UNIVERSIDADE DE SÃO PAULO

FACULDADE DE CIÊNCIAS FARMACÊUTICAS DE RIBEIRÃO PRETO

Neospora caninum:
estudo do secretoma e caracterização molecular
de três proteínas com domínios Apple

Letícia Pollo de Oliveira

Ribeirão Preto
2013

UNIVERSIDADE DE SÃO PAULO

FACULDADE DE CIÊNCIAS FARMACÊUTICAS DE RIBEIRÃO PRETO

Neospora caninum:
estudo do secretoma e caracterização molecular
de três proteínas com domínios Apple

Tese de Doutorado apresentada ao Programa de Pós-Graduação em Biociências Aplicadas à Farmácia para obtenção do Título de Doutor em Ciências

Área de Concentração: Biociências
Aplicadas à Farmácia.

Orientada: Letícia Pollo de Oliveira
Orientadora: Profa. Dra. Ana Patrícia Yatsuda Natsui

Versão corrigida da Tese de Doutorado apresentada ao Programa de Pós-Graduação Biociências Aplicadas à Farmácia em 08/11/2013. A versão original encontra-se disponível na Faculdade de Ciências Farmacêuticas de Ribeirão Preto/USP.

Ribeirão Preto
2013

FICHA CATALOGRÁFICA

AUTORIZO A REPRODUÇÃO E DIVULGAÇÃO TOTAL OU PARCIAL DESTE TRABALHO, POR QUALQUER MEIO CONVENCIONAL OU ELETRÔNICO, PARA FINS DE ESTUDO E PESQUISA, DESDE QUE CITADA A FONTE.

Pollo-Oliveira, Letícia

Neospora caninum: estudo do secretoma e caracterização molecular de três proteínas com domínios Apple. Ribeirão Preto, 2013

249 p. : Il.; 30 cm

Tese de Doutorado, apresentada à Faculdade de Ciências Farmacêuticas de Ribeirão Preto/USP – Área de concentração: Biociências Aplicadas à Farmácia.

Orientadora: Yatsuda, Ana Patrícia

1.Neospora caninum. 2. Secretoma. 3. Domínios Apple

FOLHA DE APROVAÇÃO

Letícia Pollo de Oliveira

Neospora caninum: estudo do secretoma e caracterização molecular de três proteínas com domínios Apple

Tese de Doutorado apresentada ao Programa de Pós-Graduação em Biociências Aplicadas à Farmácia para obtenção do Título de Doutor em Ciências

Área de Concentração: Biociências Aplicadas à Farmácia.

Orientadora: Profa. Dra. Ana Patrícia Yatsuda Natsui

Aprovado em:

Banca Examinadora

Prof. Dr. _____

Instituição: _____ Assinatura: _____

Prof. Dr. _____

Instituição: _____ Assinatura: _____

Prof. Dr. _____

Instituição: _____ Assinatura: _____

Prof. Dr. _____

Instituição: _____ Assinatura: _____

Prof. Dr. _____

Instituição: _____ Assinatura: _____

Prof. Dr. _____

Instituição: _____ Assinatura: _____

Dedico este trabalho aos meus pais.

AGRADECIMENTOS

Aos meus pais, Leila e Geraldo, por uma vida inteira de apoio incondicional;

Aos meus irmãos, Rodrigo e Luísa, dispostos a ajudar em todos os momentos;

Ao Pedro, pelo essencial apoio, incentivo e companheirismo;

À minha avó, Clarice, que me acompanhou por todo esse período;

À Breca, que tantas vezes me levou para passear, espairecer e refletir sobre a vida;

À técnica do Lab. de Parasitologia Molecular da FCFRP-USP Maraísa Palhão Verri,
pelo grande auxílio e por tornar os dias mais leves;

À técnica do Lab. de Parasitologia Molecular da FCFRP-USP Tânia Mara, pelo
auxílio no início do projeto;

Ao doutorando Luiz Miguel Pereira, por gentilmente ceder suas ferramentas
moleculares, pelas discussões produtivas, e por compartilhar a chuva em todos os
congressos na praia;

À doutoranda Luciana Baroni, por toda a ajuda no laboratório, pelas discussões
científicas e filosóficas, e pela grande amizade que se estende para além da
bancada;

À aluna de iniciação científica Aline Bononi, por alegrar nossa vida no laboratório;

Aos coelhos, que produziram os anticorpos tão essenciais à execução deste
trabalho;

Aos técnicos da FCFRP-USP, Marlise, Rubinho e Alcides, pelo auxílio na imunização
dos coelhos;

Às técnicas da FMRP-USP, Elizabete Milani e Roberta Rosales, pelo auxílio na
captação de imagens com o microscópio confocal;

À CAPES e à Universidade de Utrecht pelas bolsas de doutorado e doutorado-
sanduíche, respectivamente;

À Prof. Dra. Maria Cristina Nonato, e aos seus alunos Ricardo de Pádua e Renata
Reis, pela colaboração para os estudos estruturais;

Ao Prof. Dr. Ney Lemke e ao Dr. Marcio Luis Acencio, do Instituto de Biociências da UNESP de Botucatu, pela colaboração para a construção da rede de interação de proteínas;

Ao técnico Harm Post do *Biomolecular Mass Spectrometry and Proteomics Group*, da Universidade de Utrecht, Holanda, por me auxiliar na execução de todos os experimentos lá desenvolvidos;

Ao Dr. Henk van den Toorn, pela colaboração na análise proteogenômica;

Ao doutorando Vinícius Tragante do Ó, do *University Medical Center Utrecht*, Holanda, pelo auxílio com ferramentas de bioinformática e, principalmente, pela amizade e grande acolhida da cidade de Utrecht;

Ao Dr. Maarten Altelaar e ao Prof. Albert Heck, pela oportunidade de desenvolver parte deste trabalho em seu grupo e pela supervisão nas análises do secretoma;

e, em especial,

À Profa. Dra. Ana Patrícia Yatsuda Natsui, pela acolhida em seu laboratório, pela orientação sempre muito próxima e presente, e por todo incentivo, confiança, compreensão, e amizade,

Meu sincero agradecimento.

*“O tempo só anda de ida·
A gente nasce, cresce, envelhece e morre·
Para não morrer eis a ciência da poesia:
Amarrar o Tempo no Poste!”*

Manoel de Barros

*E eu diria que essa é também a poesia da ciência·
Sigamos então amarrando o tempo no poste·*

SUMÁRIO

RESUMO	i
ABSTRACT	iii
RESUMEN	v
Lista de figuras	vii
Lista de tabelas	x
Lista de abreviaturas e siglas	xii
1. INTRODUÇÃO	
1.1 Revisão bibliográfica	1
1.1.2 <i>Neospora caninum</i>	1
1.1.3 Proteínas com domínios Apple	5
1.1.4 Secretoma	7
1.1.4.1 Secreção Ca^{2+} dependente	7
1.1.4.2 Identificação de proteínas por espectrometria de massas	9
1.2 Justificativa e trabalho desenvolvido	11
1.2.1 Caracterização de proteínas com domínio Apple	12
1.2.2 Estudo do secretoma	13
2. OBJETIVOS	
2.1 Objetivos gerais	14
2.2 Objetivos específicos	14
2.2.1 Caracterização de proteínas Apple de <i>N. caninum</i>	14
2.2.2 Secretoma de <i>N. caninum</i>	14
3. MATERIAL E MÉTODOS	
3.1 Técnicas em comum	15
3.1.1 Cultura <i>in vitro</i> e purificação de taquizoítas de <i>N. caninum</i>	15
3.1.2 Extração de DNA genômico de <i>N. caninum</i>	15
3.1.3 Extração de RNA total	16
3.1.4 Reação em cadeia de polimerase (PCR)	16
3.1.5 Ligação de fragmento de DNA (inserto) a plasmídeo	17
3.1.6 Preparo de células eletrocompetentes	17
3.1.7 Transformação de células eletrocompetentes	18
3.1.8 Purificação de plasmídeos selecionados (miniprep)	18
3.1.9 Eletroforese em gel de agarose	18
3.1.10 Quantificação de proteínas pelo método de Bradford	19
3.1.11 Eletroforese em gel de acrilamida	19
3.1.11.1 Eletroforese unidimensional (SDS-PAGE 1D)	19
3.1.11.2 Eletroforese bidimensional (SDS-PAGE 2D)	20
3.1.11.2.1 Primeira dimensão ou focalização isoeletrica	20
3.1.11.2.2 Segunda dimensão	21
3.1.12 Coloração de gel de acrilamida	22

3.1.12.1 Coloração por Coomassie G-250	22
3.1.12.2 Coloração por prata.....	22
3.1.13 Captação de imagens - géis de acrilamida e filmes de <i>Western blot</i>	23
3.1.14 Detecção de proteínas por <i>Western blot</i>	23
3.1.15 Preparo de amostras para análise por espectrometria de massas	24
3.2 Caracterização das proteínas com domínios Apple	26
3.2.1 Obtenção e análise das sequências genômicas	26
3.2.2 Amplificação e sequenciamento das sequências codificadoras de proteínas hipotéticas contendo domínios PAN/Apple	26
3.2.3 Predição da estrutura terciária dos domínios Apple.....	28
3.2.3.1 Alinhamento sequencial entre os domínios das proteínas MIC17 e o domínio A5 de <i>T. gondii</i>	29
3.2.3.2 Seleção de modelos estruturais de domínios Apple.....	29
3.2.3.3 Modelagem molecular com o software Modeller 9.11	30
3.2.3.4 Refinamento das estruturas geradas usando REFMAC/CCP4	30
3.2.3.5 Cálculo da superfície eletrostática das estruturas preditas	31
3.2.3.6 Predição de <i>folding</i> de cada proteína, reunindo os quatro domínios.....	31
3.2.4 Amplificação das sequências de MIC17B ₂₀₋₃₀₃ e MIC17C ₂₀₋₃₄₇	31
3.2.5 Clonagem de MIC17B ₂₀₋₃₀₃ e MIC17C ₂₀₋₃₄₇ em plasmídeo de expressão ..	33
3.2.6 Expressão das proteínas recombinantes MIC17B ₂₀₋₃₀₃ e MIC17C ₂₀₋₃₄₇	34
3.2.7 Purificação, dessalinização e concentração de proteínas recombinantes .	36
3.2.8 Confirmação da identidade das proteínas recombinantes por espectrometria de massas.....	37
3.2.9 Preparo de extrato total de <i>N. caninum</i> para produção de anticorpos	37
3.2.10 Produção de anticorpos contra MIC17B ₂₀₋₃₀₃ , MIC17C ₂₀₋₃₄₇ e NcET	37
3.2.11 Detecção de anticorpos por ELISA	38
3.2.12 Purificação de anticorpos	39
3.2.13 Detecção das proteínas recombinantes e nativas por <i>Western blot</i>	40
3.2.13.1 Reconhecimento das proteínas recombinantes pelos anticorpos purificados	40
3.2.13.2 Preparo dos antígenos: NcET e NcESA _{SN}	40
3.2.13.3 Detecção das proteínas nativas em NcET e NcESA _{SN} por WB 1D	41
3.2.13.4 Detecção das proteínas nativas em NcET e NcESA _{SN} por WB 2D	41
3.2.14 Identificação das proteínas nativas por espectrometria de massas	42
3.2.14.1 MS por MALDI, proteínas separadas por gel 1D e 2D em ET e ESA _{SN}	42
3.2.14.2 MS por <i>electrospray</i> , em ESA	43
3.2.15 Imunolocalização <i>in situ</i>	44
3.2.15.1 Taquizoítas em cultura de células Vero.....	44
3.2.15.2 Taquizoítas purificados.....	45
3.2.16 Ensaio de invasão	46
3.2.16.1 Com taquizoíta transfectado de forma transiente (NcLacZ transiente).	46
3.2.16.2 Com taquizoíta transfectado de forma estável (NcLacZ estável)	48
3.2.17 Ensaio de adesão (<i>binding</i>).....	49
3.2.17.1 Com proteína recombinante solúvel.....	49

3.2.17.2 Com ESA _{SN} por <i>binding</i> -ELISA (a), <i>binding</i> -Western blot (b), e <i>binding</i> -FACS (c).....	51
3.2.18 Deleção (nocaute) dos genes de MIC17A, MIC17B e MIC17C.....	55
3.2.18.1 Preparo de plasmídeo e transfecção de taquizoítas	55
3.2.18.2 Transfecção e Seleção com cloranfenicol	57
3.2.18.3 Confirmação de nocaute por PCR.....	58
3.3 Estudo do secretoma	59
3.3.1 Viabilidade de taquizoítas por citometria de fluxo	59
3.3.2 Análise de ESA e ET - por MALDI	59
3.3.2.1 Preparo de extrato secretado (ESA) e extrato total (ET) de taquizoítas	59
3.3.2.2 Eletroforese bidimensional (SDS-PAGE 2D) de ET e ESA	60
3.3.2.3 Identificação de proteínas por MS-MALDI.....	60
3.3.3 Análise de ESA e quantificação no taquizoíta <i>estimulado</i> - por LC-MS.....	60
3.3.3.1 Abordagem Via Extrato Secretado	62
3.3.3.1.1 Obtenção do extrato secretado (ESA) de taquizoítas de <i>N. caninum</i>	62
3.3.3.1.2 Digestão de proteínas	62
3.3.3.1.3 Análise por espectrometria de massas: RP-nanoLC-MS/MS	62
3.3.3.1.4 Identificação de proteínas	63
3.3.3.1.5 Predições <i>in silico</i> para secreção e classificação de proteínas	64
3.3.3.2 Abordagem Quantitativa.....	64
3.3.3.2.1 Preparo de ET de taquizoítas <i>estimulados</i> e taquizoítas <i>controle</i>	64
3.3.3.2.2 Digestão dos extratos e marcação com dimetil-isótopos.....	65
3.3.3.2.3 Fracionamento dos extratos por SCX.....	65
3.3.3.2.4 Espectrometria de massas por RP-nanoHPLC-MS/MS	66
3.3.3.2.5 Identificação e quantificação de proteínas	66
3.3.3.2.6 Classificação das proteínas quantificadas por localização e vias de secreção.....	67
3.3.4 Rede de interação de proteínas de <i>N. caninum</i>	67
3.3.5 Análise proteogenômica: Peptídeos em regiões não-annotadas.....	67

4. RESULTADOS

4.1 Proteínas com domínios Apple	69
4.1.1 Análise <i>in silico</i> das sequências com domínios Apple	69
4.1.2 Amplificação e sequenciamento dos genes com domínios Apple.....	71
4.1.3 Predição da estrutura terciária dos domínios Apple e das proteínas MIC17A, MIC17B e MIC17C	73
4.1.3.1 Alinhamento estrutural entre os domínios das proteínas MIC17 e o domínio A5 de <i>T. gondii</i>	73
4.1.3.2 Modelos de domínios Apple selecionados para a predição.....	73
4.1.3.3 Predição estrutural dos quatro domínios Apple de MIC17A, MIC17B e MIC17C	75
4.1.3.4 Análise da superfície eletrostática dos domínios Apple.....	79
4.1.3.5 Predição de <i>fold</i> ing de MIC17A, MIC17B e MIC17C, reunindo os quatro domínios Apple.....	81

4.1.4 Amplificação de MIC17B ₂₀₋₃₀₃ e MIC17C ₂₀₋₃₄₇ e clonagem em plasmídeo de expressão.....	82
4.1.5 Expressão heteróloga de MIC17B ₂₀₋₃₀₃ e MIC17C ₂₀₋₃₄₇ em <i>E. coli</i> BL21 ...	83
4.1.6 Confirmação da identidade das proteínas recombinantes por espectrometria de massas.....	86
4.1.7 Detecção por ELISA de anticorpos produzidos em coelhos	87
4.1.8 Purificação de anticorpos.....	88
4.1.9 Detecção das proteínas recombinantes e nativas por <i>Western blot</i>	89
4.1.10 Identificação das proteínas MIC17B e MIC17C nativas por espectrometria de massas.....	93
4.1.11 Imunolocalização de MIC17B e MIC17C em taquizoítas de <i>N. caninum</i> .	96
4.1.12 Ensaios de invasão	99
4.1.13 Ensaios de adesão (<i>binding</i>).....	103
4.1.14 Construção do plasmídeo pΔApple e seleção de taquizoítas de <i>N. caninum</i> nocaute para MIC17A, MIC17B e MIC17C	107
4.2 Secretoma de <i>N. caninum</i>	110
4.2.1 Viabilidade de taquizoítas	110
4.2.2 Identificação de proteínas de <i>N. caninum</i> separadas por géis bidimensionais e analisadas por MALDI	110
4.2.3 Proteínas identificadas no ESA de <i>N. caninum</i> por LC-MS	112
4.2.4 Quantificação relativa de proteínas no taquizoíta <i>estimulado</i>	118
4.2.5 Rede de interação entre proteínas de <i>N. caninum</i>	127
4.2.6 Análise proteogenômica: Peptídeos em regiões não-annotadas.....	129
 5. Discussão	
5.1 Caracterização das proteínas com domínio Apple.....	131
5.2 Secretoma de <i>N. caninum</i>	140
 6. CONCLUSÃO	
6.1 Proteínas com domínios Apple de <i>N. caninum</i>	152
6.2 Secretoma de <i>N. caninum</i>	153
 7. REFERÊNCIAS BIBLIOGRÁFICAS	154

APÊNDICES

Apêndice 1 Relação das 615 proteínas identificadas por LC-MS no ESA de taquizoítas de <i>N. caninum</i>	165
---	-----

Apêndice 2 Proteínas identificadas e quantificadas no extrato total de <i>N. caninum</i> submetido ao estímulo secretório por etanol (taquizoíta <i>estimulado</i>)	176
--	-----

RESUMO

Pollo-Oliveira, L. ***Neospora caninum*: estudo do secretoma e caracterização molecular de três proteínas com domínios Apple**. 2013. 249 f. Tese (doutorado). Faculdade de Ciências Farmacêuticas de Ribeirão Preto - Universidade de São Paulo, Ribeirão Preto, São Paulo, 2013.

Neospora caninum (filo Apicomplexa) é um parasita obrigatório intracelular como todos os membros deste filo, alguns reconhecidos por causarem doenças com impacto relevante na saúde humana (*Plasmodium* e *Toxoplasma*) e veterinária (*Babesia*, *Eimeria* e *Cryptosporidium*). Causador da neosporose, *N. caninum* vem emergindo como um dos maiores causadores de abortos infecciosos em bovinos, levando a consideráveis perdas econômicas na bovinocultura mundial. Devido à sua recente descoberta, o conhecimento sobre diversos processos bioquímicos de *N. caninum* ainda é limitado, demandando novas pesquisas para a compreensão de seus mecanismos de sobrevivência e consequente identificação de alvos para intervenção terapêutica. O processo de invasão celular é bastante investigado em pesquisas envolvendo apicomplexas, uma vez que a sobrevivência desses parasitas depende do sucesso de sua entrada na célula hospedeira. Proteínas secretadas de organelas filo-específicas (micronemas, roptrias e grânulos densos) estão intimamente envolvidas com a invasão celular. Elas são responsáveis pela interação inicial com a célula hospedeira, participam da junção de movimento formada no momento da invasão, e contribuem para a estabilização do vacúolo parasitóforo. Neste trabalho as proteínas secretadas por taquizoítas de *N. caninum* foram investigadas de duas formas: (1) por caracterização molecular de proteínas com domínio Apple; e (2) por estudo do secretoma do parasita. Os domínios proteicos do tipo Apple são caracterizados pela capacidade de interação proteína-proteína e proteína-carboidrato, e estão presentes em algumas proteínas micronêmicas com propriedades adesivas. Neste trabalho três proteínas de *N. caninum* contendo domínios Apple foram caracterizadas: MIC17A, MIC17B e MIC17C. A análise das sequências proteicas e das estruturas dos domínios Apple, obtidas por modelagem molecular, mostraram alta identidade sequencial e estrutural entre MIC17A e MIC17C. Apesar de ser paráloga às outras duas, MIC17B apresenta diferenças importantes em sua sequência e estrutura. Para MIC17B e MIC17C foram realizados experimentos de detecção das proteínas nativas nos extratos total e secretado do taquizoíta que sugerem diferentes formas de processamento entre essas proteínas no parasita. Para MIC17B foi confirmada a localização em micronemas, num padrão diferente do observado para MIC17C. Os ensaios de invasão combinados aos de localização indicam que estas proteínas estejam relacionadas ao processo de invasão celular, porém, suas funções permanecem desconhecidas. O secretoma é o conjunto de proteínas secretadas pelo parasita e, para explorar a composição deste extrato (ESA) no taquizoíta de *N. caninum*, duas abordagens complementares foram utilizadas. Na primeira abordagem foram identificadas as proteínas presentes no ESA por espectrometria de massas. Na segunda abordagem realizou-se uma

quantificação relativa das proteínas, marcadas por dois isótopos, nos extratos totais de taquizoítas submetidos ou não ao estímulo secretório. O resultado esperado seria com as proteínas secretadas diminuídas no parasita estimulado. Em ambas as abordagens foram utilizadas técnicas de espectrometria de massas de alta resolução (nanoLC-MS/MS), o que resultou num alto número de identificações; 615 proteínas no ESA e 2011 proteínas quantificadas. A comparação das duas abordagens permitiu o reconhecimento de proteínas com maior probabilidade de secreção. Uma rede de interação entre as proteínas diferencialmente expressas foi predita, gerando resultados que, associados às informações sobre as proteínas aumentadas, permitiram uma investigação sobre proteínas potencialmente envolvidas com a regulação do metabolismo relacionado à secreção. Os resultados obtidos por ambos os estudos aqui demonstrados somam conhecimento acerca do parasita *N. caninum* e demonstram ser úteis para guiar a busca e seleção de alvos a serem investigados para o desenvolvimento de terapêutica contra a neosporose.

Palavras-chave: *Neospora caninum*; Apicomplexa; domínios Apple; invasão celular; micronemas; roptrias; grânulos densos; secretoma; espectrometria de massas; nanoLC-MS/MS , rede de interação de proteínas.

ABSTRACT

Pollo-Oliveira, L. ***Neospora caninum*: study of the secretome and molecular characterization of three proteins containing Apple domains**. 2013. 249 f. Tese (doutorado). Faculdade de Ciências Farmacêuticas de Ribeirão Preto - Universidade de São Paulo, Ribeirão Preto, São Paulo, 2013.

Neospora caninum (Apicomplexa phylum) is an obligatory intracellular parasite like all members from this phylum, some causing diseases with relevant impact on human (*Plasmodium* and *Toxoplasma*) and veterinary (*Babesia*, *Eimeria* and *Cryptosporidium*) health. Causative agent of neosporosis, *N. caninum* has emerged as one of the leading causes of infectious abortion in cattle, generating huge economical losses in worldwide livestock. Due to its recent discovery, knowledge of *N. caninum* biochemical processes remains scarce, demanding new research for comprehending its survival mechanisms and, consequently, identifying new targets for therapeutic intervention. The invasion process has often been investigated in apicomplexans since their survival depends on the success of their entry into the host cell. Proteins secreted from phylum-specific organelles (micronemes, rhoptries and dense granules) are deeply involved with invasion. They are responsible for the initial interaction with the host cell; participate of the moving junction formed in the moment of invasion; and contribute for the stabilization of the parasitophorous vacuole. In this study, the proteins secreted by *N. caninum* tachyzoites were investigated in two ways: (1) the molecular characterization of Apple domain-containing proteins; and (2) exploring the parasite secretome. The Apple protein domains are characterized by the ability to interact as protein-protein and protein-carbohydrate, and are present in some microneme proteins with adhesive properties. Here three *N. caninum* proteins containing Apple domains were characterized: MIC17A, MIC17B and MIC17C. Analyses of the Apple domains sequences and structures, obtained by molecular modeling, revealed high sequential and structural identities between MIC17A and MIC17C. Although being a paralog of the other two proteins, MIC17B presents significant differences in its sequence and structure. Experiments were performed for native MIC17B and MIC17C detection in the total and secreted tachyzoite extracts, suggesting different processing forms for these proteins in the parasite. For MIC17B, the microneme localization was confirmed, differently from the pattern observed for MIC17C. Invasion and localization assays indicated that these proteins are related to the cell invasion process; nevertheless, their functions remain unknown. The secretome is the set of proteins secreted by the parasite and, to explore this extract (ESA) composition in *N. caninum*, two complementary approaches were used. Firstly proteins present in ESA were identified by mass spectrometry. In the second approach, a relative quantification was performed on the proteomes of ethanol stimulated/non stimulated tachyzoites, expecting that the secreted proteins would be down regulated at the stimulated parasite. Both approaches were performed with high resolution mass spectrometry techniques (nanoLC-MS/MS), reaching a high number of identifications: 615 proteins

in ESA and 2011 quantified proteins. The comparison between both approaches allowed the recognition of the most likely secreted proteins. An interaction network was predicted, involving the differentially expressed proteins. These results, associated with the information of up regulated proteins, allowed the investigation of proteins potentially involved with the secretion metabolism regulation. The findings from our two studies add up knowledge about *N. caninum* and demonstrate to be useful in guiding the search and selection for new targets for therapeutic development against neosporosis.

Keywords: *Neospora caninum*; Apicomplexa; Apple domains; cell invasion; micronemes; rhoptries; dense granules; secretome; mass spectrometry; nanoLC-MS/MS, proteins interaction network.

RESUMEN

Pollo-Oliveira, L. ***Neospora caninum*: estudio de secretoma y la caracterización molecular de tres proteínas con dominios de Apple**. 2013. 249 f. Tese (doutorado). Faculdade de Ciências Farmacêuticas de Ribeirão Preto - Universidade de São Paulo, Ribeirão Preto, São Paulo, 2013.

Neospora caninum (phylum Apicomplexa) es un parásito intracelular obligado, como todos los miembros de este filo, algunos conocidos por causar enfermedades de gran impacto significativo en la salud humana (*Plasmodium* y *Toxoplasma*) y veterinaria (*Babesia*, *Eimeria* y *Cryptosporidium*). El agente causal de la neosporosis, *N. caninum*, ha surgido como una de las principales causas del aborto infeccioso en el ganado bovino, lo que conduce a considerables pérdidas económicas en el mercado ganadero en todo el mundo. Debido a su reciente descubrimiento, el conocimiento de los procesos bioquímicos de *N. caninum* es aún limitado, lo que requiere más investigación para la comprensión de sus mecanismos de supervivencia y posterior identificación de dianas para la intervención terapéutica. El proceso de invasión ha sido frecuentemente investigado en organismos apicomplexas, ya que la supervivencia de estos parásitos depende del éxito de su entrada en la célula huésped. Las proteínas secretadas por organelos filo específicos (micronemas, roptrias y gránulos densos), están íntimamente involucradas en el proceso de invasión celular. Estas proteínas también son responsables de la interacción inicial con la célula huésped, participan en la formación de la unión móvil al momento de la invasión y contribuyen a la estabilización de la vacuola parasitófora. En este estudio, las proteínas secretadas por los taquizoitos de *N. caninum* fueron investigadas de dos maneras: (1) Por caracterización molecular de las proteínas que contienen el dominio Apple, y (2) por el estudio del secretoma del parásito. Los dominios Apple de las proteínas se caracterizan por sus habilidades de interacción proteína-proteína y proteína-carbohidrato y se encuentran presentes en algunas proteínas del micronema, las cuales presentan propiedades adhesivas. En este estudio, se caracterizaron tres proteínas de *N. caninum* que contienen el dominio Apple, MIC17A, MIC17B y MIC17C. El análisis de las secuencias y estructuras de los dominios Apple, obtenidos por modelado molecular, reveló grandes semejanzas en secuencia y estructura entre MIC17A y MIC17C. Aunque paráloga a las otras dos, MIC17B presenta importantes diferencias en estos aspectos (secuencia y estructura). Se llevaron a cabo experimentos para detectar las proteínas MIC17B y MIC17C nativas, esto en extractos totales y secretados del taquizoito. Dichos experimentos sugieren diferentes formas de procesamiento para estas proteínas en el parásito. Para MIC17B, fue confirmada la localización en el micronema, lo cual resultó en un patrón diferente del observado para MIC17C. Los ensayos de invasión y ubicación indican que estas proteínas están relacionadas con el proceso de invasión celular, pero sus funciones continúan siendo desconocidas. El secretoma es el conjunto de proteínas secretadas por el parásito. Para explorar la composición de este extracto (ESA) en el

taquizoíto de *N.caninum*, se utilizaron dos enfoques complementarios. En el primer enfoque se identificaron proteínas presentes en el ESA por espectrometría de masas. En el segundo enfoque se realizó una cuantificación relativa de proteínas en los proteomas de taquizoitos estimulados y no estimulados con etanol. El resultado esperado sería una disminución de las proteínas secretadas en el parásito estimulado. Ambos métodos fueron llevados a cabo con técnicas de espectrometría de masas de alta resolución (nanoLC-MS/MS), con lo que se logró un gran número de identificaciones: 615 en ESA y 2.011 de proteína cuantificada. Una comparación de los dos enfoques permitió el reconocimiento de proteínas con una mayor probabilidad de ser secretadas. Así, se predice una red de interacciones entre las proteínas expresadas diferencialmente. Estos resultados, asociados a la información de las proteínas expresadas en alta proporción, permitieron la investigación de proteínas potencialmente involucradas en la regulación del metabolismo de secreción. Los resultados obtenidos en ambos estudios aquí descritos, se suman al conocimiento de *N. caninum*, demostrando además su utilidad para orientar la búsqueda y selección de nuevos blancos terapéuticos frente a la neosporosis.

Palabras clave: *Neospora caninum*; Apicomplexa, dominios Apple; invasión celular; micronemas, roptrias, gránulos densos; secretoma, espectrometría de masas; nanoLC-MS/MS, red de interacción de proteínas.

