeats (red); TcSub2 (green). Cells were stained with DAPI to locate the nuclear and kinetoplast DNA. Bars = 1 μ m. (TIF)

Figure S2 (A, B) Nuclear colocalization of TcSub2 with transcription sites in different focal planes. Analysis of colocalization of TcSub2 (red) with transcription sites (BrRNA, green) without α -amanitin. Bars = 1 μ m. (TIF)

Figure S3 Cell proliferation and viability analysis. (A) Growth curves of induced (I) and non-induced (NI) cells by direct counting in a Neubauer chamber; each point represents the mean and standard deviations of triplicate experiments. Western blot of TbSub2 during RNAi induction. Antiserum against GAPDH was used as a protein-loading control. (B) Percentage of cells in G2 phase of cell cycle calculated by flow cytometry (details in methods); (C) Percentage of non-viable cells (dead cells) calculated by flow cytometry using propidium iodide staining; (D) Comparison of R123 fluorescence intensity between induced to non induced cells plotted as fold change graph using geometric mean of FL1-H detector intensity. For flow cytometry data (B–D), each experimental point represents the average and respective standard deviation of duplicate experiments. (TIF)

Figure S4 Analysis of phosphatidylserine exposure after TbSub2 knockdown by RNAi. The density plots show the non-induced (NI, top row) or induced (I, bottom row) cells co-stained with Annexin-V-FITC and Propidium Iodide after 48 to 96 hours (indicated above the plots) of RNAi induction. The values inside the density plots represent the percentage of cells inside each region. (TIF)

References

- Du TG, Schmid M, Jansen RP (2007) Why cells move messages: the biological functions of mRNA localization. *Semin Cell Dev Biol* 18: 171–177.
- Kohler A, Hurt E (2007) Exporting RNA from the nucleus to the cytoplasm. *Nat Rev Mol Cell Biol* 8: 761–773.
- Santos-Rosa H, Moreno H, Simos G, Segref A, Fahrenkrog B, et al. (1998) Nuclear mRNA export requires complex formation between Mex67p and Mtr2p at the nuclear pores. *Mol Cell Biol* 18: 6826–6838.
- Segref A, Sharma K, Doye V, Hellwig A, Huber J, et al. (1997) Mex67p, a novel factor for nuclear mRNA export, binds to both poly(A)+ RNA and nuclear pores. *Embo J* 16: 3256–3271.
- Chavez S, Beilharz T, Rondon AG, Erdjument-Bromage H, Tempst P, et al. (2000) A protein complex containing Tho2, Hpr1, Mft1 and a novel protein, Thp2, connects transcription elongation with mitotic recombination in *Saccharomyces cerevisiae*. *Embo J* 19: 5824–5834.
- Strasser K, Masuda S, Mason P, Pfannstiel J, Oppizzi M, et al. (2002) TREX is a conserved complex coupling transcription with messenger RNA export. *Nature* 417: 304–308.
- Aguilera A (2005) Cotranscriptional mRNP assembly: from the DNA to the nuclear pore. *Curr Opin Cell Biol* 17: 242–250.
- Serpeloni M, Vidal NM, Goldenberg S, Avila AR, Hoffmann FG (2011) Comparative genomics of proteins involved in RNA nucleocytoplasmic export. *BMC Evol Biol* 11: 7.
- Simpson AG, Inagaki Y, Roger AJ (2006) Comprehensive multigene phylogenies of excavate protists reveal the evolutionary positions of “primitive” eukaryotes. *Mol Biol Evol* 23: 615–625.
- Hampfl V, Hug L, Leigh JW, Dacks JB, Lang BF, et al. (2009) Phylogenomic analyses support the monophyly of Excavata and resolve relationships among eukaryotic “supergroups”. *Proc Natl Acad Sci U S A* 106: 3859–3864.
- Hotez PJ, Molyneux DH, Fenwick A, Kumaresan J, Sachs SE, et al. (2007) Control of neglected tropical diseases. *N Engl J Med* 357: 1018–1027.
- Rassi A, Jr., Rassi A, Marin-Neto JA (2010) Chagas disease. *Lancet* 375: 1388–1402.
- Coura JR, Borges-Pereira J (2010) Chagas disease: 100 years after its discovery. A systemic review. *Acta Trop* 115: 5–13.

Figure S5 Flow cytometry analysis of cell viability and cell cycle after TbSub2 knockdown by RNAi. (A) Overlay histograms of induced (I) and non-induced (NI) cells stained with Propidium Iodide after 24 to 96 hours (h, indicated inside the graphs) of RNAi induction; the percentage of PI positive cells (“Dead” region) is indicated. (B) Overlay histograms of I and NI cells stained with rhodamine 123 for mitochondrial membrane potential analysis. (C) Analysis after RNAi induction of DNA content by staining of permeabilized cells with propidium iodide; note the gradual increase of cells in G2/S phase after 48 hours of RNAi induction. (D) Cell cycle analysis by Dean-Jett-Fox algorithm of FlowJo software; data of non-induced (left histogram) and induced (right histogram) cells after 48 hours of RNAi induction are shown; the percentage of cells in each phase of cell cycle is indicated inside the graphs. (TIF)

Figure S6 FISH of polIA⁺ mRNA in *T. cruzi*. Localization of mRNAs by FISH using DIG-labeled oligo d(T)₃₀ in wild-type (WT) and single knockout of TcSub2 (KO) epimastigotes. Cells were stained with DAPI to locate the nuclear and kinetoplast DNA. For control, cells were treated with 100 μ g ml⁻¹ of RNase A before hybridization. Bars = 1 μ m. (TIF)

Figure S7 FISH of polIA⁺ mRNA in *T. brucei*. Localization of mRNAs by FISH using DIG-labeled oligo d(T)₃₀ after 48 and 72 hours of RNAi induction. Cells were stained with DAPI to locate the nuclear and kinetoplast in non-induced cells (NI) and induced cells (I). DNA. For control, cells were treated with 100 μ g ml⁻¹ RNase A before hybridization. Bars = 1 μ m. (TIF)

Acknowledgments

We thank Newton M. Vidal and Christian M. Probst for bioinformatic assistance, Nilson Fidêncio and Vanessa M. dos Santos for technical assistance, Miriam Giardini for providing amastigote forms, André Luis Bombo for figure assistance, Lysangela Ronalte Alves, Alejandro Correa and Bruno Dallagiovana for valuable comments, Michel Baptista and Flávia S. Pereira de Souza for providing the vectors for localization and knockout analysis and Sérgio Schenkman for providing the vector with SL gene.

Author Contributions

Conceived and designed the experiments: MS CBM ASPR SPF LHF-J ARA. Performed the experiments: MS JRCM MCMM ASPR RLK WDR AHI. Analyzed the data: MS JRCM ARA. Contributed reagents/materials/analysis tools: SFY-O MCMM SG LHF-J ARA. Wrote the paper: MS ARA. Provided guidance, direction, decision-making, and oversight: MS CBM SG LHF-J WDR ARA.

14. Teixeira SM, daRocha WD (2003) Control of gene expression and genetic manipulation in the Trypanosomatidae. *Genet Mol Res* 2: 148–158.
15. Martinez-Calvillo S, Vizuet-de-Rueda JC, Florencio-Martinez LE, Manning-Cela RG, Figueroa-Angulo EE (2010) Gene expression in trypanosomatid parasites. *J Biomed Biotechnol* 2010: 525241.
16. Contreras VT, Salles JM, Thomas N, Morel CM, Goldenberg S (1985) In vitro differentiation of *Trypanosoma cruzi* under chemically defined conditions. *Mol Biochem Parasitol* 16: 315–327.
17. Camargo EP (1964) Growth and Differentiation in *Trypanosoma Cruzi*. I. Origin of Metacyclic Trypanosomes in Liquid Media. *Rev Inst Med Trop Sao Paulo* 6: 93–100.
18. Elias MC, Marques-Porto R, Freymuller E, Schenkman S (2001) Transcription rate modulation through the *Trypanosoma cruzi* life cycle occurs in parallel with changes in nuclear organization. *Mol Biochem Parasitol* 112: 79–90.
19. Avila AR, Yamada-Ogatta SF, da Silva Monteiro V, Krieger MA, Nakamura CV, et al. (2001) Cloning and characterization of the metacyclogenin gene, which is specifically expressed during *Trypanosoma cruzi* metacyclogenesis. *Mol Biochem Parasitol* 117: 169–177.
20. Hull RN, Cherry WR, Tritch OJ (1962) Growth characteristics of monkey kidney cell strains LLC-MK1, LLC-MK2, and LLC-MK2(NCTC-3196) and their utility in virus research. *J Exp Med* 115: 903–918.
21. Schenkman S, Jiang MS, Hart GW, Nussenzweig V (1991) A novel cell surface trans-sialidase of *Trypanosoma cruzi* generates a stage-specific epitope required for invasion of mammalian cells. *Cell* 65: 1117–1125.
22. Wirtz E, Leal S, Ochatt C, Cross GA (1999) A tightly regulated inducible expression system for conditional gene knock-outs and dominant-negative genetics in *Trypanosoma brucei*. *Mol Biochem Parasitol* 99: 89–101.
23. Dhalia R, Marinsek N, Reis CR, Katz R, Muniz JR, et al. (2006) The two eIF4A helicases in *Trypanosoma brucei* are functionally distinct. *Nucleic Acids Res* 34: 2495–2507.
24. McGuffin LJ, Jones DT (2003) Improvement of the GenTHREADER method for genomic fold recognition. *Bioinformatics* 19: 874–881.
25. Shi H, Cordin O, Minder CM, Linder P, Xu RM (2004) Crystal structure of the human ATP-dependent splicing and export factor UAP56. *Proc Natl Acad Sci U S A* 101: 17628–17633.
26. Sali A, Blundell TL (1993) Comparative protein modelling by satisfaction of spatial restraints. *J Mol Biol* 234: 779–815.
27. Laskowski RA, Rullmann JA, MacArthur MW, Kaptein R, Thornton JM (1996) AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *J Biomol NMR* 8: 477–486.
28. Luthy R, Bowie JU, Eisenberg D (1992) Assessment of protein models with three-dimensional profiles. *Nature* 356: 83–85.
29. Vriend G (1990) WHAT IF: a molecular modeling and drug design program. *J Mol Graph* 8: 52–56, 29.
30. Dossin Fde M, Schenkman S (2005) Actively transcribing RNA polymerase II concentrates on spliced leader genes in the nucleus of *Trypanosoma cruzi*. *Eukaryot Cell* 4: 960–970.
31. Campbell DA, Thomas S, Sturm NR (2003) Transcription in kinetoplastid protozoa: why be normal? *Microbes Infect* 5: 1231–1240.
32. Batista M, Marchini FK, Celedon PA, Fragoso SP, Probst CM, et al. (2010) A high-throughput cloning system for reverse genetics in *Trypanosoma cruzi*. *BMC Microbiol* 10: 259.
33. Lu HY, Buck GA (1991) Expression of an exogenous gene in *Trypanosoma cruzi* epimastigotes. *Mol Biochem Parasitol* 44: 109–114.
34. de Souza FS, Rampazzo Rde C, Manhaes L, Soares MJ, Cavalcanti DP, et al. (2010) Knockout of the gene encoding the kinetoplast-associated protein 3 (KAP3) in *Trypanosoma cruzi*: effect on kinetoplast organization, cell proliferation and differentiation. *Mol Biochem Parasitol* 172: 90–98.
35. Tolivia J, Navarro A, del Valle E, Perez C, Ordóñez C, et al. (2006) Application of Photoshop and Scion Image analysis to quantification of signals in histochemistry, immunocytochemistry and hybridocytochemistry. *Anal Quant Cytol Histol* 28: 43–53.
36. Redmond S, Vadvivelu J, Field MC (2003) RNAi: an automated web-based tool for the selection of RNAi targets in *Trypanosoma brucei*. *Mol Biochem Parasitol* 128: 115–118.
37. Wickstead B, Ersfeld K, Gull K (2002) Targeting of a tetracycline-inducible expression system to the transcriptionally silent minichromosomes of *Trypanosoma brucei*. *Mol Biochem Parasitol* 125: 211–216.
38. Cassola A, Frasch AC (2009) An RNA recognition motif mediates the nucleocytoplasmic transport of a trypanosome RNA-binding protein. *J Biol Chem* 284: 35015–35028.
39. Schwede A, Manful T, Jha BA, Helbig C, Bercovich N, et al. (2009) The role of deadenylation in the degradation of unstable mRNAs in trypanosomes. *Nucleic Acids Res* 37: 5511–5528.
40. Freire ER, Dhalia R, Moura DM, da Costa Lima TD, Lima RP, et al. (2011) The four trypanosomatid eIF4E homologues fall into two separate groups, with distinct features in primary sequence and biological properties. *Mol Biochem Parasitol* 176: 25–36.
41. Worthen C, Jensen BC, Parsons M (2010) Diverse effects on mitochondrial and nuclear functions elicited by drugs and genetic knockdowns in bloodstream stage *Trypanosoma brucei*. *PLoS Negl Trop Dis* 4: e678.
42. Cordin O, Banroques J, Tanner NK, Linder P (2006) The DEAD-box protein family of RNA helicases. *Gene* 367: 17–37.
43. Shankar J, Pradhan A, Tuteja R (2008) Isolation and characterization of *Plasmodium falciparum* UAP56 homolog: evidence for the coupling of RNA binding and splicing activity by site-directed mutations. *Arch Biochem Biophys* 478: 143–153.
44. Kistler AL, Guthrie C (2001) Deletion of MUD2, the yeast homolog of U2AF65, can bypass the requirement for sub2, an essential spliceosomal ATPase. *Genes Dev* 15: 42–49.
45. Strasser K, Hurt E (2001) Splicing factor Sub2p is required for nuclear mRNA export through its interaction with Yra1p. *Nature* 413: 648–652.
46. Verschure PJ, van Der Kraan I, Manders EM, van Driel R (1999) Spatial relationship between transcription sites and chromosome territories. *J Cell Biol* 147: 13–24.
47. Djikeng A, Shen S, Tschudi C, Ullu E (2004) Analysis of gene function in *Trypanosoma brucei* using RNA interference. *Methods Mol Biol* 270: 287–298.
48. Clayton CE (2002) Life without transcriptional control? From fly to man and back again. *Embo J* 21: 1881–1888.
49. Cuevas IC, Frasch AC, D’Orso I (2005) Insights into a CRM1-mediated RNA-nuclear export pathway in *Trypanosoma cruzi*. *Mol Biochem Parasitol* 139: 15–24.
50. Moore MJ, Proudfoot NJ (2009) Pre-mRNA processing reaches back to transcription and ahead to translation. *Cell* 136: 688–700.
51. Hilbert M, Karow AR, Klostermeier D (2009) The mechanism of ATP-dependent RNA unwinding by DEAD box proteins. *Biol Chem* 390: 1237–1250.
52. Liang XH, Haritan A, Uliel S, Michaeli S (2003) trans and cis splicing in trypanosomatids: mechanism, factors, and regulation. *Eukaryot Cell* 2: 830–840.
53. Vazquez M, Atorrasagasti C, Bercovich N, Volcovich R, Levin MJ (2003) Unique features of the *Trypanosoma cruzi* U2AF35 splicing factor. *Mol Biochem Parasitol* 128: 77–81.
54. Katahira J, Strasser K, Podtelejnikov A, Mann M, Jung JU, et al. (1999) The Mex67p-mediated nuclear mRNA export pathway is conserved from yeast to human. *Embo J* 18: 2593–2609.
55. Gatfield D, Le Hir H, Schmitt C, Braun IC, Kocher T, et al. (2001) The DEXH/ D box protein HEL/UAP56 is essential for mRNA nuclear export in *Drosophila*. *Curr Biol* 11: 1716–1721.
56. Kota KP, Wagner SR, Huerta E, Underwood JM, Nickerson JA (2008) Binding of ATP to UAP56 is necessary for mRNA export. *J Cell Sci* 121: 1526–1537.
57. Kiesler E, Miralles F, Visa N (2002) HEL/UAP56 binds cotranscriptionally to the Balbiani ring pre-mRNA in an intron-independent manner and accompanies the BR mRNP to the nuclear pore. *Curr Biol* 12: 859–862.
58. Spector DL (2003) The dynamics of chromosome organization and gene regulation. *Annu Rev Biochem* 72: 573–608.
59. Zhao R, Bodnar MS, Spector DL (2009) Nuclear neighborhoods and gene expression. *Curr Opin Genet Dev* 19: 172–179.
60. Iborra FJ, Pombo A, Jackson DA, Cook PR (1996) Active RNA polymerases are localized within discrete transcription “factories” in human nuclei. *J Cell Sci* 109(Pt 6): 1427–1436.
61. Cmarko D, Verschure PJ, Martin TE, Dahms ME, Krause S, et al. (1999) Ultrastructural analysis of transcription and splicing in the cell nucleus after bromo-UTP microinjection. *Mol Biol Cell* 10: 211–223.
62. Cremer T, Cremer C (2001) Chromosome territories, nuclear architecture and gene regulation in mammalian cells. *Nat Rev Genet* 2: 292–301.
63. Nardelli SC, da Cunha JP, Motta MC, Schenkman S (2009) Distinct acetylation of *Trypanosoma cruzi* histone H4 during cell cycle, parasite differentiation, and after DNA damage. *Chromosoma* 118: 487–499.
64. Siegel TN, Hekstra DR, Kemp LE, Figueiredo LM, Lowell JE, et al. (2009) Four histone variants mark the boundaries of polycistronic transcription units in *Trypanosoma brucei*. *Genes Dev* 23: 1063–1076.
65. Villanova GV, Nardelli SC, Cribb P, Magdaleno A, Silber AM, et al. (2009) *Trypanosoma cruzi* bromodomain factor 2 (BDF2) binds to acetylated histones and is accumulated after UV irradiation. *Int J Parasitol* 39: 665–673.
66. Respuela P, Ferella M, Rada-Iglesias A, Aslund L (2008) Histone acetylation and methylation at sites initiating divergent polycistronic transcription in *Trypanosoma cruzi*. *J Biol Chem* 283: 15884–15892.
67. Figueiredo LM, Cross GA, Janzen CJ (2009) Epigenetic regulation in African trypanosomes: a new kid on the block. *Nat Rev Microbiol* 7: 504–513.
68. Percipalle P, Fomproix N, Kylberg K, Miralles F, Bjorkroth B, et al. (2003) An actin-ribonucleoprotein interaction is involved in transcription by RNA polymerase II. *Proc Natl Acad Sci U S A* 100: 6475–6480.
69. Kiesler E, Hase ME, Brodin D, Visa N (2005) Hrp59, an hnRNP M protein in *Chironomus* and *Drosophila*, binds to exonic splicing enhancers and is required for expression of a subset of mRNAs. *J Cell Biol* 168: 1013–1025.
70. Zhang M, Green MR (2001) Identification and characterization of yUAP/Sub2p, a yeast homolog of the essential human pre-mRNA splicing factor hUAP56. *Genes Dev* 15: 30–35.
71. MacMorris M, Brocker C, Blumenthal T (2003) UAP56 levels affect viability and mRNA export in *Caenorhabditis elegans*. *Rna* 9: 847–857.
72. Kapadia F, Pryor A, Chang TH, Johnson LF (2006) Nuclear localization of poly(A)⁺ mRNA following siRNA reduction of expression of the mammalian RNA helicases UAP56 and URH49. *Gene* 384: 37–44.
73. Hautbergue GM, Hung ML, Walsh MJ, Snijders AP, Chang CT, et al. (2009) UIF, a new mRNA export adaptor that works together with REF/ALY, requires FACT for recruitment to mRNA. *Curr Biol* 19: 1918–1924.

74. Yamazaki T, Fujiwara N, Yukinaga H, Ebisuya M, Shiki T, et al. (2010) The closely related RNA helicases, UAP56 and URH49, preferentially form distinct mRNA export machineries and coordinately regulate mitotic progression. *Mol Biol Cell* 21: 2953–2965.
75. Dias AP, Dufu K, Lei H, Reed R (2010) A role for TREX components in the release of spliced mRNA from nuclear speckle domains. *Nat Commun* 1: 97.
76. Sahni A, Wang N, Alexis JD (2010) UAP56 is an important regulator of protein synthesis and growth in cardiomyocytes. *Biochem Biophys Res Commun* 393: 106–110.

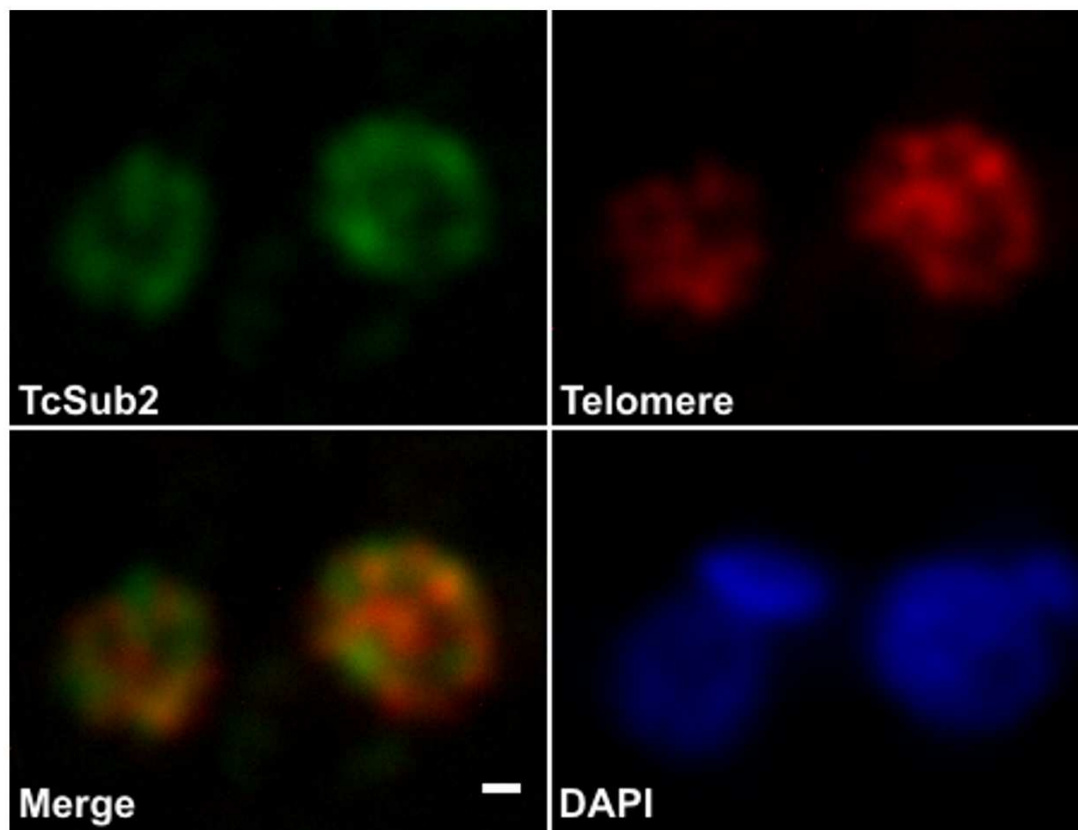


Figure S1. Nuclear colocalization of TcSub2 with telomere repeats. FISH of telomeric repeat regions followed by immunofluorescence of TcSub2 was performed in epimastigote forms. Telomeric repeats (red); TcSub2 (green). Cells were stained with DAPI to locate the nuclear and kinetoplast DNA. Bars = 1 μ m.

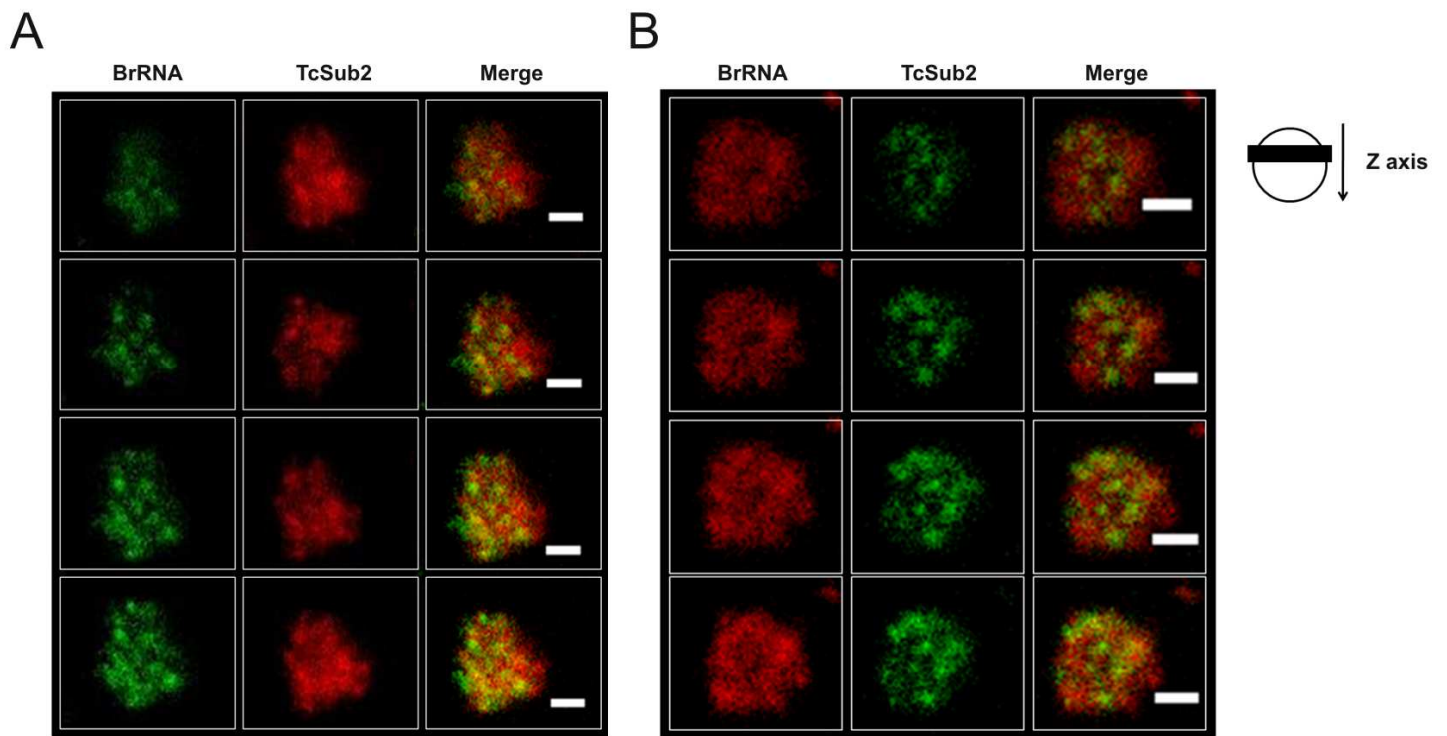


Figure S2. (A, B) Nuclear colocalization of TcSub2 with transcription sites in different focal planes. Analysis of colocalization of TcSub2 (red) with transcription sites (BrRNA, green) without α -amanitin. Bars = 1 μ m.

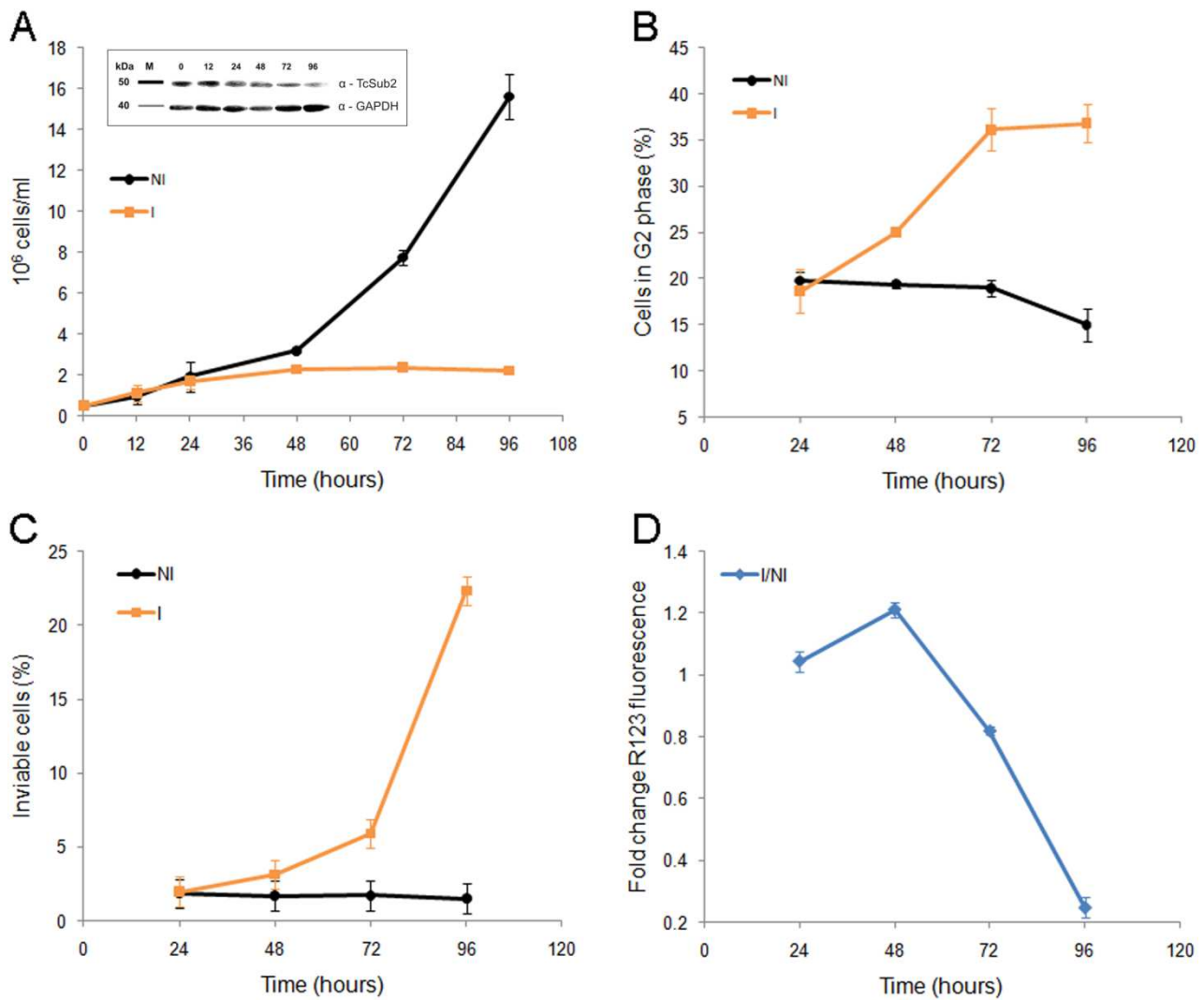


Figure S3. Cell proliferation and viability analysis. (A) Growth curves of induced (I) and non-induced (NI) cells by direct counting in a Neubauer chamber; each point represents the mean and standard deviations of triplicate experiments. Western blot of TbSub2 during RNAi induction. Antiserum against GAPDH was used as a protein-loading control. (B) Percentage of cells in G2 phase of cell cycle calculated by flow cytometry (details in methods); (C) Percentage of non-viable cells (dead cells) calculated by flow cytometry using propidium iodide staining; (D) Comparison of R123 fluorescence intensity between induced to non induced cells plotted as fold change graph using geometric mean of FL1-H detector intensity. For flow cytometry data (B–D), each experimental point represents the average and respective standard deviation of duplicate experiments.

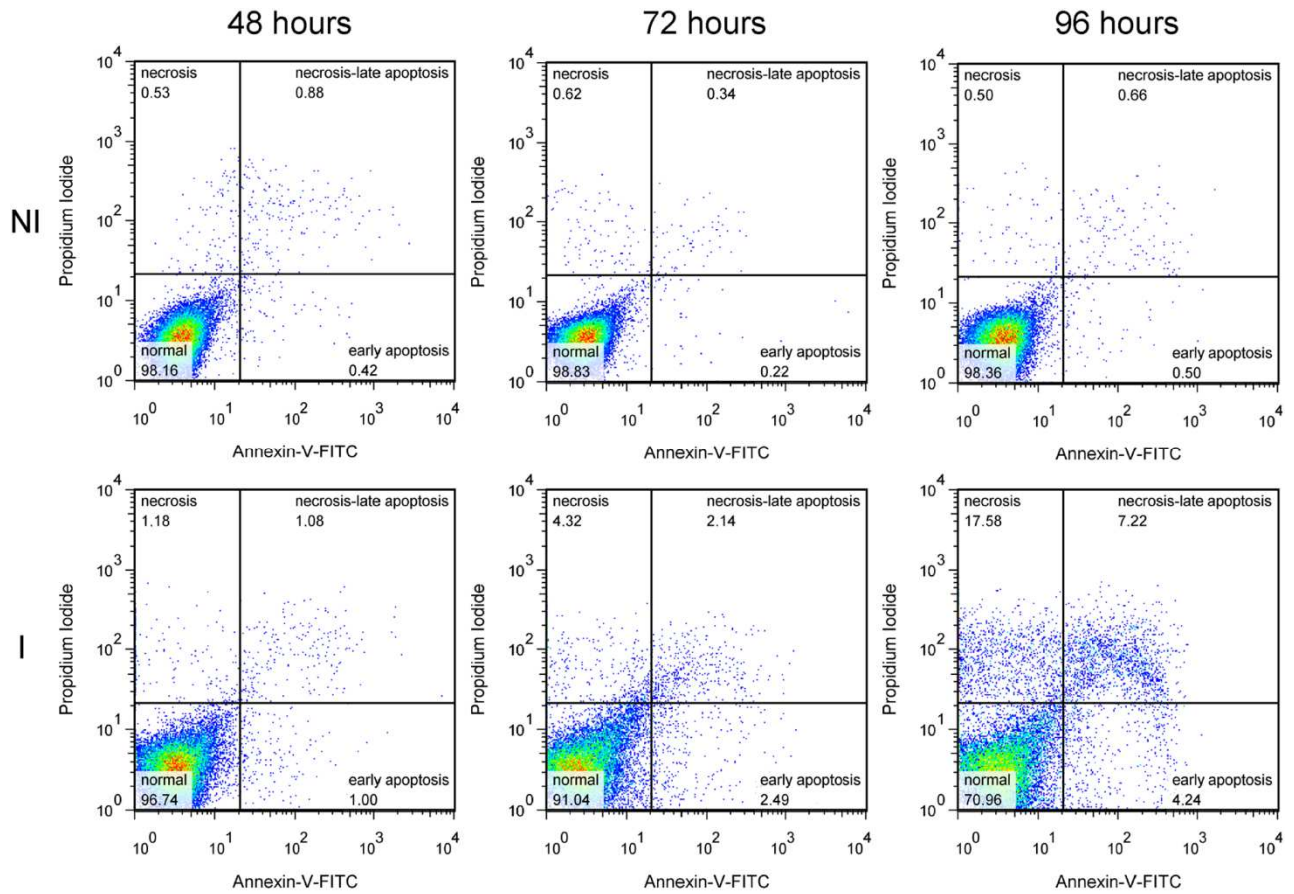


Figure S4. Analysis of phosphatidylserine exposure after *TbSub2* knockdown by RNAi. The density plots show the non-induced (NI, top row) or induced (I, bottom row) cells co-stained with Annexin-V-FITC and Propidium Iodide after 48 to 96 hours (indicated above the plots) of RNAi induction. The values inside the density plots represent the percentage of cells inside each region.

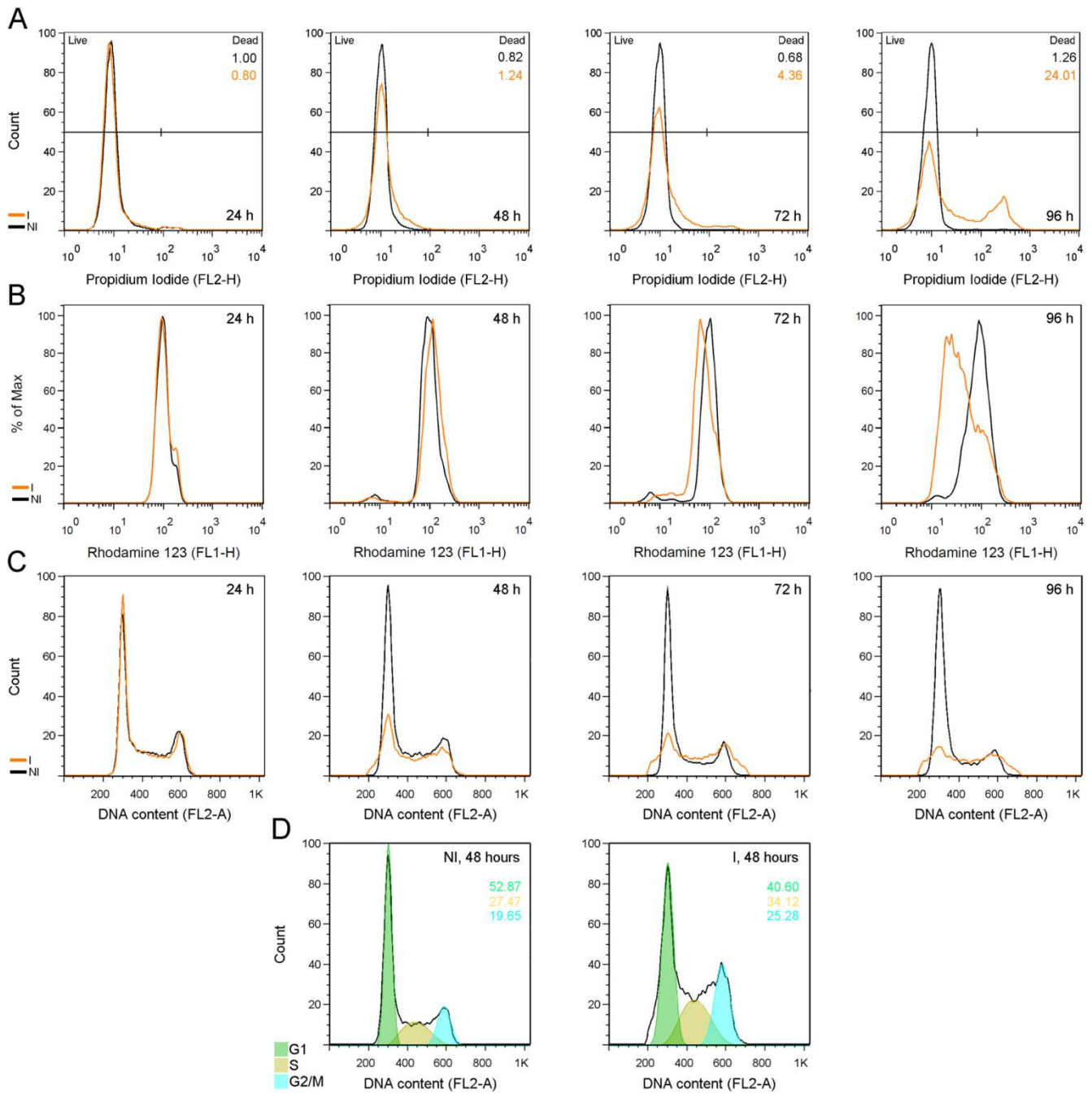


Figure S5. Flow cytometry analysis of cell viability and cell cycle after TbSub2 knockdown by RNAi. (A) Overlay histograms of induced (I) and non-induced (NI) cells stained with Propidium Iodide after 24 to 96 hours (h, indicated inside the graphs) of RNAi induction; the percentage of PI positive cells (“Dead” region) is indicated. (B) Overlay histograms of I and NI cells stained with rhodamine 123 for mitochondrial membrane potential analysis. (C) Analysis after RNAi induction of DNA content by staining of permeabilized cells with propidium iodide; note the gradual increase of cells in G2/S phase after 48 hours of RNAi induction. (D) Cell cycle analysis by Dean-Jett-Fox algorithm of FlowJo software; data of non-induced (left histogram) and induced (right histogram) cells after 48 hours of RNAi induction are shown; the percentage of cells in each phase of cell cycle is indicated inside the graphs.

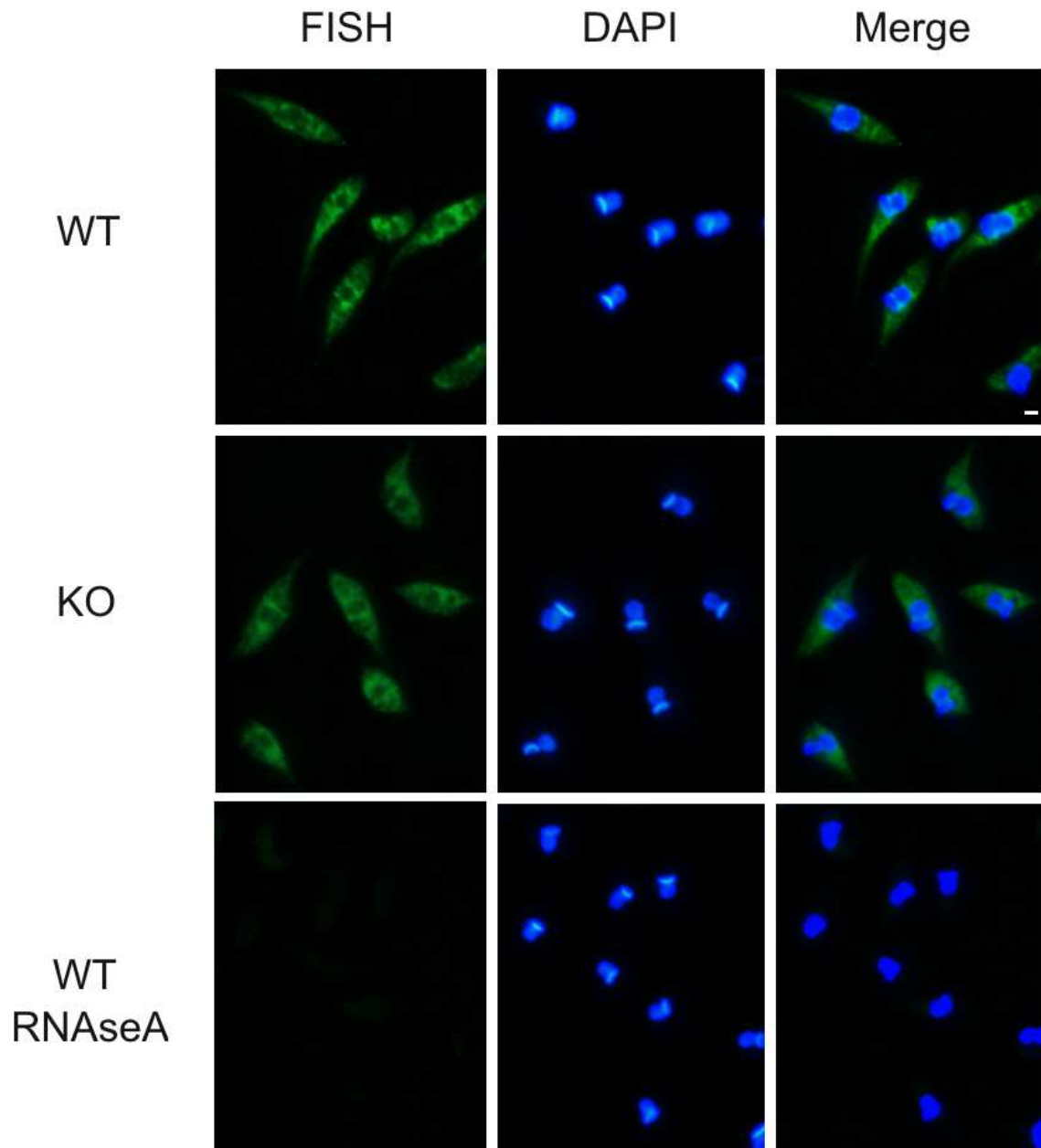


Figure S6. FISH of *polyA*⁺mRNA in *T. cruzi*. Localization of mRNAs by FISH using DIG-labeled oligo d(T)₃₀ in wild-type (WT) and single knockout of TcSub2 (KO) epimastigotes. Cells were stained with DAPI to locate the nuclear and kinetoplast DNA. For control, cells were treated with 100 $\mu\text{g ml}^{-1}$ of RNase A before hybridization. Bars = 1 μm .

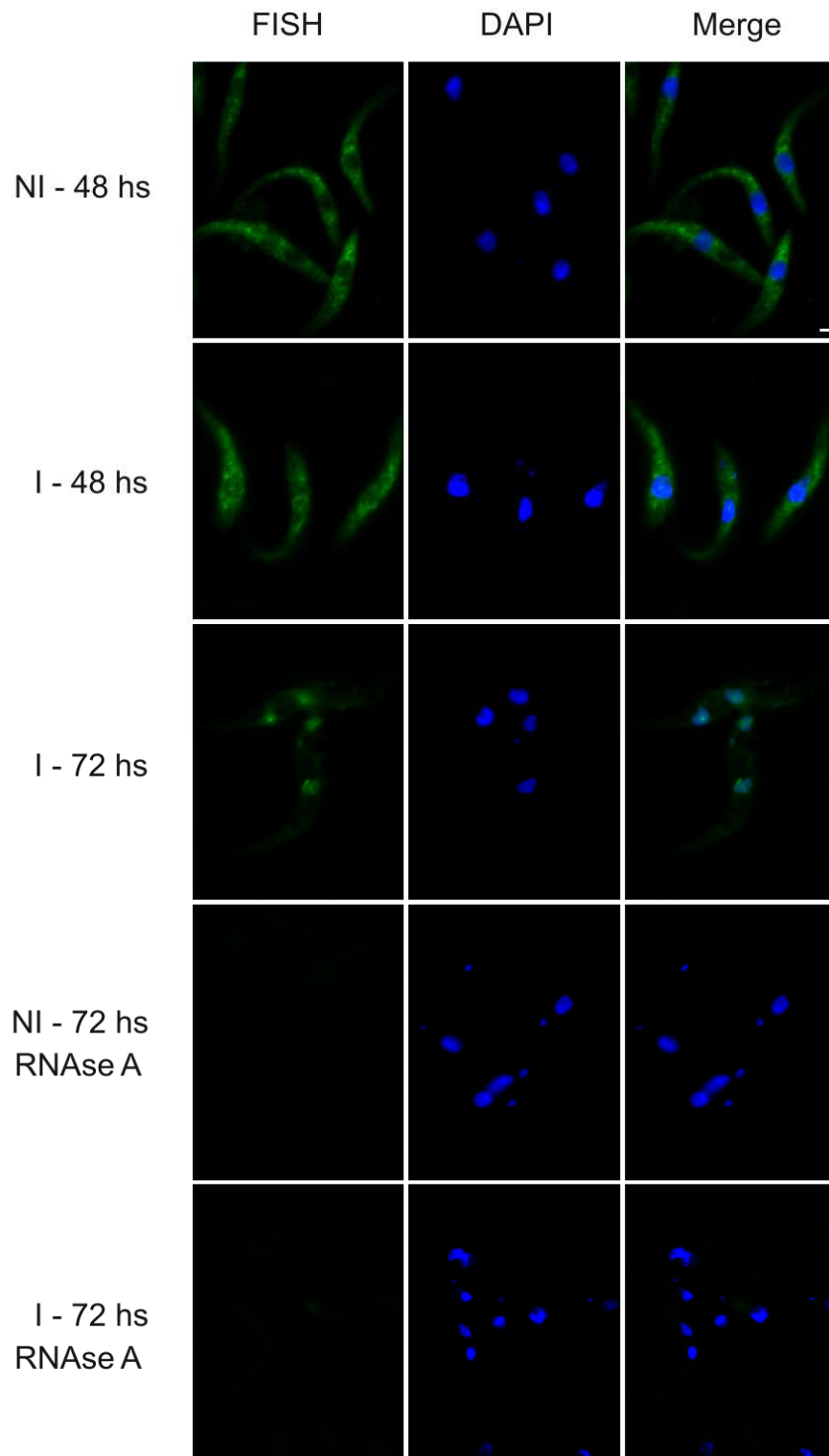


Figure S7. FISH of *poliA*⁺mRNA in *T. brucei*. Localization of mRNAs by FISH using DIG-labeled oligo d(T)₃₀ after 48 and 72 hours of RNAi induction. Cells were stained with DAPI to locate the nuclear and kinetoplast in non-induced cells (NI) and induced cells (I). DNA. For control, cells were treated with 100 $\mu\text{g ml}^{-1}$ RNase A before hybridization. Bars = 1 μm .

5 Capítulo 3.

TgUAP56: an essential nuclear protein related with mRNA export in *T. gondii* (artigo em preparação – Anexo I)

Como citado no Capítulo 1, UAP56/Sub2 é uma proteína altamente conservada ao longo da filogenia de eucariotos, inclusive em protozoários pertencentes a supergrupos que divergiram na base da árvore de filogenia de eucariotos. Dentre eles se encontra o Chromalveolata, ao qual pertencem parasitas do Filo Apicomplexa de importância médica como *P. falciparum* e *T. gondii*, agentes causadores de malária e toxoplasmose, respectivamente. Dentre os parasitas deste supergrupo, *T. gondii* tem se mostrado um importante modelo para estudo sistemático de genes essenciais, com uma ampla gama de ferramentas de genética reversa bem estabelecidas (KIM e WEISS, 2004; MEISSNER *et al.*, 2007).

Durante o período de Novembro de 2012 a Maio de 2013 realizei uma parte do Doutorado na Universidade de Glasgow sob orientação do Dr. Markus Meissner, que tem grande experiência em produção e utilização de ferramentas de genética reversa para estudo de genes essenciais em *T. gondii*.

Sendo assim, visto que foi estabelecido um método de nocaute para estudos de genes essenciais em *T. gondii*, e considerando se tratar de um parasita pertencente a um grupo diferente de *T. cruzi*, resolvemos verificar se a função de Sub2 estaria conservada nestes organismos. A proteína ortóloga de UAP56/Sub2 em *T. gondii* foi nomeada de TgUAP56. Primeiramente confirmamos a alta conservação de TgUAP56 ao longo da filogenia de eucariotos através de alinhamento múltiplo entre diferentes proteínas ortólogas e verificamos a relação da proteína com a exportação de mRNAs através de análise de distribuição celular de mRNAs após a superexpressão da proteína dd-GFP-TgUAP56 e após nocaute condicionado do respectivo gene *uap56*. Foi verificada uma relação direta entre esta proteína e exportação de mRNAs no parasita, assim como verificado em tripanossomatídeos (SERPELONI *et al.*, 2011). A outra etapa do trabalho foi o estabelecimento de uma metodologia de análise fenotípica baseado no sistema MS2 para o monitoramento do transporte de mRNA entre o núcleo e citoplasma. Futuramente, como prova de conceito desta metodologia, será utilizado o nocaute condicionado do gene *uap56*

de *T. gondii* para comprovar a aplicabilidade do método. Além disso, pretende-se usar o sistema para análises em média/larga escala através de ensaios de varredura de drogas na tentativa de buscar novos alvos para quimioterapia.

5.1 Materiais e Métodos

5.1.1 Análises de sequências de proteínas ortólogas de UAP56 e filogenia por ferramentas de bioinformática

As sequências das proteínas UAP56 e Sub2, de humano e levedura respectivamente, foram utilizadas como sementes de busca (*queries*) para identificar proteínas ortólogas em 43 espécies representativas dos diferentes grupos eucarióticos, listados na figura 11: Metazoans, Fungi, Amoebozoa, Plants, Parabasilids, Kinetoplastida e Apicomplexa. As sequências de proteínas de todas as espécies foram retiradas do banco de dados do RefSeq (*Reference Sequence*) do NCBI (*National Center for Biotechnology Information*). Para isso foi feito o *download* do proteoma completo de todas as espécies a partir do banco de dados RefSeq (*Reference Sequence*) do NCBI (*National Center for Biotechnology Information*). As buscas foram feitas utilizando o programa BLASTP e parâmetros padrão (ALTSCHUL *et al.*, 1997) e as proteínas ortólogas foram identificadas através do critério de *reciprocal best hit* (RBH). Os domínios das proteínas foram identificados através do programa hmmscan do pacote HMMER (EDDY, 1998), buscando contra a coleção de modelos ocultos de Markov (*Hidden Markov Models* - HMM) disponíveis no banco de dados do Pfam, versão 27.0 (FINN *et al.*, 2014). O alinhamento de múltiplas sequências ortólogas foi realizado utilizando o programa MAFFT versão 7 e os parâmetros: *localpair - maxiterate* 1000 – *reorder* (KATO e STANDLEY, 2013). A análise filogenética foi realizada utilizando uma aproximação do método de Máxima Verossimilhança (*maximum likelihood*) implementada no programa FastTree 2.1 (PRICE, DEHAL e ARKIN, 2010), utilizando o modelo WAG. A árvore filogenética foi desenhada utilizando o programa FigTree v1.4 (<http://tree.bio.ed.ac.uk/software/figtree/>). Para o cálculo de similaridade e identidade entre os pares de sequências de aminoácidos foi utilizado o programa needle do pacote EMBOSS (RICE, LONGDEN e BLEASBY, 2000), que permite o melhor alinhamento global entre duas sequências de proteínas.

Eukaryotic Group	Species	Code	GI
Metazoans	<i>Homo sapiens</i>	Hsa	4758112
Metazoans	<i>Mus musculus</i>	Mmu	9790069
Metazoans	<i>Ciona intestinalis</i>	Cin	198430623
Metazoans	<i>Drosophila melanogaster</i>	Dme	24581952
Metazoans	<i>Apis mellifera</i>	Ame	66532824
Metazoans	<i>Caenorhabditis briggsae</i>	Cbr	157749783
Metazoans	<i>Caenorhabditis elegans</i>	Cel	71982858
Choanoflagellate	<i>Monosiga brevicollis</i>	Mbr	167522271
Fungi	<i>Saccharomyces cerevisiae</i>	Sce	6320119
Fungi	<i>Candida glabrata</i>	Cgl	50293331
Fungi	<i>Ashbya gossypii</i>	Ago	45188133
Fungi	<i>Kluyveromyces lactis</i>	Kla	50309861
Fungi	<i>Debaryomyces hanseni</i>	Dha	50424083
Fungi	<i>Yarrowia lipolytica</i>	Yli	210075100
Fungi	<i>Magnaporthe oryzae</i>	Mor	145610002
Fungi	<i>Neurospora crassa</i>	Ncr	164423323
Fungi	<i>Fusarium graminearum</i>	Fgr	46128571
Fungi	<i>Aspergillus fumigatus</i>	Afu	70984615
Fungi	<i>Schizosaccharomyces pombe</i>	Spo	19115173
Fungi	<i>Cryptococcus neoformans</i>	Cne	58259249
Fungi	<i>Ustilago maydis</i>	Uma	71021713
Amoebozoa	<i>Dictyostelium discoideum</i>	Ddi	66826121
Plantae	<i>Arabidopsis thaliana</i>	Ath	18416493
Plantae	<i>Oryza sativa</i>	Osa	115437448
Plantae	<i>Populus trichocarpa</i>	Pop	224092470
Plantae	<i>Physcomitrella patens</i>	Ppa	168035593
Plantae	<i>Chlamydomonas reinhardtii</i>	Cre	159491657
Plantae	<i>Volvox carteri</i>	Vca	302836427
Plantae	<i>Ostreococcus lucimarinus</i>	Olu	145356615
Apicomplexa	<i>Theileria parva</i>	Tpa	71027499
Apicomplexa	<i>Plasmodium berghei</i>	Pbe	68073823
Apicomplexa	<i>Plasmodium falciparum</i>	Pfa	124801080
Apicomplexa	<i>Plasmodium knowlesi</i>	Pkn	221053125
Apicomplexa	<i>Plasmodium vivax</i>	Pvi	156093980
Apicomplexa	<i>Toxoplasma gondii</i>	Tgo	237843393
Apicomplexa	<i>Cryptosporidium hominis</i>	Cho	67623541
Apicomplexa	<i>Cryptosporidium parvum</i>	Cpa	66361130
Kinetoplastida	<i>Leishmania braziliensis</i>	Lbr	154337278
Kinetoplastida	<i>Leishmania infantum</i>	Lin	146086104
Kinetoplastida	<i>Leishmania major</i>	Lma	157869116
Kinetoplastida	<i>Trypanosoma brucei</i>	Tbr	71746514
Kinetoplastida	<i>Trypanosoma cruzi</i>	Tcr	71411514
Parabasalid	<i>Trichomonas vaginalis</i>	Tva	123449042

Figura 11. Número de identificação das proteínas ortólogas de UAP56/Sub2 do GeneBank (GI) identificadas em 43 espécies representativas dos grupos de eucariotos, pintados em cores. As sequências foram utilizadas para a análise de filogenia. Amoebozoa: Ddi – *Dictyostelium discoideum*; Plants: Osa – *Oryza sativa*, Ppa – *Physcomitrella patens*, Ath – *Arabidopsis thaliana*, Pop – *Populus trichocarpa*, Olu – *Ostreococcus lucimarinus*, Cre – *Chlamydomonas reinhardtii*, Vca – *Volvox carteri*; Fungi: Uma – *Ustilago maydis*, Cne – *Cryptococcus neoformans*, Spo – *Schizosaccharomyces pombe*, Afu – *Aspergillus fumigatus*, Fgr – *Fusarium graminearum*, Mor – *Magnaporthe oryzae*, Ncr – *Neurospora crassa*, Yli – *Yarrowia lipolytica*, Dha – *Debaryomyces hanseni*, Ago – *Ashbya gossypii*, Kla – *Kluyveromyces lactis*, Cgl – *Candida glabrata*, Sce – *Saccharomyces cerevisiae*; Metazoans: Cbr – *Caenorhabditis briggsae*, Cel – *Caenorhabditis elegans*, Cin – *Ciona intestinalis*, Hsa – *Homo sapiens*, Mmu – *Mus musculus*, Dme – *Drosophila melanogaster*, Ame – *Apis mellifera*, Mbr – *Monosiga brevicollis*; Parabasalid: Tva – *Trichomonas vaginalis*; Apicomplexans: Tpa – *Theileria parva*, Cho – *Cryptosporidium hominis*, Cpa – *Cryptosporidium parvum*, Pfa – *Plasmodium falciparum*, Pbe – *Plasmodium berghei*, Pkn – *Plasmodium knowlesi*, Pvi – *Plasmodium vivax*, Tgo – *Toxoplasma gondii*;

Kinetoplastids: Lbr – *Leishmania braziliensis*, Lin – *Leishmania infantum*, Lma - *Leishmania major*, Tcr – *Trypanosoma cruzi*, Tbr – *Trypanosoma brucei*.

5.1.2 Meios de cultura

Tabela 1. Meios de cultura para bactérias, células HFF ou *T. gondii*

Meios de cultura	Composição
LB	Bacto-triptona 10 g/L, Extrato de levedura 5 g/l, NaCl 5 g/L
LB-ágar	Ágar-ágar 1,5% (m/v) em meio LB
DMEM completo	500 mL de DMEM (Gibco # 12100046), soro fetal bovino 10% (v/v), glutamina 2 mM, gentamicina 20 µg/mL
Meio para congelamento de células HFF e <i>T. gondii</i>	Soro fetal bovino 25% (v/v), DMSO 10% em meio DMEM completo

5.1.3 Microrganismos

T. gondii

Para este trabalho foram utilizadas cepas ou linhagens que estão exemplificadas na tabela a seguir. Para o cultivo dos parasitas, as formas taquizoítas foram mantidas em monocamada de células HFF (*human foreskin fibroblast*) confluyente com meio DMEM completo a 37 °C em estufa com atmosfera de 5 % CO₂ (ROOS *et al.*, 1994).

Tabela 2. Cepas e linhagens de *T. gondii* utilizadas e obtidas neste estudo.

Cepa / Linhagem utilizada	Cassete inserido	Marca de seleção	Linhagem obtida
Cepa RH Δ <i>hxgprt</i> (Tipo I, parental) (DONALD <i>et al.</i> , 1996) “cepa RH”	Cassete DDGFPUAP56	HXGPRT	Linhagem DDGFPUAP56
	Cassete MS2-YFP-DD e cassete Dip13-SL	HXGPRT e CAT	Linhagem MS2
	Cassete NLS-MS2-YFP- DD e cassete Dip13-SL		
Linhagem DiCre <i>ku80::hxgprt</i> Pall, em preparação “linhagem DiCre-HX”	Cassete MS2-Dip13-SL	Seleção negativa para HXGPRT	Linhagem DiCre- MS2
Linhagem DiCre- Δ <i>ku80</i> Pall, em preparação	Cassete cDNA <i>uap56</i> flanqueado por <i>LoxP</i>	HXGPRT	Linhagem cKOuap56

Características das cepas/linhagens:

- **Cepa parental RH Δ *hxgprt* (DONALD *et al.*, 1996):** Os parasitas dessa cepa possuem o gene *hxgprt* nocauteado. Ao longo do texto será citada como “cepa RH”.
- **Linhagem DiCre *ku80::hxgprt* (DiCre-HX):** Os parasitas dessa linhagem possuem o gene *ku80* substituído pelo gene *hxgprt*, bem como DiCre integrada de modo randômico no genoma. Ao longo do texto será citada como “linhagem DiCre-HX”.
- **Linhagem DiCre- Δ *ku80*:** Os parasitas dessa linhagem possuem os genes *hxgprt* e *ku80* nocauteados, bem como DiCre integrada de modo randômico no genoma.

Escherichia coli

Genótipo XL-10 Gold: Tet^rΔ(*mcrA*)183 Δ (*mcrCB-hsdSMR-mrr*)173 *endA1 supE44 thi-1 recA1 gyrA96 relA1 lac* Hte [F' *proAB lac^f ZΔM15 Tn10(Tet^r)Amy Cam^r*]

5.1.4 Soluções e tampões

As soluções e tampões utilizados neste trabalho estão detalhados na tabela a seguir.

Tabela 3. Soluções e Tampões

Soluções e Tampões	Composição
Solução de bloqueio para <i>Western blot</i>	Tampão PBS, Tween 20 0,05% e leite em pó desnatado 5%
Solução de Brometo de Etídio	5,0 µg/mL de brometo de etídio em água destilada
Solução de Ponceau S	Ponceau S (Sigma P-3504) 0,5%; Ácido acético glacial 1%
Solução TBE para eletroforese de DNA (Tris- Ácido Bórico – EDTA)	Tris-base 89 mM, Ácido bórico 89 mM, EDTA 2 mM pH 8.0
Tampão de amostra para eletroforese de DNA 6x	Azul de bromofenol 0,25%; xileno cianol 0,25%; glicerol 30%
Tampão de amostra de proteínas	Tris-HCl 40 mM pH 6,8; SDS 1%, β-mercaptoetanol 2,5%; glicerol 6% e azul de bromofenol 0,005%
Tampão de eletroporação de <i>T. gondii</i> (<i>Cytomix</i>)	K ₂ HPO ₄ /KH ₂ PO ₄ 10 mM, HEPES 25 mM, EGTA 2 mM pH 7.6, KCl 120 mM, CaCl ₂ 0,15 mM; MgCl ₂ 5 mM, KOH 5 mM pH 7.6, ATP 3 mM e GSH 3 mM
Tampão de transferência para <i>Western blot</i>	Tris-base 25 mM, Glicina 192 mM, Metanol 20%
Tampão de corrida para SDS-PAGE	Tris-HCl 0,037 M pH 8.4, Glicina 192 mM pH 8.4, SDS 0,1%
Tampão PBS - solução de uso (<i>Phosphate-buffer saline</i>)	KCl 2,7 mM; KH ₂ PO ₄ 1,5 mM; NaHPO ₄ .7H ₂ O 4,3 mM; NaCl 137 mM
Tampão PBS - TWEEN 20	Tampão PBS, Tween 20 0,05%
Tampão SSC 20X	NaCl 3M e Citrato trissódico 2-hidrato 0,3M
Tampão SSPE 2X	NaCl 300 mM, Tampão fosfato 20 buffer pH 7.4 e EDTA 2 mM
Tampão de hibridação para FISH (TH)	Dextran 10%, SSPE 2X, formamida 35% e tRNA 0,5 mg/ml

5.1.5 Amplificação de DNA ou cDNA por PCR (Reação em cadeia da polimerase - *Polymerase Chain Reaction*) e subclonagem do produto em vetor pGEM-Teasy®

Para a amplificação dos genes a serem estudados em *T. gondii*, a reação foi feita em volume final de 15 µL contendo 100 ng de cDNA de *T. gondii* ou DNA

genômico dos parasitas, 37,5 U de Platinum *Taq* DNA Polimerase *High Fidelity* (Invitrogen #11304-011), 200 μ M de cada dNTP, 10 pmol de cada *primer* F e R, tampão para a enzima Platinum *Taq* DNA Polimerase *High Fidelity* e água deionizada. Para esta reação foi utilizada a ciclagem: desnaturação a 95 °C por 2 minutos seguido de 30 ciclos de 30 segundos a 95 °C (desnaturação), 30 segundos de temperatura de anelamento (específico para cada par de *primer* utilizado) e 2 a 3 minutos a 68 °C (extensão), utilizando a relação de 1 minuto de tempo de extensão para cada 1 kb de DNA a ser amplificado. Ao final da ciclagem foi adicionada uma etapa de extensão de 68 °C por 10 minutos, que permite a adição de adeninas à extremidade 3' do produto da amplificação necessário para subclonagem em vetor pGEM® – Teasy (Promega #A1360), ilustrado na figura 12. A ligação do produto de PCR com o vetor foi realizada conforme as recomendações do fabricante.

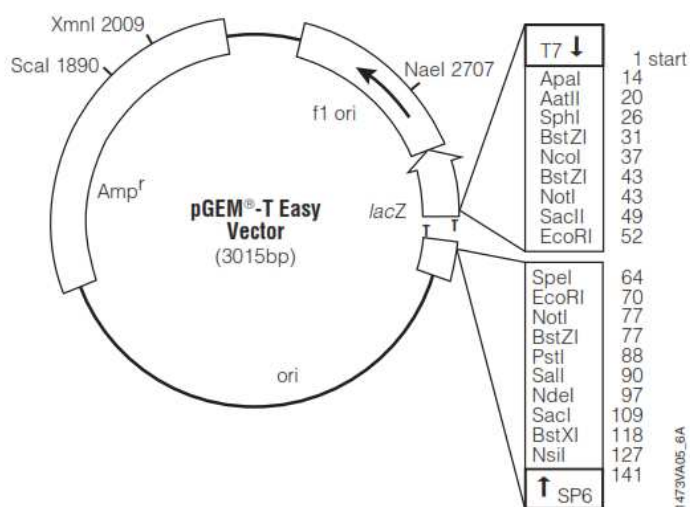


Figura 12. Mapa do vetor pGEM® – Teasy utilizado para as etapas de subclonagens. T7: Região de iniciação da transcrição por T7 RNA Polimerase. 14-141: Sítio múltiplo de clonagem. SP6: Promotor da SP6 RNA Polimerase. LacZ: Códon inicial de LacZ. Amp^r: Gene de resistência à ampicilina. F1 ori: origem de replicação de bacteriófago filamentosos (FONTE: Manual Promega).

A reação de ligação foi utilizada para transformar bactérias *E. coli* quimio-competentes, como descrito no item 5.1.7. As células transformadas foram semeadas em meio LB-ágar acrescido de 100 μ g/mL de ampicilina, IPTG 0,4 mM e X-gal 0,04 mg/mL e incubadas a 37 °C por 12 a 16 horas. Neste caso, a seleção dos clones positivos foi feita pela análise da cor das colônias, devido a presença ou ausência da expressão da enzima β -galactosidase, cuja expressão é induzida por

IPTG havendo degradação do substrato X-gal produzindo um substrato azul. Caso o fragmento seja incorporado ao vetor, a enzima β -galactosidase não é expressa e as colônias permanecem brancas, facilitando a identificação dos clones positivos, que posteriormente foram sequenciados.

Após o período de incubação, foram observadas colônias azuis (não recombinantes) e brancas (recombinantes). Algumas colônias brancas foram cultivadas em meio LB/Amp para purificação dos plasmídeos. Os plasmídeos foram purificados usando o kit Qiaprep® *Spin Miniprep Kit* (QIAGEN #27106), conforme o manual do fabricante. Os fragmentos clonados foram então analisados por sequenciamento de DNA, utilizando os *primers* específicos para cada caso.