LISTA DE FIGURAS

Figura 1	Transmissão da neosporose bovina.....	2
Figura 2.	Representação esquemática de um taquizoíta de <i>Neospora caninum</i>	3
Figura 3.	Representação do processo de invasão por taquizoítas de <i>Toxoplasma gondii</i>	4
Figura 4.	Representação esquemática de proteínas micronêmicas e seus respectivos domínios em quatro gêneros do filo Apicomplexa.....	5
Figura 5.	Representação estrutural do domínio Apple A9 da proteína MIC5 de <i>Eimeria tenella</i>	7
Figura 6.	Vias de sinalização de secreção e motilidade afetadas por aumento de Ca^{2+} intracelular em <i>Toxoplasma gondii</i>	8
Figura 7.	Exemplo de árvore de decisão usada na escolha do método de fragmentação de peptídeos em espectrômetro de massas.....	9
Figura 8.	Marcação de peptídeos para quantificação por dimetil-isótopos.....	10
Figura 9.	Quantificação relativa usando marcação com dimetil-isótopos (duplex)	11
Figura 10.	Sequências de aminoácidos dos modelos selecionados para a predição das estruturas dos domínios Apple das proteínas MIC17A, MIC17B e MIC17C.....	30
Figura 11.	Metodologia para expressão heteróloga e extração de MIC17B ₂₀₋₃₀₃ e MIC17C ₂₀₋₃₄₇	35
Figura 12.	Disposição dos genes no plasmídeo NcRPS13tetO/LacZ usado para transfecção transiente de taquizoítas utilizados em ensaios de invasão	47
Figura 13.	Construção do plasmídeo NcDHFR-DHFRM2M3/NcTub-tetO/Lac-Z, utilizado para a inserção do gene da β -galactosidase no genoma de <i>N. caninum</i>	48
Figura 14.	Esquema do ensaio de adesão por <i>Western blot</i>	53
Figura 15.	Construção do plasmídeo p Δ Apple.....	55
Figura 16.	Desenho experimental para o estudo do secretoma de <i>N. caninum</i> .	61
Figura 17.	Organização dos genes de MIC17A, MIC17B e MIC17C no genoma de <i>N. caninum</i> (isolado Liverpool).....	69
Figura 18.	Alinhamento entre as sequências de aminoácidos das proteínas hipotéticas NCLIV_038120 (MIC17A), NCLIV_038110 (MIC17B), NCLIV_038100 (MIC17C) e NCLIV_068830.....	70
Figura 19.	Organização das regiões codificadoras de MIC17A, MIC17B e MIC17C e de suas regiões adjacentes não traduzidas no cromossomo VIII de <i>N. caninum</i>	71
Figura 20.	Amplificação dos genes que codificam as proteínas Apple por PCR e clonagem em pGEM.....	72
Figura 21.	Alinhamento estrutural dos domínios Apple A1, A2, A3 e A4 de MIC17A, MIC17B e MIC17C com a sequência 2LL3.....	73
Figura 22.	RMSD por resíduo de aminoácido dos domínios Apple com base no alinhamento estrutural dos modelos 2YIL, 2F83_2 e 2F83_3 com o modelo 2LL3.....	75
Figura 23.	Domínio a3m17b obtido por modelagem molecular.....	76
Figura 24.	RMSD por resíduo de aminoácido, com base no alinhamento estrutural dos onze domínios Apple das proteínas de <i>N. caninum</i> MIC17A, MIC17B e MIC17C com o domínio a1m17a.....	77
Figura 25.	Domínios A2 das proteínas Apple MIC17A, MIC17B e MIC17C gerados por predição estrutural com o <i>software</i> Modeller 9.11.....	78

Figura 26. Superfície eletrostática dos domínios Apple das proteínas MIC17A, MIC17B e MIC17C de <i>N. caninum</i>	80
Figura 27. Tetrâmero formado por quatro domínios Apple que compõem o fator de coagulação XI (PDB 2F83).....	81
Figura 28. Amplificação e clonagem de MIC17B ₂₀₋₃₀₃ e MIC17C ₂₀₋₃₄₇ em pGEM e pET28.....	83
Figura 29. Sequências de aminoácidos das proteínas recombinantes MIC17B ₂₀₋₃₀₃ e MIC17C ₂₀₋₃₄₇	84
Figura 30. Extração de proteínas recombinantes após indução de <i>E. coli</i> BL21.....	85
Figura 31. Proteínas recombinantes purificadas, dessalinizadas e concentradas para imunização.....	86
Figura 32. Produção de anticorpos (IgG) contra as proteínas recombinantes MIC17B ₂₀₋₃₀₃ , MIC17C ₂₀₋₃₄₇ ou contra o extrato total de <i>N. caninum</i> (NcET).....	88
Figura 33. Membranas com proteínas recombinantes usadas na purificação de anticorpos.....	89
Figura 34. Detecção das proteínas recombinantes e nativas por <i>Western blot</i> 1D.....	90
Figura 35. <i>Western blots</i> bidimensionais (WB 2D) para a detecção das proteínas nativas MIC17B e MIC17C.....	92
Figura 36. Espectros de fragmentação (MS/MS) dos três peptídeos usados na identificação das proteínas MIC17 no extrato secretado (ESA) de <i>N. caninum</i>	95
Figura 37. MIC17B e MIC17C localizadas em taquizoítas de <i>N. caninum</i> fixados em cultura de células Vero.....	96
Figura 38. MIC17B em taquizoíta de <i>N. caninum</i> em possível fase inicial de invasão celular.....	97
Figura 39. MIC17B e MIC17C nativas em taquizoítas de <i>N. caninum</i> purificados.....	98
Figura 40. Efeito dos anticorpos e dos soros anti-MIC17B ₂₀₋₃₀₃ e anti-MIC17C ₂₀₋₃₄₇ , e do soro anti-NcET sobre a invasão celular por taquizoítas de <i>N. caninum</i>	100
Figura 41. Lise de taquizoítas durante o ensaio de invasão.....	101
Figura 42. Efeito dos anticorpos anti-MIC17B ₂₀₋₃₀₃ e anti-MIC17C ₂₀₋₃₄₇ sobre a invasão por taquizoítas de <i>N. caninum</i>	102
Figura 43. Ensaio de adesão das proteínas MIC17B e MIC17C, presentes em extrato secretado de <i>N. caninum</i> (NcESA _{SN}), às células Vero, por <i>binding</i> -ELISA.....	104
Figura 44. Ensaio de adesão em células Vero das proteínas MIC17B e MIC17C presentes em extrato secretado de <i>N. caninum</i> (NcESA _{SN}) pela técnica de <i>Western blot</i>	105
Figura 45. Ensaio de adesão a células Vero pelas proteínas MIC17B e MIC17C presentes em extrato secretado de <i>N. caninum</i> (NcESA _{SN}), pela técnica de <i>binding</i> -FACS.....	106
Figura 46. Etapas de construção do plasmídeo pΔApple para transfecção em <i>N. caninum</i> e nocaute dos genes de MIC17A, MIC17B e MIC17C.....	107
Figura 47. Conferência por PCR do nocaute das proteínas Apple pelo plasmídeo pΔApple.....	109
Figura 48. Avaliação por citometria de fluxo da viabilidade de taquizoítas de <i>N. caninum</i> recém-purificados.....	110
Figura 49. Géis bidimensionais dos extratos total (ET) e secretado (ESA) de taquizoítas de <i>N. caninum</i>	111

Figura 50. Extrato secretado (ESA), obtido por estímulo secretório com 1% de etanol, separado por SDS-PAGE 1D.....	113
Figura 51. Localização predita de proteínas no extrato secretado (ESA) de taquizoítas de <i>N. caninum</i> , de um total de 615 proteínas identificadas por LC-MS (em porcentagem)	113
Figura 52. Cromatograma do fracionamento por cromatografia de troca iônica (SCX) dos peptídeos marcados.....	118
Figura 53. Distribuição de proteínas no extrato total (ET) de <i>N. caninum</i> após estímulo secretório	124
Figura 54. Visualização no programa Cytoscape das redes de interação entre proteínas quantificadas no taquizoíta de <i>N. caninum</i> após estímulo secretório.....	128
Figura 55. Visualização no programa IGV (Integrative Genomics Viewer) de peptídeos identificados em regiões do genoma de <i>N. caninum</i> não anotadas como codificadoras de genes	129
Figura 56. Reconstrução filogenética das sequências das proteínas Apple MIC17 de <i>N. caninum</i> e <i>T. gondii</i>	133
Figura 57. Alinhamento sequencial das proteínas MIC17A, MIC17B e MIC17C de <i>N. caninum</i> (Nc) e <i>T. gondii</i> (Tg).....	134
Figura 58. Superfície eletrostática do domínio Apple A5 da proteína MIC4 de <i>Toxoplasma gondii</i> ligado ao dissacarídeo lacto-N-biose.....	135
Figura 59. Alinhamento de proteínas Apple de <i>N. caninum</i> e <i>T. gondii</i> (primeiros 81 aminoácidos)	137
Figura 60. Representação esquemática das identificações obtidas pelas duas abordagens do Secretoma de <i>N. caninum</i> : Extrato Secretado (ESA) e Quantificação Relativa (QUANT) sobre os extratos totais (ET) dos taquizoítas <i>estimulado</i> e <i>controle</i>	142
Figura 61. Metabolismo de inositol fosfato, adaptado de LAMP (<i>Library of Apicomplexan Metabolic Pathways</i>)	146
Figura 62. Metabolismo de purinas, recorte da via mostrando adenosina cinase diminuída e AMP deaminase aumentada	147
Figura 63. Metabolismo de purinas, recorte da via mostrando GMO sintase e guanilato cinase aumentadas	148

LISTA DE TABELAS

Tabela 1. Programas utilizados para a focalização isoeletrica de <i>strips</i>	21
Tabela 2. Sequência dos <i>primers</i> utilizados para amplificação de regiões codificadoras de proteínas contendo domínios PAN/Apple.....	27
Tabela 3. Condições de reação da polimerase em cadeia (PCR) para amplificação das sequências codificadoras de proteínas contendo domínios PAN/Apple	27
Tabela 4. <i>Primers</i> para sequenciamento dos fragmentos de MIC17A, MIC17B e MIC17C inseridos em plasmídeo pGEM	28
Tabela 5. Modelos selecionados para a predição das estruturas dos domínios Apple das proteínas MIC17A, MIC17B e MIC17C por modelagem molecular baseada em homologia.....	29
Tabela 6. <i>Primers</i> utilizados para a amplificação de fragmentos das proteínas MIC17B e MIC17C a serem expressados	32
Tabela 7. Locais de análise, por espectrometria de massas, de proteínas presentes nos extratos total (ET) e secretado (ESA _{SN3}), de <i>N. caninum</i> (Nc) e de células Vero (Ve), separados por eletroforese uni e bidimensional	43
Tabela 8. <i>Primers</i> utilizados para a amplificação de fragmentos a serem expressados em pET28-SUMO.....	50
Tabela 9. Condições variadas para ensaio de adesão por <i>binding</i> -ELISA.....	52
Tabela 10. Sequência dos <i>primers</i> utilizados na amplificação dos fragmentos PROM_M17A e 3'UTR_M17C para inserção em plasmídeo pΔApple	57
Tabela 11. <i>Primers</i> utilizados nas PCRs para a amplificação de MIC17B e das construções PROM_M17A-CAT e CAT-3'UTR para conferência de nocaute sobre DNA de taquizoíta NcLacZ transfectado com pΔApple	58
Tabela 12. Porcentagens de identidade/similaridade (%) entre modelos de domínios Apple selecionados para predição estrutural de MIC17A, MIC17B e MIC17C	74
Tabela 13. RMSD de Cα (Å) entre modelos de domínios Apple selecionados para predição estrutural de MIC17A, MIC17B e MIC17C	74
Tabela 14. Número de resíduos de aminoácidos nos intervalos entre os domínios Apple de algumas proteínas de organismos apicomplexas e do fator de coagulação XI.....	82
Tabela 15. Peptídeos das proteínas recombinantes NcMIC17B ₂₀₋₃₀₃ e NcMIC17C ₂₀₋₃₄₇ identificados por espectrometria de massas (MS).	87
Tabela 16. Descrição de peso molecular e ponto isoeletrico (pI) das proteínas detectadas por <i>Western blot</i> bidimensional com os anticorpos purificados anti-MIC17B ₂₀₋₃₀₃ e anti-MIC17C ₂₀₋₃₄₇	93
Tabela 17. Peptídeos identificados no extrato secretado de <i>N. caninum</i> (ESA) usados na identificação das proteínas MIC17B e MIC17C	94

Tabela 18. Identificação de proteínas presentes no extrato total (ET) e extrato secretado (ESA) do taquizoíta de <i>N. caninum</i>	112
Tabela 19. Proteínas identificadas por LC-MS no ESA de <i>N. caninum</i> com localização relacionada à secreção.....	114
Tabela 20. Proteínas identificadas por LC-MS no ESA de <i>N. caninum</i> com localização não relacionada à secreção.....	116
Tabela 21. Ciclofilinas identificadas no ESA de <i>N. caninum</i>	117
Tabela 22. Proteínas aumentadas no taquizoíta de <i>N. caninum</i> submetido ao estímulo secretório por etanol (<i>estimulado</i>).....	119
Tabela 23. Proteínas diminuídas no taquizoíta de <i>N. caninum</i> submetido ao estímulo secretório por etanol (<i>estimulado</i>).....	121
Tabela 24. Proteínas de organelas secretórias (micronemas, roptrias e grânulos densos) e de superfície identificadas no extrato secretado (ESA) e/ou quantificadas no extrato total (ET) de <i>N. caninum</i>	125
Tabela 25. Peptídeos em regiões não anotadas como codificadoras	130
Tabela 26. Relação de identidade e similaridade entre as proteínas Apple de <i>Neospora caninum</i> e <i>Toxoplasma gondii</i> , baseada nas sequências de aminoácidos	132
Tabela 27. Proteínas quantificadas no taquizoíta <i>estimulado</i> de <i>N. caninum</i> potencialmente envolvidas com invasão	149

LISTA DE ABREVIATURAS E SIGLAS

1D	Unidimensional
2D	Bidimensional
ACN	Acetonitrila
AM	Tampão de amostra
AMA1	Antígeno de membrana apical 1 (<i>apical membrane antigen 1</i>)
ATP	Adenosina trifosfato
BPK1	Pseudocinase 1 de bradizoíta (<i>bradyzoite pseudokinase 1</i>)
BSA	Albumina sérica bovina
cAMP	Adenosina monofosfato cíclica
CAT	Cloranfenicol acetil transferase
CCR5	CC- <i>chemokine receptor 5</i>
cDNA	DNA complementar
cGMP	Guanosina monofosfato cíclica
CHAPS	(3-[(3-Colamidopropil)-dimetil amonio]-propano-sulfonato)
CID	Dissociação induzida por colisão
CLP1	<i>Chitin-like protein 1</i>
CPRG	Clorofenolred- β -D-galactopiranoside
DHFR-TS	Diidrofolato redutase - timidilato sintase
DNA	Ácido desoxirribonucleico
dNTP	Desoxirribonucleotídeo
DTT	Ditiotreitol
dUDP	Deoxiuridina difosfato
EDTA	Ácido etilenodiamino tetra-acético
EGF	Epidermal growth factor
ELISA	<i>Enzyme-linked immunosorbent assay</i>
ESA	Extrato secretado
ESA _{SN}	Extrato secretado a partir do sobrenadante de cultura
EST	<i>Expressed sequence tag</i>
ET	Extrato total
ETD	Dissociação por transferência de elétrons
FACS	<i>Fluorescence-activated cell sorting</i>
FITC	Isotiocianato de fluoresceína
FDR	Taxa de detecção falsa (<i>false discovery rate</i>)
FT	<i>Fourier transform ion cyclotron resonance</i>
FXI	Fator de coagulação XI
GRA	Proteína de grânulo denso
HCD	Dissociação por colisão de alta energia
IMP	Inosina monofosfato
IPTG	Isopropiltiogalactosídeo
LacZ	Gene para β -galactosidase
LC	Cromatografia líquida
LC-MS	Cromatografia líquida acoplada à espectrometria de massas

LTD	<i>Linear ion trap quadrupole</i>
M2AP	Proteína associada à MIC2
MALDI	Ionização e Dessorção a Laser Assistida por Matriz
MIC	Proteína micronêmica
MJ	Junção de movimento (<i>moving junction</i>)
MS	Espectrometria de massas (<i>mass spectrometry</i>)
MS/MS	Espectrometria de massas em tandem
Nc	<i>Neospora caninum</i>
Nc-1	<i>Neospora caninum</i> linhagem 1
NL	Não linear
PBS	Tampão fosfato salino
PBS-T	PBS adicionado de Tween ou Triton X-100
PBS-GT	PBS-T adicionado de gelatina suína
PCR	Reação em cadeia de polimerase
PDB	RCSB <i>Protein Data Bank</i>
pET28	Plasmídeo de expressão
pGEM	Plasmídeo de clonagem
pI	Ponto isoelétrico
PROM	Região promotora
PS	Peptídeo sinal
PSM	<i>Peptide spectrum match</i>
pΔApple	Plasmídeo para deleção de genes Apple
RMSD	Desvio médio quadrático
RNA	Ácido ribonucleico
RNaseH	Ácido ribonuclease H
RON	Proteína de pescoço de roptrias
ROP	Proteína de roptrias
RP	Fase reversa (<i>reversed phase</i>)
RPS13	Proteína ribossomal S13
SAG	Proteína de superfície (<i>surface antigen</i>)
SCX	Cromatografia líquida de troca catiônica (<i>strong-cation exchange</i>)
SDS	Dodecil sulfato de sódio
SDS-PAGE	Eletroforese em gel de acrilamida com dodecil sulfato de sódio
Ser	Serina
SFB	Soro fetal bovino
SPATR	<i>Sporozoite protein with an altered thrombospondin repeat</i>
SRS	Proteína de superfície (<i>surface antigen-related sequences</i>)
SUB	<i>Subtilisin-like protease</i>
SUMO	<i>Small ubiquitin-like modifier</i>
TAE	Tampão Tris acetato-EDTA
TE	Tris EDTA
tetO	Operador de tetraciclina
Thr	Treonina
TM	Transmembrana

TLN	Toxolisina
TOF	Tempo de voo (<i>time of flight</i>)
TOP10	Células de <i>E. coli</i> eletrocompetentes para clonagem
Tris	2-Amino-2-hidroximetil-propano-1,3-diol
Tub	Promotor de tubulina
Tyr	Tirosina
UTR	Região não traduzida (<i>untranslated region</i>)
Ve	Célula Vero
VP	Vacúolo parasitóforo
VWFC	Fator de von Willebrand tipo C
WB	Western Blot
X-Gal	5-bromo-4-cloro-3-indol- β -D-galactopiranosida
YFP	Proteína fluorescente amarela (<i>Yellow fluorescent protein</i>)
Δ	Nocaute

1. INTRODUÇÃO

1.1 Revisão bibliográfica

1.1.2 *Neospora caninum*

Neospora caninum é um protozoário que pertence à classe Coccidia do filo Apicomplexa. Parasitas deste filo são relacionados a doenças que têm impacto relevante tanto na saúde humana, causadas por parasitas como *Plasmodium* e *Toxoplasma*, quanto na veterinária, por exemplo por *Babesia*, *Eimeria* e *Cryptosporidium*.

Em 1984, *N. caninum* foi reconhecido pela primeira vez como um suposto novo parasita, presente na forma de cistos em animais com desordens neurológicas¹. Apenas em 1988 *N. caninum* foi descrito como novo gênero e espécie, diferenciado do parasita *Toxoplasma gondii*, como era até então erroneamente diagnosticado². Apesar das semelhanças morfológicas e genéticas, *N. caninum* e *T. gondii* diferem quanto aos hospedeiros definitivos, composição de carboidratos na superfície, antigenicidade e características estruturais³. Ademais, diferentemente de *T. gondii*, *N. caninum* não causa infecção em humanos⁴.

Desde sua descoberta, *N. caninum* tem emergido como um importante causador de abortos e mortalidade neonatal em bovinos. A neosporose bovina já foi relatada em países de todos os continentes, exceto na Antártica⁵, e um cálculo com os dados de 10 países comprovadamente afetados (Argentina, Austrália, Brasil, Canadá, México, Holanda, Nova Zelândia, Espanha, Reino Unido e Estados Unidos) estimou uma perda anual média de US\$ 1,298 bilhão, com aproximadamente dois terços (US\$ 842,9 milhões) afetando o gado de leite⁶.

A alta eficiência de transmissão transplacentária de *N. caninum* em seus hospedeiros intermediários pode resultar em aborto, mortalidade neonatal, ou sobrevivência do filhote que, quando fêmea, poderá perpetuar a infecção ao se reproduzir⁷. Em seu ciclo de vida heteroxênico, *N. caninum* tem como hospedeiros definitivos cães (*Canis familiares*) e coiotes (*Canis latrans*)^{8,9,10}, dingos (*Canis lupus dingo*)¹¹ e lobos cinza (*Canis lupus*)¹².

Após a ingestão de tecidos de animais infectados com cistos de *N. caninum* (figura 1), os hospedeiros definitivos liberam oocistos em suas fezes, que podem contaminar água e alimentos que serão ingeridos pelos hospedeiros intermediários.

Uma variedade de animais de sangue quente, entre domésticos e selvagens, já foi descrita como hospedeiros intermediários, porém, os mais comuns e de maior importância econômica são os bovinos^{7,13}. Diferente da transmissão horizontal (exógena), em que a fêmea grávida ingere os oocistos do hospedeiro definitivo e então transmite a infecção para o feto, a transmissão vertical, da fêmea previamente infectada para seu feto (endógena), é mais frequente e representa um meio de propagação da doença por sucessivas gerações^{12,13}. No hospedeiro intermediário são encontrados os estágios infectantes taquizoíta e bradizoíta do parasita. O taquizoíta se multiplica rapidamente e corresponde à forma invasiva do parasita. Ele pode se diferenciar em bradizoíta, que tem multiplicação lenta e representa a forma quiescente do parasita, contida em cistos. Durante a gravidez, os bradizoítas são reativados e diferenciados novamente em taquizoítas, que se espalham por vários tecidos, incluindo o útero, quando cruzam a placenta e infectam o feto⁷.

Transmissão da neosporose bovina

Fig. 1 Transmissão da neosporose bovina. Oocistos são liberados nas fezes do hospedeiro definitivo (cão) e ingeridos por uma vaca prenha susceptível, o que leva à infecção do feto (transmissão transplacentária exógena). O feto infectado pode ser abortado ou sobreviver e permanecer infectado quando adulto, que, no caso da fêmea, pode transmitir a infecção no momento de sua reprodução (transmissão transplacentária endógena). (Adaptado de Dubey *et al* 2006¹³).

N. caninum, assim como todos os outros apicomplexas, é um parasita intracelular obrigatório, e depende do sucesso no processo de invasão celular para sobreviver. O taquizoíta (figura 2) tem capacidade de infectar uma grande variedade de tipos

celulares^{14,15}. Ele penetra na célula hospedeira de maneira ativa e conservada entre os parasitas do filo, num processo que envolve a secreção de proteínas de organelas filo-específicas, chamadas micronemas, roptrias e grânulos densos, e a ativação do sistema motor do parasita, constituído por um sistema actina/miosina. Na porção apical do taquizoíta encontra-se o conóide, região de interação inicial com a célula hospedeira no momento da invasão. Nesta região apical situam-se as organelas micronemas e roptrias e, distribuídas ao longo do taquizoíta, encontram-se os grânulos densos. As proteínas secretadas por estas organelas têm papel essencial na invasão, envolvidas com a adesão apical, formação da junção de movimento, motilidade deslizante, e formação e estabilização do vacúolo parasitóforo (VP)^{15,16,17}.

Fig. 2 Representação esquemática de um taquizoíta de *Neospora caninum*. O parasita é formado por duas membranas, a membrana plasmática e a membrana interna. Na região apical do parasita encontram-se a estrutura conóide e as organelas secretórias micronemas e roptrias. Distribuídas ao longo do parasita são encontradas as organelas secretórias grânulos densos. Também estão presentes mitocôndria, apicoplasto, complexo de Golgi, retículo endoplasmático, e o núcleo. (Adaptado de Joiner & Roos, 2002¹⁸).

O processo de invasão é caracterizado por contato e interação inicial de baixa afinidade entre o parasita e a célula hospedeira, via proteínas de superfície (SAG ou SRS) do parasita (fig. 3 passo 1). Em seguida, ocorre a reorientação do parasita e a formação de uma ligação de alta afinidade entre a porção apical do parasita e a membrana da célula hospedeira (junção de movimento), com participação de complexos formados por proteínas micronêmicas (MIC) e de roptrias (ROP e RON) (fig. 3 passo 2). O parasita então invade a célula de forma ativa, envolvendo-se na membrana da célula hospedeira, por um mecanismo chamado motilidade por deslizamento (fig. 3 passo 3). Este processo necessita do sistema motor proteico “*glideosome*” situado entre as membranas plasmática e interna do parasita (IMC). Ao final do processo ocorre a clivagem de proteínas micronêmicas ainda ligadas à

membrana da célula hospedeira, e a formação do VP, com participação das proteínas de grânulos densos (GRA) para a sua estabilização (fig. 3 passo 4). Após a multiplicação de taquizoítas, o VP é rompido e o egresso dos parasitas da célula também é dependente das proteínas secretadas e do sistema motor^{15,16,19}.

Fig. 3 Representação do processo de invasão por taquizoítas de *Toxoplasma gondii*. Na porção superior esquerda da figura é representada a região extracelular da célula hospedeira, cuja membrana plasmática é mostrada em cinza. Na porção inferior da figura é mostrado o interior da célula hospedeira, em azul. **1**, contato e interação inicial de baixa afinidade entre o parasita (verde) e a célula hospedeira, via proteínas de superfície. **2**, reorientação do parasita e formação da junção de movimento (anel vermelho) entre a porção apical do parasita e a membrana da célula hospedeira, com participação de complexos formados por proteínas micronêmicas e de roptrias. **3**, invasão de forma ativa por motilidade de deslizamento, em que o parasita envolve-se na membrana plasmática da célula hospedeira. **4**, formação do vacúolo parasitóforo e estabilização com participação de proteínas de grânulos densos. Além da invasão, estão representados na figura a multiplicação de taquizoítas no interior do vacúolo e o egresso da célula. (Ilustração adaptada de Billker *et al*, 2009²⁰)

No presente trabalho, dois estudos foram propostos para a investigação de proteínas envolvidas com o processo de invasão pelo parasita *Neospora caninum*; um para a identificação e caracterização de proteínas compostas por domínios Apple, potencialmente micronêmicas; e outro para a investigação do conjunto de proteínas secretadas pelo parasita, chamado secretoma.

1.1.3 Proteínas com domínios Apple

Sabe-se que as proteínas micronêmicas são secretadas rapidamente assim que o parasita faz contato com a célula hospedeira; são alvo de inibição por anticorpos, como relatado em *Plasmodium*, bloqueando a invasão pelo parasita; e que seus genes são essenciais para efetiva motilidade, migração e invasão do parasita²¹.

Dentre as proteínas micronêmicas são encontrados domínios estruturais conservados (figura 4) que apresentam homologia com conhecidos domínios adesivos de eucariotas superiores²². Estes domínios estão frequentemente envolvidos na formação de complexos, muitas vezes entre as próprias proteínas micronêmicas, e geralmente envolvendo proteínas solúveis ligadas a proteínas transmembrana, resultando na formação de uma ponte entre o parasita e a célula hospedeira²¹. Um entre estes domínios, conhecido como domínio Apple, foi descrito como componente de diferentes proteínas micronêmicas, destacadas por setas na figura 4.

Domínios presentes em proteínas micronêmicas

Fig. 4 Representação esquemática de proteínas micronêmicas e seus respectivos domínios em quatro gêneros do filo Apicomplexa: *Eimeria*, *Cryptosporidium*, *Plasmodium* e *Toxoplasma*. Os símbolos que representam cada domínio são mostrados na lista à direita da figura. Proteínas que

possuem domínios Apple estão indicadas por setas. Entre elas, MIC5 de *Eimeria* e MIC4 de *Toxoplasma* são exclusivamente formadas por estes domínios. (Adaptada de Carruthers & Tomley, 2008²¹)

O domínio Apple (PAN1) é um membro da superfamília PAN, cujos módulos estão distribuídos em proteínas de diferentes organismos, como as proteínas do plasma fator de coagulação XI (FXI) e plasminogênio, proteínas de nematodos, e também de protozoários^{23,24}. Uma característica comum a estas proteínas é sua propriedade adesiva, por interação com outras proteínas ou com carboidratos^{23,25,26,27}. Dentro do filo Apicomplexa, algumas proteínas contendo domínios Apple já foram descritas. As proteínas SML1, SML2 e SML3, de *Sarcocystis muris*, são lectinas com especificidade para galactose^{28,29,30}. Em *Eimeria tenella*, EtMIC5 é uma proteína micronêmica solúvel composta por 11 domínios Apple que interage com a proteína transmembrana EtMIC4, formando um complexo. Além disso, tanto EtMIC4 quanto EtMIC5 demonstraram capacidade de ligação à célula hospedeira, *in vitro*^{24,31,32}. Composta por seis domínios Apple, TgMIC4 de *Toxoplasma gondii* interage com TgMIC1 via domínio A2, formando o complexo de proteínas micronêmicas MIC4-1-6 que interage com a superfície da célula hospedeira^{33,34}, e também é uma lectina, que se liga à galactose pelo domínio A5³⁵. Essencial para a formação da junção de movimento durante a invasão celular em várias espécies apicomplexas, AMA1 (*Apical Membrane Antigen-1*) contém dois domínios PAN/Apple. Estes domínios possuem longos loops que se estendem do núcleo indicando, em *Plasmodium falciparum*, adaptações para a evasão de resposta imune por anticorpos³⁶.

Um padrão de cisteínas e *clusters* de resíduos hidrofóbicos compõem o domínio Apple²⁵, que estruturalmente é constituído por uma folha, composta por 4 a 5 fitas-beta antiparalelas, amarrada à uma alfa-hélice; enquanto na outra face da folha beta pode haver uma pequena folha adicional, composta por duas fitas-beta antiparalelas mais curtas^{35,37} (figura 5). Embora algumas proteínas, como o plasminogênio, apresentem apenas 4 cisteínas (C), que formam as pontes dissulfeto do cerne do domínio²⁴, a maioria das proteínas presente em organismos apicomplexas apresenta o padrão de 6 cisteínas, formando 3 pontes dissulfeto: C2-C5 e C3-C4 seguram a folha-beta na alfa-hélice, e C1-C6 liga as extremidades N e C terminal do domínio.

Fig. 5 Representação estrutural do domínio Apple A9 da proteína MIC5 de *Eimeria tenella* (PDB 1HKY). A alfa-hélice é mostrada em vermelho. As fitas beta antiparalelas são mostradas em azul. Estão destacadas as 5 fitas que formam a folha beta maior e as 2 fitas que formam a folha beta menor. (Figura adaptada de Brown *et al* 2003³⁸).

1.1.4 Secretoma

O secretoma é o conjunto de proteínas secretadas por um organismo em um determinado momento. Como apresentado anteriormente (seção 1.1), no filo Apicomplexa, proteínas secretadas estão intimamente relacionadas ao processo de invasão celular, essencial para a sobrevivência e replicação do parasita.

Em *N. caninum* não há estudos prévios sobre o secretoma, já que estudos proteômicos anteriores focaram na identificação de proteínas no extrato total (ET) usando técnicas de gel bidimensional (2D), com ou sem complementação por *imunoblotting*^{39,40,41,42,43,44}. Outros estudos utilizaram a técnica de 2D-DIGE seguida por análise por espectrometria de massas (MS) para comparar os perfis de taquizoítas *versus* bradizoítas⁴⁵ ou de isolados atenuados *versus* selvagens⁴⁶. A técnica de cromatografia líquida acoplada à espectrometria de massas em tandem (LC-MS/MS) foi empregada em três estudos de identificação: antígenos de uma fração organelar após geração de anticorpos monoclonais⁴⁷, antígenos capazes de estimular células T CD4+ve bovinas⁴⁸, e proteínas em uma fração enriquecida com roptrias⁴⁹.