5.1.6 Reação de Ligação de DNA (fragmento e vetor, previamente digeridos com as mesmas enzimas de restrição)

Para a reação de ligação foi utilizada a relação de 1:10 pmol de vetor e fragmento, digeridos previamente com as enzimas de restrição de interesse de acordo com o manual do fabricante (New England Biolabs). A reação ocorreu na presença de 1 U de T4 DNA ligase (Invitrogen #15224-025) em um volume final de 10 μ L. Esta reação foi incubada a 16 °C durante 18 horas e posteriormente utilizada para transformar *E. coli* XL-10 quimio-competentes, como descrito no item 5.1.7. As bactérias transformadas foram cultivadas em LB contendo 100 μ g/mL de ampicilina. Após confirmação por PCR de colônia utilizando os *primers* específicos, conforme item 5.1.8., os plasmídeos foram purificados dessas colônias por Qiaprep® *Spin Miniprep Kit* (QIAGEN #27106), conforme as recomendações do fabricante, para a análise de sequenciamento utilizando os *primers* específicos para cada caso.

5.1.7 Preparação e transformação de bactérias quimio-competentes

Para o preparo de bactérias quimio-competentes, foi utilizado o método de CaCl_2 descrito por Sambrook e Russell (2001). Uma colônia de *E. coli* da cepa desejada foi inoculada em 5,0 mL de meio LB contendo antibiótico apropriado. A cultura foi incubada a 37 °C por 18 horas sob agitação constante de 220 rpm. Um volume de 1 mL desta cultura foi transferido para 100 mL de meio LB. As células foram incubadas a 37 °C sob as mesmas condições de agitação até alcançar a fase

de crescimento exponencial (D.O. 600 de 0,6). A cultura foi centrifugada a 4.000 x g por 5 minutos a 4 °C e as células foram suspensas em 50 mL (metade do volume original) de solução gelada de CaCl₂ 100 mM tamponada com HEPES 10 mM pH 7,0 e mantidas no gelo durante 30 minutos. A suspensão foi submetida à centrifugação nas mesmas condições anteriores, as células foram suspensas em 2,0 mL (1/50 do volume da cultura original) da solução anterior acrescida de 10% de glicerol, mantidas no gelo seco por 2 horas e em seguida aliquoteadas e armazenadas a -70 °C.

Reações de recombinação ou ligação foram incubadas com 100 µL da suspensão de *E. coli* quimio-competentes por 30 minutos no gelo. Após esse período, as células foram submetidas ao choque térmico pela incubação a 42 °C por 2 minutos, seguido de incubação de 2 minutos no gelo e posterior adição de 1 mL de meio LB para incubação sob agitação constante de 220 rpm a 37 °C por uma hora. Alíquotas de 100 µL e 900 µL foram espalhadas em meio LB-ágar acrescido de antibiótico (meio seletivo) de acordo com a resistência conferida pelo vetor utilizado na transformação de bactérias e os meios com as bactérias foram incubados a 37 °C por 18 horas. As colônias cultivadas nessas condições foram analisadas por PCR de colônia, conforme descrito no item 5.1.8. No caso de clonagem em pGEM® – Teasy (Promega), o meio LB-ágar foi acrescido de 100 µg/mL de ampicilina, 0,04 mg/mL de X-gal e 0,4 mM de IPTG.

5.1.8 Confirmação de clonagens por PCR de colônia

Após o crescimento em meio seletivo, cada colônia foi coletada e misturada em 10 µL de água deionizada, sendo utilizado 1 µL dessa suspensão para cada reação de PCR. Foram acrescentados aos tubos de PCR 1,5 mM de MgCl₂, 2,5 U *Taq* DNA Polimerase (Invitrogen # 18038-042), 200 µM de cada dNTP, 10 pmol de cada *primer* F e R de interesse, tampão para a enzima *Taq* e água deionizada para volume final de 20 µL. Foi utilizada a ciclagem: *holding* de 94 °C por 10 minutos seguidos de 30 ciclos de amplificação. As temperaturas, tempo de anelamento e extensão foram padronizados para cada caso devido às características dos *primers* e tamanho do fragmento a ser amplificado. No caso do tempo extensão foi utilizada a relação de 1 minuto de tempo de extensão para cada 1 kb de DNA.

5.1.9 Imunofluorescência para a localização subcelular de proteínas por microscopia de fluorescência e deconvolução

Para a visualização da localização de proteína endógena ou de proteína fusionada a etiquetas foi utilizado o protocolo de imunofluorescência de acordo com Andenmatten e colaboradores, 2013 (ANDENMATTEN *et al.*, 2013). A etapa de fixação das proteínas foi realizada com PBS e paraformaldeído (PFA) 4% por 20 minutos. Após essa etapa os parasitas foram lavados com PBS e então foi adicionada uma solução de Triton X-100 0,1% durante 90 segundos para a permeabilização destes. Posteriormente, as lamínulas foram lavadas por imersão em PBS e bloqueadas com PBS contendo BSA 4% por 16 horas a 4 °C em câmara úmida. Após o período de bloqueio, as lamínulas foram lavadas e incubadas por uma hora a 37 °C com anticorpo primário produzido em coelho ou camundongo, indicados nas tabelas 4 e 5, diluído em PBS-BSA 4%. Em seguida, as lâminas foram lavadas com PBS e incubadas por uma hora a 37 °C com o anticorpo secundário anti-IgG de coelho ou anti-IgG de camundongo conjugado com Alexa 594 ou Alexa 488 (Tabela 5) diluído 1:600 em PBS-BSA 4%. Para a coloração da cromatina as lamínulas foram incubadas com 10 µg/mL de DAPI em PBS a temperatura ambiente por 5 minutos seguido de lavagens com PBS. Foi adicionado n-propil-galato ou VectaShield como meio de montagem à superfície de lâminas de vidro e em seguida as lamínulas foram vertidas e seladas. As imagens bidimensionais foram obtidas no microscópio de fluorescência Zeiss AxioScope 2 (*mot plus*) equipado com objetiva de 63X ou 100X de imersão a óleo e capturadas com câmera AxioCam MRm CCD, usando o *software* Volocity. As imagens foram processadas com o programa ImageJ 1.34r software ou Photoshop CS6 (Adobe Systems Inc). Imagens tridimensionais (séries Z) foram adquiridas em um microscópio de deconvolução Leica TCS SP2 AOBS.

5.1.10 Ensaio de *Western blot* (TOWBIN, STAEHELIN e GORDON, 1992) e revelação por quimioluminescência

Para verificar a presença de proteínas endógenas ou fusionadas a etiquetas em extrato protéico dos parasitas foram utilizados anticorpos específicos em ensaio de *Western blot*, indicados nas tabelas 4 e 5. As proteínas separadas em gel de poliacrilamida (SDS-PAGE) foram transferidas para a membrana de nitrocelulose

através de corrente elétrica de 20 V por 16 horas a 4 °C em tampão de transferência para *Western blot*. A membrana foi então corada com solução de Ponceau S para a verificação da transferência, posteriormente descorada em água bidestilada e incubada em solução de bloqueio por uma hora com leve agitação a temperatura ambiente. Em seguida, foi adicionado anticorpo primário específico na diluição desejada, conforme tabela 4, e a incubação foi mantida por uma hora a 37 °C com leve agitação. A membrana foi lavada três vezes por 5 minutos com PBS-Tween 0,1% e incubada por 1 hora a 37 °C com leve agitação com anticorpo secundário anti-IgG conjugado com a enzima peroxidase, indicado na tabela 5, em PBS-Tween 0,1% e após o período de incubação, foi novamente foram realizadas 3 lavagens de 15 minutos com PBS-Tween 0,1%. A revelação é feita através de sensibilização de filmes de raio-X. O produto da reação enzimática foi detectado por quimioluminescência, através do kit ECL *Western blot* (Amersham Biosciences # RPN2232) seguindo as instruções do fabricante. A exposição foi feita em filme Kodak durante 1 minuto e a revelação foi obtida na ausência de luz, em solução de revelação seguido de lavagem por 2 minutos em água e 2 minutos em solução de fixação. Os tempos de revelação foram padronizados para cada experimento.

Tabela 4. Anticorpos primários e as diluições utilizadas nesse estudo. Coelho/camundongo: animal em que o anticorpo foi produzido. Entre parênteses está a fonte comercial dos anticorpos. IF: Imunofluorescência.

Anticorpos	Coelho /camundongo	Diluição IF	Diluição <i>Western blot</i>
Anti-digoxigenina (Sigma # D8156)	camundongo	1:600	
Anti-TcSub2 (SERPELONI <i>et al.</i> , 2011)	Coelho	1:1000	1:1000
Anti-GFP (Roche # 11841460001)	camundongo	1:1000	1:2000
Anti-aldolase (STARNES <i>et al.</i> , 2006)	Coelho	1:2000	1:10000
Anti-TgIMC1 (MANN e BECKERS, 2001)	coelho	1:1500	1:2000
Anti-Ty (BASTIN <i>et al.</i> , 1996)	camundongo	1:10	

Tabela 5. Anticorpos secundários e as diluições utilizadas nesse estudo. Coelho/camundongo: reconhecimento de imunoglobulinas (IgG) de coelho ou camundongo. Entre parênteses está a fonte comercial dos anticorpos. IF: Imunofluorescência.

Anticorpos	Coelho/camundongo	Diluição IF	Diluição Western blot
Alexa Fluor® Goat anti-<i>mouse</i> IgG (H+L) (Invitrogen # A11005)	camundongo	1:600	
Alexa Fluor® 488 Goat anti-<i>mouse</i> IgG (H+L) (Invitrogen # A-11001)	camundongo	1:600	
Alexa Fluor® Goat anti-<i>rabbit</i> IgG (H+L) (Invitrogen # A11012)	coelho	1:600	
ECL Mouse IgG, HRP-linked whole Ab (from sheep #NA931)	camundongo		1:10.000
ECL Rabbit IgG, HRP-linked whole Ab (from donkey # NA934)	coelho		1:10.000

5.1.11 Linhagem de células de mamíferos HFF (*human foreskin fibroblast*)

Formas taquizoítas de *T. gondii* foram cultivadas em células HFF (LGC STANDARDS (UK) # SCRC-1041). São células de cultivo primário que possuem crescimento limitado de até 25 passagens, devido à inibição de crescimento por contato. Estas células foram utilizadas para manutenção, transfecção, subclonagem e caracterização fenotípica dos parasitas.

5.1.12 Cultivo e expansão de células HFF

Para o cultivo de células HFF, estas precisam ser desaderidas através da incubação com tripsina/EDTA (Gibco # 25300-054) para o descolamento das camadas de células do fundo da garrafa de cultura. Antes do tratamento com tripsina, o meio de cultura foi removido e as células foram lavadas com PBS estéril, para a remoção do soro fetal bovino, que bloqueia a atividade enzimática da tripsina. As células foram incubadas com uma solução 1x de tripsina/EDTA e incubadas por 3 minutos a 37 °C. Após esta incubação, foi adicionado meio de cultura *Dulbecco's modified Eagle's medium* (DMEM) completo para bloquear a atividade enzimática da

tripsina e as células em suspensão foram transferidas para novas garrafas de cultura ou placas de 6 poços contendo meio DMEM completo.

5.1.13 Cultivo de *T. gondii* em linhagens de células de mamíferos (ROOS *et al.*, 1994)

As formas taquizoítas de *T. gondii* foram cultivadas em monocamada de células HFF e mantidas em meio DMEM, suplementado com 10% de soro fetal bovino, 2 mM de glutamina e 25 µg/mL de gentamicina (DMEM completo). Os parasitas recentemente egressos (até 72 horas depois da lise dos fibroblastos) foram inoculados em uma monocamada confluyente de células HFF. Alternativamente, parasitas intracelulares foram artificialmente liberados das células-hospedeiras por lise destas com seringa de 26G e rapidamente inoculados em uma nova monocamada confluyente de células HFF.

5.1.14 Congelamento de *T. gondii*

Para o congelamento, após 36 horas de inoculação em monocamada confluyente de células HFF, o sobrenadante foi trocado por meio DMEM completo e as células infectadas foram descoladas do fundo da garrafa de cultura por ação mecânica e transferidas para um criotubo. A esse criotubo foram adicionados também 900 µL de meio de congelamento (Tabela 1) e transferidos para o freezer de -20 °C por 16 horas e então para freezer de -80 °C, onde podem ser estocados por várias semanas. No caso de um longo período de estocagem, os tubos foram transferidos para o nitrogênio líquido.

Para o congelamento das células HFF, uma garrafa de cultura de 75 cm² com monocamada confluyente foi tripsinizada e as células em suspensão foram centrifugadas a 200 x *g* por 5 minutos. O sedimento celular (*pellet*) foi suspenso em 900 µL de DMEM completo, acrescido de 900 µL de meio de congelamento (Tabela 1) e transferido para o criotubo para congelamento a -80 °C.

Para o descongelamento de células HFF e parasitas, o criotubo contendo as células foi incubado a 37 °C e após descongelamento, a cultura foi suspensa em 4 mL de DMEM completo, previamente incubado também a 37 °C. Para a remoção do DMSO, componente do meio de congelamento, as células foram centrifugadas a 600

x g por 5 minutos e o sobrenadante foi substituído por 4 mL de meio DMEM completo. Essa suspensão foi então despejada numa garrafa de cultura. No caso de parasitas, as células descongeladas foram inoculadas em uma garrafa contendo monocamada confluyente de células HFF, permitindo assim a reinfecção e a manutenção dos mesmos.

5.1.15 Transfecção de *T. gondii*

A eletroporação é a maneira mais eficiente de introduzir o DNA no parasita (SOLDATI e BOOTHROYD, 1993). Para isto, os parasitas são suspensos em *Cytomix* (Tabela 3), o qual mimetiza a composição iônica da fração citosólica de células HFF, permitindo bons resultados de sobrevivência dos parasitas (VAN DEN HOFF, MOORMAN e LAMERS, 1992). Aproximadamente 1×10^7 parasitas recentemente egressos das células HFF ou liberados mecanicamente com seringa de 26G foram coletados por centrifugação a 600 x g por 5 minutos em temperatura ambiente, lavados com 4 mL de *Cytomix* e então suspensos em 700 μ L de *Cytomix* acrescido de 24 μ L de ATP 100 mM e 24 μ L de GSH 100 mM e o DNA de interesse, circular ou linearizado. Essa mistura foi transferida para uma cubeta de eletroporação de 4 mm e a eletroporação ocorreu com dois pulsos de 1,7 kV de 200 μ s cada um, com intervalo de 5 ms entre eles. Imediatamente após a eletroporação, os parasitas foram inoculados em uma monocamada de células HFF e após 24 horas foi adicionada a droga para a seleção positiva de HXGPRT, conforme Tabela 6.

Tabela 6. Estratégias de seleção para obtenção de linhagens de parasitas transfectantes de *T. gondii*

Gene marcador de seleção	Droga	Procedimento para seleção dos parasitas
HXGPRT de <i>T. gondii</i> (DONALD <i>et al.</i>, 1996)	Seleção positiva: Ácido micofenólico (25 µg/ml) e xantina (40 µg/mL) Seleção negativa: 6-thioxantina (340 µg/ml)	Droga diluída em DMEM completo foi adicionada 16 a 24 horas depois da eletroporação. A seleção ocorreu após morte dos parasitas não transfectados e egresso da população estável de parasitas.
<i>Chloramphenical acyl transferase (CAT) de <i>E. coli</i></i>	Cloranfenicol (10 µM)	Droga diluída em DMEM completo foi adicionada 16 a 24 horas depois da eletroporação O efeito dessa droga é mais demorado (de 3 a 4 passagens) e portanto é importante realizar as passagens com 5×10^6 a cada dois dias até a morte dos parasitas não transfectados e egresso da população estável de parasitas.

5.1.16 Obtenção de linhagem clonal de parasitas por diluição limitante

Para a obtenção da linhagem clonal do parasita transfectante estável foi utilizada uma placa de 96 poços contendo monocamada de células HFF confluyente. Em cada poço da primeira coluna foram inoculados 50 µL de parasitas selecionados da população estável, egressos recentemente. A partir dessa coluna, foram feitas diluições seriadas de 1:4 para as colunas adjacentes, com pipetador multicanal. Depois de 7 a 14 dias de incubação em condições normais de crescimento (37 °C, 5 % CO₂) as culturas foram inspecionadas em busca dos poços com apenas uma placa de lise originada de apenas um parasita. Esses parasitas foram então selecionados e cultivados em cultura para as análises fenotípicas posteriores.

5.1.17 Isolamento de DNA genômico de *T. gondii*.

O kit *DNeasy Blood & Tissue kit* (Qiagen # 69506) foi utilizado para a extração de DNA genômico de parasitas intracelulares ou extracelulares. No caso de

parasitas intracelulares, células HFF infectadas foram lisadas mecanicamente utilizando uma seringa de 26G para a liberação dos parasitas. Após a obtenção de taquizoítas frescos (liberados mecanicamente ou recém egressos de culturas de placa de Petri de 6 cm), estes foram separados das células HFF por centrifugação por 5 minutos a 600 x g, o *pellet* foi suspenso em 200 µL de PBS e o DNA genômico foi isolado conforme o manual do fabricante.

5.1.18 Obtenção de cDNA de *T. gondii*

Formas taquizoítas de *T. gondii* (3×10^7 parasitas) recentemente egressos das células HFF ou liberados, por ação mecânica, de células HFF infectadas, foram utilizados para a extração de RNA total. No caso de parasitas ainda intracelulares, primeiramente foi feita a liberação mecânica de células HFF infectadas, a partir das garrafas de cultura, com o auxílio de um raspador de células estéril (*cell scraper*), seguido de liberação dos parasitas com seringa de 26G. Os parasitas foram coletados por centrifugação a 600 x g por 5 minutos em temperatura ambiente e lavados duas vezes com PBS. O RNA das amostras foi extraído com o kit RNeasy (Qiagen # 74106), conforme o manual do fabricante. O RNA foi tratado com RQ1 RNase-Free DNase (Promega # M6101) para eliminar a contaminação por DNA antes de ser utilizado para a síntese do cDNA. O cDNA foi sintetizado usando-se *random primers* (Invitrogen # 48190-011) ou oligodT (Invitrogen # AM5730G) e o sistema *ImProm-IITM Reverse Transcriptase* (Promega # A3800), como descrito no manual do fabricante. O cDNA foi utilizado nas análises de processamento de mRNAs por PCR, conforme item 5.1.5.

5.1.19 Ensaios de placa de lise (*plaque assay*)

Aproximadamente 200 parasitas foram inoculados em uma monocamada confluyente de células HFF presente em um poço de uma placa de 6 poços. A placa foi incubada por 108 horas (4 ½ dias), a 37 °C em estufa com atmosfera de 5 % CO₂. Após o tempo de incubação, a monocamada de células HFF infectadas foi lavada com PBS e fixada em metanol gelado por 10 minutos. As células foram então coradas com Giemsa (diluição de 1:10 em água destilada) por 45 minutos e lavadas 3 vezes com PBS com intervalos de 5 minutos. As imagens foram obtidas no

microscópio de fluorescência Axiovert 40 CFL com a objetiva de 5X e câmera AxioCam ICc1.

5.2 Caracterização funcional de TgUAP56 por abordagens de genética reversa

5.2.1 Superexpressão de dd-GFP-TgUAP56 em *T. gondii* pelo sistema DD

A proteína TgUAP56 foi fusionada a GFP e ao domínio de desestabilização (ddFKBP12- DD) na porção N-terminal em vetor de expressão para ensaios de localização subcelular por fluorescência bem como para a observação do fenótipo após diferentes tempos de incubação com Shield-1 (Shld-1) (Clontech). Como citado anteriormente, Shld-1 se liga ao domínio ddFKBP12 impedindo que a proteína fusionada seja degradada, permitindo controlar a expressão da proteína de interesse. Nesse sistema, a quantidade de proteína estabilizada aumenta a medida que se adiciona mais Shld-1 ao longo do tempo de incubação.

Para a clonagem em vetor de expressão da proteína fusionada a GFP e DD em *T. gondii*, o cDNA referente ao gene *uap56* (TGME49_216860 - 1305 bp) foi amplificado por reação de PCR, conforme item 5.1.5 utilizando como molde o cDNA do parasita. Os *primers* utilizados para esta amplificação podem estar mostrados na tabela 7.

Tabela 7. Primers utilizados para amplificar região de *uap56* a partir de cDNA para clonagem em vetor de superexpressão em *T. gondii*. As enzimas utilizadas estão indicadas no nome do *primer*, do lado esquerdo. As sequências reconhecidas por cada enzima estão indicadas em vermelho, à direita. *Forward* (F). *Reverse* (R).

Nome	Sequência (5' → 3')
<i>uap56</i> - <i>Pst</i> II (F)	CCCCTG CAG GAATGACGACTCTCGAGCAGAACC
<i>uap56</i> - <i>Pac</i> I (R)	CCCTTA AATTA ATCACTGATTGATGTACTGCGAGGCG

No vetor utilizado para a superexpressão da proteína dd-GFP-TgUAP56, a expressão dos genes está sob controle do promotor P5RT70, que é uma fusão entre cinco repetições do elemento 5'UTR de *sag1* (que age como realçador) e o promotor mínimo de *α -tubulina* (SOLDATI e BOOTHROYD, 1995). Este vetor já possuía as sequências ddFKB12, etiqueta myc e GFP clonadas e o cDNA foi clonado na região *downstream* e em fase com a sequência de GFP. Adicionalmente, logo após a sequência do cDNA clonado encontra-se a região 3'UTR correspondente ao gene constitutivo *dhfr*, para conferir estabilidade ao respectivo mRNA de *uap56*.

A reação de amplificação do cDNA de *uap56*, subclonagem em pGEM-Teasy® e transformação de bactérias quimio-competentes foi realizada conforme itens 5.1.5 e 5.1.7. Após a confirmação do sequenciamento do fragmento referente ao cDNA de *uap56* no clone pGEM-uap56, este foi liberado através da reação enzimática deste plasmídeo com as enzimas de restrição *Pst*I e *Pac*I, de acordo com o manual do fabricante (New England Biolabs). A amostra foi submetida a eletroforese em gel de agarose 1% para a excisão da região correspondente ao fragmento liberado. A purificação do fragmento em gel de agarose foi realizada com auxílio do kit *High Pure PCR Product Purification Kit* (Roche # 11732676001) O vetor de destino (*p5RT70ddmycGFPPMyoAtailTy-HX* (HETTMANN C *et al.*, 2000), Figura 13), também foi digerido nas mesmas condições e a ligação do fragmento ao vetor foi realizada conforme item 5.1.6. A clonagem do fragmento foi verificada por PCR de colônia conforme item 5.1.8, contendo os reagentes da reação de PCR bem como 10 pmol de cada *primer* F e R (Tabela 7). Foi utilizada a ciclagem: desnaturação a 95 °C por 2 minutos seguido de 35 ciclos de 30 segundos a 95 °C (desnaturação), 30 segundos de 56 °C (anelamento) e 2 minutos a 72 °C (extensão). Os produtos da PCR foram analisados em gel de agarose 1%.

Após a confirmação da clonagem, os plasmídeos das colônias referentes aos clones corretos foram purificados por Qiaprep® *Spin Miniprep Kit* (QIAGEN #27106), conforme o manual do fabricante e então enviados para sequenciamento para confirmar a fase aberta de leitura. Um clone sem erros na sequência de *uap56* encontrado foi então escolhido para preparação de plasmídeo em maior escala. A quantidade de 60 µg do vetor foi linearizado com a enzima *Kpn*I, purificado e utilizado para a posterior transfecção de parasitas da cepa RH de *T. gondii* (DONALD *et al.*, 1996), segundo o item 5.1.15. A seleção dos parasitas

transfectantes foi positiva para HXGPRT, conforme tabela 6. Após a obtenção de população selecionada, foi obtida a linhagem clonal DDGFPUAP56 de acordo com item 5.1.16.

A construção do vetor está esquematizada na figura 13.

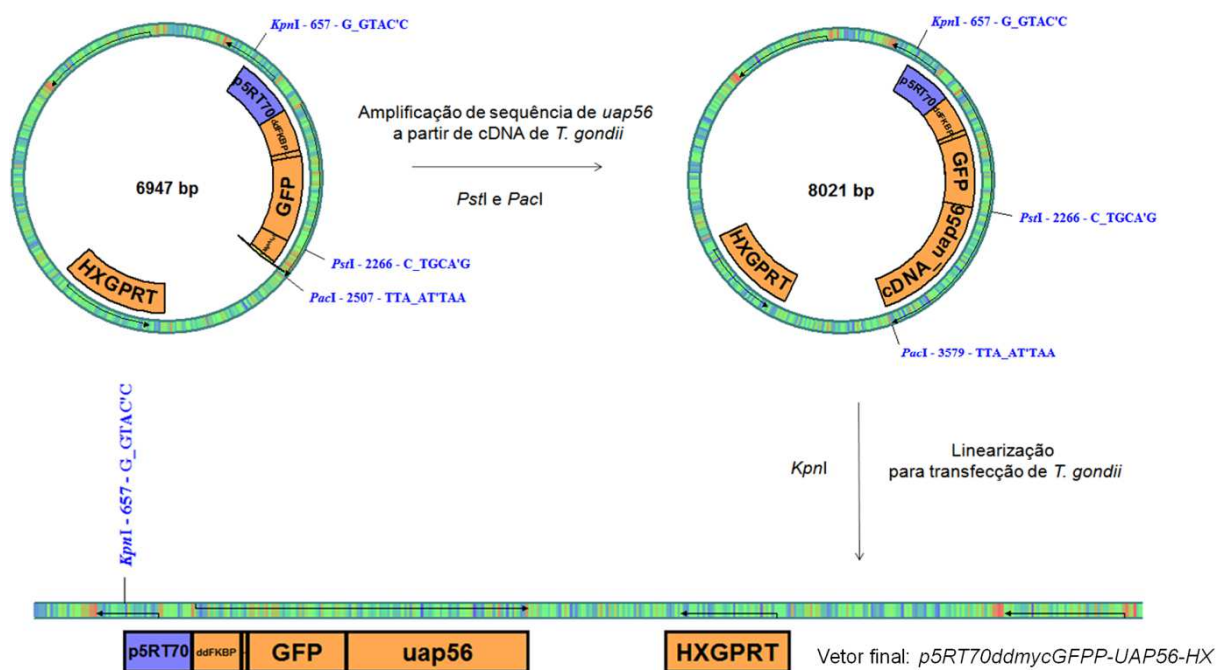


Figura 13. Estratégia para a construção do vetor de superexpressão de dd-GFP-TgUAP56 em *T. gondii*. O vetor *p5RT70ddmycGFPPfMyoAtailTy-HX* foi utilizado para a clonagem do cDNA referente ao gene *uap56* com as enzimas *PstI* e *PacI*. O vetor final, *p5RT70ddmycGFPPUAP56-HX*, foi linearizado com a enzima *KpnI* para integração randômica no genoma do parasita. A linhagem obtida foi nomeada de DDGFPUAP56.

5.2.1.1 Expressão e localização subcelular de proteína dd-GFP-TgUAP56 em *T. gondii*

Para a verificação da localização subcelular da proteína fusionada dd-GFP-TgUAP56 nas formas taquizoítas dos parasitas transfectantes foi necessária a inoculação prévia de células HFF em lamínulas redondas contidas em placas de 24 poços. Após a obtenção de monocamada confluyente de células HFF, estas foram infectadas com os parasitas e, dois dias após a infecção, foi então adicionado ao meio de cultura DMEM completo acrescido de 2 μ M de Shld-1 para incubação durante os tempos desejados. Para a fixação das células o meio de cultura foi

retirado para a fixação posterior por paraformaldeído 4% em PBS, com incubação de 20 minutos. As células foram lavadas com PBS e então foi adicionado DAPI Fluoromount (Southern Biotech # 0100-20) como meio de montagem na superfície das lâminas de vidro e em seguida as lamínulas foram vertidas e seladas.

Para os ensaios de imunofluorescência de acordo com protocolo citado no item 5.1.9, foi utilizado o anticorpo policlonal anti-TcSub2 produzido em coelho (Tabela 4). Este anticorpo reconhece TgUAP56 especificamente em extrato protéico de *T. gondii*. Neste caso foi utilizado o anticorpo secundário anti-IgG de coelho Alexa 594 (Tabela 5) para a visualização indireta da proteína TgUAP56 em comprimento de onda de 594 nm. A proteína fusionada dd-GFP-TgUAP56 foi visualizada em comprimento de onda de 488 nm.

5.2.2 Nocaute condicionado do gene *uap56* em *T. gondii* pelo sistema DiCre

Para a análise da distribuição de mRNAs na ausência de proteína TgUAP56 foi utilizado o sistema de nocaute condicionado baseado na recombinase DiCre, que permite a excisão de DNA flanqueado por sequências *LoxP* após a dimerização desta enzima por indução de rapamicina, como explicado no item 1.2.1.

Para isso, o gene *uap56* do genoma do parasita foi substituído por um cassete contendo o respectivo cDNA de *uap56*, flanqueado por sequências *LoxP*. Essa substituição ocorreu por recombinação homóloga nas regiões 5' e 3' UTRs, também presentes no cassete, e foram inseridas na construção obtida como descrito a seguir.

As sequências *LoxP* foram adicionadas nos *primers reverse* de 5' UTR e *uap56*, para a inserção das sequências após amplificação do gene por PCR. As sequências destes *primers* podem ser visualizadas na tabela 8.

Tabela 8. Primers utilizados para amplificar cDNA de *uap56* e regiões não-codificantes 5' e 3'UTRs para as clonagens em vetor de nocaute condicionado de *uap56*. As enzimas utilizadas estão indicadas no nome do oligonucleotídeo do lado esquerdo. As sequências reconhecidas por cada enzima estão indicadas em vermelho à direita. As sequências *LoxP* contidas nos primers estão sublinhadas. *Forward* (F). *Reverse* (R).

Nome	Sequência (5' → 3')
5UTR- <i>KpnI</i> (F)	GGGGTACCCTGCTTGAGGTCTTGCTTGC
5UTR- <i>Loxp-ApaI</i> (R)	GGGGGCCCATAACTTCGTATAGCATACATTATACGAAGTTATTTGTCA ACGCGGTTGAGGG
<i>uap56-ApaI</i> (F)	GGGGCCCATGACGACTCTCGAGCAGAAC
<i>uap56-LoxP-NsiI</i> (R)	GCTAATGCATTTTATAACTTCGTATAGCATACATTATACGAAGTTATTTA CTGATTGATGTACTGCGAGG
3UTR- <i>SacI</i> (F)	GGGAGCTCAGTCGAACACAGTAAACGTTTCTTG
3UTR- <i>SacI</i> (R)	GGGAGCTCCCTAGGAAGGCGTCGTGGATCCTTATT

Para as clonagens, o cDNA de *uap56* (TGME49_216860 - 1305 bp) foi amplificado por reação de PCR a partir de cDNA de *T. gondii*, conforme item 5.1.5. As regiões 5' e 3' UTRs referentes a este gene (tamanhos: 2200 e 1620 bp, respectivamente) foram amplificadas por reação de PCR a partir de DNA genômico de *T. gondii*, conforme item 5.1.5. As regiões amplificadas foram subclonadas em pGEM-Teasy® conforme item 5.1.5 e as clonagens foram confirmadas por PCR de colônia, conforme item 5.1.8, utilizando os *primers* específicos, listados na tabela 8. Para cada etapa, os clones obtidos foram sequenciados com os respectivos *primers* (Tabela 8) e após a confirmação, os fragmentos foram liberados do clone de pGEM-Teasy® após reação enzimática com as respectivas enzimas, indicadas na tabela de *primers* (Tabela 8), de acordo com o manual do fabricante (New England Biolabs). Cada amostra foi submetida individualmente a eletroforese em gel de agarose 1% para a excisão das respectivas regiões liberadas. A purificação dos fragmentos foi realizada com auxílio do kit *High Pure PCR Product Purification Kit* (Roche #11732676001) para ligação ao vetor de destino (*pT8TgMLC1Ty*), também

clivado nas mesmas condições. A ligação ocorreu conforme item 5.1.6 e a estratégia de clonagem pode ser visualizada na figura 14.

É importante ressaltar que a porção final do amplicom referente a 5' UTR termina um nucleotídeo antes do códon inicial do gene *uap56*, garantindo assim a presença do promotor endógeno nesta região clonada.

O vetor de destino escolhido foi o *pT8TgMLC1Ty*, presente no banco de vetores no Laboratório do Dr. Markus Meisser, por possuir o gene repórter que codifica a proteína *mcherry* já clonado. O cDNA de *uap56*, contendo o códon de parada, foi clonado *upstream* a sequência de *mcherry*. Logo após o gene *mcherry* se encontra a região 3'UTR correspondente ao gene *sag1*, que confere estabilidade ao respectivo mRNA de *uap56* ou *mcherry*. Após a excisão deste cDNA, o gene *mcherry* fica sob controle do promotor endógeno de *uap56*, podendo ser usado como indicador positivo de nocaute pela presença de parasitas expressando *mcherry* e conseqüentemente detectados com fluorescência em comprimento de onda do vermelho. As amostras foram visualizadas por microscopia de fluorescência em comprimento de onda de 600 nm.

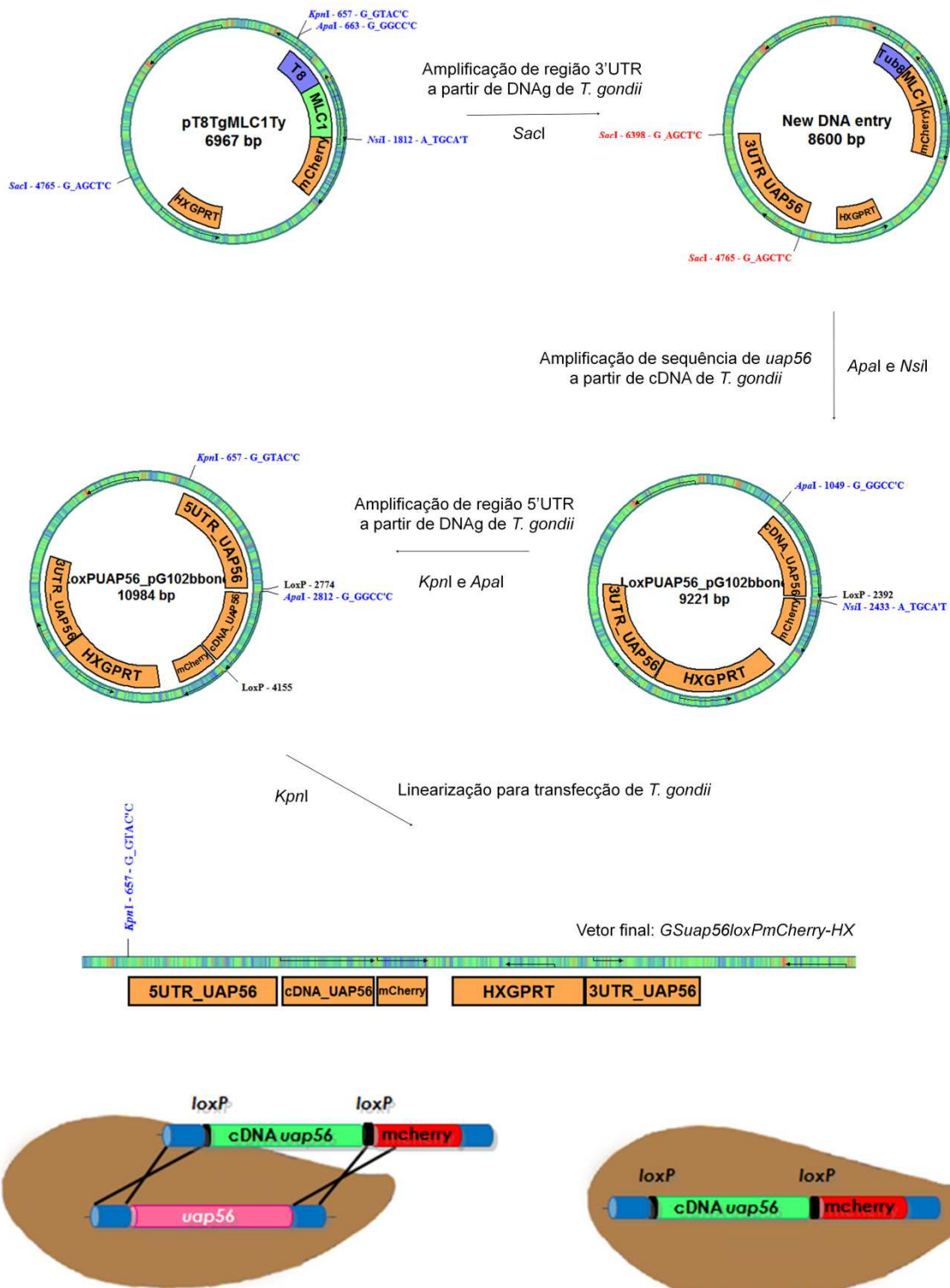


Figura 14. Estratégia de clonagem para a construção do vetor utilizado para nocaute condicionado do gene *uap56* em *T. gondii*. O vetor *pT8TgMLC1Ty* foi utilizado como molde para a construção do vetor final. 5UTR_UAP56: Região 5'UTR correspondente ao gene *uap56*. cDNA_UAP56: Região codificante do gene *uap56*. *mCherry*: gene repórter que emite fluorescência em comprimento de onda de 600 nm. HXGPRT: marca de seleção dos parasitas transfectantes.

3UTR_UAP56: região 3'UTR referente ao gene *uap56*. DNAg: DNA genômico. As regiões foram clonadas com as enzimas de restrição indicadas na figura. O vetor final, *GSuap56LoxPmCherry-HX*, foi linearizado com a enzima *KpnI* para a transfecção dos parasitas e integração no locus de *uap56* no genoma do parasita. GS: *gene-swap*. Na parte de baixo está o esquema de recombinação homóloga entre as regiões 5' e 3'UTR, em azul, relativos ao gene *uap56*. Após a recombinação, o cassete contendo o cDNA de *uap56* flanqueado por *LoxP* é inserido no locus de *uap56*. A linhagem obtida foi nomeada de cKOuap56.

O vetor final foi linearizado com *KpnI* e utilizado para a transfecção de parasitas da linhagem DiCre- $\Delta ku80$ (Tabela 2). A seleção dos parasitas transfectantes estáveis ocorreu através de seleção positiva para HXGPRT, conforme tabela 6. Após a obtenção de população selecionada, foi obtida a linhagem clonal para nocaute condicionado de do gene *uap56*, linhagem cKOuap56, de acordo com item 5.1.16.

5.2.2.1 Análise de integração do cassete para nocaute condicionado no genoma de parasitas da linhagem clonal cKOuap56

A integração do cassete no genoma dos parasitas da linhagem clonal cKOuap56, bem como a confirmação da retirada do gene do genoma foi analisada por PCR. Para isso foram utilizadas amostras de DNA genômicos referentes a 3 culturas, isolados conforme item 5.1.17: cepa RH (Tabela 2) como controle, linhagem cKOuap56 (Tabela 2) sem incubação de rapamicina ou com incubação de 50 nM de rapamicina por 24 horas.

Foram utilizados *primers* que não hibridizam no cassete inserido (mas que hibridizam em regiões presentes no genoma, tanto na região 5' quanto na região 3'UTR relativos ao gene *uap56*) pareados com *primers* que hibridizam em regiões internas ao cassete integrado. Os resultados das ampliações por PCR utilizando este conjunto de *primers* foram avaliados após eletroforese em gel de agarose 1%. Devido à presença de introns no gene *uap56* é possível verificar se o respectivo cDNA, que não possui introns, foi integrado corretamente devido à diferença de tamanho. Estes *primers* relacionados com o gene *uap56* foram utilizados também para verificar a retirada do cDNA flanqueado por sequências *LoxP* após incubação com rapamicina. Os *primers* de *mcherry* e *hxgprt* serviram como complementação da

análise, visto que estão ausentes no genoma do parasita da cepa RH. Os *primers* utilizados estão listados na tabela 9.

Tabela 9. Primers utilizados para confirmação de integração do cassete para nocaute condicionado no locus do gene *uap56* no genoma de *T. gondii*. Ext: Externo ao cassete. Int: Interno ao cassete. *Forward* (F). *Reverse* (R).

Código	Nome	Sequência (5' → 3')
I	5UTR-ext (F)	GTATTAGTACATGGTGTTC AAGCG
II	uap56-int (R)	CACTTGTGTCCAGGTT CAGCT
III	uap56 (F)	ATGACGACTCTCGAGCAGAAC
IV	mcherry-interno (R)	CACCGGCGGCATGGACGAGCTGTACAAG
V	mcherry (F)	ATGCTATACGAAGTTATAAAATGCA
VI	3UTR-ext (R)	CATGTCGGGTTAGTTCATATCGT
VII	hxgprt (F)	GCTCAATACGGCCCTTGCCCTTGC

Cada reação de PCR foi realizada em volume final de 20 µL contendo 100 ng do DNA genômico da cepa/linhagem de *T. gondii* de interesse, 1,5 mM de MgCl₂, 2,5 U *Taq* DNA Polimerase (Invitrogen # 18038-042), 200 µM de cada dNTP, 10 pmol de cada *primer* F e R (Tabela 9), tampão para a enzima *Taq* e água deionizada. Foi utilizada a ciclagem: desnaturação a 95 °C por 2 minutos seguido de 35 ciclos de 30 segundos a 95 °C (desnaturação), 30 segundos de 60 °C (anelamento) e 5 minutos a 72 °C (extensão). Os produtos da PCR foram analisados em gel de agarose 1%.

5.2.2.2 Análise de processamento por *splicing* de mRNAs após nocaute condicionado do gene *uap56*

A abordagem para análise de processamento de mRNAs por *splicing* foi escolhida com base em um trabalho anterior (SUVOROVA *et al.*, 2013). Esta abordagem utiliza método semi-quantitativo por PCR, tendo como alvos alguns mRNAs relativos a genes constitutivamente expressos em *T. gondii* que também foram analisados neste estudo. Para isso, RNA total foi extraído das linhagens DiCre- $\Delta ku80$ (controle) e cKOuap56 conforme item 5.1.18 após os tempos de incubação com rapamicina de 24 e 48 horas. Para o controle, também foi realizado o experimento com parasitas sem a incubação com rapamicina, nos mesmos tempos. Os RNAs foram utilizados para a obtenção de cDNA, como descrito no item 5.1.18. Os *primers* da tabela 10 que foram utilizados para este ensaio foram desenhados com o objetivo de amplificar regiões que possuísem introns, para verificar justamente a diferença de tamanho do produto de PCR em gel de agarose 1% referente a pré-mRNAs ou mRNAs, já processados corretamente. Como referência de tamanho foi utilizado o DNA genômico, o qual possui introns, sendo possível diferenciar os tamanhos entre pré-mRNAs (PS, *pre-spliced*) e mRNAs, já processados (S, *spliced*). Os mRNAs analisados são referentes aos genes: *RNA polymerase II p8.2 subunit* (TGME49_217560), *RNA polymerase II p19 subunit* (TGME49_271300), *RNA polymerase II p23 subunit* (TGME49_240590), IMC1 (TGME49_231640), IMC15 (TGME49_275670) e IMC5 (TGME49_224530). Como controle de qualidade de cada cDNA utilizado para estas PCR foram utilizados *primers* para tubulina (DALMASSO *et al.*, 2009).

Para a reação de PCR foram acrescentados aos tubos os respectivos cDNAs de interesse, 1,5 mM de MgCl₂, 2,5 U *Taq* DNA Polimerase (Invitrogen # 18038-042), 200 μ M de cada dNTP, 10 pmol de cada *primer* F e R (Tabela 10), tampão para a enzima *Taq* e água deionizada. Foi utilizada a ciclagem: *holding* de 94 °C por 1 minuto seguidos de 35 ciclos de amplificação. As temperatura de anelamento foi de 60 °C, por 30 segundos e a extensão ocorreu a 72 °C por 3 minutos. Os produtos da PCR foram analisados em gel de agarose 1%.

Tabela 10. Primers utilizados para análise de processamento de mRNAs após nocaute condicionado de *uap56*

Nome	Sequência (5' →3')
PoIII p8.2 (F)	ATGATCATCCCAATTCGGTGCT
PoIII p8.2 (R)	TCAGCTCTCCACAATGCGCTT
PoIII p19 (F)	GTCCTTGACGCCGTCATCACC
PoIII p19 (R)	TCAAAGCGTCGGAGAATGGGG
PoIII p23(F)	GTCAAAGAGATTCAGCTTCCAAGGATTC
PoIII p23 (R)	GTCACGAGAACTTGCAGTACACG
IMC1 (F)	GAGGTACCTGAGGTTTCAGTACGTC
IMC1 (R)	AAGACCTGCTCCACCTTGGG
IMC15 (F)	GTTCCAAAGGAATTCACGCGCATTG
IMC15 (R)	CTGTTAGCCGTTCTTTATCTAGGTGAGTG
IMC5 (F)	CTGTTAGCCGTTCTTTATCTAGGTGAGTG
IMC5 (R)	CTCCATTTTCGACAGTCAAGAGTCCAG

5.2.3 Monitoramento de mRNAs em *T. gondii* pelo sistema MS2

5.2.3.1 Obtenção de parasitas transfectantes de MS2-Dip13-SL

Para ensaios de detecção e monitoramento de mRNAs *in vivo*, foi estabelecido um sistema baseado no sistema MS2 em *T. gondii* durante o pós-doutorado da minha orientadora Andréa R. Ávila, no laboratório do Dr. Markus Meissner, o qual dei continuidade neste estudo. Os vetores foram construídos para gerar uma ferramenta de detecção de exportação de mRNAs através do rastreamento de mRNA de Dip13 contendo 24 repetições de *stem-loops* na região 3'UTR (Dip13-SL). Essas repetições são reconhecidas pela proteína MS2, que no caso deste estudo, está fusionada a algumas etiquetas: NLS de *T. gondii* (*nuclear localization signal* - sinal de localização nuclear) (BHATTI e SULLIVAN, 2005), YFP (proteína amarela fluorescente) e domínio de desestabilização ddFKBP12, resultando na proteína fusionada NLS-MS2-YFP-DD. O sistema DD, como já citado anteriormente, permite a estabilização de proteínas fusionadas ao domínio ddFKBP12 quando da presença de Shld-1, havendo aumento da quantidade da

proteína NLS-MS2-YFP-DD na célula ao longo do tempo. Logo, o que se espera desse sistema após estabilização da proteína por Shld-1 é:

- Na ausência do mRNA Dip13-SL, a proteína NLS-MS2-YFP-DD deve ser nuclear devido ao sinal NLS.
- Na presença do mRNA Dip13-SL a proteína NLS-MS2-YFP-DD está ligada ao mRNA de Dip13 no citoplasma.
- Após interromper a via de exportação de mRNA, o mRNA Dip13-SL permanece no núcleo, havendo conseqüente acúmulo nuclear da proteína NLS-MS2-YFP-DD modificando a relação de fluorescência da proteína entre núcleo e citoplasma.

A segunda parte do estudo será verificar se o gene *uap56* está relacionado com a exportação de mRNAs em *T. gondii*, assim como em tripanossomatídeos, e assim usá-lo como prova de conceito do sistema MS2. Assim, a hipótese é que após a indução do nocaute do gene *uap56* seria possível observar o acúmulo nuclear da proteína NLS-MS2-YFP-DD devido ao acúmulo nuclear do respectivo mRNA Dip13-SL, como pode ser observado na figura 15.

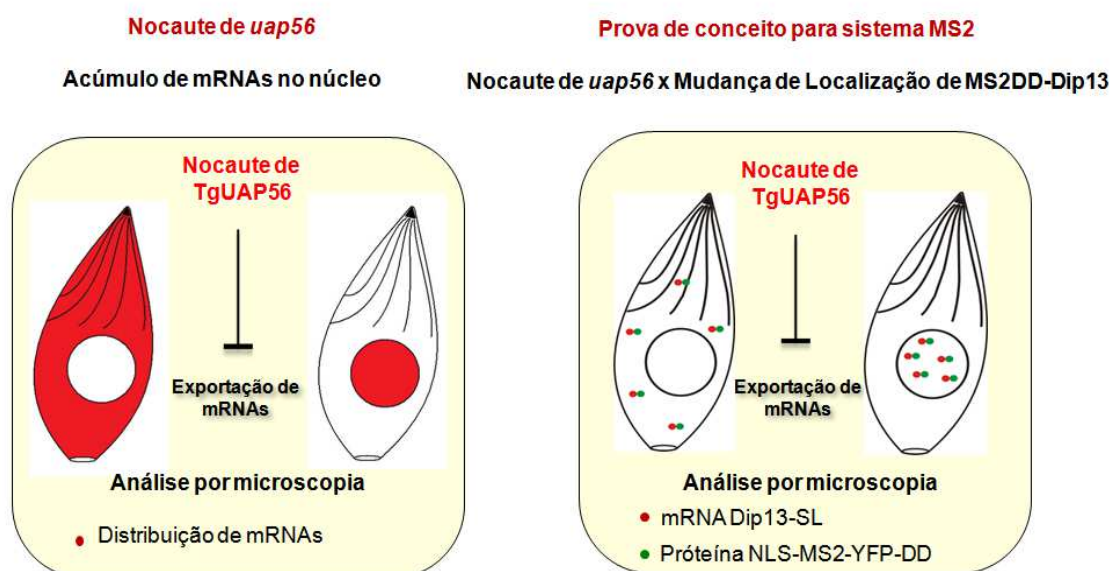


Figura 15. Distribuição de mRNAs esperada após o nocaute condicionado do gene *uap56* de *T. gondii*. Do lado esquerdo está esquematizada a dinâmica de distribuição da maioria de mRNAs poliadenilados. Em uma situação normal, a maioria dos mRNAs possui distribuição citoplasmática. Após o nocaute condicionado do gene *uap56*, espera-se o acúmulo nuclear da maioria dos mRNAs, assim como observado em tripanossomatídeos (SERPELONI *et al.*, 2011). Do lado direito está esquematizada a localização citoplasmática da proteína NLS-MS2-YFP-DD, interagindo com o mRNA

Dip13-SL, transcrito constitutivamente. Após o nocaute condicionado de *uap56* espera-se o bloqueio da exportação deste mRNA, sendo retido no núcleo, assim como a proteína NLS-MS2-YFP-DD, que é visível diretamente por microscopia de fluorescência devido à etiqueta de YFP.

Para verificar o funcionamento do sistema MS2 primeiramente foram construídos dois vetores, ambos contendo promotor tubulina para expressão constitutiva dos genes: um vetor contendo a sequência correspondente do mRNA de Dip13 com 24 repetições de *stem-loops* na região 3'UTR – Vetor Dip13-SL (figura 16 A) e outro para a expressão da proteína NLS-MS2-YFP-DD, estabilizada por Shld-1 – vetor MS2 (figura 16 B). Os vetores foram linearizados e transfectados para integração randômica no genoma dos parasitas da cepa RH (Tabela 2) e a seleção ocorreu conforme tabela 6: seleção positiva para HXGPRT e CAT. A sequência de NLS foi a última a ser clonada e o vetor contendo as sequências para MS2-YFP-DD foi também transfectado em cepa RH (Tabela 2) para análise de localização da proteína sem este sinal de localização nuclear. As linhagens obtidas por Dip13-SL / NLS-MS2-YFP-DD ou Dip13-SL / MS2-YFP-DD foram nomeadas de “linhagem MS2” (Tabela 2).

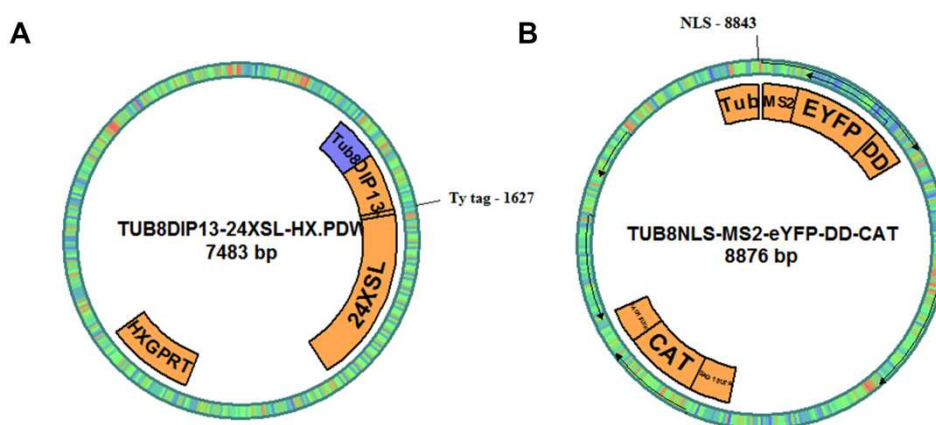


Figura 16. Vetores utilizados para implementação do sistema MS2 (MS2-Dip13) em *T. gondii*. A. Vetor Dip13-SL para expressão constitutiva de mRNA de Dip13 com 24 repetições *stem-loops*. Tub8: promotor constitutivo de tubulina. DIP13: Sequência do cDNA correspondente ao gene *dip13*. 24XSL: 24 Repetições *stem-loops*. HXGPRT: marca de seleção dos parasitas transfectantes. B. Vetor MS2 para expressão da proteína NLS-MS2-YFP-DD ou MS2-YFP-DD que será estabilizada por Shld-1. Tub: promotor constitutivo de tubulina. NLS: Sequência correspondente ao sinal de localização de *T. gondii* (BHATTI e SULLIVAN, 2005). MS2: Sequência correspondente à proteína ligadora de *stem-loops*. EYFP: proteína fluorescente amarela de eucariotos. DD: domínio de desestabilização ddFKBP12. HXGPRT e CAT: marca de seleção dos parasitas transfectantes. A linhagem obtida foi nomeada de linhagem MS2, conforme tabela 2.

Após a confirmação da obtenção da linhagem MS2 e dos resultados de localização da proteína NLS-MS2-YFP-DD e mRNA Dip13-SL, o próximo passo foi então a obtenção da linhagem do sistema MS2 compatível com a estratégia de nocaute condicionado de *uap56*, nomeada de linhagem DiCre-MS2 conforme tabela 2.

O nocaute condicionado é baseado no sistema DiCre, como descrito anteriormente e a integração do cassete contendo cDNA do gene de interesse flanqueado por sequências *LoxP* ocorre através de recombinação homóloga entre as regiões 5'/ 3' UTRs do gene de interesse. Para a futura validação do sistema MS2 com o nocaute de *uap56* foi preciso primeiramente obter uma linhagem de parasitas que além de possuir a sequência de DiCre presente no genoma do parasita não possuísse o gene *ku80* pois, ao contrário da maioria das outras células eucarióticas, o DNA é geralmente integrado de forma randômica no genoma e em apenas 5 a 10% dos casos há a integração por recombinação homóloga quando da presença do gene *ku80* (DONALD e ROOS, 1993). No entanto, quando o gene *ku80* encontra-se nocauteado, há o favorecimento para este tipo de recombinação, pois a proteína Ku80 está envolvida com reparo de DNA e integração randômica (FOX *et al.*, 2009; HUYNH e CARRUTHERS, 2009).

Para a obtenção da linhagem MS2 compatível com o nocaute condicionado de *uap56* foi utilizada a linhagem DiCre-HX (Tabela 2) que possui a sequência DiCre inserida randomicamente no genoma do parasita bem como o gene *ku80* substituído pelo gene *hxgprt*, que confere resistência a HXGPRT, conforme tabela 6. Levando em consideração que o vetor para nocaute condicionado de *uap56* possui a marca de seleção HXGPRT, conforme item 5.2.2 (Figura 14), foi inserido o cassete referente a MS2-Dip13-SL no locus de *ku80*, substituindo então o gene *hxgprt*. Para tanto, a construção do vetor foi baseada nos vetores da figura 16, onde as sequências referentes a NLS-MS2-YFP-DD foram amplificadas em uma etapa a partir do vetor MS2 (Figura 16 B) através de Platinum *Taq* DNA Polimerase *High Fidelity* conforme item 5.1.5, e clonadas no vetor Dip13-SL (Figura 16 A), que já continha a sequência para o mRNA Dip13-SL. As regiões referentes a 5' e 3' UTR de *ku80* foram também amplificadas a partir de DNA genômico de cepa RH de *T. gondii* e clonadas neste mesmo vetor, garantindo a recombinação homóloga no

locus de *ku80* e a substituição do gene *hxgprt* presente no genoma do parasita pelo cassete MS2-Dip13-SL. O cassete MS2-Dip13-SL foi utilizado para transfectar parasitas da linhagem DiCre-HX (Tabela 2) e a seleção dos parasitas foi negativa para HXGPRT através da adição do reagente 6-tioxantina (6-TX), conforme tabela 6. Essa nova linhagem foi denominada “linhagem DiCre-MS2”, conforme tabela 2.

Os *primers* utilizados estão indicados na tabela 11 e a estratégia de clonagem pode ser visualizada na figura 17.

Tabela 11. *Primers* utilizados para amplificar sequência de NLS-MS2-YFP-DD bem como regiões não codificantes 5' e 3' UTR relativas ao gene *ku80*. As enzimas utilizadas estão indicadas no nome do *primer*, do lado esquerdo. As sequências reconhecidas por cada enzima estão indicadas em vermelho, à direita. *Forward* (F). *Reverse* (R).

Nome	Sequência (5' → 3')
TUBMS2DD- <i>Xba</i> I (F)	CCTCTAGATGCCTGCATTGGGTGCCGGTTGG
TUBMS2DD- <i>Xba</i> I (R)	CCTCTAGACAATTGGAAGTAGTGGATCCCCCTCGGGGG
5UTR- <i>Kpn</i> I (F)	CCGGTACCGATATCGCAGAGAAAACG
5UTR- <i>Kpn</i> I (R)	CCGGTACCTAGTTTTGCGAGGAGAATAAGAGAAGTC
3UTR- <i>Sac</i> II (F)	CCCGCGGCTTACTCGATCTCGTTCGAGTAGAAAACG
3UTR- <i>Sac</i> II (R)	CCCGCGGGGATTGACCA

Todas as regiões amplificadas foram amplificadas por PCR e subclonadas em pGEM-Teasy® conforme item 5.1.5. As clonagens foram confirmadas por PCR de colônia conforme item 5.1.8 e sequenciados utilizando os *primers* da tabela 11. Após confirmação por sequenciamento de cada clone obtido, os fragmentos foram liberados do clone de pGEM-Teasy® após reação com as respectivas enzimas, indicadas na tabela de *primers* (Tabela 11), de acordo com o manual do fabricante (New England Biolabs). Cada amostra foi submetida individualmente a eletroforese em gel de agarose 1% para a excisão das respectivas regiões liberadas. A purificação dos fragmentos foi realizada com auxílio do kit *High Pure PCR Product Purification Kit* (Roche # 11732676001) para ligação ao vetor de destino (vetor

Dip13-SL, figura 16 A), também clivados nas mesmas condições para cada etapa de clonagem. Anterior a cada etapa de clonagem, o vetor digerido foi desfosforilado através da enzima CIP (*alkaline phosphatase, calf intestinal*) conforme manual do fabricante (New England Biolabs #M0290S) para evitar recircularização do vetor. A ligação ocorreu conforme item 5.1.6 e a estratégia de clonagem pode ser visualizada na figura 17.

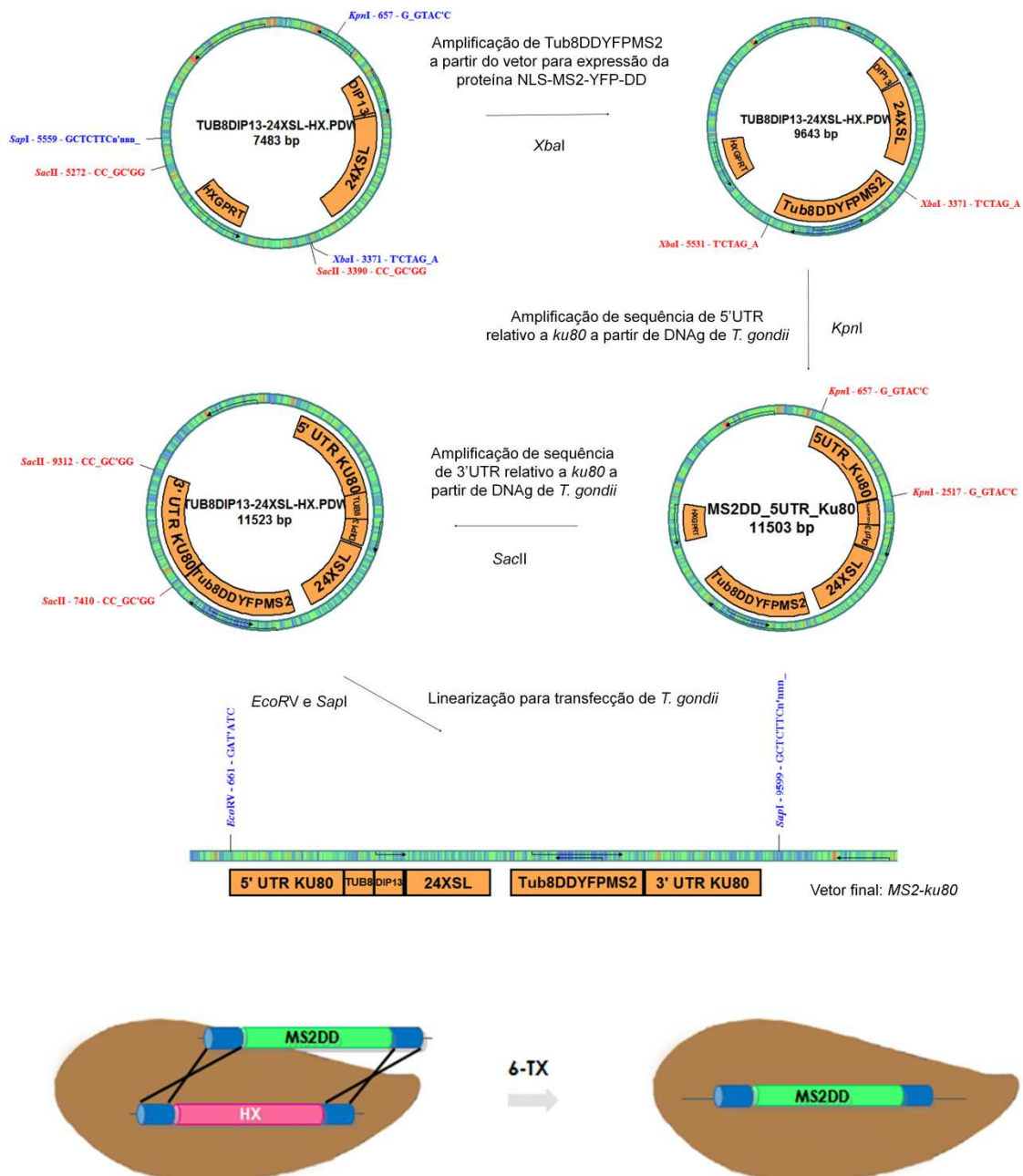


Figura 17. Estratégia de clonagem para a construção do vetor utilizado para obtenção de linhagem DiCre-MS2 em *T. gondii*. O vetor *Dip13-SL* foi utilizado como molde para a construção do vetor final. 5UTR_ku80: Região 5'UTR correspondente ao gene *ku80*. TubDDYFPMS2: Região correspondente ao cassete de MS2 do vetor MS2. HXGPRT/HX: gene de resistência. 3UTR_ku80: região 3'UTR referente ao gene *ku80*. DNAG: DNA genômico. As regiões foram clonadas com as enzimas de restrição indicadas na figura. O vetor final, *MS2-ku80* foi digerido com as enzimas *EcoRV* e *SapI* para integração no locus de *ku80* no genoma do parasita da linhagem DiCre-HX. Na parte de baixo está o esquema de recombinação homóloga entre as regiões 5' e 3'UTR, em azul, relativos ao gene *ku80*. Após a recombinação, o cassete MS2-Dip13-SL é inserido no locus de *ku80*, substituindo o gene *hxgprt*. Na parte de baixo está o esquema de recombinação homóloga entre as regiões 5' e

3'UTR, em azul, relativos ao gene *ku80*. Após a recombinação, o cassete contendo o cassete MS2-Dip13 substituiu o gene *hxprrt* no locus de *ku80*. A linhagem obtida foi nomeada de DiCre-MS2.

A integração do cassete MS2-Dip13-SL no locus do gene *ku80* foi confirmada por PCR de DNA genômico de *T. gondii* da linhagem obtida, DiCre-MS2 (Tabela 2). Os *primers* utilizados estão esquematizados na tabela 12. Para a reação de PCR foram acrescentados aos tubos os reagentes já citados anteriormente e *primers* que hibridizam em regiões do DNA a ser amplificado, em um volume final de reação de 20 µL. As amostras foram incubadas inicialmente com um *holding* de 94 °C por 5 minutos seguidos de 30 ciclos de amplificação. As temperaturas de anelamento foram de 58 °C, por 30 segundos e a extensão ocorreu a 72 °C por 6 minutos. Os produtos da PCR foram analisados em gel de agarose 1%.

Tabela 12. Primers utilizados para confirmação de integração do cassete MS2-Dip13-SL no locus do gene *ku80* no genoma de *T. gondii*. Ext: Externo ao cassete. Int: Interno ao cassete. Forward (F). Reverse (R).

Nome	Sequência (5' → 3')
946 (ku80 ext – F)	AAGACCACGTTCTTCATGATAGCTGTACACCACGAGTTGGACGACTGATTTAGATCTTG
811 (MS2 – F)	CCCAATTGCCTAGGATCGAAGGCCGCCATATGCTAGCC
118 (Dip13 – R)	CCCGAATTCTCCTGCGCCAGCTGTGGG
766 (ku80 ext – R)	GGAACCGTGAAACTGACGTCTGTGAC
940 (ku80 int – R)	AGTTCACCGCTATTTACGC
812 (YFP int - R)	CCATGCATGCAGCTCGTCCATGCCGAGA

5.2.4 Hibridação fluorescente *in situ* (FISH) para detecção de mRNAs em *T. gondii*

O protocolo utilizado foi adaptado de Lirussi e Matrajt, 2011 e Serpeloni *et al.*, 2011 (LIRUSSI e MATRAJT, 2011; SERPELONI *et al.*, 2011). Para isso foi necessária a inoculação prévia de células HFF em lamínulas redondas contidas em

placas de 24 poços. Após a obtenção de monocamada confluyente de HFF, as células foram infectadas com os parasitas e, dois dias após a infecção, foi então adicionado ao meio de cultura rapamicina ou Shld-1, dependendo do objetivo do experimento. Após o tempo determinado de incubação com a droga, as células foram fixadas por 20 minutos com paraformaldeído 4% em PBS, lavadas com PBS e então permeabilizadas com triton 0,2 % em SSPE 2X por 30 minutos a temperatura ambiente. As células foram bloqueadas com tampão de hibridação (TH, Tabela 3) por 30 minutos a 37 °C, em câmara úmida. Em seguida foi adicionado 1 ng/μl da sonda em TH. Para detecção do mRNA com os sítios de ligação de MS2, foi usada sonda complementar a repetições do sítio de ligação de MS2: (5' TGATTTGAAGATTCGACCTGGAGGCACGGCTGATGCTCGTAATTTCTTGGCAGT AAGT 3') (BERTRAND *et al.*, 1998). Para detecção de mRNAs poliadenilados foi utilizada sonda de oligodT, contendo 30 repetições de timidinas. Ambas as sondas foram conjugadas a digoxigenina na região 5'. A lâmina contendo o tampão com sonda foi previamente aquecida a 65 °C por 3 minutos. Em seguida, a temperatura foi diminuída para 37 °C e o material incubado por 16 horas em câmara úmida para hibridação. Depois as lâminas foram lavadas sucessivamente com SSC 2X e 1X por 15 minutos cada. Para a detecção de digoxigenina foi necessário realizar o protocolo de imunofluorescência, como citado no item 5.1.9, utilizando o anticorpo primário monoclonal anti-digoxigenina (Tabela 4), com diluição 1:600 em PBS-BSA 4%.

5.3 Resultados

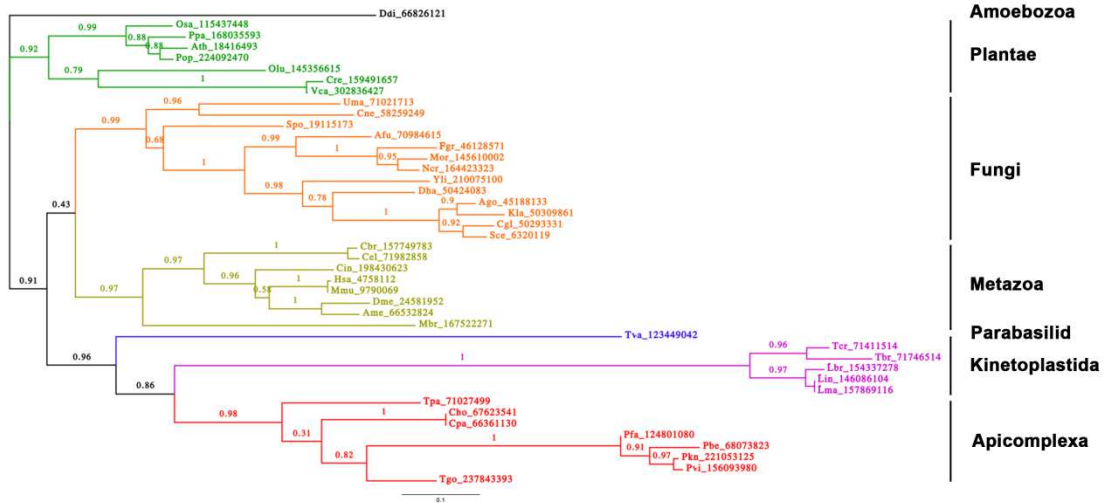
5.3.1 Análise de conservação da proteína TgUAP56 na filogenia de eucariotos

Foi mostrado previamente que a proteína UAP56/Sub2 está presente na maioria dos grupos de eucariotos (SERPELONI *et al.*, 2011) e que provavelmente já estava presente em LECA (*Last Eukaryotic Common Ancestor*). Para compreender a história evolutiva dessa proteína nos eucariotos, foi feita a reconstrução filogenética utilizando sequências ortólogas de UAP56 de 43 espécies representativas de diferentes grupos, incluindo parasitas dos grupos Kinetoplastida e Apicomplexa (Figura 18 A). As sequências foram agrupadas em clados que correspondem aos grupos de eucariotos. Os valores de suporte dos ramos (*bootstrap*) foram 0.92, 0.99, 0.97, 1.00, 0.98, para os grupos Plantae, Fungi, Metazoa, Kinetoplastida e Apicomplexa, respectivamente. Parasitas dos grupos Apicomplexa e Kinetoplastida agruparam como clados irmãos com valor de suporte de 0.86, mostrando que as sequências ortólogas de UAP56/Sub2 destes dois grupos de parasitas estão intimamente relacionadas. *Trichomonas vaginalis* (Parabasilid), *Trypanosoma* spp, *Leishmania* spp (Kinetoplastida), e *Plasmodium* spp (Apicomplexans) são as sequências mais divergentes, o que pode ser observado pelos ramos maiores na árvore filogenética.

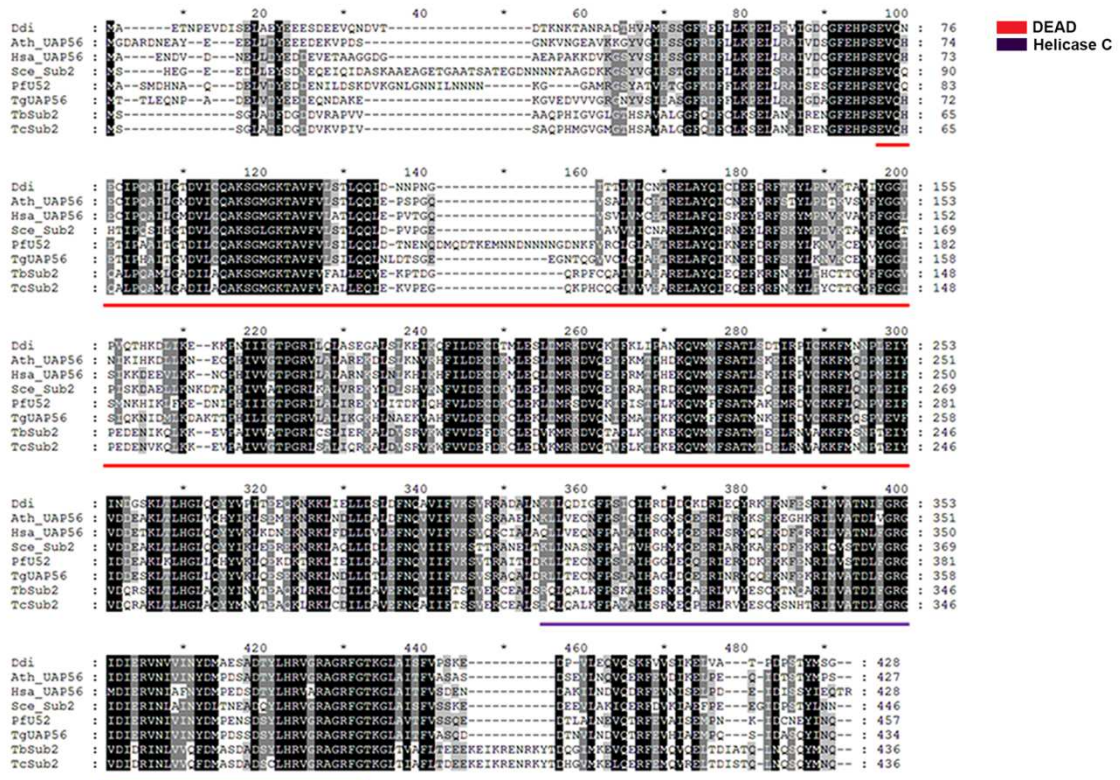
O alinhamento entre as sequências representativas dos diferentes grupos de eucariotos mostra a alta conservação ao longo de toda proteína bem como a presença dos dois domínios típicos de RNA helicases canônicas (ROCAK e LINDER, 2004): domínio na porção N-terminal DEAD-box (PF00270) e domínio na porção C-terminal Helicase_C (PF00271), como pode ser observado na figura 18 B. A proteína ortóloga em *T. gondii* foi nomeada de TgUAP56 porque já existe uma proteína com o nome TgSub2 em *T. gondii* (TGME49_314500 - *subtilisin* SUB2), uma protease trans-membrana localizada em roptrias envolvida na via secretória deste parasita (MILLER *et al.*, 2003). TgUAP56 está anotada no banco de dados do ToxoDB como *DEAD-box polypeptide DDX39* (ID: TGME49_216860, GI 237843393 no NCBI). Em relação às proteínas ortólogas conhecidas de TgUAP56 a identidade/similaridade vai de 47.1/62.2% em *Trypanosoma cruzi* (SERPELONI *et al.*, 2011), a 68.6/81.4% em *Plasmodium falciparum*, PfU52 (SHANKAR, PRADHAN

e TUTEJA, 2008), como mostrado na figura 18 C.

A



B



C

	Ddi	Ath UAP56	Hsa UAP56	Sce Sub2	PfU52	TgUAP56	TbSub2	TcSub2
Ddi	x	78.0	77.4	71.8	70.5	74.0	65.1	65.5
Ath UAP56	63.4	x	83.3	76.5	72.8	77.6	64.1	64.9
Hsa UAP56	58.1	70.1	x	77.1	71.0	78.6	66.4	66.4
Sce Sub2	54.4	61.4	61.5	x	70.9	72.8	62.7	61.5
PfU52	53.9	58.5	56.9	54.6	x	81.4	63.1	62.2
TgUAP56	57.0	65.1	64.6	56.1	68.6	x	66.1	66.1
TbSub2	44.5	48.1	48.5	43.7	44.2	47.3	x	96.1
TcSub2	45.0	49.4	49.0	44.2	44.7	47.1	89.4	x

Figura 18. UAP56 é uma proteína conservada em eucariotos. A. Reconstrução filogenética de sequências ortólogas de UAP56 em diferentes grupos de eucariotos. Números nos nodos correspondem ao valor de *bootstrap*. Amoebozoa (preto): Ddi – *Dictyostelium discoideum*; Plants (green): Osa – *Oryza sativa*, Ppa – *Physcomitrella patens*, Ath – *Arabidopsis thaliana*, Pop – *Populus trichocarpa*, Olu – *Ostreococcus lucimarinus*, Cre – *Chlamydomonas reinhardtii*, Vca – *Volvox carteri*;

Fungi (laranja): Uma – *Ustilago maydis*, Cne – *Cryptococcus neoformans*, Spo – *Schizosaccharomyces pombe*, Afu – *Aspergillus fumigatus*, Fgr – *Fusarium graminearum*, Mor – *Magnaporthe oryzae*, Ncr – *Neurospora crassa*, Yli – *Yarrowia lipolytica*, Dha – *Debaryomyces hansenii*, Ago – *Ashbya gossypii*, Kla – *Kluyveromyces lactis*, Cgl – *Candida glabrata*, Sce – *Saccharomyces cerevisiae*; Metazoans (amarelo): Cbr – *Caenorhabditis briggsae*, Cel – *Caenorhabditis elegans*, Cin – *Ciona intestinalis*, Hsa – *Homo sapiens*, Mmu – *Mus musculus*, Dme – *Drosophila melanogaster*, Ame – *Apis mellifera*, Mbr – *Monosiga brevicolis*; Parabasilid (blue): Tva – *Trichomonas vaginalis*; Kinetoplastids (roxo): Lbr – *Leishmania braziliensis*, Lin – *Leishmania infantum*, Lma – *Leishmania major*, Tbr – *Trypanosoma brucei*, Tcr – *Trypanosoma cruzi*; Apicomplexans (vermelho): Tpa – *Theileria parva*, Cho – *Cryptosporidium hominis*, Cpa – *Cryptosporidium parvum*, Pfa – *Plasmodium falciparum*, Pbe – *Plasmodium berghei*, Pkn – *Plasmodium knowlesi*, Pvi – *Plasmodium vivax*, Tgo – *Toxoplasma gondii*. B. Alinhamento de múltiplas sequências ortólogas de UAP56. Esta análise mostra a alta conservação da proteína ao longo da sequência completa. Domínio DEAD na porção N-terminal das proteínas (PF00270). Domínio Helicase_C na porção C-terminal das proteínas (PF00271). Ambos os domínios estão destacados. O número de identificação do Genbank (GI) para estas proteínas são: (Ddi 66826121), (Ath_UAP56 18416493), (Hsa_UAP56 4758112), (Sce_Sub2 6320119), (PfU52 124801080), (TgUAP56 237843393), (TbSub2 71746514), (TcSub2 71411514). C. Porcentagem de similaridade (diagonal superior) e de identidade (diagonal inferior) baseadas em alinhamento global ótimo entre os pares de sequências mostrados em B.

5.3.2 Localização celular de TgUAP56 em *T. gondii* na linhagem DDGFPUAP56

5.3.2.1 Obtenção de linhagem DDGFPUAP56 e estabilização da proteína dd-GFP-TgUAP56 após incubação com Shld-1

O sistema DD, como citado anteriormente, permite regulação rápida e reversível dos níveis de proteínas fusionadas ao domínio ddFKBP12 e vem sendo utilizado para estudos de proteínas em *T. gondii* (KREMER *et al.*, 2013; PIEPERHOFF *et al.*, 2013). Foram iniciadas as análises de localização da proteína TgUAP56 utilizando sistema de superexpressão da proteína fusionada à GFP e ddFKBP12 na porção N-terminal. Para tanto, a sequência codificante do gene, foi amplificada a partir de cDNA e clonada em vetor contendo as sequências de DD e GFP. Após a confirmação das clonagens por sequenciamento, o vetor final foi linearizado com *KpnI* para a transfecção de parasitas da cepa RH, conforme item 5.2.1. A seleção dos parasitas transfectantes foi feita através de seleção positiva para HXGPRT, conforme tabela 6, e a integração no genoma ocorreu de modo randômico. Após a seleção dos parasitas, foi observado que TgUAP56 é uma proteína nuclear. No entanto, notamos que a indução da estabilização da proteína utilizando 2 μ M de Shld-1 por 108 horas causava a morte do parasita, indicando que esta “superexpressão” de TgUAP56 causa um fenótipo dominante negativo que é

letal. Por isso, foi feito um ensaio de crescimento por método de placa de lise (*plaque assay*), com incubação de diferentes concentrações de Shld-1 por 108 horas (4 ½ dias), conforme item 5.1.19. Os parasitas transfectantes que foram incubados com as concentrações de 1 e 2 µM não formaram placas de lise, confirmando que a superexpressão é letal ao parasita (Figura 19 A), e mostrando uma forte evidência de que TgUAP56 é uma proteína de função essencial. A confirmação da superexpressão de dd-GFP-TgUAP56 foi realizada por ensaio de *western blot* revelado por quimioluminescência, conforme item 5.1.10, após a incubação com Shld-1 2 µM em diferentes tempos, utilizando o anticorpo primário monoclonal que reconhece especificamente GFP, indicado na tabela 4 (Figura 19 B). Logo após 6 horas de incubação com Shld-1 foi detectada a proteína dd-GFP-TgUAP56 no *western blot*, atingindo níveis máximos após 36 horas. Para esta análise de quantificação foi utilizado o programa ImageJ *densitometry software* (Version 1.6, National Institutes of Health, Bethesda, MD), tendo como referência a proteína constitutiva aldolase como controle de quantidade de proteína aplicada em cada poço. Em virtude desses resultados, para a localização subcelular da proteína, estes parasitas foram incubados por apenas 6 horas com Shld-1 2 µM e posteriormente visualizados por microscopia de fluorescência. Da mesma forma que as proteínas ortólogas de outros eucariotos, TgUAP56 é uma proteína exclusivamente nuclear em formas taquizoítas do parasita, como mostrado na figura 19 C.

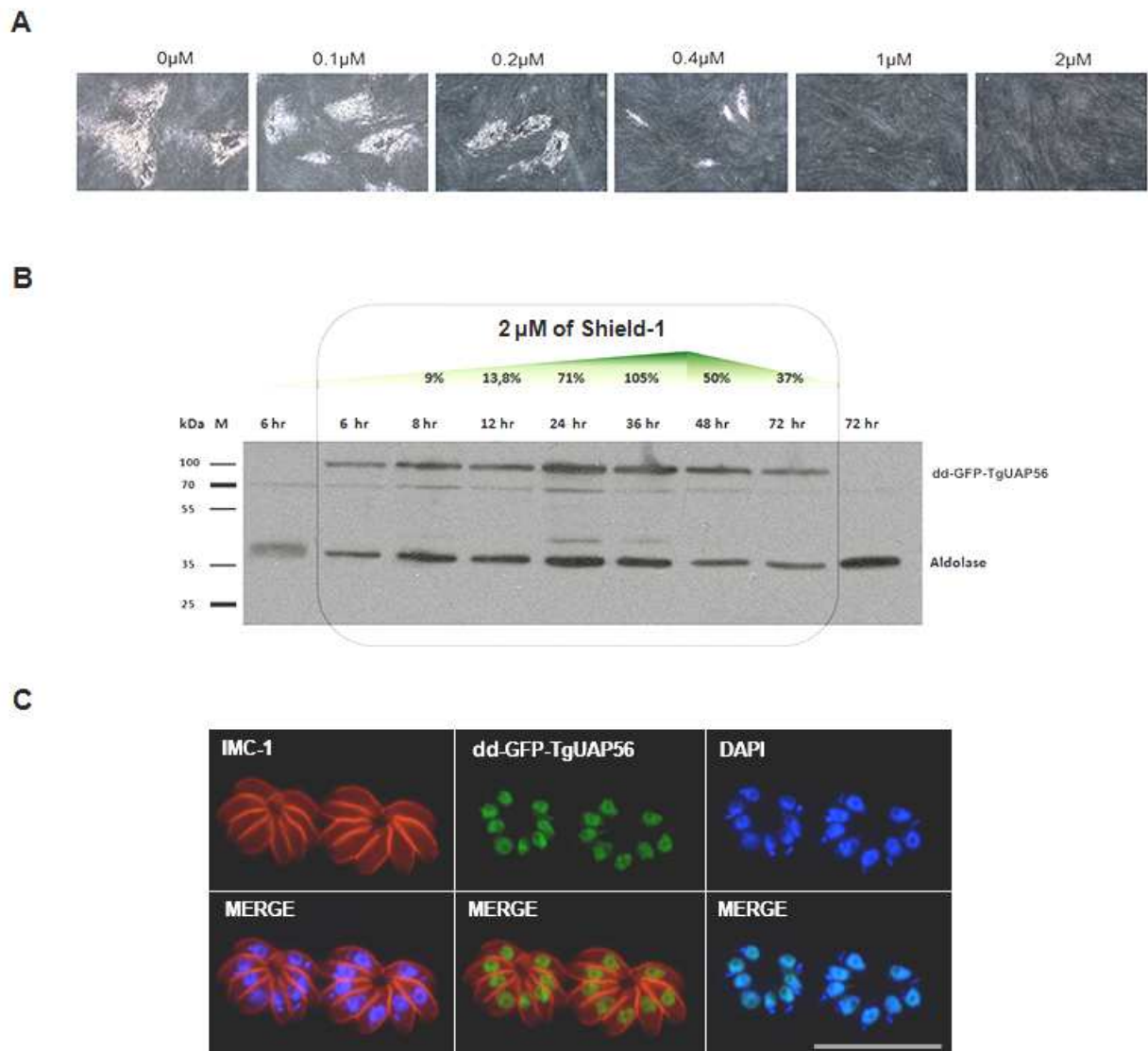


Figura 19. TgUAP56: Proteína nuclear e essencial em *T. gondii*. A. Ensaio de placas de lise (*plaque assay*) de parasitas transfectantes de dd-GFP-TgUAP56 em células HFF na presença de diferentes concentrações de Shld-1 por 108 horas, indicado acima de cada figura. Para a visualização das placas de lise, as células foram fixadas e coradas com Giemsa, conforme item 5.1.19. B. Análise de superexpressão de dd-GFP-TgUAP56 por *western blot* após incubação com 2 μ M de Shld-1 em diferentes tempos, indicados na parte superior da figura. Foi utilizado o anticorpo primário anti-GFP para a detecção da proteína fusionada e anti-Aldolase para a detecção da proteína constitutiva aldolase, utilizada com controle de quantidade de proteína aplicada em cada poço neste experimento. Para a quantificação da quantidade de proteína dd-GFP-TgUAP56 no *western blot* foi utilizado o programa ImageJ, tendo como referência a quantidade de proteína aldolase para cada poço (Version 1.6, National Institutes of Health, Bethesda, MD). A quantificação para cada ponto se encontra na região superior da figura, em porcentagem. C. Localização subcelular de dd-GFP-TgUAP56 por microscopia de fluorescência, em verde. IMC = proteína de superfície de membrana (*inner membrane complex*), em vermelho. DAPI: Marcação da cromatina nuclear e de apicoplasto, em azul. Merge: Sobreposição de imagens. Barra = 10 μ m.

Aparentemente, a fusão com GFP não alterou a localização da proteína, visto que a proteína endógena também foi localizada utilizando anticorpo policlonal anti-

TcSub2 (Tabela 4), por imunofluorescência conforme descrito no item 5.1.9. Através de ensaio de *western blot* conforme item 5.1.10 foi verificado que este anticorpo também reconhece especificamente a proteína TgUAP56 endógena de massa molecular de aproximadamente 50 kDa no extrato protéico de parasitas transfectantes, não incubados com Shld-1, devido à alta conservação entre TcSub2 e TgUAP56 (Figura 20 A, poço 1). Após a incubação com Shld-1 2 μ M por 6 horas foi possível detectar a proteína fusionada dd-GFP-TgUAP56 de massa molecular aproximada de 80 kDa no extrato protéico de parasitas transfectantes (figura 20 A, poço 2). Outra observação foi que após 12 horas de incubação com Shld-1, notamos uma concentração pontual tanto de TgUAP56 quanto dd-GFP-UAP56, próximos à região do nucléolo, com diminuição de sinal geral no nucleoplasma. Estes pontos de dd-GFP-TgUAP56 aparentemente permanecem por 24 e 48 horas, como pode ser observado na figura 20 B, e ainda é preciso avaliar a relevância deste resultado.

A



B

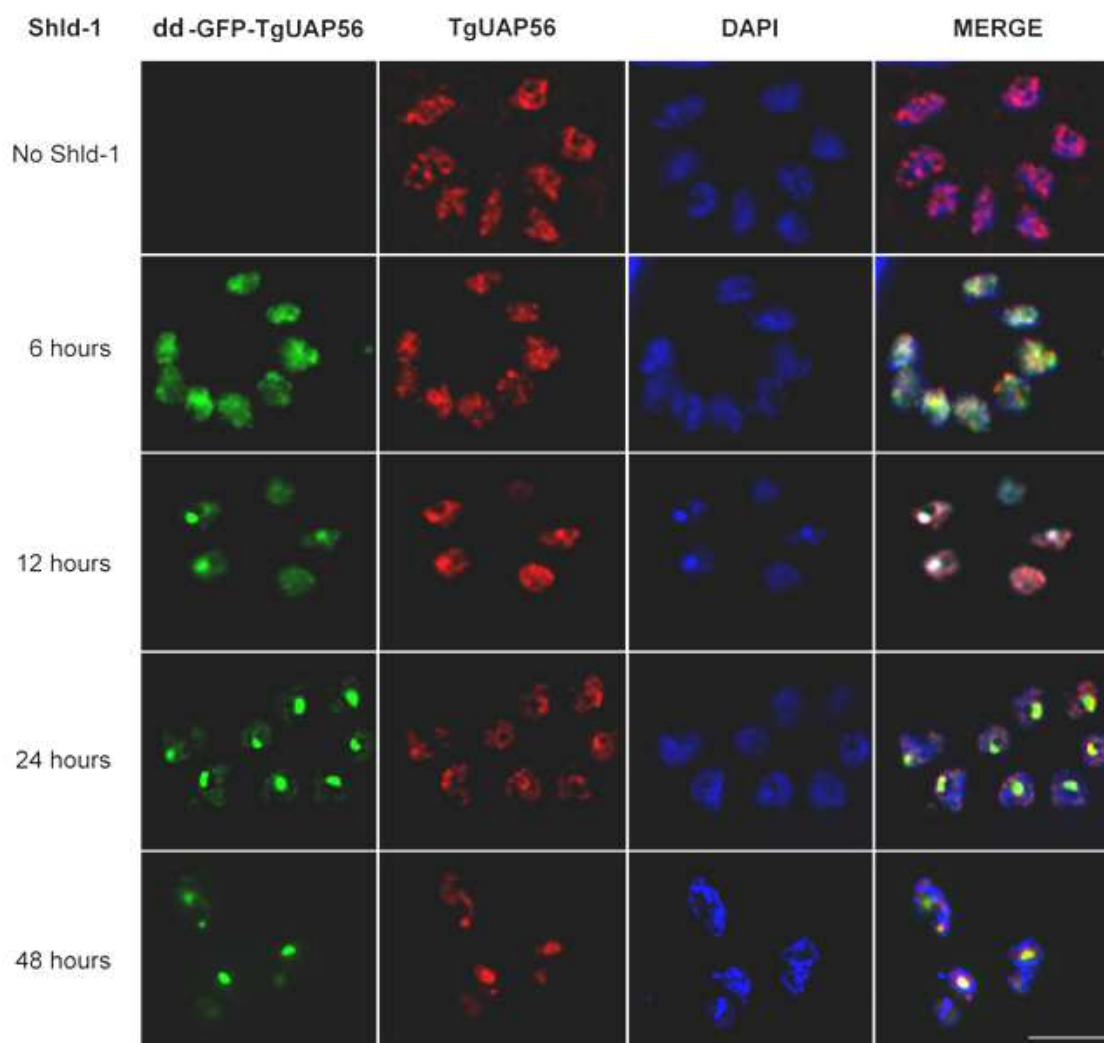


Figura 20. Localização de dd-GFP-TgUAP56 e TgUAP56 por imunofluorescência na linhagem DDGFPUAP56. A. Confirmação da presença de TgUAP56 endógena e dd-GFP-TgUAP56 em extrato protéico de parasitas transfectantes por *Western blot* revelado por quimioluminescência, utilizando o anticorpo primário anti-TcSub2 (Tabela 4). 1. Extrato protéico de parasitas não incubados com Shld-1. 2. Extrato protéico de parasitas incubados previamente com 2 μ M de Shld-1 por 6 horas. Foi utilizada a proteína Aldolase como controle de quantidade de proteína aplicada, detectada por anticorpo primário anti-aldolase (Tabela 4). B. Análise de colocalização entre a proteína endógena TgUAP56,

detectada pelo anticorpo anti-TcSub2, e a proteína fusionada, dd-GFP-TgUAP56, detectada por fluorescência após incubação com Shld-1. Imunolocalização de TgUAP56 endógena, em vermelho. Localização de dd-GFP-TgUAP56, em verde. DAPI: Marcação da cromatina nuclear e de apicoplasto, em azul. Merge: Sobreposição de imagens. Barra = 5 μ m.

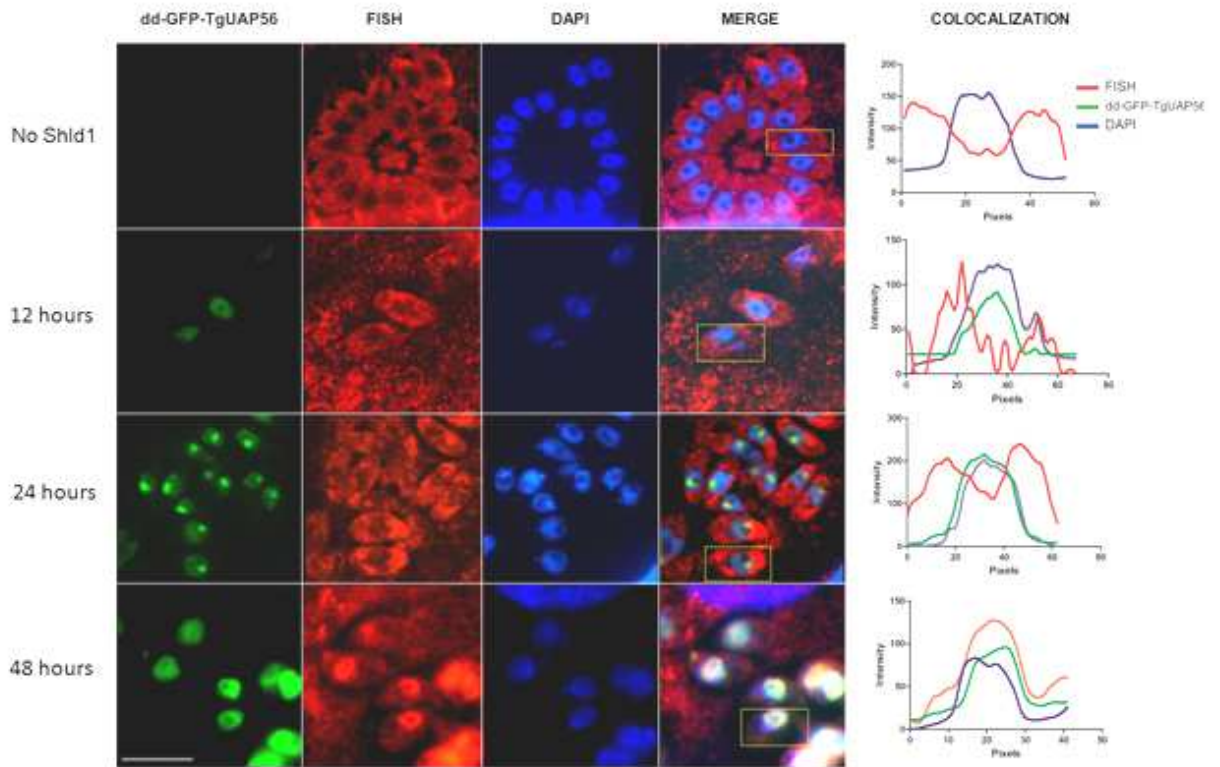
5.3.2.2 Análise de distribuição de mRNAs após indução da superexpressão de dd-GFP-TgUAP56

Visto o fenótipo dominante negativo letal causado após superexpressão de dd-GFP-TgUAP56, o próximo passo foi analisar se isso seria devido ao bloqueio na exportação de mRNA, resultando em morte dos parasitas. Para isso foram realizados ensaios de hibridação *in situ* seguido de fluorescência (FISH), utilizando uma sonda oligodT conjugada a digoxigenina para detecção de mRNAs poliadenilados após a indução da superexpressão, conforme item 5.2.4. Neste experimento os parasitas foram incubados com Shld-1 2 μ M em diferentes tempos e o perfil de distribuição de mRNAs foi observado a cada ponto. Foram também realizadas análises de colocalização através da detecção de pixels referentes aos picos de intensidade de fluorescência de dd-GFP-TgUAP56, mRNAs (FISH) e núcleo (DAPI) em uma região delimitada (em pixels) da imagem obtida de um parasita selecionado figura 21 A, lado direito.

Após 12 horas de incubação com Shld-1 foi detectada a presença da proteína dd-GFP-TgUAP56 no núcleo, mas ainda com uma distribuição citoplasmática dos mRNAs (figura 21 A -12 horas). Após 24 horas, foi observado um acúmulo de dd-GFP-TgUAP56 em região específica do núcleo, como observado anteriormente quando ocorre a superexpressão da proteína, sendo que a maioria dos mRNAs possui uma distribuição citoplasmática, assim como em parasitas da cepa RH, mesmo após 48 horas (Figura 21 A, B). No entanto, notamos que também existe o início de um acúmulo de mRNAs na mesma região onde ocorre acúmulo de dd-GFP-TgUAP56, como destacado nas figuras 21, C e D. Já após 48 horas foi observado o acúmulo da maioria dos mRNAs nos núcleos dos parasitas (Figura 21 - 48 horas) e após esse período, foi observado que este acúmulo, tanto da proteína quanto dos mRNAs é letal ao parasita. Como controle do experimento, foi observada a distribuição de mRNAs na cepa RH após incubação de 48 horas com mesma

concentração de Shld-1, onde não foram observados esses pontos específicos em verde ou mRNAs poliadenilados (Figura 21 B).

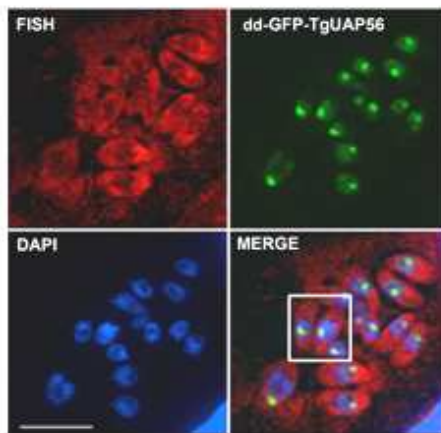
A



B



C



D

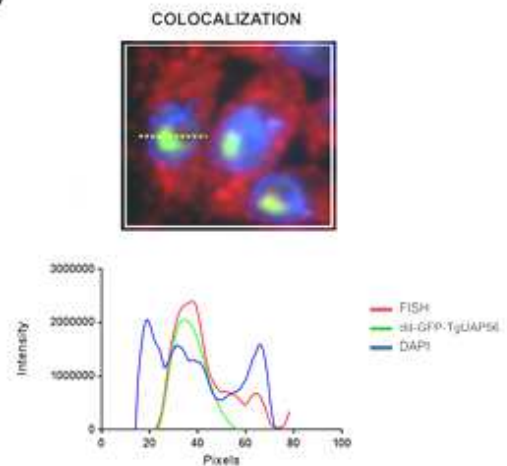


Figura 21. Análise de distribuição de mRNAs após superexpressão de dd-GFP-TgUAP56 na linhagem DDGFPUAP56. A. A superexpressão de dd-GFP-TgUAP56 foi induzida por Shld-1 2 μ M

em diferentes tempos, quando indicado. B. Distribuição de mRNAs em parasitas selvagens (cepa RH) após incubação com 2 μ M de Shld-1 por 48 horas. C. Distribuição de mRNAs após incubação com 2 μ M de Shld-1 por 24 horas. D. Análise de colocalização de dd-GFP-TgUAP56, mRNAs (FISH) e núcleo (DAPI) do parasita presente na região selecionada em C. Os mRNAs poliadenilados foram detectados por hibridação *in situ* utilizando sonda oligodT conjugada com digoxigenina seguido de imunofluorescência (FISH), utilizando anticorpo primário anti-digoxigenina (Tabela 4), em vermelho. Localização de dd-GFP-TgUAP56, em verde. DAPI = Marcação da cromatina nuclear e de apicoplasto, em azul. Merge: Sobreposição de imagens. *Colocalization*: Análise de colocalização de dd-GFP-TgUAP56, mRNAs (FISH) e núcleo (DAPI) foi realizada utilizando ImageJ (Version 1.6, National Institutes of Health, Bethesda, MD) para cada parasita, destacado em amarelo na região de sobreposição das imagens (merge). Barra = 10 μ m.

5.3.3 Nocaute o gene causa acúmulo de mRNAs no núcleo

Levando em consideração a essencialidade desta proteína em tripanossomatídeos, bem como os resultados de superexpressão em *T. gondii*, a simples retirada do gene *uap56* seria letal ao parasita e, assim como ocorreu durante a seleção do nocaute duplo em *T. cruzi*, não conseguiríamos obter uma linhagem nocaute deste gene em *T. gondii*. Por isso optamos pela estratégia de nocaute condicionado baseado no sistema DiCre, recentemente publicado em *T. gondii* (ANDENMATTEN *et al.*, 2013; BARGIERI *et al.*, 2013), sendo possível assim obter a linhagem cKOuap56 para análise da distribuição de mRNAs ao longo do tempo após indução do nocaute de *uap56*. Esse sistema nos permitiu confirmar o envolvimento da proteína com a exportação de mRNAs e de que o fenótipo observado pela superexpressão não seria um evento indireto.

5.3.3.1 Obtenção da linhagem clonal cKOuap56 e indução do nocaute do gene *uap56* por rapamicina

Para a estratégia de nocaute foi construído um vetor contendo um cassete que consiste na sequência codificante de *uap56* flanqueada por sequências *LoxP* situada *upstream* ao gene repórter *mcherry*, e que nas suas extremidades possui as regiões 5' e 3' UTRs relacionadas ao gene *uap56*. A estratégia de clonagem está detalhada no item 5.2.2. Neste sistema de nocaute condicionado o que ocorre é que após a reconstituição da atividade enzimática de DiCre pela incubação dos parasitas com rapamicina, há a circularização do DNA flanqueado por *LoxP* e consequente retirada deste do genoma do parasita. Logo, primeiramente, parasitas da linhagem

DiCre- $\Delta ku80$ foram transfectados com o cassete contendo cDNA de *uap56* flanqueado por *LoxP*, para a obtenção da linhagem cKOuap56 (Tabela 2). Nesta linhagem, após a reconstituição de DiCre por incubação com rapamicina, o gene marcador *mcherry* é então posicionado sob regulação do promotor endógeno de *uap56*, originando parasitas que emitem fluorescência no comprimento de onda de 600 nm, podendo ser utilizado para monitoramento da ocorrência do nocaute do gene.

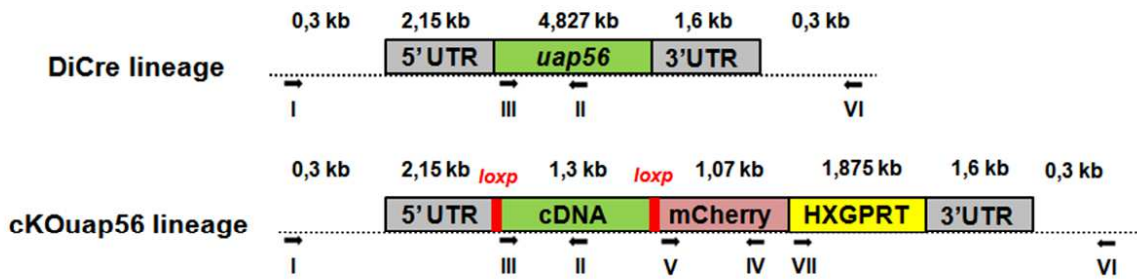
A linhagem clonal para o nocaute condicionado cKOuap56 apresentou crescimento normal em cultura de células HFF quando na ausência de rapamicina. No entanto, a incubação com rapamicina 50 nM por 3 dias foi letal, indicando que se trata realmente de uma proteína essencial para a sobrevivência do parasita.

Para avaliar se a recombinação do cassete ocorreu corretamente no genoma dos parasitas da linhagem cKOuap56, foram feitas análises por PCR com DNA provenientes de: DiCre- $\Delta ku80$ (controle) e linhagem cKOuap56 sem incubação de rapamicina e cKOuap56 com incubação de 50 nM de rapamicina por 24 horas, como explicado no item 5.2.2.1. Os resultados desta análise estão mostrados na figura 22. Na figura 22 A está esquematizada uma tabela com os tamanhos esperados dos produtos de PCR nas diferentes linhagens analisadas para cada par de *primer*. Na figura 22 B pode ser observada a região de hibridação de cada *primer* no genoma das linhagens DiCre- $\Delta ku80$ (DiCre *strain*) e cKOuap56, bem como o tamanho das sequências em análise. Na figura 22 C está o resultado do gel de agarose 1% dos produtos de PCR obtidos com os diferentes pares de *primers*.

A

	PRIMERS		cKOuap56		DiCre strain
			A No rapa	B 50 nM rapa	C 50 nM rapa
1	I / II	5UTRext F / uap56 int R	2,79 kb		4,4 kb
2	III / II	uap56 F / uap56 int R	0,334 kb		2,012 kb
3	I / IV	5UTRext F / mcherry R	4,5 kb	3,2 kb	
4	V / VI	mcherry F / 3UTRext R	4,6 kb	4,6 kb	
5	VII / VI	hxgprt F / 3UTRext R	2,4 kb	2,4 kb	

B



C

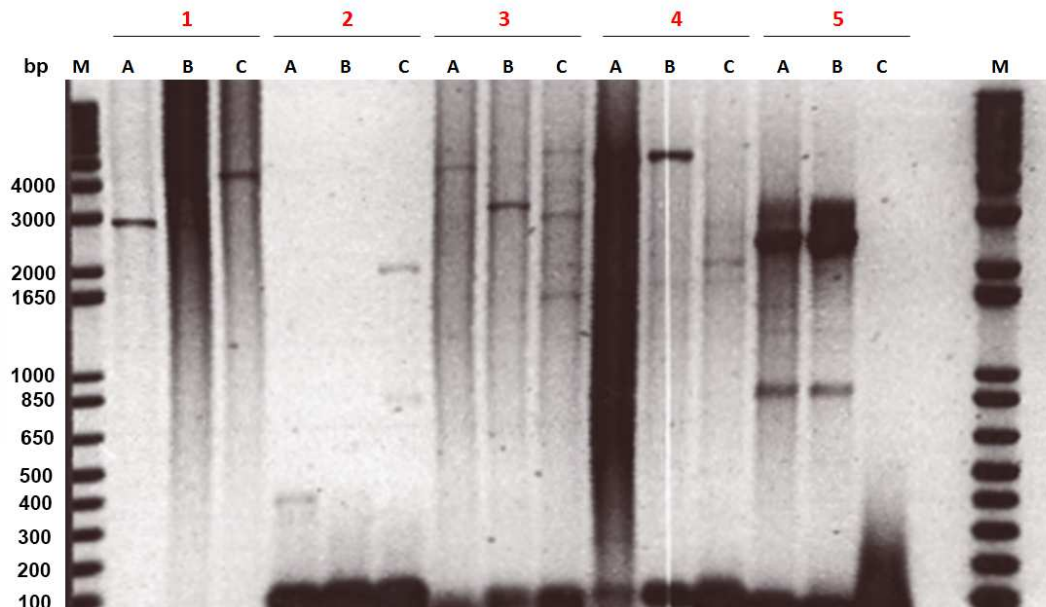


Figura 22. Análise por PCR de integração do cassete para nocaute condicionado do gene *uap56* no genoma de parasitas da linhagem clonal cKOuap56. A. Tabela com o código e nome dos *primers* do lado esquerdo. As sequências dos *primers* estão na Tabela 9. As colunas com as letras A, B e C indicam as populações analisadas: cKOuap56: Linhagem clonal cKOuap56. DiCre strain: Linhagem parental DiCre- $\Delta ku80$. “no rapa/50nM rapa” indica ausência/presença de incubação

prévia com 50 nM de rapamicina por 24 horas. Para cada par de *primer*, o tamanho esperado está indicado na tabela em Kb (quilobases). As células pintadas de cinza indicam obrigatoriamente a ausência de amplificação, pela ausência dos genes nos parasitas da linhagem em análise. Após a adição de rapamicina, espera-se a ausência do cDNA de *uap56* no genoma dos parasitas e portanto ausência de amplificação. No caso da linhagem parental DiCre- $\Delta ku80$ (DiCre *strain*), não há os genes *mcherry* e *hxgprt*. B. Desenho esquemático de região de anelamento dos *primers* indicados em A em ambas as linhagens. O tamanho de cada região esta indicada acima de cada esquema, em kb (quilobases). cDNA: cDNA relativo ao gene *uap56*. C. Resultado da PCR em gel de agarose 1%. Os números indicam os pares de *primers* utilizados e as letras indicam as linhagens analisadas, como mostrados em A. M: 1 kb Plus DNA (Invitrogen # 10787-018).

De acordo com esses resultados foi possível confirmar a integração do cassete para o nocaute induzido. De acordo com a tabela, nas colunas A e B estão os resultados para a linhagem cKOuap56, que deve conter o cDNA referente a *uap56* integrado. A diferença entre elas é que em A não houve indução do nocaute, enquanto houve em B. A coluna C se refere à linhagem DiCre- $\Delta ku80$ (DiCre *strain*) que deve conter gene *uap56*, pois foi a linhagem parental utilizada para a integração do cassete com cDNA de *uap56*.

O *primer* interno para *uap56* foi desenhado justamente para hibridizar em uma região que contem o quarto exon para verificar a diferença de tamanho entre o gene e o cDNA de *uap56*, o qual não possui introns. Os resultados referentes a esse *primer* estão nas colunas 1A e 1C, bem como 2A e 2C da figura 22 C, que confirmam que em parasitas da linhagem clonal de cKOuap56 ocorreu a substituição do gene *uap56* endógeno pelo cDNA presente no cassete. No caso da coluna 1, foram utilizados os *primers* 5UTR externo e *uap56* interno. Quando comparados os produtos de PCR entre as colunas 1A (linhagem cKOuap56, que contem o cDNA de *uap56*) com a coluna 1C (linhagem parental, DiCre- $\Delta ku80$, que contem o gene *uap56*) é possível ver esta diferença de tamanho. Para a coluna 2 (resultados 2A e 2C), esta confirmação foi através da utilização dos *primers* específicos para o gene *uap56* sendo um deles este *primer* interno reverso de *uap56*, que também confirma esta diferença de tamanho entre gene e cDNA.

Ainda em relação a figura 22 C, a retirada do cDNA de *uap56* após indução por rapamicina foi confirmada pelos resultados das colunas 2B e 3B. No caso do resultado de 2B não foi observada amplificação do cDNA, conforme esperado, visto que foram usados os *primers* específicos para o gene *uap56*. Para o resultado 3B foram utilizados os *primers* 5UTR externo e *mcherry* R, que flanqueiam o cDNA referente ao gene *uap56*. Conforme o esperado, houve uma diminuição do tamanho

do produto de PCR referente a 1300 bp quando comparado a 3A, utilizando os mesmos *primers*, resultado da retirada do cDNA. A integração correta do cassete foi confirmada também pela presença dos genes *mcherry* e *hxgprt* nas linhagens cKOuap56, como é possível observar nas colunas correspondentes a estes *primers*: 3, 4 e 5, ausentes na linhagem DiCre *strain*. É importante ressaltar que foi utilizado a mesma quantidade de DNA genômico para as reações de PCR.

5.3.3.2 Análise de acúmulo de mRNAs na linhagem cKOuap56 após nocaute do gene *uap56*

Após a confirmação tanto da obtenção da linhagem cKOuap56 quanto da possibilidade de induzir a retirada do cDNA referente a *uap56* por rapamicina, foram analisados os níveis da proteína TgUAP56 em diferentes tempos após a indução do nocaute com 50 nM de rapamicina, bem como o perfil de distribuição de mRNAs poliadenilados nos parasitas, conforme item 5.2.2. Foram também realizadas análises de colocalização através da detecção de pixels referentes aos picos de intensidade de fluorescência de mRNAs (FISH) e núcleo (DAPI) em uma região delimitada (em pixels) da imagem obtida de um parasita selecionado figura 23 A, lado direito.

Através de ensaios de *western blot* revelado por quimioluminescência, conforme descrito em item 5.1.10, verificamos a diminuição do nível de proteína quando utilizado anticorpo anti-TcSub2 (Tabela 4) para detecção de TgUAP56. Foi possível observar uma diminuição drástica dos níveis desta proteína logo após 24 horas. Neste ensaio a detecção da proteína aldolase foi usada como controle da quantidade de extrato protéico aplicado em cada poço (Figura 23 A). Pela quantificação, foi verificada uma diminuição de 25 % do nível de proteína aldolase após 24 horas de indução do nocaute, provavelmente consequência do acúmulo dos respectivos mRNAs e diminuição da tradução desta no parasita. Mesmo com a diminuição da aldolase, a diminuição de TgUAP56 foi drástica após 24 horas do nocaute, sendo superior a 95 %.

Quando analisado o perfil de distribuição de mRNAs, foi possível observar o acúmulo nuclear em alguns parasitas após 24 horas da indução do nocaute, sendo este resultado mais evidente em 48 horas, onde praticamente todos os parasitas

apresentam acúmulo de mRNA nuclear (Figura 23 B - 24 e 48 horas). Após 48 horas, também foi possível observar parasitas fluorescendo em comprimento de onda de 600 nm, na cor vermelha, indicando que houve a retirada do cDNA de *uap56* e consequente reposicionamento do gene *mcherry* no locus de *uap56*.

Por fim, confirmamos através de ensaio de placas de lise (*plaque assay*, item 5.1.19) com os parasitas da linhagem cKOuap56 que o gene *uap56* tem função essencial no parasita e portando os dados observados para a superexpressão são resultado direto da inativação de TgUAP56 endógena. Neste caso, os parasitas foram incubados com meio DMEM e rapamicina 50 nM por 24 horas, sendo então substituído por meio DMEM sem a droga e os parasitas foram mantidos em cultivo por 7 dias. No fim, não foi verificado crescimento de placas de lise com parasitas da linhagem cKOuap56, provavelmente pelo fenótipo letal causado pelo nocaute de *uap56* (Figura 23 C).

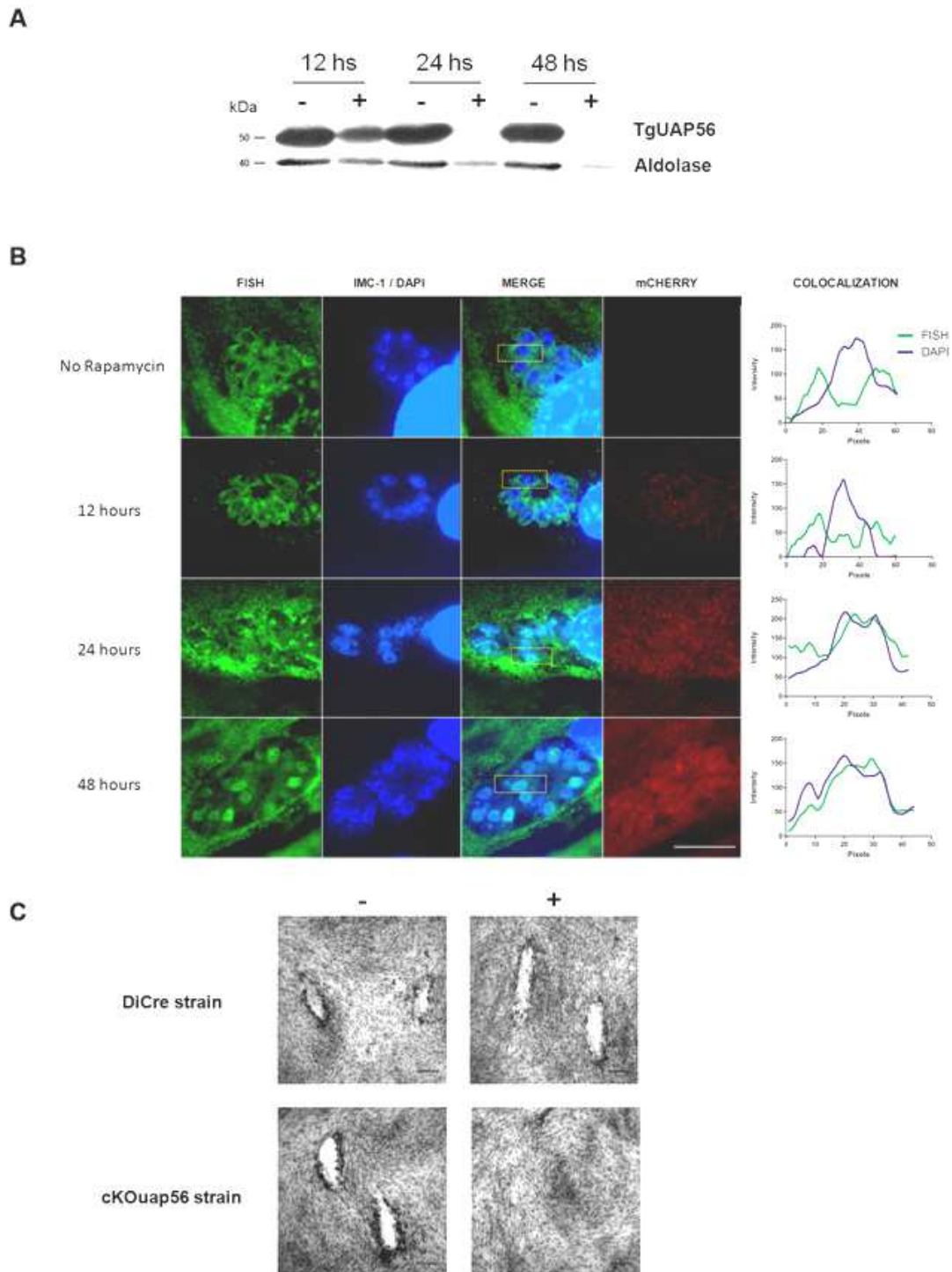


Figura 23. Análise de distribuição de mRNAs após nocaute condicionado do gene *uap56* induzido por rapamicina em parasitas da linhagem clonal cKOuap56. A indução do nocaute ocorreu após incubação com 50 nM de rapamicina e os parasitas foram observados em diferentes tempos, conforme indicado. Neste ensaio foi utilizado anticorpo anti-TgIMC1 (*inner membrane complex*) (Tabela 4), em azul, para observar a possível mudança de morfologia dos parasitas. A. Análise da quantidade de proteína TgUAP56 ao longo da incubação com rapamicina. -, não induzido. +, induzido por rapamicina. Foram utilizados os anticorpos anti-TcSub2 para a detecção da proteína TgUAP56 e anti-Aldolase para a detecção da proteína constitutiva aldolase (Tabela 4), sendo controle de quantidade de proteína aplicada em cada poço neste experimento. B. Distribuição de mRNAs em

parasitas da linhagem cKOuap56 após incubação com 50 nM de rapamicina. Os mRNAs poliadenilados foram detectados por hibridação *in situ* seguido de imunofluorescência (FISH) utilizando sonda oligodT conjugada com digoxigenina, detectada por anticorpo primário anti-digoxigenina (Tabela 4), em verde. (FISH) Localização dos mRNAs. IMC/DAPI: Marcação da superfície dos parasitas bem como da cromatina nuclear e de apicoplasto, em azul. *mcherry*: proteína usada como repórter que emite fluorescência em comprimento de onda de 600 nm. Merge: Sobreposição de imagens. *Colocalization*: Análise de colocalização de mRNAs (FISH) e núcleo (DAPI) foi realizada utilizando ImageJ (Version 1.6, National Institutes of Health, Bethesda, MD) para cada parasita, destacado em amarelo na região de sobreposição das imagens (merge). Barra = 10 µm. C. Ensaio de placas de lise (*plaque assay*) de linhagem DiCre- $\Delta ku80$ (controle) e cKOuap56 em células HFF na presença de 50 nM rapamicina por 7 dias.

5.3.3.3 Análise do efeito do nocaute do gene *uap56* no processamento por *splicing* em alguns mRNAs

Como mostrado anteriormente, após a indução do nocaute de *uap56* foi observado que a quantidade de proteína diminui drasticamente e os mRNAs poliadenilados acumulam no núcleo após 24 horas, sendo mais proeminente após 48 horas. A próxima etapa foi verificar se a proteína TgUAP56 está envolvida diretamente com o processamento de alguns mRNAs por *splicing* analisando alguns mRNAs abundantes e expressos constitutivamente em *T. gondii*. As análises foram feitas após 24 ou 48 horas da indução do nocaute *uap56*. Como a indução do nocaute ocorre por rapamicina, esses alvos foram também analisados na cepa parental, DiCre- $\Delta ku80$, para verificar se a droga poderia alterar algo que não fosse específico de *uap56*, como controle do experimento.

A escolha dos alvos foi de acordo com Suvorova e colaboradores (SUVOROVA *et al.*, 2013) que observou mudanças na quantidade de mRNA processado e pré-mRNA para cada alvo na ausência de TgRRM1, um fator relacionado a *splicing* em *T. gondii*. Para isso estes alvos foram analisados em parasitas da cepa DiCre- $\Delta ku80$ e da linhagem cKOuap56 foram analisados na presença ou ausência de 50 nM de rapamicina por 24 e 48 horas. O RNA total foi extraído para a obtenção posterior de cDNAs, conforme item 5.1.18. A partir do cDNA foram realizados os ensaios semi-quantitativos por PCR conforme item 5.2.2.2, utilizando os *primers* da Tabela 10. Para o desenho dos *primers* foram escolhidas regiões entre éxons que garantissem a presença de íntrons na região a ser amplificada e que possibilitassem a observação da diferença de tamanho entre

os produtos de PCR relativos ao pré-mRNA e o mRNA processado de cada alvo por gel de agarose 1%. Como referência para o tamanho esperado de pré-mRNA foi utilizado o DNA genômico, que possui introns. Esta análise preliminar foi feita através de PCR com 35 ciclos de amplificação e para todos os casos não foi observado maior quantidade de pré-mRNA quando comparado ao respectivo mRNA processado. A próxima etapa será confirmar estes resultados por PCR utilizando menos ciclos, e tubulina como normalizador para uma idéia da quantificação dos pré-mRNAs. Os resultados podem ser observados na figura 24.

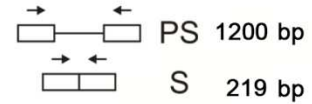
A

Selected genes	Pre-Spliced (PS)	Spliced (S)
PolII p8.2	1200 bp	219 bp
PolII p19	918 bp	375 bp
PolII p23	850 bp	538 bp
IMC1	919 bp	530 bp
IMC15	1113 bp	442 bp
IMC5	780 bp	356 bp

B

PolII p8.2 analysis

Pre-Spliced x Spliced mRNA



C

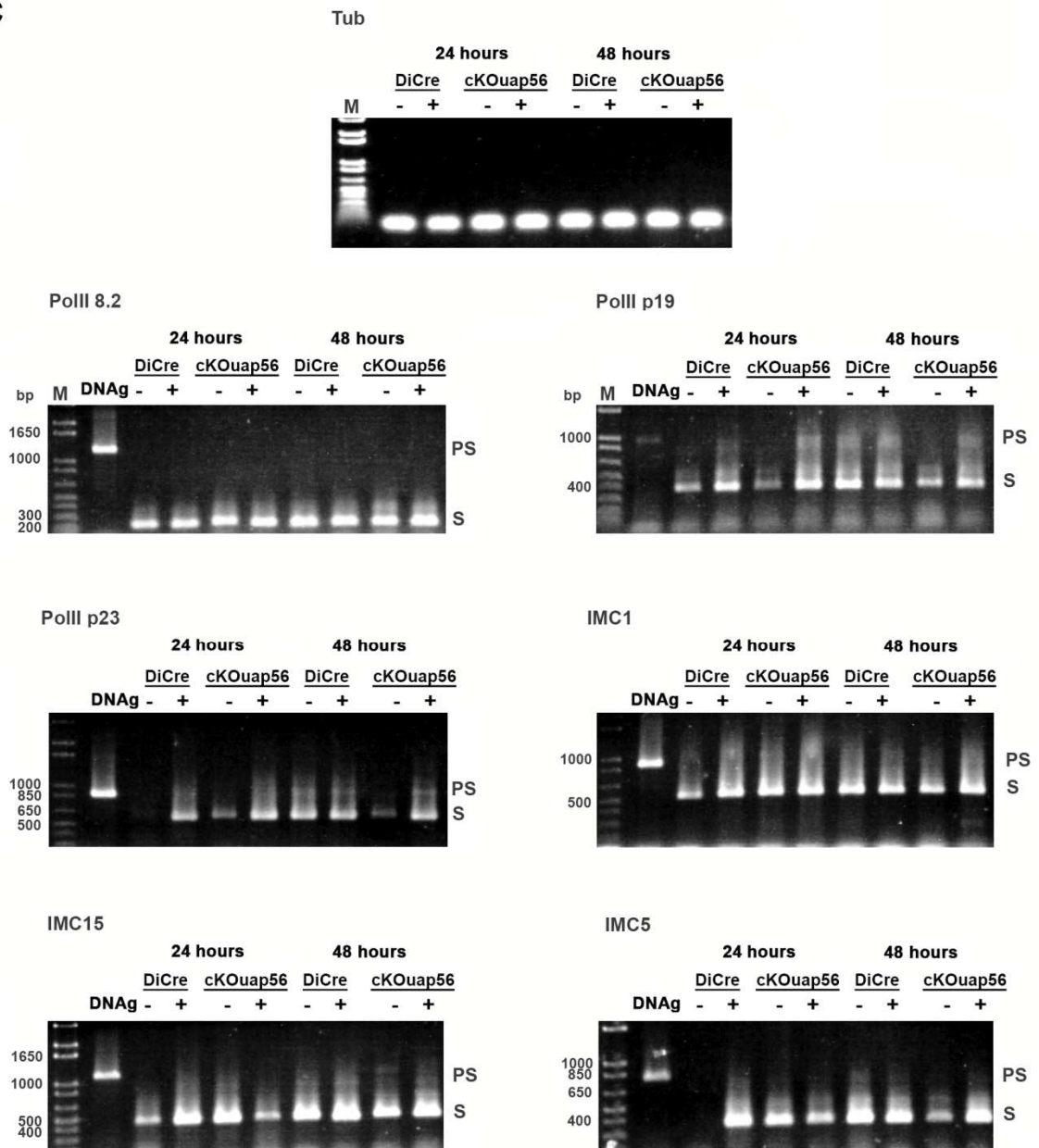


Figura 24. Análise de processamento de alguns mRNAs de *T. gondii* após 24 e 48 horas da indução do nocaute do gene *uap56*. A. Tabela indicando o nome dos *primers* utilizados e os tamanhos esperados para os produtos de PCR de cada alvo, relativos ao tamanho do pré-mRNA

(*Pre-spliced*, PS) e do mRNA processado (*Spliced*, PS). B. Exemplo de resultados esperados para pré-mRNA (*Pre-spliced*, PS) e do mRNA processado (*Spliced*, PS) respectivo ao gene *RNA polymerase II p8.2 subunit* (TGME49_217560). C. Resultados PCR a partir de cDNA de cada alvo indicado na tabela A. RNA total foi purificado de parasitas da linhagem DiCre (DiCre- $\Delta ku80$) e cKOuap56 na ausência ou presença de 50 nM de rapamicina por 24 ou 48 horas. A partir desses RNAs foram obtidos os cDNAs que foram analisados por PCR utilizando os *primers* listados na Tabela 10, conforme item 5.2.2.2. Os seguintes alvos foram analisados: *RNA polymerase II p8.2 subunit* (TGME49_217560), *RNA polymerase II p19 subunit* (TGME49_271300), *RNA polymerase II p23 subunit* (TGME49_240590), IMC1 (TGME49_231640), IMC15 (TGME49_275670), IMC5 (TGME49_224530). Tubulin was used as loading control (DALMASSO *et al.*, 2009). PS: mRNAs não-processados (PS, *pre-spliced*). S: processados (S, *spliced*). DNAg: DNA genômico de parasitas da linhagem cKOuap56. M: 1 kb Plus DNA (Invitrogen # 10787-018).

5.3.4 Monitoramento de mRNAs em *T. gondii* pelo sistema MS2

5.3.4.1 Análise de localização de proteína NLS-MS2-YFP-DD e mRNAs Dip13-SL

Foi verificado que MS2-YFP-DD e NLS-MS2-YFP-DD estão em locais diferentes no parasita após a incubação com 0,1 μ M de Shld-1 por 24 horas. O sinal de localização nuclear (NLS) garantiu a localização nuclear da proteína NLS-MS2-YFP-DD, como pode ser observado na figura 25 A. As proteínas foram vistas diretamente em microscopia de fluorescência devido à fusão de YFP na proteína MS2, sem necessidade de utilização de anticorpo.

Na linhagem MS2 contendo NLS-MS2-YFP-DD foi confirmada também a presença do mRNA Dip13-SL no citoplasma do parasita por ensaio de FISH, utilizando sonda MS2 conforme item 5.2.4. Foi verificado que NLS-MS2-YFP-DD reconheceu os *stem-loops* dos mRNAs Dip13-SL sendo que, mesmo na presença de NLS, a proteína se encontra no citoplasma na mesma localização de mRNAs de Dip13 (Figura 25 B).

Estes resultados mostram que a ferramenta MS2 é funcional em *T. gondii* e permite a tentativa de testes futuros para experimentos em média/larga-escala de varredura de drogas que inibam a exportação de mRNAs, visto que se trata de uma via essencial ao parasita. Levando em consideração que a retirada do gene *uap56* resulta em inibição de exportação de mRNAs de modo geral, o nocaute condicionado deste gene será utilizado como prova de conceito do sistema, para a

confirmação da mudança de localização da proteína. A hipótese é que após a indução do nocaute de *uap56*, a proteína fica acumulada no núcleo como consequência do acúmulo nuclear do mRNAs de Dip13-SL ao longo do tempo, como proposto na figura 15.

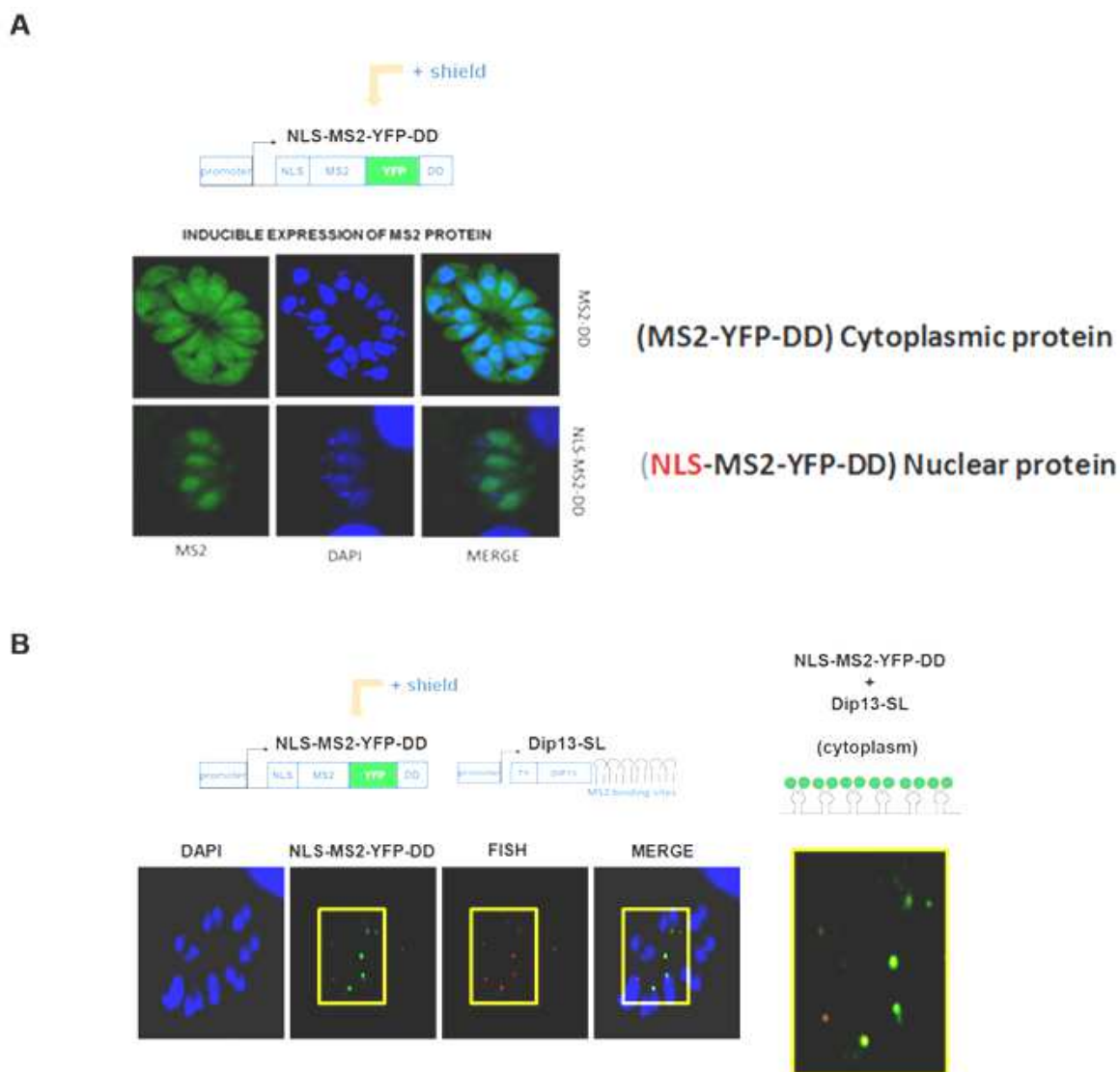


Figura 25. Análise do sistema MS2 em *T. gondii*. A. Localização subcelular de proteína MS2-YFP-DD (sem NLS) e NLS-MS2-YFP-DD após incubação com 0,1 μ M de Shld-1 por 24 horas. B. Localização subcelular de NLS-MS2-YFP-DD e mRNAs de Dip13 contendo 24 repetições de *stem-loops* (Dip13-SL). As proteínas MS2-YFP-DD e NLS-MS2-YFP-DD foram detectadas por fluorescência, em comprimento de onda de 488 nm. Os mRNAs de Dip13-SL foram detectados por hibridação *in situ* seguido de imunofluorescência (FISH) utilizando uma sonda MS2, conforme item 5.2.4. A sonda foi detectada por anticorpo primário anti-digoxigenina (Tabela 4). (FISH) Localização dos mRNAs, em vermelho. DAPI: Marcação da cromatina nuclear e de apicoplasto, em azul. Merge: Sobreposição de imagens. NLS (*nuclear localization signal*), YFP (*eukaryotic yellow fluorescent protein*) e DD, domínio de desestabilização - dFKBP12.

Para a prova de conceito do sistema, o próximo passo foi então a obtenção de uma linhagem MS2 compatível com o nocaute condicionado do gene *uap56*. Para isso foi escolhida a linhagem DiCre-HX, que possui a sequência de DiCre integrada de modo randômico no genoma de *T. gondii* e não possui o gene *ku80*, pois este foi substituído pelo gene *hxgprt*. A estratégia escolhida foi transfectar essa linhagem com um único vetor linearizado contendo as sequências NLS-MS2-YFP-DD e Dip13-SL, bem como as regiões 5' e 3' UTR de *ku80* para então substituir o gene *hxgprt* pelo cassete MS2-Dip13-SL no locus de *ku80*, por recombinação homóloga. A substituição do gene *hxgprt* pelo cassete MS2-Dip13-SL foi estratégica, pois o vetor para nocaute condicionado de *uap56* que será futuramente transfectado nesta linhagem possui resistência para HXGPRT.