1.1.4.1 Secreção Ca^{2+} dependente

Foi verificado para *Toxoplasma gondii* e *Plasmodium falciparum* que, quando ocorre o contato inicial entre o parasita e a célula hospedeira, há um aumento na

concentração de Ca^{2+} no citoplasma do parasita. Este aumento está relacionado a eventos críticos para o parasita, como secreção de adesinas, motilidade por deslizamento, invasão celular e egresso^{20,50,51}. Estudos com *P. falciparum* e *T. gondii* indicam que, seguindo-se o contato com a célula hospedeira (figura 6), os mensageiros inositol trifosfato (IP_3) e ADP ribose cíclica (cADPR) são gerados e ligam-se a canais de cálcio putativamente localizados no retículo endoplasmático, liberando o Ca^{2+} ali armazenado^{20,52}. Eventos decorrentes da mobilização de cálcio, como ativação de CDPKs (proteínas cinases dependentes de cálcio) e ativação de outras cinases por cAMP e cGMP, regulam a exocitose de organelas secretórias e a ativação do motor de invasão do parasita^{20,52,53}.

Fig. 6 Vias de sinalização de secreção e motilidade afetadas por aumento de Ca^{2+} intracelular em *Toxoplasma gondii*. O retículo endoplasmático (ER) é a principal forma de armazenamento de Ca^{2+} , preenchido por ação da Ca^{2+} ATPase chamada SERCA. A estimulação de receptores ainda desconhecidos (R) na membrana plasmática do parasita geram inositol trifosfato (IP_3) e diacilglicerol (DAG) a partir do fosfatidil inositol bifosfato (PIP_2) por ativação da fosfolipase C (PLC). O estímulo, por receptor exclusivo ou comum, resulta também na produção de ADP ribose cíclica (cADPR) por ativação da ADPR ciclase. Evidências suportam a existência de canais de liberação de cálcio em resposta a cADPR, os RyR (ryanodine receptors), e a IP_3 (IP_3 receptors, IP_3R). O cálcio liberado ativa proteínas cinases cálcio-dependentes (CDPKs), implicadas com a secreção das micronemas. De forma independente, a ativação da guanilil ciclase (GC) para gerar GMP cíclico (cGMP) é importante para a ativação da proteína cinase G (PKG), que também tem controle sobre a secreção de micronemas. Estudos em *Plasmodium falciparum* indicam que as CDPKs podem agir sobre a miosina do complexo motor, que é ancorado ao complexo de membrana interna (IMC). (Adaptado de Billker et al 2009²⁰).

In vitro, o etanol foi identificado como uma das substâncias capazes de desencadear a secreção de organelas apicais por mobilização de Ca^{2+} intracelular^{51,54}. O emprego de etanol como estímulo secretório já foi utilizado em outros estudos proteômicos de *T. gondii*, como na identificação do secretoma⁵⁵, e na quantificação de sítios de fosforilação cálcio-dependentes⁵².

1.1.4.2 Identificação de proteínas por espectrometria de massas

Na identificação de proteínas por espectrometria de massas, uma estratégia que pode ser usada para obter melhores resultados é a árvore de decisão que guia a escolha do método de fragmentação de peptídeos. Foi observado que alguns métodos de fragmentação e leitura resultavam em um maior número de identificações para peptídeos de determinada carga e determinada massa⁵⁶. Assim, utilizando a árvore de decisão, ao serem detectadas a carga e a relação massa/carga (m/z) de um peptídeo, este é enviado para fragmentação e leitura pelo método que garantirá um melhor resultado para as suas características, como mostrado na figura 7.

Fig. 7 Exemplo de árvore de decisão usada na escolha do método de fragmentação de peptídeos em espectrômetro de massas LTQ-Orbitrap. Dependendo da carga (2+, 3+, 4+ ou maior que 5+) e da relação massa/carga (m/z) do peptídeo precursor (*precursor charge state*) este será enviado para o método de fragmentação e análise mais apropriado para seu caso específico. HCD, Dissociação por Colisão de Alta Energia; ETD, Dissociação por Transferência de Elétrons; ETD FT, ETD e analisador com Transformada de Fourier. (Adaptado de Frese *et al*, 2011⁵⁶).

Para a quantificação relativa entre amostras proteicas, diferentes marcações (multiplex) com dimetil-isótopos estáveis podem ser utilizadas⁵⁷. Neste tipo de marcação química, que é uma alternativa rápida e econômica a outros métodos de marcação como iTRAQ, as aminas primárias (N terminal e cadeia lateral de resíduos de lisina) do peptídeo são convertidas em dimetilaminas, numa reação com um dos isótopos na presença de cianoboroidreto. Três isótopos do formaldeído podem ser usados, levando à marcação “leve”, “intermediária” ou “pesada” do peptídeo (figura

8). Em relação ao peptídeo não marcado, são adicionados 28 Da, 32 Da, e 36 Da, respectivamente, para cada uma das marcações. No espectro de massas MS podem ser observadas essas diferenças de peso na relação m/z .

Marcação de peptídeos com dimetil-isótopos

Fig. 8 Marcação de peptídeos para quantificação multiplex por dimetil-isótopos. À esquerda são mostradas as reações que levam às três diferentes marcações. R-NH_2 representa o peptídeo. **Marcação leve**, a combinação entre formaldeído regular e cianoboroidreto (NaBH_3CN) levam a um aumento de 28 Da por amina primária do peptídeo. **Marcação intermediária**, formaldeído deuterado com cianoboroidreto, aumento de 32 Da por amina primária. **Marcação pesada**, formaldeído deuterado com ^{13}C e cianoboroidreto deuterado, aumento de 36 Da por amina primária. **A**, peptídeo não marcado. **B**, peptídeo com marcação leve (vermelho), ou intermediária (verde), ou pesada (azul). São mostrados os cromatogramas dos peptídeos com seus tempos de retenção (min) em coluna C18 e os espectros de massas MS com suas relações massa/carga (m/z). Após marcação, o tempo de retenção sofre um aumento na coluna C18 em relação ao peptídeo não marcado, mas é o mesmo para as três diferentes marcações. Peptídeos com marcação leve, intermediária e pesada são aumentados em 28.03, 32.06 e 36.08 m/z , respectivamente, o que corresponde a 56.06, 64.11 e 72.15 Da, já que o peptídeo é duplamente carregado. As setas indicam os sítios de dimetilação no peptídeo. (Adaptado de Boersema *et al*, 2009⁵⁷).

As amostras marcadas são misturadas e, na análise por LC-MS, peptídeos iguais, porém com diferentes marcações, são reconhecidos pela diferença de massa entre eles adicionada pelos isótopos. A quantificação pode ser então realizada por comparação entre a intensidade de sinal destes peptídeos (figura 9).

Fig. 9 Quantificação relativa usando marcação com dimetil-isótopos (duplex). Curva de distribuição das razões (Log2 ratio) entre as amostras A (azul) e B (rosa). Para cada ponto a razão foi obtida por divisão da intensidade (área do pico) do peptídeo na amostra A sobre a intensidade na amostra B. Na região 1 da curva estão os peptídeos de proteínas diminuídas na amostra A (*down regulated*); na região 2 estão os peptídeos de proteínas não diferencialmente expressas; na região 3 estão os peptídeos de proteínas aumentadas na amostra B (*up regulated*). (Adaptado de Boersema *et al*, 2009⁵⁷).

1.2 Justificativa e trabalho desenvolvido

A descoberta de *N. caninum* é recente e até hoje não existe um método terapêutico eficaz para o seu combate. Informações bioquímicas sobre os processos de invasão, multiplicação, egresso, metabolismo e resposta imune ainda são muito limitadas em relação a este parasita; e a compreensão destes processos é fundamental para o desenvolvimento de uma estratégia terapêutica efetiva.

O processo de invasão, por ser crucial para a sobrevivência de parasitas apicomplexas e por envolver organelas e sistemas específicos, que não são compartilhados com seus hospedeiros intermediários, tem sido investigado para a identificação de potenciais alvos terapêuticos^{58,59}.

Dois estudos foram aqui propostos para a investigação de proteínas envolvidas com o processo de invasão pelo parasita *Neospora caninum*; um para a identificação e caracterização de proteínas compostas por domínios Apple; e outro para a investigação de seu secretoma.

1.2.1 Caracterização de proteínas com domínio Apple

Com base nas propriedades adesivas descritas para as proteínas com domínios Apple (chamadas aqui 'proteínas Apple'), e por existirem importantes proteínas micronêmicas com estes domínios, nós procuramos por novas proteínas de *N. caninum* contendo domínios Apple, que poderiam potencialmente estar relacionadas ao processo de invasão celular.

Este trabalho focou na caracterização de três proteínas Apple de *N. caninum*, incluindo a caracterização molecular e estrutural por modelagem molecular baseada em homologia. Para duas delas foi realizada também a caracterização físico-química e funcional.

Em *T. gondii*, quatro homólogas destas proteínas foram identificadas, primeiramente *in silico*. A única caracterização experimental para elas foi a imunolocalização no parasita. Porém, o resultado não foi conclusivo, devido à colocalização parcial em micronemas e roptrias, provavelmente causada pela cauda de YFP incorporada às proteínas. Apesar do depósito das sequências no GenBank, os autores não nomearam estas proteínas Apple⁶⁰.

Durante a execução deste projeto, três anos após seu início, uma entre essas três proteínas Apple de *N. caninum* foi identificada após produção de anticorpos monoclonais contra um extrato de organelas do parasita, com identificação dos antígenos após eluição dos anticorpos⁴⁷. Esta proteína foi localizada nas micronemas de *N. caninum* por detecção com seu anticorpo monoclonal, e os autores a nomearam MIC17B. Devido à alta identidade com as outras duas proteínas Apple e à disposição de seus genes no genoma, esses autores expandiram a nomenclatura às duas outras proteínas, como MIC17A e MIC17C, que, entretanto, não foram por eles identificadas. Esta nomenclatura também foi estendida por eles às proteínas homólogas em *T. gondii*, que haviam sido previamente identificadas⁶⁰.

No presente trabalho apresentamos as peculiaridades de cada uma destas três proteínas Apple de *N. caninum*, mostrando que, apesar da alta similaridade, podem ter funções e localizações diferentes no parasita, o que questionaria inclusive a nomenclatura previamente adotada.

1.2.2 Estudo de secretoma

O interesse na investigação do secretoma de parasitas apicomplexas reside no envolvimento das proteínas secretadas com o processo de invasão celular, essencial para a sobrevivência e replicação do parasita. Em nosso estudo sobre o secretoma, inicialmente, a análise do extrato secretado (ESA) foi feita com separação por gel 2D e identificação de proteínas por MS-MALDI (Ionização e Dessorção a Laser Assistida por Matriz), sem o emprego de etanol para o estímulo secretório.

Posteriormente, com a oportunidade de realização das análises junto ao *Biomolecular Mass Spectrometry and Proteomics Group*, numa colaboração com a Universidade de Utrecht, Holanda, o secretoma foi investigado por duas abordagens diferentes, utilizando técnicas de espectrometria massas de alta resolução (nanoLC-MS/MS). Na primeira abordagem foram identificadas as proteínas presentes no extrato secretado de *N. caninum* (ESA) após estímulo secretório com emprego de etanol. Na segunda abordagem foi realizada uma quantificação relativa do proteoma dos taquizoítas estimulados e não estimulados para a secreção por etanol, e a construção de uma rede de interação, por predição, com as proteínas diferencialmente expressas. No presente estudo, um grande número de identificações foi alcançado e os resultados foram explorados numa busca por proteínas envolvidas na invasão celular e no mecanismo de secreção do parasita. Os achados desta análise contribuem com novas informações sobre *N. caninum* e poderão auxiliar na seleção futura de alvos em potencial a serem investigados para o desenvolvimento de tratamento terapêutico contra a neosporose.

2. OBJETIVOS

2.1 Objetivos gerais

- 1) Caracterizar molecularmente proteínas com domínios Apple em *N. caninum*;
- 2) Investigar secretoma de *N. caninum*.

2.2 Objetivos específicos

2.2.1 Caracterização de proteínas Apple em *N. caninum*

- 1) Identificar, clonar e expressar de forma heteróloga os genes codificadores de proteínas com domínio Apple;
- 2) Caracterizar estruturalmente os domínios Apple por modelagem molecular;
- 3) Detectar, identificar e localizar as proteínas Apple nativas de *N. caninum*;
- 4) Avaliar funcionalmente por ensaios de invasão, de adesão e por obtenção de taquizoíta nocaute os genes das proteínas Apple.

2.2.2 Estudo do secretoma de *N. caninum*

- 1) Identificar proteínas no extrato secretado (ESA) do taquizoíta de *N. caninum*;
- 2) Quantificar relativamente os extratos totais de taquizoítas submetidos e não submetidos ao estímulo secretório;
- 3) Construir uma rede de interações entre proteínas diferencialmente expressas no taquizoíta estimulado;
- 4) Identificar alvos em potencial para o desenvolvimento de tratamento terapêutico.

3. MATERIAL E MÉTODOS

3.1 Técnicas em comum

3.1.1 Cultura *in vitro* e purificação de taquizoítas de *N. caninum*

Células Vero, cultivadas em garrafas de 25 cm² e 75 cm² em meio RPMI 1640 (Sigma-Aldrich) contendo 5% de soro fetal bovino (SFB, Gibco), 2mM de glutamina (Gibco) e 10 µg/mL de kanamicina, mantidas a 37°C com 5% de CO₂, foram usadas como células hospedeiras para a manutenção de taquizoítas de *N. caninum*.

Taquizoítas da linhagem Nc-1⁶¹ foram mantidos em meio RPMI 1640 (Sigma-Aldrich) contendo 2mM de glutamina e 10 µg/mL de kanamicina a 37°C com 5% de CO₂. Aproximadamente uma semana após a inoculação, tempo para a destruição parcial das células Vero, os taquizoítas foram purificados. Para a remoção das células Vero aderidas à garrafa foi utilizado um *scraper*, e o sobrenadante, contendo células e taquizoítas, foi submetido à lise por passagens repetidas em agulha 26G1/2. Os debris de células Vero foram removidos por centrifugação a 1000 x g, por 3 min, a 4°C; o pellet contendo taquizoítas foi ressuspensionado em 10 mL de PBS e passado em filtro de 5 µm. Os taquizoítas foram novamente coletados por centrifugação e ressuspensionados em 2,5 mL de PBS, e então submetidos à purificação em coluna PD10 (*Sephadex G-25*, GE Healthcare). As frações eluídas foram centrifugadas a 3500 x g por 10 min a 4°C, e os pellets de taquizoítas foram reunidos em um único tubo, diluídos para um volume final de 500 µL e contados em câmara de Neubauer.

3.1.2 Extração de DNA genômico de *N. caninum*

Aproximadamente 10⁸ taquizoítas de *N. caninum* foram incubados com 500 µL de tampão de lise (Tris 40 mM pH 8,0; EDTA 80 mM pH 8,0; SDS 2%) e 100 µL de proteinase K (0,1 mg/mL) por 1 hora a 37°C. Após centrifugação a 1000 x g por 5 minutos, foram adicionados ao sobrenadante 500 µL de solução fenol:clorofórmio (1:1) e esta mistura foi centrifugada a 3500 x g por 8 minutos. À fase aquosa adicionou-se 1/10 de seu volume de acetato de sódio 3 M pH 5,5, e 2,5 vezes seu volume de etanol absoluto. A incubação para precipitação do DNA foi feita por 16 horas a -20°C, o qual foi recuperado por centrifugação por 15 minutos. O *pellet* de

DNA foi lavado com 500 µL de etanol 70%, centrifugado por 10 minutos, seco a temperatura ambiente, e dissolvido em água Milli-Q autoclavada ou em tampão TE (Tris 10 mM pH 8,0; EDTA 1mM). Para melhor dissolução do DNA, a solução final foi aquecida a 56°C por 1 hora, e depois armazenada a -20°C.

3.1.3 Extração de RNA total

A extração de RNA total de aproximadamente 10⁸ taquizoítas de *N. caninum* foi feita com 1,0 mL de Trizol (Gibco) em incubação a temperatura ambiente por 5 min. Para a separação de fases foram adicionados 0,2 mL de clorofórmio, com agitação por inversão, e incubação por 3 minutos a temperatura ambiente. Após centrifugação a 3500 x g, por 15 minutos a 4°C, o RNA presente na fase aquosa foi precipitado por adição de 0,5 mL de álcool isopropílico absoluto e incubação a temperatura ambiente por 10 minutos. O *pellet* de RNA foi coletado por centrifugação a 3500 x g, por 10 minutos a 4°C, e lavado com etanol 75%. Após nova centrifugação, o *pellet* foi seco a temperatura ambiente e dissolvido em água livre de RNase.

Tanto DNA quanto RNA foram quantificados em espectrofotômetro *GeneQuant* (GE Healthcare), considerando também a relação entre as detecções em 260 nm (nucleotídeos) e 280 nm (proteínas) com valor entre 1.6 e 2.0 para obtenção de adequado grau de pureza do material genético.

3.1.4 Reação em cadeia de polimerase (PCR)

Para as reações em cadeia de polimerase, diferentes kits de Taq polimerase foram utilizados e serão discriminados a cada seção, mantendo-se as concentrações dos reagentes: 1x tampão para Taq polimerase; 1,5 mM de MgCl₂ (Invitrogen) - quando não presente no tampão; 200 µM de DNTP Mix (Qiagen); 2 pmol de cada primer (IDT); 2,5 U da enzima Taq polimerase; aproximadamente 100 ng do material genético a ser amplificado (DNA, cDNA, plasmídeo ou produto de PCR); e água Milli-Q autoclavada para 25 µL (amplificação para conferência) ou 50 µL (amplificação para purificação).

As reações foram realizadas em termociclador (*MasterCycler*, Eppendorf), iniciando-se com pré-desnaturação a 95°C por 2 minutos, seguida por 40 ciclos de: desnaturação a 94°C por 30 segundos; anelamento com temperatura adequada aos

primers por 1 minuto; e extensão a 72°C por 2 minutos. Uma extensão final foi adicionada após a finalização dos 40 ciclos, a 72°C por 10 minutos.

3.1.5 Ligação de fragmento de DNA (inserto) a plasmídeo

Fragmentos de DNA, oriundos de produtos de PCR purificados ou de plasmídeos digeridos, foram ligados a plasmídeos de clonagem (pGEM), plasmídeos de expressão (pET28 e pET28-SUMO) ou plasmídeo para deleção (pΔApple). Para ligação ao plasmídeo pGEM, foi utilizado o kit *pGEM-T Easy Vector Systems* (Promega), seguindo orientações do fabricante. A reação de ligação foi incubada a 4°C por 16 horas, e depois utilizada para a transformação de células eletrocompetentes TOP10.

A reação para ligação de fragmentos aos plasmídeos de expressão ou deleção foi realizada com a enzima *T4 DNA Ligase* (Fermentas), na presença de 10 mM de DTT, 0,5 mM de ATP, em tampão para *T4 DNA Ligase*. A reação foi incubada a 4°C por 24 a 48 horas, e a ligação usada para a transformação de células de *E. coli* eletrocompetentes TOP10 ou BL21.

3.1.6 Preparo de células eletrocompetentes

Cepas de *Escherichia coli* TOP10 ou BL21 foram plaqueadas em agar Luria-Bertani (LB: 1% tripton, 0,5% extrato de levedura, 1% NaCl, 1,5% agar em pH 7,0) a 37°C por 18 horas. Uma colônia desta placa foi repicada em 50 mL de meio LB e incubada a 37°C por 18 horas, sob agitação a 250 rpm em Incubadora de Bancada Refrigerada CT-712R (Cientec). Após incubação, 5 mL do caldo foram inoculados em 500 mL de meio LB e incubados a 37°C, sob agitação, até a densidade óptica em 600 nm (OD_{600}) atingir entre 0,5 e 0,8 (medido no aparelho *Genequant*, GE Healthscience).

Após incubação em gelo por 30 minutos, o conteúdo foi centrifugado (4000 x g, 4°C, 15 minutos) e o sobrenadante descartado. O pellet foi gentilmente ressuspenso, a 4°C, em 500 mL de água deionizada esterilizada, centrifugado e o sobrenadante foi descartado. O procedimento anterior foi repetido e o pellet resultante foi homogeneizado em 20 mL de glicerol 10% esterilizado, centrifugado, e o sobrenadante descartado. O pellet foi resuspenso, desta vez em 1 mL de glicerol 10% esterilizado, e esta suspensão foi diluída 10 vezes em água deionizada para avaliação

de OD₆₀₀. A concentração de células para transformação (10^8 UFC) foi considerada ideal com OD maior que 0,5.

3.1.7 Transformação de células eletrocompetentes

Para a multiplicação e seleção de plasmídeos, cepas de *E. coli* eletrocompetentes TOP10 ou BL21 foram transformadas por eletroporação com o plasmídeo previamente ligado ao fragmento de interesse. A 50 µL de células eletrocompetentes, descongeladas e mantidas em gelo, foram adicionados 1 a 2 µL da ligação e o choque foi dado no eletroporador *GenePulser Xcell* (BIO-RAD) nas seguintes condições: voltagem 2,5kV; resistência 200 Ω; capacitância 25 µF; em cubeta de 2 mm previamente mantida em gelo. Em seguida foram adicionados 250 µL do meio de cultura SOC (2% de triptona, 0,5% de extrato de levedura, 10mM de NaCl, 2,5mM de KCl, 10mM de MgCl₂, 10mM MgSO₄ e 20mM de glucose) ou meio NYZ+ (mesma composição de SOC, com 1% de NZ amina e sem triptona) e todo o material recuperado da cubeta foi incubado a 37°C por 1 hora, sob agitação (350 rpm). Após a incubação todo o conteúdo foi plaqueado em meio de cultura LB com reagentes para a seleção, e incubado a 37°C por 16 horas.

3.1.8 Purificação de plasmídeos selecionados (miniprep)

As colônias transformadas e selecionadas após o plaqueamento foram separadamente inoculadas em 3 mL de meio LB contendo antibiótico de seleção e incubadas a 37°C sob agitação a 200 rpm por 16 horas. Uma parcela da cultura foi armazenada em glicerol a -70°C e o restante foi utilizado para a purificação dos plasmídeos. Os kits *illustra plasmidPrep Mini Spin Kit* (GE Healthcare) e *PureYield Plasmid Miniprep System* (Promega) foram utilizados para o isolamento dos plasmídeos, seguindo-se as recomendações dos fabricantes.

3.1.9 Eletroforese em gel de agarose

Géis de agarose (0,8%) foram preparados em tampão TAE (0,04 M Tris acetato, 1 mM EDTA, pH 8,3) contendo 0,5 µg/mL de brometo de etídeo (Promega). As eletroforeses foram realizadas em cuba *Easy Cast Mini* (Owl Scientific) ou cuba *EH5-130N* (Bioagency) contendo tampão TAE, usando voltagem constante de 100 V e amperagem livre. As amostras foram diluídas seis vezes em tampão de corrida

Orange DNA Loading Dye (Fermentas) e foram utilizados os marcadores moleculares (5 µL/poço) *GeneRuler DNA Ladder Mix* ou *GeneRuler 1kb DNA Ladder* (ambos da Fermentas). As fotografias dos géis de agarose foram feitas em fotodocumentador Digidoc-BA (Bioagency).

3.1.10 Quantificação de proteínas pelo método de Bradford

Proteínas recombinantes ou presentes em extratos proteicos foram quantificadas pelo método de Bradford⁶². Uma curva de calibração foi construída com BSA (albumina sérica bovina acetilada), diluída na mesma solução em que as amostras de proteína estavam solubilizadas, nas concentrações 1 mg/mL, 0,75 mg/mL, 0,5 mg/mL, 0,25 mg/mL, 0,125 mg/mL e 0,0 mg/mL. As amostras foram diluídas nas proporções 1:2, 1:4, 1:10 e 1:20 e, em uma placa de ELISA com fundo chato, 20 µL de cada diluição de BSA e das amostras foram adicionados a 200 µL do reagente de Bradford (Bioagency) diluído 5 vezes em água. A leitura espectrofotométrica foi feita em duplicata, utilizando comprimento de onda de 595 nm, em leitor de microplacas *SpectraMax Plus* (Molecular Devices).

3.1.11 Eletroforese em gel de acrilamida

A separação de proteínas para fins de visualização, determinação de peso molecular, determinação de ponto isoelétrico, excisão para análise por espectrometria de massas ou detecção por *Western blot* (WB) foi realizada por eletroforese uni ou bidimensional em gel de acrilamida.

3.1.11.1 Eletroforese unidimensional (SDS-PAGE 1D)

A separação de proteínas por peso molecular foi feita por eletroforese em gel de acrilamida a 12,5% (sistema SE260, GE Healthcare), sob voltagem fixa de 120 V e amperagem livre, em tampão para eletroforese (Tris 25 mM; glicina 192 mM; SDS 0,1%). As amostras foram diluídas duas vezes em tampão de corrida (Tris-HCl 187,5 mM pH 6,8; SDS 6%; glicerol 30%; azul de bromofenol 0,03%; DTT 26 mg/mL) e os marcadores de peso molecular *Protein MW Marker* (Fermentas) ou *Precision Plus Protein All Blue Standards* (BioRad) foram utilizados em caso de géis submetidos à coloração, e o marcador pré-corado *Precision Plus Protein Dual Color* (BioRad) em caso de transferência para membrana de *Western blot*.

3.1.11.2 Eletroforese bidimensional (SDS-PAGE 2D)

A separação de proteínas por eletroforese bidimensional permite a determinação do peso molecular e do ponto isoelétrico (pI) destas. Na primeira dimensão, as proteínas são submetidas a uma focalização isoelétrica, que permite que as proteínas migrem por um gradiente de pH imobilizado em gel (*strip*), por atuação de um campo elétrico, até atingirem o pH em que sua carga se iguala a zero, correspondente ao seu pI, e precipitem. Após essa separação, as proteínas têm suas cargas igualadas e suas pontes dissulfeto reduzidas para a separação por peso molecular, que corresponde à segunda dimensão.

3.1.11.2.1 Primeira dimensão ou focalização isoelétrica

Para serem submetidas à focalização isoelétrica, as amostras proteicas foram preparadas em tampão de amostra (AM) contendo agentes desnaturantes e um detergente zwitteriônico (ureia 7M, tioureia 2M e CHAPS 4%) e adicionadas à solução de reidratação (RHS: ureia 7M, tioureia 2M, CHAPS 4%, traços de azul de bromofenol, 0,5% de mistura anfólitos carreadores adequados para determinada faixa de pH (IPG *Buffer*, GE Healthcare), 0,3% de DTT (p/v), e 2% coquetel de inibidores de proteases (Sigma-Aldrich)). O volume final de amostra em RHS foi de 125 µL para a *strip* de 7 cm e de 250 µL para a *strip* de 13 cm. Foram utilizadas *strips* de pH imobilizado 7-11 e 3-11 NL (*Rehydrate Immobiline DryStrip gel*, GE Healthcare).

Em bandeja para reidratação (*DryStrip Reswelling Tray*, GE Healthcare), as *strips* foram posicionadas sobre as amostras contidas em solução de reidratação, cobertas com óleo mineral (*DryStrip Cover Fluid*, GE Healthcare) e incubadas a temperatura ambiente por 16 horas. Hidratadas, as *strips* foram transferidas para o equipamento *Ettan IPGphor3* (GE Healthcare), cobertas com óleo mineral e submetidas à corrente elétrica, com a programação de acordo com as recomendações do fabricante (Tabela 1). Finalizada a corrida, as *strips* foram armazenadas a -70°C até o momento da corrida em segunda dimensão.

Tabela 1 Programas utilizados para a focalização isoeétrica de *strips*. As focalizações ocorreram a 20°C com correntes de 50 μ A e 75 μ A, por *strip*, para as *strips* de 7 cm e 13 cm, respectivamente.

Etapa	Modo de voltagem	13 cm / 3-11 NL		7 cm / 3-11 NL		7 cm / 7-11 NL	
		Voltagem (V)	kVh	Voltagem (V)	kVh	Voltagem (V)	kVh
1	<i>Step and Hold</i>	500	0.5	300	0.2	300	0.2
2	<i>Gradient</i>	1000	0.8	1000	0.3	1000	0.7
3	<i>Gradient</i>	8000	11.3	5000	4.0	5000	4.5
4	<i>Step and Hold</i>	8000	4.4	5000	2.5	5000	3.0
	Total		17.0		7.0		8.4

3.1.11.2.2 Segunda dimensão

Após a focalização isoeétrica, as *strips* foram equilibradas, o que corresponde à redução das pontes dissulfeto das proteínas com DTT e alquilação das cisteínas reduzidas com iodoacetamida, para que as pontes não voltem a se formar. Primeiramente as *strips* foram incubadas em solução de equilíbrio (Tris 50 mM, ureia 6 M, glicerol 30%, SDS 2%) contendo 10 mg/mL de DTT, por 15 minutos, a temperatura ambiente, sob agitação. Em seguida, esta solução foi trocada por nova solução de equilíbrio contendo 25 mg/mL de iodoacetamida, e nova incubação foi feita por 15 minutos, sob agitação.

Equilibradas, as *strips* foram posicionadas na horizontal sobre o gel de acrilamida a 12,5%. À esquerda da *strip*, também em contato com o gel, foi posicionado o marcador de peso molecular, adicionado a um papel de filtro para *Western blot*. A *strip* e o marcador foram selados com tampão de eletroforese adicionado de 0.5% de agarose e traços de azul de bromofenol, para que, solidificado, permita a passagem de corrente elétrica e a visualização da corrida. A eletroforese foi realizada com tampão Tris-glicina-SDS, em que o SDS garante que todas as proteínas fiquem com carga negativa e sejam separadas exclusivamente por seu peso molecular.

Para *strips* de 7 cm, a corrida foi feita em sistema SE 260 (GE Healthcare) sob amperagem de 15 mA por gel durante os primeiros 15 minutos, e o restante do tempo sob 30 mA por gel. Para *strips* de 13 cm foi utilizado o sistema SE 600 Ruby (GE Healthcare), com amperagem de 60 mA por gel por 15 min, e 120 mA por gel no restante da corrida.

3.1.12 Coloração de gel de acrilamida

3.1.12.1 Coloração por Coomassie G-250

O uso do corante coloidal *Coomassie Brilliant Blue* G-250 (USB) permite detecção de pelo menos 10 ng de proteína, separada em gel de acrilamida uni ou bidimensional⁶³. Completada a eletroforese, o gel foi incubado em solução para fixação de proteínas (ácido fosfórico 3%, etanol 50%) por no mínimo 1 hora. Foram feitas três lavagens de 5 minutos com água Milli-Q, sob agitação, e foi então adicionada a solução de Coomassie (Coomassie G-250 0,1%, sulfato de amônio 17%, metanol 35%, ácido fosfórico 3%), com incubação a temperatura ambiente, sob agitação, por 1 a 3 dias.