Após a seleção negativa para HXGPRT, conforme tabela 6, e clonagem dos parasitas foi possível a obtenção da linhagem DiCre-MS2 e a integração do cassete MS2-Dip13-SL no locus de *ku80* foi verificada por reação de PCR utilizando os *primers* da Tabela 12. Na figura 26 estão os resultados da análise dos produtos de PCR por eletroforese em gel de agarose 1% durante diferentes tempos de exposição em luz ultravioleta. Os resultados 2 e 4 são referentes à linhagem DiCre-MS2 e confirmam a integração do cassete MS2-Dip13-SL no locus de *ku80*. No caso do resultado relativo ao número 2 foi utilizado o *primer* 766, que hibridiza em uma região de 3' UTR do genoma do parasita, externa à região que está presente no cassete MS2-Dip13-SL. Para o resultado relativo ao número 4 foi utilizado o *primer* 946 que hibridiza em uma região de 5' UTR do genoma do parasita, externo ao cassete integrado. Em ambos os casos estes *primers* foram pareados com *primers* que hibridizam em regiões presentes no cassete (766/811 e 946/118) e os tamanhos dos produtos de PCR foram de acordo com o esperado. O resultado relativo ao número 3 não mostrou amplificação provavelmente pelo tempo de extensão utilizado de 6 minutos. O fragmento esperado deveria apresentar um tamanho de 6,43 kb o ideal seria ter repetido com o tempo de 7 minutos de extensão. Como os resultados 2 e 4 confirmaram a integração, esta PCR relativa ao número 3 não foi repetida. Os resultados 1 e 5 foram obtidos com o mesmo par de *primer* (811/940), onde ambos hibridizam em regiões presentes no cassete e foram utilizados como controle positivo das reações de PCR. O resultado 1 é referente à linhagem DiCre-MS2

enquanto o resultado 5 é para o vetor circular MS2-*ku80* (Figura 17), que após a linearização foi utilizado para transfecção dos parasitas.

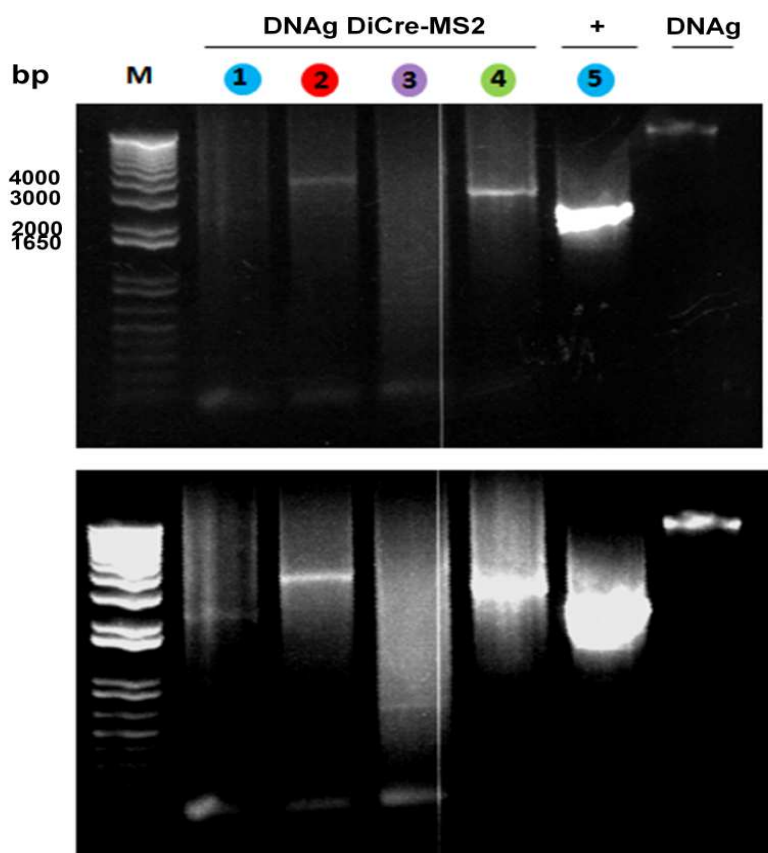
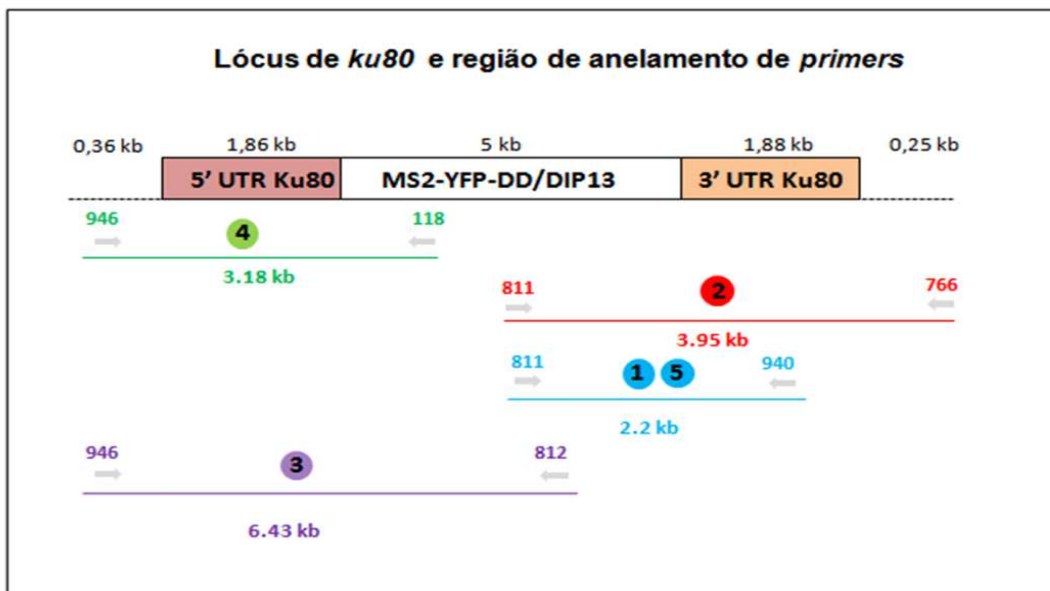


Figura 26. Análise da integração do cassete MS2-Dip13-SL no lócus de *ku80* em parasitas da linhagem DiCre-MS2. Na parte de cima está esquematizado o lócus de *ku80*, com o cassete MS2-Dip13-SL integrado. Os *primers* utilizados estão indicados por números e as respectivas sequências

estão na Tabela 12. O tamanho do produto de PCR esperado para cada primer está indicado em kb (quilobase). DNAg: DNA genômico da linhagem DiCre-MS2. +: plasmídeo não linearizado, utilizado como controle positivo na reação de PCR.

Após a confirmação da obtenção da linhagem DiCre-MS2 foi verificada a presença da proteína NLS-MS2-YFP-DD através de estabilização com Shld-1 em diferentes concentrações por 24 horas em meio de cultura. Ao contrário do obtido com a linhagem MS2, na presença de 0,1 μ M só foi possível visualizar a proteína utilizando o anticorpo anti-GFP (Tabela 4), que também reconhece a proteína YFP presente na proteína fusionada. Uma explicação é que a região 3' UTR de *ku80*, agora presente no mRNA MS2-Dip13-SL, possa estar interferindo com a expressão da proteína, sendo necessária a visualização desta apenas por imunofluorescência para a amplificação do sinal. Nessa concentração de Shld-1 a proteína tem distribuição citoplasmática, assim como na linhagem MS2, provavelmente pela interação com os mRNAs citoplasmáticos de Dip13-SL (Figura 27). Já na concentração de 1 μ M de Shld-1 não é necessária a utilização de anti-GFP para visualizar a proteína, que possui localização nuclear. Esse acúmulo nuclear pode ser o reflexo do aumento da quantidade desta proteína pelo aumento da quantidade do estabilizador Shld-1, visto que a proteína originalmente é nuclear devido ao NLS. Este resultado, embora não mostrado, foi obtido também para a linhagem MS2, onde na presença de 1 μ M de Shld-1 há acúmulo nuclear da proteína.

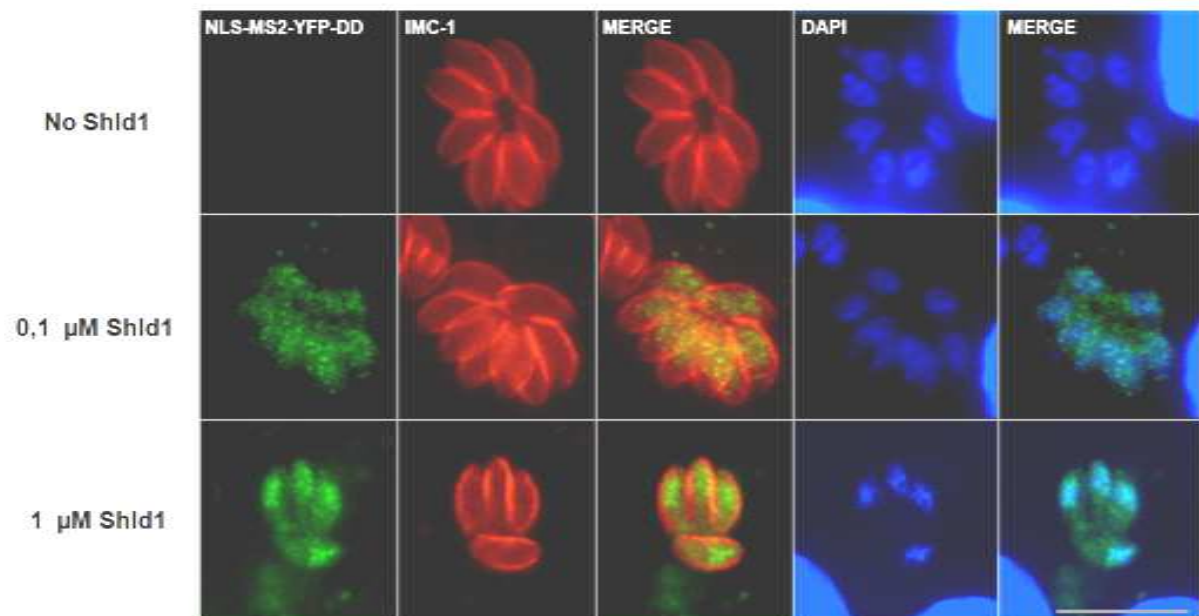


Figura 27. Localização de NLS-MS2-YFP-DD após indução por 24 horas com diferentes concentrações de Shld-1. Para a detecção de YFP a 0,1 μM de Shld-1 foi utilizado o anticorpo monoclonal anti-GFP (Tabela 4). Para 1 μM de Shld a proteína foi visualizada diretamente por fluorescência, sem incubação com anti-GFP. Para delimitação dos parasitas foi utilizado o anticorpo anti-TgIMC1 (*inner membrane complex*) que reconhece uma proteína que está presente na membrana celular. NLS-MS2-YFP-DD: Localização da proteína NLS-MS2-YFP-DD, em verde. DAPI: Marcação da cromatina nuclear e de apicoplasto, em azul. Merge: Sobreposição de imagens. Barra: 10 μm .

Para a confirmação da presença de mRNAs de Dip13-SL nesta linhagem foi realizado o ensaio de FISH utilizando a sonda MS2, conforme item 5.2.4. Como o anti-GFP (para detecção de NLS-MS2-YFP-DD) e anti-digoxigenina (para detecção dos mRNAs Dip13-SL) foram ambos produzidos em camundongo (ver Tabela 4), não foi possível o ensaio de colocalização entre eles, pois o anticorpo secundário seria o mesmo. Pelo fato da proteína poder ser visualizada diretamente por fluorescência na presença de 1 μM de Shld-1, foi realizado o ensaio tanto na ausência quanto na presença desta concentração de Shld-1 por 24 horas e o resultado pode ser visualizado na figura 28. Na ausência de Shld-1 foi observado que os mRNAs Dip13-SL apresentam localização citoplasmática, de acordo com o resultado obtido para a linhagem MS2, o que explica a distribuição citoplasmática também da proteína NLS-MS2-YFP-DD em 0,1 μM de Shld-1 da figura 27. Já na presença de 1 μM de Shld-1 foi observada mudança da distribuição dos mRNAs Dip13-SL, sendo possível observar alguns pontos ao redor do núcleo. Embora seja especulação do resultado, a presença de grande quantidade de proteína NLS-MS2-

YFP-DD no núcleo pode ter sequestrado os mRNAs Dip13-SL, resultando nessa diferença de localização destes mRNAs.

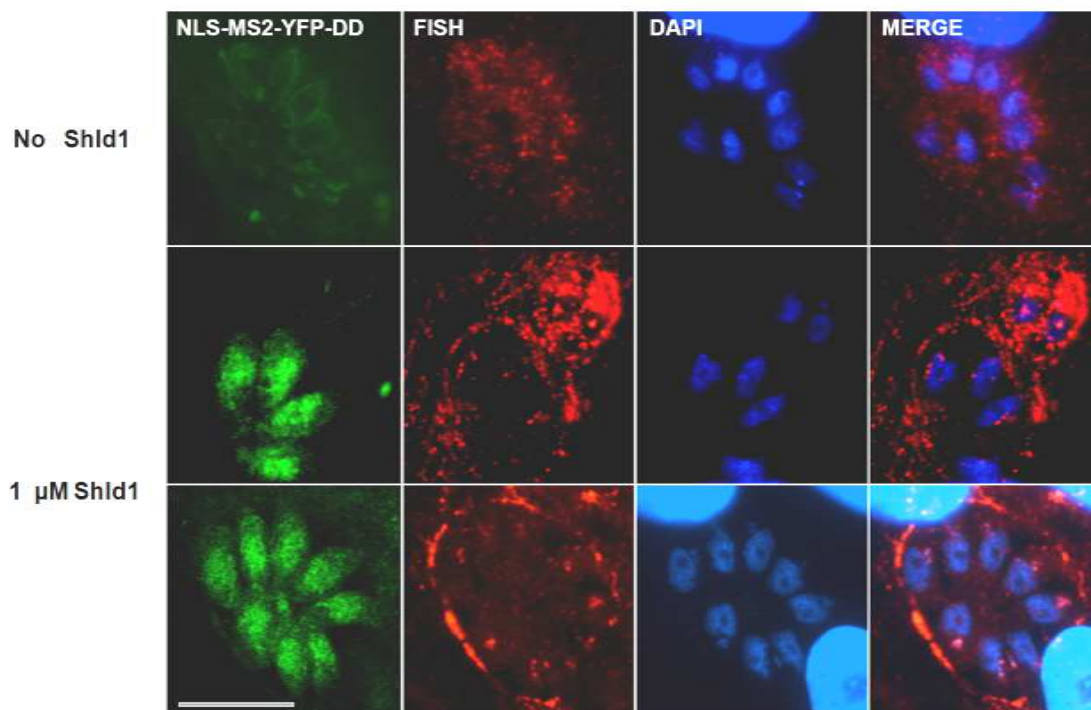


Figura 28. Localização de mRNAs Dip13-SL e proteína NLS-MS2-YFP-DD na ausência ou presença de 1 μ M de Shld-1 por 24 horas. NLS-MS2-YFP-DD: Localização da proteína NLS-MS2-YFP-DD, em verde. FISH: Localização de mRNAs de Dip13-SL por FISH utilizando sonda MS2, conforme item 5.2.4, em vermelho. DAPI: Marcação da cromatina nuclear e de apicoplasto, em azul. Merge: Sobreposição de imagens. Barra: 10 μ m.

Ainda em relação a esta linhagem DiCre-MS2 foi verificada a presença da proteína Dip13, traduzida a partir do mRNA de Dip13-SL, que possui na região *downstream* a sequência nucleotídica de Ty, o que resulta numa etiqueta Ty fusionada na porção C-terminal da proteína (Ty-Dip13). A detecção foi feita por imunofluorescência utilizando o anticorpo monoclonal anti-Ty (Tabela 4) e a localização da proteína é citoplasmática. O experimento então foi realizado da mesma forma que o anterior: ausência ou incubação com Shld-1 1 μ M por 24 horas. Um resultado interessante foi que na presença da proteína nuclear NLS-MS2-YFP-DD após indução com Shld-1 1 μ M não foi observada a proteína Ty-Dip13. Este resultado pode ser uma consequência do sequestro de mRNAs Dip13-SL pelo excesso de proteína nuclear NLS-MS2-YFP-DD, não havendo exportação desses

mRNAs e conseqüentemente ausência de tradução de proteína Ty-Dip13 nestes parasitas. O resultado pode ser observado na figura 29.

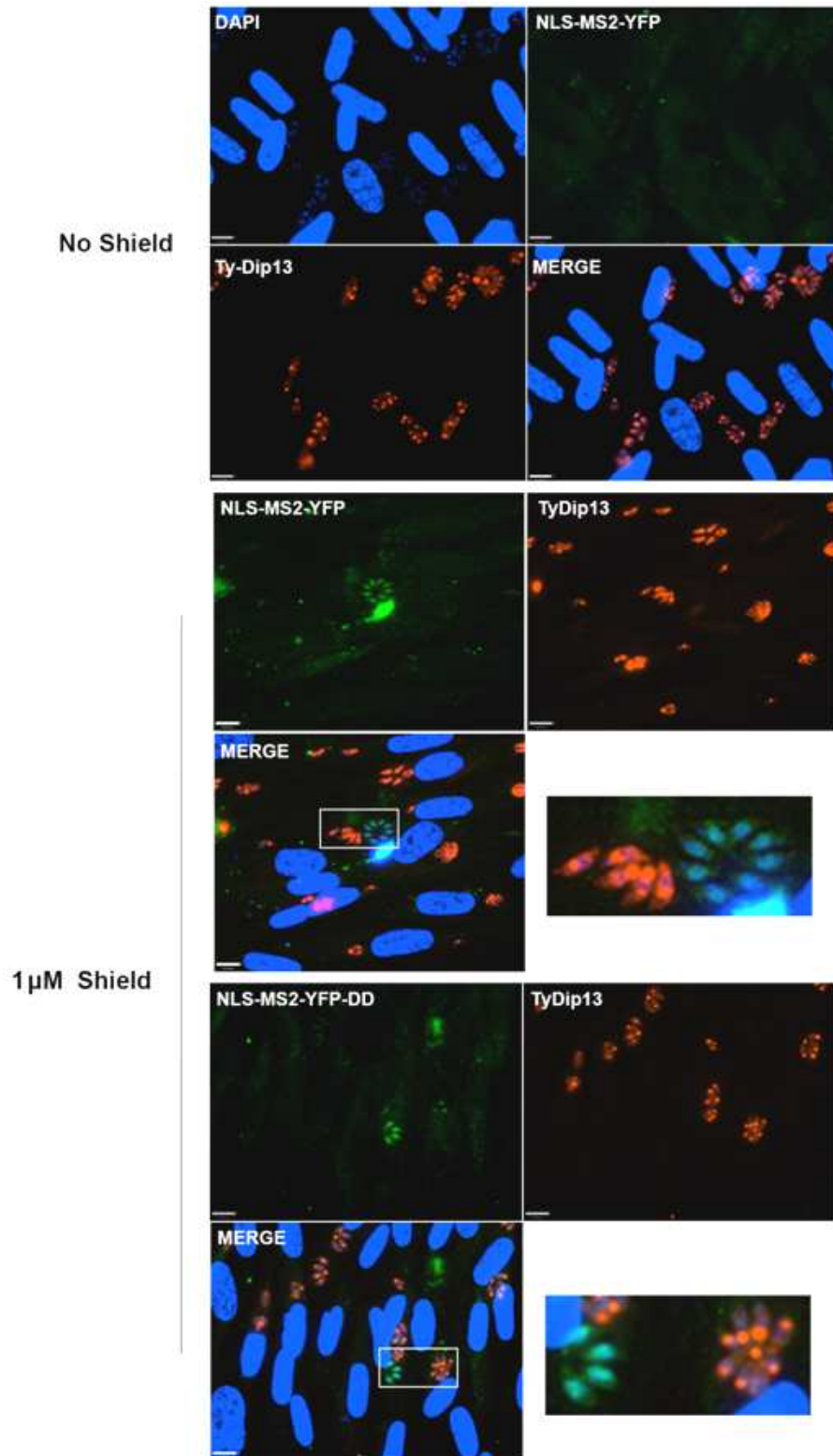


Figura 29. Localização de proteína Ty-Dip13 na linhagem DiCre-MS2 por imunofluorescência. A. proteína NLS-MS2-YFP-DD pode ser observada por fluorescência quando na presença de 1 μM de

Shld-1 após 24 horas. NLS-MS2-YFP-DD: Proteína NLS-MS2-YFP-DD, em verde. Ty-Dip13: Proteína Dip13 fusionada à etiqueta Ty, em vermelho. DAPI: Marcação da cromatina nuclear e de apicoplasto, em azul. Merge: Sobreposição de imagens. Barra: 10 µm.

5.4 Discussão

Em eucariotos, o processo pelo qual o mRNA transcrito e processado no núcleo é exportado para o citoplasma está fisicamente separado pela membrana do envelope nuclear. Este transporte ocorre através de interação de proteínas adaptadoras e de ligação ao mRNA resultando em mRNPs que atingem o citoplasma, onde ocorre a tradução. A exportação de mRNAs é um ponto importante para a expressão gênica e é bem estudada em organismos-modelo sendo que na maioria dos casos envolve proteínas do complexo THO/TREX como UAP56/Sub2, bem como proteínas adaptadoras (revisado por (MULLER-MCNICOLL e NEUGEBAUER, 2013)). No entanto o mecanismo de exportação de mRNAs ainda não é bem compreendido em eucariotos que divergiram na base da árvore filogenética, incluindo parasitas do grupo Kinetoplastida e Apicomplexa (SERPELONI *et al.*, 2011). Alguns estudos mostram que eventos pós-transcricionais são importantes para o desenvolvimento de parasitas do Filo Apicomplexa, mas ainda não há estudos sobre a exportação de mRNAs nestes eucariotos. Há apenas um trabalho relacionado com a proteína ortóloga de UAP56/Sub2 de *P. falciparum*, PfU52, caracterizada como RNA helicase e ATPase dependente de RNA (SHANKAR, PRADHAN e TUTEJA, 2008), no entanto a relação com a exportação de mRNA ainda precisa ser estudada.

Neste presente estudo foram utilizadas ferramentas bem estabelecidas de genética reversa (ANDENMATTEN *et al.*, 2013; HERM-GOTZ *et al.*, 2007) para a análise da função da proteína ortóloga de UAP56/Sub2 de *T. gondii*, como modelo do grupo Apicomplexa. Primeiramente foi confirmada a conservação da proteína UAP56 através de análises de filogenia de eucariotos e de sequência de proteínas ortólogas. A proteína foi nomeada de TgUAP56 e apresenta os dois domínios típicos de helicases com alta similaridade de sequência em relação as sequências das proteínas ortólogas de outros eucariotos. A proteína mais similar é a PfU52, ortóloga

de *P. falciparum* (~80 % de similaridade) enquanto as mais divergentes são TcSub2 e TbSub2 (TrypSub2) (~60 % similares) e foi observado que sequências referentes a UAP56/Sub2 nos grupos Apicomplexa e Kinetoplastida agruparam juntas, mostrando que as sequências estão intimamente relacionadas. Uma vez que TbSub2 é um componente da via de exportação de mRNAs em *T. brucei* (SERPELONI *et al.*, 2011) seria razoável supor que TgUAP56 tivesse um papel nesta via em *T. gondii*. No entanto, esta relação encontrada na análise de filogenia não é uma garantia de conservação funcional e por isso, foi iniciada a análise funcional desta proteína.

Foi observado que TgUAP56 é exclusivamente nuclear, localizada por todo o núcleo, assim como em outros eucariotos (GATFIELD *et al.*, 2001; SAHNI, WANG e ALEXIS, 2010; SERPELONI *et al.*, 2011) e a superexpressão é letal para *T. gondii*. Após 12 horas da superexpressão da proteína há um acúmulo de TgUAP56 em uma região específica do núcleo, aparentemente perto do nucléolo, com diminuição do sinal geral da proteína no nucleoplasma. Em células de humanos a região perinucleolar está relacionada ativamente com maturação e modificação de proteínas relacionadas a *splicing* e transporte de mRNAs. Na ausência de transcrição por RNA Polimerase II há a redistribuição de proteínas relacionadas ao metabolismo mRNAs, originárias de *speckles*, para regiões perinucleolares denominadas “*perinucleolar caps*” (PNC) (HUANG *et al.*, 1998), uma região do nucléolo que foi descrita durante a caracterização de hnRNPI/PTB (*polypyrimidine tract binding*) (revisto por Pollock e Huang, 2010 (POLLOCK e HUANG, 2010)). Dentre estas proteínas foram encontradas as proteínas SR que além de acoplar transcrição de mRNAs com *splicing* atuam como adaptadores para a exportação de mRNAs e após a inibição da transcrição de mRNAs permanecem acumuladas e hipofosforiladas nesta região (ANDERSEN *et al.*, 2005; BUBULYA *et al.*, 2004; SHAV-TAL *et al.*, 2005). Em outro trabalho, a parada geral do metabolismo celular na presença de luz ultravioleta ou danos químicos ao DNA também causou a realocação de diversas proteínas relacionadas com metabolismo de mRNA para os *perinucleolar caps* entre elas o receptor majoritário de exportação de mRNAs, TAP/Mex67 (SAKASHITA e ENDO, 2010), que interage diretamente com as proteínas SRs quando desfosforiladas durante a exportação de mRNAs que não possuem introns (HUANG *et al.*, 2003; HUANG, YARIO e STEITZ, 2004; LAI e

TARN, 2004). Embora pareça ser uma região relacionada com metabolismo de mRNAs, este resultado de realocização de TgUAP56 em *T. gondii* ainda precisa ser estudado.

Esta realocização de TgUAP56 na periferia nucleolar após a superexpressão é mais evidente após 24 horas e há colocalização de mRNAs poliadenilados nestas regiões, podendo ser resultado de sequestro de mRNAs pelo excesso desta proteína nesta região. No entanto, após 48 horas, foi notado o acúmulo nuclear da maioria dos mRNAs resultando em morte celular. Esse resultado indica que foi obtido um fenótipo dominante negativo, provavelmente por bloquear a exportação de mRNAs demonstrando que TgUAP56 deve ser uma proteína essencial envolvida nesta via, assim como em outros eucariotos. A superexpressão de UAP56 é letal em todas as etapas de crescimento de *C. elegans* devido ao bloqueio da exportação de mRNAs (MACMORRIS, BROCKER e BLUMENTHAL, 2003). Em células humanas e fungos o excesso de UAP56/Sub2 bloqueia a exportação de mRNAs de modo dominante negativo, sequestrando outros fatores da via e assim impedindo o recrutamento destes para os mRNAs processados, resultando em retenção nuclear dos mRNAs poliadenilados (LUO *et al.*, 2001; STRASSER e HURT, 2001).

Para confirmar se o acúmulo de mRNAs após a superexpressão é um efeito diretamente relacionado com TgUAP56 foi realizado o nocaute condicionado do gene *uap56* através da ferramenta DiCre em *T. gondii* (ANDENMATTEN *et al.*, 2013). A eficiência do nocaute foi confirmada e após 24 horas da indução do nocaute já foi observado acúmulo de mRNAs em alguns parasitas, sendo ainda mais evidente após 48 horas, em praticamente todos os parasitas, também resultando em morte celular. Este resultado mostrou que TgUAP56 está diretamente relacionada com a via de exportação de mRNAs em *T. gondii*, e está de acordo com o encontrado na literatura. Em fungos e células de *Drosophila melanogaster* a diminuição de UAP56/Sub2 resultou em parada de crescimento celular e acúmulo geral de mRNAs poliadenilados no núcleo (GATFIELD *et al.*, 2001; STRASSER e HURT, 2001). Em *T. brucei*, pertencente ao grupo Kinetoplastida, a diminuição de TbSub2 também causou acúmulo de mRNAs, resultando em diminuição drástica da tradução, resultando em morte dos parasitas (ver capítulo 2, (SERPELONI *et al.*, 2011)).

Foi também investigado se o nocaute de *uap56* poderia afetar o processamento de mRNAs por *splicing* uma vez que proteínas ortólogas de TgUAP56 foram originalmente associadas à maquinaria de processamento (revisado por Linder and Stutz, 2001 (LINDER e STUTZ, 2001)). Em *S. cerevisiae*, Sub2 pode estar envolvida em montagem do spliceossomo, eventos de *splicing* bem como na exportação de mRNAs (FLECKNER *et al.*, 1997; LIBRI *et al.*, 2001), (JENSEN *et al.*, 2001; STRASSER e HURT, 2001; THAKURTA *et al.*, 2007) e em *P. falciparum* foi observado que a proteína ortóloga purificada PfU52 tem atividade de *splicing in vitro* embora não tenha sido confirmada esta atividade *in vivo* (SHANKAR, PRADHAN e TUTEJA, 2008). Para isso nos baseamos em ferramentas utilizadas para a caracterização do fator de *splicing* em *T. gondii*, TgRRM1 onde foram selecionados alguns genes cujos mRNAs não são processados quando esta proteína está mutada (SUVOROVA *et al.*, 2013). Na ausência da proteína TgUAP56 não foram observadas modificações de processamento dos mesmos alvos embora tenha sido observado o bloqueio na exportação de mRNAs poliadenilados de modo geral. Embora seja necessário avaliar o processamento de mRNAs totais, a hipótese atual é que TgUAP56 possa atuar na liberação de mRNAs processados para as proteínas adaptadoras, para então serem exportados para o citoplasma. Essa hipótese corrobora com estudos anteriores em organismos-modelo que indicam que UAP56 e Sub2 tem um papel essencial na exportação de mRNAs (GATFIELD *et al.*, 2001; JENSEN *et al.*, 2001; LUO *et al.*, 2001; STRASSER e HURT, 2001) embora tenha sido originalmente identificada como relacionada ao *splicing* (revisado por (LINDER e STUTZ, 2001)).

No fungo *S. pombe*, por exemplo, Sub2 não é recrutada para o *splicing*, mas o papel na exportação de mRNAs se mantém conservada (THAKURTA *et al.*, 2007). Em células de *Drosophila*, HEL/UAP56 se associa preferencialmente a mRNAs processados e está envolvida com a exportação de mRNAs que não possuam introns. Quando a proteína é diminuída, a exportação de mRNAs é bloqueada mas não é uma consequência da inibição geral do *splicing* (GATFIELD *et al.*, 2001). Em células de humanos, o recrutamento de UAP56 é através de um mecanismo acoplado ao *splicing* no qual a proteína interage diretamente com mRNAs já processados (MASUDA *et al.*, 2005) e um estudo recente mostrou que UAP56 se

liga preferencialmente a mRNAs já processados em relação aos pré-mRNAs, o que pode explicar a interação necessária de UAP56 para a liberação de mRNAs poliadenilados de *speckles* nucleares (DIAS *et al.*, 2010; STECKELBERG e GEHRING, 2014), estabelecendo uma relação funcional entre *splicing* e exportação de mRNAs.

A conservação da função de UAP56 em parasitas de grupos diferentes como Kinetoplastida e Apicomplexa indica que é um componente básico da exportação de mRNAs em parasitas. Uma vez identificada a presença de UAP56 como um componente específico para a via de exportação de mRNAs em metazoários e fungos, a identificação das proteínas ortólogas em tripanossomatídeos (TrypSub2) e em *T. gondii*, TgUAP56, aponta para a existência de uma via especializada de mRNAs também nestes parasitas. No entanto, parasitas do grupo Apicomplexa, entre outros, adquiriram características peculiares em relação ao metabolismo de mRNA durante a evolução e não podemos descartar a necessidade de estudo de outros componentes da via de exportação de mRNAs em parasitas para investigar o quão similar esta via é em relação às vias de metazoários e fungos.

Uma outra vertente deste trabalho foi o desenvolvimento de uma ferramenta para o monitoramento do transporte de mRNAs em *T. gondii* baseado no sistema MS2. Foi verificado que o sistema MS2 é funcional no parasita, sendo possível relacionar a localização da proteína NLS-MS2-YFP-DD com a do mRNA Dip13-SL. No entanto, após a obtenção da linhagem com o sistema MS2 e que permitisse uma análise fenotípica após nocaute de genes (DiCre-MS2) não foi possível a detecção direta da proteína por fluorescência, sendo necessária a utilização do anticorpo para detecção de GFP e conseqüentemente amplificação do sinal. Por isso, esta linhagem não será útil para ensaios em média/larga escala. Ainda que as evidências obtidas indiquem a mobilidade do mRNA Dip13-SL nesta linhagem, a prova de conceito deverá ser realizada através do nocaute condicionado de *uap56*. Caso haja o acúmulo nuclear do mRNA Dip13-SL, e conseqüentemente da proteína NLS-MS2-YFP-DD conforme o esperado, irá ser analisado se este acúmulo ocorre em uma região preferencial no núcleo por imunocitoquímica e microscopia eletrônica, utilizando o anticorpo anti-GFP. Em células de mamíferos é sabido que fatores de exportação estão presentes na vizinhança de *speckles*, que são estruturas

subnucleares contidas em espaços intercromatínicos envolvidas em biogênese de mRNPs, processamento e exportação de mRNAs (LAMOND e SPECTOR, 2003; TENG e WILSON, 2013). Logo, o sistema MS2 também poderá ser utilizado para colocalização com proteínas de *splicing*, como por exemplo TgRRM1 (SUVOROVA *et al.*, 2013) e assim iniciar estudos de verificação da existência destas estruturas também em *T. gondii*.

Já na linhagem MS2 a visualização de NLS-MS2-YFP-DD pode ser diretamente por fluorescência e, após prova de conceito quanto à alteração da localização de MS2 em virtude do bloqueio da exportação de mRNA, esta linhagem anterior pode ser usada para a aplicabilidade do sistema em varredura de drogas, buscando aquelas que inibam especificamente a exportação de mRNA do parasita.

6 **Capítulo 4. Busca de evidências de proteínas associadas a TcSub2 bem como de proteínas relacionadas com a via de exportação de mRNAs em *T. cruzi*.**

Nos Capítulos 2 e 3 foi verificada a função das proteínas ortólogas de Sub2 na exportação de mRNAs em tripanossomatídeos e *T. gondii*. Pelos dados obtidos *in silico* no Capítulo 1, a maioria dos componentes da via de exportação de mRNA não possui alto grau de conservação em protozoários parasitas. A próxima etapa, portanto, foi tentar identificar proteínas presentes na via de exportação através de outros métodos. Neste caso, optamos por análise de interação entre proteínas por duplo-híbrido (Y2H) a partir de uma colaboração com o grupo do Dr. Christian M. Probst e Dr. Henrique Preti, do laboratório de genômica do Instituto Carlos Chagas – ICC (Fiocruz Paraná), baseado no trabalho de análise do interatoma de proteínas ligadoras de RNA em *Trypanosoma cruzi*. Em paralelo foram iniciados ensaios de imunoprecipitação associados à espectrometria de massas para identificar proteínas do complexo que contém TcSub2, visto que é um componente da via em tripanossomatídeos. No ensaio de duplo-híbrido foram verificadas as interações de TcSub2, bem como de proteínas candidatas de *T. cruzi* possivelmente relacionadas com a via de exportação de mRNAs encontradas em nosso trabalho anterior (ver capítulo 1, (SERPELONI *et al.*, 2011)), com um conjunto de 480 proteínas envolvidas em vias funcionais de RNA (ribossomal, *splicing*, tradução, transcrição, transporte, dentre outras classes funcionais) (Preti, em preparação). Tanto nos resultados de duplo-híbrido e de imunoprecipitação de TcSub2, não foi observada uma relação direta entre TcSub2 com o provável receptor de exportação Mex67 ou exportina TcCRM1, e sim, com uma proteína muito similar a Imp3 de outros eucariotos, nomeada de “TcImp3” neste estudo. Além da interação com TcSub2, TcImp3 também interage diretamente com TcCRM1, com a provável proteína ortóloga de Mex67, bem como com outras proteínas provavelmente relacionadas a via de exportação de mRNAs e possíveis fatores de transcrição de diferentes tipos de RNA Polimerases, gerando a hipótese de que esta proteína pode ter uma função não canônica em *T. cruzi*. Em outros eucariotos, Imp3 está relacionada com processamento de pré-rRNAs e com localização preferencial no nucléolo. Em *T. cruzi*, embora não existam evidências funcionais sobre a proteína, identificamos que esta apresenta distribuição citoplasmática e perinuclear. Estes resultados, embora

promissores, ainda necessitam de ensaios funcionais complementares para a confirmação das interações, bem como a investigação da função de TcImp3 em tripanossomatídeos.

6.1 Materiais e Métodos

As soluções e tampões utilizados neste estudo estão detalhados na Tabela 13.

Tabela 13. Soluções e Tampões

Soluções e Tampões	Composição
Solução de bloqueio para <i>Western blot</i>	Tampão PBS, Tween 20 0,05% e leite em pó desnatado 5%
Solução de Ponceau S	Ponceau S (Sigma P-3504) 0,5%; Ácido acético glacial 1%
Tampão de amostra de proteínas	Tris-HCl 40 mM pH 6,8; SDS 1%, β -mercaptoetanol 2,5%; glicerol 6% e azul de bromofenol 0,005%
Tampão de transferência para <i>Western blot</i>	Tris-base 25 mM, Glicina 192 mM, Metanol 20%
Tampão de corrida para SDS-PAGE	Tris-HCl 0,037 M pH 8.4, Glicina 192 mM pH 8.4; SDS 0,1%
Tampão PBS - solução de uso (<i>Phosphate-buffer saline</i>)	KCl 2,7 mM; KH_2PO_4 1,5 mM; $\text{NaHPO}_4 \cdot 7\text{H}_2\text{O}$ 4,3 mM; NaCl 137 mM
Tampão Tris-EDTA-Noct (TE-Noct)	Tris 50 mM pH 8.0, EDTA 10 mM, Noct 1% (Sigma # O8001)
Tampão de digestão (solução ABC)	bi-carbonato de amônio 50 mM (NH_4HCO_3 , ABC) (pH 8.0)
Tampão de descoloração	NH_4HCO_3 25 mM, etanol 50%
Tampão de redução	Ditiotreitol (DTT) 10 mM em solução ABC 50 mM
Tampão de alquilação	Iodoacetamida 55 mM em ABC
Solução de tripsina	Tripsina 12,5 ng/ μL modified sequencing grade, Promega (#cat. V5111)), em solução ABC
Solução de extração	Ácido trifluoroacético (TFA) 3%, acetonitrila 30%

6.1.1 Meios de cultura

Tabela 14. Meio de cultura para *T. cruzi*

Meio de cultura	Composição
LIT (<i>Liver Infusion Tryptose</i>) (CAMARGO, 1964)	Infusão de fígado 5g/L, NaCl 4,4 g/L, KCl 0,4 g/L, Glicose 2,2 g/L, Tryptose 5 g/L, Fosfato básico de sódio 11,56 g/L pH 7.2, Extrato de levedura 15 g/L, Hemina 0,02 g/L, Soro fetal bovino 10%, Penicilina 10.000 U.

6.1.2 Microrganismos

T. cruzi clone Dm28c (GOLDENBERG *et al.*, 1984)

Formas epimastigotas foram mantidas em cultura axênica com meio LIT (CAMARGO, 1964) a 28 °C com passagens a cada quatro dias.

6.1.3 Preparo dos extratos protéicos de *T. cruzi* para a imunoprecipitação de TcSub2

6.1.3.1 Obtenção do extrato protéico para a imunoprecipitação de TcSub2 em tampão de lise com Tris-EDTA e Noct

Embora não mostrados, foram realizados alguns testes de preparo do extrato protéico de *T. cruzi* e foi verificado por cromatografia de gel-filtração que a proteína TcSub2 se apresenta majoritariamente na forma descomplexada, mesmo na presença de tampão com baixa força iônica (tampão hipotônico – HEPES 10 mM pH 7.4, KCl 10 mM, MgCl₂ 1.5 mM (ROBERTS *et al.*, 1998)). Levando estes resultados em consideração, a alternativa para a busca de complexos protéicos de TcSub2 foi a obtenção do extrato protéico através da estabilização prévia destes através de fixação de proteínas por formaldeído (*cross-linking*), segundo Villanova e colaboradores, 2009 (VILLANOVA *et al.*, 2009), revisito por Sutherland e colaboradores, 2008 (SUTHERLAND, TOEWS e KAST, 2008), para então realizar o

ensaio de imunoprecipitação de TcSub2. O detergente SDS do protocolo original foi substituído pelo detergente Noct (n-octyl- β -D-glucoside, Sigma #O8001) a 1%, que é compatível com as análises usando o espectrômetro de massas. O tampão utilizado para obtenção dos extratos protéicos de *T. cruzi* bem como para as lavagens das etapas da imunoprecipitação de TcSub2 foi denominado TE-Noct e a composição está mostrada na Tabela 13.

Algumas etapas foram modificadas resultando em protocolos diferentes do original. As mudanças foram relacionadas com a forma de como o extrato protéico foi obtido ou como as lavagens foram feitas durante a imunoprecipitação de TcSub2. Por questões didáticas os protocolos foram nomeados de A a D, sendo o protocolo D o extrato protéico obtido de acordo com o original (VILLANOVA *et al.*, 2009).

Para a obtenção do extrato protéico de acordo com o protocolo original (D), 1×10^9 parasitas em crescimento exponencial no meio LIT foram coletados por centrifugação (6000 x g, 5 minutos a temperatura ambiente), lavados com PBS e suspensos em meio LIT sem soro fetal bovino, seguido da adição de formaldeído (Sigma # F8775) para concentração final de 1 %. Os parasitas foram então incubados por 15 minutos a 28 °C e novamente coletados por centrifugação, lavados com PBS e suspensos em 1 mL de tampão TE-Noct. A incubação neste tampão foi de 2 horas a 4 °C, com agitação. Após o tempo de incubação o lisado foi submetido a 3 pulsos de sonicações no gelo (100 Joules) por 15 segundos cada. A lise celular foi monitorada por microscopia de luz ou de fluorescência, após coloração da cromatina com DAPI. A amostra obtida foi centrifugada a 20.000 x g por 10 minutos a 4 °C, obtendo então a fração solúvel (sobrenadante) e insolúvel (*pellet*), sendo a fração solúvel de cada extrato protéico utilizado para a imunoprecipitação de TcSub2.

As explicações relativas às mudanças pontuais do protocolo original (D) seguem descritas abaixo e todos os protocolos estão resumidos na tabela 15.

- **Protocolo A: Extrato protéico a partir de parasitas retirados do meio LIT e sem fixação por formaldeído**

O protocolo original exige incubação prévia dos parasitas em meio LIT sem soro fetal bovino (SFB). A retirada do SFB pode interferir no rearranjo de complexos e o resultado não condizer com complexos protéicos presentes em uma situação

normal dos parasitas. Pensando nisso foi também realizada a imunoprecipitação de TcSub2 utilizando um extrato protéico a partir de parasitas retirados diretamente do meio LIT normal (que contém soro fetal bovino) sem fixação por formaldeído. A fixação não foi utilizada neste caso, pois poderia permitir fixação entre proteínas presentes SFB com as proteínas dos parasitas, aumentando a probabilidade de proteínas contaminantes na amostra.

- **Protocolo B. Etapas de lavagens de imunoprecipitação de TcSub2 com TE-Noct-50**

É importante ressaltar que o tampão TE-Noct utilizado para a obtenção do extrato protéico também foi utilizado durante as lavagens da imunoprecipitação de TcSub2. Este tampão não possui sal e, portanto possui baixa força iônica. Como uma alternativa de aumentar a força iônica do tampão utilizado para estas lavagens e com isso diminuir as proteínas contaminantes na amostra final, o extrato protéico relativo ao protocolo A foi também submetido em paralelo ao ensaio de imunoprecipitação de TcSub2 e as lavagens foram realizadas com tampão TE-Noct-50, que possui KCl 50 mM (TE-Noct e KCl 50 mM).

Para este protocolo foi escolhido o extrato protéico obtido pelo protocolo A devido à mínima interferência durante a preparação, pois os parasitas foram apenas retirados do meio de cultura e as proteínas não foram previamente fixadas por formaldeído.

- **Protocolo C: Extrato protéico a partir de parasitas incubados em meio LIT sem soro e sem formaldeído**

O protocolo original (D) exige fixação prévia de proteínas por formaldeído para obtenção do extrato protéico para os ensaios de imunoprecipitação. Esta estabilização de complexos protéicos por formaldeído pode permitir também a fixação entre proteínas que não necessariamente interagem entre si, podendo haver proteínas contaminantes nos complexos protéicos de TcSub2 que estarão presentes na amostra final da imunoprecipitação. Como controle para este experimento foi

obtido um extrato protéico de parasitas incubados previamente em meio LIT sem soro, conforme D, porém sem fixação por formaldeído.

As mudanças relacionadas a cada protocolo foram resumidas na tabela 15.

Tabela 15. Protocolos relacionados a obtenção de extratos protéicos de *T. cruzi* e lavagens referentes aos ensaios de imunoprecipitação de TcSub2. SFB: Soro fetal bovino.

Extrato protéico	Meio LIT	Fixação por formaldeído (<i>cross-linking</i>)	Tampão de lavagem na imunoprecipitação
A	Com SFB	Não	TE-Noct
B	Com SFB	Não	TE-Noct-50
C	Sem SFB	Não	TE-Noct
D	Sem SFB	Sim	TE-Noct

As imunoprecipitações foram realizadas para cada protocolo, em triplicata técnica, utilizando o anticorpo anti-TcSub2 (Tabela 4). Como controle deste ensaio foi utilizado o anticorpo pré-imune, também produzido em coelho. Este ensaio será descrito detalhadamente no item a seguir.

6.1.4 Ensaios de Imunoprecipitação de TcSub2 e preparo das frações eluídas para identificação por espectrometria de massas

Após a obtenção dos extratos protéicos solúveis dos protocolos A, C e D, conforme citado no item acima, foram realizados os ensaios de imunoprecipitação utilizando o anticorpo anti-TcSub2 (Tabela 4) produzido em coelho que reconhece especificamente a proteína TcSub2 em extrato protéico de *T. cruzi* (SERPELONI *et al.*, 2011). Como controle foi utilizado o anticorpo pré-imune, também produzido em coelho e que não reconhece a proteína TcSub2 em extrato protéico de *T. cruzi*. Para cada experimento, 30 µL das *beads* magnéticas (Invitrogen - *Dynabeads® Protein A*, # 10001D) foram bloqueadas previamente em tampão PBS-BSA 4% por 15 minutos

e em seguida 10 µL dos anticorpos (imune e pré-imune) foram incubados com as *beads* em PBS-BSA 4 % por uma hora, em temperatura ambiente com agitação. O excesso de anticorpo foi retirado através de lavagens com o mesmo tampão, seguido de uma lavagem com o tampão TE-Noct, presente nos extratos protéicos. Cada extrato protéico referente a cada protocolo do item 6.1.3.1 (ver Tabela 15) foi incubado com as *beads* ligadas aos anticorpos por duas horas a 4 °C, com agitação, seguido de 4 lavagens, de 1 mL cada, com o tampão TE-Noct, com exceção do protocolo B em que as lavagens foram realizadas com TE-Noct-50, conforme explicado no item 6.1.3.1. As lavagens foram realizadas a 4 °C com agitação. Após as lavagens, os complexos ligados aos anticorpos foram eluídos com 50 µL de glicina 0,2 M pH 2.5. Todas as etapas de imunoprecipitação foram coletadas para análise em *western blot* utilizando o anticorpo anti-TcSub2 (Tabela 4) para a detecção da proteína, conforme item 5.1.10.

Para a tripsinização de proteínas fixadas, presentes nas frações eluídas referentes ao protocolo original (D, item 6.1.3.1) foi necessário desfazer a reação de *cross-linking* com tampão redutor 4X (500 mM Tris-HCl pH 6.8, 8% SDS, 40% glicerol, 20% β-mercaptoetanol, 5 mg/mL azul de bromofenol) e incubação a 99 °C por 20 minutos (KLOCKENBUSCH e KAST, 2010). Para retirar estes reagentes do tampão e preparar a amostra para sequenciamento por espectrômetro de massas as proteínas foram separadas por eletroforese em gel de poliácridamida 10% SDS-PAGE. Após a corrida, o gel foi cortado em pedaços de aproximadamente 1 × 1 mm e desidratados em etanol 100% por 10 minutos a 25 °C. Entre todas as etapas detalhadas a seguir houve a etapa de desidratação por etanol 100% por 10 min a 25 °C, para garantir que a solução incubada posteriormente reidratasse o gel. Os pedaços foram reidratados em tampão de redução e incubados por 60 minutos a 56 °C. O líquido foi retirado e então adicionado o tampão de alquilação por 45 min a 25 °C, no escuro. Os pedaços foram lavados com tampão de digestão por 20 minutos a 25 °C e desidratados novamente. Os pedaços foram então lavados com tampão de digestão por 20 minutos a 25 °C e desidratados novamente, garantindo a retirada do tampão utilizado anteriormente. Os pedaços foram reidratados com solução de tripsina a 4 °C, por 20 minutos. O excesso de solução de tripsina foi retirado e foi adicionado o tampão de digestão para incubação a 37 °C por 16 horas. Para interromper a digestão, a amostra foi acidificada com 2 µL de TFA 100%. Os

peptídeos da matriz do gel foram extraídos após 2 ciclos de incubação com tampão de extração por 10 minutos, agitando vigorosamente a 25 °C. Os pedaços foram desidratados duas vezes com acetonitrila 100% por 10 minutos cada, agitando vigorosamente a 25 °C.

Para as frações eluídas dos protocolos A, B e C foram realizadas as mesmas etapas de redução. Alquilação e tripsinização das proteínas em solução. Todas as soluções utilizadas nessas etapas estão listadas na tabela 13. Após a retirada de acetonitrila, a purificação dos peptídeos foi através de StageTips. Para isso, uma ponteira (2-200 µL) foi preenchida com uma porção de membrana de C18 (Sigma, *Empore SPE Disks C18 # 66883-U*), a qual foi ativada com 100 µL de metanol e lavada com 200 µL de ácido fórmico 0,1% (solução A) por centrifugação a 1.000 x g. A amostra foi então passada na ponteira por centrifugação e lavada com 200 µL de solução A. A eluição foi feita com acetonitrila 80%, ácido fórmico 0,1%. O eluído foi concentrado em speed vac e ressuspenso em uma concentração de 0,5 µg/µL de solução A.

6.1.4.1 Identificação de proteínas por espectrometria de massas

Os peptídeos tripticos foram submetidas à cromatografia líquida de fase reversa (RP), acoplada à espectrometria de massas de alta resolução. Os experimentos foram realizados com um Easy- nLC 1000 (Thermo Scientific, EUA), conectado ao espectrômetro de massas LTQ Orbitrap XL ETD (espectrometria de massas RPT02H PDTIS / Instituto Carlos Chagas – ICC (Fiocruz Paraná)), equipado com uma fonte iônica nanoeletrospray (S/T Phoenix). A separação cromatográfica dos peptídeos ocorreu em um capilar de sílica fundido de 15 centímetros (75 µm de diâmetro interno) empacotada com resina de fase reversa ReproSilPur-C18-AQ 3 µm (Dr. Maisch GmbH, Ammerbuch - Entringen, Alemanha). Os peptídeos foram injetados na coluna com um fluxo de 250 nL/minutos e, subsequentemente, eluídos com 5 a 40% de acetonitrila em 0,1% de ácido fórmico em um gradiente de 120 minutos. O espectrômetro de massas foi operado no modo dependente de dados para alternar automaticamente entre aquisição MS e MS/MS (MS2). Um perfil completo dos espectros MS (de 350-1.650 gama m/z) foram adquiridos no

analisador Orbitrap com resolução $R= 60.000$ em m/z 400 (após a acumulação de um valor-alvo de 1.000.000 no ion *trap* linear). Os dez íons mais intensos foram sequencialmente isolados e fragmentados no ion *trap* linear usando a dissociação induzida por colisão em um valor-alvo de 10.000. Ex-íons alvos selecionados para o MS/MS foram excluídos dinamicamente durante 90 segundos. O tempo total do ciclo foi de aproximadamente 3 segundos. As condições gerais de espectrometria de massas foram: spray de tensão, de 2,4 kV; temperatura no tubo de transferência de íon de 100 °C; pressão de gás de colisão e 1,3 mTorr; energia de colisão normalizada usando de banda larga no modo de ativação de 35% para MS2. O limiar de seleção de íons foi 250 contagens para MS2. Uma ativação $q=0,25$ e tempo de ativação de 30 ms foi aplicado em aquisições MS2. A opção de "lock mass" foi ativado em todas as análises completas para melhorar a precisão das massa de íons precursores (OLSEN *et al.*, 2005).

Espectros de massas em *tandem* foram convertidos para o formato MS2 (MCDONALD *et al.*, 2004) utilizando o RawReader do *PatternLab for proteomics* v3.0.0.3 (CARVALHO *et al.*, 2012). A identificação de peptídeos foi realizada usando o programa ProLuCID v 1.3 (XU *et al.*, 2006). Resumidamente, este programa compara espectros experimentais contra teóricos, gerados a partir de um banco de sequências e seleciona aquela cujo espectro teórico mais se assemelha ao experimental de acordo com a correlação cruzada entre eles. O banco de sequências do *T. cruzi* foi composto por 66,602 sequências obtidas do NCBI somadas a outras 127 de contaminantes comuns à espectrometria de massas (queratina, BSA, entre outras). Em seguida, para cada sequência foram incluídas uma sequência artificial (*decoy*) gerada a partir da inversão da ordem dos aminoácidos da sequência original. Isto permite adiante modelar uma função de densidade de probabilidade de *scores* das identificações fornecidas pelo ProLuCID tanto para as sequências alvo quanto artificiais, ou seja, que não deveriam ocorrer. O espaço de busca considerou sequências de peptídeos trípticos e semi-trípticos com massa variando entre 600 e 6000 Da obtidas a partir do banco de sequências. Os parâmetros de identificação foram: acurácia de 50 ppm no MS1, 550 ppm para o MS2, carbamidometilação da cisteína e oxidação da metionina como modificações fixa e variável, respectivamente. A validade estatística dos *peptide sequence*

matches (PSM) foram calculadas usando o programa Search Engine Processor (CARVALHO *et al.*, 2012). Resumidamente, identificações foram agrupadas de acordo com o seu estado de carga (+2 and \geq +3) e status tríptico (tríptico ou semi-tríptico) resultando em quatro subgrupos distintos. Para cada subgrupo, os *scores* XCorr, DeltaCN e Zscore, provenientes do ProLuCID apontados para cada PSM, foram utilizados na modelagem de uma função discriminante Bayesiana. Isso permitiu estabelecer um critério de forma a convergir para um *false-discovery rate* (FDR) de 1% baseado no número de sequências *decoy*. Este procedimento foi realizado de forma independente para cada sub-conjunto de forma que o FDR seja independente do estado de carga ou status tríptico. Os resultados foram pós-processados de forma a eliminar PSMs com menos de 6 aminoácidos, ou tendo mais de 5ppm entre a massa experimental e a teórica. Uma lista de identificações com o conjunto mínimo de proteínas que explicam a todos os peptídeos (lista de máxima parsimônia) também foi gerada de acordo com a técnica do grafo bipartido (ZHANG, CHAMBERS e TABB, 2007). As análises subsequentes foram geradas usando o *PatternLab for proteomics* (CARVALHO *et al.*, 2012).

Após a identificação das proteínas foi realizada a verificação de sintenia genômica dos respectivos genes de interesse através dos resultados contidos em banco de dados de tripanossomatídeos - TritrypDB (<http://TriTrypDB.org>) e a identificação de possíveis proteínas ortólogas de *S. cerevisiae* pelo banco de dados OrthoMCL (<http://orthomcl.org/orthomcl/>). Foi também possível observar a rede de interações da proteína Imp3 de *S. cerevisiae* pela ferramenta STRING – *Search Tool for the Retrieval of Interacting Genes/Proteins* (<http://string-db.org/>) que através de bases de dados de interações conhecidas e *text mining* de artigos científicos integra informações de resultados já obtidos para o organismo de interesse com valores estatísticos sobre cada interação, atribuindo valores de confiança (*score*) em ordem crescente, de 0 a 1, sendo o valor 1 considerado de alta confiança (FRANCESCHINI *et al.*, 2013). Os resultados do ensaio de duplo-híbrido foram integrados para a visualização da rede de interações entre proteínas através da utilização do programa Cytoscape (SHANNON *et al.*, 2003).

6.1.5 Cromatografia de gel-filtração (cromatografia de exclusão por tamanho)

Esse procedimento foi utilizado para verificar se TcSub2 estaria em complexos nos extratos protéicos de *T. cruzi* obtidos através dos protocolos A, C e D, conforme item 6.1.3.1. Para isso foi utilizada a coluna analítica de filtração em gel Superdex® 200 10/300 GL (GE Healthcare Life Sciences # 17-5175-01) em sistema FPLC. Esta coluna permite o fracionamento de complexos protéicos, sendo possível a coleta das diferentes frações ao longo do tempo da corrida de cromatografia, sendo os complexos protéicos maiores os primeiros a deixar a coluna. Na figura 30 pode ser visualizada um cromatograma padrão para esta coluna, retirado no manual do fabricante. Este perfil de cromatograma padrão foi obtido através proteínas de tamanhos conhecidos que foram detectadas ao longo do tempo da cromatografia de gel-filtração em um fluxo padrão de 0,4 mL/min. Para as análises de TcSub2, a fração solúvel dos extratos protéicos obtidos dos protocolos A, C e D foram fracionados utilizando a coluna citada acima. Para isso foi utilizado o tampão TE-Noct e frações de 500 µL foram coletadas ao longo do tempo da cromatografia e analisadas por *western blot* com anticorpo anti-TcSub2 (Tabela 4), conforme item 5.1.10, para a detecção da proteína nestas frações.

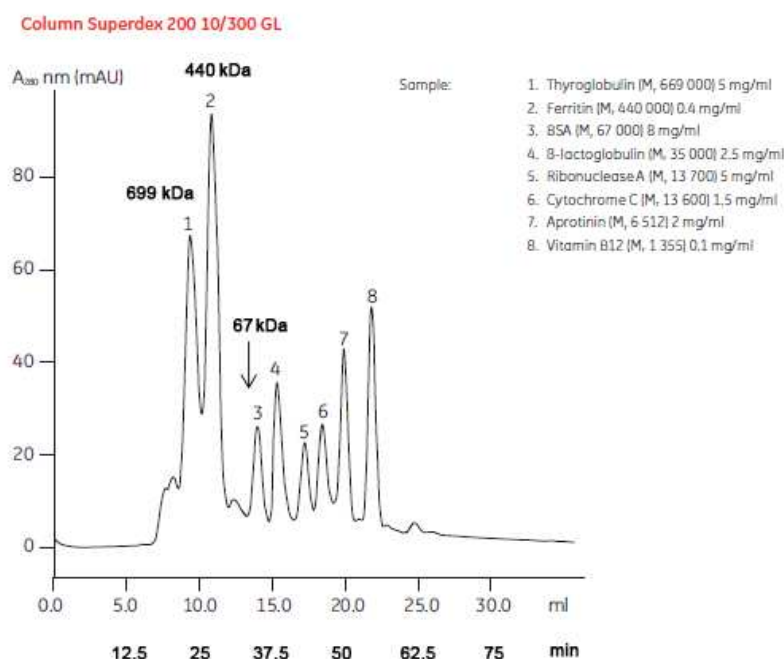


Figura 30. Cromatograma padrão de um teste funcional da coluna de gel-filtração Superdex 200 10/300 GL (GE Healthcare). Proteínas de tamanhos conhecidos estão indicadas do lado direito da figura. Algumas proteínas estão indicadas nos respectivos picos detectados. mAU: leitura da

absorbância das proteínas eluídas. ml: mililitros. Min: minutos. Fonte: Adaptado do manual Superdex 200 10/300 GL (GE Healthcare).

6.2 Resultados e Discussão

6.2.1 Análise de interações de proteínas por ensaios de duplo-híbrido em leveduras

Visando obter dados de interação proteica de componentes da via de exportação de mRNAs foi utilizada a metodologia de duplo-híbrido em leveduras, visto a disponibilidade de dados produzidos por outro grupo do Instituto Carlos Chagas – ICC (Fiocruz, Paraná).

Esta técnica foi concebida por FIELDS e SONG, 1989, e é realizada geralmente em leveduras ou bactérias, sendo um experimento *in vivo*. Classicamente é a mais utilizada para ensaios binários de interação proteína-proteína sendo melhorada constantemente em termos de função e eficiência (DURFEE *et al.*, 1993; GYURIS *et al.*, 1993; VIDAL *et al.*, 1996). Serve para identificar interações binárias através de transcrição de genes repórteres. Como desvantagem pode-se citar o alto número de falsos positivos e negativos e a necessidade da localização nuclear para o funcionamento do sistema. Apesar destes pontos negativos, esta metodologia é a mais eficiente e conveniente para se utilizar na exploração de interações individuais ou de interatomas inteiros, sendo a mais utilizada até o momento (SUTER, KITTANAKOM e STAGLJAR, 2008).

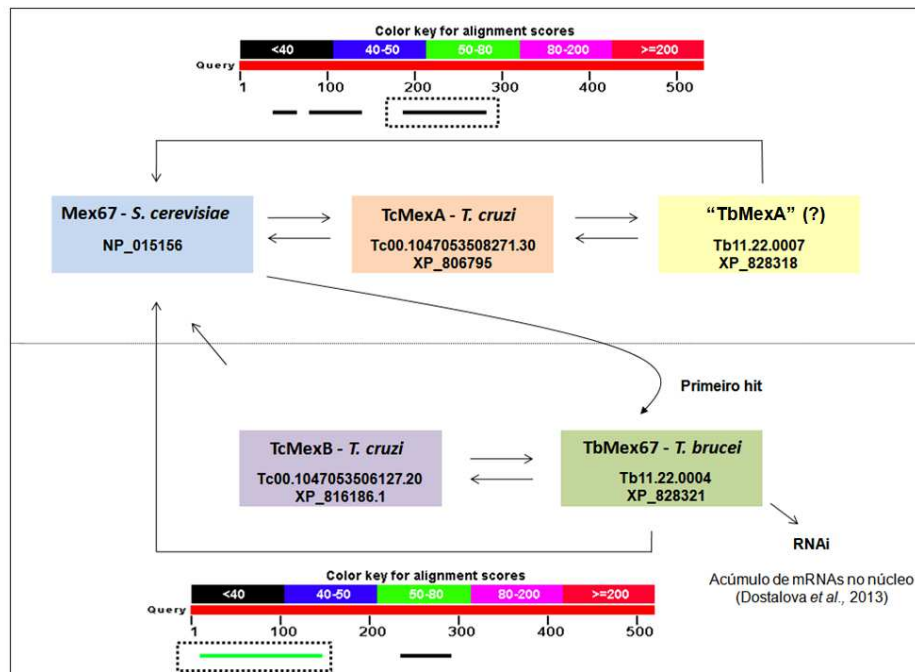
Para este ensaio foram selecionadas algumas proteínas candidatas a componentes da via de exportação de mRNAs em *T. cruzi* e o grau de conservação de cada uma delas estão descritas ao longo do texto (para classificação do grau de conservação ver artigo anexado no Capítulo 1, (SERPELONI *et al.*, 2011)).

Vale ressaltar que algumas das proteínas de *T. cruzi* incluídas neste estudo já foram relacionadas com a via de transporte nucleocitoplasmático em tripanossomatídeos, como é o caso de Tc/TbSub2 (SERPELONI *et al.*, 2011), TcCrm1 (CUEVAS, FRASCH e D'ORSO, 2005), TbNmd3 (PROHASKA e WILLIAMS, 2009), TbMex67 (DOSTALOVA *et al.*, 2013) e Lmj/TbRan (CASANOVA *et al.*, 2008). De acordo com o nosso estudo anterior (Capítulo 1) pode ser observado que a

maioria destas proteínas já estudadas funcionalmente estão bem conservadas ao longo da filogenia de eucariotos, classificadas de 1 a 2 em tripanossomatídeos como *Leishmania major*, *T. cruzi* e *T. brucei*.

Já a provável proteína ortóloga do receptor de Mex67 em tripanossomatídeos, ainda não estudada em *T. cruzi*, foi classificada de 3 a 5, o que mostra que não é uma proteína tão bem conservada. A falta de conservação em termos de sequência entre proteínas não permite inferir qual seria a provável ortóloga de Mex67 em *T. cruzi* e por isso foram utilizados para alvos deste estudo os dois primeiros *hits* resultantes dos alinhamentos entre TAP/Mex67 e proteínas de *T. cruzi*, sendo nomeadas de TcMexA e TcMexB. Um trabalho recente em *T. brucei* (DOSTALOVA *et al.*, 2013) confirmou a função da proteína ortóloga de Mex67, TbMex67, na exportação de mRNAs neste parasita, sendo que a proteína apresenta baixa cobertura de similaridade (apenas 22% de cobertura de alinhamento) com Mex67 de *S. cerevisiae*. Além da baixa similaridade entre as sequências, TbMex67 possui um motivo *zinc-finger* na porção N-terminal com papel essencial na exportação de mRNAs que está ausente em Mex67 de *S. cerevisiae* e outros eucariotos, parecendo ser específico de tripanossomatídeos. A busca a partir de TbMex67 em banco de dados de *T. cruzi* teve como primeiro *hit* a proteína TcMexB, com 79% de similaridade e cobertura de alinhamento de 96%. Os resultados da busca estão esquematizados na figura 31 A. Através de análises de sequências de proteínas de TbMex67 e prováveis ortólogas de *T. cruzi* (TcMexA e TcMexB) foi verificado que apenas TcMexB possui o motivo *zinc-finger*, que está ausente em TcMexA, como mostrado na figura 31 B. Com isso pode ser um forte indício de que TcMexB seja a proteína ortóloga de TbMex67 em *T. cruzi*.

A



B

```

TbMexA : NSRPVYKRG--GGPNRADSRRR--RKGLD--DRDVLVTRGGHNSWNAKPDRSRR--SRGSGPPGKGTGGSDNVQPQSTMMPTLLDLLFQRSSSVTFNRSGMLNLSN : 101
TcMexA : NSRPVYKHRGRDGGRCNPGRCNTGRGPFQETIANE--RFSRWNRGRN--SGGDRNNKSNNSNKKSGDRDVAAVFVMTPTLLDLLFQRSSAGAFDPNTGMLNLSR : 102
TbMexB : MPNRYRPN--KDFSRILVY-----YKRGSGANK--HDFYLRG-----CEKVGHEP--GKAPGD-----VGTNLLCPMLKLVYFEKQQQRVYNAEFTGVNLLHQ : 83
TcMexB : NSAPYRKN--KNTQNVFPI-----YKRGATNQ--RDFYLRGK--EIGENADAAPIGNGSTD-----AATNLLSFMKLVYFEKQQRVYNAEFTGVNLLDSE : 86

TbMexA : FSQSCDILLVRSVDFENNVIFCRSIVSVIRSHYGSMLRAVAANDKTKLSVFLSAVVEADLHICHTALSATGNATIDDLSEFVIGELKRRFVNIGELVLRKRNPFVTRRE : 206
TcMexA : FKDSFDLLSVQRVDFENSVIFCKSLADVIRKNGTSLRILVINEKNTKRLTLLTALTEGADTHRGVTAISATGNQISDFAFVIGELKRYEGEGELVLRKRNITITRA : 207
TbMexB : FKELDULLVAGSLNEMTAPCAQALCSLIRSLVPPPSAICLQKNGILSVSHLAMLEKRALHLLRALISLEANLIKSVESLQELKKHTNDRLEIVLRNLFANS : 188
TcMexB : LRKFFDLKDVSNINENTQSCRAHCSSIRSLVPPPSAICLQKNDITTSFFHFASCENAGLHLSIRALSLEANIKTMEPAICQTKKPSNIGELVIVGAVARFD : 191

TbMexA : DYNSSIQKVKSLMMDDKVINRAPIRLNEN--IPSSLSQLCFQVLYFLAIFSAAAEG-----KWTETGVIYAKNATLGSVRSSEPIACRIEFDKHKSDT : 302
TcMexA : DYRTQVRSIPKLMMDGELVHRGILHLPNE--VPVSLNESCLGVIRFLSNIFAAAASR-----NYDPAVNIYTSGTIVSVSRAEEP IPRRIEFDTAHNGDL : 303
TbMexB : DYRTTKKLMPSLIGDGESECVPHSLPWFQFSPTGYSEACKHVLCFCGGLNPLEGDGASGVAGVAVASDIYAAANAVTLLSLSPEVAVSTVMTMNGAAG : 293
TcMexB : DYRAEYKKAPELLICDGEAMAPPELLMERTNSDYTDAGQHVLCVVCSSLNPLQAGDAAGVPPQGVAVASDFYACKATPSISLSSPEAVAVSTVARTVKSPPV : 296

TbMexA : LNPTRVWMTDDEVFILRNVVPW--RNHVEHSLRSFACCCQTKAVAKANILGGGERKFFPSSEELNGNANVVVETSDN-----MNVFPCVITTHGRFHW--HWS : 396
TcMexA : LSESQRNVMTDFLILRQVWR--RNQNNIGLSLHVAACGRASSLAQEVGGGERKFFISSEELNGNANVVVETSDN-----MNVFPCVITTHGRFHW--HWS : 397
TbMexB : -VEQQRNVIRDIVSLRLKQTESNHNLLHGVRN--TVVACRTRVCSQMEH--WLYPKTFRWHYHDSASTVELDNSYLTGFSPTMVRVETVTVLHGRMVMNNTG : 394
TcMexB : -GGTQRNVIREIVSFRILQTESNHNLLHGVRN--SVVACRTRVCSQMEH--WLYPKTFRWHYHDSASTVELDNSYLTGFSPTMVRVETVTVLHGRMVMNHTLS : 397

TbMexA : -FI--DGGEKIFARDDAPEVSCFPHRRMSITLDPNGTITSVHNMMPFRP---DRLLRQESGEKALPMFFA-----NDPMRHEHRRRYLPQARTDVMKT : 485
TcMexA : -FKGNSGADLIPREKAPFVTCFPHRRVSLILLDSVSNVHICQEMHEMRP---DHTLVHENGASAPLFFA-----NEATRVESNRRRYLPNASLEVMQS : 487
TbMexB : VEGDATITGPM-----LHRRNRRVLSVTGGDAGR--RLITNEMVSYLYLTPSGSSASSKSGNVPEP--TQLSECRILLDPRSDRSRREVLRSK--KDVVPEVIMA : 489
TcMexB : -EQNQSNDRV-----VHYRNSRWISVQCNDAGR--RLITNEMVSYLYPFGDAASPSTGCDANASTQISLSECRILFSPREDRSRREVLRSK--KNVPLEVILA : 492

TbMexA : IVEAVSSDADVYGFQNCFAHLEPDQLI--IATASASDVMAN-----LGG----- : 528
TcMexA : VYBHFSGDADVHAFVTCFHTLEADRV--GAKDRSVMLS-----LQR----- : 530
TbMexB : FCGHVNDAEFTVLLDD--IRCVLSAFPHCANFAGENIMESIQCRLVNLFGMAPQNALBMRHNNGNWSDTVAAVAVTAPVAAPVQ-- : 575
TcMexB : ICGHVSNDADITVIND--IGQVLSMYPHCATFAGENILESIQVCRIGNRFGLAPEEGLELRRVGGNWCVAEEAMAAA--VGVPVQAS : 578

```

Figura 31. Busca por prováveis proteínas ortólogas de Mex67 em *T. cruzi*. A. As proteínas TcMexA e TcMexB foram encontradas a partir de alinhamentos por BLASTp com sequência de Mex67 de *S. cerevisiae* (NP_015156). Nesta busca TcMexA foi o primeiro *hit* encontrado e TcMexB o segundo *hit*. A busca dos ortólogos de TcMexA e TcMexB no banco de dados de proteínas de *T. brucei* resultam em proteínas diferentes, as quais foram nomeadas de "TbMexA" e "TbMexB" correlacionadas com TcMexA e TcMexB, respectivamente. TbMexB é a TbMex67 de *T. brucei*

(DOSTALOVA *et al.*, 2013) enquanto “TbMexA” ainda não foi estudada. **B.** TbMex67 (TbMexB) e as prováveis ortólogas desta proteína em *T. cruzi* (TcMexA e TcMexB) foram alinhadas, assim como a proteína relacionada a TcMexA, TbMexA. Em vermelho está destacado o motivo *zinc-finger* encontrado apenas em TcMexB e TbMexB, que está diretamente relacionado com a exportação de mRNAs em *T. brucei* (DOSTALOVA *et al.*, 2013). TcMexA: XP_806795. TcMexB: XP_816186. TbMexA: XP_828318. TbMexB (TbMex67): XP_828321.

Para as prováveis proteínas ortólogas de Mtr2/p15, componente do heterodímero receptor de exportação Mex67-Mtr2 / TAP-p15, os resultados de conservação em proteínas de tripanossomatídeos com a proteína Mtr2 de *S. cerevisiae* indicaram a baixa conservação das proteínas encontradas, não sendo classificadas para as análises (NC). Já a busca a partir de p15 (NXT1) de humanos em bancos de dados de proteínas de *T. cruzi*, o primeiro *hit* foi classificado como 2 ou 3, o que indica a boa conservação da proteína de *T. cruzi*. De fato o primeiro *hit* apresentou 74% de cobertura de alinhamento e 48% de similaridade e foi então denominada de Tcp15A. Para o segundo *hit* (Tcp15B) foi obtido também um bom resultado, com 82% de cobertura de alinhamento e 41% de similaridade e então ambas foram utilizadas como alvos deste estudo (Tcp15 A e B). O trabalho citado acima que identificou TbMex67 de *T. brucei* identificou também uma proteína associada, similar à proteína p15 de humanos, denominada de TbMtr2, também relacionada com a exportação de mRNAs no parasita (DOSTALOVA *et al.*, 2013). Através de alinhamentos de TbMtr2 por BLASTp no banco de dados de *T. cruzi* o resultado confirmou que Tcp15A é a provável proteína ortóloga pois apresentou 100% de cobertura no alinhamento e 98% de similaridade entre elas. Já a proteína Tcp15B não foi identificada nesta busca.

Algumas outras proteínas relacionadas com a via de exportação de mRNAs que ainda não foram caracterizadas em tripanossomatídeos mas possuem classificações de 1 a 3 em *T. cruzi* e *T. brucei* também foram alvos deste estudo. São elas as prováveis proteínas ortólogas de Dbp5, Gle2 e Gbp2. Outras proteínas com baixa conservação em tripanossomatídeos como Yra1 e Lea, classificadas como 4, também foram inseridas.

Em relação as proteínas relacionadas com a via de exportação de RNAs dependentes de RanGTP estão TcCRM1, TcRan e TcRanbp1, que foram classificadas de 1 a 3 em tripanossomatídeos e consideradas altamente conservadas ao longo da evolução de eucariotos. Com isso foram utilizadas como

alvos neste estudo para verificar as interações com possíveis componentes da exportação de mRNAs em *T. cruzi*.

Algumas proteínas que não foram analisadas no trabalho anterior também foram incluídas, como as prováveis proteínas ortólogas de nucleoporinas relacionadas com a exportação de mRNAs em fungos como Nup155 (FRANZ *et al.*, 2005), Nic96 (YOON *et al.*, 1997) e Nup145/Nup98 (FABRE *et al.*, 1994), bem como proteínas relacionadas ao processamento de pré-mRNAs por *splicing* já caracterizadas em tripanossomatídeos: U2AF35 (U2AF26), U2AF65 e SF1 (VAZQUEZ *et al.*, 2003; VAZQUEZ *et al.*, 2009), TcU2B (PREUSSER, PALFI e BINDEREIF, 2009), SmD1, Sm15k, Sm16,5k (WANG *et al.*, 2006), Lsm4 (LIU *et al.*, 2004), Ssm4 (TKACZ *et al.*, 2007), U2-40k (CROSS *et al.*, 1993) e Sf3b155 (AVILA *et al.*, 2007). Em relação aos nomes das proteínas foi utilizado o termo “Tc” antes do nome predito, mesmo para aquelas que ainda não foram caracterizadas funcionalmente.

Estas proteínas de *T. cruzi*, listadas na Tabela 16, foram utilizadas como alvos no ensaio de duplo-híbrido em leveduras e incorporadas ao conjunto, totalizando 480 proteínas que participam direta ou indiretamente na via do RNA para a identificação das interações. Estas proteínas ligadoras de RNAs (RBPs) foram previamente selecionadas pelo grupo do Dr. Christian M. Probst através de anotação pelo termo *RNA binding*, bem como pela identificação de domínios protéicos relacionados à ligação a ácidos nucléicos presentes no banco de dados Pfam (RRM, PUF, Zf-CCCH, Zf-CCHC, KH, OB, etc) (FINN *et al.*, 2014).

Tabela 16. Lista de proteínas de *T. cruzi* selecionadas para o ensaio de duplo-híbrido. A identificação das proteínas foi através de alinhamentos de sequência por BLASTp utilizando proteínas já caracterizadas em *S. cerevisiae*, *H. sapiens* ou *T. brucei*. Nesta busca, o primeiro *hit* encontrado foi considerado como resultado e então nomeado com o termo “Tc” para fins didáticos. As exceções foram para Mex67 e p15, em que foram utilizados os dois primeiros *hits* encontrados após os alinhamentos: “TcMexA” e “TcMexB”/ “Tc15_A” e “Tc15_B”. As proteínas estão coloridas de acordo com a provável função no parasita: Verde, via de exportação de mRNAs; Vermelho, vias de exportação de RNAs dependentes de RanGTP; Azul, processamento de mRNAs por *splicing*; Roxo, Eventos de tradução.

Organismo utilizado - Query		<i>T. cruzi</i>		Organismo utilizado - Query		<i>T. cruzi</i>		
Organismo	(ID / Anotação)	Alvo	Trityp ID / Anotação	Organismo	(ID / Anotação)	Alvo	Trityp ID / Anotação	
<i>S. cerevisiae</i>	NP_014689.1 ATP-dependent RNA helicase DBP5	TcDbp5	TcCLB.506587.40 ATP-dependent RNA helicase FAL1, putative	<i>S. cerevisiae</i>	NP_013396 Gsp1p (Ran)	TcRan	TcCLB.509455.80 GTP-binding nuclear protein rtb2, putative	
	NP_015156.1 Mex67p	TcMexA	TcCLB.506127.50 hypothetical protein, conserved		NP_010285.1 Yrb1p (Ranbp)	TcRanBP1	TcCLB.507099.30 Ran-binding protein 1, putative	
		TcMexB	TcCLB.506127.20 hypothetical protein, conserved		NP_012040 Nmd3p	TcNmd3	TcCLB.503843.10 hypothetical protein, conserved	
	NP_011033.1 Gle2p	TcGle2	TcCLB.511577.130 poly(A) export protein, putative		NP_011734 CRM1p	TcCRM1	TcCLB.511725.150 exportin 1, putative, RNA-nuclear export factor, putative (CRM1)	
	NP_010199 Sub2p	TcSub2	TcCLB.508319.40 RNA helicase, putative		EGA87168 Pab1p	TcPABP	TcCLB.506885.70 polyadenylate-binding protein 1, putative	
	NP_015111.1 Lea1p	TcLea	TcCLB.506945.280 hypothetical protein, conserved		NP_012717 Los1p	TcXPOt	TcCLB.509395.30 tRNA exportin, putative	
	NP_009916.1 Gbp2p	TcGbp2	TcCLB.510755.120 RNA-binding protein, putative		NP_009676.1 Tef2p	TcEF1	TcCLB.511367.360 elongation factor 1-alpha (ef-1- alpha), putative	
	NP_010669.1 Yra1p	TcYra1	TcCLB.508461.140 poly(A)-binding protein, putative, polyadenylate- binding protein, putative		<i>T. brucei</i>	Tb927.3.3480 U2 small nuclear ribonucleoprotein B, putative (RBP36 U2SNRNPB)	TcU2B	TcCLB.507951.140 U2 small nuclear ribonucleoprotein B, putative
	NP_014900.1 CDC31p	TcCDC31	TcCLB.506559.380 centrin, putative			Tb927.10.9400 branch point binding protein, putative	TcSF1	TcCLB.508717.40 branch point binding protein, putative
	EGA79018 Nic96p (Nup93)	TcNup95	TcCLB.507609.10 nucleoporin interacting component (NUP93), putative			Tb927.11.13960 U6 snRNA-associated Sm-like protein LSM4p (TbLSM4)	TcLsm4	TcCLB.504643.40 U6 snRNA-associated Sm-like protein LSM4, putative
CAY77705 Nup155/ Nup170p	TcNup147	TcCLB.506247.70 nuclear pore complex protein (NUP155), putative, nucleoporin, putative	Tb927.10.3500 RNA-binding protein, putative (RBSR4)	TcU2AF65		TcCLB.510265.40 hypothetical protein, conserved		
<i>T. brucei</i>	Tb927.11.980 Nucleoporin (TbNup158)	TcNup98	TcCLB.506925.440 nucleoporin (pseudogene), putative, serine peptidase, Clan SP, family S59, putative	Tb927.10.3200 U2 splicing auxiliary factor, putative		TcU2AF26	TcCLB.503577.20 U2 splicing auxiliary factor, putative (u2af26)	
				Tb927.10.4950 U2 small nuclear ribonucleoprotein 16.5K	TcSm16,5k	TcCLB.506583.10 U2 small nuclear ribonucleoprotein 16.5K, putative		
<i>H. sapiens</i>	NP_037380 p15 (NTF2-related export protein 1)	Tcp15A	TcCLB.508173.180 nuclear transport factor 2 protein (NTF2), putative	Tb927.6.4340 U6 snRNA-associated Sm-like protein LSM5p	TcSm15k	TcCLB.506943.114 U6 snRNA-associated Sm-like protein LSM5p		
		Tcp15B	TcCLB.511407.10 nuclear transport factor 2, putative	Tb927.10.14360 U2 small nuclear ribonucleoprotein 40K (U2_40K)	TcU2-40k	TcCLB.511283.190 U2 small nuclear ribonucleoprotein 40K, putative		

A partir dos resultados de interações pelo ensaio de duplo-híbrido foi gerada uma lista das proteínas identificadas. Um dado que chamou a atenção foi a interação da proteína TcImp3 encontrada com 8 proteínas selecionadas como alvos: TcSub2, TcDbp5, TcMexA, TcMexB, TcGle2, TcU2AF65, TcGbp2 e TcCRM1. Todos os resultados encontrados estão na tabela 17, sendo as interações por TcImp3 destacadas na tabela e comentadas ao longo do texto.

Tabela 17. Lista de interações obtidas por ensaio de duplo-híbrido para TcSub2 e outras proteínas candidatas a componentes das vias de exportação de RNAs de *T. cruzi*. As proteínas foram classificadas em cores de acordo com a provável função no parasita: Verde, via de exportação de mRNAs. Azul, processamento de mRNAs por *splicing*. Vermelho, vias de exportação de RNAs dependentes de RanGTP. Roxo, eventos de tradução. A coluna ao lado do nome predito "Tc" indica o código utilizado no ensaio de duplo-híbrido de cada proteína encontrada (presa). Logo ao lado está a descrição das proteínas encontradas (presas) em bancos de dados de proteínas do NCBI ou TritypDB. A proteína TcImp3 está destacada em cinza escuro. HPC: *Hypothetic Protein Conserved*.

Isca	Cód	Presa: Anotação Trityp	Isca	Cód	Presa: Anotação Trityp
TcDbp5	1C06	hypothetical protein, conserved (RBP25)	TcNup95	2A11	HPC (ELF-1 LIKE)
	2A06	60S ribosomal protein L12, putative		2D02	U2 splicing auxiliary factor zf-CCHC
	2F06	proteasome beta 5 subunit, putative (PRCE)		2A06	60S ribosomal protein L12
	4B09	U3 small nucleolar ribonucleoprotein protein IMP3, putative		2C05	60S ribosomal protein L10a
	3E01	hypothetical protein, conserved		2C03	60S ribosomal protein L10
TcMexB	2G05	metallo-peptidase, clan MP, family M67	TcNup98	3D02	eukaryotic translation initiation factor 3 subunit 8
	4B09	U3 small nucleolar ribonucleoprotein protein IMP3, putative		1B08	zf-CCHC universal minicircle binding protein (UMSBP)
	1H01	ribosomal protein S26, putative (ribosomal protein S26, putative)		1E10	RBP RRM (Brucei UBP1)
	2G12	proteasome regulatory non-ATPase subunit, putative		2E12	Proteasome Alpha 3 Subunit
	4F10	pre-mRNA cleavage complex II Clp1 protein		1F03	TcYra1
TcGle2	3E09	cleavage and polyadenylation specificity factor	TcU2AF65	2C10	transcription elongation factor
	3H08	40S ribosomal protein S13		3B06	TcNup147
	4F04	HPC		1C11	RBP RRM (DRBD2)
	4B09	U3 small nucleolar ribonucleoprotein protein IMP3, putative		2D01	proteasome alpha 2 subunit
	3D05	ATP-dependent RNA helicase, putative		4E11	splicing factor PTSR1 interacting protein
TcSm16,5k	4E07	ribosomal protein L11, putative	TcU2AF26	4G08	RNA editing 3' terminal uridylyl transferase 2 (KRET2)
	1H01	elongation factor 1-gamma (EF-1-gamma), putative		1E01	RBP (RBP29)
	4A02	eukaryotic initiation factor 5a, putative		1E12	HPC (PPCT1)
	1B02	HPC (RBP19)		5G04	HPC
	3A08	small nuclear ribonucleoprotein		3H09	40S ribosomal protein S16
TcSF1	4G02	proteasome regulatory ATPase subunit 5	TcU2B	4H07	HPC (RBSR4)
	1A06	RBP (Brucei RBP5)		2D01	proteasome alpha 2 subunit+A1:H306
	2F05	proteasome regulatory non-ATPase subunit 11		5G06	arginine N-methyltransferase
	3B11	TcNMD3		3E09	cleavage and polyadenylation specificity factor
	3C08	translation initiation factor eIF2B subunit-like protein		4E08	60S ribosomal protein L11
TcLsm4	3D02	60S ribosomal protein L37	TcLsm4	2E12	Proteasome Alpha 3 Subunit
	5C07	transcription modulator/accessory protein		3A11	U2 small nucleolar ribonucleoprotein 40K
	2F06	Proteasome beta 5 subunit		1F02	HPC
	2A02	40S ribosomal protein S15		1E10	UBP
	3H06	50S ribosomal protein L7Ae		5D05	small nuclear ribonucleoprotein SmD2
TcSub2	4B09	U3 small nucleolar ribonucleoprotein protein IMP3, putative	TcSm15k	4H03	TcGle2
	2F06	proteasome beta 5 subunit, putative (PRCE)		4A09	HPC
	2A06	60S ribosomal protein L12, putative		4E11	splicing factor PTSR1 interacting protein
	2C06	elongation factor 1-gamma (EF-1-gamma), putative		4F03	HPC
	1C09	hypothetical protein, conserved (RNA-binding protein, putative)		2A02	40S ribosomal protein S15
TcU2AF65	4C03	60S ribosomal protein L2, putative	TcU2-40k	1G10	U6 snRNA-associated Sm-like protein LSm4
	5H08	translation initiation factor eIF-2b beta subunit		5H11	HPC
	4C07	hypothetical protein, conserved		5B04	pre-mRNA splicing factor ATP-dependent RNA helicase
	3E08	RNA helicase, putative		2C10	transcription elongation factor
	1F03	poly(A)-binding protein, putative, polyadenylate-binding protein,		4C02	HPC
TcPabp1	3F12	ribosomal protein S25, putative	TcXp0t	4B09	U3 small nucleolar ribonucleoprotein protein IMP3, putative
	3A10	*TcSF1 branch point binding protein, putative		2D02	U2 splicing auxiliary factor zf-CCHC
	4B09	U3 small nucleolar ribonucleoprotein protein IMP3, putative		2H01	proteasome regulatory non-ATPase subunit 7
	3C02	splicing factor 3a, putative		1F01	PUF9
	4E05	eukaryotic translation initiation factor 3 subunit, putative		1C09	HPC (Brucei RBP34)
TcCRM1	3E02	pre-mRNA splicing factor ATP-dependent RNA helicase, putative	TcCRM1	4B09	U3 small nucleolar ribonucleoprotein protein IMP3, putative
	4B09	U3 small nucleolar ribonucleoprotein protein IMP3, putative		1B07	hypothetical protein, conserved
	4E08	U3 snoRNA-associated protein UTP11, putative		3G05	ribosomal protein L24, putative
	4C03	60S ribosomal protein L2, putative		2F03	translation initiation factor EIF-2b alpha subunit
	1F01	pumilio/PUF RNA binding protein 9, putative (PUF9)		3B12	tRNA nucleotidyltransferase
TcRan	4G04	DNA polymerase sigma-like protein, putative	TcRan	3H08	40S RIBOSOMAL PROTEIN S13
	1C09	hypothetical protein, conserved (RBP, putative (RBP34)		2G02	Translation initiation factor
	1E02	RNA-binding protein, putative (RBP, putative (RBP37)		5A12	ATP-dependent DEAD/H RNA helicase
	2A01	ubiquitin/ribosomal protein S27a, putative		1C06	HPC (RBP25)
	2H03	eukaryotic translation initiation factor, putative		2D07	proteasome alpha 5 subunit
TcYra1-PABP	3B03	hypothetical protein, conserved	TcNmd3	5D09	DNA-directed RNA polymerase subunit
	1C06	hypothetical protein, conserved (RBP, putative (RBP25)		2B12	HPC (TFIIF-stimulated CTD phosphatase)
	4B07	25 kDa translation elongation factor 1-beta		5G09	HPC
	4C05	IF4E		1A08	HPC (ZFP2)
	4E07	ribosomal protein L11		4H03	TcGLE2
TcCDC31	3G09	ribosomal protein S6	TcRanBP1	4B11	DNA-directed RNA polymerase II subunit 2
	1E04	RBP (Brucei DRBD6)		1H02	elongation factor 1-gamma (EF-1-gamma)
	1A03	HPC		2E10	HPC
	1B07	HPC		4F06	proteasome regulatory ATPase subunit
	1A10	ZF protein PDZ5/universal minicircle sequence (UMSBP)		4G07	eukaryotic translation initiation factor 3 (eIF-3)
TcEF1	1D11	RBP (TRRM1)	TcEF1	1B05	PUF3
	1G03	HPC		4D01	HPC (30S ribosomal protein S8)
	2D04	proteasome alpha 7 subunit		1B12	ALBA
	2D12	exosome complex exonuclease RRP41A		5G11	arginine N-methyltransferase, type III
	3A10	SF1		1F09	RBP (TRRM3)
3F06	ribosomal protein S27				

Estes dados de interações obtidos, listados na Tabela 17, foram integrados para a construção da rede de interação entre proteínas através do programa Cytoscape (SHANNON *et al.*, 2003), preservando as cores para facilitar a identificação destas na rede de interações, que pode ser observada na figura 32. Devido ao número de interações encontradas para a proteína TcImp3 como citado anteriormente, esta pode ser considerada um nó importante para essa rede.

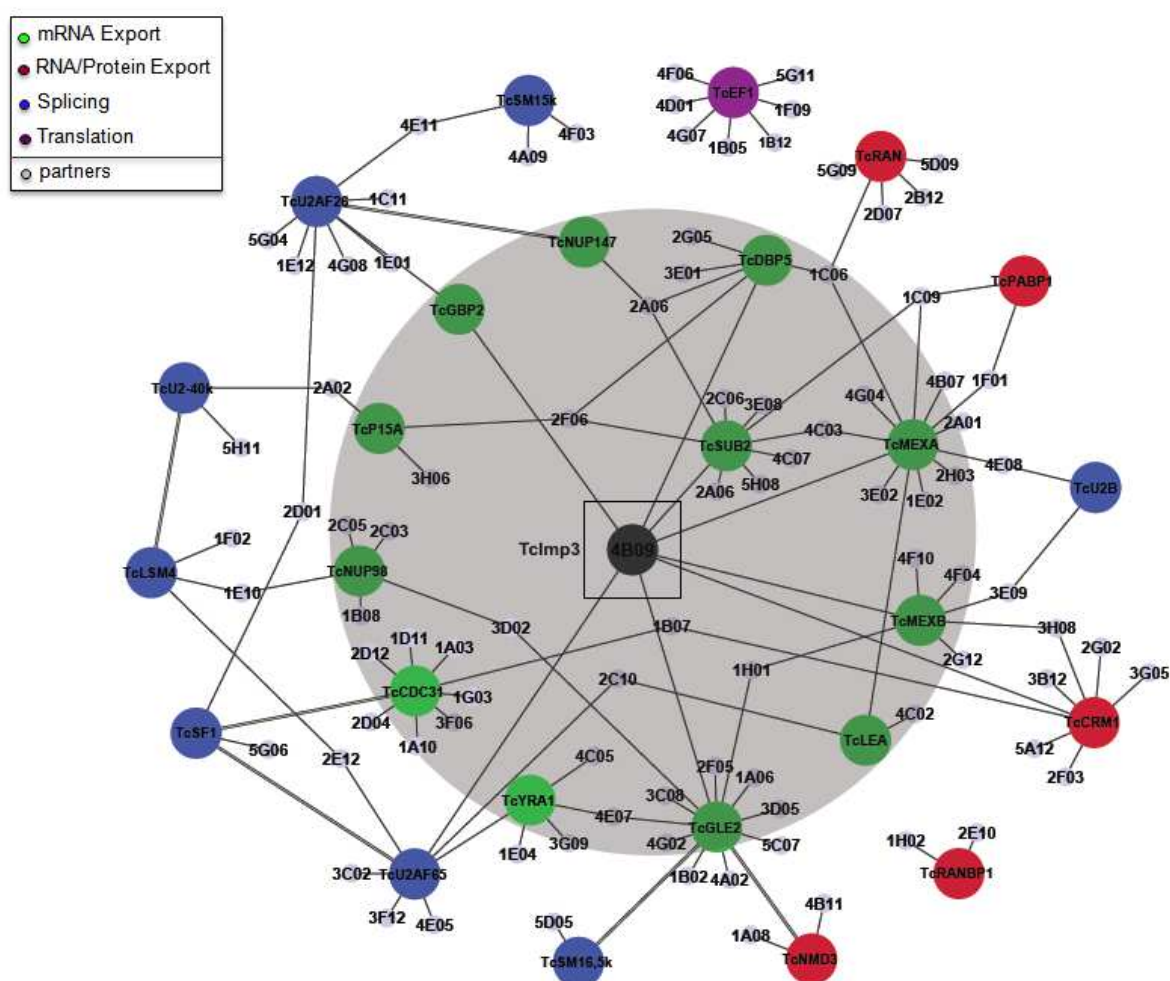


Figura 32. Rede de interações obtida através de resultados de ensaio de duplo-híbrido para TcSub2 e proteínas candidatas a componentes da via de exportação de RNAs de *T. cruzi*. Para a integração dos dados foi utilizado o programa Cytoscape (SHANNON *et al.*, 2003). A legenda indica as cores e a quais vias as proteínas provavelmente pertençam em *T. cruzi*: mRNA Export, via de exportação de mRNAs; RNA/protein export, vias de exportação de RNAs e/ou proteínas dependente de RanGTP; Splicing, Processamento de mRNAs por splicing; Translation, eventos de tradução; partners; proteínas encontradas (presas) neste ensaio. Os códigos das proteínas encontradas (presas) estão relacionadas com a Tabela 17. TcImp3 está destacada, no meio da rede de interações.

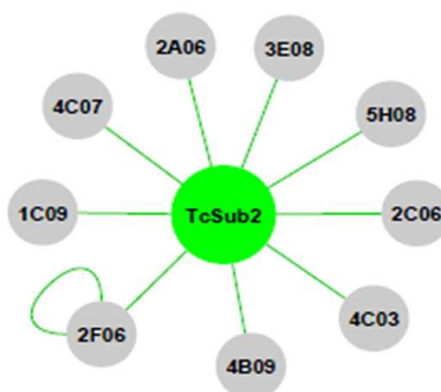
Para TcSub2 não foram identificadas evidências de interação direta com proteínas da via de exportação de mRNAs como TcCRM1 ou com a provável

proteína ortóloga de Mex67, TcMexB. As interações dependentes de mRNAs entre essas proteínas ainda não podem ser descartadas e aparentemente a interação entre TcSub2 e outras prováveis envolvidas na via de exportação de mRNAs possa ser mediada por TcImp3 (B09 – Figura 33). Foram ainda encontradas interações com proteínas ribossomais ou de proteossomo bem conservadas em eucariotos. Algumas proteínas como 4C07 e 1C09 (Figura 33) podem ser específicas de tripanossomatídeos e ter um papel importante na via de exportação de mRNAs de *T. cruzi*, mas estas proteínas ainda precisam ser estudadas.

Figura 33. Interações identificadas entre TcSub2 e proteínas RBPs por ensaio de duplo-híbrido.

A. A partir dos dados de interação obtidos do duplo-híbrido foi verificada a rede de interações de TcSub2, sendo esta a referência da busca na rede. Em verde está indicada TcSub2, alvo desta análise. Em cinza estão as interações encontradas com as RBPs, presentes no conjunto estudado. B. Tabela com análises das proteínas referentes às interações encontradas para as RBPs. Código (presa): Nome utilizado para a identificação da proteína no banco de RBPs. Tritryp (ID) Anotação: Anotação da proteína de *T. cruzi* no bancos de dados do Tritryp. As prováveis proteínas ortólogas de *T. brucei* e *S. cerevisiae* foram identificadas pela presença no mesmo banco de dados do OrthoMCL da proteína de *T. cruzi* em questão. Caso tenha sido encontrada a provável proteína ortóloga de *S. cerevisiae*, foram verificados a anotação, tamanho da proteína e domínios encontrados nas proteínas. Ao lado direito está demonstrada a porcentagem de similaridade entre as sequências de *T. cruzi* e as prováveis proteínas ortólogas de *S. cerevisiae*, encontradas no OrthoMCL.

A



B

Código (presa)	<i>T. cruzi</i>			<i>T. brucei</i>	<i>S. cerevisiae</i>			% S Tc x Sc
	Tritryp (ID) / Anotação	Tamanho da proteína (aa)	domínios encontrados (NCBI Conserved Domains)	Tritryp Anotação	Anotação no OrthoMCL	Tamanho da proteína (aa)	domínios encontrados (NCBI Conserved Domains)	
4C03	TcCLB.511527.34 60S ribosomal protein L2, putative	260 aa	Ribosomal Proteins L2	60S ribosomal protein L2, putative, 60S ribosomal protein L8, putative	Protein component of the large (60S) ribosomal subunit,	255 aa	Ribosomal Proteins L2	76%
4C07	TcCLB.510595.40 hypothetical protein, conserved	266 aa	Maf1 regulator; negative regulator of RNA polymerase III	hypothetical protein, conserved				
1C09	TcCLB.503999.90 hypothetical protein, conserved	270 aa	RRM_SF	RNA-binding protein, putative (RBP34)				
5H08	TcCLB.506865.40 guanine nucleotide exchange factor subunit GCD7, translation initiation factor 2b, delta subunit, putative, eIF-2B GDP-GTP exchange factor, putative	519 aa	IF-2B super family	translation initiation factor eif-2b beta subunit, putative (GCD2)	Beta subunit of the translation initiation factor eIF2B, the guanine-nucleotide exchange factor for eIF2	382 aa	Translation initiation factor 2B subunit	47%
2F06	TcCLB.507639.40 proteasome beta 5 subunit, putative	311aa	proteasome_beta_type_5	proteasome beta 5 subunit, putative, proteasome e beta 5 subunit (PRCE)	Beta 5 subunit of the 20S proteasome	288 aa	Proteasome subunit beta type-5	74%
4B09	TcCLB.504643.20 hypothetical protein, conserved	226 aa	S4/Hsp/ tRNA synthetase RNA-binding domain	U3 small nucleolar ribonucleoprotein protein IMP3, putative	Component of the SSU processome, which is required for pre-18S rRNA processing	184 aa	S4/Hsp/ tRNA synthetase RNA-binding domain - Ribosomal protein S4 and related proteins	62%
3. E08	TcCLB.504051.20 RNA helicase, putative	1324 aa	Sec63 Brl / DEAD-like helicases superfamily	RNA helicase, putative, DEAD/DEAH box helicase, putative	Hfm1p / Meiosis specific DNA helicase involved in the conversion of double-stranded breaks to later recombination intermediates	1188 aa	Domain of unknown function in Sec63p, Brr2p and other proteins / DEAD-like helicases superfamily	47%
2C06	TcCLB.510625.140 elongation factor 1-gamma (EF-1-gamma), putative	208 aa	Gamma subunit of Elongation Factor 1B					
2A06	TcCLB.508413.44 60S ribosomal protein L12, putative	164 aa	60S ribosomal protein L12	60S ribosomal protein L12, putative	Protein component of the large (60S) ribosomal subunit	166 aa	60S ribosomal protein L12	77%

Em relação as interações encontradas na tabela 17 foi possível observar que a provável proteína ortóloga de Mex67 (TcMexB) bem como TcMexA interagiram com TcImp3. No entanto as outras interações identificadas foram diferentes entre elas não sendo possível decifrar a função destas na via de exportação de mRNAs em *T. cruzi*. No entanto, assim como os dados obtidos para TcSub2, podem ajudar a nortear os estudos futuros de componentes específicos desta via.

Para TcMexA foi identificada uma proteína possivelmente relacionada com processamento de pré-mRNAs (E02: TcCLB.507681.220 *pre-mRNA splicing factor ATP-dependent RNA helicase, putative*), com 65% de cobertura de alinhamento e 68% de similaridade com a Prp22 de *S. cerevisiae*, uma DExD/H-box ATPase pertencente a spliceossomo (MAYAS, MAITA e STALEY, 2006)).

Já para TcMexB, que como citado anteriormente é similar a TbMex67 de *T. brucei*, dentre as interações identificadas há proteínas relacionadas com clivagem e poliadenilação de mRNAs: 3E09 (3E09: TcCLB.506871.140 - *cleavage and polyadenylation specificity factor, putative*) que parece ser específica de tripanossomatídeos pela ausência de similaridade de sequências com outros eucariotos, enquanto 4E10 (4E10: TcCLB.507027.59 - *pre-mRNA cleavage complex II Clp1 protein, putative*) tem tamanho similar à provável proteína ortóloga de *S. cerevisiae* Clp1p, cujo alinhamento entre as sequências resultou em 92% de cobertura e 37% de similaridade. Clp1 é uma das subunidades do fator de clivagem IA (*Cleavage factor IA - CF IA*), que faz parte do complexo de maturação do final 3' de mRNAs (DE VRIES *et al.*, 2000; HADDAD *et al.*, 2012) bem como relacionado com a exportação de mRNAs, conforme revisto por Saguez e Jensen, 2009 (SAGUEZ e JENSEN, 2009).

Dentre os resultados obtidos entre as próprias proteínas utilizadas como alvos foi encontrada a interação entre TcNmd3 (89% similar, 95% de cobertura com TbNmd3) e TcGle2 (82% similar, 94% de cobertura com TbGle2) que podem estar relacionadas com a exportação de rRNAs em *T. cruzi*. Em *T. brucei* TbNmd3 está relacionada com a exportação de subunidades 60S de ribossomos (PROHASKA e WILLIAMS, 2009) enquanto TbGle2, a provável proteína ortóloga de *T. brucei*, foi denominada nucleoporina, mas a função ainda não foi observada no parasita. O papel de Gle2 na exportação de rRNA, embora não canônica, já foi observada anteriormente em fungos (MOY e SILVER, 2002).

Em relação as interações identificadas entre dois alvos foram encontradas proteínas conservadas em fungos devido à alta similaridade de sequência e tamanhos entre elas, bem como proteínas que possuem prováveis ortólogos apenas em tripanossomatídeos e que provavelmente sejam mais específicas destes parasitas. Das proteínas bem conservadas estão proteínas de proteossomo e ribossomais e as provavelmente específicas de tripanossomatídeos pela ausência de similaridade de sequências com outros eucariotos estão possíveis fatores de transcrição, clivagem de pré-mRNAs ou as que possuem domínio RRM_SF, geralmente envolvido com processamento de mRNAs e rRNAs e estabilidade de RNAs.

Dos resultados entre dois alvos, uma interação interessante foi entre TcNup98, uma provável nucleoporina de *T. cruzi* e Lsm4 (cuja provável proteína ortóloga de *T. brucei* é nuclear e interage com U6 snRNA) (JAE *et al.*, 2010; TKACZ *et al.*, 2008), com a proteína TcUBP1 (Código: 1E10, Tabela 18) que, de acordo com a literatura, além de estabilizar alguns mRNA ricos em U na porção 3'UTR participa indiretamente da exportação de mRNAs em *T. cruzi* (CASSOLA e FRASCH, 2009; D'ORSO e FRASCH, 2001; D'ORSO e FRASCH, 2002). Estes resultados podem ser observados na Tabela 18.

Tabela 18. Interações identificadas entre 2 alvos por ensaio de duplo-híbrido. Tabela com análises das proteínas referentes às interações encontradas entre proteínas utilizadas como alvos. Interação encontrada - Código (presa): Código utilizado para a identificação da proteína no banco de RBPs. Trityp (ID) Anotação: Anotação da proteína de *T. cruzi* no bancos de dados do Trityp. As prováveis proteínas ortólogas de *T. brucei* e *S. cerevisiae* foram identificadas pela presença no mesmo banco de dados do OrthoMCL da proteína de *T. cruzi* em questão. Caso tenha sido encontrada a provável proteína ortóloga em *S. cerevisiae*, foram verificados a anotação, tamanho da proteína e domínios encontrados. Ao lado direito está demonstrada a porcentagem de similaridade entre as sequências de *T. cruzi* e as prováveis proteínas ortólogas de *S. cerevisiae*, encontradas no OrthoMCL.

Proteínas utilizadas como alvos	Interação encontrada Código (presa)	<i>T. cruzi</i>			<i>T. brucei</i>	<i>S. cerevisiae</i>			% S Tc x Sc
		Trityp (ID) / Anotação	Tamanho da proteína (aa)	domínios encontrados (NCBI Conserved Domains)	Trityp Anotação	Anotação no OrthoMCL	Tamanho da proteína (aa)	domínios encontrados (NCBI Conserved Domains)	
TcSF1 TcCRM1	1B07	TcCLB.506797.120 hypothetical protein, conserved	185 aa	RRM_SF Superfamily	RNA-binding protein, putative (RBP20)				
TcDbp5 TcMexA	1C06	TcCLB.508879.80 hypothetical protein, conserved	257 aa	RRM_SF Superfamily	RNA-binding protein, putative (RBP25)				
TcSub2 TcPABP	1C09	TcCLB.503999.90 hypothetical protein, conserved	270 aa	RRM_SF Superfamily	RNA-binding protein, putative (RBP34)				
TcLsm4 TcNup98	1E.10	TcCLB.507093.220 RNA-binding protein, putative	224 aa	RRM1_Hu_Like	RNA-binding protein, UBP1				
TcGle2 TcMexB	1H01	TcCLB.503801.20 ribosomal protein S26, putative	112 aa	40S ribosomal protein S26	ribosomal protein S26, putative	Protein component of the small (40S) ribosomal subunit; nearly identical to Rps26A	120 aa	Ribosomal protein S26e	63%
TcMexA TcPABP	1F01	TcCLB.506563.10 pumilio/PUF RNA binding protein 9, putative (PUF9)	535 aa	Pumilio-family RNA binding domain; Puf repeats	pumilio/PUF RNA binding protein 9 (PUF9)				
Tcp15A TcU2-40k	2A02	TcCLB.511809.9 40S ribosomal protein S15, putative9	152 aa	40S ribosomal protein S15	40S ribosomal protein S15, putative	Mitochondrial ribosomal protein of the small subunit, has similarity to E. coli S19 ribosomal protein	222 aa	Ribosomal protein S19	69%
TcNu147 TcDbp5 TcSub2	2A06	TcCLB.508413.44 60S ribosomal protein L12, putative	164 aa	60S ribosomal protein L12	60S ribosomal protein L12, putative	Protein component of the large (60S) ribosomal subunit	166 aa	60S ribosomal protein L12	77%
TcU2AF65 TcLea	2C10	TcCLB.507669.104 transcription elongation factor, putative	219 aa	Transcription factor S-II (TFIIS)					
TcSF1 TcU2AF26	2D01	TcCLB.504069.10 proteasome alpha 2 subunit, putative	231 aa	proteasome_alpha_type_2	proteasome alpha 2 subunit, putative	Alpha 2 subunit of the 20S proteasome	251 aa	proteasome_alpha_ty pe_2	70%
TcLsm4 TcU2AF65	2E.12	TcCLB.504213.120 proteasome alpha 3 subunit, putative	286 aa	proteasome_alpha_type_4	proteasome alpha 3 subunit, putative	Alpha 3 subunit of the 20S proteasome	259 aa	proteasome_alpha_ty pe_4	66%
TcSub2 Tcp15A	2F06	TcCLB.507639.40 proteasome beta 5 subunit, putative	311 aa	proteasome beta type-5 subunit	proteasome beta 5 subunit, putative, proteasom e beta 5 subunit (PRCE)	Beta 5 subunit of the 20S proteasome	288 aa	proteasome beta type- 5 subunit	74%
TcNup98 TcGle2	3D02	TcCLB.506885.14 ribosomal protein L37, putative	84 aa	60S ribosomal protein L37	60S ribosomal protein L37, putative	Protein component of the large (60S) ribosomal subunit	89 aa	60S ribosomal protein L37	71%
TcMexB TcU2B	3E.09	TcCLB.506871.140 cleavage and polyadenylation specificity factor, putative	1436 aa	Pre-mRNA cleavage and polyadenylation specificity factor	cleavage and polyadenylation specificity factor-like protein, putative				
TcMexB TcCRM1	3H08	TcCLB.511291.90 40S ribosomal protein S13, putative	151 aa	40S ribosomal protein S13	40S ribosomal protein S13, putative	Protein component of the small (40S) ribosomal subunit	152 aa	40S ribosomal protein S13	78%
TcGbp2 TcDbp5 TcMexA TcMexB TcCRM1 TcGle2 TcU2AF65 TcSub2	4B09	TcCLB.504643.20 hypothetical protein, conserved	226 AA	S4/Hsp/ tRNA synthetase RNA-binding domain	U3 small nuclear ribonucleoprotein protein IMP3, putative	Component of the SSU processome, which is required for pre-18S rRNA processing	184 aa	S4/Hsp/ tRNA synthetase RNA-binding domain - Ribosomal protein S4 and related proteins	62%
TcSub2 TcMexA	4C03	TcCLB.511527.34 60S ribosomal protein L2, putative	260 aa	60S ribosomal protein L8	60S ribosomal protein L2, putative, 60S ribosomal protein L8, putative	Protein component of the large (60S) ribosomal subunit	255 aa	60S ribosomal protein L8	76%
TcYra TcGle2	4E.07	TcCLB.509317.40 ribosomal protein L11, putative	341 aa	Ribosomal protein L11	ribosomal protein L11, putative	Mitochondrial ribosomal protein of the large subunit	159 aa	50S ribosomal protein L11	55%
TcMexA TcU2B	4E.08	TcCLB.508197.10 60S ribosomal protein L11, putative	192 aa	60S ribosomal protein L11	60S ribosomal protein L11, putative	Mitochondrial ribosomal protein of the large subunit	293 aa	Ribosomal protein L5	79%

Como pode ser observado na tabela 18, foi verificada a interação da provável proteína ortóloga de Imp3 (Código 4B09) de *T. cruzi* com TcSub2, TcCRM1, TcDbp5, TcGle2, TcMexA, TcMexB, TcGbp2 e TcU2AF65. Levando em consideração esses resultados para TcImp3, esta proteína foi analisada com mais detalhes em *T. cruzi*.

A proteína TcImp3 (GI: 407850254, TcSYLVIO_004230) está anotada como hipotética no banco de dados do TriTrypDB e está incluída no grupo OG5_128165 do banco de dados do OrthoMCL, assim como as proteínas ortólogas de tripanossomatídeos identificadas pelo TriTrypDB. A provável proteína ortóloga de *T. brucei* (Tb927.11.13930/Tb11.01.5500 – “TbImp3”), por exemplo, está anotada como *U3 small nucleolar ribonucleoprotein protein IMP3, putative* e apresenta 90% de similaridade, com tamanho similar a TcImp3. Neste grupo também estão presentes as proteínas de *S. cerevisiae* e *H. sapiens* denominadas Imp3 e relacionadas com a produção de 18S rRNA maduros (COSNIER *et al.*, 2011; GRANNEMAN *et al.*, 2003; LEE e BASERGA, 1999) e apresentam 62% de similaridade e tamanhos similares com a proteína TcImp3. Foi encontrado apenas um domínio em TcImp3 que pertence a superfamília S4, também presente nas prováveis proteínas ortólogas de fungos e humanos, que está relacionado a ligação a RNAs, e também pode estar presente em proteínas relacionadas a estresse, proteínas ribossomais ou tRNA sintetases (ARAVIND e KOONIN, 1999; MARKUS *et al.*, 1998). Este domínio pode ser observado na figura 34.

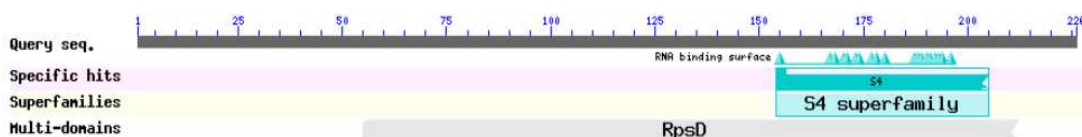


Figura 34. Esquema representativo do domínio encontrado na proteína “TcImp3”. S4 Superfamily (cd00165) - S4/Hsp/ tRNA synthetase RNA-binding domain. RpsD (COG0522). Ribosomal protein S4 and related proteins. Fonte: NCBI Conserved Domain Database and Search Service (MARCHLER-BAUER *et al.*, 2011).

Visto a alta similaridade com a provável proteína ortóloga de *S. cerevisiae* foram então observadas as interações já conhecidas de Imp3 neste fungo. Através da ferramenta STRING foi possível obter a rede de interações conhecidas de Imp3 de *S. cerevisiae* cujas interações já foram comprovadas por dados experimentais.

Foi observado que as proteínas associadas a Imp3 estão relacionadas com biogênese de ribossomos e processamento de pré-rRNA 18S, associadas ao pré-ribossomo (LEARY, TERNS e HUANG, 2004), bem como com a exportação de pré-40S dependente de CRM1 para a maturação no citoplasma, como é o caso da proteína ENP1 (WILD *et al.*, 2010; ZEMP *et al.*, 2009). Estes resultados estão mostrados na figura 35.

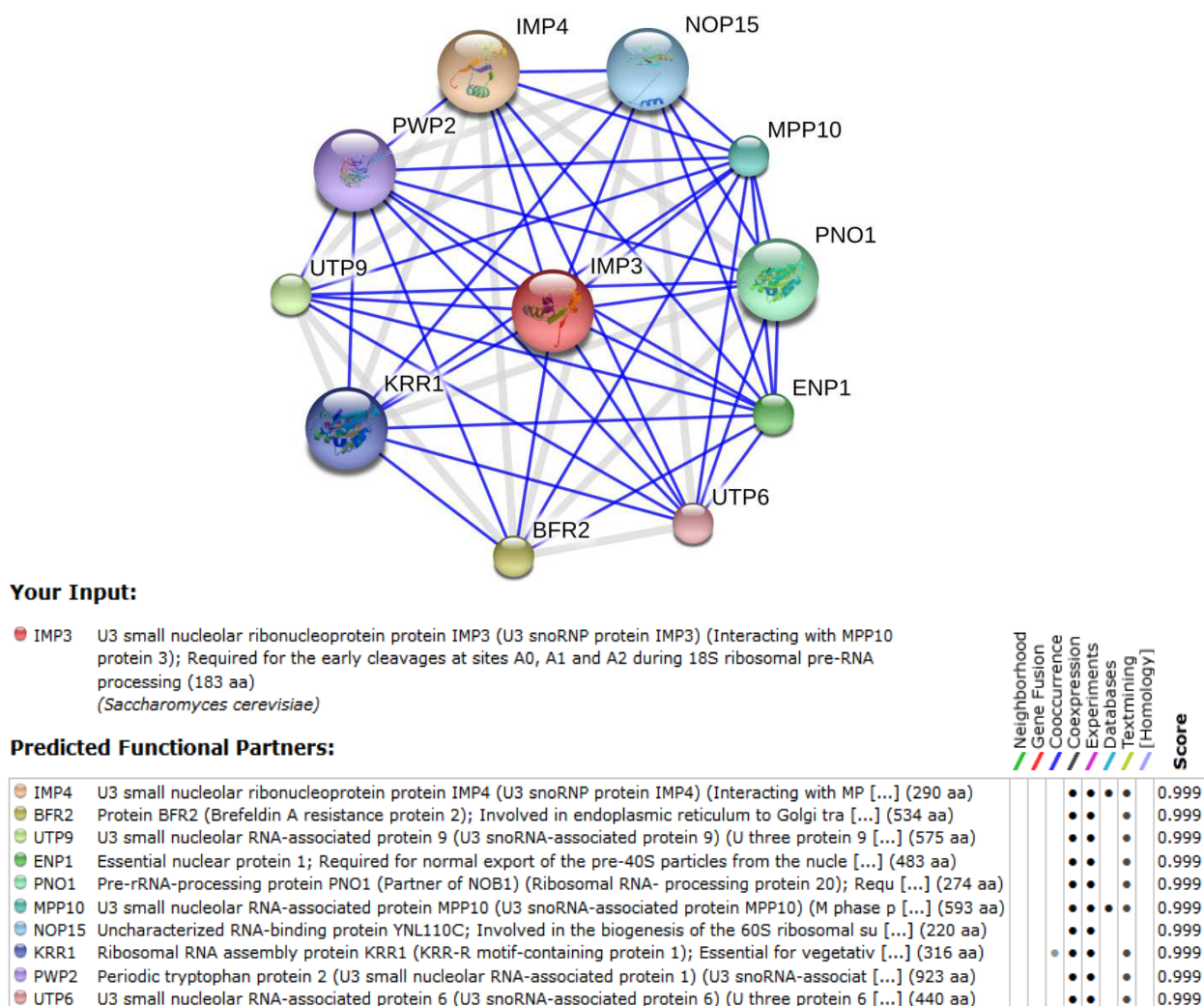
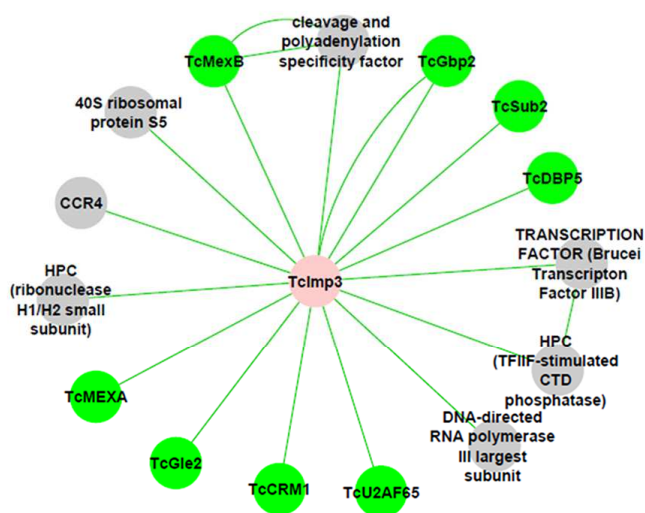


Figura 35. Rede de interação de Imp3 de *S. cerevisiae* pela ferramenta STRING (FRANCESCHINI *et al.*, 2013). As proteínas identificadas na rede de interação estão listadas logo abaixo. Score: Valor de confiança para a predição dos resultados. Esse valor varia de 0 a 1, sendo o valor 1 considerado de alta confiança.

Visto na figura 32 que TcImp3 interagiu em grande parte com os alvos deste estudo, previamente selecionadas neste estudo, resolvemos então verificar as interações desta proteína com o conjunto total das 480 RBPs e checar a presença

de interações com proteínas preditas da via de processamento de rRNAs, assim como o que vem sendo descrito em *S. cerevisiae*. Para isso foram utilizados todos os resultados obtidos do duplo-híbrido para a integração dos dados pelo Cytoscape, tendo como referência de busca a TcImp3. Conforme mostrado na figura abaixo, embora presentes no conjunto de RBPs, não foram encontradas interações diretas entre TcImp3 com as proteínas ortólogas TcImp4 (FRAGOSO *et al.*, 2003), TcSof1 (NARDELLI *et al.*, 2007) ou TcMpp10 que seriam as interações esperadas pois estas proteínas provavelmente participem da via de processamento de rRNA. Foram observadas, além das interações com os alvos deste estudo (em verde, Figura 36) outras proteínas RBPs (Figura 36 A). Através de análises de bioinformática foi possível observar a anotação destas proteínas de *T. cruzi* no banco de dados do TritypDB, tamanhos das sequências, os domínios presentes e identificação de prováveis proteínas ortólogas de *S. cerevisiae* no banco de dados do OrthoMCL (Figura 36 B). Foi verificado que estas proteínas de *T. cruzi* que interagem com TcImp3 estão provavelmente relacionadas a fatores de transcrição, clivagem de pré-mRNAs, ribonucleases bem como possivelmente associados a fatores da maquinaria de transcrição por RNA Polimerases II ou III devido à identificação dos prováveis ortólogos de Brf1(GEIDUSCHEK e KASSAVETIS, 2001) e Rpo31p (MOYLE, HOFMANN e INGLES, 1986) de *S. cerevisiae* (Figura 36 B). Ainda em relação às interações de TcImp3 foi também verificada a presença da provável proteína ortóloga de CCR4 de *T. cruzi* que participa do complexo CCR4-Not em *S. cerevisiae* cuja função está envolvida em etapas da regulação da expressão gênica através de ubiquitinação, deadenilação, associação com fator de transcrição TFIID e fatores de exportação de mRNAs (KERR *et al.*, 2011; SCHWEDE *et al.*, 2008) (revisto por Collart e Panasencko, 2010 (COLLART e PANASENKO, 2010). Em parasitas, existem poucos dados relacionados a função desta proteína. Apenas um trabalho em *T. brucei* descreve que TbCCR4 (a candidata à proteína ortóloga) não parece fazer parte do complexo Caf1/Not, mas a função ainda é desconhecida (SCHWEDE *et al.*, 2008).

A



B

Código (presa)	<i>T. cruzi</i>			<i>T. brucei</i>	<i>S. cerevisiae</i>			% S Tc x Sc
	Trityp (ID) / Anotação	Tamanho da proteína (aa)	domínios encontrados (NCBI Conserved Domains)	Trityp Anotação	Anotação no OrthoMCL	Tamanho da proteína (aa)	domínios encontrados (NCBI Conserved Domains)	
TRANSCRIPTION FACTOR (Transcription Factor IIIB)	TcCLB.507093.180 transcription factor, putative	602 aa	Transcription initiation factor TFIIB, Cyclin box fold	transcription factor, putative,transcription factor IIIb, putative	TFIIB B-related factor, one of three subunits of RNA polymerase III transcription initiation factor TFIIB (Brf1p)	597 aa	Brf1-like TBP-binding domain / Cyclin box fold / TFIIB zinc-binding	50%
40S ribosomal protein S5	TcCLB.506297.160 40S ribosomal protein S5, putative	190	40S ribosomal protein S5	40S ribosomal protein S5, putative	Protein component of the small (40S) ribosomal subunit, the least basic of the non-acidic ribosomal proteins	226 aa	40S ribosomal protein S5	83%
DNA-directed RNA polymerase III largest subunit	TcCLB.504105.30 DNA-directed RNA polymerase III largest subunit, putative	1553 aa	Largest subunit (RPC1) of eukaryotic RNA polymerase III (RNAP III)/Largest subunit (Rpc1) of Eukaryotic RNA polymerase III (RNAP III)	DNA-directed RNA polymerase iii largest subunit,RNA polymerase III largest subunit (RPC160)	RNA polymerase III subunit C160, part of core enzyme (Rpo31p)	1461 aa	Largest subunit (RPC1) of eukaryotic RNA polymerase III (RNAP III) N and C - terminal	60%
HPC (TFIIF-stimulated CTD phosphatase)	TcCLB.507857.70 hypothetical protein, conserved	327 aa	NLI interacting factor-like phosphatase	TFIIF-stimulated CTD phosphatase, putative				
cleavage and polyadenylation specificity factor	TcCLB.506871.140 cleavage and polyadenylation specificity factor, putative	1436 aa	CPSF A subunit region / Pre-mRNA cleavage and polyadenylation specificity factor	cleavage and polyadenylation specificity factor-like protein, putative				
HPC (ribonuclease H1/H2 small subunit)	TcCLB.508593.69 hypothetical protein, conserved	144 aa	RNase_H2_suC	(ribonuclease H1/H2 small subunit)				
CCR4	TcCLB.510535.60 CCR4 associated factor, putative	336 aa	CAF1 family ribonuclease	CCR4 associated factor, putative (CAF1)	RNase of the DEDD superfamily, subunit of the Ccr4-Not complex that mediates 3 to 5 mRNA deadenylation (Pop2)	434 aa	CAF1 superfamily mRNA deadenylase subunit	54%

Figura 36. Interações identificadas para TcImp3 por ensaio de duplo-híbrido. A partir dos dados de interação obtidos entre as RBPs foi verificada a rede de interações de TcImp3, sendo a referência da busca na rede. A. Em verde estão as proteínas de interesse desse estudo, utilizadas como alvos, listadas na Tabela 16. Em cinza estão as interações encontradas com as RBPs, presentes no

conjunto estudado. B. Tabela com análises das proteínas referentes às interações encontradas para as RBPs. Código (presa): Nome utilizado para a identificação da proteína no banco de RBPs. Trityp (ID) Anotação: Anotação da proteína de *T. cruzi* no bancos de dados do Trityp. As prováveis proteínas ortólogas de *T. brucei* e *S. cerevisiae* foram identificadas pela presença no mesmo banco de dados do OrthoMCL da proteína de *T. cruzi* em questão. Caso tenha sido encontrada a provável proteína ortóloga de *S. cerevisiae*, foram verificados a anotação, tamanho da proteína e domínios encontrados. Ao lado direito está demonstrada a porcentagem de similaridade entre as sequências de *T. cruzi* e as prováveis proteínas ortólogas de *S. cerevisiae*, encontradas no OrthoMCL.

Um outro resultado interessante foi a identificação de TcImp3 de *T. cruzi* também nos ensaios de imunoprecipitação de TcSub2 que serão descritos abaixo.

6.2.2 Imunoprecipitação para isolamento de complexos protéicos de *T. cruzi* contendo TcSub2

Após a incubação dos parasitas com o tampão TE-Noct para a obtenção do extrato protéico de parasitas de acordo com os protocolos descritos no item 6.1.3.1 (ver Tabela 15) foi observado que praticamente não há *pellet* após a centrifugação, havendo solubilização total da amostra. Os extratos protéicos de *T. cruzi* referentes aos protocolos A, C e D foram submetidas à cromatografia de gel-filtração para verificar se TcSub2 estaria presente em complexos protéicos, de acordo com o item 6.1.5. Após a cromatografia, as frações foram coletadas ao longo do tempo e então analisadas por *western blot* utilizando o anticorpo anti-TcSub2 (Tabela 4) para detecção da proteína por quimioluminescência, conforme item 5.1.10. Foi observado em todos os casos que TcSub2 se encontra nas frações 16 e 17 do cromatograma, o que indica que está eluindo em torno de 15 a 20 minutos da corrida. Estes resultados podem ser observados nas figuras 37, 38 e 39, referentes aos extratos protéicos obtidos a partir dos protocolos A, C e D, respectivamente. Este resultado mostra a presença desta proteína em complexos protéicos entre 400 e 699 kDa, quando comparado com os tamanhos das proteínas eluídas no cromatograma padrão mostrado na figura 30, neste mesmo tempo de eluição e utilizando o mesmo fluxo.

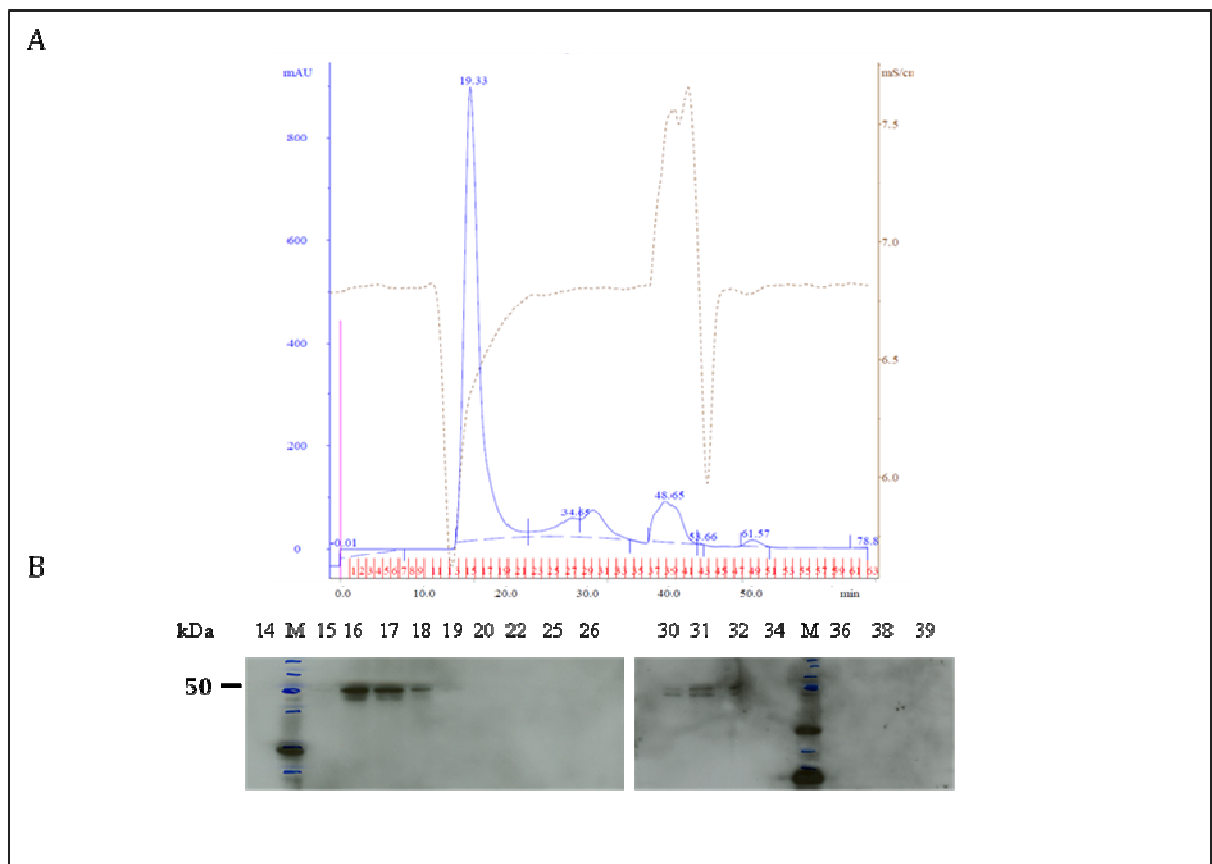


Figura 37. Cromatograma e análise das frações coletadas do extrato protéico referente ao protocolo A A. Cromatograma do perfil de distribuição de proteínas após cromatografia de gel-filtração – Superdex 200 10/300 GL (GE Healthcare). Em azul está a leitura da absorbância das proteínas eluídas. Em vermelho estão indicadas as frações coletadas ao longo do tempo. Em marrom a condutividade do tampão utilizado. B. *Western blot* das frações 14 a 40 utilizando o anticorpo anti-TcSub2 para a detecção desta proteína nos complexos protéicos. M: Marcador de massa molecular representado em kDa - *Benchmark* (Invitrogen).

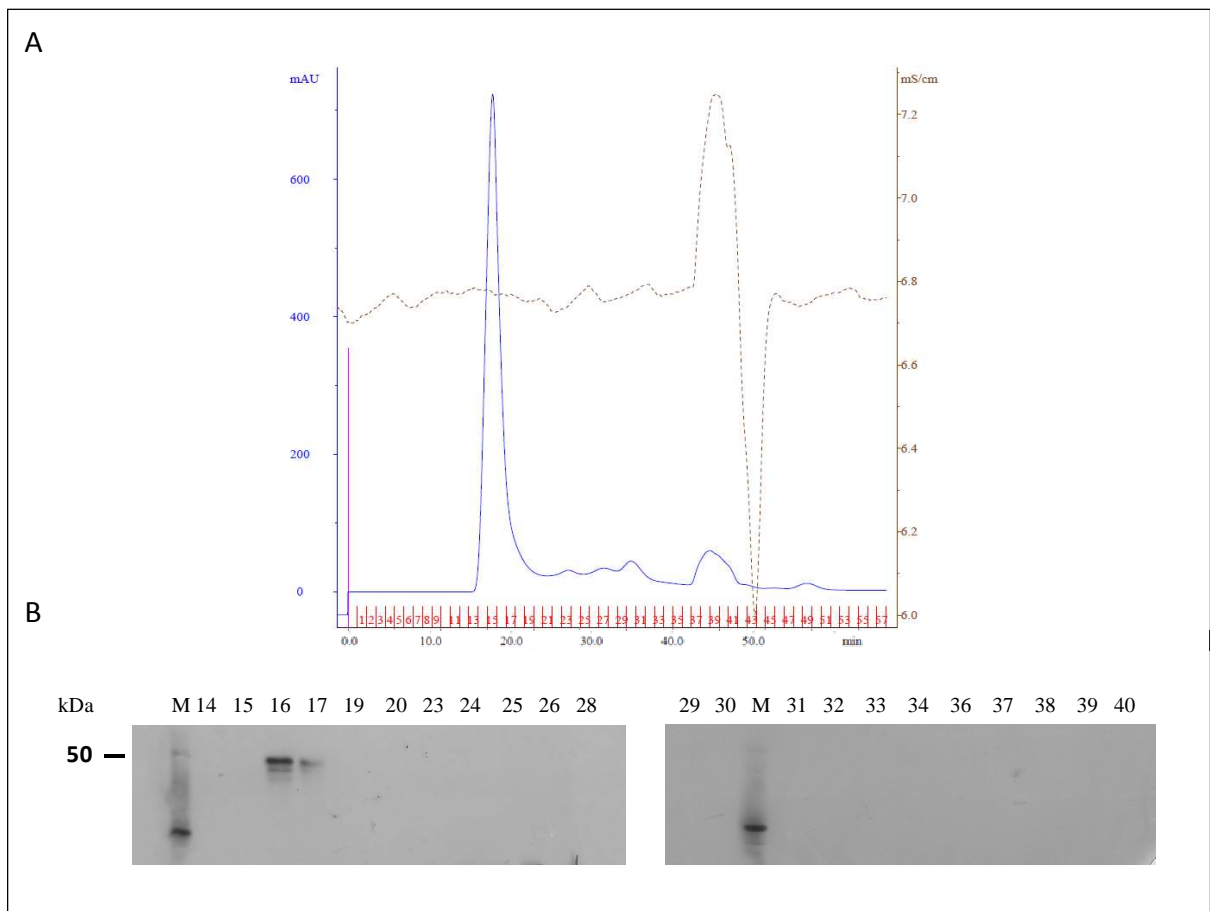


Figura 38. Cromatograma e análise das frações coletadas do extrato protéico referente ao protocolo C. A. Cromatograma do perfil de distribuição de proteínas após cromatografia de gel-filtração – Superdex 200 10/300 GL (GE Healthcare). Em azul está a leitura da absorbância das proteínas eluídas. Em vermelho estão indicadas as frações coletadas ao longo do tempo. Em marrom a condutividade do tampão utilizado. B. *Western blot* das frações 14 a 40 utilizando o anticorpo anti-TcSub2 para a detecção desta proteína nos complexos protéicos. M: Marcador de massa molecular representado em kDa - *Benchmark* (Invitrogen).