3.1.12.2 Coloração por prata

A coloração com nitrato de prata apresenta uma sensibilidade superior à coloração com Coomassie, com limite de detecção de menos que 1 ng de proteína, e este protocolo também apresenta compatibilidade com análise por espectrometria de massas⁶⁴. Esta coloração foi priorizada para os géis contendo amostras escassas. Todas as incubações e lavagens descritas a seguir foram feitas a temperatura ambiente e sob agitação.

Após a eletroforese, o gel foi incubado por no mínimo 1 hora com solução de fixação (metanol 50%, ácido acético 10%) e em seguida por 15 minutos com solução contendo 5% de metanol. Foram feitas três lavagens de 5 minutos com água Milli-Q, e o gel foi sensibilizado com 0,02% (p/v) de tiosulfato de sódio pentaidratado por 2 minutos, e enxaguado três vezes com água Milli-Q, por 20 segundos. A incubação com a solução de nitrato de prata 0,2% (p/v) foi feita por 23 minutos, seguida por três lavagens rápidas (20 segundos) com água em abundância. A solução de revelação (carbonato de sódio 3%, formaldeído 0,0185%, tiosulfato de sódio pentaidratado 0,0004%) foi adicionada primeiramente para uma breve lavagem, e descartada. Em seguida, o gel foi incubado com a solução de revelação até que fosse obtida a visualização desejada das bandas ou *spots* (entre 4 e 10 minutos). A revelação foi interrompida com uma lavagem em solução de EDTA a 1.4% (p/v) e o gel foi submetido a uma lavagem final com água Milli-Q.

3.1.13 Captação de imagens de géis de acrilamida e de filmes de *Western blot*

Os géis de acrilamida e filmes de *Western blot* foram digitalizados no equipamento *Image Scanner*, utilizando o software *LabScan* (GE, Healthcare).

3.1.14 Detecção de proteínas por *Western blot*

A técnica de *Western blot* (WB) foi utilizada para a detecção de proteínas presentes em géis de acrilamida 1D ou 2D por anticorpos específicos.

Finalizada a eletroforese uni ou bidimensional, foi realizada transferência das proteínas do gel de acrilamida para membrana de fluoreto de polivinilideno (PVDF) (*Immobilon TM-P*, Millipore). Para a transferência, foi utilizado o sistema semi-seco (TE PWR 77, GE), no qual o gel de acrilamida foi acomodado sobre a membrana de PVDF, previamente ativada com metanol por 5 minutos e equilibrada com tampão de transferência (Tris 0,25 M, glicina 0,2 M, SDS 0,001%, metanol 15%) por 5 minutos. Abaixo da membrana e acima do gel foram adicionadas três folhas de papel de filtro (BioRad) previamente embebidas em tampão de transferência. Após a adição de cada componente, possíveis bolhas foram removidas com a passagem de um tubo cilíndrico. A corrente aplicada foi de 65 mA por gel (8,6 cm x 10 cm) por 1 hora.

Terminada a transferência, foram marcados com lápis sobre a membrana o início e o fim da corrida eletroforética e a distribuição do marcador de peso molecular. A membrana foi corada por 5 minutos em solução DB71 (*Direct Blue 71* (Sigma-Aldrich) 0,008%, etanol 40%, ácido acético 10%), digitalizada, e então descorada com solução de descoloração (bicarbonato de sódio 0,15 M em etanol 50%).

Depois de seca, a membrana foi bloqueada por 1 hora, sob agitação, em PBS-GT 0.8% (gelatina suína (Sigma) 0,8%, Triton X-100 (USB) 0,05%, em PBS), e lavada com PBS-T (Triton X-100 0,05% em PBS) em quatro etapas: um enxágue rápido, uma lavagem de 10 minutos e duas lavagens de 5 minutos.

A incubação da membrana com anticorpos primários foi feita utilizando soro ou anticorpos purificados diluídos em PBS-GT 0,08%, por 16 horas, a temperatura ambiente, sob agitação. Após as quatro etapas de lavagem com PBS-T, a membrana foi incubada com anticorpo de cabra anti-coelho conjugado à enzima peroxidase (*Goat Anti-Rabbit IgG (H+L) HRP Conjugate* ZyMax, Invitrogen) diluído em PBS-GT 0,08%, por 1 hora, a temperatura ambiente, sob agitação. A membrana foi submetida às quatro etapas de lavagem com PBS-T e então incubada por 5 minutos com os reagentes ativadores e estabilizadores da peroxidase (*SuperSignal*

West Pico Chemiluminescent Substrate, Pierce). Envolta em filme plástico e dentro de um cassete, a membrana foi exposta a filme fotográfico (AGFA - *Ortho CP-G Plus, Medical X-Ray Film*) por determinado tempo que variou de 30 segundos a 5 minutos. O filme foi revelado em sala escura com incubação de 1 minuto em reagente revelador (Exsil MX), lavagem por 1 minuto em água, incubação de 1 minuto em reagente fixador (Exsil MX), e enxágue em água.

3.1.15 Preparo de amostras para análise por espectrometria de massas

As proteínas presentes em bandas e *spots* de géis de acrilamida 1D ou 2D foram excisadas para serem digeridas e submetidas à análise por espectrometria de massas (MS).

A descoloração de bandas e *spots* provenientes de gel corado com Coomassie G-250 foi feita com três lavagens em solução de NH_4HCO_3 25 mM pH 8,0 preparada em 50% acetonitrila (ACN). No caso de géis corados com prata, foram feitas duas lavagens com as soluções $\text{Na}_2\text{S}_2\text{O}_3$ 50mM e $\text{K}_3[\text{Fe}(\text{CN})_{16}]$ 15mM, na proporção 1:1, seguidas por duas lavagens com água Milli-Q. Para ambos os casos, após a descoloração, os pedaços de gel foram lavados duas vezes com NH_4HCO_3 25 mM pH 8,0, preparado em água Milli-Q. Em todas as lavagens houve agitação do tubo em *vortex* seguida por 10 minutos de descanso. No intervalo entre as lavagens, os pedaços de gel foram incubados por 10 minutos com ACN 100%, para que fossem desidratados, favorecendo assim a absorção da próxima solução a ser adicionada, mimetizando um efeito-esponja.

No caso de géis 1D, após a descoloração foram feitas duas etapas adicionais: uma para a redução das pontes dissulfeto, em incubação com DTT 6.5 mM preparado em NH_4HCO_3 25 mM pH 8.0, a 60°C, por 1 hora, sob agitação; e outra, após desidratação de gel com ACN 100%, para a alquilação das cisteínas reduzidas, com iodoacetamida 54 mM também preparada em NH_4HCO_3 25 mM pH 8.0, e incubação a temperatura ambiente por 25 minutos, no escuro. Após a redução e a alquilação, nova lavagem com NH_4HCO_3 25 mM pH 8.0 foi feita, e os géis foram desidratados com ACN 100%, até que ficassem com coloração branca opaca, e secos em centrífuga a vácuo (*CentriVap Concentrator*, Labconco).

Para a digestão das proteínas, 15 a 20 μL de solução de tripsina 25 ng/ μL (*Trypsin Gold*, Promega), preparada em NH_4HCO_3 25 mM pH8,0, foram adicionados aos pedaços de gel, que foram reidratados por 30 minutos a temperatura ambiente.

Em seguida, foi adicionado volume suficiente de solução NH_4HCO_3 25 mM pH8,0 para cobrir o gel, e a digestão foi feita em estufa a 37°C por 16 horas.

A solução contendo tripsina e peptídeos foi transferida para um novo tubo de 1,5 mL e duas extrações adicionais foram feitas, reunidas depois no mesmo tubo. Em caso de análise por MALDI, foi feita uma extração com solução de ácido trifluoroacético (TFA) 0,1% em ACN 50%, e outra apenas com ACN 100%. Para análise por eletrospray, as duas lavagens foram feitas com ACN 100%. A solução contendo peptídeos foi seca em centrífuga a vácuo e armazenada a 4°C até sua utilização. Quando analisados por MALDI, os peptídeos foram ressuspensos em matriz de ácido alfa-ciano-4-hidroxicinâmico e, quando analisados por eletrospray, em ácido fórmico 10%.

3.2 Caracterização das proteínas com domínios Apple

3.2.1 Obtenção e análise das sequências genômicas

As sequências parciais de duas proteínas transcritas (EST - *expressed sequence tags*) depositadas em banco de dados (GenBank) com os códigos CF660241.1 e BF249278.1 foram reconhecidas por nosso grupo como portadoras de domínios PAN/Apple. Estas sequências foram submetidas a uma busca contra o banco de dados genômico não-annotado (à época) de *N. caninum* do *Sanger Institute*⁶⁵. Com as sequências genômicas encontradas, a organização predita dos genes codificadores foi verificada no banco de dados anotado do projeto ToxoDB⁶⁶, que é membro do conjunto de bancos de dados de patógenos EuPathDB (*Eukaryotic Pathogens Database, Bioinformatics Resource Center*). Originalmente planejado para ser um banco de dados do genoma anotado do protozoário *Toxoplasma gondii*, atualmente o ToxoDB inclui também os dados genômicos de *N. caninum* e *Eimeria tenella*. Os domínios PAN/Apple das sequências foram preditos com a ferramenta *Pfam* disponibilizada na página do *Sanger Institute*. Um alinhamento foi feito com as sequências de aminoácidos das proteínas encontradas para identificação de regiões conservadas entre elas. Foi utilizada a ferramenta *Align* do site www.uniprot.org.

3.2.2 Amplificação e sequenciamento das sequências codificadoras de proteínas hipotéticas contendo domínios PAN/Apple

A partir de 1,95 µg de RNA total de taquizoítas de *N. caninum*, o seu DNA complementar (cDNA) foi preparado utilizando o kit *GoScript Reverse Transcription System* (Promega). Juntamente com 1 µL do *primer* oligo(dT), o RNA foi desnaturado a 70°C por 5 minutos e, em seguida, foram adicionados os outros reagentes para a polimerização de cDNA: tampão 1x, 3,5 mM de MgCl₂, 0,5 mM de cada dNTP, 20 U de inibidor de ribonuclease, 1 µL da enzima transcriptase reversa, e água Milli-Q autoclavada para 20 µL. A reação de polimerização foi realizada nas seguintes condições: anelamento a 25°C por 5 minutos; extensão a 42°C por 1 hora; e inativação a 70°C por 15 minutos. O RNA restante na reação foi degradado com 2 U de RNase H, em reação a 37°C por 20 minutos, e o cDNA armazenado a -20°C.

Foram realizadas PCRs (seção 3.1.4) para a amplificação, a partir de cDNA, de quatro sequências possivelmente codificadoras de proteínas contendo domínios

PAN/Apple, utilizando *kit GoTaq Hot Star* (Promega). As sequências dos *primers* utilizados nestas reações encontram-se na tabela 2 e as condições de cada uma das reações são mostradas na tabela 3. Devido à alta identidade entre sequências codificadoras de MIC17A, MIC17B e MIC17C, os *primers* foram desenhados na região promotora do gene (UTR 5') ou após o códon de parada (UTR 3').

Tabela 2. Sequência dos *primers* utilizados para amplificação de regiões codificadoras de proteínas contendo domínios PAN/Apple.

Nome do <i>primer</i>	Sequência (5' - 3')
V1	CCACGCCTCTTGTCCATCCA
AV2	CTAGCTGGGTGTGAACCAAGT
AR1	ATCTAATCTGAATGCTATCACCT
R2	GCCAGCAACACATATTTCTTGT
VAz1	CAGAAAGCATGTGGGTGGTAAGT
Az2	GGTACATAGAATATGTGGCTGCA

Tabela 3. Condições de reação da polimerase em cadeia (PCR) para amplificação das sequências codificadoras de proteínas contendo domínios PAN/Apple.

Genes Apple a serem amplificados	NCLIV_038120 (MIC17A)	NCLIV_038110 (MIC17B)	NCLIV_038100 (MIC17C)	NCLIV_068830
primer <i>forward</i>	V1	AR1	AR1	VAz1
primer reverso	AV2	R2	AV2	Az2
temp. anelamento	58°C	54°C	54°C	58°C
produto de PCR	1593 pb	1543 pb	1407 pb	1662 pb

Os produtos da amplificação foram conferidos por eletroforese em gel de agarose e as sequências de MIC17A, MIC17B e MIC17C foram selecionadas para amplificação com o *HotStar HiFidelity Polymerase Kit* (Qiagen), utilizando na PCR os mesmos *primers* descritos na tabela 3, e as temperaturas de anelamento 54°C para MIC17A e MIC17B, e 51°C para MIC17C. Os produtos de PCR foram separados por eletroforese em gel de agarose e purificados com o *illustra GFX PCR DNA and Gel Band Purification Kit* (GE Healthcare), segundo orientações do fabricante.

Purificadas, as sequências amplificadas de MIC17A, MIC17B e MIC17C foram então ligadas ao plasmídeo pGEM (seção 3.1.5). Após incubação a 4°C por 16 horas, células eletrocompetentes TOP10 foram transformadas com os plasmídeos previamente ligados (seção 3.1.7) e selecionadas por plaqueamento em meio LB agar contendo 75 µg/mL de ampicilina, adicionado em sua superfície de 2 mg de X-

Gal (5-bromo-4-cloro-3-indolil-beta-D-galactopiranosideo, Promega). A placa foi incubada a 37°C por 16 horas e colônias brancas foram selecionadas para crescimento em meio LB com 50 µg/mL de ampicilina, para posterior purificação dos plasmídeos por miniprep (seção 3.1.8). Os plasmídeos purificados foram digeridos para conferência da presença de insertos em reações contendo aproximadamente 2 µg do plasmídeo, 10 U da enzima EcoRI, 2,5 µL de tampão para EcoRI, e água Milli-Q autoclavada suficiente para 25 µL. A digestão foi realizada a 37°C por 16 horas, e o resultado foi verificado por eletroforese em gel de agarose. Os plasmídeos contendo os insertos de MIC17A, MIC17B e MIC17C foram enviados para sequenciamento com os *primers* complementares às regiões do pGEM, T7 e SP6, no laboratório de Biologia Molecular da Prof^a. Maria Helena Goldman, da FFCLRP-USP.

Devido às sequências serem muito longas, o sequenciamento não foi completo. Assim, os mesmos plasmídeos foram novamente enviados para sequenciamento, desta vez no laboratório LAEG do Hemocentro de Ribeirão Preto, utilizando *primers* para regiões mais internas das sequências, descritos na tabela 4.

Tabela 4. *Primers* para sequenciamento dos fragmentos de MIC17A, MIC17B e MIC17C inseridos em plasmídeo pGEM. No nome dos *primers* “ *f* ” designa orientação *forward* e “ *r* ” orientação reversa.

Gene	Nome do <i>primer</i>	Sequencia (5' - 3')
NCLIV_038120 (MIC17A)	Vaz1_ <i>f</i>	CAGAAAGCATGTGGGTGGTAAGT
	AppleVeAm_ <i>r</i>	TCAGCATGTTACATCCCCTGCTT
NCLIV_038110 (MIC17B)	AR1_ <i>f</i>	ATCTAATCTGAATGCTATCACCT
	AppleVeAm_ <i>f</i>	TGCCTTGAAGCAGGCAAAGAGT
	AppleVeAm_ <i>r</i>	TCAGCATGTTACATCCCCTGCTT
NCLIV_038100 (MIC17C)	AR1_ <i>f</i>	ATCTAATCTGAATGCTATCACCT
	AppleRo_ <i>f</i>	TGCATGGTGCAAGGCAAAGAGTA
	AppleRo_ <i>r</i>	TCAGCATGTTACATCGCCTCCTT

3.2.3 Predição da estrutura terciária dos domínios Apple

A predição estrutural dos domínios Apple que compõem as proteínas de *N. caninum* MIC17A, MIC17B e MIC17C foi feita por modelagem molecular por homologia utilizando o programa Modeller 9.11⁶⁷, com base nas estruturas

cristalográficas de proteínas homólogas previamente depositadas no RCSB Protein Data Bank⁶⁸. Além da predição da estrutura de cada domínio isolado, foram realizadas predições das três proteínas Apple completas, incluindo os quatro domínios presentes em cada uma.

3.2.3.1 Alinhamento sequencial entre os domínios das proteínas MIC17 e o domínio A5 de *T. gondii*

O alinhamento entre as sequências de aminoácidos de todos os domínios Apple das proteínas MIC17A, MIC17B e MIC17C e o domínio Apple A5 de MIC4 de *T. gondii*, cuja estrutura foi elucidada e depositada sob código 2LL3 no PDB, foi realizado com o auxílio do programa Multalin⁶⁹. O resultado foi visualizado com base na distribuição das estruturas secundárias encontradas no modelo 2LL3 com o auxílio do programa ESPript (<http://esript.ibcp.fr>).

3.2.3.2 Seleção de modelos estruturais de domínios Apple

As estruturas cristalográficas de quatro domínios Apple, pertencentes a proteínas de outros organismos e previamente depositadas no PDB, foram inicialmente selecionadas como modelos para a predição das estruturas de interesse. A tabela 5 contém as informações sobre cada um dos modelos: código PDB, proteína e organismo ao qual pertencem, e o domínio Apple selecionado. A figura 10 exibe as sequências de aminoácidos dos quatro modelos, alinhadas pelo programa Multalin⁶⁹ e visualizadas com as estruturas secundárias de 2LL3 no programa *on line* ESPript (<http://esript.ibcp.fr>).

Tabela 5. Modelos selecionados para a predição das estruturas dos domínios Apple das proteínas MIC17A, MIC17B e MIC17C por modelagem molecular baseada em homologia.

identificação do modelo	código PDB	proteína	organismo	domínio Apple	cadeia	resíduos
2YIL	2YIL	SML-2	<i>Sarcocystis muris</i>	A1	C	9-78
2LL3	2LL3	MIC4	<i>Toxoplasma gondii</i>	A5	A	10-79
2F83_2	2F83	F XI	<i>Homo sapiens</i>	A2	A	92-175
2F83_3	2F83	F XI	<i>Homo sapiens</i>	A3	A	182-265

Figura 10 Sequências de aminoácidos dos modelos selecionados para a predição das estruturas dos domínios Apple das proteínas MIC17A, MIC17B e MIC17C. O alinhamento das sequências foi realizado pelo programa *on line* Multalin⁶⁹ e visualizado com as estruturas secundárias de 2LL3 no programa *on line* ESript (<http://esript.ibcp.fr>). 2YIL, domínio A1 da proteína SML2 de *Sarcocystis muris*; 2LL3, domínio A5 da proteína MIC4 de *Toxoplasma gondii*; 2F83_2 e 2F83_3, domínios A2 e A3, respectivamente, do fator de coagulação XI de *Homo sapiens*. Os retângulos azuis delimitam as regiões conservadas por similaridade. Os resíduos conservados por identidade são destacados em vermelho.

As estruturas foram avaliadas com relação a estereoquímica, avaliada pelo software *Molprobit*⁷⁰ com a ferramenta *Analyze geometry without all-atom contacts* e com relação a flexibilidade estrutural baseada no cálculo do RMSD (desvio médio quadrático) de carbonos-alfa (C α) entre os modelos, que foi calculado por alinhamento estrutural com o software PYMOL⁷¹. O modelo 2LL3 foi usado como referência para o cálculo de RMSD de C α entre resíduos. O perfil de distribuição do RMSD em função dos resíduos de aminoácido foi realizado no programa Protein3Dfit⁷².

3.2.3.3 Modelagem molecular com o software Modeller 9.11⁶⁷

A modelagem foi realizada para cada um dos quatro domínios das proteínas MIC17A, MIC17B e MIC17C. Os domínios foram denominados A1, A2, A3 e A4, e a estrutura resultante recebe a descrição do domínio combinada à proteína (ex. a1m17a para o domínio A1 de MIC17A). Foram seguidas as instruções do Tutorial *Advanced Modeling*, para modelagem baseada em múltiplos modelos. No alinhamento múltiplo entre modelos (*salin.py*), as cisteínas C1 foram alinhadas manualmente. O alinhamento resultante dos modelos com a sequência a ser modelada foi conferido no *site* Esript (<http://esript.ibcp.fr>). Foram gerados 5 modelos para cada sequência de interesse e, após verificação da qualidade da estereoquímica no *site* Molprobit, as melhores estruturas foram selecionadas para o refinamento.

3.2.3.4 Refinamento das estruturas geradas usando REFMAC/CCP4⁷³

As estruturas geradas foram refinadas usando a ferramenta *Structure idealization* do programa REFMAC⁷³ e foram feitos 10 ciclos de idealização para cada estrutura.

A melhora na qualidade estereoquímica da estrutura foi conferida no *site* Molprobability⁷⁰.

3.2.3.5 Cálculo da superfície eletrostática das estruturas preditas

As estruturas preditas para os quatro domínios Apple (A1, A2, A3 e A4) das proteínas MIC17A, MIC17B e MIC17C tiveram suas superfícies de potencial eletrostático calculadas com o software *Adaptive Poisson-Boltzmann Solver* (APBS Web Solver⁷⁴). Primeiramente, o arquivo PDB de cada domínio foi convertido para arquivo PQR, selecionando campo de força AMBER, através do servidor PDB2PQR⁷⁵. Neste mesmo site foi realizado então o cálculo de superfície de potencial com APBS Web Solver, e o resultado foi visualizado no programa PYMOL, com intervalo de carga delimitado entre -5 e +5.

3.2.3.6 Predição de *folding* de cada proteína, reunindo os quatro domínios

Com a sequência de aminoácidos que compreende os quatro domínios de cada uma das três proteínas MIC17A, MIC17B e MIC17C, excluindo-se a região N terminal que determina o peptídeo sinal, foi feita a predição de *folding* entre estes domínios por modelagem molecular, usando como modelo os quatro domínios Apple do fator de coagulação XI (2F83 - resíduos 1 a 356). Desta vez foi utilizado o Tutorial *Modeller Basic*, para modelagem baseada em apenas um modelo. O alinhamento entre as sequências das proteínas de *N. caninum* com a sequência de 2F83 foi editado manualmente, com base no alinhamento múltiplo descrito anteriormente (salign.py), para que fosse mantido o alinhamento de regiões importantes para a formação de estruturas secundárias.

3.2.4 Amplificação das sequências de MIC17B₂₀₋₃₀₃ e MIC17C₂₀₋₃₄₇

As proteínas hipotéticas MIC17B (NCLIV_038110) e MIC17C (NCLIV_038100) foram selecionadas para expressão em suas formas recombinantes. MIC17A (NCLIV_038120), além de possuir 98% de identidade com a sequência genômica de MIC17C, apresenta menor expressão gênica que esta, segundo dados do banco ToxoDB, portanto não foi selecionada para expressão.

Novamente foi extraído o RNA total de taquizoítas de *N. caninum* (seção 3.2.2), porém o cDNA foi produzido com o *primer* reverso NcR2, complementar à região 3'UTR de MIC17C (tabela 6) e com a enzima transcriptase reversa *SuperScript III*

(Invitrogen). Primeiramente, 2 µg de RNA total, 20 pmol de *primer* NcR2 e 0,5 mM de cada dNTP foram desnaturados a 65°C por 5 minutos, e então adicionados de 1x tampão para *SuperScript III*, 1 µL de DTT 0,1 M, 1 µL da enzima inibidora de RNase (*RNAse Out*), e 1 µL da enzima *SuperScript III*. A síntese de cDNA ocorreu a 50°C por 45 minutos, e a inativação da enzima foi feita a 70°C por 15 minutos. Para a degradação do RNA restante, 1 µL da enzima *RNAse H* foi adicionado e a reação incubada a 37°C por 20 minutos.

Para a construção da proteína recombinante MIC17B, a região que engloba os três domínios PAN/Apple, A1, A2 e A3, (posições 58 a 909 da sequência nucleotídica) foi amplificada com o par de *primers* NcF2/NcR1; e, para a recombinante MIC17C, foi amplificada a região que abrange os três domínios PAN/Apple, A1, A2 e A3, e quase totalmente o domínio Apple-like, A4, (posições 58 a 1042 da sequência nucleotídica) com o par de *primers* NcF2/NcR4; resultando nas sequências que codificam as proteínas recombinantes MIC17B₂₀₋₃₀₃ e MIC17C₂₀₋₃₄₇. Os *primers* utilizados para a amplificação dos fragmentos e posterior inserção em plasmídeo de expressão foram adicionados em sua extremidade 5' de sítios de restrição enzimática e de uma sequência extra de três citosinas para facilitar a ligação da enzima ao sítio de clivagem (tabela 6).

Tabela 6. *Primers* utilizados para a amplificação de fragmentos das proteínas MIC17B e MIC17C a serem expressados. NcR2, para obtenção de cDNA; NcF2_BamHI e NcR1_HindIII para amplificação MIC17B₂₀₋₃₀₃; e NcF2_Bam HI e NcR4_Hind III para amplificação de MIC17C₂₀₋₃₄₇. As sequências referentes aos sítios de restrição encontram-se sublinhadas.

<i>Primer</i>	Orientação	Sequência (5' - 3')
NcR2	reverso	GCCTGTCCCGCTGTACCGAAA
NcF2_BamHI	<i>forward</i>	CCCGGATCCGCGGTGTTGCTGCAGGATGTTAC
NcR1_HindIII	reverso	CCCAAGCTTCACTTCCGTCTTGGGGTAGGTG
NcR4_HindIII	reverso	CCCAAGCTTCCCCTGCTTTGGAGTTCGC

A PCR foi realizada com a enzima *Taq Polimerase Expand High Fidelity Plus* (Roche) com temperatura de anelamento de 59°C, e o resultado das amplificações conferido por eletroforese em gel de agarose. Os produtos obtidos foram purificados após excisão das bandas do gel com o *illustra GFX PCR DNA and Gel Band Purification Kit* (GE Healthcare), e os produtos purificados também foram conferidos em gel de agarose.

3.2.5 Clonagem de MIC17B₂₀₋₃₀₃ e MIC17C₂₀₋₃₄₇ em plasmídeo de expressão

Antes da ligação ao plasmídeo de expressão pET28 (Novagen), os fragmentos de MIC17B₂₀₋₃₀₃ e MIC17C₂₀₋₃₄₇ foram pré-clonados em plasmídeo pGEM. A transformação em células TOP10, seleção de colônias, e purificação de plasmídeos foi feita da mesma forma descrita na seção 3.2.2.

Os plasmídeos foram então digeridos com as enzimas BamHI e HindIII (Fermentas) para a formação de extremidades complementares aos sítios de restrição do plasmídeo de expressão. Aproximadamente 0,8 µg de cada plasmídeo pGEM purificado foram digeridos com 10 U de BamHI e 10 U de HindIII, na mesma reação, em tampão Orange 1X, num volume final de 20 µL, por 3 horas a 37°C. Os fragmentos digeridos foram purificados após eletroforese em gel de agarose e novamente conferidos em gel de agarose.

Os plasmídeos pET28 foram submetidos a uma seleção por meio de “auto-ligação”, processo este que aumenta posteriormente o sucesso de ligação aos fragmentos. Para isso, 8,0 µg de pET28 foram digeridos com 27 U de EcoRI em tampão Orange (Fermentas), a 37°C por 2 horas. Após purificação, os plasmídeos digeridos foram religados (“auto-ligação”) com a enzima *T4 DNA Ligase* (Fermentas) em reação contendo 10 mM de DTT e 0,5 mM de ATP, incubada a 4°C por 16 horas. Células eletrocompetentes TOP10 foram transformadas com a ligação e plaqueadas em meio LB ágar contendo 75 µg/mL de kanamicina. Após incubação da placa a 37°C por 16 horas, colônias foram selecionadas para crescimento em meio LB com 50 µg/mL de kanamicina, para posterior purificação dos plasmídeos por miniprep. Após a seleção, aproximadamente 0,8 µg de pET28 foram digeridos com as enzimas BamHI e HindIII em tampão Orange, por 3 horas a 37°C, purificados e conferidos em gel de agarose.

Os fragmentos de MIC17B₂₀₋₃₀₃ e MIC17C₂₀₋₃₄₇ foram então ligados em plasmídeo pET28 selecionado e digerido em reação contendo 1,5 µL de pET28, 3 µL do inserto, 1 µL de *T4 DNA Ligase* (Fermentas), 10 mM de DTT, 0,5 mM de ATP, em tampão para *T4 DNA Ligase*. A reação foi incubada a 4°C por 48 horas, e a ligação usada para a transformação de células TOP10. As células transformadas foram selecionadas com kanamicina (como descrito acima para a seleção por “auto-ligação” de pET28), e os plasmídeos foram isolados por miniprep e purificados de gel de agarose. Foi realizada então a digestão dos plasmídeos simultaneamente

com as enzimas BamHI e HindIII para a verificação, em gel de agarose, da presença dos insertos corretos.

3.2.6 Expressão das proteínas recombinantes MIC17B₂₀₋₃₀₃ e MIC17C₂₀₋₃₄₇

Células eletrocompetentes BL21 foram transformadas com os plasmídeos pET28 contendo os genes de MIC17B₂₀₋₃₀₃ ou com pET28 contendo MIC17C₂₀₋₃₄₇. As colônias foram selecionadas com kanamicina e os plasmídeos isolados por miniprep.

A expressão das proteínas recombinantes foi realizada primeiramente em pequena escala, para testes de extração com diferentes soluções. Células de BL21 transformadas com os pET28 contendo MIC17B₂₀₋₃₀₃ ou MIC17C₂₀₋₃₄₇ foram inoculadas em 100 mL de meio LB adicionado de 50 µg/mL de kanamicina, e incubadas a 37°C por 16 horas, sob agitação. Destas suspensões, 3,5 mL foram inoculados em 60 mL de LB com 50 µg/mL de kanamicina para crescimento das células por aproximadamente 3 horas a 37°C, sob agitação, até atingirem OD₆₀₀ entre 0,5 e 0,8. Dos 60 mL de cada cultura, 45 mL foram induzidos à expressão de proteínas com 5,0 mM de IPTG (isopropiltiogalactosídeo), e 15 mL não receberam IPTG, como controle não-induzido. Tanto as culturas induzidas como não-induzidas foram incubadas a 37°C por 2 horas, sob agitação, e então centrifugadas a 4000 x g por 15 min. Os *pellets* das culturas induzidas com IPTG foram ressuspensos em 4,5 mL de PBS, dos quais, 1/3 foi centrifugado e submetido à extração total com 0,5 mL de tampão ureia 8 M (ureia 8 M, NaH₂PO₄ 100 mM, Tris-Cl 100 mM) com sonicação do *pellet* em equipamento *Sonic Dismembrator 100* (Fisher Scientific). Após a sonicação, o conteúdo foi centrifugado a 4000 x g por 10 minutos, o sobrenadante transferido para outro tubo, e o processo de extração repetido com mais 0,5 mL de ureia 8 M sobre o *pellet* remanescente. Os outros 2/3 da cultura induzida, ressuspensa em 4,5 mL, foram centrifugados e o *pellet* submetido à extração sequencial com diferentes soluções, ordenadas em poder crescente de extração de proteínas insolúveis. Portanto, primeiro o *pellet* foi submetido à extração com Tris-NaCl 100 mM pH 7,5 com duas vezes de 0,5 mL, sob sonicação. Separado o sobrenadante, o *pellet* remanescente foi ressuspensa em detergente Triton X-100 2,0% e também sonicado duas vezes com 0,5 mL. O *pellet* insolúvel foi então submetido a duas extrações sob sonicação com 0,5 mL de tampão ureia 8M

(esquema na figura 11). O sobrenadante de cada extração foi verificado por eletroforese em gel de acrilamida 1D com coloração por Coomassie Blue G-250.

Figura 11. Metodologia para expressão heteróloga e extração de MIC17B₂₀₋₃₀₃ e MIC17C₂₀₋₃₄₇. Células de *E. coli* BL21 transformadas com plasmídeo pET28 contendo um dos insertos foram induzidas com 5.0 mM de IPTG. Um controle não-induzido foi submetido à extração total de proteínas em tampão ureia 8M. Do conteúdo induzido, uma parte foi submetida diretamente à extração com tampão ureia 8M, e outra à extração sequencial com Tris-NaCl 100mM pH 7,5, em seguida com Triton X-100 2%, e finalmente com tampão ureia 8M. Símbolo ☉ = centrifugação a 4000 x g por 10 minutos; sonic. = sonicagem.

Outro teste de extração foi realizado para verificar a solubilização das proteínas recombinantes em diferentes concentrações de ureia. A indução à expressão foi feita como descrito anteriormente (figura 11), e a extração sequencial com Tris-NaCl e Triton também antecedeu a extração final com cada uma das concentrações de tampão ureia: 2M, 4M, 6M e 8M, todos os tampões de ureia contendo 100 mM de NaH₂PO₄ e 100 mM de Tris-Cl, em pH 8,0. O sobrenadante de cada extração também foi verificado por eletroforese em gel de acrilamida 1D corado com Coomassie Blue G-250.

Para a indução à expressão em grande escala de MIC17B₂₀₋₃₀₃ e MIC17C₂₀₋₃₄₇, as células BL21 transformadas com os plasmídeos pET28 contendo seus respectivos insertos foram inoculadas em 50 mL de meio LB com 50 µg/mL de kanamicina. Após incubação por 16 horas a 37°C, sob agitação, 15 mL desta suspensão foram inoculados em 750 mL de meio LB adicionado de 50 µg/mL de kanamicina. O crescimento a 37°C, sob agitação, por 3 horas levou a uma OD₆₀₀ de 0,55. Todo o conteúdo da cultura foi submetido à indução de expressão com 5.0 mM de IPTG a 37°C, sob agitação, por 2 horas, e então centrifugado dividido em 8 tubos *falcon* a

4000 x g, por 15 minutos. Seis destes tubos tiveram seus *pellets* transferidos para tubos *ependorf* de 1,5 mL e, para cada um deles, procedeu-se a extração sequencial como demonstrado na figura 11, utilizando ao final tampão ureia 8 M.

3.2.7 Purificação, dessalinização e concentração das proteínas recombinantes

Por conterem uma cauda de aminoácidos histidina, adicionada no momento da expressão em plasmídeo pET28, as proteínas recombinantes são passíveis de purificação pela técnica de IMAC (Cromatografia de Afinidade por Íons Metálicos Imobilizados), na qual as proteínas são adsorvidas à uma matriz de agarose carregada com íons Ni^{2+} . A dessorção das proteínas foi realizada por adição do agente competidor imidazol, que possui maior afinidade pelo sítio de ligação do que a molécula que está adsorvida. Isto faz com que a biomolécula seja dessorvida para o meio, enquanto o competidor permanece adsorvido à matriz⁷⁶.

Foram usados 0,75 mL de resina *Ni Sepharose 6 Fast Flow* (GE Healthcare) para o empacotamento de colunas vazias PD 10 (GE Healthcare). A resina foi lavada com 3,0 mL de tampão ureia 8 M em pH 8,0 adicionado de 10 mM de imidazol e, com a base da coluna fechada, procedeu-se a adição de 3,0 mL da amostra de proteína produzida em grande escala, adicionada de 10 mM de imidazol, para evitar a adsorção de outras proteínas que não as recombinantes. A coluna foi tampada e mantida na horizontal sob agitação a 150 rpm, em temperatura ambiente, por 16 horas, para adsorção da proteína à matriz.

Para a dessorção das proteínas recombinantes, primeiramente foi coletada a fração não ligada, e a coluna foi lavada com 5 mL de tampão ureia 8M pH 8,0 contendo 10mM de imidazol. As proteínas recombinantes foram eluídas com 10 mL de tampão ureia 8 M pH8,0 contendo 500 mM de imidazol, em 10 frações de 1,0 mL. A presença de proteína nas frações foi verificada em reagente de Bradford e por eletroforese em gel de acrilamida corado com prata.

As proteínas purificadas foram submetidas à dessalinização para a remoção de imidazol, que é incompatível com a imunização dos coelhos. Colunas PD-10 empacotadas com *Sephadex G-25 Medium* (GE Healthcare) foram equilibradas com 25 mL de tampão ureia 8M pH 8,0, e receberam 2,5 mL de proteína recombinante purificada. Após a coleta da fração não ligada, as proteínas foram eluídas em 10 frações de 1,0 mL de tampão ureia 8M pH 8,0. A presença de proteína nas frações foi verificada com reagente de Bradford e eletroforese em gel de acrilamida, com

coloração em prata. Após a dessalinização, as amostras foram quantificadas pelo método de Bradford e o seu volume reduzido para atingir a concentração desejada para a imunização. As amostras foram concentradas em sistema *Amicon Ultra – Centrifugal Filter Devices 5K* (Millipore) por centrifugação a 4000 x g, e a concentração final das proteínas foi calculada pelo método de Bradford.

3.2.8 Confirmação da identidade das proteínas recombinantes por espectrometria de massas

As proteínas recombinantes MIC17B₂₀₋₃₀₃ e MIC17C₂₀₋₃₄₇ foram separadas por eletroforese em gel de acrilamida 1D corado com Coomassie Blue G-250, e suas bandas foram excisadas para identificação por espectrometria de massas. As proteínas foram digeridas com tripsina (seção 3.1.15) e os peptídeos analisados em espectrômetro LTQ-Orbitrap-XL (3.2.14.1). A busca com os resultados foi feita contra o banco de dados anotado de *N. caninum* (ToxoDB versão 6.4), incluindo uma lista de contaminantes comuns, usando o programa Mascot 2.3.02 (Matrix Science). A busca foi limitada com tolerância de 0,9 Da tanto para peptídeo quanto para fragmentos de íon, e mínimo score de 20.

3.2.9 Preparo de extrato total de *N. caninum* (NcET) para produção de anticorpos

O extrato total foi obtido por sonicação de 4×10^8 taquizoítas de *N. caninum* em tampão ureia 8M pH8,0. Após a sonicação, foi realizada centrifugação a 4000 x g, por 10 minutos a 4°C, e o sobrenadante (NcET) quantificado pelo método de Bradford.

3.2.10 Produção de anticorpos contra MIC17B₂₀₋₃₀₃, MIC17C₂₀₋₃₄₇ e NcET

Os anticorpos foram produzidos em coelhos machos da linhagem Nova Zelândia contra as proteínas recombinantes MIC17B₂₀₋₃₀₃ e MIC17C₂₀₋₃₄₇, e contra o extrato total de taquizoítas de *N. caninum* (NcET). Um total de quatro imunizações, a intervalos de três semanas, foi feito com aproximadamente 200-300 µg do antígeno; sendo a primeira imunização com o antígeno emulsificado em Adjuvante Completo de Freund (ACF), injetado por via subcutânea, e as três imunizações seguintes em Adjuvante Incompleto de Freund (AIF), por via intramuscular. O sangue pré-imune foi coletado no dia da primeira imunização e, a cada imunização, foi coletado o

sangue referente à imunização anterior. Antes da coleta do sangue e da imunização, cada animal foi sedado com acepromazina (0,25 mg/kg), por via intramuscular e, ao final de todo o procedimento, a eutanásia dos animais ocorreu por choque hipovolêmico, com anestesia prévia com ketamina (25 mg/kg) e xilazina (5 mg/kg), via intramuscular, seguida de colheita de sangue por punção cardíaca. Todos os procedimentos envolvendo animais foram aprovados pela Comissão de Ética no Uso de Animais (CEUA) do Campus Ribeirão Preto da Universidade de São Paulo (Processo no. 08.1.1134.53.0).

3.2.11 Detecção de anticorpos por ELISA

A detecção dos anticorpos produzidos foi realizada ao longo das imunizações, por ELISA (*enzyme-linked immunosorbent assay*). Os antígenos (0,05 µg/mL das proteínas recombinantes ou 5,0 µg/mL do extrato total) foram diluídos em tampão carbonato ou *coating buffer* (Na₂CO₃ 0,06 M pH 9,5) e adsorvidos à placa de ELISA (96 poços, NUNC MaxiSorp) por incubação a temperatura ambiente por 16 horas. A placa foi lavada 3 vezes com PBS-T (PBS com 0,05% *Tween 20*, Sigma-Aldrich) e bloqueada com PBS-GT (PBS-T com 2 mg/mL de gelatina suína, Sigma) a 37°C por 1 hora. Após lavagem por três vezes com PBS-T, o soro foi adicionado na diluição 1:12000, em PBS-GT, para a detecção das proteínas recombinantes e na diluição 1:4000 para a detecção das proteínas do extrato total. A placa foi incubada a 37°C por 1 hora, e novamente lavada três vezes com PBS-T. O anticorpo secundário de cabra anti-coelho conjugado à enzima peroxidase (*Goat Anti-Rabbit IgG (H+L) HRP Conjugate ZyMax*, Invitrogen) foi adicionado na diluição, em PBS-GT, de 1:20000 para detecção de proteínas recombinantes e 1:10000 para a detecção de extrato total. A placa foi incubada a 37°C por 1 hora e, em seguida, lavada por três vezes com PBS-T. O reagente TMB (*TMB (ELISA) Single solution. Ready-to-use*, Zymed-Invitrogen) foi adicionado para a detecção da atividade de peroxidase, em reação incubada a temperatura ambiente, no escuro, por 15 minutos. A reação foi interrompida pela adição de HCl 2 M e a absorbância de cada poço foi lida a 450 nm.

3.2.12 Purificação de anticorpos

Os anticorpos contra as proteínas recombinantes MIC17B₂₀₋₃₀₃ e MIC17C₂₀₋₃₄₇ foram purificados dos soros, coletados três semanas após a terceira imunização dos coelhos, por cromatografia de afinidade⁷⁷. As proteínas recombinantes foram separadas por SDS-PAGE 1D em géis de 18 cm x 16 cm (sistema SE600 Ruby, GE Healthcare). Neste caso, ao invés de em poços, as proteínas foram adicionadas (750 µg de MIC17B₂₀₋₃₀₃; 560 µg de MIC17C₂₀₋₃₄₇) juntamente com tampão de corrida, em uma tira de 14 cm x 1 cm de papel de filtro para *Western blot* (GE Healthcare), posicionada sobre o gel. O marcador de peso molecular *Precision Plus Protein Dual Color* (BioRad) foi adicionado em outra tira e posicionado ao lado da tira com proteína. As tiras foram seladas com tampão de agarose e a corrida eletroforética desenvolvida com amperagem de 15 mA por gel por 40 minutos, e 30 mA por gel no restante da corrida, de aproximadamente 4 horas. Terminada a corrida, as proteínas foram transferidas para membranas de nitrocelulose em sistema semi-seco (seção 3.1.14). As membranas foram coradas com solução de DB71 e a banda correspondente a cada proteína recombinante foi recortada, descolorida e bloqueada com 0,8% de gelatina suína (Sigma-Aldrich) dissolvida em solução de TBS-T (20 mM Tris, 500 mM NaCl, 0,05% Tween20, pH 7,4), por 1 hora, a temperatura ambiente, sob agitação. As membranas foram lavadas três vezes por 5 minutos com TBS-T e então incubadas com os seus respectivos anti-soros (1,0 mL do soro anti-MIC17B₂₀₋₃₀₃; 1,5 mL do soro anti-MIC17C₂₀₋₃₄₇), por 16 horas, a 4°C, sob agitação a 70 rpm. Os anticorpos específicos foram eluídos das membranas por incubação com 900 µL de tampão de eluição (Tris 50 mM, glicina 50 mM, pH 2,6), alternando-se por 10 minutos entre gelo, agitação a 160 rpm, e vortex. O eluído foi transferido para um novo tubo e imediatamente neutralizado com 100 µL de tampão Tris 1 M pH 8,0. O processo foi repetido com mais 900 µL de tampão de eluição que, ao final, foi neutralizado com 100 µL de tampão Tris 1 M pH 8,0, e transferido para o tubo com o primeiro eluído. Alíquotas das soluções contendo os anticorpos purificados foram separadas para quantificação por Bradford e verificação em gel de acrilamida 1D, e todo o conteúdo restante foi estabilizado com 1 mg/mL de BSA, que funciona como antioxidante.

3.2.13 Detecção das proteínas recombinantes e nativas por *Western blot* (WB)

3.2.13.1 Reconhecimento das proteínas recombinantes pelos anticorpos purificados

A especificidade dos anticorpos purificados a cada uma de suas proteínas recombinantes MIC17B₂₀₋₃₀₃ e MIC17C₂₀₋₃₄₇ foi checada pela técnica de *Western blot* unidimensional (WB 1D). MIC17B₂₀₋₃₀₃ ou MIC17C₂₀₋₃₄₇ (0,25 µg) foram separadas por eletroforese em gel de acrilamida 1D e transferidas para membrana de PVDF. Depois de bloqueada, 3 ng/mL de anticorpo purificado anti-MIC17B₂₀₋₃₀₃ ou anti-MIC17C₂₀₋₃₄₇ foram incubados com a membrana por 16 horas, que foi então lavada e incubada com o anticorpo secundário de cabra anti-coelho conjugado à peroxidase (*Goat anti-rabbit IgG HRP*, ZyMax, Invitrogen) por 1 hora, na diluição 1:100.000. Após lavagem, a membrana foi revelada como descrito na seção 3.1.14.

3.2.13.2 Preparo dos antígenos: NcET e NcESA_{SN}

O extrato total de *N. caninum* (NcET) foi preparado a partir de $1,6 \times 10^8$ taquizoítas sonicados em duas etapas com 215 µL tampão de amostra (AM: ureia 7M, tiourea 2M e CHAPS 4%) e adicionado de 10 µL de inibidor de protease (*Protease Inhibitor Cocktail*, Sigma). Para ser usado como controle, o extrato total de células Vero (VeET) foi preparado da mesma forma descrita para os taquizoítas, com as células recuperadas de uma garrafa de 25 cm².

O extrato secretado NcESA_{SN} foi preparado a partir dos sobrenadantes da cultura de taquizoítas produzidos durante os 3 dias após infecção (p.i). Para a infecção, os taquizoítas presentes em uma garrafa média (75 cm²) com células Vero há 7 dias, foram purificados em fluxo laminar, contados ($1,5 \times 10^8$) e inoculados em nova garrafa média de células Vero, preparada 7 dias antes, em 13 mL de meio RPMI contendo 10 µg/mL de kanamicina. A garrafa inoculada foi previamente enxaguada por três vezes com meio RPMI para a remoção de SFB do meio de cultura anterior. Após 24 horas de incubação em estufa a 37°C com 5% de CO₂, o sobrenadante da cultura foi coletado (SN₁) e mais 13 mL de meio RPMI/kanamicina foram adicionados à cultura, que foi incubada por mais 24 horas em estufa a 37°C com 5% de CO₂. O sobrenadante SN₁ foi primeiramente filtrado em filtro de 5 µm e, em seguida, em filtro de 0,2 µm. As proteínas do filtrado foram precipitadas por adição do mesmo

volume (13 mL) de ácido tricloroacético (TCA) 30% em acetona contendo 10 mM de DTT, e incubação a -20°C por 1 hora. Após centrifugação a 4000 x g, por 15 min, a 4°C, o *pellet* de proteínas precipitadas foi lavado com acetona contendo 10 mM de DTT, novamente centrifugado por 10 minutos, seco a temperatura ambiente, e finalmente ressuspenso em 150 µL de tampão AM. No segundo dia p.i., o sobrenadante SN₂ foi coletado e substituído por mais 13 mL de meio RPMI/kanamicina, e a garrafa incubada por mais 24 horas. O mesmo tratamento para SN₁ foi realizado para SN₂. No terceiro dia p.i., quando as células Vero já estavam quase totalmente destruídas, o sobrenadante SN₃ foi coletado e tratado da mesma forma que os anteriores, e a garrafa descartada. Todos os passos foram realizados simultaneamente com cultura apenas de células Vero, para a coleta dos SN₁, SN₂ e SN₃, e produção dos controles (VeESA_{SN}) a serem usados no desenvolvimento de *Western blot*. Todos os extratos foram quantificados por método de Bradford e visualizados em gel de acrilamida 1D corado com prata.

3.2.13.3 Detecção das proteínas nativas em NcET e NcESA_{SN} por WB 1D

Os extratos ET e ESA_{SN3} de *N. caninum* (Nc), assim como seus respectivos controles de células Vero (Ve), foram selecionados como antígenos para a detecção das proteínas MIC17B e MIC17C nativas usando os anticorpos purificados anti-MIC17B₂₀₋₃₀₃ e anti-MIC17C₂₀₋₃₄₇. Por eletroforese em gel de acrilamida 1D foram separados 12 µg de NcET, 12 µg de VeET, 12 µg de NcESA_{SN3}, e 2 µg de VeESA_{SN3}, que foram em seguida transferidos para membrana de PVDF. Após bloqueio com 0,8% de gelatina suína, a membrana foi incubada com 0,2 µg/mL de anticorpo purificado anti-MIC17B₂₀₋₃₀₃ ou anti-MIC17C₂₀₋₃₄₇ por 16 horas e, após lavagem, incubada com o anticorpo secundário de cabra anti-coelho conjugado à peroxidase (*Goat anti-rabbit IgG peroxidase*, KPL) por 1 hora. A revelação da membrana foi feita como descrito na seção 3.1.14.

3.2.13.4 Detecção das proteínas nativas em NcET e NcESA_{SN} por WB 2D

O ponto isoelétrico (pI) teórico das proteínas MIC17B e MIC17C é de 8,0 a 8,5, porém, inicialmente, buscou-se uma distribuição em faixa de pH ampla das proteínas do extrato total de *N. caninum* para a detecção das proteínas nativas. Assim, 50 µg de extratos totais de *N. caninum* e de células Vero (NcET e VeET) foram separados

por eletroforese bidimensional, cuja separação em primeira dimensão foi feita com *strips* de 7 cm e pH imobilizado 3-11 NL. Após a corrida em segunda dimensão, as proteínas foram transferidas para membrana PVDF e detectadas com 0,12 µg/mL dos anticorpos purificados anti-MIC17B₂₀₋₃₀₃ ou anti-MIC17C₂₀₋₃₄₇. Utilizou-se como anticorpo secundário o anticorpo de cabra anti-coelho conjugado à peroxidase (*Goat anti-rabbit IgG HRP*, ZyMax, Invitrogen) na diluição 1:20000, e a revelação foi desenvolvida como descrito em seção 3.1.14.

Novos *Western blots* 2D foram realizados focando-se a faixa de pI predita para as proteínas nativas, e incluindo o extrato secretado como antígeno. Para tal, os extratos ET e ESA_{SN3} de *N. caninum* e de Vero foram separados por eletroforese bidimensional usando *strips* de 7 cm e pH imobilizado de 7-11 NL. As *strips* de NcET, VeET e NcESA_{SN3}, foram preparadas com 70 µg do extrato, enquanto a *strip* de VeESA_{SN3}, foi preparada com 5 µg, devido à escassez do material. As proteínas do gel 2D foram transferidas para membrana de PVDF que, após bloqueio, foram incubadas com o anticorpo purificado anti-MIC17B₂₀₋₃₀₃ e anti-MIC17C₂₀₋₃₄₇ nas seguintes concentrações: 0,2 µg/mL nas membranas com NcESA_{SN3} e VeESA_{SN3}; e 0,25 µg/mL nas membranas com NcET e VeET. O anticorpo secundário (*Goat anti-rabbit IgG peroxidase*, KPL) foi usado nas diluições 1:15000 no caso de NcESA_{SN3} e VeESA_{SN3}, e 1:12000 para NcET e VeET. As membranas foram reveladas como descrito na seção 3.1.14.

3.2.14 Identificação das proteínas nativas por espectrometria de massas (MS)

3.2.14.1 MS por MALDI, de proteínas separadas por gel 1D e 2D, em ET e ESA_{SN}

Géis uni e bidimensionais foram preparados com 25 µg por poço das amostras NcET e NcESA_{SN3} (gel 1D) e 150 µg dos extratos NcET, VeET e NcESA_{SN3} (gel 2D *strips* de 7 cm, pH 7-11 NL). De cada gel foram excisados as bandas ou os *spots* das regiões correspondentes aos sinais detectados por *Western blot* 1D ou 2D com os anticorpos anti-MIC17B₂₀₋₃₀₃ e anti-MIC17C₂₀₋₃₄₇. As amostras foram preparadas para a análise por espectrometria de massas como descrito na seção 3.1.15 e analisadas em diferentes laboratórios, listados na tabela 7. A identificação de proteínas foi realizada contra o banco de dados anotado de *N. caninum* (ToxoDB versão 6.4; 7344 entradas). Nas análises por MALDI, foi acessado o programa

Mascot 2.3.02 (Matrix Science) licenciado para o Laboratório Nacional de Luz Sincrotron (LNLS) e, para a busca, selecionou-se tolerância para íons precursores de 0.6 Da e para fragmentos de íons de 100 ppm. No caso das análises em LTQ-Orbitrap-XL, a busca foi realizada no Mascot 2.3.02 por meio do programa *Proteome Discoverer 1.3* (Thermo Fischer Scientific), selecionando-se tolerância de 0.9 Da tanto para íons precursores quanto para fragmentos de íons. Em todas as buscas foram incluídas oxidação de metionina e carbamidometilação de cisteína como modificações variável e fixa, respectivamente.

Tabela 7. Locais de análise, por espectrometria de massas, de proteínas presentes nos extratos total (ET) e secretado (ESA_{SN3}), de *N. caninum* (Nc) e de células Vero (Ve), separados por eletroforese uni e bidimensional. Foram excisados as bandas e os spots das regiões correspondentes aos sinais detectados por *Western blot* 1D e 2D com os anticorpos anti-MIC17B₂₀₋₃₀₃ e anti-MIC17C₂₀₋₃₄₇. *, MALDI-TOF-TOF; **, LTQ-Orbitrap-XL; ***, TOF-TOF Ultraflex II Bruker.

Amostra	Separação	Coloração	Laboratório
NcET	2D	Coomassie	Centro de Química de Proteínas - FMRP - USP*
NcET e VeET	2D	Prata	Biomolecular Mass Spectrometry Group - Utrecht University**
NcET e NcESA _{SN3}	2D	Prata	Central de Espectrometria de Massas - FCFRP - USP***
NcET e NcESA _{SN3}	1D	Coomassie	Central de Espectrometria de Massas - FCFRP - USP***

3.2.14.2 MS por *electrospray*, em ESA

O extrato secretado de *N. caninum* (ESA) analisado por *electrospray* foi estimulado via aumento de Ca²⁺ intracelular, induzido por etanol (seção 3.3). Brevemente, taquizoítas recém-purificados foram incubados em meio RPMI adicionado de 1% de etanol a 37°C por 15 minutos e, em seguida, resfriados por 10 minutos em gelo. O sobrenadante (ESA) foi coletado por centrifugação e as alíquotas foram reunidas para concentração em 700 µL de água Milli-Q, usando concentrador *Amicon Ultra Centrifugal Filter Devices 5K* (Millipore). A amostra foi seca em concentrador a vácuo (*CentriVap Concentrator*, Labconco) e após concentração e secagem a amostra foi levada para análise no Biomolecular Mass Spectrometry Group, Universidade de Utrecht, Holanda. O extrato seco ESA foi ressuspensão em 50 mM de bicarbonato de amônio e separado por eletroforese unidimensional, em gel posteriormente corado com *Coomassie Blue G-250* (*Gel Code Blue Stain Reagent*, Thermo Scientific) e recortado em 13 bandas para digestão e preparo das amostras para análise por espectrometria de massas (seção 3.1.15). Os peptídeos extraídos das 13 bandas foram analisados em sistema com separação por cromatografia líquida de fase reversa acoplado a espectrômetro de massas (RP-nanoLC-MS/MS). Os dados foram

buscados contra o banco de dados anotado de *N. caninum* (ToxoDB 6.4), usando o programa Mascot 2.3.02 (Matrix Science) com auxílio de busca utilizando o programa *Proteome Discoverer 1.3* (Thermo Fischer Scientific).

3.2.15 Imunolocalização *in situ*

Com os anticorpos purificados anti-MIC17B₂₀₋₃₀₃ e anti-MIC17C₂₀₋₃₄₇, as proteínas nativas MIC17B e MIC17C foram localizadas nos taquizoítas de *N. caninum* por microscopia confocal. Por reconhecimento com anticorpos de cabra anti-coelho conjugados a fluoróforos, as proteínas foram visualizados em Microscópio de Fluorescência Confocal Leica TCS SP5 (Laboratório Multiusuário de Microscopia Confocal - FMRP - USP). Os fluoróforos utilizados isotiocianato de fluoresceína (FITC) e succinimidil éster (Alexa 488) têm seus espectros de excitação em 495 nm e emissão em 520 nm, e são visualizados na cor verde. O iodeto de propídeo, que se intercala entre as bases de DNA, foi usado para a marcação de núcleos, e possui excitação em 538 nm e emissão em 617 nm, sendo visualizado na cor vermelha. Foi realizada aquisição sequencial de imagens, para os marcadores não serem excitados simultaneamente. A edição de imagens, para modificação de brilho e contraste, foi realizada no programa *Image J 1.41* (National Institutes of Health, USA).

3.2.15.1 Taquizoítas em cultura de células Vero

Monocamadas de células Vero foram cultivadas em placas de 8 poços com uma câmara divisória removível acoplada à lâmina de microscópio (*Lab-Tek Chamber*, Nunc). As células Vero foram cultivadas por 5 dias e então inoculadas com aproximadamente 10^6 taquizoítas por poço, completados com 250 μ L de meio RPMI. A placa foi incubada em estufa a 37°C com 5% CO₂ por 8 horas e 30 minutos, tempo suficiente para a invasão e para que os taquizoítas ainda permanecessem intracelulares no momento de detecção das proteínas. Entre todos os passos descritos a seguir, três lavagens com PBS foram realizadas em cada poço. Assim, após lavagem, a cultura foi fixada com formaldeído 4% por 30 minutos; as células foram permeabilizadas com 0,1% de Triton X-100 por 20 minutos; e o bloqueio foi feito com 2% de BSA e 20 mM de glicina em PBS, a 4°C por 16 horas.

Os poços foram incubados com 400 µL de anticorpos purificados anti-MIC17B₂₀₋₃₀₃ ou anti-MIC17C₂₀₋₃₄₇ 0,5 µg/mL, diluídos em PBS contendo 0,2% de BSA; e um poço foi incubado apenas com 2% de BSA em PBS, para ser usado como controle de ligação do anticorpo secundário às proteínas. A incubação com os anticorpos primários foi feita a temperatura ambiente, por 1 hora (40 minutos com agitação a 50 rpm e 20 minutos parado). Após três lavagens de 5 minutos a 60 rpm com PBS, foi realizada incubação com o anticorpo secundário de cabra anti-coelho conjugado à FITC (*FITC Goat Anti-Rabbit IgG (H+L) Conjugate*, Zymed, Invitrogen) diluído 1:500 em PBS contendo 0,2% de BSA, no escuro por 1 hora (45 minutos a 50 rpm e 15 minutos parado) a temperatura ambiente. Após duas lavagens com PBS, os núcleos foram marcados com 0,02 µg/µL de iodeto de propídeo (Santa Cruz Biotechnology), por 10 minutos a 50 rpm, no escuro. Os poços foram lavados três vezes, a câmara foi removida e a lâmina preparada em glicerol 60%, coberta por lamínula e selada com esmalte.

3.2.15.2 Taquizoítas purificados

Neste ensaio a localização não foi feita com os taquizoítas em células Vero, mas fixados na placa após processo de purificação (seção 3.1.1) de garrafas parcialmente destruídas. Placas com 2 poços e câmara removível acoplada à lâmina para microscópio (*Lab-Tek Chamber Slide System*, Sigma-Aldrich) foram previamente recobertas com 700 µL de bromidrato de poli-L-lisina 1 mg/mL (Sigma-Aldrich) e incubadas para secagem da solução em estufa a 37°C por 3 horas. Em seguida, foram lavadas duas vezes com água MilliQ e novamente incubadas em estufa para secagem. Em tubo de 1,5 mL, taquizoítas purificados foram fixados com 4% de formaldeído por 25 minutos, centrifugados, e ressuspensos em PBS. Aproximadamente 10⁷ taquizoítas já fixados foram adicionados por poço e incubados a temperatura ambiente por 15 minutos, para aderência à placa. Os taquizoítas aderidos foram permeabilizados com 0,1% de Triton X-100 por 20 minutos e, após lavagem, a lâmina foi bloqueada com 2% de BSA e 20 mM de glicina em PBS, por 16 horas a 4°C. Cada poço foi incubado com 1,0 mL de anti-MIC17B₂₀₋₃₀₃ ou anti-MIC17C₂₀₋₃₄₇ na concentração 0,5 µg/mL em PBS contendo 0,2% de BSA, por 1 hora, com agitação a 50 rpm, a temperatura ambiente. Os poços foram lavados três vezes por 5 minutos com PBS e incubados com o anticorpo secundário de cabra

anti-coelho conjugado ao marcador Alexa 488 (*Alexa Fluor 488 Goat Anti-Rabbit IgG (H+L)*, Invitrogen) na diluição 1:500 em PBS contendo 0,2% de BSA, por 1 hora, no escuro, com agitação a 50 rpm. Após três lavagens com PBS, os poços foram incubados com 0,01 µg/µL de iodeto de propídeo por 10 minutos, no escuro, a 50 rpm. Após três lavagens, a câmara foi removida e a lâmina preparada em glicerol 60%, coberta com lamínula e selada com esmalte.