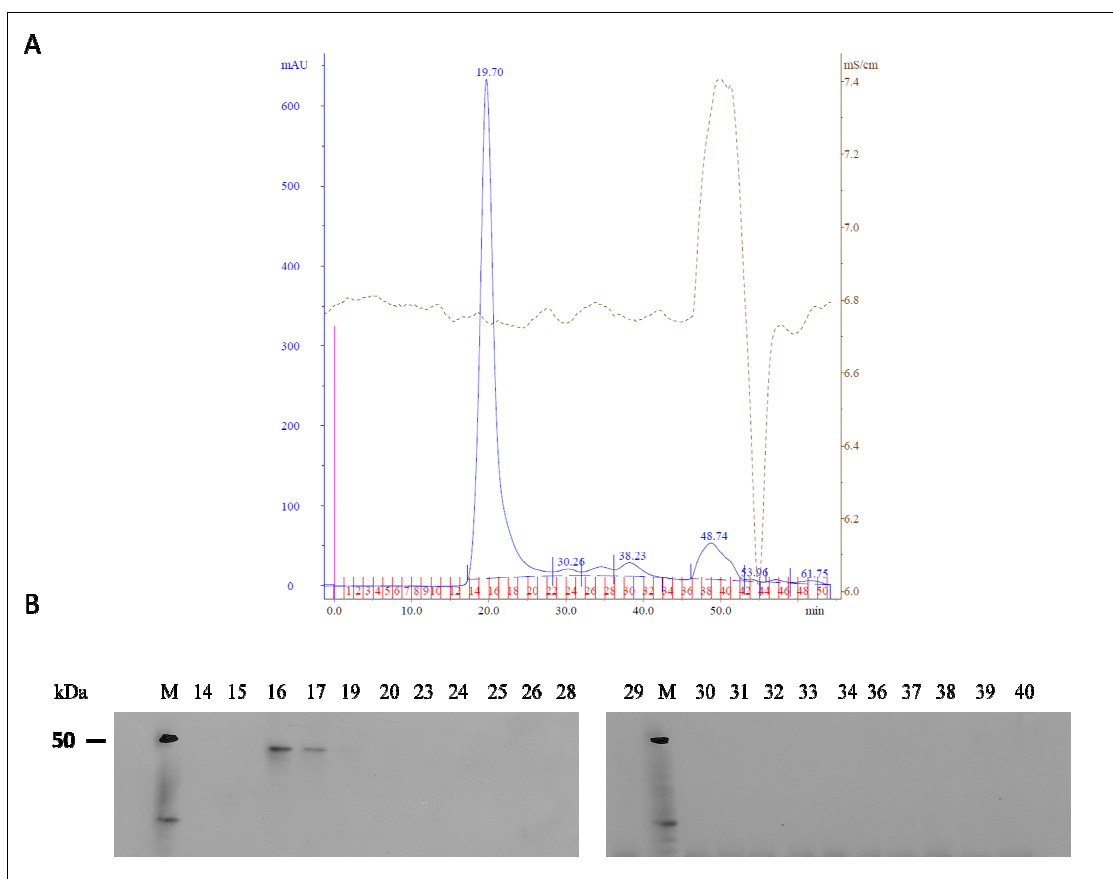


Figura 39. Cromatograma e análise das frações coletadas do extrato protéico referente ao protocolo D. A. Cromatograma do perfil de distribuição de proteínas após cromatografia de gel-filtração – Superdex 200 10/300 GL (GE Healthcare). Em azul está a leitura da absorvância das proteínas eluídas. Em vermelho estão indicadas as frações coletadas ao longo do tempo. Em marrom a condutividade do tampão utilizado. B. *Western blot* das frações 14 a 40 utilizando o anticorpo anti-TcSub2 para a detecção desta proteína nos complexos protéicos. M: Marcador de massa molecular representado em kDa - *Benchmark* (Invitrogen).

Com esses resultados foi possível observar que em todos os protocolos para obtenção do extrato protéico, TcSub2 se encontra complexada, independentemente da incubação dos parasitas em meio LIT sem soro fetal bovino ou da fixação prévia do extrato com formaldeído (*cross-linking*). De qualquer forma, foram realizados os ensaios de imunoprecipitação em triplicata técnica utilizando os protocolos A, B, C e D para a busca de proteínas associadas a TcSub2, conforme item 6.1.4. Todas as etapas das imunoprecipitações foram analisadas por *western blot* utilizando o anticorpo anti-TcSub2 para a detecção da proteína, conforme item 5.1.10. A ausência de TcSub2 na última lavagem (poço 3, Figura 40) bem como a presença de TcSub2 nas frações eluídas da triplicata técnica utilizando o anticorpo anti-TcSub2 (Tabela 4), “experimentos imunes”, confirmaram a imunoprecipitação de

TcSub2 (poços 4, 5 e 6, Figura 40). Nas frações eluídas relacionadas ao anticorpo pré-imune, “experimentos pré-ímenes”, não foi observada TcSub2, conforme o esperado (poços 9,10 e 11, Figura 40). Os resultados mostrados são relativos ao ensaio de imunoprecipitação de TcSub2 com o protocolo A e foram similares aos encontrados com os protocolos B, C e D. Estes resultados podem ser observados na figura 40. As proteínas presentes nas frações eluídas correspondentes aos ensaios realizados com anticorpo pré-imune e imune (experimentos pré-ímenes e ímenes) de todos os protocolos foram tripsinizadas e sequenciadas pelo espectrômetro de massas Orbitrap do Instituto Carlos Chagas – ICC (Fiocruz Paraná), conforme item 6.1.4.1.

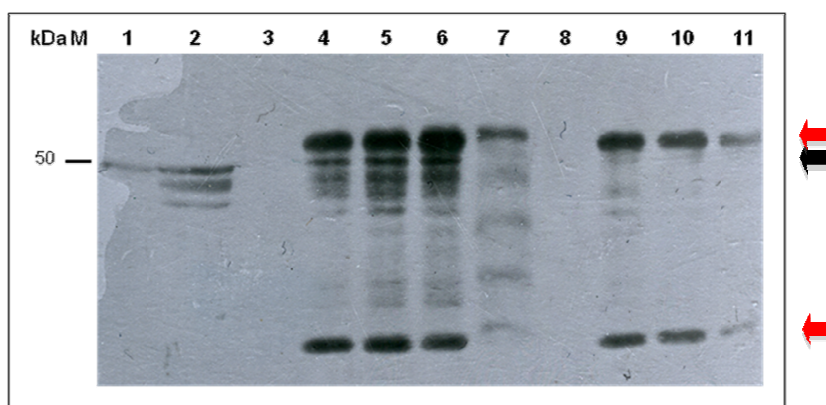


Figura 40. Imunoprecipitação da proteína TcSub2 em extrato protéico de *T. cruzi* obtido através de lise celular por TENoct (protocolo A). 1. Extrato protéico de 5×10^6 células, utilizado como controle positivo para TcSub2. 2. Fração do extrato protéico não ligada ao anticorpo após incubação de 16 horas a 4 °C. 3. Última lavagem (lavagem 4) de um dos experimentos relacionados ao anticorpo imune anti-TcSub2. 4-6. Frações eluídas de cada imunoprecipitação (triplicata) relacionada ao anticorpo imune anti-TcSub2. 7. Resina de um dos experimentos relacionado ao anticorpo imune anti-TcSub2. 8. Última lavagem (lavagem 4) de um dos experimentos relacionados ao pré-imune. 9-11. Frações eluídas de cada imunoprecipitação (triplicata) relacionada ao pré-imune. TcSub2 está indicada pela seta preta. Frações pesada e leve de IgG está indicada pelas setas vermelhas. M: Marcador de massa molecular representado em kDa - *Benchmark* (Invitrogen).

Embora os experimentos de imunoprecipitação tenham sido realizados em triplicata técnica, foram sequenciadas as amostras referentes a um experimento imune e um pré-imune de cada protocolo (A, B, C e D), pois dentro do tempo disponibilizado, ajudaria a obter dados para verificar a eficiência dos protocolos testado como confirmação da presença/ausência de TcSub2 nos experimentos imune/pré-imune e da ausência de polímeros nas amostras. Os espectros obtidos referentes aos peptídeos (SpecCount) foram utilizados para a identificação das

proteínas em bancos de dados de proteínas de *T. cruzi*, conforme item 6.1.4.1, obtendo uma lista de proteínas de *T. cruzi* identificadas.

Mesmo que sem valor estatístico, foi realizada uma primeira análise com a quantidade de espectros obtidos (SpecCount) de cada protocolo (apenas 1 experimento imune e do experimento pré-imune de cada triplicata técnica), para principalmente observar se TcSub2 foi identificada em todos os experimentos imunes mas ausentes no pré-imune. Para isso todas as proteínas identificadas nos experimentos imunes relacionados aos 4 protocolos (A-D, ver item 6.1.3.1 – Tabela 15) foram agrupadas em uma tabela, em ordem decrescente de SpecCounts. A partir dessa tabela, as proteínas comuns em pelo menos dois experimentos pré-imunes foram retiradas, independente do valor de leitura de SpecCount, obtendo assim a tabela final com 47 proteínas. Conforme o esperado, TcSub2 foi a proteína mais abundante sendo a primeira proteína da linha da tabela. Com essa tabela também é possível especular sobre algumas proteínas associadas a TcSub2. Entre as proteínas que apresentaram maior leitura de SpecCount e identificadas no banco de dados de *T. cruzi* estão: proteína ribossomal 40S (*40S ribosomal protein S11, putative*), fator de alongação da tradução (*elongation factor 1-alpha* (EF-1-alpha)), histona H2A (*Histone H2A*) e proteínas hipotéticas. Após a confirmação da presença de TcSub2 em todas as amostras bem como a ausência de polímeros, as proteínas presentes nas amostras restantes estão sendo sequenciadas para a obtenção dos dados e análises estatísticas.

Na figura 41 estão listadas as 47 proteínas identificadas na tabela final. As células correspondentes às proteínas que foram identificadas em cada um dos protocolos foram pintadas de verde. As 8 proteínas mais abundantes encontradas nos 4 protocolos estão destacadas por um quadrado preto. Foi verificada a conservação das proteínas identificadas em eucariotos através de comparação entre as primeiras sequências encontradas em 43 espécies representativas de eucariotos (Figura 41 - lado direito) utilizando abordagens já estabelecidas pelo trabalho anterior do nosso grupo para a classificação de 1 a 5/NC (SERPELONI, VIDAL *et al.*, 2011). Por motivos de espaço estão mostrados os resultados encontrados para *H. sapiens*, *C. elegans*, *S. cerevisiae*, *T. brucei* e *T. cruzi*, que nestes casos foram representativos para os resultados encontrados.

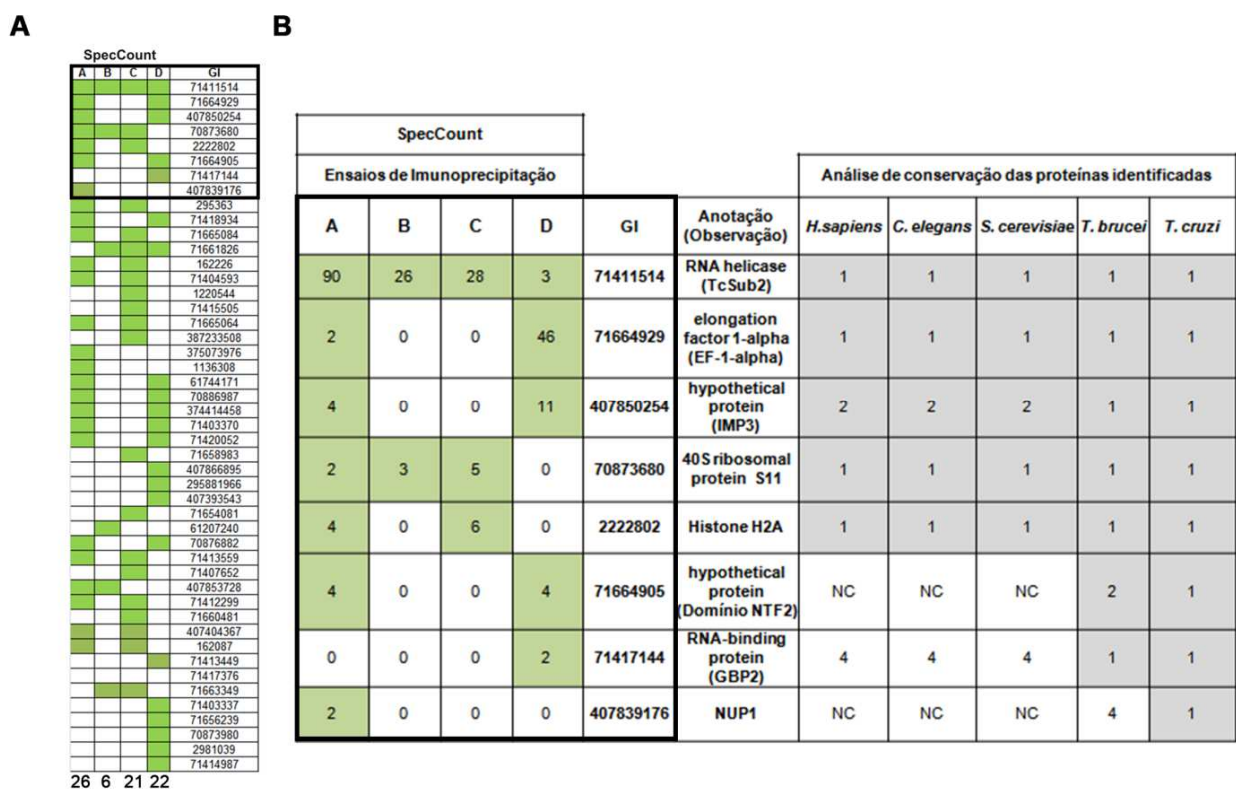


Figura 41. Lista de proteínas identificadas por espectrometria de massas a partir da imunoprecipitação de TcSub2 em diferentes extratos protéicos de *T. cruzi*. Os extratos protéicos foram obtidos na ausência ou presença de formaldeído 1%, seguido de lise celular com tampão TE-Noct. As letras indicam os protocolos utilizados para os ensaios de imunoprecipitação, listados na Tabela 15. A. Tabela final com 47 proteínas no total. As células pintadas em verde indicam que foram encontrados espectros referentes a peptídeos (SpecCount) para a identificação das proteínas. As proteínas estão identificadas pelo GI, em ordem decrescente de peptídeos identificados por proteína (SpecCount). Abaixo da coluna referente a cada protocolo estão indicados os números de proteínas identificadas por experimento. B. Lista com as 8 proteínas mais abundantes de todos os protocolos até o momento (A-D). A conservação das proteínas identificadas foi analisada através de busca em bancos de dados de proteínas de 43 eucariotos, sendo mostradas na tabela apenas os resultados *H. sapiens*, *C. elegans*, *S. cerevisiae*, *T. brucei* e *T. cruzi* que foram representativos para o resultado geral. Os números indicam a classificação da conservação de acordo com critérios previamente estabelecidos (SERPELONI, VIDAL *et al.*, 2011): 1 - altamente similar; 2 – similar; 3 - pouco similar; 4 e 5 – parcialmente similar, NC – não analisada pois o *e-value* está acima da linha de corte. As classificações 1 e 2 estão em destaque e mostram os melhores resultados de conservação neste tipo de análise. GI: Número de identificação das proteínas do *GeneBank*. SpecCount: Quantidade de espectros identificados de peptídeos referentes à proteína identificada no banco de dados de proteínas de *T. cruzi*. A informação obtida durante as análises considerada relevante estão entre parênteses, logo abaixo do GI.

Como citado anteriormente ainda é preciso incluir os resultados referentes às identificações das proteínas presentes nas outras réplicas de experimentos. Mesmo assim, foi possível observar que a amostra referente ao protocolo B, o qual as lavagens foram feitas com tampão TE-Noct-50 (ver item 6.1.3.1, Tabela 15), teve um número menor de proteínas identificadas (Comparar colunas A, C e D com coluna B da figura 41 A). Este resultado pode indicar que o aumento da quantidade de KCl

durante as lavagens tenha interferido na interação das proteínas com TcSub2, que podem ser muito dinâmicas ou até mesmo instáveis. Visto que ainda faltam dados, não foram realizadas análises detalhadas de bioinformática para as proteínas identificadas até o momento. No entanto, entre as 8 proteínas mais abundantes existem algumas proteínas com características interessantes:

- GI: 71417144, embora não tão conservada em eucariotos, possui domínios em comum com a provável proteína ortóloga Gbp2, uma proteína SR relacionada com a exportação de mRNAs em fungos (HURT *et al.*, 2004; LEE, HENRY e SILVER, 1996).
- GI: 71664905, uma proteína que possui domínio NTF2 – *Nuclear Transport Factor 2* (código: cl09109). Este domínio pode se ligar a Ran-GDP e nucleoporinas durante o transporte de moléculas nas células (STEWART, KENT e MCCOY, 1998) e também está presente na proteína p15, que forma o heterodímero receptor de exportação de mRNAs em humanos (FRIBOURG *et al.*, 2001).
- GI: 407839176, TcNup1, presente na periferia nuclear, relacionada com o ancoramento da cromatina ao envelope nuclear em *T. cruzi* (PICCHI *et al.*, 2011).

Visto que TcImp3 foi a única proteína cujas interações com TcSub2 foram identificadas por análises de duplo híbrido e imunoprecipitação até o momento, o grupo resolveu iniciar os estudos com esta proteína de *T. cruzi*. Neste caso, o projeto foi desenvolvido por uma aluna de iniciação científica, que permitiu a apresentação neste trabalho dos dados de localização de TcImp3. Foi verificada a localização por microscopia de fluorescência de TcImp3 fusionada a GFP na porção N ou C-terminal e a proteína não está localizada no nucléolo, sendo a distribuição citoplasmática e perinuclear (Figura 42). A localização de Imp3 fusionada a GFP é preferencialmente nuclear e nucleolar em outros eucariotos (COSNIER *et al.*, 2011; GRANNEMAN *et al.*, 2003), diferente do observado para TcImp3. Mesmo assim, esta localização de TcImp3 ainda será confirmada em *T. cruzi* através de imunofluorescência e imunocitoquímica utilizando anticorpo específico para esta proteína. O resultado de localização pode ser observado na figura 42.

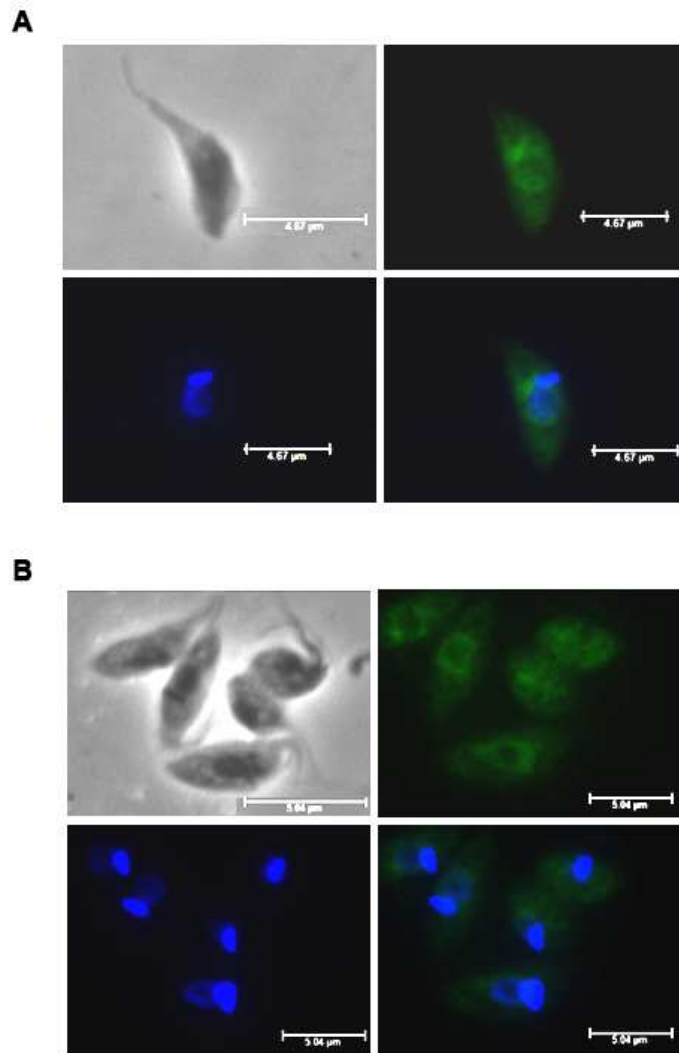


Figura 42. Localização de TcImp3 fusionada a GFP em *T. cruzi*. A visualização foi por microscopia de fluorescência, sem incubação com anticorpo. A. Localização de TcImp3 fusionada a GFP na porção C-terminal. B. Localização de TcImp3 fusionada a GFP N-terminal. Em verde: Localização da proteína TcImp3 fusionada a GFP. DAPI: Marcação da cromatina nuclear e de cinetoplasto, em azul. Merge: Sobreposição de imagens. Barra: 5 µm.

A proteína Imp3 foi identificada primeiramente em fungos como essencial na etapa de processamento de pré-18S rRNA através da interação com U3snoRNA (U3 *small nucleolar RNA*) formando complexos U3 snoRNPs ou *small-subunit* (SSU) *processome* (LEE e BASERGA, 1999). Estes complexos possuem além de Imp3 e U3 snoRNAs, outras proteínas como Sof1p, Mpp10p, Imp4, Lcp5p e fatores de processamento de subunidades 40S (DLAKIC, 2005; DRAGON *et al.*, 2002; DUNBAR *et al.*, 1997; GRANDI *et al.*, 2002; JANSEN, TOLLERVEY e HURT, 1993; WIEDERKEHR, PRETOT e MINVIELLE-SEBASTIA, 1998) que são necessários para

as etapas de clivagens do pré-rRNA 35S nucleolar na produção de 18S rRNA maduros (KASS *et al.*, 1990; VENEMA e TOLLERVEY, 1995). Embora ainda não tenha sido estudado o papel de U3 snoRNA em *T. cruzi*, já foram caracterizadas algumas proteínas ortólogas de Imp4 e Sof1 no parasita e foi verificado que não possuem localizações estritamente nucleolares, como em outros eucariotos (GRANNEMAN *et al.*, 2003; LEARY, TERNS e HUANG, 2004). A localização de TcImp4 é perinucleolar (FRAGOSO *et al.*, 2003) e TcSof1 tem uma distribuição nuclear, principalmente na região da periferia do núcleo (NARDELLI *et al.*, 2007). Logo, pela falta de dados e características diferentes das proteínas ortólogas de fungos, a função destas proteínas como componentes do complexo de processamento de pré-rRNAs ainda precisa ser investigada.

TcImp3 foi identificada tanto nos ensaios de imunoprecipitação de TcSub2 bem como com interações diretas por ensaios de duplo-híbrido. Além desta interação, foram identificadas interações de TcImp3 por ensaios de duplo-híbrido com TcCRM1, com a provável proteína ortóloga de Mex67, TcMexB, entre outros possíveis fatores envolvidos com a exportação de mRNAs em *T. cruzi*. Levando em consideração os resultados de interações e a localização da proteína em *T. cruzi* torna-se interessante verificar se TcImp3 é um componente da via de exportação de mRNAs e/ou participa da biogênese de ribossomos no parasita.

7 Conclusões e Perspectivas

Através de análises de conservação de proteínas relacionadas com exportação de RNAs em organismos-modelo ao longo da filogenia de eucariotos foi possível observar que componentes básicos das vias dependentes de RanGTP estão bem conservados sendo a via de exportação de mRNAs, independente de RanGTP, a que menos apresenta proteínas altamente conservadas principalmente em parasitas. A partir desses resultados foi proposto um modelo evolutivo da exportação de RNAs, incluindo papéis de exportinas ancestrais e evolução das vias de RNAs, sendo a via de exportação de mRNAs a mais especializada em eucariotos que se divergiram no topo da árvore filogenética. Neste estudo de análise de conservação foi encontrada uma exceção: Sub2 (ou UAP56, de humanos), uma proteína altamente conservada em eucariotos inclusive em parasitas. Sub2/UAP56 é uma RNA helicase relacionada com processamento por *splicing* de pré-mRNAs e com papel essencial na exportação de mRNAs em organismos-modelo. Através de análises funcionais foi verificado o papel essencial da proteína ortóloga de Sub2 de tripanossomatídeos na via de exportação de mRNAs, sendo denominada de Tryp-Sub2, compreendendo TbSub2 e TcSub2, de *T. brucei* e *T. cruzi*, respectivamente. Foi observada que a ausência de TcSub2 em *T. cruzi* é letal e a diminuição gradual da proteína ortóloga, TbSub2, em *T. brucei* resultou em acúmulo nuclear de mRNAs ao longo do tempo, resultando também em morte dos parasitas. Em *T. cruzi* TcSub2 apresenta distribuição nuclear preferencialmente em bordas de cromatina densa e não-densa, assim como em outros eucariotos, colocalizando com alguns sítios de transcrição por RNA Polimerase II. A conservação da função desta proteína em parasitas foi confirmada através de estudos funcionais em *T. gondii*, pertencente ao Filo Apicomplexa, sendo denominada de TgUAP56. Através de análises de superexpressão ou nocaute condicionado foi confirmado o papel essencial de TgUAP56 na exportação de mRNAs, embora não necessariamente relacionado com o processamento por *splicing* de alguns mRNAs. Baseado nos experimentos para verificar a presença de mRNAs precursores não processados e o acúmulo de mRNAs no núcleo dos parasitas após 48 horas da ausência de TgUAP56, acreditamos que a proteína provavelmente atue na liberação de mRNAs maduros para algum receptor de exportação de mRNAs, assim como em células humanas.

Ainda assim, para confirmar o real papel da proteína no *splicing* deverão ser realizados experimentos de sequenciamento em larga-escala destes mRNAs acumulados e buscas por regiões não processadas para uma avaliação global do papel da proteína no processamento de mRNAs. *T. gondii* se mostrou um bom modelo de estudo, permitindo a análise funcional de TgUAP56 de um modo rápido, podendo ser utilizado para estudo funcional de outras proteínas essenciais e possivelmente relacionadas com a exportação de mRNAs de parasitas através de nocaute condicionado ou através de *screening* funcional utilizando a superexpressão como ferramenta. O sistema MS2 em *T. gondii* após a validação funcional, poderá ser utilizado para a análise da confirmação do papel de proteínas nesta via, sendo um sistema de visualização rápida da localização de mRNAs, bem como para experimentos futuros de varreduras de drogas em média/larga escala.

Em relação às proteínas que interagem com TcSub2 por imunoprecipitação, os resultados são ainda preliminares pois as análises não estão finalizadas. No entanto os resultados preliminares indicaram a interação com TcImp3, a provável proteína ortóloga de Imp3 em *T. cruzi*, que também foi encontrada em ensaios de duplo-híbrido com interação direta com TcSub2. Além desta interação foram observadas interações de TcImp3 com TcCRM1 e com as proteínas conservadas em parasitas: TcDbp5 e TcGle2. Foram identificadas interações com proteínas menos conservadas como a provável proteína ortóloga de TbMex67, TcMexB, reforçando a necessidade de estudar o papel destas proteínas na exportação de mRNAs em *T. cruzi*.

Foi observado que TcImp3 em *T. cruzi* possui uma distribuição citoplasmática e perinuclear, diferente da observada em organismos-modelo onde é preferencialmente nucleolar. Levando em consideração os resultados de interações obtidos para TcImp3 bem como a localização celular torna interessante o estudo desta proteína não só no papel da biogênese de ribossomos bem como na via de exportação de mRNAs no parasita.

8 Referências Bibliográficas

- ABRUZZI, K. C., S. LACADIE e M. ROSBASH. Biochemical analysis of TREX complex recruitment to intronless and intron-containing yeast genes. *EMBO J*, v.23, n.13, Jul 7, p.2620-31, (2004).
- AGUILERA, A. Cotranscriptional mRNP assembly: from the DNA to the nuclear pore. *Curr Opin Cell Biol*, v.17, n.3, Jun, p.242-50, (2005).
- AKEF, A., H. ZHANG, S. MASUDA e A. F. PALAZZO. Trafficking of mRNAs containing ALREX-promoting elements through nuclear speckles. *Nucleus*, v.4, n.4, Jul-Aug, p.326-40, (2013).
- ALTSCHUL, S. F., T. L. MADDEN, A. A. SCHAFFER, J. ZHANG, Z. ZHANG, W. MILLER e D. J. LIPMAN. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res*, v.25, n.17, Sep 1, p.3389-402, (1997).
- ANDENMATTEN, N., S. EGARTER, A. J. JACKSON, N. JULLIEN, J. P. HERMAN e M. MEISSNER. Conditional genome engineering in *Toxoplasma gondii* uncovers alternative invasion mechanisms. *Nat Methods*, v.10, n.2, Feb, p.125-7, (2013).
- ANDERSEN, J. S., Y. W. LAM, A. K. LEUNG, S. E. ONG, C. E. LYON, A. I. LAMOND e M. MANN. Nucleolar proteome dynamics. *Nature*, v.433, n.7021, Jan 6, p.77-83, (2005).
- ANDRADE, Z. A. Immunopathology of Chagas disease. *Mem Inst Oswaldo Cruz*, v.94 Suppl 1, p.71-80, (1999).
- ANDREWS, N. W. Living dangerously: how *Trypanosoma cruzi* uses lysosomes to get inside host cells, and then escapes into the cytoplasm. *Biol Res*, v.26, n.1-2, p.65-7, (1993).
- ANSELMINI, A., F. PIFANO, J. A. SUAREZ e O. GURDIEL. Myocardiopathy in Chagas' disease. I. Comparative study of pathologic findings in chronic human and experimental Chagas' myocarditis. *Am Heart J*, v.72, n.4, Oct, p.469-81, (1966).
- ARAVIND, L. e E. V. KOONIN. Novel predicted RNA-binding domains associated with the translation machinery. *J Mol Evol*, v.48, n.3, Mar, p.291-302, (1999).
- ARMSTRONG, C. M. e D. E. GOLDBERG. An FKBP destabilization domain modulates protein levels in *Plasmodium falciparum*. *Nat Methods*, v.4, n.12, Dec, p.1007-9, (2007).
- AVILA, A. R., S. F. YAMADA-OGATTA, V. DA SILVA MONTEIRO, M. A. KRIEGER, C. V. NAKAMURA, W. DE SOUZA e S. GOLDENBERG. Cloning and characterization of the metacyclogenin gene, which is specifically expressed during *Trypanosoma cruzi* metacyclogenesis. *Mol Biochem Parasitol*, v.117, n.2, Oct, p.169-77, (2001).

AVILA, M. L., N. BERCOVICH, G. WESTERGAARD, M. J. LEVIN e M. P. VAZQUEZ. Mapping of the protein-binding interface between splicing factors SF3b155 and p14 of *Trypanosoma cruzi*. *Biochem Biophys Res Commun*, v.364, n.1, Dec 7, p.26-32, (2007).

BAHIA-OLIVEIRA, L. M., J. L. JONES, J. AZEVEDO-SILVA, C. C. ALVES, F. OREFICE e D. G. ADDISS. Highly endemic, waterborne toxoplasmosis in north Rio de Janeiro state, Brazil. *Emerg Infect Dis*, v.9, n.1, Jan, p.55-62, (2003).

BANASZYNSKI, L. A., L. C. CHEN, L. A. MAYNARD-SMITH, A. G. OOI e T. J. WANDLESS. A rapid, reversible, and tunable method to regulate protein function in living cells using synthetic small molecules. *Cell*, v.126, n.5, Sep 8, p.995-1004, (2006).

BARGIERI, D. Y., N. ANDENMATTEN, V. LAGAL, S. THIBERGE, J. A. WHITELAW, I. TARDIEUX, M. MEISSNER e R. MENARD. Apical membrane antigen 1 mediates apicomplexan parasite attachment but is dispensable for host cell invasion. *Nat Commun*, v.4, p.2552, (2013).

BASTIN, P., Z. BAGHERZADEH, K. R. MATTHEWS e K. GULL. A novel epitope tag system to study protein targeting and organelle biogenesis in *Trypanosoma brucei*. *Mol Biochem Parasitol*, v.77, n.2, May, p.235-9, (1996).

BEHNKE, M. S., J. B. RADKE, A. T. SMITH, W. J. SULLIVAN, JR. e M. W. WHITE. The transcription of bradyzoite genes in *Toxoplasma gondii* is controlled by autonomous promoter elements. *Mol Microbiol*, v.68, n.6, Jun, p.1502-18, (2008).

BELTRAO HDE, B., P. CERRONI MDE, D. R. FREITAS, A. Y. PINTO, C. VALENTE VDA, S. A. VALENTE, G. COSTA EDE e J. SOBEL. Investigation of two outbreaks of suspected oral transmission of acute Chagas disease in the Amazon region, Para State, Brazil, in 2007. *Trop Doct*, v.39, n.4, Oct, p.231-2, (2009).

BERTRAND, E., P. CHARTRAND, M. SCHAEFER, S. M. SHENOY, R. H. SINGER e R. M. LONG. Localization of ASH1 mRNA particles in living yeast. *Mol Cell*, v.2, n.4, Oct, p.437-45, (1998).

BEYER, A. L., M. E. CHRISTENSEN, B. W. WALKER e W. M. LESTOURGEON. Identification and characterization of the packaging proteins of core 40S hnRNP particles. *Cell*, v.11, n.1, May, p.127-38, (1977).

BHATTI, M. M. e W. J. SULLIVAN, JR. Histone acetylase GCN5 enters the nucleus via importin-alpha in protozoan parasite *Toxoplasma gondii*. *J Biol Chem*, v.280, n.7, Feb 18, p.5902-8, (2005).

BLOBEL, G. Gene gating: a hypothesis. *Proc Natl Acad Sci U S A*, v.82, n.24, Dec, p.8527-9, (1985).

BOGERD, H. P., R. A. FRIDELL, R. E. BENSON, J. HUA e B. R. CULLEN. Protein sequence requirements for function of the human T-cell leukemia virus type 1 Rex

nuclear export signal delineated by a novel in vivo randomization-selection assay. *Mol Cell Biol*, v.16, n.8, Aug, p.4207-14, (1996).

BONALDO, M. C., T. SOUTO-PADRON, W. DE SOUZA e S. GOLDENBERG. Cell-substrate adhesion during *Trypanosoma cruzi* differentiation. *J Cell Biol*, v.106, n.4, Apr, p.1349-58, (1988).

BOOTHROYD, J. C. e G. A. CROSS. Transcripts coding for variant surface glycoproteins of *Trypanosoma brucei* have a short, identical exon at their 5' end. *Gene*, v.20, n.2, Dec, p.281-9, (1982).

BOUGDOUR, A., L. BRAUN, D. CANNELLA e M. A. HAKIMI. Chromatin modifications: implications in the regulation of gene expression in *Toxoplasma gondii*. *Cell Microbiol*, v.12, n.4, Apr 1, p.413-23, (2010).

BROWN, J. M., J. GREEN, R. P. DAS NEVES, H. A. WALLACE, A. J. SMITH, J. HUGHES, N. GRAY, S. TAYLOR, W. G. WOOD, D. R. HIGGS, F. J. IBORRA e V. J. BUCKLE. Association between active genes occurs at nuclear speckles and is modulated by chromatin environment. *J Cell Biol*, v.182, n.6, Sep 22, p.1083-97, (2008).

BUBULYA, P. A., K. V. PRASANTH, T. J. DEERINCK, D. GERLICH, J. BEAUDOUIN, M. H. ELLISMAN, J. ELLENBERG e D. L. SPECTOR. Hypophosphorylated SR splicing factors transiently localize around active nucleolar organizing regions in telophase daughter nuclei. *J Cell Biol*, v.167, n.1, Oct 11, p.51-63, (2004).

CAMARGO, E. P. Growth and Differentiation in *Trypanosoma cruzi*. I. Origin of Metacyclic Trypanosomes in Liquid Media. *Rev Inst Med Trop Sao Paulo*, v.6, May-Jun, p.93-100, (1964).

CAMPBELL, D. A., S. THOMAS e N. R. STURM. Transcription in kinetoplastid protozoa: why be normal? *Microbes Infect*, v.5, n.13, Nov, p.1231-40, (2003).

CARO, F., N. BERCOVICH, C. ATORRASAGASTI, M. J. LEVIN e M. P. VAZQUEZ. *Trypanosoma cruzi*: analysis of the complete PUF RNA-binding protein family. *Exp Parasitol*, v.113, n.2, Jun, p.112-24, (2006).

CARVALHO, P. C., J. S. FISCHER, T. XU, D. COCIORVA, T. S. BALBUENA, R. H. VALENTE, J. PERALES, J. R. YATES, 3RD e V. C. BARBOSA. Search engine processor: Filtering and organizing peptide spectrum matches. *Proteomics*, v.12, n.7, Apr, p.944-9, (2012).

CARVALHO, P. C., J. S. FISCHER, T. XU, J. R. YATES, 3RD e V. C. BARBOSA. PatternLab: from mass spectra to label-free differential shotgun proteomics. *Curr Protoc Bioinformatics*, v.Chapter 13, Dec, p.Unit13 19, (2012).

CASANOVA, M., P. PORTALES, C. BLAINEAU, L. CROBU, P. BASTIEN e M. PAGES. Inhibition of active nuclear transport is an intrinsic trigger of programmed cell death in trypanosomatids. *Cell Death Differ*, v.15, n.12, Dec, p.1910-20, (2008).

CASSOLA, A. e A. C. FRASCH. An RNA recognition motif mediates the nucleocytoplasmic transport of a trypanosome RNA-binding protein. *J Biol Chem*, v.284, n.50, Dec 11, p.35015-28, (2009).

CASTELLANI, O., L. V. RIBEIRO e J. F. FERNANDES. Differentiation of *Trypanosoma cruzi* in culture. *J Protozool*, v.14, n.3, Aug, p.447-51, (1967).

CHAGAS, C. New human trypanosomiasis. Morphology and life cycle of *Schyzotrypanum cruzi*, the cause of a new human disease. *Mem. Inst. Oswaldo Cruz*. v. 1, p. 159-218, (1909).

CHAVEZ, S., T. BEILHARZ, A. G. RONDON, H. ERDJUMENT-BROMAGE, P. TEMPST, J. Q. SVEJSTRUP, T. LITHGOW e A. AGUILERA. A protein complex containing Tho2, Hpr1, Mft1 and a novel protein, Thp2, connects transcription elongation with mitotic recombination in *Saccharomyces cerevisiae*. *EMBO J*, v.19, n.21, Nov 1, p.5824-34, (2000).

CHEN, C. Y. e A. B. SHYU. AU-rich elements: characterization and importance in mRNA degradation. *Trends Biochem Sci*, v.20, n.11, Nov, p.465-70, (1995).

CLAYTON, C. e M. SHAPIRA. Post-transcriptional regulation of gene expression in trypanosomes and leishmanias. *Mol Biochem Parasitol*, v.156, n.2, Dec, p.93-101, (2007).

CLAYTON, C. E. Life without transcriptional control? From fly to man and back again. *EMBO J*, v.21, n.8, Apr 15, p.1881-8, (2002).

COLLART, M. A. e O. O. PANASENKO. The Ccr4--not complex. *Gene*, v.492, n.1, Jan 15, p.42-53, (2010).

CONTRERAS, V. T., C. M. MOREL e S. GOLDENBERG. Stage specific gene expression precedes morphological changes during *Trypanosoma cruzi* metacyclogenesis. *Mol Biochem Parasitol*, v.14, n.1, Jan, p.83-96, (1985).

CONTRERAS, V. T., J. M. SALLES, N. THOMAS, C. M. MOREL e S. GOLDENBERG. In vitro differentiation of *Trypanosoma cruzi* under chemically defined conditions. *Mol Biochem Parasitol*, v.16, n.3, Sep, p.315-27, (1985).

COOK, A., F. BONO, M. JINEK e E. CONTI. Structural biology of nucleocytoplasmic transport. *Annu Rev Biochem*, v.76, p.647-71, (2007).

COSNIER, B., M. KWAPISZ, I. HATIN, O. NAMY, S. HERMANN-LE DENMAT, A. MORILLON, J. P. ROUSSET e C. FABRET. A viable hypomorphic allele of the essential IMP3 gene reveals novel protein functions in *Saccharomyces cerevisiae*. *PLoS One*, v.6, n.4, p.e19500, (2011).

CROSS, M., B. WIELAND, Z. PALFI, A. GUNZL, U. ROTH LISBERGER, H. W. LAHM e A. BINDEREIF. The trans-spliceosomal U2 snRNP protein 40K of *Trypanosoma brucei*: cloning and analysis of functional domains reveals homology to a mammalian snRNP protein. EMBO J, v.12, n.3, Mar, p.1239-48, (1993).

CUEVAS, I. C., A. C. FRASCH e I. D'ORSO. Insights into a CRM1-mediated RNA-nuclear export pathway in *Trypanosoma cruzi*. Mol Biochem Parasitol, v.139, n.1, Jan, p.15-24, (2005).

CULLEN, B. R. Nuclear mRNA export: insights from virology. Trends Biochem Sci, v.28, n.8, Aug, p.419-24, (2003a).

CULLEN, B. R. Nuclear RNA export. J Cell Sci, v.116, n.Pt 4, Feb 15, p.587-97, (2003b).

CUSTODIO, N., C. CARVALHO, I. CONDADO, M. ANTONIOU, B. J. BLENCOWE e M. CARMO-FONSECA. In vivo recruitment of exon junction complex proteins to transcription sites in mammalian cell nuclei. RNA, v.10, n.4, Apr, p.622-33, (2004).

D'ORSO, I., J. G. DE GAUDENZI e A. C. FRASCH. RNA-binding proteins and mRNA turnover in trypanosomes. Trends Parasitol, v.19, n.4, Apr, p.151-5, (2003).

D'ORSO, I. e A. C. FRASCH. TcUBP-1, a developmentally regulated U-rich RNA-binding protein involved in selective mRNA destabilization in trypanosomes. J Biol Chem, v.276, n.37, Sep 14, p.34801-9, (2001).

D'ORSO, I. e A. C. FRASCH. TcUBP-1, an mRNA destabilizing factor from trypanosomes, homodimerizes and interacts with novel AU-rich element- and Poly(A)-binding proteins forming a ribonucleoprotein complex. J Biol Chem, v.277, n.52, Dec 27, p.50520-8, (2002).

DALMASSO, M. C., D. O. ONYANGO, A. NAGULESWARAN, W. J. SULLIVAN, JR. e S. O. ANGEL. Toxoplasma H2A variants reveal novel insights into nucleosome composition and functions for this histone family. J Mol Biol, v.392, n.1, Sep 11, p.33-47, (2009).

DARZACQ, X., R. H. SINGER e Y. SHAV-TAL. Dynamics of transcription and mRNA export. Curr Opin Cell Biol, v.17, n.3, Jun, p.332-9, (2005).

DE GAUDENZI, J., A. C. FRASCH e C. CLAYTON. RNA-binding domain proteins in Kinetoplastids: a comparative analysis. Eukaryot Cell, v.4, n.12, Dec, p.2106-14, (2005).

DE GAUDENZI, J. G., I. D'ORSO e A. C. FRASCH. RNA recognition motif-type RNA-binding proteins in *Trypanosoma cruzi* form a family involved in the interaction with specific transcripts in vivo. J Biol Chem, v.278, n.21, May 23, p.18884-94, (2003).

DE LANGE, T., T. M. BERKVEN, H. J. VEERMAN, A. C. FRASCH, J. D. BARRY e P. BORST. Comparison of the genes coding for the common 5' terminal sequence of

messenger RNAs in three trypanosome species. *Nucleic Acids Res*, v.12, n.11, Jun 11, p.4431-43, (1984).

DE SOUZA, W. A short review on the morphology of *Trypanosoma cruzi*: from 1909 to 1999. *Mem Inst Oswaldo Cruz*, v.94 Suppl 1, p.17-36, (1999).

DE SOUZA, W. Basic cell biology of *Trypanosoma cruzi*. *Curr Pharm Des*, v.8, n.4, p.269-85, (2002).

DE VRIES, H., U. RUEGSEGGER, W. HUBNER, A. FRIEDLEIN, H. LANGEN e W. KELLER. Human pre-mRNA cleavage factor II(m) contains homologs of yeast proteins and bridges two other cleavage factors. *EMBO J*, v.19, n.21, Nov 1, p.5895-904, (2000).

DEGRASSE, J. A., K. N. DUBOIS, D. DEVOS, T. N. SIEGEL, A. SALI, M. C. FIELD, M. P. ROUT e B. T. CHAIT. Evidence for a shared nuclear pore complex architecture that is conserved from the last common eukaryotic ancestor. *Mol Cell Proteomics*, v.8, n.9, Sep, p.2119-30, (2009).

DIAS, E., F. S. LARANJA, A. MIRANDA e G. NOBREGA. Chagas' disease; a clinical, epidemiologic, and pathologic study. *Circulation*, v.14, n.6, Dec, p.1035-60, (1956).

DIAS, J.C.P. Epidemiologia. Em: *Trypanosoma cruzi* e a Doença de Chagas. Zigman Brener; Zilton Andrade; Manoel Barral-Neto 2. ed. Rio de Janeiro: Editora Guanabara Koogan, (2000).

DIAS, A. P., K. DUFU, H. LEI e R. REED. A role for TREX components in the release of spliced mRNA from nuclear speckle domains. *Nat Commun*, v.1, p.97, (2010).

DLAKIC, M. The ribosomal subunit assembly line. *Genome Biol*, v.6, n.10, p.234, (2005).

DONALD, R. G., D. CARTER, B. ULLMAN e D. S. ROOS. Insertional tagging, cloning, and expression of the *Toxoplasma gondii* hypoxanthine-xanthine-guanine phosphoribosyltransferase gene. Use as a selectable marker for stable transformation. *J Biol Chem*, v.271, n.24, Jun 14, p.14010-9, (1996).

DONALD, R. G. e D. S. ROOS. Stable molecular transformation of *Toxoplasma gondii*: a selectable dihydrofolate reductase-thymidylate synthase marker based on drug-resistance mutations in malaria. *Proc Natl Acad Sci U S A*, v.90, n.24, Dec 15, p.11703-7, (1993).

DOSSIN FDE, M. e S. SCHENKMAN. Actively transcribing RNA polymerase II concentrates on spliced leader genes in the nucleus of *Trypanosoma cruzi*. *Eukaryot Cell*, v.4, n.5, May, p.960-70, (2005).

DOSTALOVA, A., S. KASER, M. CRISTODERO e B. SCHIMANSKI. The nuclear mRNA export receptor Mex67-Mtr2 of *Trypanosoma brucei* contains a unique and essential zinc finger motif. *Mol Microbiol*, v.88, n.4, May, p.728-39, (2013).

DOSTALOVA, A., S. KASER, M. CRISTODERO e B. SCHIMANSKI. The nuclear mRNA export receptor Mex67-Mtr2 of *Trypanosoma brucei* contains a unique and essential zinc finger motif. *Mol Microbiol*, v.88, n.4, May, p.728-39, (2013).

DRAGON, F., J. E. GALLAGHER, P. A. COMPAGNONE-POST, B. M. MITCHELL, K. A. PORWANCHER, K. A. WEHNER, S. WORMSLEY, R. E. SETTLAGE, J. SHABANOWITZ, Y. OSHEIM, A. L. BEYER, D. F. HUNT e S. J. BASERGA. A large nucleolar U3 ribonucleoprotein required for 18S ribosomal RNA biogenesis. *Nature*, v.417, n.6892, Jun 27, p.967-70, (2002).

DREYFUSS, G., V. N. KIM e N. KATAOKA. Messenger-RNA-binding proteins and the messages they carry. *Nat Rev Mol Cell Biol*, v.3, n.3, Mar, p.195-205, (2002).

DU, T. G., M. SCHMID e R. P. JANSEN. Why cells move messages: the biological functions of mRNA localization. *Semin Cell Dev Biol*, v.18, n.2, Apr, p.171-7, (2007).

DUBEY, J. P. Advances in the life cycle of *Toxoplasma gondii*. *Int J Parasitol*, v.28, n.7, Jul, p.1019-24, (1998).

DUBEY, J. P. The history of *Toxoplasma gondii*--the first 100 years. *J Eukaryot Microbiol*, v.55, n.6, Nov-Dec, p.467-75, (2008).

DUBEY, J. P. History of the discovery of the life cycle of *Toxoplasma gondii*. *Int J Parasitol*, v.39, n.8, Jul 1, p.877-82, (2009).

DUBEY, J. P., E. G. LAGO, S. M. GENNARI, C. SU e J. L. JONES. Toxoplasmosis in humans and animals in Brazil: high prevalence, high burden of disease, and epidemiology. *Parasitology*, v.139, n.11, Sep, p.1375-424, (2012).

DUNBAR, D. A., S. WORMSLEY, T. M. AGENTIS e S. J. BASERGA. Mpp10p, a U3 small nucleolar ribonucleoprotein component required for pre-18S rRNA processing in yeast. *Mol Cell Biol*, v.17, n.10, Oct, p.5803-12, (1997).

DURFEE, T.; BECHERER, K.; CHEN, P. L.; YEH, S. H.; YANG, Y.; KILBURN, A. E.; LEE, W. H.; ELLEDGE, S. J. The retinoblastoma protein associates with the protein phosphatase type 1 catalytic subunit. *Genes Dev*, v. 7, n. 4, p. 555-69, (1993).

EBERL, D. F., L. J. LORENZ, M. B. MELNICK, V. SOOD, P. LASKO e N. PERRIMON. A new enhancer of position-effect variegation in *Drosophila melanogaster* encodes a putative RNA helicase that binds chromosomes and is regulated by the cell cycle. *Genetics*, v.146, n.3, Jul, p.951-63, (1997).

EDDY, S. R. Profile hidden Markov models. *Bioinformatics*, v.14, n.9, p.755-63, (1998).

FABRE, E., W. C. BOELEN, C. WIMMER, I. W. MATTAJ e E. C. HURT. Nup145p is required for nuclear export of mRNA and binds homopolymeric RNA in vitro via a novel conserved motif. *Cell*, v.78, n.2, Jul 29, p.275-89, (1994).

FIELDS, S.; SONG, O. A novel genetic system to detect protein-protein interactions. *Nature*, v. 340, n. 6230, p. 245-6, (1989).

FINN, R. D., A. BATEMAN, J. CLEMENTS, P. COGGILL, R. Y. EBERHARDT, S. R. EDDY, A. HEGER, K. HETHERINGTON, L. HOLM, J. MISTRY, E. L. SONNHAMMER, J. TATE e M. PUNTA. Pfam: the protein families database. *Nucleic Acids Res*, v.42, n.Database issue, Jan, p.D222-30, (2014).

FISCHER, T., K. STRASSER, A. RACZ, S. RODRIGUEZ-NAVARRO, M. OPPIZZI, P. IHRIG, J. LECHNER e E. HURT. The mRNA export machinery requires the novel Sac3p-Thp1p complex to dock at the nucleoplasmic entrance of the nuclear pores. *EMBO J*, v.21, n.21, Nov 1, p.5843-52, (2002).

FLECKNER, J., M. ZHANG, J. VALCARCEL e M. R. GREEN. U2AF65 recruits a novel human DEAD box protein required for the U2 snRNP-branchpoint interaction. *Genes Dev*, v.11, n.14, Jul 15, p.1864-72, (1997).

FOUTS, D. E., H. L. TRUE e D. W. CELANDER. Functional recognition of fragmented operator sites by R17/MS2 coat protein, a translational repressor. *Nucleic Acids Res*, v.25, n.22, Nov 15, p.4464-73, (1997).

FOX, B. A., J. G. RISTUCCIA, J. P. GIGLEY e D. J. BZIK. Efficient gene replacements in *Toxoplasma gondii* strains deficient for nonhomologous end joining. *Eukaryot Cell*, v.8, n.4, Apr, p.520-9, (2009).

FRAGOSO, S. P., C. PLAZANET-MENUT, M. A. CARREIRA, M. C. MOTTA, B. DALLAGIOVANA, M. A. KRIEGER e S. GOLDENBERG. Cloning and characterization of a gene encoding a putative protein associated with U3 small nucleolar ribonucleoprotein in *Trypanosoma cruzi*. *Mol Biochem Parasitol*, v.126, n.1, Jan, p.113-7, (2003).

FRANCESCHINI, A., D. SZKLARCZYK, S. FRANKILD, M. KUHN, M. SIMONOVIC, A. ROTH, J. LIN, P. MINGUEZ, P. BORK, C. VON MERING e L. J. JENSEN. STRING v9.1: protein-protein interaction networks, with increased coverage and integration. *Nucleic Acids Res*, v.41, n.Database issue, Jan, p.D808-15, (2013).

FRANKEL, M. B. e L. J. KNOLL. Functional analysis of key nuclear trafficking components reveals an atypical Ran network required for parasite pathogenesis. *Mol Microbiol*, v.70, n.2, Oct, p.410-20, (2008).

FRANKEL, M. B. e L. J. KNOLL. The ins and outs of nuclear trafficking: unusual aspects in apicomplexan parasites. *DNA Cell Biol*, v.28, n.6, Jun, p.277-84, (2009).

FRANZ, C., P. ASKJAER, W. ANTONIN, C. L. IGLESIAS, U. HASELMANN, M. SCHELDER, A. DE MARCO, M. WILM, C. ANTONY e I. W. MATTAJ. Nup155 regulates nuclear envelope and nuclear pore complex formation in nematodes and vertebrates. *EMBO J*, v.24, n.20, Oct 19, p.3519-31, (2005).

FRASER, P. e W. BICKMORE. Nuclear organization of the genome and the potential for gene regulation. *Nature*, v.447, n.7143, May 24, p.413-7, (2007).

FRIBOURG, S., I. C. BRAUN, E. IZAURRALDE e E. CONTI. Structural basis for the recognition of a nucleoporin FG repeat by the NTF2-like domain of the TAP/p15 mRNA nuclear export factor. *Mol Cell*, v.8, n.3, Sep, p.645-56, (2001).

FUSCO, D., N. ACCORNERO, B. LAVOIE, S. M. SHENOY, J. M. BLANCHARD, R. H. SINGER e E. BERTRAND. Single mRNA molecules demonstrate probabilistic movement in living mammalian cells. *Curr Biol*, v.13, n.2, Jan 21, p.161-7, (2003).

GARCIA, E. S. e P. AZAMBUJA. Development and interactions of *Trypanosoma cruzi* within the insect vector. *Parasitol Today*, v.7, n.9, Sep, p.240-4, (1991).

GATFIELD, D., H. LE HIR, C. SCHMITT, I. C. BRAUN, T. KOCHER, M. WILM e E. IZAURRALDE. The DExH/D box protein HEL/UAP56 is essential for mRNA nuclear export in *Drosophila*. *Curr Biol*, v.11, n.21, Oct 30, p.1716-21, (2001).

GEIDUSCHEK, E. P. e G. A. KASSAVETIS. The RNA polymerase III transcription apparatus. *J Mol Biol*, v.310, n.1, Jun 29, p.1-26, (2001).

GILBERT, R. E., K. FREEMAN, E. G. LAGO, L. M. BAHIA-OLIVEIRA, H. K. TAN, M. WALLON, W. BUFFOLANO, M. R. STANFORD e E. PETERSEN. Ocular sequelae of congenital toxoplasmosis in Brazil compared with Europe. *PLoS Negl Trop Dis*, v.2, n.8, p.e277, (2008).

GILBERT, R. E. e M. R. STANFORD. Is ocular toxoplasmosis caused by prenatal or postnatal infection? *Br J Ophthalmol*, v.84, n.2, Feb, p.224-6, (2000).

GLISOVIC, T., J. L. BACHORIK, J. YONG e G. DREYFUSS. RNA-binding proteins and post-transcriptional gene regulation. *FEBS Lett*, v.582, n.14, Jun 18, p.1977-86, (2008).

GOLDENBERG, S.; CONTRERAS, V.T.; SALLES, J.M.; FRANCO, M.P.A.L.; BONALDO, M.C.; VALLE, D.; GONÇALVES, A.M.; MOREL, C.M. Perspectives of vaccination against Chagas Disease through biotechnology. II – Gene expression in *T. cruzi* trypomastigotes and cell-free translation of mRNAs coding for relevant surface antigens. *New Approaches to Vaccine Development*. TORRIGIANI, O. (Ed). Schwabe e Co., Basel, Switzerland, p. 442-459, (1984).

GOLDENBERG, S.; CONTRERAS, V.T.; BONALDO, M.C.; SALLES, J.M.; FRANCO, M.P.A.L.; LAFAILLE, J.; GONZALES-PERDOMO, M.; LINSS, J.; MOREL, C.M. *In vitro* differentiation systems for the study of differential gene expression during *Trypanosoma cruzi* development. *Molecular Strategies of Parasitic Invasion*. p. 203-212, (1987).

GONZALEZ-AGUILERA, C., C. TOUS, B. GOMEZ-GONZALEZ, P. HUERTAS, R. LUNA e A. AGUILERA. The THP1-SAC3-SUS1-CDC31 complex works in transcription elongation-mRNA export preventing RNA-mediated genome instability. *Mol Biol Cell*, v.19, n.10, Oct, p.4310-8, (2008).

GONZALEZ, M. S., N. F. NOGUEIRA, C. B. MELLO, W. DE SOUZA, G. A. SCHAUB, P. AZAMBUJA e E. S. GARCIA. Influence of brain and azadirachtin on *Trypanosoma cruzi* development in the vector, *Rhodnius prolixus*. *Exp Parasitol*, v.92, n.2, Jun, p.100-8, (1999).

GRANDI, P., V. RYBIN, J. BASSLER, E. PETFALSKI, D. STRAUSS, M. MARZIOCH, T. SCHAFFER, B. KUSTER, H. TSCHOCHNER, D. TOLLERVEY, A. C. GAVIN e E. HURT. 90S pre-ribosomes include the 35S pre-rRNA, the U3 snoRNP, and 40S subunit processing factors but predominantly lack 60S synthesis factors. *Mol Cell*, v.10, n.1, Jul, p.105-15, (2002).

GRANNEMAN, S., J. E. GALLAGHER, J. VOGELZANGS, W. HORSTMAN, W. J. VAN VENROOIJ, S. J. BASERGA e G. J. PRUIJN. The human Imp3 and Imp4 proteins form a ternary complex with hMpp10, which only interacts with the U3 snoRNA in 60-80S ribonucleoprotein complexes. *Nucleic Acids Res*, v.31, n.7, Apr 1, p.1877-87, (2003).

GRAVELEY, B. R. Sorting out the complexity of SR protein functions. *RNA*, v.6, n.9, Sep, p.1197-211, (2000).

GUNZL, A. The pre-mRNA splicing machinery of trypanosomes: complex or simplified? *Eukaryot Cell*, v.9, n.8, Aug, p.1159-70, (2010).

GYURIS, J.; GOLEMIS, E.; CHERTKOV, H.; BRENT, R. Cdi1, a human G1 and S phase protein phosphatase that associates with Cdk2. *Cell*, v. 75, n. 4, p. 791-803, (1993).

HADDAD, R., F. MAURICE, N. VIPHAKONE, F. VOISINET-HAKIL, S. FRIBOURG e L. MINVIELLE-SEBASTIA. An essential role for Clp1 in assembly of polyadenylation complex CF IA and Pol II transcription termination. *Nucleic Acids Res*, v.40, n.3, Feb, p.1226-39, (2012).

HAILE, S. e B. PAPADOPOULOU. Developmental regulation of gene expression in trypanosomatid parasitic protozoa. *Curr Opin Microbiol*, v.10, n.6, Dec, p.569-77, (2007).

HARPER, J. T., E. WAANDERS e P. J. KEELING. On the monophyly of chromalveolates using a six-protein phylogeny of eukaryotes. *Int J Syst Evol Microbiol*, v.55, n.Pt 1, Jan, p.487-96, (2005).

HAUTBERGUE, G. M., M. L. HUNG, M. J. WALSH, A. P. SNIJDERS, C. T. CHANG, R. JONES, C. P. PONTING, M. J. DICKMAN e S. A. WILSON. UIF, a New mRNA export adaptor that works together with REF/ALY, requires FACT for recruitment to mRNA. *Curr Biol*, v.19, n.22, Dec 1, p.1918-24, (2009).

HERM-GOTZ, A., C. AGOP-NERSESIAN, S. MUNTER, J. S. GRIMLEY, T. J. WANDLESS, F. FRISCHKNECHT e M. MEISSNER. Rapid control of protein level in the apicomplexan *Toxoplasma gondii*. *Nat Methods*, v.4, n.12, Dec, p.1003-5, (2007).

HEROLD, A., L. TEIXEIRA e E. IZAURRALDE. Genome-wide analysis of nuclear mRNA export pathways in *Drosophila*. *EMBO J*, v.22, n.10, May 15, p.2472-83, (2003).

HETTMANN C., H. A., GEITER A., FRANK B., SCHWARZ E., SOLDATI T., SOLDATI D. A dibasic motif in the tail of a class XIV apicomplexan myosin is an essential determinant of plasma membrane localization. *Mol Biol Cell*, v.11, n.4, (2000).

HIDE, G. History of sleeping sickness in East Africa. *Clin Microbiol Rev*, v.12, n.1, Jan, p.112-25, (1999).

HILL, D. E., S. CHIRUKANDOTH e J. P. DUBEY. Biology and epidemiology of *Toxoplasma gondii* in man and animals. *Anim Health Res Rev*, v.6, n.1, Jun, p.41-61, (2005).

HO, J. H., G. KALLSTROM e A. W. JOHNSON. Nmd3p is a Crm1p-dependent adapter protein for nuclear export of the large ribosomal subunit. *J Cell Biol*, v.151, n.5, Nov 27, p.1057-66, (2000).

HOARE, C. A. [Rationalization of the terminology for the developmental stages of trypanosomatid flagellates]. *Med Parazitol (Mosk)*, v.40, n.3, May-Jun, p.307-9, (1971).

HOCINE, S., P. RAYMOND, D. ZENKLUSEN, J. A. CHAO e R. H. SINGER. Single-molecule analysis of gene expression using two-color RNA labeling in live yeast. *Nat Methods*, v.10, n.2, Feb, p.119-21, (2013).

HOESS, R. H. e K. ABREMSKI. Mechanism of strand cleavage and exchange in the Cre-lox site-specific recombination system. *J Mol Biol*, v.181, n.3, Feb 5, p.351-62, (1985).

HOLLAND, G. N. Reconsidering the pathogenesis of ocular toxoplasmosis. *Am J Ophthalmol*, v.128, n.4, Oct, p.502-5, (1999).

HOLLAND, G. N. Ocular toxoplasmosis: a global reassessment. Part I: epidemiology and course of disease. *Am J Ophthalmol*, v.136, n.6, Dec, p.973-88, (2003).

HSU, C. K., G. R. CAMPBELL e N. D. LEVINE. A check-list of the species of the genus *Leucocytozoon* (Apicomplexa, Plasmodiidae). *J Protozool*, v.20, n.2, May, p.195-203, (1973).

HUANG, M., J. E. RECH, S. J. NORTHINGTON, P. F. FLICKER, A. MAYEDA, A. R. KRAINER e W. M. LESTOURGEON. The C-protein tetramer binds 230 to 240 nucleotides of pre-mRNA and nucleates the assembly of 40S heterogeneous nuclear ribonucleoprotein particles. *Mol Cell Biol*, v.14, n.1, Jan, p.518-33, (1994).

- HUANG, S., T. J. DEERINCK, M. H. ELLISMAN e D. L. SPECTOR. The perinucleolar compartment and transcription. *J Cell Biol*, v.143, n.1, Oct 5, p.35-47, (1998).
- HUANG, Y., R. GATTONI, J. STEVENIN e J. A. STEITZ. SR splicing factors serve as adapter proteins for TAP-dependent mRNA export. *Mol Cell*, v.11, n.3, Mar, p.837-43, (2003).
- HUANG, Y. e J. A. STEITZ. Splicing factors SRp20 and 9G8 promote the nucleocytoplasmic export of mRNA. *Mol Cell*, v.7, n.4, Apr, p.899-905, (2001).
- HUANG, Y. e J. A. STEITZ. SRprises along a messenger's journey. *Mol Cell*, v.17, n.5, Mar 4, p.613-5, (2005).
- HUANG, Y., T. A. YARIO e J. A. STEITZ. A molecular link between SR protein dephosphorylation and mRNA export. *Proc Natl Acad Sci U S A*, v.101, n.26, Jun 29, p.9666-70, (2004).
- HURT, E., M. J. LUO, S. ROTHER, R. REED e K. STRASSER. Cotranscriptional recruitment of the serine-arginine-rich (SR)-like proteins Gbp2 and Hrb1 to nascent mRNA via the TREX complex. *Proc Natl Acad Sci U S A*, v.101, n.7, Feb 17, p.1858-62, (2004).
- HUTTEN, S. e R. H. KEHLENBACH. CRM1-mediated nuclear export: to the pore and beyond. *Trends Cell Biol*, v.17, n.4, Apr, p.193-201, (2007).
- HUYNH, M. H. e V. B. CARRUTHERS. Tagging of endogenous genes in a *Toxoplasma gondii* strain lacking Ku80. *Eukaryot Cell*, v.8, n.4, Apr, p.530-9, (2009).
- IGLESIAS, N. e F. STUTZ. Regulation of mRNP dynamics along the export pathway. *FEBS Lett*, v.582, n.14, Jun 18, p.1987-96, (2008).
- IKUTA, T., H. EGUCHI, T. TACHIBANA, Y. YONEDA e K. KAWAJIRI. Nuclear localization and export signals of the human aryl hydrocarbon receptor. *J Biol Chem*, v.273, n.5, Jan 30, p.2895-904, (1998).
- IZAURRALDE, E., A. JARMOLOWSKI, C. BEISEL, I. W. MATTAJ, G. DREYFUSS e U. FISCHER. A role for the M9 transport signal of hnRNP A1 in mRNA nuclear export. *J Cell Biol*, v.137, n.1, Apr 7, p.27-35, (1997).
- JAE, N., P. WANG, T. GU, M. HUHN, Z. PALFI, H. URLAUB e A. BINDEREIF. Essential role of a trypanosome U4-specific Sm core protein in small nuclear ribonucleoprotein assembly and splicing. *Eukaryot Cell*, v.9, n.3, Mar, p.379-86, (2010).
- JANI, D., S. LUTZ, E. HURT, R. A. LASKEY, M. STEWART e V. O. WICKRAMASINGHE. Functional and structural characterization of the mammalian TREX-2 complex that links transcription with nuclear messenger RNA export. *Nucleic Acids Res*, v.40, n.10, May, p.4562-73, (2012).

JANSEN, R., D. TOLLERVEY e E. C. HURT. A U3 snoRNP protein with homology to splicing factor PRP4 and G beta domains is required for ribosomal RNA processing. *EMBO J*, v.12, n.6, Jun, p.2549-58, (1993).

JEKELY, G. Small GTPases and the evolution of the eukaryotic cell. *Bioessays*, v.25, n.11, Nov, p.1129-38, (2003).

JENSEN, T. H., J. BOULAY, M. ROSBASH e D. LIBRI. The DECD box putative ATPase Sub2p is an early mRNA export factor. *Curr Biol*, v.11, n.21, Oct 30, p.1711-5, (2001).

JOHNSON, P. J., J. M. KOOTER e P. BORST. Inactivation of transcription by UV irradiation of *T. brucei* provides evidence for a multicistronic transcription unit including a VSG gene. *Cell*, v.51, n.2, Oct 23, p.273-81, (1987).

JULLIEN, N., I. GODDARD, S. SELMI-RUBY, J. L. FINA, H. CREMER e J. P. HERMAN. Conditional transgenesis using Dimerizable Cre (DiCre). *PLoS One*, v.2, n.12, p.e1355, (2007).

JULLIEN, N., F. SAMPIERI, A. ENJALBERT e J. P. HERMAN. Regulation of Cre recombinase by ligand-induced complementation of inactive fragments. *Nucleic Acids Res*, v.31, n.21, Nov 1, p.e131, (2003).

JURICA, M. S. e M. J. MOORE. Pre-mRNA splicing: awash in a sea of proteins. *Mol Cell*, v.12, n.1, Jul, p.5-14, (2003).

KAMMEL, C., M. THOMAIER, B. B. SORENSEN, T. SCHUBERT, G. LANGST, M. GRASSER e K. D. GRASSER. Arabidopsis DEAD-box RNA helicase UAP56 interacts with both RNA and DNA as well as with mRNA export factors. *PLoS One*, v.8, n.3, p.e60644, (2013).

KAPADIA, F., A. PRYOR, T. H. CHANG e L. F. JOHNSON. Nuclear localization of poly(A)⁺ mRNA following siRNA reduction of expression of the mammalian RNA helicases UAP56 and URH49. *Gene*, v.384, Dec 15, p.37-44, (2006).

KASS, S., K. TYC, J. A. STEITZ e B. SOLLNER-WEBB. The U3 small nucleolar ribonucleoprotein functions in the first step of preribosomal RNA processing. *Cell*, v.60, n.6, Mar 23, p.897-908, (1990).

KATAHIRA, J., K. STRASSER, A. PODTELEJNIKOV, M. MANN, J. U. JUNG e E. HURT. The Mex67p-mediated nuclear mRNA export pathway is conserved from yeast to human. *EMBO J*, v.18, n.9, May 4, p.2593-609, (1999).

KATAOKA, N., J. YONG, V. N. KIM, F. VELAZQUEZ, R. A. PERKINSON, F. WANG e G. DREYFUSS. Pre-mRNA splicing imprints mRNA in the nucleus with a novel RNA-binding protein that persists in the cytoplasm. *Mol Cell*, v.6, n.3, Sep, p.673-82, (2000).

KATO, K. e D. M. STANDLEY. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol*, v.30, n.4, Apr, p.772-80, (2013).