3.2.16 Ensaio de invasão

Com o objetivo de verificar a influência das proteínas MIC17B e MIC17C no processo de invasão celular pelo taquizoíta de *N. caninum*, foram utilizados os anti-soros e os anticorpos purificados anti-MIC17B₂₀₋₃₀₃ e anti-MIC17C₂₀₋₃₄₇ para neutralização destas proteínas no taquizoíta, que depois foi incubado com as células Vero. Para quantificar a invasão, foram utilizados taquizoítas modificados por inserção do gene que codifica a enzima β-galactosidase. Quando o análogo de galactosídeo CPRG (chlorophenol red-β-D-galactopyranoside) é hidrolisado pela β-galactosidase, sua cor amarelo-alaranjada é convertida em vermelho escuro, que pode ser lida a 570 nm. Portanto, a absorbância obtida neste ensaio é proporcional à quantidade de taquizoítas que invadiram as células^{78,79}.

3.2.16.1 Com taquizoíta transfectado de forma transiente (NcLacZ transiente)

Taquizoítas de *N. caninum* foram transfectados de forma transiente com o plasmídeo NcRPS13tetO/LacZ⁸⁰ que contém gene para a enzima β-galactosidase (LacZ), cuja expressão depende do promotor do gene que codifica a proteína ribossomal RPS13 de *N. caninum*. Este plasmídeo foi construído por substituição da região promotora de RPS13 de *T. gondii* pela de *N. caninum* no plasmídeo RPS13sub23⁸¹. A construção final no plasmídeo foi a seguinte: região promotora do gene RPS13 de *N. caninum*; sequência nucleotídica *tetO*, usada em sistemas para controle de expressão por tetraciclina (que não se aplica neste ensaio); e o gene LacZ, que codifica a proteína β-galactosidase (figura 12).

Figura 12. Disposição dos genes no plasmídeo NcRPS13tetO/LacZ usado para transfecção transiente de taquizoítas utilizados em ensaios de invasão. (Adaptado de Pereira *et al*, 2013⁸⁰).

Para o ensaio, $1,2 \times 10^8$ taquizoítas selvagens recém-purificados foram eletroporados com 75 μ g do plasmídeo NcRPS13tetO/LacZ em 750 μ L de tampão de eletroporação (120 mM KCl; 0,15 mM CaCl_2 ; 100 mM $\text{K}_2\text{HPO}_4/\text{KH}_2\text{PO}_4$ pH 7,6; 5 mM MgCl_2 ; 2 mM EDTA; 25 mM Hepes 7,6) adicionado de 2 mM de ATP e 5 mM de glutathiona. A eletroporação foi realizada em cubetas de 4 mm sob 1,8 kV, 100 Ω , e 25 μ F (*GenePulser*, BioRad). Os taquizoítas eletroporados foram imediatamente reinoculados em uma garrafa de 75 cm^2 com crescimento confluyente de células Vero. Após três dias de incubação em estufa a 37°C com 5% de CO_2 , os taquizoítas NcLacZ foram purificados e utilizados no ensaio de invasão, que foi realizado em placas de 6 poços com crescimento confluyente de células Vero. Antes da incubação na placa, os taquizoítas NcLacZ foram pré-incubados com os soros (diluição 1:50) ou com os anticorpos purificados (concentração 1,2 μ g/mL) diluídos em meio RPMI, por 10 minutos a 37°C sob agitação, e então adicionados na placa, em duplicata, com $1,3 \times 10^6$ taquizoítas por poço. Como controles, foram usados os respectivos soros pré-imune e também o soro anti-NcTE. Como controle positivo foram utilizados os taquizoítas NcLacZ incubados apenas em meio RPMI, enquanto o controle negativo foi a incubação de apenas RPMI, sem taquizoítas. Para todas as amostras o ensaio foi realizado com e sem inativação do sistema complemento, a 54°C por 20 minutos. A placa foi incubada a 37°C com 5% de CO_2 por 2 horas, para invasão. Após incubação, o sobrenadante de cada poço foi removido e 200 μ L de tampão de ensaio (100 mM Hepes pH8,0; 1% Triton X-100; 1 mM $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$) contendo 5 mM de DTT foram distribuídos em cada poço, seguindo-se uma incubação a 50°C por 1 hora, sob agitação a 125 rpm, para a lise das células e obtenção dos taquizoítas que invadiram as células. O conteúdo lisado foi transferido para tubos de 1,5 mL e centrifugado a 10000 x g por 4 minutos. Em uma microplaca de 96 poços, 20 μ L do sobrenadante de cada tubo foi adicionado a 190 μ L de tampão de ensaio contendo 5 mM de DTT, em duplicata para cada tubo. Finalmente, o CPRG foi adicionado a cada poço para uma concentração final de 1 mM. A reação foi medida após 18 horas de incubação a 37°C, em comprimento de onda de 570 nm (*ELISA reader*

SpectraMax Plus, Molecular Devices). Para verificar a possível lise de taquizoítas durante a incubação, a medida da atividade de β -galactosidase também foi realizada no sobrenadante retirado após as 2 horas de incubação para invasão, previamente centrifugado. As análises estatísticas foram elaboradas no programa Graph Pad PRISM 4.0 (Graph Pad, USA), em que foi aplicada a análise de variância ANOVA, e em que a comparação entre amostras foi feita com o teste Bonferroni, considerando significância estatística para erro menor que 5%.

3.2.16.2 Com taquizoíta transfectado de forma estável (NcLacZ estável)

Os taquizoítas utilizados neste ensaio tiveram o gene da β -galactosidase inserido em seu genoma junto com o gene para resistência à pirimetamina, por transfecção com o plasmídeo NcDHFR-DHFRM2M3/NcTub-tetO/Lac-Z⁸⁰. Neste plasmídeo, o gene da β -galactosidase vem após o promotor da proteína tubulina (NcTub), que também contém a sequência *tetO*. À frente do promotor de NcTub foi inserido o gene mutado da diidrofolato redutase – timidilato sintase (DHFR-TS). As duas mutações pontuais neste gene, chamadas M2M3, possibilitam a obtenção de taquizoítas resistentes à pirimetamina. Esta construção foi flanqueada na região 5' com o promotor de DHFR-TS de *N. caninum* e na região 3' com a porção 3' UTR também de DHFR-TS de *N. caninum* (figura 13), possibilitando a recombinação homóloga destas extremidades com as extremidades do gene DHFR-TS do genoma de *N. caninum*, por pressão seletiva com 1 mM de pirimetamina.

Figura 13. Construção do plasmídeo NcDHFR-DHFRM2M3/NcTub-tetO/Lac-Z, utilizado para a inserção do gene da β -galactosidase no genoma de *N. caninum*. DHFR-TS, diidrofolato redutase – timidilato sintase; NcTub, promotor de tubulina com sequência *tetO*; LacZ, gene para β -galactosidase; 3'UTR, região 3' não traduzida. (Figura adaptada de Pereira *et al*, 2013⁸⁰).

Ensaio preliminares com NcLacZ estável, utilizando apenas anticorpos purificados para a pré-incubação com os taquizoítas, resultaram em promoção da invasão. Este efeito poderia ser devido ao BSA adicionado aos anticorpos após sua purificação, para estabilização por ação antioxidante. Portanto, para o próximo

ensaio, foi utilizado um controle positivo contendo a mesma concentração de BSA presente nos anticorpos purificados. Assim, os taquizoítas LacZ foram pré-incubados com 1,2 µg/mL de anticorpo purificado anti-MIC17B₂₀₋₃₀₃ ou anti-MIC17C₂₀₋₃₄₇, e para o controle positivo os taquizoítas foram incubados em RPMI adicionado de solução de BSA 1 mg/mL, referente ao mesmo volume usado de anticorpos purificados. O controle negativo foi de apenas RPMI, sem taquizoítas. A pré-incubação foi feita por 30 minutos a 37°C, sob agitação, seguindo-se a distribuição na placa de 24 poços coberta com células Vero, em triplicata, de 10⁶ taquizoítas por poço. A placa foi incubada a 37°C com 5% de CO₂ por 2 horas, tempo para ocorrer a invasão. Após incubação, o sobrenadante foi retirado e a placa lavada três vezes com PBS. As células e taquizoítas invasores foram lisados por adição de 200 µL de tampão de ensaio adicionado de 3 mg/mL de DTT e 0,1% de SDS e incubação a 50°C, por 1 hora, com agitação a 125 rpm. O conteúdo lisado foi transferido para tubos de 1,5 mL e, após centrifugação, 20 µL do sobrenadante foram adicionados a 190 µL de tampão de ensaio contendo 3 mg/mL de DTT, em placa de 96 poços, em triplicata para cada tubo. Foram adicionados 40 µL de CPRG 50 mM e a placa foi incubada a 37°C por 22 horas. A leitura das absorbâncias foi realizada em 570 nm (Sunrise Tecan). A análise estatística foi elaborada no programa Graph Pad PRISM 4.0 (Graph Pad, USA), com análise de variância ANOVA, e comparação entre amostras foi feita com o teste Tukey, considerando significância estatística para erro menor que 5%.

3.2.17 Ensaio de adesão (*binding*)

Os domínios Apple estão presentes em proteínas reconhecidas por sua propriedade adesiva que pode ser devida à interação proteína-proteína ou proteína carboidrato²³. Portanto foram desenvolvidos ensaios para verificar a possível adesão das proteínas MIC17B e MIC17C às células hospedeiras, no caso, às células Vero.

3.2.17.1 Com proteína recombinante solúvel

Os genes das proteínas MIC17A, MIC17B e MIC17C foram clonados para expressão heteróloga, usando o plasmídeo pET28-SUMO. Este plasmídeo, gentilmente cedido pela Profa. Dra. Maria Cristina Nonato, possui uma sequência para a proteína SUMO (*small ubiquitin-like modifier*) inserida no plasmídeo pET28

entre os sítios NheI e BamHI, o que faz com que a proteína seja expressada com a cauda de histidina ligada à porção da proteína SUMO, que aumenta a solubilidade de proteínas parcialmente insolúveis. Assim, as proteínas expressadas por este sistema podem ser extraídas com soluções não-desnaturantes e purificadas em coluna de níquel com eluição após clivagem da porção SUMO por protease SUMO-específica.

Devido à alta identidade entre as sequências codificadoras de MIC17A, MIC17B e MIC17C, o que dificulta o desenho de *primers* específicos para cada sequência, primeiramente foram amplificadas as sequências específicas de cada uma das proteínas com *primers* desenhados para a região promotora do gene (UTR 5') e para a região após o códon de parada (UTR 3'), que são um pouco menos conservadas. Em seguida, apenas a região codificadora, incluindo-se os três domínios PAN/Apple (A1, A2, A3) mais o domínio Apple-like (A4) (aminoácidos 55 a 351) e excluindo-se a região de peptídeo sinal, foi amplificada por PCR sobre os produtos de PCR previamente obtidos (*Nested* PCR).

As sequências que seriam usadas como substrato para a amplificação por *Nested* PCR (promotora até 3' UTR), foram amplificadas e clonadas em plasmídeo pGEM, com os *primers* descritos na tabela 2.

Por *Nested* PCR, as sequências de MIC17A₅₅₋₃₅₁, MIC17B₅₅₋₃₅₁ e MIC17C₅₅₋₃₅₁ foram amplificadas usando como substrato o plasmídeo pGEM previamente ligado às sequências completas de cada proteína. Os *primers* utilizados nas amplificações foram adicionados dos sítios de restrição BamHI e HindIII e estão descritos na tabela 8. As PCRs foram realizadas com o *HotStar HiFidelity Polymerase Kit* (Qiagen) e 2,0 µL de substrato, com temperatura de anelamento 58°C. Os produtos de PCR foram separados por eletroforese em gel de agarose e purificados com o *illustra GFX PCR DNA and Gel Band Purification Kit* (GE Healthcare).

Tabela 8. *Primers* utilizados para a amplificação de fragmentos a serem expressados em pET28-SUMO. Os *primers* AppleVeAm foram usados na amplificação de MIC17A₅₅₋₃₅₁ e MIC17C₅₅₋₃₅₁, e os *primers* AppleRo na amplificação de MIC17B₅₅₋₃₅₁. As sequências referentes aos sítios de restrição encontram-se sublinhadas.

<i>Primer</i>	Orientação	Sequência (5' - 3')
AppleVeAmF - BamHI	forward	CCC <u>GGATCC</u> TGCCTTGAAGCAGGCAAAGAGT
AppleVeAmR - HindIII	reverso	CCC <u>AAGCTT</u> TCAGCATGTTACATCCCCTGCTT
AppleRoF - BamHI	forward	CCC <u>GGATCC</u> TGCATGGTGCAAGGCAAAGAGTA
AppleRoR - HindIII	reverso	CCC <u>AAGCTT</u> TCAGCATGTTACATCGCCTCCTT

Antes da inserção no plasmídeo pET28-SUMO, as sequências de MIC17A₅₅₋₃₅₁, MIC17B₅₅₋₃₅₁ e MIC17C₅₅₋₃₅₁ foram inseridas no plasmídeo pGEM, exatamente como descrito na seção 3.2.4. Após a transformação em células TOP10, purificação e digestão dos plasmídeos com BamHI e HindIII, verificou-se que apenas a sequência de MIC17C₅₅₋₃₅₁ foi ligada com sucesso ao plasmídeo. Este plasmídeo pGEM-MIC17C₅₅₋₃₅₁ foi então digerido com as enzimas BamHI e HindIII e o fragmento purificado.

O plasmídeo pET28-SUMO foi purificado de células TOP10 e também digerido com as enzimas BamHI e HindIII para a ligação ao fragmento MIC17C₅₅₋₃₅₁, em reação com a enzima T4 DNA Ligase (Fermentas) por 16 horas a 16°C, e mais 6 horas a 4°C. Células BL21 foram transformadas com a ligação e plaqueadas em meio LB ágar contendo 75 µg/mL de kanamicina. Após crescimento a 37°C por 16 horas, oito colônias foram inoculadas em meio LB com 50 µg/mL de kanamicina para crescimento por 16 horas a 37°C, sob agitação. Os plasmídeos foram purificados por miniprep com o kit *PureYield Plasmid Miniprep System* (Promega) e digeridos com as enzimas BamHI e HindIII para conferência da inserção do fragmento.

O mesmo processo de ligação de MIC17C₅₅₋₃₅₁ foi realizado com um plasmídeo pET28-SUMO previamente selecionado. Para esta seleção, o DNA genômico de *N. caninum* foi digerido com as enzimas BamHI e HindIII e todo o produto da digestão foi submetido à ligação com pET28. O plasmídeo em que a inserção ocorreu com sucesso, chamado pET28-tag1, foi digerido com as enzimas BamHI e HindIII e submetido à ligação com MIC17C₅₅₋₃₅₁.

3.2.17.2 Com ESA_{SN} por *binding*-ELISA (a), *binding*-Western blot (b), e *binding*-FACS (c)

Nestes ensaios foi avaliada a capacidade de adesão das proteínas nativas MIC17B e MIC17C, presentes no ESA_{SN}, às células Vero, por detecção com os anticorpos purificados anti-MIC17B₂₀₋₃₀₃ ou anti-MIC17C₂₀₋₃₄₇.

O ESA_{SN} usado nestes ensaios foi produzido da mesma forma descrita na seção 3.2.13.2, porém, ao invés de ser precipitado com TCA em acetona e ressuspenso em tampão AM, o sobrenadante da cultura em seu terceiro dia (ESA_{SN3}) foi centrifugado para a remoção de taquizoítas e passado em filtro de 5 µm, e depois de 0,2 µm. As proteínas que estavam em meio RPMI foram lavadas com PBS e

concentradas para um volume final de 1,0 mL, com concentração de 1,5 µg/µL. Dessa forma trabalhou-se com as proteínas em sua forma nativa, e não desnaturada.

a) *Binding-ELISA*

Uma placa de 96 poços (TPP) foi preparada com crescimento confluyente de células Vero, com alguns poços sem células. Nos passos a seguir, foram realizadas três lavagens com PBS entre todos eles. As células foram primeiramente fixadas com formaldeído 3,65% por 1 hora; não foram permeabilizadas; e foram bloqueadas com BSA 1% em PBS contendo 10 mM de glicina, por 16 horas a 4°C, em câmara úmida. Foram testadas 3 condições diferentes, com duas concentrações de ESA_{SN} e duas concentrações de anticorpos purificados (tabela 9). Os poços foram incubados com ESA_{SN} por 2 horas a temperatura ambiente, para permitir a adesão das proteínas às células Vero. Após remoção de ESA_{SN} e lavagem com PBS, os poços foram incubados com os anticorpos purificados por 1 hora a 37°C, e depois com o anticorpo secundário de cabra anti-coelho conjugado à peroxidase (*Goat anti-rabbit IgG peroxidase*, KPL) também por 1 hora a 37°C. A reação com TMB (*Single solution, Ready-to-use*, Zymed-Invitrogen) foi incubada por 15 minutos a temperatura ambiente, no escuro, e interrompida com HCl 2 M. A leitura foi realizada a 450 nm em leitor *Sunrise Tecan*. Foram utilizados controles negativos (para cada uma das condições) sem células Vero, ou sem ESA_{SN}, ou sem anticorpo; além de um controle positivo com o soro anti-NcET (1:500).

Tabela 9. Condições variadas para ensaio de adesão por *binding-ELISA*.

Condição	Quantidade de ESA _{SN} por poço (µg)	Concentração de anticorpo purificado anti-MIC17B ₂₀₋₃₀₃ ou anti-MIC17C ₂₀₋₃₄₇ (µg/mL)
1	0,5	0,2
2	0,5	0,05
3	1,0	0,05

b) *Binding-Western blot*

Este ensaio foi adaptado de Garcia-Requet *et al*, 2000⁸² e um esquema geral do ensaio é demonstrado na figura 14.

Figura 14. Esquema do ensaio de adesão por Western-blot. 1) O extrato secretado de *N. caninum* composto pelo sobrenadante da cultura de taquizoítas (ESA_{SN}) foi tratado (t), ou não tratado (nt), com os anticorpos purificados anti-MIC17B₂₀₋₃₀₃ (M17B) ou anti-MIC17C₂₀₋₃₄₇ (M17C). 2) ESA_{SN} tratado ou não tratado, e o controle com PBS, foram incubados na placa para adesão das proteínas de ESA_{SN} às células Vero. 3) Após lavagem, o conteúdo dos poços foi solubilizado em T.A.M. 4) Separação do conteúdo solubilizado (proteínas de células Vero e proteínas de ESA_{SN} aderidas) por SDS-PAGE 1D. 5) Transferência das proteínas para membrana de PVDF. 6) Incubações das tiras de PVDF com os anticorpos anti-MIC17B₂₀₋₃₀₃ (Ac M17B), anti-MIC17C₂₀₋₃₄₇ (Ac M17C) ou com o soro anti-NcET (Ac NcET). M, marcador de peso molecular.

Uma placa de 24 poços (TPP) foi preparada com crescimento confluyente de células Vero. Antes da incubação com as células Vero, 100 μ g de ESA_{SN} foram pré-incubados com 0,25 μ g/mL de anticorpos purificados anti-MIC17B₂₀₋₃₀₃ ou anti-MIC17C₂₀₋₃₄₇ por 1 hora a 37°C, sob agitação. Após este tratamento, o t- ESA_{SN} (tratado) foi adicionado à placa para incubação por 2 horas a 37°C. Foram adicionados também controles por incubação das células com ESA_{SN} não tratado (nt- ESA_{SN}) ou sem ESA_{SN} (apenas com PBS). Após lavagem da placa com PBS, o conteúdo dos poços foi solubilizado em 60 μ L de tampão AM, com sonicação, e centrifugado a 10000 x g por 10 minutos, guardando-se o sobrenadante. Em um gel de acrilamida unidimensional preparado com um pente para 5 poços (seção 3.1.11.1), foram adicionados e separados por eletroforese 20 μ L do conteúdo

solubilizado de cada poço. Após a separação, as proteínas foram transferidas do gel para membrana de PVDF (seção 3.1.14). As membranas foram coradas com solução de DB71, digitalizadas, descoloridas, bloqueadas e então recortadas em tiras para incubação com os anticorpos primários. As tiras foram incubadas por 1 hora, sob agitação, com 0,25 µg/mL de anti-MIC17B₂₀₋₃₀₃ ou anti-MIC17C₂₀₋₃₄₇, ou com soro anti-NcTE 1:500, diluídos em PBS-GT 0,08%. Após lavagem com PBS-T, as tiras foram incubadas com o anticorpo secundário de cabra anti-coelho conjugado à peroxidase (*Goat anti-rabbit IgG peroxidase*, KPL) também por 1 hora, e reveladas como descrito na seção 3.1.14.

c) **Binding-FACS**

Neste ensaio, o ESA_{SN} foi incubado com células Vero tripsinizadas que, após lavagem, foram incubadas com os anticorpos purificados para detectar se as proteínas MIC17B e MIC17C ficaram aderidas às células Vero. Depois do anticorpo primário, as células foram incubadas com anticorpo secundário conjugado a FITC, e a leitura foi realizada por citometria de fluxo.

Seis garrafas com crescimento confluyente de células Vero ($1,3 \times 10^6$ células) foram tripsinizadas por incubação com um volume final de 16 mL de tripsina (0,5% *Trypsin-EDTA*, Gibco) por 30 minutos em estufa a 37°C com 5% de CO₂. As células foram reunidas num tubo *falcon* de 50 mL e um volume igual, de 16 mL, de meio RPMI contendo 5% de soro fetal bovino (SFB) foi adicionado para neutralização da tripsina. Com centrifugações de 5 minutos a 1000 x g, as células foram lavadas três vezes com meio RPMI/SFB. Ao final foram ressuspensas em 9 mL de RPMI/SFB e divididas em 9 tubos de 1,5 mL (1 mL/tubo = $1,44 \times 10^5$ células). Aos tubos foram adicionados 115 µg de ESA_{SN} para incubação com as células por 45 minutos a 37°C, sob agitação. As células foram lavadas, também por centrifugação, duas vezes com PBS, e incubadas com 0,25 µg/mL de anticorpo purificado anti-MIC17B₂₀₋₃₀₃ ou anti-MIC17C₂₀₋₃₄₇, diluídos em PBS, por 45 minutos a 25°C, sob agitação. Após duas lavagens com PBS, as células foram incubadas com 1:500 do anticorpo secundário de cabra anti-coelho conjugado à FITC (*FITC Goat Anti-Rabbit IgG (H+L) Conjugate*, Zymed, Invitrogen) por 30 minutos a 25°C, sob agitação, no escuro, e então lavadas três vezes com PBS. Ao final, as células foram ressuspensas em 100 µL de PBS e a leitura de fluorescência foi realizada em

citômetro de fluxo (*FACSCanto*, *FACSDiva software*, BD), com número de eventos detectados igual a 10.000.

3.2.18 Deleção (nocaute) dos genes de MIC17A, MIC17B e MIC17C

O plasmídeo p Δ Apple foi construído para nocautear os genes das proteínas de *N. caninum* MIC17A, MIC17B e MIC17C por recombinação homóloga induzida por seleção com o antibiótico cloranfenicol (figura 15). Os taquizoítas nocaute (Δ) para os genes das proteínas Apple seriam submetidos à avaliação funcional por ensaios de invasão e de adesão.

Fig 15. Construção do plasmídeo p Δ Apple. A partir do plasmídeo pSAG-CAT, o plasmídeo p Δ Apple foi construído por substituição do promotor de SAG1 de *N. caninum* (PROM_SAG1) pela região promotora do gene para MIC17A de *N. caninum* (PROM_M17A), e também por substituição da região 3' de SAG1 (3'_SAG1) pela região 3' não traduzida de MIC17C de *N. caninum* (3'UTR_M17C). Para a deleção dos genes das três proteínas Apple, a recombinação homóloga entre as regiões PROM_M17A e 3'UTR_M17C do plasmídeo p Δ Apple e as mesmas regiões presentes no cromossomo VIII (Crm_VIII) de *N. caninum*, foi induzida por pressão seletiva com o antibiótico cloranfenicol, cuja resistência depende da expressão da proteína cloranfenicol acetiltransferase (CAT). Sítios de restrição: 1, Hind III; 2, Bgl II; 3, Nde I, Xba I.

3.2.18.1 Preparo de plasmídeo e transfecção de taquizoítas

O plasmídeo pSAG-CAT foi construído sobre o plasmídeo TgRPS13tetOLacZ, como descrito em Pereira *et al*, 2013⁸⁰. Para a construção do plasmídeo p Δ Apple (figura 15), as regiões promotora (PROM_SAG1) e terminal (3'_SAG1) da proteína de superfície SAG1 de *N. caninum* foram removidas do plasmídeo pSAG-CAT por digestão com enzimas de restrição específicas e, com as extremidades formadas

pelas mesmas enzimas de restrição, os fragmentos referentes à região promotora do gene para MIC17A de *N. caninum* (PROM_M17A) e à região 3' não traduzida de MIC17C de *N. caninum* (3'UTR_M17C) foram inseridos antecedendo e sucedendo, respectivamente, o gene para cloranfenicol acetiltransferase (CAT).

O fragmento PROM_M17A foi amplificado por PCR a partir de DNA genômico de *N. caninum* com os *primers* adicionados de sítios de restrição PromVeF_HindIII e PromVeR_BglII, enquanto o fragmento 3'UTR_M17C foi amplificado com os *primers* StopAmF_NdeI e StopAmR_XbaI (tabela 10). A reação foi feita com o *HotStar HiFidelity Polymerase Kit* (Qiagen), empregando temperatura de anelamento de 58°C.

Os fragmentos amplificados foram purificados de gel de agarose com *illustra GFX PCR DNA and Gel Band Purification Kit* (GE Healthcare) e ligados ao plasmídeo pGEM (*pGEM-T Easy Vector Systems*, Promega). Células TOP10 foram transformadas com os plasmídeos pGEM-PROM_M17A ou pGEM-3'UTR_M17C, e selecionadas com ampicilina e X-Gal. As colônias brancas foram inoculadas em meio LB com 50 µg/mL de ampicilina para crescimento a 37°C por 16 horas, com agitação. Os plasmídeos foram purificados das células com o *AxyPrep Plasmid Miniprep Kit* (Axygen) e digeridos com as enzimas HindIII e BglII para pGEM-PROM_M17A (em tampão *R*, Fermentas), e as enzimas NdeI e XbaI para pGEM-3'UTR_M17C (em tampão *D*, Promega). Os fragmentos PROM_M17A e 3'UTR_M17C foram purificados de gel de agarose com o *illustra GFX PCR DNA and Gel Band Purification Kit* (GE Healthcare).

O plasmídeo pSAG-CAT foi purificado de células TOP10 e digerido primeiramente com as enzimas NdeI e XbaI para retirada de 3'UTR de SAG1 e ligação com o fragmento 3'UTR_M17C previamente digerido de pGEM com as mesmas enzimas. Células TOP10 foram transformadas com a ligação e selecionadas com 50 µg/mL de ampicilina. Os plasmídeos foram purificados das células com o kit *Pure Yield Plasmid Miniprep System* (Promega) e digeridos com as enzimas NdeI e XbaI para conferência da inserção de 3'UTR_M17C. A presença de 3'UTR_M17C no plasmídeo também foi conferida por PCR com os *primers* StopAmF e StopAmR e a disposição correta de CAT-3'UTR_M17C foi verificada por PCR com os *primers* CAT_F e StopAmR (tabela 10).

O plasmídeo pSAG-CAT já com 3'UTR_M17C (pSAGCAT-3'UTR) foi então digerido com as enzimas HindIII e BglII para retirada do promotor de SAG1 e ligação com o fragmento PROM_M17A previamente digerido de pGEM com as mesmas enzimas. A ligação foi usada para transformar células TOP10 que foram selecionadas com ampicilina. Os plasmídeos foram purificados das células com o kit *Pure Yield Plasmid Miniprep System* (Promega) e digeridos com as enzimas HindIII e BglII para conferência da inserção de PROM_M17A. Foi realizada PCR para verificar a correta disposição de PROM_M17A-CAT com os primers PromVeF e CAT_R. O plasmídeo final com a construção PROM_CAT_3'UTR foi nomeado pΔApple.

Tabela 10. Sequência dos *primers* utilizados na amplificação dos fragmentos PROM_M17A e 3'UTR_M17C para inserção em plasmídeo pΔApple. Os sítios de restrição estão especificados no nome do *primer* e suas sequências encontram-se sublinhadas. CAT_F e CAT_R foram usados juntamente com os outros *primers* para conferência de inserção correta no plasmídeo pΔApple.

<i>primer</i>	orientação	sequência 5' - 3'
PromVeF - HindIII	forward	CCCAAGCTTCTTTTCCGTGTGCAGTCTTC
PromVeR - BglII	reverso	CCCAGATCTGGTGCTTTTCCAATGCAGGGAA
StopAmF - NdeI	forward	CCCCATATGGGACTTCTACGATGCTCAGACTA
StopAmR - XbaI	reverso	CCCTCTAGACTTTCCCTGTCTGCCCGTTCTT
CAT_F	forward	ATGCATGAGAAAAAATC
CAT_R	reverso	TTATGCCCCGCCCTGCCA

3.2.18.2 Transfecção e seleção com cloranfenicol

A transfecção com o plasmídeo pΔApple foi realizada no taquizoíta de *N. caninum* NcLacZ, que possui o gene para β-galactosidase inserido de forma estável em seu genoma, por seleção com DHFR-TS mutado (seção 3.2.16.2). Assim, após o nocaute, os taquizoítas poderiam ser avaliados quanto à invasão pelo método de quantificação com CPRG.

Os taquizoítas NcLacZ foram purificados de uma garrafa pequena (25 cm²) em fluxo laminar, resultando num total de $4,5 \times 10^7$ taquizoítas, que foram ressuspensos em 1,0 mL de tampão de eletroporação (120 mM KCl; 0,15 mM CaCl₂; 100 mM K₂HPO₄/KH₂PO₄ pH 7,6; 5 mM MgCl₂; 2 mM EDTA; 25 mM Hepes 7,6) adicionado de 5 mM de glutatona, e eletroporados com 2,8 µg de pΔApple, em cubetas de 4 mm, sob 1,8 kV, 100 Ω, e 25 uF (*GenePulser*, BioRad). Após o choque o conteúdo foi inoculado em uma garrafa pequena com crescimento confluyente de células Vero e adicionado de 6,0 mL de meio RPMI contendo 10 µg/mL de kanamicina e 1 mM de

pirimetamina. Após 1 dia de incubação a 37°C com 5% de CO₂, o meio da garrafa foi trocado, adicionando-se desta vez, além da kanamicina e da pirimetamina, 20 mM de cloranfenicol.

Os taquizoítas foram mantidos sob pressão seletiva com 20 mM de cloranfenicol e reinoculação aproximadamente a cada 2 semanas. Após dois meses de seleção, o cloranfenicol foi retirado de uma das garrafas, enquanto outras continuaram sob seleção por até 5 meses (nomeada NcΔ1).

Uma nova transfecção foi realizada, também com taquizoítas NcLacZ, porém, desta vez, com o plasmídeo pΔApple linearizado por digestão com a enzima HindIII, para aumentar as chances de recombinação homóloga. A eletroporação foi realizada da mesma forma descrita anteriormente, utilizando-se diretamente todo o conteúdo da digestão. Após um dia da eletroporação, adicionou-se 15 mM de cloranfenicol à cultura, que foi reinoculada em novas garrafas, em média a cada duas semanas. Após 2,5 meses de seleção, o cloranfenicol foi retirado de uma das garrafas e mantido em outra. Dois meses depois, o DNA foi extraído da cultura ainda sob seleção (NcΔ2) com o *Wizard Genomic DNA Purification Kit* (Promega).

3.2.18.3 Confirmação de nocaute por PCR

Com o DNA genômico de NcΔ1 e NcΔ2, foram realizadas PCRs com o kit *GoTaq Hot Start* (Promega) para verificar a presença do gene para a proteína Apple MIC17B e a presença das construções PROM_M17A-CAT e CAT-3'UTR, com os *primers* relacionados na tabela 11. Para as reações foi usada temperatura de anelamento de 56°C e foram usados como controle o DNA de NcLacZ não transfectado e o plasmídeo purificado pΔApple.

Tabela 11. *Primers* utilizados nas PCRs para a amplificação de MIC17B e das construções PROM_M17A-CAT e CAT-3'UTR para conferência de nocaute sobre DNA de taquizoíta NcLacZ transfectado com pΔApple.

	MIC17B (1543 pb)	PROM_M17A-CAT (1512 pb)	CAT-3'UTR (2.280 pb)
<i>Primer forward</i>	AR1	PromVeF	CAT_F
<i>Primer reverso</i>	R2	CAT_R	StopAmR

3.3 Estudo do secretoma

3.3.1 Viabilidade de taquizoítas por citometria de fluxo

Logo após o processo de purificação (seção 3.1.1), os taquizoítas foram submetidos à avaliação de viabilidade por citometria de fluxo. Como controle, foram utilizados taquizoítas mortos por aquecimento a 85°C por 10 minutos. A viabilidade dos taquizoítas na amostra e no controle foi avaliada por incorporação de 20 µg/mL de iodeto de propídeo (Santa Cruz Biotechnology) e leitura no citômetro de fluxo FACSCanto com *software* FACSDiva 6.1.3 (BD, equipamento multiusuário da FCFRP-USP). O laser de argônio foi ajustado para 488 nm e utilizou-se 535 nm e 617 nm para excitação e detecção, respectivamente. A abertura correspondente às características de *N. caninum* foi desenhada por detectores situados na linha do feixe de luz (*forward scatter* ou FSC) para análise do volume celular, e detectores perpendiculares ao feixe de luz (*side scatter* ou SSC) para análise da complexidade celular. A porcentagem de taquizoítas vivos e mortos foi determinada com base no resultado negativo e positivo, respectivamente, ao iodeto de propídeo.

3.3.2 Análise de NcESA e NcET - por MALDI

3.3.2.1 Preparo de extrato secretado (ESA) e extrato total (ET) de taquizoítas

Na primeira análise, por MALDI (Ionização e Dessorção a Laser Assistida por Matriz), o extrato secretado de *N. caninum* (NcESA) foi preparado sem o estímulo por etanol. Assim, taquizoítas recém-purificados (seção 3.1.1) foram incubados em meio RPMI (10⁷/mL) a 37°C por 10 minutos e resfriados em gelo por mais 10 minutos. O sobrenadante (ESA) foi coletado após duas etapas de centrifugação e as proteínas foram precipitadas com adição do mesmo volume de ácido tricloroacético (TCA) 30% em acetona e incubação a -20°C por 45 minutos. Após centrifugação, o pellet com proteínas do ESA foi ressuspenso em 100 µL de tampão AM (ureia 7M, tiourea 2M e CHAPS 4%). O controle foi preparado sem taquizoítas (VeESA), seguindo-se os mesmos passos de purificação e incubação para obtenção de NcESA, mas a partir de garrafas contendo apenas células Vero.

O extrato total de *N. caninum* (NcET) foi preparado a partir de 5 x 10⁸ taquizoítas sonificados em 350 µL tampão AM. Para ser usado como controle, o extrato total de

células Vero (VeET) foi preparado da mesma forma descrita para os taquizoítas, com as células recuperadas de uma garrafa de 25 cm².

3.3.2.2 Eletroforese bidimensional (SDS-PAGE 2D) de ET e ESA

Géis bidimensionais foram preparados como descrito anteriormente (3.1.11.2). Foram separados 250 µg dos extratos total (ET) e secretado (ESA) de *N. caninum* e de células Vero em *strips* de 13 cm e pH imobilizado de 3 a 11 não linear (NL) e, em seguida em gel de acrilamida a 12,5%. Os géis foram corados com Coomassie-G250 e então digitalizados (seções 3.1.12.1 e 3.1.13).

3.3.2.3 Identificação de proteínas por MS-MALDI

Após comparação entre os géis de *N. caninum* e de células Vero, spots de interesse foram selecionados, excisados e preparados para análise por espectrometria de massas (seção 3.1.15). Os peptídeos foram analisados em equipamento MALDI-TOF-TOF do Centro de Química de Proteínas de Ribeirão Preto e também no equipamento MALDI-Q-TOF do Laboratório Nacional de Luz Síncrotron, em Campinas. A busca para a identificação de proteínas foi realizada contra o banco de dados de *Neospora caninum* (ToxoDB versão 6.0) e o banco de dados SwissProt (www.ebi.ac.uk/swissprot) através do *software on line* Mascot (Matrix Science). Foi incluída como modificação fixa a carbamidometilação de cisteína e como modificação variável a oxidação de metionina. Os limites de tolerância foram 1.2 Da para íons precursores e de 0.6 Da para fragmentos de íons.

3.3.3 Análise de ESA e quantificação no taquizoíta estimulado - por LC-MS

Foram utilizadas duas abordagens diferentes e complementares para a investigação do secretoma de *N. caninum* usando cromatografia líquida acoplada à espectrometria de massas (LC-MS):

1. identificação de proteínas presentes no extrato secretado (Abordagem Via Extrato Secretado);
2. quantificação relativa de proteínas, por marcação com isótopos, no taquizoíta submetido à estimulo secretório por etanol (Abordagem Quantitativa).

O desenho experimental para ambas as abordagens é mostrado na figura 16.

Fig. 16. Desenho experimental para o estudo do secretoma de *N. caninum*. Os taquizoítas purificados foram submetidos ao estímulo secretório por incubação com 1% de etanol em meio RPMI, por 15 minutos a 37°C. Após centrifugação, as proteínas do ESA foram separadas do taquizoíta *estimulado*. A, **Abordagem Quantitativa**, extratos totais foram produzidos com os taquizoítas não-estimulados (*controle*) e com os taquizoítas recuperados após a coleta do ESA (*estimulados*), e digeridos com as enzimas lisina-C e tripsina. Os peptídeos foram marcados com isótopos dimetil, com marcação leve (*controle*) e intermediária (*estimulado*). As duas amostras foram adicionadas na proporção 1:1 e a mistura foi fracionada por cromatografia de troca iônica, com resina catiônica (SCX). As frações 9 a 24 foram analisadas em LTQ-Orbitrap-XL usando árvore de decisão para os métodos de fragmentação: Dissociação Induzida por Colisão (CID) e Dissociação por Transferência de Elétrons (ETD) e analisador *Ion Trap* (ETD-IT). As proteínas foram identificadas usando o programa Mascot contra o banco de dados *decoy* de *N. caninum*, e a quantificação foi realizada com o programa Proteome Discoverer. A localização putativa e os sinais para secreção foram avaliados *in silico*. A interação entre proteínas identificadas foi predita com o programa UNISPPi, e os arquivos

MS gerados foram usados na busca por regiões codificadoras não-annotadas com a ferramenta Pepsiplice. **B, Abordagem via Extrato Secretado**, o ESA foi lavado e concentrado em filtros para centrífuga e então secos em concentrador a vácuo (SpeedVac). Após separação por SDS-PAGE 1D, as proteínas do ESA foram digeridas, e os peptídeos extraídos. Os peptídeos tripticos foram analisados em dois espectrômetros de massa: um LTQ-Orbitrap-XL, com fragmentação por dissociação induzida por colisão (CID), e um LTQ-Orbitrap-Velos, com a escolha do método de fragmentação baseada numa árvore de decisão entre (1) ETD-IT, (2) ETD e analisador com Transformada de Fourier (ETD-FT), e (3) Dissociação por Colisão de Alta Energia (HCD) e analisador com FT (HCD-FT). Usando o programa Mascot, os dados de espectrometria de massas (MS) foram buscados contra o banco de dados de *N. caninum*, e os resultados filtrados com o programa Scaffold. Análises *in silico* foram realizadas para a predição de sinais de secreção e para a classificação das proteínas identificadas.

3.3.3.1 Abordagem Via Extrato Secretado

3.3.3.1.1 Obtenção do extrato secretado (ESA) de taquizoítas de *N. caninum*

O extrato secretado (ESA) foi obtido por incubação de taquizoítas recém-purificados (10^8 /mL) em meio RPMI contendo 1% de etanol (Merck) por 15 minutos a 37°C, e em seguida por 10 minutos em gelo. O sobrenadante foi purificado por duas etapas de centrifugação e armazenado em alíquotas a -20°C. Empregando concentradores com filtro para centrífuga (*Amicon Ultra Centrifugal Filter Devices 5K*, Millipore), alíquotas combinadas de ESA foram lavadas, a 4°C, com bicarbonato de amônio 25 mM pH 7,4 e então concentradas em água MilliQ para 700 µL, e o ESA foi finalmente seco em concentrador a vácuo.

3.3.3.1.2 Digestão de proteínas

O ESA foi ressuspenso por sonicação (Labsonic M, Sartorius) em 120 µL de bicarbonato de amônio 50 mM; 25% desta amostra (38,5 µg) foram separados por SDS-PAGE 1D (3.1.11.1) e o gel, corado com Coomassie G-250 (*Gel Code Blue Stain Reagent*, Thermo Scientific), foi recortado em 13 partes. Cada parte foi tratada com DTT 6,5 mM por 1 hora a 60°C para redução, e com iodoacetamida 54 mM por 30 minutos para alquilação. As proteínas foram digeridas com 50 µg/mL de tripsina (Promega) por 16 horas a 37°C, e os peptídeos extraídos com acetonitrila (ACN) e secos em concentrador a vácuo (seção 3.1.15).

3.3.3.1.3 Análise por espectrometria de massas: RP-nanoLC-MS/MS

Os dados foram adquiridos em dois espectrômetros diferentes com a mesma configuração de cromatografia líquida de fase reversa (RP-nanoLC) acoplada ao

espectrômetro de massa. Antes da análise por espectrometria de massas (MS), os peptídeos foram separados por cromatografia líquida (LC) em sistema Agilent 1200 equipado com pré-coluna (*trap*) Reprosil C18-AQ (100 μm $\times$ 20 mm, 3 μm , 120 Å) e coluna analítica Reprosil C18-AQ (50 μm $\times$ 400 mm, 3 μm , 120 Å), ambas *Dr. Maisch, Ammerbuch, Germany*. Na pré-coluna um gradiente foi realizado com solvente A (ácido acético 0.1 M em água) a 5 $\mu\text{L}/\text{min}$ por 10 minutos; 40% de solvente B (ácido acético 0.1 M em 8:2 (v/v) de ACN/água) por 30 minutos; 100% solvente B por 2 minutos; e solvente A por 15 minutos. O fluxo foi alterado para 100 nL/min durante as análises.

Os peptídeos trípticos foram analisados em dois espectrômetros de massa: LTQ-Orbitrap-XL e LTQ-Orbitrap-Velos (Thermo). No LTQ-Orbitrap-XL a fragmentação de peptídeos foi realizada por Dissociação Induzida por Colisão (CID). No LTQ-Orbitrap-Velos utilizou-se um sistema de fragmentação guiado por árvore de decisão, que, com base no estado de carga e na relação massa/carga (m/z) do peptídeo precursor⁵⁶, escolhe o melhor método de fragmentação. Os métodos de fragmentação foram Dissociação por Transferência de Elétrons (ETD) e analisador *Ion Trap* (ETD-IT), ETD e analisador com Transformada de Fourier (ETD-FT), e Dissociação por Colisão de Alta Energia e analisador com FT (HCD-FT). Foram adquiridos espectros MS variando de 350 a 1500 (m/z) e os 5 íons mais intensos foram selecionados para fragmentação.

3.3.3.1.4 Identificação de proteínas

A busca com os dados de MS foi feita contra o banco de dados de *N. caninum* (ToxoDB versão 6.4, 7344 entradas), incluindo uma lista de contaminantes comuns, com o programa *Mascot* versão 2.3.02 (Matrix Science). Carbamidometilação de cisteína e oxidação de metionina foram selecionadas como modificações fixa e variável, respectivamente. A busca foi realizada por meio do programa *Proteome Discoverer* 1.3 (Thermo Fischer Scientific) com tolerância para peptídeo de 7 ppm e para fragmento de íon de 0.6 Da (leitura IT) ou 0.05 Da (leitura FT), permitindo duas clivagens perdidas e respeitando o *score* de no mínimo 20. Os resultados em arquivos “.mgf” foram convertidos para “.dat” e filtrados no programa Scaffold versão 3. Para as identificações de peptídeos e proteínas, a probabilidade deveria ser maior que 95.0%, com o mínimo de 1 peptídeo único, como especificado pelo algoritmo *Peptide and Protein Prophet*⁸³.

3.3.3.1.5 Predições *in silico* para sinais de secreção e classificação de proteínas

As proteínas identificadas no ESA de *N. caninum* foram classificadas de acordo com a localização no parasita e vias de secreção. Para as proteínas ainda não caracterizadas, a localização foi predita principalmente com base em homólogos presentes em organismos Apicomplexas relacionados, buscados pelo site do ToxoDB⁶⁶. Dessa forma, a lista de proteínas com localização putativa nas micronemas foi obtida por meio da palavra-chave “micronem*”, e as proteínas previamente identificadas em *N. caninum* ou em organismo Apicomplexa relacionado, mas ausentes desta lista, foram inseridas manualmente à lista final. Para outros grupos a busca foi realizada seguindo o mesmo método: “rhoptr*” para proteínas de roptrias, “dense granul*” para proteínas de grânulos densos, “surface*” ou “SAG*” ou “SRS*” para proteínas de superfície, “cycloph*” para ciclofilinas e proteínas relacionadas, “mitochond*” para proteínas mitocondriais, e “nuclear*” ou “nucleus” para proteínas de núcleo. Não foram obtidos resultados em *Neospora* com busca por “apicoplast*”, portanto, a busca foi feita sobre o banco de *T. gondii* ME49 e obteve-se uma lista de homólogos em *N. caninum*. Para a predição de secreção não-clássica e de secreção clássica (presença de peptídeo sinal), as sequências das proteínas identificadas foram submetidas à análise no programa SecretomeP 2.0 Server⁸⁵. A nomenclatura para proteínas ainda não caracterizadas segue o trabalho de Reid *et al*, 2012⁸⁶, porém, uma nova nomenclatura pode ser eventualmente concedida quando estas proteínas forem caracterizadas.

3.3.3.2 Abordagem Quantitativa

3.3.3.2.1 Preparo de ET de taquizoítas *estimulados* e taquizoítas *controle*

O tratamento para a obtenção de taquizoítas *estimulados* foi o mesmo para a obtenção de ESA (seção 3.3.3.1.1). Após a purificação, os parasitas (10^8 /mL) foram submetidos ao estímulo secretório com 1% de etanol em meio RPMI por 15 minutos a 37°C, e resfriados em gelo por 10 minutos. Os taquizoítas *estimulados* foram coletados por centrifugação e armazenados a -20°C. O controle foi composto por taquizoítas purificados e não estimulados (taquizoítas *controle*). Os extratos totais (ET) dos taquizoítas *controle* e *estimulado* foram obtidos por sonicação (Sonic Dismembrator 100 - Fisher Scientific) de aproximadamente $8,5 \times 10^8$ parasitas em

200 µL de tampão AM (ureia 7M, tiourea 2M, CHAPS 4%), resultando em 1 mg de proteínas em cada extrato. As proteínas foram precipitadas com o dobro de volume de ácido tricloroacético 30% em acetona e secas em concentrador a vácuo.

3.3.3.2.2 Digestão dos extratos e marcação com dimetil-isótopos

Cada extrato total foi ressuspensão, por sonicação (LABSONIC M - Sartorius), em 300 µL de ureia 8M, reduzido com DTT 0,4 mM por 25 minutos a 56°C, e alquilado com iodoacetamida 8 mM por 30 minutos a temperatura ambiente, no escuro. As proteínas foram inicialmente digeridas com a enzima lisina-C (1 µg para 100 µg de substrato) a 37°C por 4 horas e, após a amostra ser diluída 8 vezes em bicarbonato de amônio 50 mM, com a enzima tripsina (1,2 µg para 100 µg de substrato) a 37°C por 16 horas.

A marcação de peptídeos com isótopos estáveis dimetil foi realizada de acordo com protocolo padrão⁵⁷. Quantidades iguais de peptídeos obtidos das amostras *controle* e *estimulado* foram separadamente marcados com reagentes para marcação *leve* e *intermediária*, respectivamente, usando método de marcação em solução, por meio de cartuchos SepPak C18 (Waters). A porcentagem de peptídeos marcados e a correta proporção entre as amostras *controle* e *estimulado* foram checados por espectrometria de massas em LTQ-Orbitrap-XL, usando árvore de decisão para fragmentação por CID ou ETD. As amostras marcadas *controle* e *estimulado* foram então misturadas na proporção 1:1.

3.3.3.2.3 Fracionamento dos extratos por cromatografia líquida de troca catiônica

A cromatografia líquida de troca catiônica (SCX) para o fracionamento da mistura 1:1 com as amostras marcadas *estimulado* e *controle* foi realizada como descrito previamente⁸⁷. Utilizou-se uma coluna Zorbax BioSCX-Series II column (0.8 mm diâmetro interno x 50 mm comprimento, 3.5 µm tamanho da partícula) e, como solventes, 0,05% de ácido fórmico em 20% de ACN (solvente A), e 0,05% de ácido fórmico em solução com 0,5 M de NaCl e 20% de ACN (solvente B). A ordem do gradiente de eluição foi a seguinte: 0-0.01 min (0-2% B); 0.001-8.01 min (2-3% B); 8.01-18.01 min (3-8% B); 18.01-28 min (8-20% B); 28-38 min (20-40% B); 38-44 min

(40-100% B); 44-48 min (100% B); e 48-50 min (100-0% B), com fluxo de 40 μ L/min. Entre 40 e 90 minutos, 50 frações foram coletadas.

3.3.3.2.4 Espectrometria de massas por RP-nanoHPLC-MS/MS

As frações mais intensas, contendo peptídeos dupla e triplamente carregados, foram selecionadas para análise por LC-MS e reconstituídas em 10% de ácido fórmico contendo 5% de DMSO. Antes da análise espectrométrica (MS), os peptídeos foram separados por cromatografia líquida (LC) usando a mesma configuração descrita na secção 3.3.3.1.3. A análise dos peptídeos foi realizada em espectrômetro de massas LTQ-Orbitrap-XL, utilizando o sistema de fragmentação guiado por árvore de decisão para os métodos Dissociação Induzida por Colisão (CID) e Dissociação por Transferência de Elétrons (ETD) com analisador *Ion Trap* (ETD-IT). As frações foram eluídas num gradiente de 3 horas de 0 a 40% de solvente B (ácido acético 1M em 80% de ACN) a 100 nL/min. Os 10 íons mais intensos foram selecionados para fragmentação.

3.3.3.2.5 Identificação e quantificação de proteínas

Os dados de MS foram analisados com o programa *Proteome Discoverer* versão 1.3 (Thermo Fischer Scientific). Os espectros MS/MS foram buscados, usando o programa *Mascot* 2.3.02 (Matrix Science), contra o banco de dados de *N. caninum* (ToxoDB 6.4) incluindo uma lista de contaminantes comuns e um banco de dados *decoy* (composto por versões reversas de todas as sequências). Para a busca foi selecionada a enzima trypsin, permitindo duas clivagens perdidas. Carbamidometilação de cisteína e oxidação de metionina foram configuradas como modificações fixa e variável, respectivamente. Devido às marcações com dimetil-isótopos, também foram selecionadas como modificações variáveis: dimethyl (K), dimethyl (N-term), dimethyl-2H(4) (K) e dimethyl-2H(4) (N-term). Foram permitidas tolerâncias para peptídeo de 5 ppm e para fragmento de íon de 0.6 Da, aplicando os filtros de *score* maior que 20, posição 1 no ranking e comprimento do peptídeo de 7 a 35 resíduos, resultando numa taxa de detecção falsa (FDR, do inglês *false discovery rate*) de 1,2%.

A quantificação foi realizada no programa *Proteome Discoverer* 1.3, utilizando-se a área do pico no modo '*precursor ions quantifier*'. Apenas peptídeos únicos foram quantificados e a razão reportada foi configurada para intermediária/leve

(*estimulado/controle*). As razões foram normalizadas sobre a mediana da proteína, e um limite de duas vezes a razão (*fold change threshold of 2.0*) foi selecionado para expressão diferencial. Portanto, razão maior que 2.0 configurou regulação para cima (*up regulation*) e razão menor que 0.5 regulação para baixo (*down regulation*), indicando proteínas potencialmente não secretadas e secretadas, respectivamente.

3.3.3.2.6 Classificação das proteínas quantificadas por localização e vias de secreção

As proteínas quantificadas foram classificadas como descrito na seção 3.3.3.1.5, de acordo com a localização no parasita (micronemas, roptrias, grânulos densos, e superfície) e vias de secreção (clássica e não-clássica).

3.3.4 Rede de interação de proteínas de *N. caninum*

A predição de interações proteína-proteína entre as 2011 proteínas quantificadas foi realizada com o programa Universal Predictor of Protein-Protein Interactions (UNISPPi)⁸⁸. UNISPPi é um método computacional baseado em aprendizado de máquina que prediz a probabilidade de interação física entre um par de proteínas, com base apenas nas características físico-químicas associadas à sequência de aminoácidos. As características físico-químicas das sequências pertencentes às proteínas identificadas foram calculadas com o programa PROFEAT⁸⁹.

Os pares de proteínas interagentes, com probabilidade maior que 90%, compostos por proteínas *up* ou *down* reguladas foram selecionados para visualização com o programa Cytoscape versão 2.8⁹⁰, empregando o *layout* Edge-Weighted Force-Directed (BioLayout), cujo parâmetro “*edge betweenness*” gera uma centralidade definida pela fração de “caminhos mais curtos” entre os nodos da rede. As redes de interação envolvendo proteínas *up* e *down* reguladas foram combinadas para a construção de uma rede única contendo as interações das duas redes.

3.3.5 Análise proteogenômica: Identificação de peptídeos em regiões não-annotadas

A busca por peptídeos presentes em regiões não-annotadas como codificadoras de proteínas foi feita contra o banco de dados genômico de *N. caninum* (ToxoDB 7.2) com a ferramenta *PepsplICE*⁹¹, utilizando os arquivos de MS gerados por CID tanto

da amostra *controle* quanto *estimulada*, assim como os arquivos gerados por CID e HCD para análise do ESA de *N. caninum*.

Para a busca, um limite máximo de 1% foi aplicado para a taxa de detecção falsa (FDR) e, como modificações, foram selecionadas carbamidometilação (C), oxidação (M), dimethyl (K), dimethyl (N-term), dimethyl-2H(4) (K) e dimethyl-2H(4) (N-term). Os resultados foram convertidos para o formato GFF e visualizados com o programa Integrative Genomics Viewer⁹² (IGV) versão 1.5 contra o banco de dados predito de *N. caninum* (ToxoDB 7.2).

4. RESULTADOS

4.1 Proteínas com domínios Apple

4.1.1 Análise *in silico* das sequências com domínios Apple

A busca contra o banco de dados genômico não-annotado de *N. caninum* do Sanger Institute⁶⁵ com as ESTs CF660241.1 e BF249278.1 gerou (em agosto/2008) dois *contigs* (824 e 514). No *contig* 824, foram encontradas três sequências, cada uma com 1056 pb, em intervalos de aproximadamente 3000 pb, localizadas no cromossomo VIII. A busca no ToxoDB⁶⁶ mostrou que três proteínas eram previstas pelas sequências NCLIV_038120, NCLIV_038110 e NCLIV_038100, que posteriormente foram nomeadas⁴⁷ MIC17A, MIC17B e MIC17C, respectivamente (figura 17). No *contig* 514 foi encontrada uma sequência com 1053 pb que, no ToxoDB, prediz a proteína NCLIV_068830 (localizada em cromossomo não atribuído (“not assigned”)), sem homólogo entre os organismos incluídos no ToxoDB (*T. gondii* e *E. tenella*). Já as proteínas MIC17 possuem quatro homólogos em *T. gondii* (cepa ME49): TGME49_200230, TGME49_200240, TGME49_200250 e TGME49_200270 (figura 17).

Fig. 17 Organização dos genes de MIC17A, MIC17B e MIC17C no genoma de *N. caninum* (isolado Liverpool). Na parte superior da figura estão representadas as três proteínas por setas vermelhas, com indicação de transcrição no sentido reverso. Na parte inferior da figura é representada a relação das três proteínas entre si e com homólogos sintênicos de *T. gondii* (cepas ME49, GT1 e VEG). (Adaptado da página do ToxoDB⁶⁶).

Em um alinhamento (www.uniprot.org/align⁹³) com as sequências de aminoácidos das quatro proteínas, formadas exclusivamente por um exon, é verificado que todas apresentam (figura 18) uma região de peptídeo sinal (PS); um domínio de direcionamento para a organela micronema⁶⁰ (M); três domínios Apple ou PAN1 (A1, A2, A3); e um domínio Apple-like ou PAN3 (A4). Todos os domínios possuem seis cisteínas que possivelmente formam três pontes dissulfeto.

	PS	M	1 A1							
1	MEALKTVLSRRIALPVVAAVLLQDVAEAGQYAGLRSQLVNLHLSFAETEATSSGCL	EAGK	60	NCLIV_038120						
1	MEALKTVLSRRIALPVVAAVLLQDVT	EAGQYAGLRSQLVNLHLSFAETD	STNYS	CMVQ	GK 60 NCLIV_038110					
1	MEALKTVLSRRIALPVVAAVLLQDVAEAGQYAGLRSQLVNLHLSFAETEATSSGCL	EAGK	60	NCLIV_038100						
1	MEALKTVLSRRIALPVVAAVLLQDVAEAGQYAGLRSQLVNLHLSFAETEATSSGCL	EAGK	60	NCLIV_068830						
	*****	*****	*****	*****	*****					
	2	3	3	2						
61	EYVGFALAEFSNVSDADTCQKRCSQHSDCV	FEFFVSSGNRCLLHSTKPT	EVKNNS	NA	TAG 120 NCLIV_038120					
61	EYSGYVLAELSDVSDAASCQKNC	DQHPQCR	FEFFVSD	KRCRLQVR	QPS	EVDRD	NSNA	FSG 120 NCLIV_038110		
61	EYVGFALAEFSNVSDADTCQKRCSQHSDCV	FEFFVSSGNRCLLHSTKPT	EVKNNS	NA	TAG 120 NCLIV_038100					
61	EYVGFALAEFSNVSDADTCQKRCSQHSDCV	FEFFVSSGNRCLLHSTKPT	EVKNNS	NA	TAG 120 NCLIV_068830					
	**	.*	***	.*	***	.*	***	.*	***	.*
	1	1 A2	2	3	3	2				
121	PKRCHVCVLDGDFKGP	NLVENGQPGIK	ILSACQK	AKAKES	RCYGYL	FAPR	INK	CH	LKT	180 NCLIV_038120
121	PKRCFVCVIEGYDFKGP	PNINNGV	PGVITL	AACQV	ACQAE	PRCYG	FLF	NS	SKTC	FFK 180 NCLIV_038110
121	PKRCHVCVLDGDFKGP	NLVENGQPGIK	ILSACQK	AKAKES	RCYGYL	FAPQ	GK	CH	FKT	180 NCLIV_038100
121	PKRCHVCVLDGDFKGP	NLVENGQPGIK	ILSACQK	AKAKES	RCYGYL	FAPQ	GK	CH	CYLKT	180 NCLIV_068830
	****	***	.*	***	.*	***	.*	***	.*	***
	1	1 A3	2	3	3	2				
181	TQQCVGALAPDANYMAGPKT	CTDEHWCIMK	DVDPK	NDLR	KVKAD	TAAACQ	DACL	NDE	KC	240 NCLIV_038120
181	GKMYVSALVPDSKYIAGPKT	CANERWCV	MSGIQ	FNGG	DKRT	LKTS	AAACQ	DECL	NDG	DC 240 NCLIV_038110
181	TQHCVGALAPDANYMAGPKT	CTDEHWCIMK	DVDPK	NDLR	KVKAD	TAAACQ	DACL	NDE	KC	240 NCLIV_038100
181	TQQCVGALAPDANYMAGPKT	CTDEHWCIMK	NVDP	PGHA	-RK	VITQ	SSAACQ	DACL	NDE	KC 239 NCLIV_068830
	:	.*	***	.*	***	.*	***	.*	***	.*
	2	1	1 A4							
241	EYFSLEGPKKDCWLKAGAKVP	SSKTKKAS	VFSGPK	KH	CGLPTT	CIKERT	KYVDD	TIV	IY	PK 300 NCLIV_038120
241	DYFSWIESSNECLLKAGSAV	GAAKMTK	SGVHSG	PKH	CGLPTT	CIKEQ	TKYVDD	TIV	IY	PK 300 NCLIV_038110
241	EYFSLEGPKKDCWLKAGAKVP	SSKTKKAS	VFSGPK	KH	CGLPTT	CIKERT	KYVDD	TIV	IY	PK 300 NCLIV_038100
240	DFFSWEAKTKACWLKEGV	KSASTKKT	QANFES	GP	KH	CGLPTT	CVKERT	KYVDD	TIV	IY 299 NCLIV_068830
	::**	.*	***	.*	***	.*	***	.*	***	.*
	2	3	3	2	1					
301	TEVGSFAACQLKCWNKCV	FMHFNSSG	CTLSG	SSATE	VADAN	SKAGD	VTC	351	NCLIV_038120	
301	TEVGSFAACQRRCKWKT	GNCFM	HFNDEG	CTLSG	INATE	VADAN	SKGGD	VTC	351	NCLIV_038110
301	TEVGSFAACQLKCWNKCV	FMHFNSSG	CTLSG	SSATE	VADAN	SKAGD	VTC	351	NCLIV_038100	
300	TEVGSFAACQLKCWNKCV	FMHFNSSG	CTLSG	SSATE	VADAN	SKAGD	VTC	350	NCLIV_068830	
	*****	.*	***	.*	***	.*	***	.*	***	.*

Fig. 18 Alinhamento entre as sequências de aminoácidos das proteínas hipotéticas NCLIV_038120 (MIC17A), NCLIV_038110 (MIC17B), NCLIV_038100 (MIC17C) e NCLIV_068830. PS, peptídeo sinal; M, domínio de direcionamento para micronema; A1, A2 e A3, domínios Apple (PAN1); A4, domínio Apple-like (PAN3); numeração 1, 2 e 3, cisteínas envolvidas em pontes dissulfeto; “*”, resíduos idênticos em todas as sequências; “:”, substituições conservadas; “.” substituições semi-conservadas.