KEENE, J. D. Ribonucleoprotein infrastructure regulating the flow of genetic information between the genome and the proteome. *Proc Natl Acad Sci U S A*, v.98, n.13, Jun 19, p.7018-24, (2001).

KERR, S. C., N. AZZOUZ, S. M. FUCHS, M. A. COLLART, B. D. STRAHL, A. H. CORBETT e R. N. LARIBEE. The Ccr4-Not complex interacts with the mRNA export machinery. *PLoS One*, v.6, n.3, p.e18302, (2011).

KHAN, A., C. JORDAN, C. MUCCIOLI, A. L. VALLOCHI, L. V. RIZZO, R. BELFORT, JR., R. W. VITOR, C. SILVEIRA e L. D. SIBLEY. Genetic divergence of *Toxoplasma gondii* strains associated with ocular toxoplasmosis, Brazil. *Emerg Infect Dis*, v.12, n.6, Jun, p.942-9, (2006).

KIESLER, E., F. MIRALLES e N. VISA. HEL/UAP56 binds cotranscriptionally to the Balbiani ring pre-mRNA in an intron-independent manner and accompanies the BR mRNP to the nuclear pore. *Curr Biol*, v.12, n.10, May 14, p.859-62, (2002).

KIM, K. e L. M. WEISS. *Toxoplasma gondii*: the model apicomplexan. *Int J Parasitol*, v.34, n.3, Mar 9, p.423-32, (2004).

KLOCKENBUSCH, C. e J. KAST. Optimization of formaldehyde cross-linking for protein interaction analysis of non-tagged integrin beta1. *J Biomed Biotechnol*, v.2010, p.927585, (2010).

KOHLER, A. e E. HURT. Exporting RNA from the nucleus to the cytoplasm. *Nat Rev Mol Cell Biol*, v.8, n.10, Oct, p.761-73, (2007).

KOHLER, S., C. F. DELWICHE, P. W. DENNY, L. G. TILNEY, P. WEBSTER, R. J. WILSON, J. D. PALMER e D. S. ROOS. A plastid of probable green algal origin in Apicomplexan parasites. *Science*, v.275, n.5305, Mar 7, p.1485-9, (1997).

KOTA, K. P., S. R. WAGNER, E. HUERTA, J. M. UNDERWOOD e J. A. NICKERSON. Binding of ATP to UAP56 is necessary for mRNA export. *J Cell Sci*, v.121, n.Pt 9, May 1, p.1526-37, (2008).

KRAMER, S., N. C. KIMBLIN e M. CARRINGTON. Genome-wide in silico screen for CCCH-type zinc finger proteins of *Trypanosoma brucei*, *Trypanosoma cruzi* and *Leishmania major*. *BMC Genomics*, v.11, p.283, (2010).

KREMER, K., D. KAMIN, E. RITTWEGER, J. WILKES, H. FLAMMER, S. MAHLER, J. HENG, C. J. TONKIN, G. LANGSLEY, S. W. HELL, V. B. CARRUTHERS, D. J. FERGUSON e M. MEISSNER. An overexpression screen of *Toxoplasma gondii* Rab-GTPases reveals distinct transport routes to the micronemes. *PLoS Pathog*, v.9, n.3, Mar, p.e1003213, (2013).

KROPF, S.P.; AZEVEDO, N.; FERREIRA, L.O. Doença de Chagas: a construção de um fato científico e de um problema de saúde pública no Brasil. *Ciênc.Saúde Coletiva*. v. 5, n. 2. (2000).

KUDO, N., N. MATSUMORI, H. TAOKA, D. FUJIWARA, E. P. SCHREINER, B. WOLFF, M. YOSHIDA e S. HORINOUCI. Leptomycin B inactivates CRM1/exportin 1 by covalent modification at a cysteine residue in the central conserved region. *Proc Natl Acad Sci U S A*, v.96, n.16, Aug 3, p.9112-7, (1999).

KUMAR, A. e T. PEDERSON. Comparison of proteins bound to heterogeneous nuclear RNA and messenger RNA in HeLa cells. *J Mol Biol*, v.96, n.3, Aug 15, p.353-65, (1975).

LAI, M. C. e W. Y. TARN. Hypophosphorylated ASF/SF2 binds TAP and is present in messenger ribonucleoproteins. *J Biol Chem*, v.279, n.30, Jul 23, p.31745-9, (2004).

LAMOND, A. I. e D. L. SPECTOR. Nuclear speckles: a model for nuclear organelles. *Nat Rev Mol Cell Biol*, v.4, n.8, Aug, p.605-12, (2003).

LANAR, D. E. Growth and differentiation of *Trypanosoma cruzi* cultivated with a *Triatoma infestans* embryo cell line. *J Protozool*, v.26, n.3, Aug, p.457-62, (1979).

LARANJA, F. S. e Z. A. ANDRADE. [Chronic cardiac form of chagas disease in dogs]. *Arq Bras Cardiol*, v.35, n.5, Nov, p.377-80, (1980).

LAUFER, G., G. SCHAAF, S. BOLLGONN e A. GUNZL. In vitro analysis of alpha-amanitin-resistant transcription from the rRNA, procyclic acidic repetitive protein, and variant surface glycoprotein gene promoters in *Trypanosoma brucei*. *Mol Cell Biol*, v.19, n.8, Aug, p.5466-73, (1999).

LEARY, D. J., M. P. TERNS e S. HUANG. Components of U3 snoRNA-containing complexes shuttle between nuclei and the cytoplasm and differentially localize in nucleoli: implications for assembly and function. *Mol Biol Cell*, v.15, n.1, Jan, p.281-93, (2004).

LEBOWITZ, J. H., H. Q. SMITH, L. RUSCHE e S. M. BEVERLEY. Coupling of poly(A) site selection and trans-splicing in *Leishmania*. *Genes Dev*, v.7, n.6, Jun, p.996-1007, (1993).

LEE, M. S., M. HENRY e P. A. SILVER. A protein that shuttles between the nucleus and the cytoplasm is an important mediator of RNA export. *Genes Dev*, v.10, n.10, May 15, p.1233-46, (1996).

LEE, S. J. e S. J. BASERGA. Imp3p and Imp4p, two specific components of the U3 small nucleolar ribonucleoprotein that are essential for pre-18S rRNA processing. *Mol Cell Biol*, v.19, n.8, Aug, p.5441-52, (1999).

LENZI, H. L., D. N. OLIVEIRA, M. T. LIMA e C. R. GATTASS. *Trypanosoma cruzi*: paninfectivity of CL strain during murine acute infection. *Exp Parasitol*, v.84, n.1, Oct, p.16-27, (1996).

LEVINE, N. D. Progress in taxonomy of the Apicomplexan protozoa. *J Protozool*, v.35, n.4, Nov, p.518-20, (1988).

LEVINE, N. D., J. O. CORLISS, F. E. COX, G. DEROUX, J. GRAIN, B. M. HONIGBERG, G. F. LEEDALE, A. R. LOEBLICH, 3RD, J. LOM, D. LYNN, E. G. MERINFELD, F. C. PAGE, G. POLJANSKY, V. SPRAGUE, J. VAVRA e F. G. WALLACE. A newly revised classification of the protozoa. *J Protozool*, v.27, n.1, Feb, p.37-58, (1980).

LIANG, X. H., A. HARITAN, S. ULIEL e S. MICHAELI. trans and cis splicing in trypanosomatids: mechanism, factors, and regulation. *Eukaryot Cell*, v.2, n.5, Oct, p.830-40, (2003).

LIBRI, D., N. GRAZIANI, C. SAGUEZ e J. BOULAY. Multiple roles for the yeast SUB2yUAP56 gene in splicing. *Genes Dev*, v.15, n.1, Jan 1, p.36-41, (2001).

LINDER, P. e F. STUTZ. mRNA export: travelling with DEAD box proteins. *Curr Biol*, v.11, n.23, Nov 27, p.R961-3, (2001).

LIONNET, T., K. CZAPLINSKI, X. DARZACQ, Y. SHAV-TAL, A. L. WELLS, J. A. CHAO, H. Y. PARK, V. DE TURRIS, M. LOPEZ-JONES e R. H. SINGER. A transgenic mouse for in vivo detection of endogenous labeled mRNA. *Nat Methods*, v.8, n.2, Feb, p.165-70, (2011).

LIRUSSI, D. e M. MATRAJT. RNA granules present only in extracellular *Toxoplasma gondii* increase parasite viability. *Int J Biol Sci*, v.7, n.7, p.960-7, (2011).

LIU, Q., X. H. LIANG, S. ULIEL, M. BELAHCEN, R. UNGER e S. MICHAELI. Identification and functional characterization of lsm proteins in *Trypanosoma brucei*. *J Biol Chem*, v.279, n.18, Apr 30, p.18210-9, (2004).

LONG, J. C. e J. F. CACERES. The SR protein family of splicing factors: master regulators of gene expression. *Biochem J*, v.417, n.1, Jan 1, p.15-27, (2009).

LUCKE, S., K. JURCHOTT, L. H. HUNG e A. BINDEREIF. mRNA splicing in *Trypanosoma brucei*: branch-point mapping reveals differences from the canonical U2 snRNA-mediated recognition. *Mol Biochem Parasitol*, v.142, n.2, Aug, p.248-51, (2005).

LUO, M. L., Z. ZHOU, K. MAGNI, C. CHRISTOFORIDES, J. RAPPSILBER, M. MANN e R. REED. Pre-mRNA splicing and mRNA export linked by direct interactions between UAP56 and Aly. *Nature*, v.413, n.6856, Oct 11, p.644-7, (2001).

MA, Y. F., L. M. WEISS e H. HUANG. A method for rapid regulation of protein expression in *Trypanosoma cruzi*. *Int J Parasitol*, v.42, n.1, Jan, p.33-7, (2011).

MACMORRIS, M., C. BROCKER e T. BLUMENTHAL. UAP56 levels affect viability and mRNA export in *Caenorhabditis elegans*. *RNA*, v.9, n.7, Jul, p.847-57, (2003).

MADEIRA DA SILVA, L., K. L. OWENS, S. M. MURTA e S. M. BEVERLEY. Regulated expression of the *Leishmania major* surface virulence factor

lipophosphoglycan using conditionally destabilized fusion proteins. Proc Natl Acad Sci U S A, v.106, n.18, May 5, p.7583-8, (2009).

MAIR, G., H. SHI, H. LI, A. DJIKENG, H. O. AVILES, J. R. BISHOP, F. H. FALCONE, C. GAVRILESCU, J. L. MONTGOMERY, M. I. SANTORI, L. S. STERN, Z. WANG, E. ULLU e C. TSCHUDI. A new twist in trypanosome RNA metabolism: cis-splicing of pre-mRNA. RNA, v.6, n.2, Feb, p.163-9, (2000).

MANN, T. e C. BECKERS. Characterization of the subpellicular network, a filamentous membrane skeletal component in the parasite *Toxoplasma gondii*. Mol Biochem Parasitol, v.115, n.2, Jul, p.257-68, (2001).

MARCHLER-BAUER, A., S. LU, J. B. ANDERSON, F. CHITSAZ, M. K. DERBYSHIRE, C. DEWEESE-SCOTT, J. H. FONG, L. Y. GEER, R. C. GEER, N. R. GONZALES, M. GWADZ, D. I. HURWITZ, J. D. JACKSON, Z. KE, C. J. LANCZYCKI, F. LU, G. H. MARCHLER, M. MULLOKANDOV, M. V. OMELCHENKO, C. L. ROBERTSON, J. S. SONG, N. THANKI, R. A. YAMASHITA, D. ZHANG, N. ZHANG, C. ZHENG e S. H. BRYANT. CDD: a Conserved Domain Database for the functional annotation of proteins. Nucleic Acids Res, v.39, n.Database issue, Jan, p.D225-9, (2011).

MARECHAL, E. e M. F. CESBRON-DELAUW. The apicoplast: a new member of the plastid family. Trends Plant Sci, v.6, n.5, May, p.200-5, (2001).

MARKUS, M. A., R. B. GERSTNER, D. E. DRAPER e D. A. TORCHIA. The solution structure of ribosomal protein S4 delta41 reveals two subdomains and a positively charged surface that may interact with RNA. EMBO J, v.17, n.16, Aug 17, p.4559-71, (1998).

MASUDA, S., R. DAS, H. CHENG, E. HURT, N. DORMAN e R. REED. Recruitment of the human TREX complex to mRNA during splicing. Genes Dev, v.19, n.13, Jul 1, p.1512-7, (2005).

MATOS, S. B., A. L. JESUS, K. C. PEDROZA, H. R. SODRE, T. L. FERREIRA e F. W. LIMA. Prevalence of serological markers and risk factors for bloodborne pathogens in Salvador, Bahia state, Brazil. Epidemiol Infect, v.141, n.1, Jan, p.181-7, (2013).

MATTHEWS, K. R., C. TSCHUDI e E. ULLU. A common pyrimidine-rich motif governs trans-splicing and polyadenylation of tubulin polycistronic pre-mRNA in trypanosomes. Genes Dev, v.8, n.4, Feb 15, p.491-501, (1994).

MAYAS, R. M., H. MAITA e J. P. STALEY. Exon ligation is proofread by the DExD/H-box ATPase Prp22p. Nat Struct Mol Biol, v.13, n.6, Jun, p.482-90, (2006).

MCAFEE, J. G., L. SHAHIED-MILAM, S. R. SOLTANINASSAB e W. M. LESTOURGEON. A major determinant of hnRNP C protein binding to RNA is a novel bZIP-like RNA binding domain. RNA, v.2, n.11, Nov, p.1139-52, (1996).

MCCLOSKEY, A., I. TANIGUCHI, K. SHINMYOZU e M. OHNO. hnRNP C tetramer measures RNA length to classify RNA polymerase II transcripts for export. *Science*, v.335, n.6076, Mar 30, p.1643-6, (2012).

MCDONALD, W. H., D. L. TABB, R. G. SADYGOV, M. J. MACCOSS, J. VENABLE, J. GRAUMANN, J. R. JOHNSON, D. COCIORVA e J. R. YATES, 3RD. MS1, MS2, and SQT-three unified, compact, and easily parsed file formats for the storage of shotgun proteomic spectra and identifications. *Rapid Commun Mass Spectrom*, v.18, n.18, p.2162-8, (2004).

MEISSNER, M., M. S. BREINICH, P. R. GILSON e B. S. CRABB. Molecular genetic tools in *Toxoplasma* and *Plasmodium*: achievements and future needs. *Curr Opin Microbiol*, v.10, n.4, Aug, p.349-56, (2007).

MEISSNER, M., D. SCHLUTER e D. SOLDATI. Role of *Toxoplasma gondii* myosin A in powering parasite gliding and host cell invasion. *Science*, v.298, n.5594, Oct 25, p.837-40, (2002).

MEYER, H. e W. DE SOUZA. Electron microscopic study of *Trypanosoma cruzi* periplast in tissue cultures. I. Number and arrangement of the peripheral microtubules in the various forms of the parasite's life cycle. *J Protozool*, v.23, n.3, Aug, p.385-90, (1976).

MILLER, S. A., V. THATHY, J. W. AJIOKA, M. J. BLACKMAN e K. KIM. TgSUB2 is a *Toxoplasma gondii* rhoptry organelle processing proteinase. *Mol Microbiol*, v.49, n.4, Aug, p.883-94, (2003).

MONTOYA, J. G. e O. LIESENFELD. Toxoplasmosis. *Lancet*, v.363, n.9425, Jun 12, p.1965-76, (2004).

MORRISSETTE, N. S., V. BEDIAN, P. WEBSTER e D. S. ROOS. Characterization of extreme apical antigens from *Toxoplasma gondii*. *Exp Parasitol*, v.79, n.3, Nov, p.445-59, (1994).

MOY, T. I. e P. A. SILVER. Requirements for the nuclear export of the small ribosomal subunit. *J Cell Sci*, v.115, n.Pt 14, Jul 15, p.2985-95, (2002).

MOYLE, M., T. HOFMANN e C. J. INGLES. The RPO31 gene of *Saccharomyces cerevisiae* encodes the largest subunit of RNA polymerase III. *Biochem Cell Biol*, v.64, n.8, Aug, p.717-21, (1986).

MULLER-MCNICOLL, M. e K. M. NEUGEBAUER. How cells get the message: dynamic assembly and function of mRNA-protein complexes. *Nat Rev Genet*, v.14, n.4, Apr, p.275-87, (2013).

NARDELLI, S. C., A. R. AVILA, A. FREUND, M. C. MOTTA, L. MANHAES, T. C. DE JESUS, S. SCHENKMAN, S. P. FRAGOSO, M. A. KRIEGER, S. GOLDENBERG e B. DALLAGIOVANNA. Small-subunit rRNA processome proteins are translationally

regulated during differentiation of *Trypanosoma cruzi*. *Eukaryot Cell*, v.6, n.2, Feb, p.337-45, (2007).

NELSON, R. G., M. PARSONS, P. J. BARR, K. STUART, M. SELKIRK e N. AGABIAN. Sequences homologous to the variant antigen mRNA spliced leader are located in tandem repeats and variable orphans in *Trypanosoma brucei*. *Cell*, v.34, n.3, Oct, p.901-9, (1983).

NICHOLS, B. A. e M. L. CHIAPPINO. Cytoskeleton of *Toxoplasma gondii*. *J Protozool*, v.34, n.2, May, p.217-26, (1987).

NICOLLE, C. AND L. MANCEAUX. Sur une infection à corps de Leishman (ou organismes voisins) du gondi." *C R. Seances Acad. Sci.* v.147, p. 763-766, (1908).

NOZAKI, T. e G. A. CROSS. Effects of 3' untranslated and intergenic regions on gene expression in *Trypanosoma cruzi*. *Mol Biochem Parasitol*, v.75, n.1, Dec, p.55-67, (1995).

OHNO, M., A. SEGREF, A. BACHI, M. WILM e I. W. MATTAJ. PHAX, a mediator of U snRNA nuclear export whose activity is regulated by phosphorylation. *Cell*, v.101, n.2, Apr 14, p.187-98, (2000).

OLSEN, J. V., L. M. DE GODOY, G. LI, B. MACEK, P. MORTENSEN, R. PESCH, A. MAKAROV, O. LANGE, S. HORNING e M. MANN. Parts per million mass accuracy on an Orbitrap mass spectrometer via lock mass injection into a C-trap. *Mol Cell Proteomics*, v.4, n.12, Dec, p.2010-21, (2005).

PALAZZO, A. F., M. SPRINGER, Y. SHIBATA, C. S. LEE, A. P. DIAS e T. A. RAPOPORT. The signal sequence coding region promotes nuclear export of mRNA. *PLoS Biol*, v.5, n.12, Dec, p.e322, (2007).

PATIL, V., P. J. LESCAULT, D. LIRUSSI, A. B. THOMPSON e M. MATRAJT. Disruption of the Expression of a Non-Coding RNA Significantly Impairs Cellular Differentiation in *Toxoplasma gondii*. *Int J Mol Sci*, v.14, n.1, p.611-24, (2012).

PICCHI, G. F., A. M. FERREIRA, F. S. SOUZA, E. E. LOURENCO, P. R. ARAUCO, A. LORUSSO, J. BORDIGNON, M. A. KRIEGER, S. GOLDENBERG e S. P. FRAGOSO. *Trypanosoma cruzi*: identification of DNA targets of the nuclear periphery coiled-coil protein TcNUP-1. *Exp Parasitol*, v.127, n.1, Jan, p.147-52, (2011).

PIEPERHOFF, M. S., M. SCHMITT, D. J. FERGUSON e M. MEISSNER. The role of clathrin in post-Golgi trafficking in *Toxoplasma gondii*. *PLoS One*, v.8, n.10, p.e77620, (2013).

POLLARD, V. W., W. M. MICHAEL, S. NAKIELNY, M. C. SIOMI, F. WANG e G. DREYFUSS. A novel receptor-mediated nuclear protein import pathway. *Cell*, v.86, n.6, Sep 20, p.985-94, (1996).

POLLOCK, C. e S. HUANG. The perinucleolar compartment. *Cold Spring Harb Perspect Biol*, v.2, n.2, Feb, p.a000679, (2010).

PREUSSER, C., Z. PALFI e A. BINDEREIF. Special Sm core complex functions in assembly of the U2 small nuclear ribonucleoprotein of *Trypanosoma brucei*. *Eukaryot Cell*, v.8, n.8, Aug, p.1228-34, (2009).

PRICE, M. N., P. S. DEHAL e A. P. ARKIN. FastTree 2--approximately maximum-likelihood trees for large alignments. *PLoS One*, v.5, n.3, p.e9490, (2010).

PROHASKA, K. e N. WILLIAMS. Assembly of the *Trypanosoma brucei* 60S ribosomal subunit nuclear export complex requires trypanosome-specific proteins P34 and P37. *Eukaryot Cell*, v.8, n.1, Jan, p.77-87, (2009).

RADKE, J. R., M. S. BEHNKE, A. J. MACKEY, J. B. RADKE, D. S. ROOS e M. W. WHITE. The transcriptome of *Toxoplasma gondii*. *BMC Biol*, v.3, p.26, (2005).

RASSI, A.; RASSI. JÚNIOR; A.; RASSI, G. G. Fase aguda. Em: *Trypanosoma cruzi* e a Doença de Chagas. Zigman Brener; Zilton Andrade; Manoel Barral-Neto 2. ed. Rio de Janeiro: Editora Guanabara Koogan, (2000).

REED, R. e H. CHENG. TREX, SR proteins and export of mRNA. *Curr Opin Cell Biol*, v.17, n.3, Jun, p.269-73, (2005).

RICE, P., I. LONGDEN e A. BLEASBY. EMBOSS: the European Molecular Biology Open Software Suite. *Trends Genet*, v.16, n.6, Jun, p.276-7, (2000).

ROBERT-GANGNEUX, F. e M. L. DARDE. Epidemiology of and diagnostic strategies for toxoplasmosis. *Clin Microbiol Rev*, v.25, n.2, Apr, p.264-96, (2012).

ROBERTS, T. G., N. R. STURM, B. K. YEE, M. C. YU, T. HARTSHORNE, N. AGABIAN e D. A. CAMPBELL. Three small nucleolar RNAs identified from the spliced leader-associated RNA locus in kinetoplastid protozoans. *Mol Cell Biol*, v.18, n.8, Aug, p.4409-17, (1998).

ROCAK, S. e P. LINDER. DEAD-box proteins: the driving forces behind RNA metabolism. *Nat Rev Mol Cell Biol*, v.5, n.3, Mar, p.232-41, (2004).

RODRIGUEZ-NAVARRO, S., T. FISCHER, M. J. LUO, O. ANTUNEZ, S. BRETTSCHEIDER, J. LECHNER, J. E. PEREZ-ORTIN, R. REED e E. HURT. Sus1, a functional component of the SAGA histone acetylase complex and the nuclear pore-associated mRNA export machinery. *Cell*, v.116, n.1, Jan 9, p.75-86, (2004).

RODRIGUEZ, A. J., J. CONDEELIS, R. H. SINGER e J. B. DICTENBERG. Imaging mRNA movement from transcription sites to translation sites. *Semin Cell Dev Biol*, v.18, n.2, Apr, p.202-8, (2007).

ROGER, A. J. e L. A. HUG. The origin and diversification of eukaryotes: problems with molecular phylogenetics and molecular clock estimation. *Philos Trans R Soc Lond B Biol Sci*, v.361, n.1470, Jun 29, p.1039-54, (2006).

RONDINELLI, E., R. SILVA, J. F. CARVALHO, C. M. DE ALMEIDA SOARES, E. F. DE CARVALHO e F. T. DE CASTRO. *Trypanosoma cruzi*: an in vitro cycle of cell differentiation in axenic culture. *Exp Parasitol*, v.66, n.2, Aug, p.197-204, (1988).

ROOS, D. S., R. G. DONALD, N. S. MORRISSETTE e A. L. MOULTON. Molecular tools for genetic dissection of the protozoan parasite *Toxoplasma gondii*. *Methods Cell Biol*, v.45, p.27-63, (1994).

ROUT, M. P. e M. C. FIELD. Isolation and characterization of subnuclear compartments from *Trypanosoma brucei*. Identification of a major repetitive nuclear lamina component. *J Biol Chem*, v.276, n.41, Oct 12, p.38261-71, (2001).

SAGUEZ, C. e T. H. JENSEN. Assembly of export-competent mRNP: it's all about being connected. *Mol Cell*, v.33, n.2, Jan 30, p.139-40, (2009).

SAHNI, A., N. WANG e J. D. ALEXIS. UAP56 is an important regulator of protein synthesis and growth in cardiomyocytes. *Biochem Biophys Res Commun*, v.393, n.1, Feb 26, p.106-10, (2010).

SAKASHITA, E. e H. ENDO. SR and SR-related proteins redistribute to segregated fibrillar components of nucleoli in a response to DNA damage. *Nucleus*, v.1, n.4, Jul-Aug, p.367-80, (2010).

SAKHARKAR, M. K. e P. KANGUEANE. Genome SEGE: a database for 'intronless' genes in eukaryotic genomes. *BMC Bioinformatics*, v.5, Jun 2, p.67, (2004).

SAMBROOK, J. e RUSSEL, D. T. *Molecular cloning: a laboratory manual*. Cold Spring Harbor Laboratory. 2ndEd., (2001).

SANTOS-ROSA, H., H. MORENO, G. SIMOS, A. SEGREF, B. FAHRENKROG, N. PANTE e E. HURT. Nuclear mRNA export requires complex formation between Mex67p and Mtr2p at the nuclear pores. *Mol Cell Biol*, v.18, n.11, Nov, p.6826-38, (1998).

SCHUTZ, S., J. CHEMNITZ, C. SPILLNER, M. FROHME, J. HAUBER e R. H. KEHLENBACH. Stimulated expression of mRNAs in activated T cells depends on a functional CRM1 nuclear export pathway. *J Mol Biol*, v.358, n.4, May 12, p.997-1009, (2006).

SCHWEDE, A., L. ELLIS, J. LUTHER, M. CARRINGTON, G. STOECKLIN e C. CLAYTON. A role for Caf1 in mRNA deadenylation and decay in trypanosomes and human cells. *Nucleic Acids Res*, v.36, n.10, Jun, p.3374-88, (2008).

SCHWEDE, A., T. MANFUL, B. A. JHA, C. HELBIG, N. BERCOVICH, M. STEWART e C. CLAYTON. The role of deadenylation in the degradation of unstable mRNAs in trypanosomes. *Nucleic Acids Res*, v.37, n.16, Sep, p.5511-28, (2009).

SEGREF, A., K. SHARMA, V. DOYE, A. HELLWIG, J. HUBER, R. LUHRMANN e E. HURT. Mex67p, a novel factor for nuclear mRNA export, binds to both poly(A)+ RNA and nuclear pores. *EMBO J*, v.16, n.11, Jun 2, p.3256-71, (1997).

SERPELONI, M., C. B. MORAES, J. R. MUNIZ, M. C. MOTTA, A. S. RAMOS, R. L. KESSLER, A. H. INOUE, W. D. DAROCHA, S. F. YAMADA-OGATTA, S. P. FRAGOSO, S. GOLDENBERG, L. H. FREITAS-JUNIOR e A. R. AVILA. An essential nuclear protein in trypanosomes is a component of mRNA transcription/export pathway. *PLoS One*, v.6, n.6, p.e20730, (2011).

SERPELONI, M., N. M. VIDAL, S. GOLDENBERG, A. R. AVILA e F. G. HOFFMANN. Comparative genomics of proteins involved in RNA nucleocytoplasmic export. *BMC Evol Biol*, v.11, p.7, (2011).

SHANKAR, J., A. PRADHAN e R. TUTEJA. Isolation and characterization of *Plasmodium falciparum* UAP56 homolog: evidence for the coupling of RNA binding and splicing activity by site-directed mutations. *Arch Biochem Biophys*, v.478, n.2, Oct 15, p.143-53, (2008).

SHANNON, P., A. MARKIEL, O. OZIER, N. S. BALIGA, J. T. WANG, D. RAMAGE, N. AMIN, B. SCHWIKOWSKI e T. IDEKER. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res*, v.13, n.11, Nov, p.2498-504, (2003).

SHAPIRO, T. A. e P. T. ENGLUND. The structure and replication of kinetoplast DNA. *Annu Rev Microbiol*, v.49, p.117-43, (1995).

SHAV-TAL, Y., J. BLECHMAN, X. DARZACQ, C. MONTAGNA, B. T. DYE, J. G. PATTON, R. H. SINGER e D. ZIPORI. Dynamic sorting of nuclear components into distinct nucleolar caps during transcriptional inhibition. *Mol Biol Cell*, v.16, n.5, May, p.2395-413, (2005).

SHAV-TAL, Y., X. DARZACQ, S. M. SHENOY, D. FUSCO, S. M. JANICKI, D. L. SPECTOR e R. H. SINGER. Dynamics of single mRNPs in nuclei of living cells. *Science*, v.304, n.5678, Jun 18, p.1797-800, (2004).

SHIBATA, S., M. SASAKI, T. MIKI, A. SHIMAMOTO, Y. FURUICHI, J. KATAHIRA e Y. YONEDA. Exportin-5 orthologues are functionally divergent among species. *Nucleic Acids Res*, v.34, n.17, p.4711-21, (2006).

SIEGEL, T. N., K. S. TAN e G. A. CROSS. Systematic study of sequence motifs for RNA trans splicing in *Trypanosoma brucei*. *Mol Cell Biol*, v.25, n.21, Nov, p.9586-94, (2005).

SILVEIRA, C., R. BELFORT, JR., C. MUCCIOLI, M. T. ABREU, M. C. MARTINS, C. VICTORA, R. B. NUSSENBLATT e G. N. HOLLAND. A follow-up study of *Toxoplasma gondii* infection in southern Brazil. *Am J Ophthalmol*, v.131, n.3, Mar, p.351-4, (2001).

SMETANA, K., AND BUSCH, H. The nucleolus and nucleolar DNA. *The Cell Nucleus*, v. 1, ed. H. Busch, New York: Academic Press, p. 73–147, (1974).

SOLDATI, D. e J. C. BOOTHROYD. Transient transfection and expression in the obligate intracellular parasite *Toxoplasma gondii*. *Science*, v.260, n.5106, Apr 16, p.349-52, (1993).

SOLDATI, D. e J. C. BOOTHROYD. A selector of transcription initiation in the protozoan parasite *Toxoplasma gondii*. *Mol Cell Biol*, v.15, n.1, Jan, p.87-93, (1995).

SPEESE, S. D., J. ASHLEY, V. JOKHI, J. NUNNARI, R. BARRIA, Y. LI, B. ATAMAN, A. KOON, Y. T. CHANG, Q. LI, M. J. MOORE e V. BUDNIK. Nuclear envelope budding enables large ribonucleoprotein particle export during synaptic Wnt signaling. *Cell*, v.149, n.4, May 11, p.832-46, (2012).

STARNES, G. L., T. J. JEWETT, V. B. CARRUTHERS e L. D. SIBLEY. Two separate, conserved acidic amino acid domains within the *Toxoplasma gondii* MIC2 cytoplasmic tail are required for parasite survival. *J Biol Chem*, v.281, n.41, Oct 13, p.30745-54, (2006).

STECKELBERG, A. L. e N. H. GEHRING. Studying the composition of mRNPs in vitro using splicing-competent cell extracts. *Methods*, v.65, n.3, Feb, p.342-9, (2014).

STEINDEL, M., L. KRAMER PACHECO, D. SCHOLL, M. SOARES, M. H. DE MORAES, I. EGER, C. KOSMANN, T. C. SINCERO, P. H. STOCO, S. M. MURTA, C. J. DE CARVALHO-PINTO e E. C. GRISARD. Characterization of *Trypanosoma cruzi* isolated from humans, vectors, and animal reservoirs following an outbreak of acute human Chagas disease in Santa Catarina State, Brazil. *Diagn Microbiol Infect Dis*, v.60, n.1, Jan, p.25-32, (2008).

STEWART, M. Insights into the molecular mechanism of nuclear trafficking using nuclear transport factor 2 (NTF2). *Cell Struct Funct*, v.25, n.4, Aug, p.217-25, (2000).

STEWART, M., H. M. KENT e A. J. MCCOY. Structural basis for molecular recognition between nuclear transport factor 2 (NTF2) and the GDP-bound form of the Ras-family GTPase Ran. *J Mol Biol*, v.277, n.3, Apr 3, p.635-46, (1998).

STOCKLEY, P. G., N. J. STONEHOUSE, J. B. MURRAY, S. T. GOODMAN, S. J. TALBOT, C. J. ADAMS, L. LILJAS e K. VALEGARD. Probing sequence-specific RNA recognition by the bacteriophage MS2 coat protein. *Nucleic Acids Res*, v.23, n.13, Jul 11, p.2512-8, (1995).

STOFFLER, D., B. FAHRENKROG e U. AEBI. The nuclear pore complex: from molecular architecture to functional dynamics. *Curr Opin Cell Biol*, v.11, n.3, Jun, p.391-401, (1999).

STRASSER, K. e E. HURT. Splicing factor Sub2p is required for nuclear mRNA export through its interaction with Yra1p. *Nature*, v.413, n.6856, Oct 11, p.648-52, (2001).

STRASSER, K., S. MASUDA, P. MASON, J. PFANNSTIEL, M. OPPIZZI, S. RODRIGUEZ-NAVARRO, A. G. RONDON, A. AGUILERA, K. STRUHL, R. REED e E. HURT. TREX is a conserved complex coupling transcription with messenger RNA export. *Nature*, v.417, n.6886, May 16, p.304-8, (2002).

STRIEPEN, B. Switching parasite proteins on and off. *Nat Methods*, v.4, n.12, Dec, p.999-1000, (2007).

STROM, A. C. e K. WEIS. Importin-beta-like nuclear transport receptors. *Genome Biol*, v.2, n.6, p.REVIEWS3008, (2001).

SU, C., D. EVANS, R. H. COLE, J. C. KISSINGER, J. W. AJIOKA e L. D. SIBLEY. Recent expansion of *Toxoplasma* through enhanced oral transmission. *Science*, v.299, n.5605, Jan 17, p.414-6, (2003).

SULLIVAN JR, W. J., A. T. SMITH e B. R. JOYCE. Understanding mechanisms and the role of differentiation in pathogenesis of *Toxoplasma gondii*: a review. *Mem Inst Oswaldo Cruz*, v.104, n.2, Mar, p.155-61, (2009).

SUTER, B.; KITTANAKOM, S.; STAGLJAR, I. Interactive proteomics: what lies ahead? *Biotechniques*, v. 44, n. 5, p. 681-91, 2008.

SUTHERLAND, B. W., J. TOEWS e J. KAST. Utility of formaldehyde cross-linking and mass spectrometry in the study of protein-protein interactions. *J Mass Spectrom*, v.43, n.6, Jun, p.699-715, (2008).

SUVOROVA, E. S., M. CROKEN, S. KRATZER, L. M. TING, M. C. DE FELIPE, B. BALU, M. L. MARKILLIE, L. M. WEISS, K. KIM e M. W. WHITE. Discovery of a splicing regulator required for cell cycle progression. *PLoS Genet*, v.9, n.2, p.e1003305, (2013).

TAKAHASHI, N., Y. TAKAHASHI e F. W. PUTNAM. Periodicity of leucine and tandem repetition of a 24-amino acid segment in the primary structure of leucine-rich alpha 2-glycoprotein of human serum. *Proc Natl Acad Sci U S A*, v.82, n.7, Apr, p.1906-10, (1985).

TALCOTT, B. e M. S. MOORE. Getting across the nuclear pore complex. *Trends Cell Biol*, v.9, n.8, Aug, p.312-8, (1999).

TANOWITZ, H. B., L. V. KIRCHHOFF, D. SIMON, S. A. MORRIS, L. M. WEISS e M. WITTNER. Chagas' disease. *Clin Microbiol Rev*, v.5, n.4, Oct, p.400-19, (1992).

TEIXEIRA, A. R., R. J. NASCIMENTO e N. R. STURM. Evolution and pathology in chagas disease-a review. Mem Inst Oswaldo Cruz, v.101, n.5, Aug, p.463-91, (2006).

TEIXEIRA, S. M. Control of gene expression in Trypanosomatidae. Braz J Med Biol Res, v.31, n.12, Dec, p.1503-16, (1998).

TEIXEIRA, S. M. e W. D. DAROCHA. Control of gene expression and genetic manipulation in the Trypanosomatidae. Genet Mol Res, v.2, n.1, p.148-58, (2003).

TELLERIA, J., B. LAFAY, M. VIRREIRA, C. BARNABE, M. TIBAYRENC e M. SVOBODA. *Trypanosoma cruzi*: sequence analysis of the variable region of kinetoplast minicircles. Exp Parasitol, v.114, n.4, Dec, p.279-88, (2006).

TENG, I. F. e S. A. WILSON. Mapping interactions between mRNA export factors in living cells. PLoS One, v.8, n.6, p.e67676, (2013).

TENTER, A. M., A. R. HECKEROTH e L. M. WEISS. *Toxoplasma gondii*: from animals to humans. Int J Parasitol, v.30, n.12-13, Nov, p.1217-58, (2000).

THAKURTA, A. G., S. P. SELVANATHAN, A. D. PATTERSON, G. GOPAL e R. DHAR. The nuclear export signal of splicing factor Uap56p interacts with nuclear pore-associated protein Rae1p for mRNA export in *Schizosaccharomyces pombe*. J Biol Chem, v.282, n.24, Jun 15, p.17507-16, (2007).

THAKURTA, A. G., W. A. WHALEN, J. H. YOON, A. BHARATHI, L. KOZAK, C. WHITEFORD, D. C. LOVE, J. A. HANOVER e R. DHAR. Crp79p, like Mex67p, is an auxiliary mRNA export factor in *Schizosaccharomyces pombe*. Mol Biol Cell, v.13, n.8, Aug, p.2571-84, (2002).

TIMMERS, H. T. e L. TORA. SAGA unveiled. Trends Biochem Sci, v.30, n.1, Jan, p.7-10, (2005).

TKACZ, I. D., S. COHEN, M. SALMON-DIVON e S. MICHAELI. Identification of the heptameric Lsm complex that binds U6 snRNA in *Trypanosoma brucei*. Mol Biochem Parasitol, v.160, n.1, Jul, p.22-31, (2008).

TKACZ, I. D., Y. LUSTIG, M. Z. STERN, M. BITON, M. SALMON-DIVON, A. DAS, V. BELLOFATTO e S. MICHAELI. Identification of novel snRNA-specific Sm proteins that bind selectively to U2 and U4 snRNAs in *Trypanosoma brucei*. RNA, v.13, n.1, Jan, p.30-43, (2007).

TOWBIN, H., T. STAEBELIN e J. GORDON. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. 1979. Biotechnology, v.24, p.145-9, (1992).

TUTEJA, R. e J. MEHTA. A genomic glance at the components of the mRNA export machinery in *Plasmodium falciparum*. Commun Integr Biol, v.3, n.4, Jul, p.318-26, (2010).

TYAGI, S. Imaging intracellular RNA distribution and dynamics in living cells. *Nat Methods*, v.6, n.5, May, p.331-8, (2009).

VALLOCHI, A. L., C. MUCCIOLI, M. C. MARTINS, C. SILVEIRA, R. BELFORT, JR. e L. V. RIZZO. The genotype of *Toxoplasma gondii* strains causing ocular toxoplasmosis in humans in Brazil. *Am J Ophthalmol*, v.139, n.2, Feb, p.350-1, (2005).

VAN DE VOSSE, D. W., Y. WAN, R. W. WOZNIAK e J. D. AITCHISON. Role of the nuclear envelope in genome organization and gene expression. *Wiley Interdiscip Rev Syst Biol Med*, v.3, n.2, Mar-Apr, p.147-66, (2011).

VAN DEN HOFF, M. J., A. F. MOORMAN e W. H. LAMERS. Electroporation in 'intracellular' buffer increases cell survival. *Nucleic Acids Res*, v.20, n.11, Jun 11, p.2902, (1992).

VAN DER PLOEG, L. H., A. Y. LIU, P. A. MICHELS, T. DE LANGE, P. BORST, H. K. MAJUMDER, H. WEBER, G. H. VEENEMAN e J. VAN BOOM. RNA splicing is required to make the messenger RNA for a variant surface antigen in trypanosomes. *Nucleic Acids Res*, v.10, n.12, Jun 25, p.3591-604, (1982).

VAN DUYN, G. D. A structural view of cre-*LoxP* site-specific recombination. *Annu Rev Biophys Biomol Struct*, v.30, p.87-104, (2001).

VANHAMME, L. e E. PAYS. Control of gene expression in trypanosomes. *Microbiol Rev*, v.59, n.2, Jun, p.223-40, (1995).

VAZQUEZ, M., C. ATORRASAGASTI, N. BERCOVICH, R. VOLCOVICH e M. J. LEVIN. Unique features of the *Trypanosoma cruzi* U2AF35 splicing factor. *Mol Biochem Parasitol*, v.128, n.1, Apr 25, p.77-81, (2003).

VAZQUEZ, M. P., D. MUALEM, N. BERCOVICH, M. Z. STERN, B. NYAMBEGA, O. BARDA, D. NASIGA, S. K. GUPTA, S. MICHAELI e M. J. LEVIN. Functional characterization and protein-protein interactions of trypanosome splicing factors U2AF35, U2AF65 and SF1. *Mol Biochem Parasitol*, v.164, n.2, Apr, p.137-46, (2009).

VENEMA, J. e D. TOLLERVEY. Processing of pre-ribosomal RNA in *Saccharomyces cerevisiae*. *Yeast*, v.11, n.16, Dec, p.1629-50, (1995).

VIDAL, M.; BRACHMANN, R. K.; FATTAEY, A.; HARLOW, E.; BOEKE, J. D. Reverse two-hybrid and one-hybrid systems to detect dissociation of protein-protein and DNA-protein interactions. *Proc Natl Acad Sci U S A*, v. 93, n. 19, p. 10315-20, 1996.

VILLANOVA, G. V., S. C. NARDELLI, P. CRIBB, A. MAGDALENO, A. M. SILBER, M. C. MOTTA, S. SCHENKMAN e E. SERRA. *Trypanosoma cruzi* bromodomain factor 2 (BDF2) binds to acetylated histones and is accumulated after UV irradiation. *Int J Parasitol*, v.39, n.6, May, p.665-73, (2009).

VONLAUFEN, N., S. M. KANZOK, R. C. WEK e W. J. SULLIVAN, JR. Stress response pathways in protozoan parasites. *Cell Microbiol*, v.10, n.12, Dec, p.2387-99, (2008).

WAHL, M. C., C. L. WILL e R. LUHRMANN. The spliceosome: design principles of a dynamic RNP machine. *Cell*, v.136, n.4, Feb 20, p.701-18, (2009).

WANG, P., Z. PALFI, C. PREUSSER, S. LUCKE, W. S. LANE, C. KAMBACH e A. BINDEREIF. Sm core variation in spliceosomal small nuclear ribonucleoproteins from *Trypanosoma brucei*. *EMBO J*, v.25, n.19, Oct 4, p.4513-23, (2006).

WARNER, J. R. The economics of ribosome biosynthesis in yeast. *Trends Biochem Sci*, v.24, n.11, Nov, p.437-40, (1999).

WEIGHARDT, F., G. BIAMONTI e S. RIVA. The roles of heterogeneous nuclear ribonucleoproteins (hnRNP) in RNA metabolism. *Bioessays*, v.18, n.9, Sep, p.747-56, (1996).

WICKRAMASINGHE, V. O., P. I. MCMURTRIE, A. D. MILLS, Y. TAKEI, S. PENRHYN-LOWE, Y. AMAGASE, S. MAIN, J. MARR, M. STEWART e R. A. LASKEY. mRNA export from mammalian cell nuclei is dependent on GANP. *Curr Biol*, v.20, n.1, Jan 12, p.25-31, (2010).

WIEDERKEHR, T., R. F. PRETOT e L. MINVIELLE-SEBASTIA. Synthetic lethal interactions with conditional poly(A) polymerase alleles identify LCP5, a gene involved in 18S rRNA maturation. *RNA*, v.4, n.11, Nov, p.1357-72, (1998).

WILD, T., P. HORVATH, E. WYLER, B. WIDMANN, L. BADERTSCHER, I. ZEMP, K. KOZAK, G. CSUCS, E. LUND e U. KUTAY. A protein inventory of human ribosome biogenesis reveals an essential function of exportin 5 in 60S subunit export. *PLoS Biol*, v.8, n.10, p.e1000522, (2010).

WOLYNIAK, M. J. e C. N. COLE. Harnessing genomics to explore the processes and evolution of mRNA export. *RNA Biol*, v.5, n.2, Apr-Jun, p.68-72, (2008).

XU, T., VENABLE, J. D., PARK, S. K., COCIORVA, D., LU, B., LIAO, L., WOHLSCHLEGEL, J., HEWEL, J., YATES, J. R. I. ProLuCID, a Fast and Sensitive Tandem Mass Spectra-based Protein Identification Program. *Mol Cell Proteomics*, 5, S174, (2006).

YAMAZAKI, T., N. FUJIWARA, H. YUKINAGA, M. EBISUYA, T. SHIKI, T. KURIHARA, N. KIOKA, T. KAMBE, M. NAGAO, E. NISHIDA e S. MASUDA. The closely related RNA helicases, UAP56 and URH49, preferentially form distinct mRNA export machineries and coordinately regulate mitotic progression. *Mol Biol Cell*, v.21, n.16, Aug 15, p.2953-65, (2010).

YANG, J., H. P. BOGERD, P. J. WANG, D. C. PAGE e B. R. CULLEN. Two closely related human nuclear export factors utilize entirely distinct export pathways. *Mol Cell*, v.8, n.2, Aug, p.397-406, (2001).

YAO, W., D. ROSER, A. KOHLER, B. BRADATSCH, J. BASSLER e E. HURT. Nuclear export of ribosomal 60S subunits by the general mRNA export receptor Mex67-Mtr2. *Mol Cell*, v.26, n.1, Apr 13, p.51-62, (2007).

YONEDA, Y. Nucleocytoplasmic protein traffic and its significance to cell function. *Genes Cells*, v.5, n.10, Oct, p.777-87, (2000).

YONEDA, Y., M. HIEDA, E. NAGOSHI e Y. MIYAMOTO. Nucleocytoplasmic protein transport and recycling of Ran. *Cell Struct Funct*, v.24, n.6, Dec, p.425-33, (1999).

YOON, J. H., W. A. WHALEN, A. BHARATHI, R. SHEN e R. DHAR. Npp106p, a *Schizosaccharomyces pombe* nucleoporin similar to *Saccharomyces cerevisiae* Nic96p, functionally interacts with Rae1p in mRNA export. *Mol Cell Biol*, v.17, n.12, Dec, p.7047-60, (1997).

ZEMP, I., T. WILD, M. F. O'DONOHUE, F. WANDREY, B. WIDMANN, P. E. GLEIZES e U. KUTAY. Distinct cytoplasmic maturation steps of 40S ribosomal subunit precursors require hRio2. *J Cell Biol*, v.185, n.7, Jun 29, p.1167-80, (2009).

ZHANG, B., M. C. CHAMBERS e D. L. TABB. Proteomic parsimony through bipartite graph analysis improves accuracy and transparency. *J Proteome Res*, v.6, n.9, Sep, p.3549-57, (2007).

ZHANG, M., B. R. JOYCE, W. J. SULLIVAN, JR. e V. NUSSENZWEIG. Translational control in Plasmodium and toxoplasma parasites. *Eukaryot Cell*, v.12, n.2, Feb, p.161-7, (2013).

9 Anexo I

***TgUAP56: an essential nuclear protein related with mRNA export
in T. gondii (artigo em preparação)***

TgUAP56: an essential nuclear protein related with mRNA export in *T. gondii*

Mariana Serpeloni^{1, 2}, Newton Vidal², Nicole Andenmatten³, Gurman S. Pall³, Markus Meissner³, Andrea R. Avila²

1. Universidade Federal do Paraná (UFPR), Curitiba, Brazil, 2. Instituto Carlos Chagas (ICC) Curitiba, Brasil; 3. University of Glasgow, Glasgow, Scotland

Abstract

T.gondii is the causative agent of Toxoplasmosis. The symptoms of disease are more aggressive for Brazilian patients in comparison to patients from EUA and Europe and no treatment against the chronic infection, the tissue cysts, is available. For this reason, understanding the essential factors and the biological specialized mechanisms for parasite survival could lead to new targets for treatment of parasitic disease. Besides, *T.gondii* has emerged as a major model for the study of apicomplexan biology. Cell development is a highly regulated process in eukaryotic organisms and the cell differentiation of *Toxoplasma* is no exception, involving gene expression pathways and regulatory mechanisms operating at the transcriptional and post-transcriptional level. The nucleocytoplasmic RNA export is an essential post-transcriptional pathway for gene expression control in eukaryotic cells, but it is still poorly understood in apicomplexan parasites. In Metazoan and Fungi, the general model of RNA export involves exportins as transport receptors that carry RNA through the nuclear pore complex (NPC) in a RanGTP-dependent manner. However, nuclear export of most mRNAs does not follow the RanGTP-exportin pathway. Excluding model eukaryotic organisms, the export machinery of several species of parasites remains unknown. Previous work from our group, using bioinformatic comparisons, suggest that the basic components of the RanGTP-dependent RNA pathways are conserved throughout the eukaryotic phylogeny, whereas proteins involved in RanGTP-independent mRNA export are less conserved. The exception is UAP56, a human DEAD-box helicase related to mRNA export, that is highly conserved throughout the eukaryotic phylogeny, including in organisms that early diverged during evolution, like Kinetoplastids and Apicomplexan parasites. *T.gondii* has emerged as a major model for the study of apicomplexan biology and we demonstrate that UAP56 has a functional ortholog in *T. gondii*, the protein TgUAP56, which is localized all over the nucleus apart from the nucleolus. Overexpression and conditional knockout assays showed that TgUAP56 is an essential component of mRNA export pathway in *T. gondii*. The depletion of TgUAP56 did not affect the pre-mRNA splicing but exhibited a strong bulk mRNA export defect, leading to conclusion that TgUAP56 is involved in mRNA export, probably for releasing of spliced mRNA to adaptor proteins. Moreover, the conserved function of UAP56 in parasites from different groups, indicates that it is also a basic component of mRNA export in protozoa species.

Introduction

Toxoplasma gondii is an opportunistic pathogen which is the causative agent of toxoplasmosis, a worldwide zoonosis. Primary infection is usually subclinical but cervical lymphadenopathy, ocular toxoplasmosis (OT), encephalitis or systemic infections can be present in the immunocompromised individuals, particularly the ones with HIV/AIDS. The infection acquired during pregnancy also may cause severe damage to the fetus [1, 2], although epidemiologic data show that many cases of OTs are acquired after birth rather than congenitally [1-3]. *T.gondii* is a versatile parasite and the prevalence of infection varies between countries, geographical areas and ethnic groups living within a specific region [4]. In Brazil, high levels of ocular disease are associated with toxoplasmosis and the highest values of serologic prevalence of *T. gondii* infection is in northern and southern states, including 50-80% of adult population [5, 6]. Furthermore, the symptoms of the disease are more aggressive for Brazilian patients in comparison to patients from EUA and Europe, probably due to pathogenicity of the infecting strain and/or the immune competence of the host [7, 8]. Studies have shown that type I strains can be related to OTs in Brazilian patients [9], and most of Brazilian strains has different genotypes from the clonal lineages that predominate in North America and Europe [10]. It could be a reason for more several OT in Brazilian infants with congenital toxoplasmosis compared to European infants [11].

Another important aspect is that *T. gondii* belongs to Apicomplexan phylum that includes important human and veterinary pathogens on global human and animal health as *Plasmodium* spp., the causative agent of malaria. Among the species of the Apicomplexa, *T. gondii* is the most experimentally tractable. Methodologies for classic and reverse genetics are well established with a broad range of tools and experimental advantages. Thus, *T. gondii* has emerged as a major model for the study of apicomplexan biology [12, 13].

Toxoplasma has a relatively complex life cycle, containing both asexual and sexual stages [14]. It is an intracellular parasite that uses essential and specialized pathways in order to survive inside the host cell and differentiate into distinct development forms. The acute infection can be treated using

pyrimethamine and sulphadiazine as therapeutic agents [1]. However, no treatment against the chronic infection, the tissue cysts, is available at present. Hence, understanding the essential factors and the biological mechanisms specialized for parasite survival could lead to new targets for treatment of toxoplasmosis.

Development is a highly regulated process in eukaryotic organisms and the cell differentiation of *Toxoplasma* is no exception, involving gene expression pathways and regulatory mechanisms operating at the transcriptional and post-transcriptional level [15]. At the transcriptional level, the mechanisms of gene regulation require transcriptional factors, epigenetics factors and chromatin modifications [16-18]. However, at post-transcriptional level, most of the data published involves translation control with a particular emphasis in cellular stress response encountered during the life cycle of the parasite [17, 19, 20]. Other post-transcriptional mechanisms may be involved in *Toxoplasma* development, such as differential mRNA splicing or stability, but these have been understudied. Recent pieces of data, have shown the role of splicing factors, RNA granules and non-coding RNAs in regulating the cell cycle, extracellular viability and cell differentiation of this parasite [21-23], reassuring the importance of post-transcriptional pathways for control of gene expression in *Toxoplasma*.

The nucleocytoplasmic RNA export is an essential post-transcriptional pathway for gene expression control in eukaryotic cells, but it is poorly understood in apicomplexan parasites. In Metazoan and Fungi, the general model of RNA export involves exportins as transport receptors that carry RNA through the nuclear pore complex (NPC) in a RanGTP-dependent manner [24]. The nuclear export of most RNA species such as microRNAs (miRNAs), rRNAs, small nuclear RNAs (snRNAs), and tRNAs, has been shown to follow the RanGTP-exportin model of transport, and specific exportins are involved with the different export pathways. However, nuclear export of most mRNAs does not follow the RanGTP-exportin pathway [25-27]. The mRNA export machinery is highly integrated with mRNA processing and includes a different set of nuclear transport adaptors, mRNA binding proteins, RNA helicases, and NPC associated proteins [26-28]. In a general manner, RNA-binding proteins of the

THO complex associate with mRNA and initiate the formation of a ribonucleoprotein complex that recruits processing and export factors, such as Sub2/UAP56 and yra1/Aly (yeast and metazoans, respectively), resulting in the formation of the Transcription/Export (TREX) complex. The TREX complex interacts with the spliceosome, and processed mRNAs are exported through the NPC by the receptor Mex67/TAP. The mRNA is released for translation in the cytoplasm by the ATP-dependent helicase Dbp5/DDX19 [26, 29].

Apart from eukaryotic model organisms, the export machinery of several species of parasites remains unknown. Previous work from our group, using bioinformatic comparisons, suggest that the basic components of the RanGTP-dependent RNA pathways are conserved throughout the eukaryotic phylogeny, whereas proteins involved in RanGTP-independent mRNA export are less conserved. The exception is UAP56, a human DEAD-box helicase related to mRNA export, that is highly conserved throughout the eukaryotic phylogeny, including in organisms that early diverged during evolution, like trypanosomatids and apicomplexan parasites [30]. We have shown that is an essential protein and a component of mRNA transcription/export pathway in trypanosomes [31]. The only study involving UAP56 ortholog in the phylum Apicomplexa, PfU52 protein of *Plasmodium falciparum*, showed the involvement of this protein in mRNA splicing in vitro, although the putative mRNA export function remains to be seen [32].

In the present work, we demonstrate that UAP56 has a functional ortholog in *T. gondii*, the protein TgUAP56, which is localized all over the nucleus apart from the nucleolus. The overexpression and conditional knockout assays showed that TgUAP56 is an essential component of mRNA export pathway in *T. gondii* causing a lethal nuclear mRNA accumulation phenotype.

Material and Methods

Sequence and phylogenetic analysis

Aminoacid sequences of UAP56 from Human and Sub2 from yeast were used as query sequences to identify orthologs in 43 representative species from major eukaryotic groups as listed on Supp. Table 1: Metazoans, Fungi, Amoebozoa, Plants, Parabasilids, Kinetoplastids and Apicomplexans. Proteome sequences from all species were downloaded from the National Center for Biotechnology Information (NCBI) Reference Sequence (RefSeq) database. Bioinformatic searches were performed using BLASTP algorithm [33] with default parameters. Orthologous sequences were identified by reciprocal best hit (RBH). Protein domain searches were performed running hmmscan program from HMMER package [34] against the collection of hidden markov models downloaded from the Pfam database version 27.0 [35]. Multiple sequence alignment of orthologous sequences were done using MAFFT version 7 with the following parameters: -localpair -maxiterate 1000 -reorder [36]. Phylogenetic analysis was conducted using the approximately maximum likelihood method implemented in the FastTree 2.1 program [37] under WAG model of amino acids evolution. The tree was rendered using FigTree v1.4 (<http://tree.bio.ed.ac.uk/software/figtree/>). Identity and similarity percentages were obtained using needle program from the EMBOSS package [38], which finds the optimal global alignment of two sequences.

Parasite cell lines and transfectants

T. gondii tachyzoites (RH Δ hxgprt – RH strain [39] and *ku80 clean up* DiCre strain - DiCre strain (Pall, unpublished), and transfectants obtained in this study) were grown in human foreskin fibroblasts (HFF) and maintained in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal calf serum, 2 mM glutamine and 25 μ g/ml gentamicin.

To generate stable lines, 5×10^7 freshly released parasites were transfected and selected in presence of mycophenolic acid and xanthine as previously described [39].

For obtaining the ddGFPTgUAP56 stable line, RH Δ *hxgprt* parasites were transfected with *p5RT70ddmycGFPuap56-HX* vector (Supp Figure 1A). The open reading frame (ORF) of *uap56* gene was amplified from cDNA with primers UAP56 F (forward) and UAP56 R (reverse) containing the *Pst*I and *Pac*I restriction sites respectively and the PCR fragment was cloned downstream the dd-GFP sequence of *p5RT70DDmycGFPHXGPRT* vector [35]. The vector was linearized with *Kpn*I before transfection.

For cKOuap56 stable line, parasites of *ku80 clean up* DiCre strain (Pall, unpublished) were transfected with the *GSuap56loxPmCherry-HX* vector (Supp Figure 1B) for a gene-swap strategy [40]. The *uap56* ORF was amplified from cDNA using primers UAP56 F and R containing *Apa*I and *Nsi*I restriction sites respectively and *loxP* sequence was added in the reverse primer. The PCR product was cloned upstream the *mcherry* sequence, replacing *tub8* of parental vector (unpublished vector). Further cloning steps were performed for insertion of *uap56* 5' and 3' UTRs. *Uap56* 5' UTR was amplified from genomic DNA using 5' UTR F and 5' UTR R containing *Kpn*I and *Apa*I restriction sites respectively, and *LoxP* sequence was added in the 5' UTR reverse primer. The PCR fragment was cloned upstream the *uap56* cDNA. The transcription of *uap56* gene is driven by the endogenous promoter present in 5' UTR. Finally the 3' UTR of *uap56* was amplified from genomic DNA using the primer pair 3' UTR F and R containing *Sac*I restriction site for cloning. The plasmid was linearized with *Kpn*I before transfection. All primers are listed on Supp. Table 2.

Immunoblot assays

Intracellular parasites or freshly lysed extracellular parasites were incubated in culture media in the absence or presence of 2 μ M Shd1 or 50 nM of rapamycin. Subsequently parasites were harvested and washed in PBS. SDS-PAGE and Western Blot analysis were performed as described previously [41] using 12% polyacrylamide gels under reducing condition with 100 mM DTT. Equal number of parasites was loaded per experiment. For protein detection, polyclonal anti-TcSub2 was used [31]. Monoclonal anti-aldolase [42] was used as internal control.

Fluorescent *in situ* hybridization (FISH) and Immunofluorescence assays

For polyA⁺ detection in *T. gondii* was performed FISH assays as described previously [21, 31]. Tachyzoite-infected HFF cells on glass coverslips were fixed with 4% formaldehyde in PBS for 20 minutes and permeabilized with 0.2% Triton X-100 in SSPE 2X for 20 minutes (SSPE 2X: 300 mM NaCl, 20 mM phosphate buffer pH 7.4, 2 mM EDTA). The parasites were washed 3 times with SSPE 2X and blocked with hybridization solution (HS) (10% Dextran, SSPE 2X, 35 % formamide, 0,5 mg/ml tRNA) for 30 minutes at 37 °C in wet chamber. Further, 1ng/μl of the probe (oligodT conjugated with digoxigenin) was added in HS and heated for 3 minutes at 65 °C. The cells were incubated with the probe during 16 hours at 37 °C in wet chamber. After the hybridization, the cells were washed with SSC 2X and SSC 1X for 15 minutes each one (SSC 20X: 3M NaCl, Na₃Citratex-2H₂O pH 7.4). The FISH was followed by immunofluorescence (IF) assay for probe detection.

For IF assays in general, the cells were blocked with PBS containing 3 % bovine serum albumin for 20 minutes before incubation with the primary antibody for one hour: anti-digoxigenin (Roche #cat. number 11333062910) for FISH assays; anti-TgUAP56 for immunolocalization [31]. Further, the cells were incubated with the secondary antibodies (Alexa 594-conjugated anti-rabbit Invitrogen #cat. number A11012) and Alexa 488 or 594-conjugated goat anti-mouse antibodies (Invitrogen #cat. numbers A-11001 and A11005, respectively). The slides were mounted in DAPI Fluoromount (Southern Biotech (UK) cat# 0100-20). Image acquisition was conducted using a 100X and 63X oil objective on a Zeiss Axioskop 2 (mot plus) fluorescence microscope with Axiocam MRm CCD camera using Volocity software. Further picture processing was performed using ImageJ 1.34r software and Photoshop CS6 (Adobe Systems Inc).

Growth assays

The plaque assay was performed as described before [43]. Monolayers of human foreskin fibroblasts (HFF) grown in 6 well plates were infected with 200 tachyzoites per well. Parasites were incubated with 50 nM of rapamycin and the drug was washed out after 24 hours of incubation. After 108 hours of incubation at normal growth conditions (37°C, 5% CO₂), cells were fixed 10 minutes with –

20°C methanol 100%, dyed with Giemsa stain for 10 minutes and washed once with PBS. Images were taken using a Zeiss microscope (Axiovert 200M) with a 10x objective and plaque size was determined using Axiovision software (Zeiss).

Analysis of mRNA processing by splicing

The analysis of mRNA splicing was performed by analytical PCR for seven selected genes as previously described [23]. Total RNA was purified with RNeasy (Qiagen) from DiCre strain and cKOtguap56 line after 24 and 48 hours of knockout induction by rapamycin. The RNA was reverse transcribed using random primers (Invitrogen) and the target sequences were amplified by PCR using specific primers that span an intron, listed on Supp. Table 2. Genomic DNA was used as reference to distinguish between properly spliced (S) and pre-spliced (PS). mRNAs from the following genes were analyzed: RNA polymerase II p8.2 subunit (TGME49_217560), RNA polymerase II p19 subunit (TGME49_271300), RNA polymerase II p23 subunit (TGME49_240590), IMC1 (TGME49_231640), IMC15 (TGME49_275670), IMC5 (TGME49_224530), and transcription factor IID (TGME49_258680). Tubulin was used as control [44].

Results

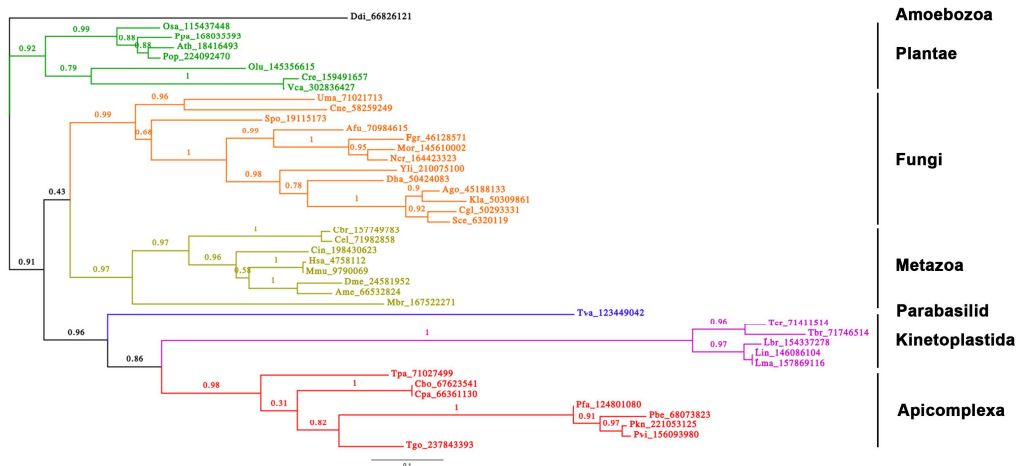
UAP56 is a highly conserved protein throughout the eukaryotic phylogeny

We have previously shown that orthologues of human UAP56 and yeast Sub2 proteins are present in most eukaryotic lineages [30] and thus it is likely to have been already present in the last eukaryotic common ancestor (LECA). Here, we surveyed 43 representative species of different eukaryotic groups in order to reconstruct the phylogeny of UAP56 orthologous sequences across eukaryotes. UAP56 protein phylogeny resembles to the eukaryotic species phylogeny (Figure 1A). The sequences were grouped into well supported clades, which correspond to eukaryotic groups. Support values are 0.92, 0.99, 0.97, 1.00, 0.98, for Plants, Fungi, Metazoa, Kinetoplastida and Apicomplexans, respectively. Apicomplexans and Kinetoplastida clustered as sister-clades with

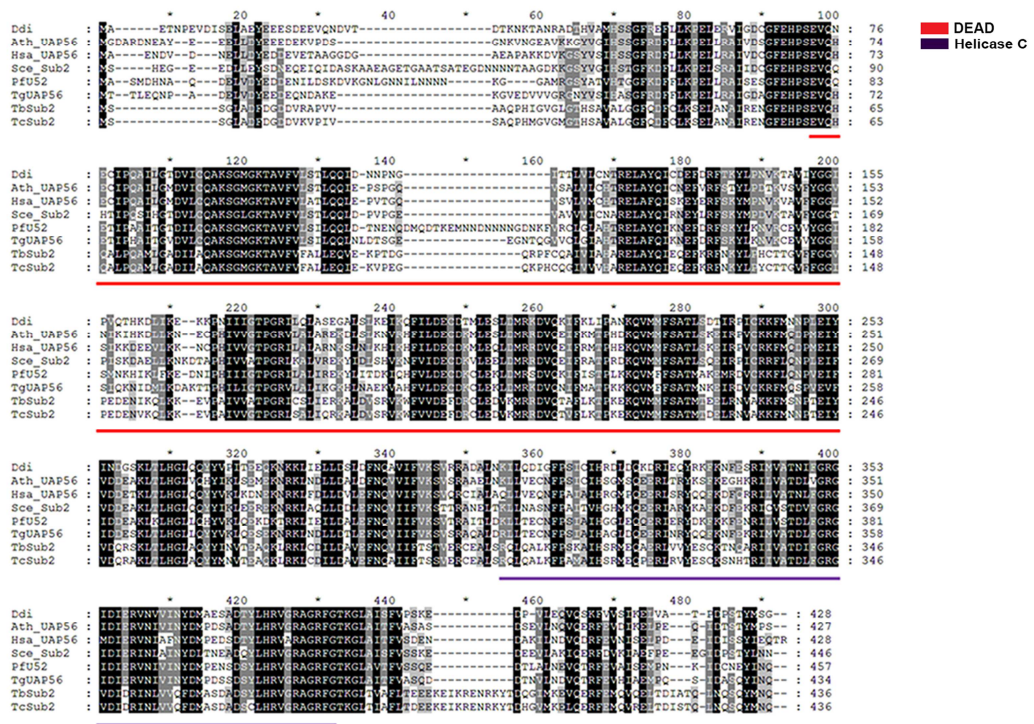
good support (0.86), showing that UAP56 sequences of different parasites from these two groups are closely related. The ortholog protein of UAP56 in *T. gondii*, named here TgUAP56, is annotated as a DEAD-box polypeptide DDX39 (ID TGME49_216860 at ToxoDB, GI 237843393 at NCBI) and is 78.6 and 72.8 % similar to the orthologs in human cells (Hsa_UAP56) and yeast (Sce_Sub2), respectively (Figure 1C). Multiple sequence alignment of TgUAP56 and representative sequences from other eukaryotic groups show high sequence conservation along the entire protein (Figure 1B). All sequences present two RNA helicases domains: a N-terminal DEAD box domain (PF00270) and a C-terminal Helicase_C (PF00271) domain, which is a structural feature of canonical RNA helicases [45] (Figure 1B).

Regarding well known orthologs, TgUAP56 identity/similarity ranges from 47.1/62.2% in *Trypanosoma cruzi*, TcSub2 [31], to 68.6/81.4% in *Plasmodium falciparum*, PfU52 [32], as showed in Figure 1C.

A



B



C

	Ddi	Ath_UAP56	Hsa_UAP56	Sce_Sub2	Pfu52	TgUAP56	TbSub2	TcSub2
Ddi	x	78.0	77.4	71.8	70.5	74.0	65.1	65.5
Ath_UAP56	63.4	x	83.3	76.5	72.8	77.6	64.1	64.9
Hsa_UAP56	58.1	70.1	x	77.1	71.0	78.6	66.4	66.4
Sce_Sub2	54.4	61.4	61.5	x	70.9	72.8	62.7	61.5
Pfu52	53.9	58.5	56.9	54.6	x	81.4	63.1	62.2
TgUAP56	57.0	65.1	64.6	56.1	68.6	x	66.1	66.1
TbSub2	44.5	48.1	48.5	43.7	44.2	47.3	x	96.1
TcSub2	45.0	49.4	49.0	44.2	44.7	47.1	89.4	x

Figure 1. UAP56 is a highly conserved RNA helicase. A. Phylogenetic reconstruction of UAP56 orthologous sequences across different groups of eukaryotes. Numbers at nodes are the bootstrap values. Amoebozoa (black): Ddi – *Dictyostelium discoideum*; Plants (green): Osa – *Oryza sativa*, Ppa – *Physcomitrella patens*, Ath – *Arabidopsis thaliana*, Pop – *Populus trichocarpa*, Olu – *Ostreococcus lucimarinus*, Cre – *Chlamydomonas reinhardtii*, Vca – *Volvox carteri*; Fungi (orange): Uma – *Ustilago maydis*, Cne – *Cryptococcus neoformans*, Spo –

Schizosaccharomyces pombe, Afu – *Aspergillus fumigatus*, Fgr – *Fusarium graminearum*, Mor – *Magnaporthe oryzae*, Ncr – *Neurospora crassa*, Yli – *Yarrowia lipolytica*, Dha – *Debaryomyces hansenii*, Ago – *Ashbya gossypii*, Kla – *Kluyveromyces lactis*, Cgl – *Candida glabrata*, Sce – *Saccharomyces cerevisiae*; Metazoans (yellow): Cbr – *Caenorhabditis briggsae*, Cel – *Caenorhabditis elegans*, Cin – *Ciona intestinalis*, Hsa – *Homo sapiens*, Mmu – *Mus musculus*, Dme – *Drosophila melanogaster*, Ame – *Apis mellifera*, Mbr – *Monosiga brevicolis*; Parabasilid (blue): Tva – *Trichomonas vaginalis*; Kinetoplastids (purple): Lbr – *Leishmania braziliensis*, Lin – *Leishmania infantum*, Lma – *Leishmania major*, Tbr – *Trypanosoma brucei*, Tcr – *Trypanosoma cruzi*; Apicomplexans (red): Tpa – *Theileria parva*, Cho – *Cryptosporidium hominis*, Cpa – *Cryptosporidium parvum*, Pfa – *Plasmodium falciparum*, Pbe – *Plasmodium berghei*, Pkn – *Plasmodium knowlesi*, Pvi – *Plasmodium vivax*, Tgo – *Toxoplasma gondii*. B. Multiple sequence alignment of UAP56 ortholog proteins. This analysis shows a high conservation along the entire length of representative sequences of eukaryotes, including apicomplexans and kinetoplastids. N-terminal region containing DEAD (PF00270) domain. C-terminal region containing Helicase_C (PF00271) domain. Both domains are evidenced. The Genbank Identifier (GI) number for the entries are listed as follows: (Ddi 66826121), (Ath_UAP56 18416493), (Hsa_UAP56 4758112), (Sce_Sub2 6320119), (PfU52 124801080), (TgUAP56 237843393), (TbSub2 71746514), (TcSub2 71411514). C. Percentage of similar (upper diagonal) and identical (lower diagonal) amino acids based on optimal global alignment between pairs of sequences showed in B.

TgUAP56 is an essential nuclear protein in *T. gondii*

The DD system allows rapid and reversible regulation of protein abundance in apicomplexan parasites [46, 47] and it has used for functional studies of proteins in *T. gondii* [48, 49]. Herein, we generated parasites expressing TgUAP56 fused with GFP and destabilizing domain (ddFKBP) in N-terminal portion (dd-GFP-TgUAP56). The ddFKBP domain results in constitutively degradation of the fused protein in normal situation but the presence of the synthetic cell-permeable ligand, Shield-1 (Shld1), allows the stabilization of fused protein in a reversible manner [50]. We noted that incubation of 1 μ M of Shld1 in the medium for more than 48 hours was lethal to the parasite, probably as consequence of a dominant negative phenotype. For this reason, we analyzed the overexpression effect in the growth of parasite by plaque assay. In this assay, parasites were incubated with distinct concentrations of Shld1 for 108 hours (Figure 2A). No plaques were observed when 1 or 2 μ M of Shld1 were added to the culture, suggesting that the overexpression of dd-GFP-TgUAP56 is lethal for *T. gondii*. Based on that, we decided to perform a time course experiment using 2 μ M of Shld1 to analyze the expression of fused protein by western blot. The results showed that dd-GFP-TgUAP56 is stabilized just after 6 hours of Shld1 incubation and reached the highest level after 36 hours, while the levels of the constitutive protein aldolase does not change during all the time course (Figure 2B).

Immunolocalization assays with cells treated with 2 μM of Shld1 for 6 hours showed that dd-GFP-TgUAP56 localize exclusively to the nucleus in intracellular tachyzoite forms (Figure 2C).

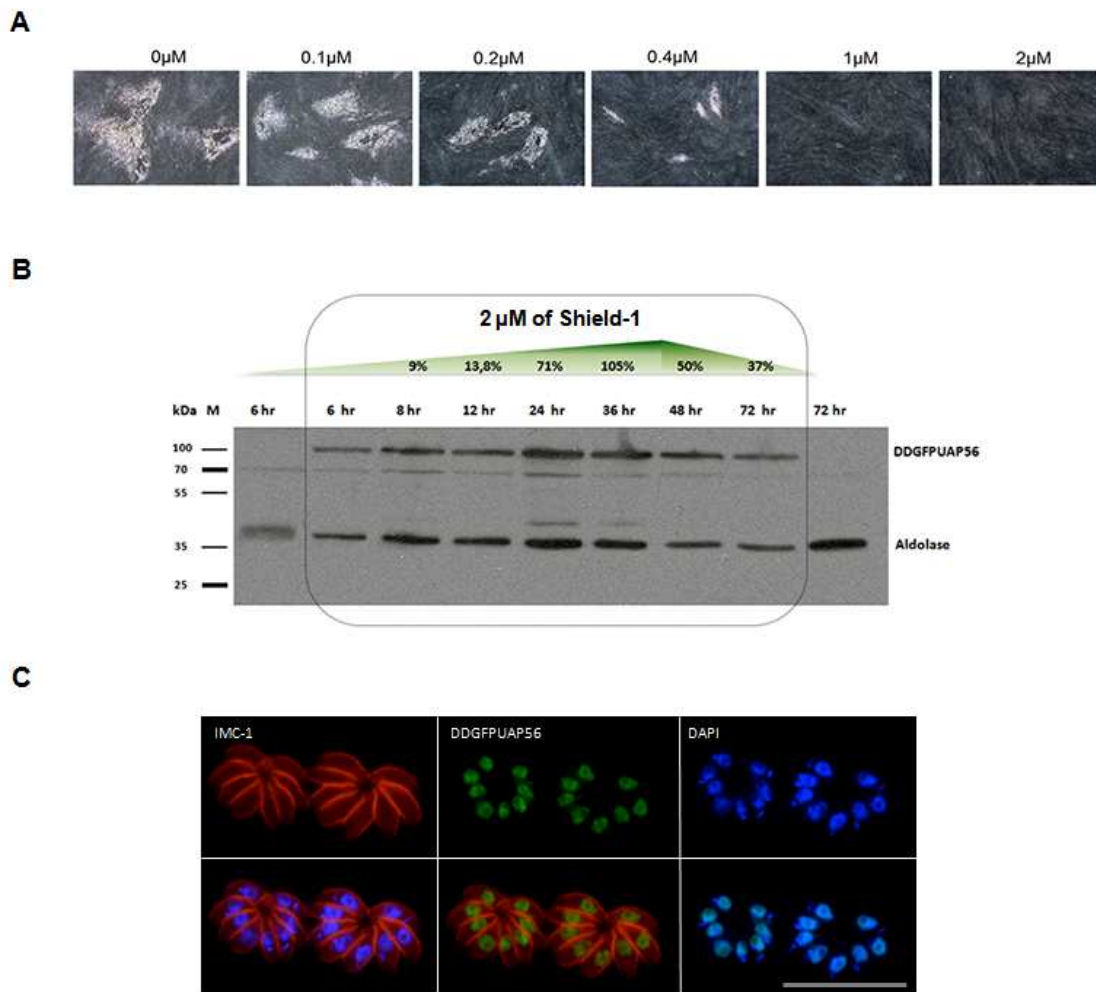
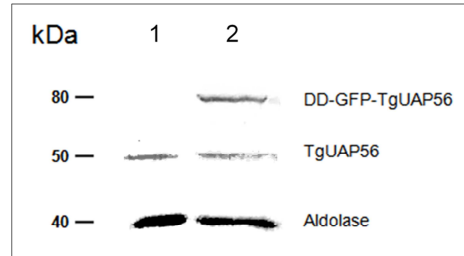


Figure 2. TgUAP56 is a nuclear essential protein in *T. gondii*. A. Plaque assay for ddGFPTgUAP56 lineage. Transfectants parasites were grown on human foreskin fibroblasts in the presence of different concentration of Shld1. After 108 hours of incubation, the cells were fixed and stained with Giemsa. B. Western blot analysis of dd-GFP-TgUAP56 expression using 2 μM of Shield-1 for different times. Monoclonal antibody against GFP (anti-GFP-Roche) was used for detection of dd-GFP-TgUAP56. Polyclonal antibody against aldolase was as control. Protein quantification was made using ImageJ densitometry software (Version 1.6, National Institutes of Health, Bethesda, MD). C. Subcellular localization of protein dd-GFP-TgUAP56. IMC = inner membrane complex, in red. DAPI = nuclear and apicoplast DNA staining. Bars = 10 μm .

The polyclonal antibody raised against the ortholog protein of trypanosome [31] recognizes dd-GFP-TgUAP56 and the endogenous protein TgUAP56, confirming the high structural conservation of this protein in parasites species from different phylogenetic groups (Trypanosoma and Apicomplexa). The co-

localization of both proteins was observed by immunofluorescence assays (Figure 3).

A



B

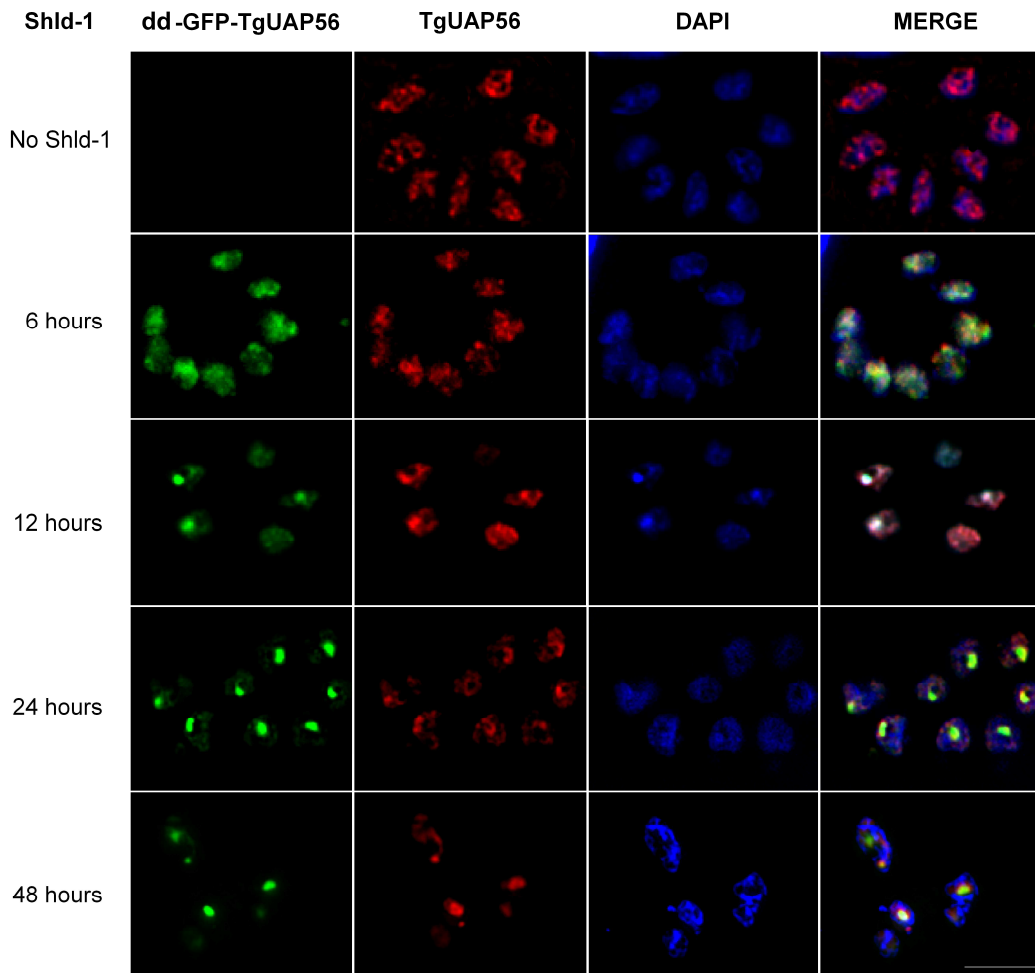


Figure 3 Expression and localization analysis of dd-GFP-TgUAP56. A. Analysis of TgUAP56 and dd-GFP-TgUAP56 expression by Western blot using a specific polyclonal antibody against TcSub2 [31]. (1) Only the endogenous protein is detected when the parasite is not incubated with Shld-1 (2) Shld-1 incubation after 6 hours induced the overexpression of fused protein (dd-GFP-TgUAP56). Aldolase: Loading control. B. dd-GFP-TgUAP56 colocalizes with endogenous TgUAP56 protein. TgUAP56 (in red) is detected in parasites not incubated

with Shld-1 while the dd-GFP-UAP56 is not detected. The detection of dd-GFP-UAP56 (in green) is observed only after incubation of parasites with 2 μ M of Shld-1. Nuclear and apicoplast DNA staining with DAPI: in blue. Bar = 5 μ m.

TgUPA56 is involved in mRNA export pathway of *T.gondii*

We decided to analyze if the lethal dominant negative phenotype was a direct consequence of mRNA export blocking. For this purpose, we performed FISH assays using oligodT as probe to detect polyA⁺ mRNAs after different periods of 2 μ M Shld1 incubation. After 12 hours of Shld1 incubation we observed a cytoplasmic polyA⁺ mRNA distribution similar to RH strain. After 24 hours of Shld1 incubation, polyA⁺ mRNA molecules accumulated into the nucleus but a clear accumulation is observed only after 48h. After 24h of overexpression, we also noted the presence signals that correspond to overexpression of dd-GFP-TgUPA56 in probably perinucleolar regions. Interestingly, those signals co-localize with a signal correspondent to polyA⁺ mRNAs accumulation (Figure 4).

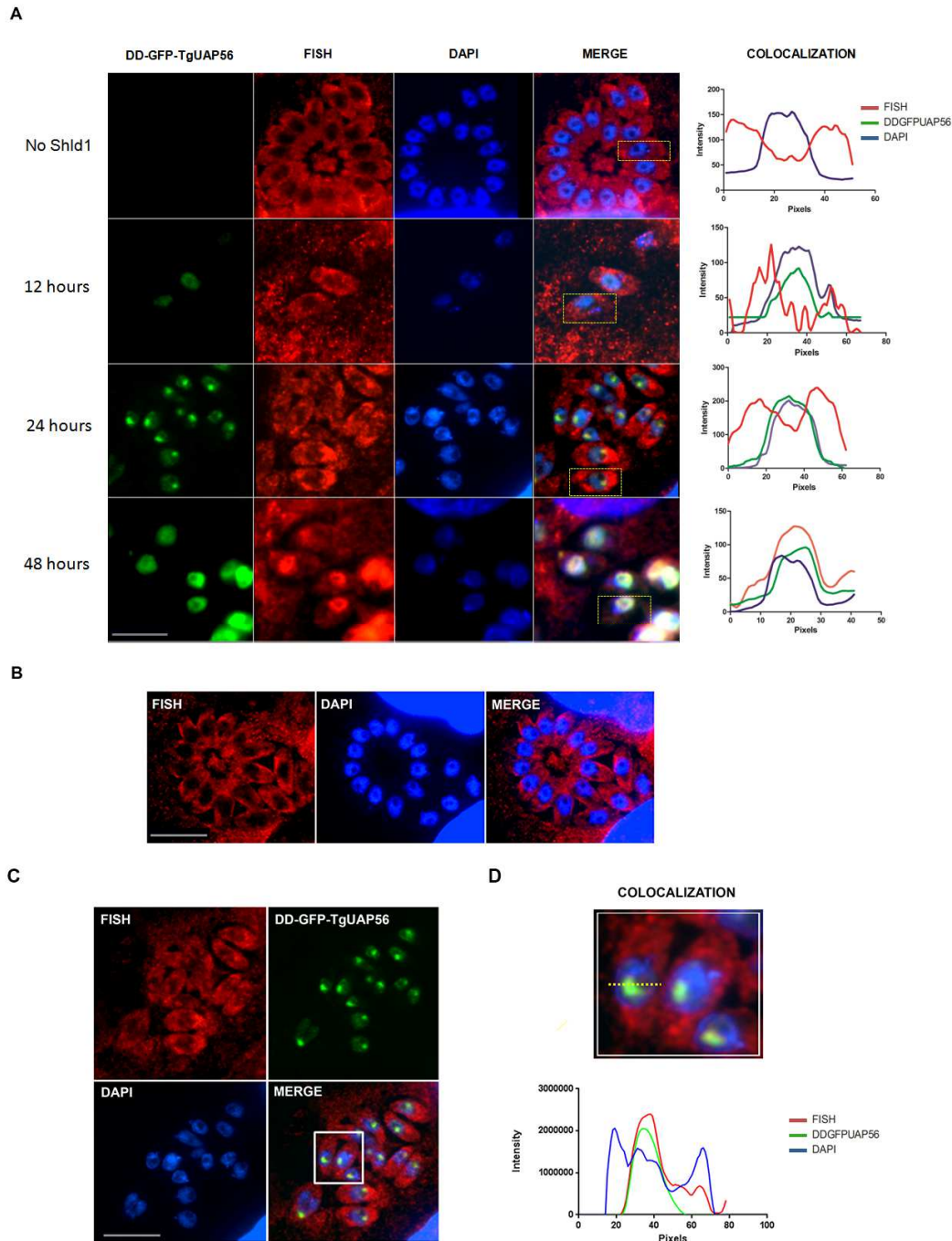


Figure 4. dd-GFP-TgUAP56 overexpression results in mRNA accumulation in the nucleus of *T. gondii*. The polyA⁺ mRNAs were detected by fluorescent *in situ* hybridization (FISH) using oligo(dT) as probe. The oligonucleotide was conjugated to DIG, and a monoclonal antibody against DIG was used as primary antibody: in Red. Nuclear and apicoplast DNA staining with DAPI: in blue. dd-GFP-TgUAP56: in green. A. The stabilization of protein dd-GFP-TgUAP56 was induced by 2 μ M of Shld1 for different time points, as indicated. In the right: quantification of signals of immunofluorescence and DAPI in selected parasites using ImageJ densitometry software (Version 1.6, National Institutes of Health, Bethesda, MD). B. polyA⁺ mRNA distribution in wild-type parasites (RH strain) after 2 μ M of Shld1 incubation for 24 hours. C. polyA⁺ mRNA distribution in ddGFP-TgUAP56 lineage after incubation with 2 μ M of Shld1 for 24 hours. D. Colocalization analysis of dd-GFP-TgUAP56 specific single-dots, mRNAs (FISH) and nucleus (DAPI) of parasite selected on C.

To confirm if the phenotype observed after the overexpression of dd-GFP-TgUAP56 is due to direct function of the protein in mRNA export of *T. gondii*, we used an inducible gene-swap strategy to knock out the gene *uap56* [40]. This strategy is based on site-specific recombination by DiCre which catalyzes excision of DNA flanked by *loxP* sites after incubation with rapamycin, and it has been used in functional studies of *T. gondii* genes [40, 51]. We constructed a vector containing: *uap56* 5'UTR including the endogenous promoter; *uap56* ORF flanked by *loxP* sites; *uap56* 3'UTR; the selection marker HXGPTR and the reporter gene *mcherry*. The DiCre strain was transfected with this vector to generate a stable lineage for conditional knockout of *uap56* gene (cKOuap56). In this system, *uap56* gene flanked by *loxP* is removed by DiCre recombinase after the activation by rapamycin, and the *mcherry* gene is positioned downstream the *uap56* endogenous promoter, resulting in red parasites. Further, we confirmed the correct recombination in the clonal cKOuap56 line by analytical PCR and immunoblot assays, using samples from DiCre strain parasites for comparisons. After 24 hours of rapamycin incubation, the *uap56* ORF was not detected by PCR (Supp Figure 2, lane B1-3; primers uap56F/R) and consequently the protein was not expressed (Figure 5A).

The cKOuap56 line was used to follow the polyA⁺ mRNA distribution during different times after incubation with rapamycin. The mRNA was detected by FISH using oligodT as probe. The figure 5B shows nuclear accumulation of mRNA in some parasites after 24 hours of incubation. The accumulation is evident in most of parasites after 48 hours. The red signal corresponding to *mcherry* expression was visible as the mRNA accumulation occurred, confirming that *uap56-loxp* was excised successfully. Once the knockout is induced, it was not possible to keep parasites alive in culture. This lethal phenotype was confirmed by plaque assays. There were not lysis plaques after 108h of incubation with rapamycin in the cKOuap56 culture, while several lysis plaques were observed in DiCre parental strain, confirming that TgUAP56 is an essential protein in *T. gondii* (Figure 5C).

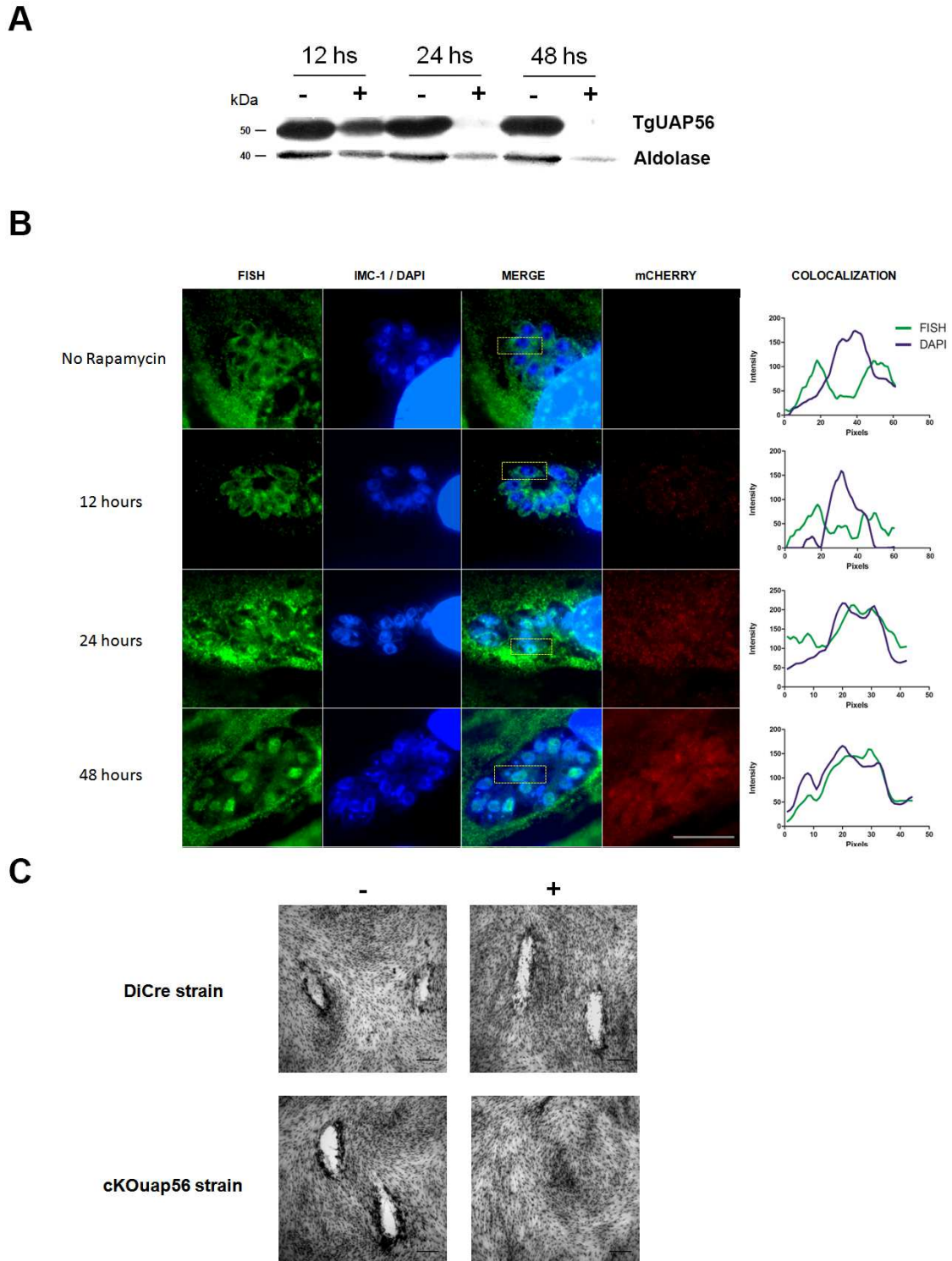


Figure 5. *uap56* knockout is lethal and results in mRNA accumulation into nucleus of *T. gondii*. The conditional knockout of *uap56* gene was performed using a conditional site-specific recombination system based on dimerizable Cre recombinase-mediated recombination [40]. A. TgUAP56 analysis in protein extract of cKOuap56 lineage incubated with 50 nM of rapamycin in different times, as indicated, by Western blot. -, not induced; +, induced. B. Analysis of mRNA distribution in cKOuap56 lineage after incubation with rapamycin in different times, as indicated. The polyA⁺ mRNAs were detected by fluorescent *in situ* hybridization (FISH) using oligodT as probe. The oligonucleotide was conjugated to DIG, and a monoclonal antibody against DIG was used as primary antibody: in green. Nuclear and apicoplast DNA staining with

DAPI: in blue. In the right: quantification of signals of immunofluorescence and DAPI in selected parasites using ImageJ densitometry software (Version 1.6, National Institutes of Health, Bethesda, MD). Bar = 10 μ m. C. Plaque assays for DiCre strain and cKOuap56 lineage. Both parasite lineages were grown on human foreskin fibroblasts, in the presence of 50 nM of rapamycin, as indicated, for 108 hours and then fixed and stained with Giemsa. Rapamycin was washed off after 24 hours of incubation. -, not induced; +, induced Bars = 500 μ m.

TgUAP56 knockout does not affect mRNA splicing

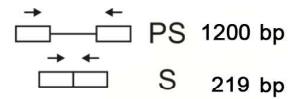
In order to examine if *uap56* knockout would affect mRNA splicing, we surveyed a selection of genes by a semi-quantitative analysis developed by Suvorova *et al*, 2013 [23], using primers spanning an intron as listed in Supp. Table 2. Parasites of cKOuap56 line were treated with 50 nM of rapamycin for 24 and 48 hours before extraction of total RNA for analysis. Parasites of DiCre strain treated with rapamycin were used as control of this experiment. The PCR data did not show accumulation of pre-mRNA for all the genes analyzed (Figure 6), different from that observed for TgRRM1, a known splicing factor [23]. Genomic DNA extracted from cKOuap56 parasites was used as a reference for intron-containing pre-spliced mRNAs.

A

Selected genes	Pre-Spliced (PS)	Spliced (S)
PolII p8.2	1200 bp	219 bp
PolII p19	918 bp	375 bp
PolII p23	850 bp	538 bp
IMC1	919 bp	530 bp
IMC15	1113 bp	442 bp
IMC5	780 bp	356 bp

B

PolII p8.2 analysis
Pre-Spliced x Spliced mRNA



C

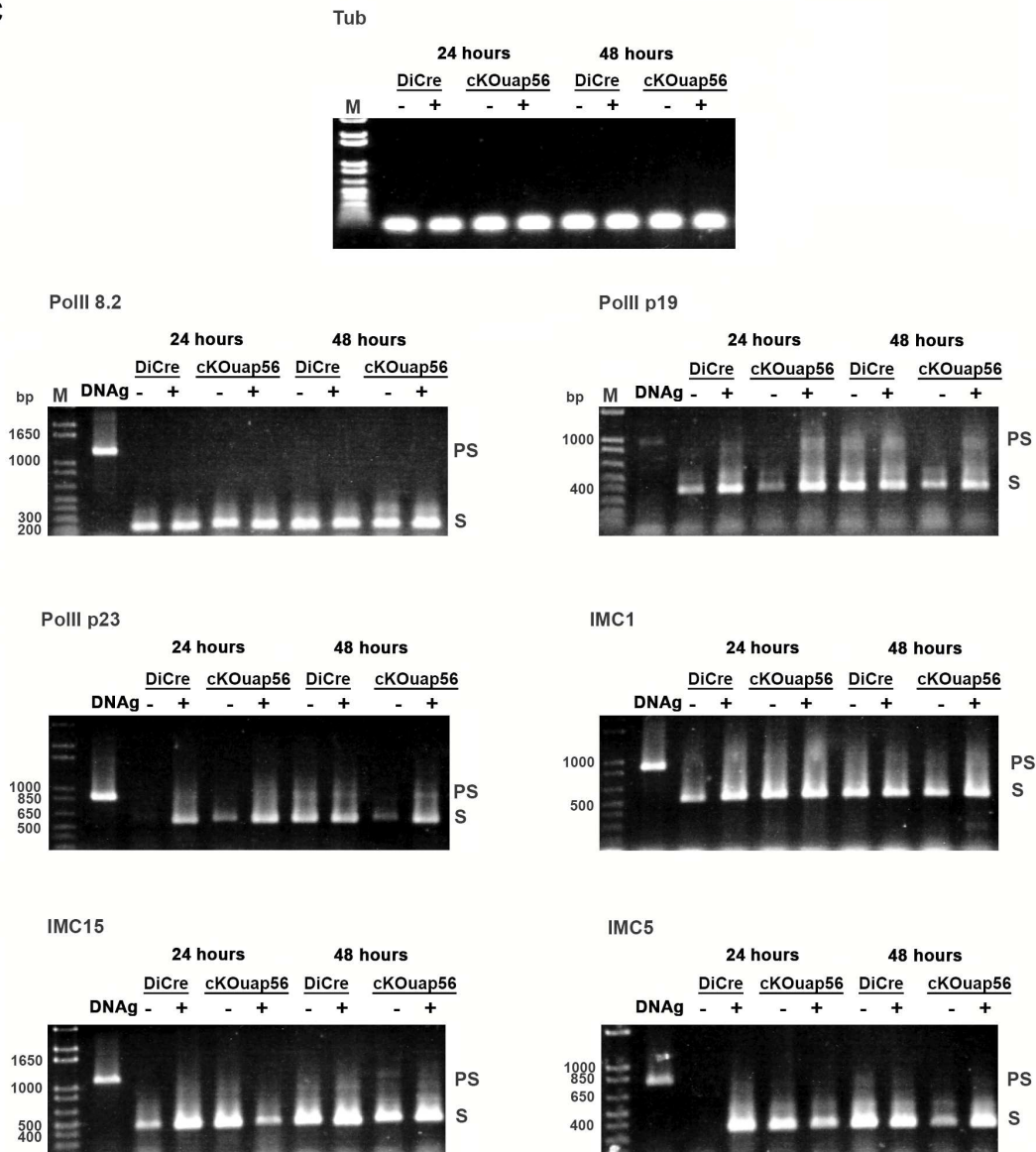


Figure 6. PCR analysis of mRNA splicing for seven selected genes. Total RNA was purified from DiCre strain and cKOuap56 incubated with 50 nM of rapamycin, as indicated. The total RNA was reverse transcribed and PCR amplified using primers that span an intron. PCR of genomic DNA was included as a reference to distinguish between properly spliced (S) and pre-spliced (PS) forms of each gene. A. The following genes were analyzed: RNA polymerase II p8.2 subunit (TGME49_217560), RNA polymerase II p19 subunit (TGME49_271300), RNA polymerase II p23 subunit (TGME49_240590), IMC1 (TGME49_231640), IMC15 (TGME49_275670), IMC5 (TGME49_224530), and transcription factor IID (TGME49_258680).

Expected sizes for pre-mRNA (Pre-spliced) and mRNA (spliced) are shown for each selected gene. B. Example of possible results for RNA polymerase II p8.2 subunit (TGME49_217560). C. PCR analysis of mRNA splicing for selected genes. Tubulin was used as loading control [44].

Discussion

In eukaryotes, transcription and translation are physically separated by the nuclear envelope. The mRNAs are exported from the nucleus as messenger ribonucleoprotein complex (mRNPs) for translation in the cytoplasm. Even if post-transcriptional events are important to development of apicomplexan parasites, no description is available about the components or mechanisms of mRNA export in these organisms. The mRNA export pathway is well studied in model organisms and the basic proteins are the components of THO/TREX complex which involves UAP56 and adaptor proteins (for review see [52]). However, in parasites, several proteins involved in this pathway are not conserved [30]. The exception is UAP56, a component of TREX in human cells. We previously demonstrated that this protein is also a component of mRNA export pathway in Trypanosomes [31]. In apicomplexa, there is only an article describing the ortholog of UAP56 in *P. falciparum*, PfU52, an RNA-dependent ATPase and RNA helicase [32]. However, the authors did not investigate the relationship of PfU52 with mRNA export.

Firstly, we confirmed by sequence and phylogenetic analysis that UAP56 is conserved across major eukaryotic groups. The UAP56 ortholog in *T. gondii*, named TgUAP56, has two RNA helicase domains and it is highly conserved along its entire sequence to the other eukaryotic sequences. The most similar sequence to TgUAP56 is the *P. falciparum* ortholog PfU52 (82% similar), while the most divergent sequence is *T. cruzi* ortholog TcSub2. Anyhow, TcSub2 has high degree of conservation with TgUAP56 (62.2% similar) and we observed that sequences of Apicomplexa and Kinetoplastida clustered together showing they are closely related. Because TcSub2 is a component of the mRNA export pathway [30], it would be reasonable to hypothesize that TgUAP56 has the same role as basic component of the mRNA export pathway. However, phylogenetic relationship is not a guarantee that proteins has the same function. In this work we presented experimental data supporting the idea that UAP56 is also an essential factor for mRNA export in *T.gondii*.

TgUAP56 is an exclusively nuclear dispersed all over the nuclei in a punctuate pattern, as in other eukaryotes [31, 53, 54]. Overexpression of this protein results in a dominant negative phenotype due to mRNA export blocking, indicating that TgUAP56 is an essential protein involved in mRNA export. Similarly, overexpression of UAP56 is lethal to the worms at all stages during *C. elegans* development because it blocks mRNA export [55]. In human cells and yeast, the excess of UAP56/Sub2 impairs mRNA export in a dominant negative manner, by titrating one or more other mRNA export factors, preventing its recruitment to spliced mRNPs and leading to nuclear retention of polyA+ mRNAs [56, 57].

To confirm if this overexpression effect is directly related to TgUAP56, we created conditional knockout of the *uap56* gene. The *T. gondii* DiCre recombinase strategy [40] was successfully used to remove *uap56* gene in this work. The ongoing of the polyA+ mRNA accumulation phenotype was observed just after 24 hours of DiCre activation by rapamycin. The nuclear accumulation of polyA+ mRNA was robust after 48 hours, corresponding to growth inhibition of parasites and, thus, confirming the direct involvement of TgUAP56 in mRNA export of *T.gondii*. These results are in accordance with previous publications, for example, the depletion of homologs of UAP56 in yeast and *D. melanogaster* resulted in growth arrest and a robust accumulation of polyA+ mRNAs within the nucleus [53, 56]. Also, in trypanosomes, the knockdown of TbSub2 caused polyA+ mRNA accumulation in the nucleus leading to a general decrease in translation levels and cell death [31].

We also decided to investigate if the knockout of *uap56* could affect the splicing of mRNAs since the orthologs were originally associated to splicing machinery (reviewed in Linder and Stutz, 2001 [58]). In *S. cerevisiae*, Sub2 can be involved in spliceosome assembly and splicing events, in addition to function in export of mRNAs [59, 60], [56, 61, 62]. In *P. falciparum*, the ortholog PfU52 has in vitro splicing activity and its function in mRNA export was not studied yet [32]. Our approach to study the putative function of TgUAP56 in splicing was based on the strategies used previously for the characterization of the splicing factor TgRRM1 [23], where the authors showed that a selected list of genes of *T.gondii* was miss-spliced in the absence of TgRRM1. Herein we did not

observe any interference in these targets with mRNA splicing absence of TgUAP56. Indeed, the depletion of TgUAP56 did not affect the pre-mRNA splicing of these genes but exhibited a strong bulk mRNA export defect. All these data, led us to conclude that TgUAP56 is involved in mRNA export, probably releasing spliced mRNAs to adaptor proteins. This idea is corroborated by several studies in model organisms that indicate that UAP56 and Sub2 play an essential role in mRNA export [53, 56, 57, 62] even though they were originally implicated in splicing (for review see [58]). In *S. pombe*, Sub2 is not required for pre-mRNA splicing [61]. In *Drosophila* HEL/UAP56 has a preferential association with spliced mRNAs and is involved in the export of both intron-containing and intronless mRNAs. Moreover, the mRNA export block is not a consequence of general inhibition of splicing when this protein is depleted [53]. In human cells, a splicing-coupled mechanism is related to the recruitment of UAP56 complex for spliced mRNAs [63]. A preferential UAP56 binding of the spliced mRNA over pre-mRNA can explain the direct function in releasing spliced polyA⁺ mRNA from nuclear speckles [64, 65], providing a functional link between splicing and export.

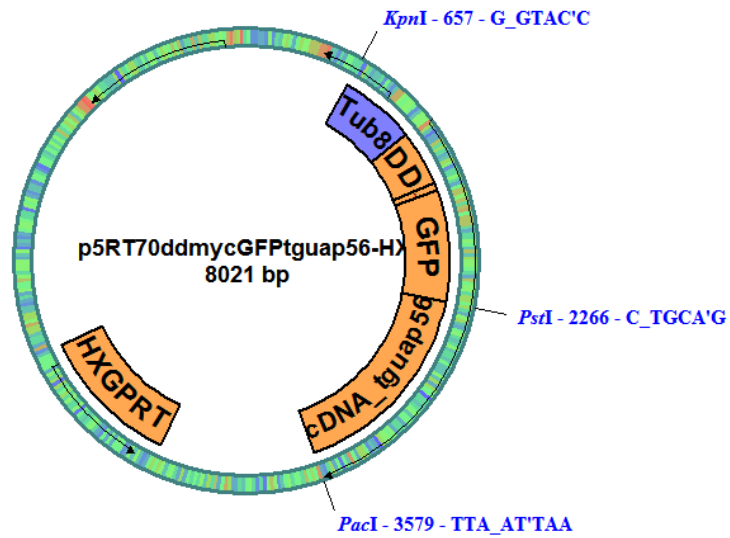
Taking together, our experimental evidences showed that TgUAP56 protein has similar function to its orthologs in other eukaryotes. By using suitable tools to identify the function of this essential gene, we were able to conclude that TgUAP56 is involved in mRNA export pathway in *T.gondii*. Moreover, the conserved function of UAP56 in parasites from different groups, as Kinetoplastida and Apicomplexa, indicates that it is a basic component of mRNA export in protozoa species. Since UAP56 is a basic and specific component of a specialized mRNA export in mammalian, the identification of homologue proteins in parasites would point out the presence of a specialized pathway in these organisms. Naturally, apicomplexas and other parasites acquired peculiar features in relation to mRNA metabolism during evolution and we cannot rule out that it is still necessary to get further insights into function of others component of mRNA export pathway to investigate how similar it is in relation to mammals.

1. Gilbert, R.E. and M.R. Stanford, *Is ocular toxoplasmosis caused by prenatal or postnatal infection?* Br J Ophthalmol, 2000. **84**(2): p. 224-6.
2. Holland, G.N., *Reconsidering the pathogenesis of ocular toxoplasmosis.* Am J Ophthalmol, 1999. **128**(4): p. 502-5.
3. Holland, G.N., *Ocular toxoplasmosis: a global reassessment. Part I: epidemiology and course of disease.* Am J Ophthalmol, 2003. **136**(6): p. 973-88.
4. Tenter, A.M., A.R. Heckeroth, and L.M. Weiss, *Toxoplasma gondii: from animals to humans.* Int J Parasitol, 2000. **30**(12-13): p. 1217-58.
5. Bahia-Oliveira, L.M., et al., *Highly endemic, waterborne toxoplasmosis in north Rio de Janeiro state, Brazil.* Emerg Infect Dis, 2003. **9**(1): p. 55-62.
6. Silveira, C., et al., *A follow-up study of Toxoplasma gondii infection in southern Brazil.* Am J Ophthalmol, 2001. **131**(3): p. 351-4.
7. Dubey, J.P., et al., *Toxoplasmosis in humans and animals in Brazil: high prevalence, high burden of disease, and epidemiology.* Parasitology, 2012. **139**(11): p. 1375-424.
8. Matos, S.B., et al., *Prevalence of serological markers and risk factors for bloodborne pathogens in Salvador, Bahia state, Brazil.* Epidemiol Infect, 2013. **141**(1): p. 181-7.
9. Vallochi, A.L., et al., *The genotype of Toxoplasma gondii strains causing ocular toxoplasmosis in humans in Brazil.* Am J Ophthalmol, 2005. **139**(2): p. 350-1.
10. Khan, A., et al., *Genetic divergence of Toxoplasma gondii strains associated with ocular toxoplasmosis, Brazil.* Emerg Infect Dis, 2006. **12**(6): p. 942-9.
11. Gilbert, R.E., et al., *Ocular sequelae of congenital toxoplasmosis in Brazil compared with Europe.* PLoS Negl Trop Dis, 2008. **2**(8): p. e277.
12. Kim, K. and L.M. Weiss, *Toxoplasma gondii: the model apicomplexan.* Int J Parasitol, 2004. **34**(3): p. 423-32.
13. Meissner, M., et al., *Molecular genetic tools in Toxoplasma and Plasmodium: achievements and future needs.* Curr Opin Microbiol, 2007. **10**(4): p. 349-56.
14. Hill, D. and J.P. Dubey, *Toxoplasma gondii: transmission, diagnosis and prevention.* Clin Microbiol Infect, 2002. **8**(10): p. 634-40.
15. Sullivan Jr, W.J., A.T. Smith, and B.R. Joyce, *Understanding mechanisms and the role of differentiation in pathogenesis of Toxoplasma gondii: a review.* Mem Inst Oswaldo Cruz, 2009. **104**(2): p. 155-61.
16. Behnke, M.S., et al., *The transcription of bradyzoite genes in Toxoplasma gondii is controlled by autonomous promoter elements.* Mol Microbiol, 2008. **68**(6): p. 1502-18.
17. Bougdour, A., et al., *Chromatin modifications: implications in the regulation of gene expression in Toxoplasma gondii.* Cell Microbiol, 2010. **12**(4): p. 413-23.
18. Radke, J.R., et al., *The transcriptome of Toxoplasma gondii.* BMC Biol, 2005. **3**: p. 26.
19. Vonlaufen, N., et al., *Stress response pathways in protozoan parasites.* Cell Microbiol, 2008. **10**(12): p. 2387-99.
20. Zhang, M., et al., *Translational control in Plasmodium and toxoplasma parasites.* Eukaryot Cell, 2013. **12**(2): p. 161-7.
21. Lirussi, D. and M. Matrajt, *RNA granules present only in extracellular toxoplasma gondii increase parasite viability.* Int J Biol Sci, 2011. **7**(7): p. 960-7.
22. Patil, V., et al., *Disruption of the Expression of a Non-Coding RNA Significantly Impairs Cellular Differentiation in Toxoplasma gondii.* Int J Mol Sci, 2012. **14**(1): p. 611-24.
23. Suvorova, E.S., et al., *Discovery of a splicing regulator required for cell cycle progression.* PLoS Genet, 2013. **9**(2): p. e1003305.
24. Fried, H. and U. Kutay, *Nucleocytoplasmic transport: taking an inventory.* Cell Mol Life Sci, 2003. **60**(8): p. 1659-88.
25. Cullen, B.R., *Nuclear RNA export.* J Cell Sci, 2003. **116**(Pt 4): p. 587-97.
26. Kohler, A. and E. Hurt, *Exporting RNA from the nucleus to the cytoplasm.* Nat Rev Mol Cell Biol, 2007. **8**(10): p. 761-73.

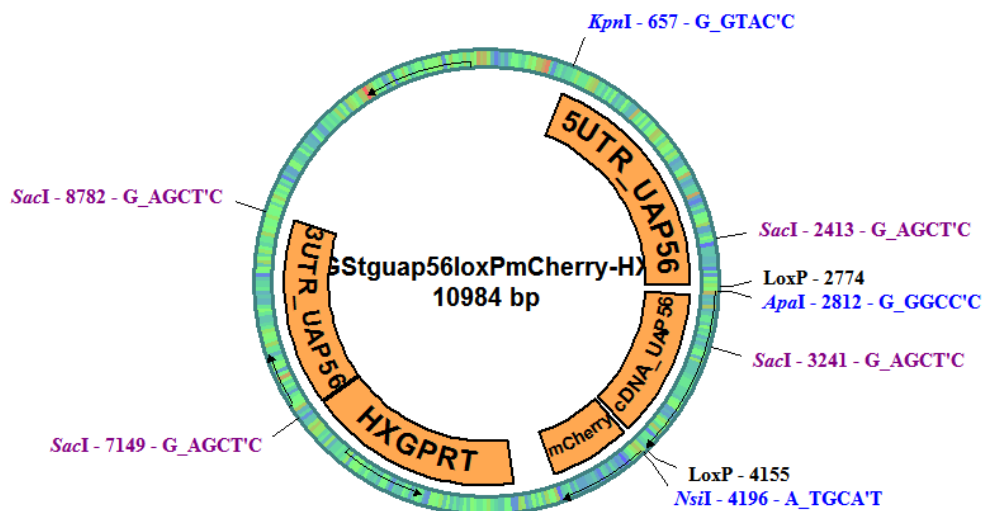
27. Rodriguez-Navarro, S., et al., *Sus1, a functional component of the SAGA histone acetylase complex and the nuclear pore-associated mRNA export machinery*. *Cell*, 2004. **116**(1): p. 75-86.
28. Wolyniak, M.J. and C.N. Cole, *Harnessing genomics to explore the processes and evolution of mRNA export*. *RNA Biol*, 2008. **5**(2): p. 68-72.
29. Nino, C.A., et al., *mRNA nuclear export in yeast*. *Chem Rev*, 2013. **113**(11): p. 8523-45.
30. Serpeloni, M., et al., *Comparative genomics of proteins involved in RNA nucleocytoplasmic export*. *BMC Evol Biol*, 2011. **11**: p. 7.
31. Serpeloni, M., et al., *An essential nuclear protein in trypanosomes is a component of mRNA transcription/export pathway*. *PLoS One*, 2011. **6**(6): p. e20730.
32. Shankar, J., A. Pradhan, and R. Tuteja, *Isolation and characterization of Plasmodium falciparum UAP56 homolog: evidence for the coupling of RNA binding and splicing activity by site-directed mutations*. *Arch Biochem Biophys*, 2008. **478**(2): p. 143-53.
33. Altschul, S.F., et al., *Basic local alignment search tool*. *J Mol Biol*, 1990. **215**(3): p. 403-10.
34. Eddy, S.R., *Profile hidden Markov models*. *Bioinformatics*, 1998. **14**(9): p. 755-63.
35. Finn, R.D., et al., *Pfam: the protein families database*. *Nucleic Acids Res*, 2014. **42**(Database issue): p. D222-30.
36. Katoh, K. and D.M. Standley, *MAFFT multiple sequence alignment software version 7: improvements in performance and usability*. *Mol Biol Evol*, 2013. **30**(4): p. 772-80.
37. Price, M.N., P.S. Dehal, and A.P. Arkin, *FastTree 2--approximately maximum-likelihood trees for large alignments*. *PLoS One*, 2010. **5**(3): p. e9490.
38. Rice, P., I. Longden, and A. Bleasby, *EMBOSS: the European Molecular Biology Open Software Suite*. *Trends Genet*, 2000. **16**(6): p. 276-7.
39. Donald R.G., C.D., Ullman B., Roos D.S., *Insertional tagging, cloning, and expression of the Toxoplasma gondii hypoxanthine-xanthine-guanine phosphoribosyltransferase gene. Use as a selectable marker for stable transformation*. *J Biol Chem*, 1996. **271**(24).
40. Andenmatten N., E.S., Jackson A.J., Jullien N., Herman J.P., Meissner M., *Conditional genome engineering in Toxoplasma gondii uncovers alternative invasion mechanisms*. *Nature Methods*, 2013. **10**.
41. Hettmann C., H.A., Geiter A., Frank B., Schwarz E., Soldati T., Soldati D., *A dibasic motif in the tail of a class XIV apicomplexan myosin is an essential determinant of plasma membrane localization*. *Mol Biol Cell*, 2000. **11**(4).
42. Starnes, G.L., et al., *Two separate, conserved acidic amino acid domains within the Toxoplasma gondii MIC2 cytoplasmic tail are required for parasite survival*. *J Biol Chem*, 2006. **281**(41): p. 30745-54.
43. Roos DS, D.R., Morrissette NS, Moulton AL, *Molecular tools for genetic dissection of the protozoan parasite Toxoplasma gondii*. *Methods Cell Biol*, 1994. **45**.
44. Dalmaso, M.C., et al., *Toxoplasma H2A variants reveal novel insights into nucleosome composition and functions for this histone family*. *J Mol Biol*, 2009. **392**(1): p. 33-47.
45. Rocak, S. and P. Linder, *DEAD-box proteins: the driving forces behind RNA metabolism*. *Nat Rev Mol Cell Biol*, 2004. **5**(3): p. 232-41.
46. Armstrong, C.M. and D.E. Goldberg, *An FKBP destabilization domain modulates protein levels in Plasmodium falciparum*. *Nat Methods*, 2007. **4**(12): p. 1007-9.
47. Herm-Gotz, A., et al., *Rapid control of protein level in the apicomplexan Toxoplasma gondii*. *Nat Methods*, 2007. **4**(12): p. 1003-5.
48. Pieperhoff, M.S., et al., *The role of clathrin in post-Golgi trafficking in Toxoplasma gondii*. *PLoS One*, 2013. **8**(10): p. e77620.
49. Kremer, K., et al., *An overexpression screen of Toxoplasma gondii Rab-GTPases reveals distinct transport routes to the micronemes*. *PLoS Pathog*, 2013. **9**(3): p. e1003213.
50. Banaszynski, L.A., et al., *A rapid, reversible, and tunable method to regulate protein function in living cells using synthetic small molecules*. *Cell*, 2006. **126**(5): p. 995-1004.

51. Bargieri, D.Y., et al., *Apical membrane antigen 1 mediates apicomplexan parasite attachment but is dispensable for host cell invasion*. Nat Commun, 2013. **4**: p. 2552.
52. Muller-McNicoll, M. and K.M. Neugebauer, *How cells get the message: dynamic assembly and function of mRNA-protein complexes*. Nat Rev Genet, 2013. **14**(4): p. 275-87.
53. Gatfield, D., et al., *The DExH/D box protein HEL/UAP56 is essential for mRNA nuclear export in Drosophila*. Curr Biol, 2001. **11**(21): p. 1716-21.
54. Sahni, A., N. Wang, and J.D. Alexis, *UAP56 is an important regulator of protein synthesis and growth in cardiomyocytes*. Biochem Biophys Res Commun, 2010. **393**(1): p. 106-10.
55. MacMorris, M., C. Brocker, and T. Blumenthal, *UAP56 levels affect viability and mRNA export in Caenorhabditis elegans*. RNA, 2003. **9**(7): p. 847-57.
56. Strasser, K. and E. Hurt, *Splicing factor Sub2p is required for nuclear mRNA export through its interaction with Yra1p*. Nature, 2001. **413**(6856): p. 648-52.
57. Luo, M.L., et al., *Pre-mRNA splicing and mRNA export linked by direct interactions between UAP56 and Aly*. Nature, 2001. **413**(6856): p. 644-7.
58. Linder, P. and F. Stutz, *mRNA export: travelling with DEAD box proteins*. Curr Biol, 2001. **11**(23): p. R961-3.
59. Libri, D., et al., *Multiple roles for the yeast SUB2/yUAP56 gene in splicing*. Genes Dev, 2001. **15**(1): p. 36-41.
60. Fleckner, J., et al., *U2AF65 recruits a novel human DEAD box protein required for the U2 snRNP-branchpoint interaction*. Genes Dev, 1997. **11**(14): p. 1864-72.
61. Thakurta, A.G., et al., *The nuclear export signal of splicing factor Uap56p interacts with nuclear pore-associated protein Rae1p for mRNA export in Schizosaccharomyces pombe*. J Biol Chem, 2007. **282**(24): p. 17507-16.
62. Jensen, T.H., et al., *The DECD box putative ATPase Sub2p is an early mRNA export factor*. Curr Biol, 2001. **11**(21): p. 1711-5.
63. Masuda, S., et al., *Recruitment of the human TREX complex to mRNA during splicing*. Genes Dev, 2005. **19**(13): p. 1512-7.
64. Dias, A.P., et al., *A role for TREX components in the release of spliced mRNA from nuclear speckle domains*. Nat Commun, 2010. **1**: p. 97.
65. Steckelberg, A.L. and N.H. Gehring, *Studying the composition of mRNPs in vitro using splicing-competent cell extracts*. Methods, 2014. **65**(3): p. 342-9.

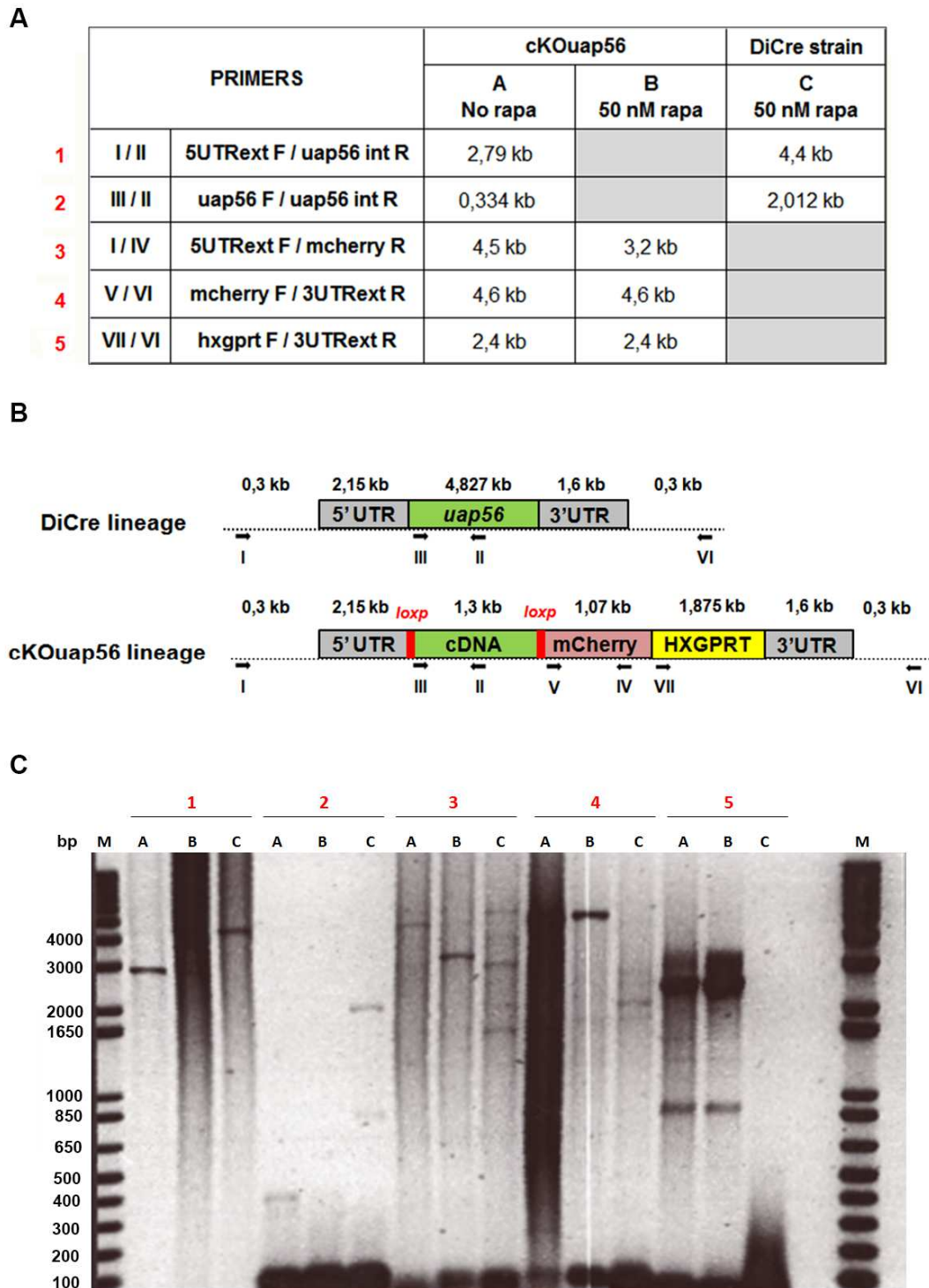
A



B



Supp. Fig. 1. Schematic map of vectors used for dd-GFP-TgUAP56 overexpression and *uap56* gene-swap strategies. A. Vector *p5RT70ddmycGFPTguap56-HX*, used for dd-GFP-TgUAP56 overexpression. B. Vector *GSuap56loxPmCherry-HX*, used for *uap56* Gene-swap strategy. 5'UTR: *uap56* 5'untranslated region upstream of *uap56* (5'UTR), containing the endogenous promoter. cDNA_UAP56: coding sequence of *uap56*, flanked by *loxP* sequences in same orientation. *mcherry*: reporter gene. 3'UTR: *uap56* 3'untranslated region downstream of *uap56* (3'UTR). *hxgprt*: selection marker. The restriction enzymes used are indicated in both figures.



Supp. Fig. 2. Genomic PCR analyzes of cKOuap56 lineage. A. Primers and expected sizes of PCR products. Primers sequences are available on Supp. Table 2. Letters indicate parasites analyzed: A, B: cKOuap56 lineage. C: DiCre strain. 50 nM of rapamycin was incubated for 24 hours, as indicated. Expected sizes are shown for each primer pair in kb. Gray cells indicate absence of expected amplification. B. Schematic draw of *uap56* locus in DiCre strain and cKOuap56 lineage and region for annealing primers. Size of each region is indicated above. cDNA: *uap56* coding region. C. PCR results on 1% Agarose gel. Numbers and letters are related with A. M. 1 kb Plus DNA Marker (Invitrogen).

Supplementary Table 1: Supplementary Table 1: Genebank Identifier (GI) numbers of UAP56 ortholog proteins identified in 43 representative species of different eukaryotic groups. Amoebozoa: Ddi – *Dictyostelium discoideum*; Plants: Osa – *Oryza sativa*, Ppa – *Physcomitrella patens*, Ath – *Arabidopsis thaliana*, Pop – *Populus trichocarpa*, Olu – *Ostreococcus lucimarinus*, Cre – *Chlamydomonas reinhardtii*, Vca – *Volvox carteri*; Fungi: Uma – *Ustilago maydis*, Cne – *Cryptococcus neoformans*, Spo – *Schizosaccharomyces pombe*, Afu – *Aspergillus fumigatus*, Fgr – *Fusarium graminearum*, Mor – *Magnaporthe oryzae*, Ncr – *Neurospora crassa*, Yli – *Yarrowia lipolytica*, Dha – *Debaryomyces hansenii*, Ago – *Ashbya gossypii*, Kla – *Kluyveromyces lactis*, Cgl – *Candida glabrata*, Sce – *Saccharomyces cerevisiae*; Metazoans: Cbr – *Caenorhabditis briggsae*, Cel – *Caenorhabditis elegans*, Cin – *Ciona intestinalis*, Hsa – *Homo sapiens*, Mmu – *Mus musculus*, Dme – *Drosophila melanogaster*, Ame – *Apis mellifera*, Mbr – *Monosiga brevicollis*; Parabasalid: Tva – *Trichomonas vaginalis*; Apicomplexans: Tpa – *Theileria parva*, Cho – *Cryptosporidium hominis*, Cpa – *Cryptosporidium parvum*, Pfa – *Plasmodium falciparum*, Pbe – *Plasmodium berghei*, Pkn – *Plasmodium knowlesi*, Pvi – *Plasmodium vivax*, Tgo – *Toxoplasma gondii*; Kinetoplastids: Lbr – *Leishmania braziliensis*, Lin – *Leishmania infantum*, Lma – *Leishmania major*, Tcr – *Trypanosoma cruzi*, Tbr – *Trypanosoma brucei*

Eukaryotic Group	Species	Code	GI
Metazoans	<i>Homo sapiens</i>	Hsa	4758112
Metazoans	<i>Mus musculus</i>	Mmu	9790069
Metazoans	<i>Ciona intestinalis</i>	Cin	198430623
Metazoans	<i>Drosophila melanogaster</i>	Dme	24581952
Metazoans	<i>Apis mellifera</i>	Ame	66532824
Metazoans	<i>Caenorhabditis briggsae</i>	Cbr	157749783
Metazoans	<i>Caenorhabditis elegans</i>	Cel	71982858
Choanoflagellate	<i>Monosiga brevicollis</i>	Mbr	167522271
Fungi	<i>Saccharomyces cerevisiae</i>	Sce	6320119
Fungi	<i>Candida glabrata</i>	Cgl	50293331
Fungi	<i>Ashbya gossypii</i>	Ago	45188133
Fungi	<i>Kluyveromyces lactis</i>	Kla	50309861
Fungi	<i>Debaryomyces hansenii</i>	Dha	50424083
Fungi	<i>Yarrowia lipolytica</i>	Yli	210075100
Fungi	<i>Magnaporthe oryzae</i>	Mor	145610002
Fungi	<i>Neurospora crassa</i>	Ncr	164423323
Fungi	<i>Fusarium graminearum</i>	Fgr	46128571
Fungi	<i>Aspergillus fumigatus</i>	Afu	70984615
Fungi	<i>Schizosaccharomyces pombe</i>	Spo	19115173
Fungi	<i>Cryptococcus neoformans</i>	Cne	58259249
Fungi	<i>Ustilago maydis</i>	Uma	71021713
Amoebozoa	<i>Dictyostelium discoideum</i>	Ddi	66826121
Plantae	<i>Arabidopsis thaliana</i>	Ath	18416493
Plantae	<i>Oryza sativa</i>	Osa	115437448
Plantae	<i>Populus trichocarpa</i>	Pop	224092470
Plantae	<i>Physcomitrella patens</i>	Ppa	168035593
Plantae	<i>Chlamydomonas reinhardtii</i>	Cre	159491657
Plantae	<i>Volvox carteri</i>	Vca	302836427
Plantae	<i>Ostreococcus lucimarinus</i>	Olu	145356615
Apicomplexa	<i>Theileria parva</i>	Tpa	71027499
Apicomplexa	<i>Plasmodium berghei</i>	Pbe	68073823
Apicomplexa	<i>Plasmodium falciparum</i>	Pfa	124801080
Apicomplexa	<i>Plasmodium knowlesi</i>	Pkn	221053125
Apicomplexa	<i>Plasmodium vivax</i>	Pvi	156093980
Apicomplexa	<i>Toxoplasma gondii</i>	Tgo	237843393
Apicomplexa	<i>Cryptosporidium hominis</i>	Cho	67623541
Apicomplexa	<i>Cryptosporidium parvum</i>	Cpa	66361130
Kinetoplastida	<i>Leishmania braziliensis</i>	Lbr	154337278
Kinetoplastida	<i>Leishmania infantum</i>	Lin	146086104
Kinetoplastida	<i>Leishmania major</i>	Lma	157869116
Kinetoplastida	<i>Trypanosoma brucei</i>	Tbr	71746514
Kinetoplastida	<i>Trypanosoma cruzi</i>	Tcr	71411514
Parabasalid	<i>Trichomonas vaginalis</i>	Tva	123449042

Supplementary Table 2: Primers used in this study. Name and Sequence of individual primers are listed. The restriction enzymes used are in the name of primer and its respective recognized sequence in *italic*. *loxP* sequences added to reverse primers are underlined.

TgUAP56 Overexpression vector	
Name	Sequence 5' to 3'
<i>uap56 - PstI</i> (F)	CCCCTGCAAGGAATGACGACTCTCGAGCAGAACC
<i>uap56 - PacI</i> (R)	TTAATTAATCACTGATTGATGACTGCGAGGCG

<i>uap56</i> Gene swap vector	
Name	Sequence 5' to 3'
5UTR- <i>KpnI</i> (F)	GGGGTACCCTGCTTGAGGTCTTGCTTGC
5UTR- <i>Loxp-ApaI</i> (R)	GGGGGCCCATAACTTCGTATAGCATACATTATACGAAGTTA <u>TTTGTC</u> AACGCGTTGAGGG
<i>uap56-ApaI</i> (F)	GGGGGCCCATGACGACTCTCGAGCAGAACC
<i>uap56-Loxp-NsiI</i> (R)	GCTAATGCATTTTATAACTTCGTATAGCATACATTATACGAAGTTA <u>TTTACTG</u> ATTGATGACTGCGAGG
3UTR- <i>SacI</i> (F)	GGGAGCTCAGTCGAACACAGTAAACGTTTCTTG
3UTR- <i>SacI</i> (R)	GGGAGCTCCCTAGGAAGGCGCTGGATCCTTATT

Confirmation of cKOuap56 lineage		
Number	Name	Sequence 5' to 3'
I	5UTR ext F (external)	GTATTAGTACATGGTGTCAAGCG
II	UAP56 int R (internal)	CACTTGTGTCCAGGTTCAAGCT
III	UAP56 F	ATGACGACTCTCGAGCAGAACC
IV	mcherry R (internal)	CACCGCGGCATGGACGAGCTGTACAAG
V	mcherry F	ATGCTATACGAAGTTATAAAATGCA
VI	3UTR R ext (external)	CATGTCGGGTTAGTTCATATCGT
VII	HX F	GCTCAATACGGCCCTTGCCCTTGC

Analysis of mRNA processing by splicing	
PolII p8.2 (F)	ATGATCATCCCAATTCGGTGCT
PolII p8.2 (R)	TCAGCTCTCCACAATGCGCTT
PolII p19 (F)	GTCCTTGACGCCGTCATCACC
PolII p19 (R)	TCAAAGCGTCGGAGAATGGGG
PolII p23 (F)	GTCAAAGAGATTGAGCTTCCAAGGATTC
PolII p23 (R)	GTCACGAGAAGTTCAGTACACG
IMC1 (F)	GAGGTACCTGAGGTTGAGTACGTC
IMC1 (R)	AAGACCTGCTCCACCTTGGG
IMC15 (F)	GTTCCAAAGGAATTCACGCGCATTG
IMC15 (R)	CTGTTAGCCGTTCTTTATCTAGGTGAGTG
IMC5 (F)	CTGTTAGCCGTTCTTTATCTAGGTGAGTG
IMC5 (R)	CTCCATTCGACAGTCAAGAGTCCAG
TFIID (F)	ATGACAACCTGAGGACGACGAGACA
TFIID (R)	CTACTGATGGTAGCGAACAAGCACG