As sequências de NCLIV_038120 (MIC17A) e NCLIV_038100 (MIC17C) são as mais próximas entre si, com 97,2% de identidade e 99,7% de similaridade. NCLIV_068830 também é mais parecida com MIC17A e MIC17C, com 89,7% de identidade e 97% de similaridade em relação a ambas. NCLIV_038110 (MIC17B) tem a sequência que mais difere das outras três, com 63,5%, 64,7% e 65,0% de

identidade e 86%, 87,2% e 87.5% de similaridade em relação a NCLIV_068830, MIC17A e MIC17C, respectivamente.

Para as proteínas MIC17A, MIC17B e MIC17C, cujas sequências codificadoras encontram-se próximas e no mesmo cromossomo (cromossomo VIII, região FR823390: 6,487,905 a 6,497,164), foram comparadas também as sequências não transcritas 5'UTR (região promotora) e 3'UTR (região após o códon de parada). As regiões 3'UTR (1400 pb) de MIC17A e MIC17C são extremamente parecidas, com 98% de identidade. Já em relação às regiões 5' UTR, essa alta identidade é observada entre MIC17B e MIC17C (2000 pb), também com 98% de identidade. As regiões 5'UTR de MIC17A e 3'UTR de MIC17B são as que mais diferem das outras, têm apenas 38% de identidade com as demais regiões 5'UTR ou 3'UTR (Figura19).

Fig. 19 Organização das regiões codificadoras de MIC17A, MIC17B e MIC17C (representadas por blocos coloridos) e de suas regiões adjacentes não traduzidas (5'UTR e 3'UTR) no cromossomo VIII de *N. caninum*. As letras coloridas referem-se às regiões não traduzidas (5'UTR ou 3'UTR) de cada uma das proteínas: **A**, de MIC17A; **B**, de MIC17B; **C**, de MIC17C. O número abaixo das letras equivale ao número de pares de base dessas regiões não traduzidas. As regiões 5'UTR de MIC17A (**A** 1000 pb, à esquerda do bloco verde) e 3'UTR de MIC17B (**B** 1000 pb, à direita do bloco rosa) apresentam apenas 38% de identidade com as demais regiões 5'UTR ou 3'UTR. As regiões 3'UTR de MIC17A e MIC17C compartilham 98% de identidade entre si, portanto foram representadas por **A/C** para enfatizar a semelhança entre essas regiões (**A/C** 1400 pb, à direita dos blocos verde e amarelo). As regiões 5' UTR de MIC17B e MIC17C também compartilham 98% de identidade entre si, e são representadas da mesma forma (**B/C** 2000 pb, à esquerda dos blocos rosa e amarelo).

4.1.2 Amplificação e sequenciamento dos genes com domínios Apple

Foram realizadas PCRs para amplificação, sobre cDNA de *N. caninum*, das sequências completas (desde 5'UTR até 3'UTR) dos genes com domínios Apple NCLIV_038120, NCLIV_038110, NCLIV_038100 e NCLIV_068830 (seção 3.2.2). Foram amplificadas com sucesso (figura 20 A e B) as sequências de NCLIV_038120 (MIC17A), NCLIV_038110 (MIC17B) e NCLIV_038100 (MIC17C). Não foi obtida a amplificação de NCLIV_068830 mesmo após tentativas de variação de temperatura de anelamento ou mesmo com PCR sobre o DNA genômico. As bandas obtidas com aproximadamente 600 pb e 2500 pb não corresponderam ao tamanho esperado

para NCLIV_068830, de 1662 pb (figura 20 B). Assim, apenas as três sequências citadas acima foram clonadas em pGEM e sequenciadas. Os plasmídeos pGEM submetidos à sequenciamento com *primers* específicos são mostrados, após digestão para conferência de fragmentos, na figura 20 C.

Fig. 20 Amplificação dos genes que codificam as proteínas Apple por PCR e clonagem em pGEM. Géis de agarose (0,8%). **M**, marcador de tamanho molecular (pb) *GeneRuler 1kb DNA Ladder* (Fermentas). **A**, produtos da amplificação por PCR de NCLIV_038110 e NCLIV_038100. **B**, produtos da amplificação por PCR de NCLIV_038120 e NCLIV_068830. **C**, produto da digestão de plasmídeo pGEM para conferência da inserção de MIC17A (NCLIV_038120), MIC17B (NCLIV_038110) e MIC17C (NCLIV_038100). O conteúdo dos poços e o tamanho esperado de cada sequência estão especificados na figura.

Os resultados de sequenciamento confirmaram tanto a região codificadora de cada proteína, por comparação com as sequências depositadas no ToxoDB⁶⁶, quanto as regiões não traduzidas (5'UTR e 3'UTR) por comparação com o *contig* 824 obtido no banco do Sanger Institute⁶⁵. Para a sequência de MIC17C foram encontradas duas variáveis; uma sequência que possui na posição 319 o nucleotídeo adenina (A), que corresponde à sequência depositada no ToxoDB, e outra em que uma guanina (G) ocupa esta posição 319. A sequência com adenina geraria, na posição 107 da sequência proteica, o aminoácido lisina (AAG), e com guanina geraria o aminoácido ácido glutâmico (GAG). Foi confirmado que a

sequência amplificada para expressão heteróloga de MIC17C₂₀₋₃₄₇ possuía o nucleotídeo adenina.

4.1.3 Predição da estrutura terciária dos domínios Apple e das proteínas MIC17A, MIC17B e MIC17C

4.1.3.1 Alinhamento estrutural entre os domínios das proteínas MIC17 e o domínio A5 de *T. gondii*

Por um alinhamento estrutural entre os domínios Apple das proteínas MIC17 com um domínio Apple cuja estrutura foi depositada no PDB (2LL3) verificou-se que, apesar da baixa identidade sequencial, o domínio Apple é bastante conservado e que as principais diferenças encontradas nos diferentes domínios Apple residiam nas regiões flexíveis desses domínios e que as regiões necessárias à formação de estruturas secundárias são conservadas (figura 21). Esse resultado foi crucial para apoiar a predição estrutural dos domínios Apple das MIC17 por modelagem molecular baseada nas estruturas depositadas no PDB.

Fig. 21 Alinhamento estrutural dos domínios Apple A1, A2, A3 e A4 de MIC17A, MIC17B e MIC17C com a sequência 2LL3 (domínio A5 da proteína MIC4 de *T. gondii*, cuja estrutura foi elucidada por RMN e depositada no PDB sob o código 2LL3). As estruturas secundárias são representadas acima de suas regiões correspondentes na sequência: setas, fitas-beta (β); espiral, alfa-hélice (α). Os retângulos azuis delimitam as regiões conservadas por similaridade. As cisteínas conservadas são destacadas em vermelho, e as pontes dissulfeto nas quais elas estão envolvidas estão numeradas em verde, abaixo das fileiras de cisteínas (1, 2 e 3). Nomenclatura dos domínios: a1m17a, domínio Apple A1 pertencente à proteína MIC17A, e assim sucessivamente para os outros domínios e para as proteínas MIC17B e MIC17C.

4.1.3.2 Modelos de domínios Apple selecionados para a predição

Os quatro domínios Apple, previamente depositados no PDB e descritos na tabela 5, foram avaliados quanto à viabilidade para uso como modelos na predição estrutural dos domínios Apple das proteínas MIC17A, MIC17B e MIC17C. Uma vez

que a porcentagem de identidade sequencial entre os quatro domínios Apple mostrou-se baixa (tabela 12, de 15% a 45%), decidiu-se inicialmente avaliar o grau de conservação estrutural deste tipo de domínio, de forma a validar o método de predição estrutural por modelagem molecular baseada em homologia.

Tabela 12. Porcentagens de identidade/similaridade (%) entre modelos de domínios Apple selecionados para predição estrutural de MIC17A, MIC17B e MIC17C. 2YIL, domínio A1 da proteína SML2 de *Sarcocystis muris*; 2LL3, domínio A5 da proteína MIC4 de *Toxoplasma gondii*; 2F83_2 e 2F83_3, domínios A2 e A3, respectivamente, do fator de coagulação XI de *Homo sapiens*.

	2LL3	2YIL	2F83_A2
2YIL	45/71	—	—
2F83_A2	19/50	21/46	—
2F83_A3	15/39	18/45	38/70

Uma superposição entre os modelos estruturais, realizada no software PYMOL⁷¹ utilizando a estrutura 2LL3 como referência, mostrou que, apesar da baixa porcentagem de identidade/similaridade sequencial (tabela 12), o RMSD (desvio médio quadrático) para a posição dos carbonos-alfa (C α) entre estes modelos mostrou-se baixo (tabela 13) (em torno de 1 Å).

Tabela 13. RMSD de C α (Å), calculados por alinhamento estrutural no PYMOL, entre modelos de domínios Apple selecionados para predição estrutural de MIC17A, MIC17B e MIC17C. 2YIL, domínio A1 da proteína SML2 de *Sarcocystis muris*; 2LL3, domínio A5 da proteína MIC4 de *Toxoplasma gondii*; 2F83_2 e 2F83_3, domínios A2 e A3, respectivamente, do fator de coagulação XI de *Homo sapiens*.

	2LL3	2YIL	2F83_A2
2YIL	0.58	—	—
2F83_A2	0.61	0.68	—
2F83_A3	0.71	1.60	0.75

Além dessas análises iniciais, com o intuito de verificar onde residiam as diferenças estruturais entre os modelos, o valor de RMSD foi plotado em função do resíduo (a numeração se refere à estrutura 2LL3) (figura22). Para facilitar a observação dos resultados, a sequência de aminoácidos de 2LL3 e a localização de suas estruturas secundárias foram adicionadas abaixo do eixo X do gráfico, assim como o alinhamento sequencial entre os quatro modelos. Verificou-se então que as regiões de hélices e fitas apresentam RMSD mais baixo (de aproximadamente 1 Å), e a maior diferença entre os modelos (maior RMSD) reside nas regiões flexíveis, de loops.

Fig. 22 RMSD por resíduo de aminoácido dos domínios Apple com base no alinhamento estrutural dos modelos 2YIL, 2F83_2 e 2F83_3 com o modelo 2LL3. O cálculo de RMSD foi realizado com o software Protein3Dfit. **Eixo X**, posição dos aminoácidos na sequência de 2LL3. **Eixo Y**, RMSD de C α por resíduo das sequências de 2YIL, 2F83_2 e 2F83_3 em relação aos resíduos da sequência de 2LL3 (Å). Acompanhando o eixo X, são mostradas a sequência de aminoácidos de 2LL3 e as estruturas secundárias formadas em cada região deste modelo. Para fins de comparação, o alinhamento sequencial entre os modelos foi adicionado abaixo do eixo X. Estruturas secundárias: espiral, alfa-hélice; setas, fitas-beta. 2YIL, domínio A1 da proteína SML2 de *Sarcocystis muris*; 2F83_2 e 2F83_3, domínios A2 e A3, respectivamente, do fator de coagulação XI de *Homo sapiens*; 2LL3, domínio A5 da proteína MIC4 de *Toxoplasma gondii*.

Esse resultado reforça a hipótese de que a construção de modelos para os domínios Apple em estudo neste trabalho é viável, e que as regiões flexíveis, como esperado, encontram-se na superfície da proteína. Desta forma, foram escolhidos os modelos cristalográficos que apresentaram boa estereoquímica, inicialmente avaliada pelo software *Molprobit*⁷⁰ (ferramenta: *Analyze geometry without all-atom contacts*).

4.1.3.3 Predição estrutural dos quatro domínios Apple de MIC17A, MIC17B e MIC17C

A configuração padrão encontrada para os 12 domínios, tanto Apple quanto Apple-like, segue a estrutura representada na figura 23. Um folha-beta, formada por 4 fitas antiparalelas, é estabilizada por duas pontes dissulfeto (C2-C5 e C3-C4) que ligam a folha à alfa-hélice. Perpendicularmente à folha-beta, uma pequena folha é formada por duas fitas antiparalelas, no lado oposto ao da hélice. Outra ponte dissulfeto (C1-C6 ou C1-C7) conecta as extremidades N e C terminal do domínio. Em alguns domínios verifica-se a formação de apenas três fitas da folha-beta maior e/ou ausência das duas fitas da folha beta menor.

Figura 23. Domínio a3m17b obtido por modelagem molecular. A estrutura é formada por uma folha beta, com 4 fitas anti-paralelas, ligada a uma alfa-hélice; há outra folha beta menor, formada por 2 fitas antiparalelas. Duas pontes dissulfeto amarram a folha beta à hélice, e outra ponte dissulfeto liga as extremidade N e C terminal do domínio. As cisteínas que formam as pontes dissulfeto estão destacadas em amarelo. **a3m17b**, domínio Apple 3 (A3) da proteína MIC17B de *N. caninum*.

As estruturas geradas para os 12 domínios Apple foram comparadas por alinhamento estrutural com cálculo do RMSD de C α por resíduo, em relação à sequência do domínio a1m17a (A1 de MIC17A), com o programa Protein3Dfit⁷². A figura 24 mostra os RMSD dos onze domínios em relação aos resíduos de aminoácidos do domínio a1m17a. A sequência de aminoácidos de a1m17a encontra-se abaixo do eixo X do gráfico, onde também estão desenhadas as estruturas secundárias que se formam em cada região da sequência. Nesta visualização ficou evidente que os maiores valores de RMSD correspondem a regiões flexíveis, e que nas regiões importantes para a formação de estrutura secundária o RMSD é bastante baixo, em geral menor que 1.0 Å. O resultado é compatível com a variabilidade estrutural previamente observada entre os domínios Apple utilizados como modelo.

Fig. 24 RMSD por resíduo de aminoácido, com base no alinhamento estrutural dos onze domínios Apple das proteínas de *N. caninum* MIC17A, MIC17B e MIC17C (na legenda) com o domínio a1m17a. Eixo X, posição dos aminoácidos na sequência de a1m17a. Eixo Y, RMSD de Ca por resíduo em relação aos resíduos da sequência de a1m17a (Angstroms). Acompanhando o eixo X, são mostradas a sequência de aminoácidos de a1m17a e as estruturas secundárias formadas em cada região deste domínio. Estruturas secundárias: espiral, alfa-hélice; setas, fitas-beta. Os conectores em amarelo representam as três pontes dissulfeto intra-domínio, formadas pelas cisteínas numeradas.

Uma diferença importante foi observada nos domínios A2 das proteínas MIC17A e MIC17C. Enquanto MIC17B apresenta o padrão normal de 6 cisteínas, em MIC17A e MIC17C são verificadas 7 cisteínas, onde, além das três pontes dissulfetos formadas entre C1-C7, C2-C5, C3-C4, há a cisteína C6, que fica livre (figura 25).

Fig. 25 Domínios A2 das proteínas Apple MIC17A, MIC17B e MIC17C gerados por predição estrutural com o software Modeller 9.11. **A.** a2m17a, domínio Apple A2 de MIC17A. **B.** a2m17b, domínio Apple A2 de MIC17B. **C.** a2m17c, domínio Apple A2 de MIC17C. As pontes dissulfeto estão representadas em amarelo, com as cisteínas numeradas. Em a2m17b (B), as três pontes dissulfeto são formadas pelas cisteínas C1-C6, C2-C5, C3-C4, enquanto em a2m17a e a2m17c (C) as três pontes dissulfeto são formadas por C1-C7, C2-C5, C3-C4. Nestes domínios, a cisteína livre C6 está destacada em cor diferente. Em a2m17b (B) não existe esta cisteína adicional.

4.1.3.4 Análise da superfície eletrostática dos domínios Apple

Para todos os domínios Apple das proteínas MIC17A, MIC17B e MIC17C de *N. caninum* foi calculada a superfície eletrostática sobre as estruturas preditas. Na figura 26 estão representados os perfis de distribuição de carga, com uma comparação entre os domínios A1 (painel A), A2 (painel B), A3 (painel C) e A4 (painel D) das três proteínas NcMIC17. Em vermelho são mostradas as regiões de carga negativa e em azul as regiões de carga positiva; regiões neutras são mostradas em branco. Em cada painel foi adicionada a imagem do domínio representado na forma de *cartoon*, mostrando a sua disposição espacial (como padrão foram escolhidos os domínios da proteína MIC17A). Na parte superior de cada painel os domínios foram mantidos com a α -hélice situada à direita da folha- β antiparalela, e a ponte dissulfeto que une as extremidades N e C terminal está na região anterior do domínio. Na parte inferior de cada painel os domínios sofreram uma rotação vertical de 180°, portanto, a ponte dissulfeto N-C terminal está na região posterior do domínio.

A distribuição de cargas varia tanto entre os diferentes domínios (A1, A2, A3 e A4) quanto entre os mesmos domínios de diferentes proteínas (exemplo, A1 de MIC17A, MIC17B e MIC17C). Como esperado, o perfil de distribuição de cargas, e também a curvatura da superfície, dos domínios da proteína MIC17B diferem mais dos domínios de MIC17A e MIC17C. Mesmo com a alta identidade compartilhada entre MIC17A e MIC17C, são encontradas regiões com diferente distribuição de cargas entre essas proteínas, como nos domínios A2 e A3 (fig. 26 B e C).

No domínio A2 (fig. 26 B), em que uma cisteína adicional existe apenas em MIC17A e MIC17C, são observadas regiões com carga negativa (vermelho) ao redor desta cisteína. Estas regiões poderiam estar envolvidas na atração de outra molécula para a formação de um possível complexo, via interação com a cisteína livre. No domínio A2 de MIC17B, que não possui esta cisteína adicional, a região correspondente é formada por cargas positivas (azul).

Fig. 26 Superfície eletrostática dos domínios Apple das proteínas MIC17A, MIC17B e MIC17C de *N. caninum*. Os perfis de distribuição de carga são comparados entre os domínios A1 (painel A), A2 (painel B), A3 (painel C) e A4 (painel D) das três proteínas. Em vermelho, regiões de carga negativa; em azul, regiões de carga positiva; em branco, regiões neutras. O domínio representado na forma de *cartoon* permite visualizar sua disposição espacial. Na parte superior de cada painel os domínios foram mantidos com a α -hélice à direita da folha- β antiparalela, e a ponte dissulfeto que une as extremidades N e C terminal do domínio está na região anterior. Na parte inferior de cada painel os domínios sofreram uma rotação vertical de 180° , portanto, a ponte dissulfeto N-C terminal está na região posterior do domínio. No domínio A2, a cisteína adicional em MIC17A e MIC17C está destacada no *cartoon* (seta).

4.1.3.5 Predição de *folding* de MIC17A, MIC17B e MIC17C, reunindo os quatro domínios Apple

As estruturas completas das proteínas MIC17A, MIC17B e MIC17C, preditas com base no modelo do fator de coagulação XI (PDB 2F83), não apresentaram boa estereoquímica e mostraram-se estruturalmente instáveis mesmo após o refinamento. Este resultado mostra que a organização entre os domínios das proteínas MIC17 não é o mesmo de FXI. A provável causa para a instabilidade das estruturas obtidas por predição é o baixo número de aminoácidos que compõem os intervalos entre os domínios. Enquanto no FXI há 6 ou 7 aminoácidos que separam dois domínios (figura 27), nas MIC17 são apenas 2 aminoácidos que separam A1 de A2, e 5 aminoácidos nos intervalos entre A2 e A3, e A3 e A4.

Fig. 27 Tetrâmero formado por quatro domínios Apple que compõem o fator de coagulação XI (PDB 2F83). Na estrutura os domínios estão coloridos em azul (A1), verde (A2), amarelo (A3) e vermelho (A4). As cisteínas que formam as pontes dissulfeto que unem as extremidades N e C terminal de cada domínio estão em amarelo com representação em *sticks*. As regiões que compõem os intervalos entre os domínios estão destacadas em preto e, junto a elas é mostrado o número de aminoácidos (aa) que as compõem.

Uma comparação dos intervalos entre domínios Apple foi realizada entre proteínas de organismos apicomplexas e o fator de coagulação XI (FXI), mostrando quantos domínios Apple existem em cada uma das proteínas listadas e o número de aminoácidos situados entre os domínios (Tabela 14). Provavelmente o tamanho da

sequência que separa os domínios tem forte influência na organização final destes na proteína.

Tabela 14. Número de resíduos de aminoácidos nos intervalos entre os domínios Apple de algumas proteínas de organismos apicomplexas e do fator de coagulação XI. As colunas representam os números de resíduos de aminoácidos situados entre os domínios Apple (A). **FXI**, fator de coagulação XI; **NcMIC17A/B/C**, proteínas MIC17 de *Neospora caninum*; **TgMIC4**, MIC4 de *Toxoplasma gondii*; **EtMIC5**, MIC5 de *Eimeria tenella*; **SmSML2**, SML2 de *Sarcocystis muris*.

Proteína/ domínio	A1/A2	A2/A3	A3/A4	A4/A5	A5/A6	A6/A7	A7/A8	A8/A9	A9/A10	A10/A11
FXI	6	6	7							
NcMIC17A	2	5	5							
NcMIC17A	2	5	5							
NcMIC17A	2	5	5							
TgMIC4	3	17	3	38	3					
EtMIC5	3	12	3	12	2	3	2	5	6	2
SmSML2	3									

4.1.4 Amplificação de MIC17B₂₀₋₃₀₃ e MIC17C₂₀₋₃₄₇ e clonagem em plasmídeo de expressão

As sequências codificadoras de MIC17B₂₀₋₃₀₃ e MIC17C₂₀₋₃₄₇ foram amplificadas por PCR sobre cDNA de *N. caninum* (figura 28.A), purificadas de gel de agarose e inseridas em plasmídeo pGEM. Após digestão com enzimas específicas para os sítios de restrição adicionados às sequências (figura 28.B), os fragmentos foram purificados e clonados com sucesso em plasmídeo pET28 (figura 28.C).

Fig. 28 Amplificação e clonagem de MIC17B₂₀₋₃₀₃ e MIC17C₂₀₋₃₄₇ em pGEM e pET28. Géis de agarose (0,8%). **M**, marcador de tamanho molecular (pb) *GeneRuler DNA Ladder Mix* (Fermentas). **A**, produtos da amplificação por PCR das sequências de MIC17B₂₀₋₃₀₃ e MIC17C₂₀₋₃₄₇ sobre cDNA de *N. caninum*. Após purificação estes fragmentos foram clonados em pGEM. **B**, produtos da digestão de pGEM com as enzimas BamHI e HindIII. Os fragmentos obtidos foram purificados e clonados em pET28. **C**, produtos da digestão de pET28 com BamHI e HindIII.

4.1.5 Expressão heteróloga de MIC17B₂₀₋₃₀₃ e MIC17C₂₀₋₃₄₇ em *E. coli* BL21

O peso molecular predito das proteínas recombinantes, considerando as porções adicionadas pelo plasmídeo pET28, foi de 35.7 kDa para MIC17B₂₀₋₃₀₃ e de 40.2 kDa para MIC17C₂₀₋₃₄₇ (figura 29).

Figura 29 Sequências de aminoácidos das proteínas recombinantes MIC17B₂₀₋₃₀₃ e MIC17C₂₀₋₃₄₇. Os pesos moleculares preditos foram de 35.7 kDa para MIC17B₂₀₋₃₀₃ e de 40.2 kDa para MIC17C₂₀₋₃₄₇. Em negrito, regiões referentes às proteínas com domínios Apple; Grifadas em cinza, regiões adicionadas pelo plasmídeo pET28.

Após a indução da expressão de células BL21 transformadas com os plasmídeos pET28, a extração das proteínas recombinantes foi testada por extração sequencial (seção 3.2.6). Em comparação com o controle não-induzido, ambas as proteínas foram expressas apenas sob indução com IPTG. As proteínas recombinantes não foram solubilizadas nas soluções de Tris-NaCl e de Triton, sendo extraídas apenas com o tampão ureia 8M (figura 30.A). Num segundo teste de extração, aplicando a extração sequencial, as proteínas recombinantes foram pouco extraídas em tampão ureia 2M e 4M. Houve significativa extração com o tampão 6M, mas optou-se pelo tampão ureia 8M para garantir uma extração mais completa destas proteínas (figura 30.B). Os pesos moleculares das proteínas recombinantes observadas nos géis SDS-PAGE corresponderam ao predito (figura 29).

Fig. 30. Extração de proteínas recombinantes após indução de *E. coli* BL21. SDS-PAGE-1D corados com Coomassie G-250. **A, extração sequencial de MIC17B₂₀₋₃₀₃ e MIC17C₂₀₋₃₄₇.** Poços 1, 2, 6, e 7, comparação entre controle não-induzido e induzido, ambas as extrações em tampão ureia 8M pH 8,0. Poços 3, 4, 5, 8, 9 e 10, cultura induzida com extração sequencial em Tris-NaCl 100mM pH7,5; Triton X-100 2%; e tampão ureia 8M pH8,0. Nos poços 5 e 10, MIC17B₂₀₋₃₀₃ e MIC17C₂₀₋₃₄₇, condizentes com os pesos moleculares preditos de 35,7 kDa e 40,2 kDa, respectivamente. **B, extração com diferentes tampões de ureia.** Poço 1, sobrenadante após centrifugação das células (antes de iniciar processo de extração). Extrações sequenciais após indução de MIC17C₂₀₋₃₄₇, com Tris-NaCl (poços 2, 5, 8 e 11) e Triton X-100 (poços 3, 6, 9 e 12) seguidas por tampões ureia 2M (poço 4), ureia 4M (poço 7), ureia 6M (poço 10), e ureia 8M (poço 13). **M**, marcador de peso molecular (kDa) *Protein MW Marker* (Fermentas).

Após a extração em tampão ureia 8M pH 8,0, as proteínas recombinantes foram purificadas em colunas de níquel, com eluição em tampão ureia 8M adicionado de 500 mM de imidazol (figura 31 A e B). Foram coletadas em média 15 frações a cada purificação. As frações purificadas foram reunidas e dessalinizadas em colunas PD10, com eluição em tampão ureia 8M (figura 31 C e D). A primeira fração eluída na dessalinização foi descartada e as próximas frações foram reunidas e concentradas para 1,5 mg/mL em sistema *Amicon Ultra – Centrifugal Filter Devices*

5K (Millipore) (seção 3.2.7). As proteínas recombinantes e o extrato total de *N. caninum* utilizados na imunização dos coelhos são mostrados na figura 31 E.

Fig. 31. Proteínas recombinantes purificadas, dessalinizadas e concentradas para imunização. Géis de SDS-PAGE-1D (12,5%) corados com prata. **A e B**, purificação das proteínas recombinantes MIC17B₂₀₋₃₀₃ (gel A) e MIC17C₂₀₋₃₄₇ (gel B) em coluna de níquel. 1, extrato de BL21 não submetido à coluna; 2, fração não-ligada à coluna de níquel após incubação por 16 horas; 3, lavado em tampão ureia 8M/imidazol 10mM; 4-7, exemplo de frações eluídas em tampão ureia 8M/imidazol 500 mM. **C e D**, dessalinização das proteínas recombinantes MIC17B₂₀₋₃₀₃ (gel C) e MIC17C₂₀₋₃₄₇ (gel D) em coluna PD10. 1, proteína recombinante purificada; 2, fração não ligada à coluna PD10; 3-6, frações eluídas em tampão ureia 8M pH8,0. **E**, proteínas recombinantes concentradas e extrato total de *N. caninum* (NcET) utilizados na imunização.

4.1.6 Confirmação da identidade das proteínas recombinantes por espectrometria de massas

As identidades de MIC17B₂₀₋₃₀₃ e MIC17C₂₀₋₃₄₇ foram confirmadas por identificação de 11 peptídeos (score 1245.86) e 7 peptídeos (score 383.40), respectivamente. As sequências dos peptídeos e informações sobre o número de PSM's (do inglês, *Peptide Spectrum Matches*), score e carga são mostrados na tabela 15.

Tabela 15. Peptídeos das proteínas recombinantes NcMIC17B₂₀₋₃₀₃ e NcMIC17C₂₀₋₃₄₇ identificados por espectrometria de massas (MS). Na tabela estão indicadas: a Sequência do peptídeo identificado; No. de PSM's, que corresponde ao número de espectros de fragmentação experimental em relação ao espectro teórico gerado a partir do banco de dados; Score de identificação baseado nos parâmetros do programa Mascot; e Carga do peptídeo detectado por MS. Entre os códigos de aminoácidos, "m" representa metionina oxidada.

Proteína	Sequência do peptídeo	No. de PSM's	Score	Carga
MIC17B₂₀₋₃₀₃	MYVSALVPDSK	10	87	2
	YVDDTIATYPK	15	68	2
	mYVSALVPDSK	4	66	2
	DNSNAFSGPK	2	50	1
	DNSNAFSGPKR	2	42	2
	SGVHSGPK	1	41	2
	FFTFVSDKK	2	39	2
	FFTFVSDK	9	38	2
	QPSEVRDNSNAFSGPK	1	34	3
	FFTFVSDKKR	1	33	3
	LQVRQPSEVR	1	32	3
score total 1245.86				
MIC17C₂₀₋₃₄₇	TKYVDDTIVIYPK	3	71	2
	YVDDTIVIYPK	7	69	2
	CVFmHFNSSGCTLSGSSATEVADANSK	1	45	3
	KASLFSGPK	2	41	2
	CLLHSTKPTEVK	2	39	2
	ASLFSGPK	3	37	1
	GQPNLVEK	2	30	2
score total 383.40				

4.1.7 Detecção por ELISA de anticorpos produzidos em coelhos

A soro-conversão ocorreu após a primeira imunização para MIC17B₂₀₋₃₀₃ e MIC17C₂₀₋₃₄₇ ($p < 0,05$), e após a segunda imunização para o extrato total de *N. caninum* ($p < 0,05$). Foram selecionados os soros produzidos após a terceira imunização, coletados no dia 63, para a purificação de anticorpos (figura 32).

Fig. 32 Produção de anticorpos (IgG) contra as proteínas recombinantes MIC17B₂₀₋₃₀₃, MIC17C₂₀₋₃₄₇ ou contra o extrato total de *N. caninum* (NcET). Detecção por ELISA: 0,05 µg/mL de proteína recombinante MIC17B₂₀₋₃₀₃ ou MIC17C₂₀₋₃₄₇ detectada por seus respectivos anti-soros na diluição 1:12000; 5,0 µg/mL de NcET detectado por soros anti-NcET na diluição 1:4000. Dia 0 (zero), primeira (1a) imunização; Dia 21, segunda (2a) imunização; Dia 42, terceira (3a) imunização; Dia 63, quarta (4a) imunização; Dia 84, coleta três semanas após a última imunização. *, soro conversão para MIC17B₂₀₋₃₀₃ e MIC17C₂₀₋₃₄₇ após a primeira imunização ($p < 0,05$); **, soro-conversão para NcET após a segunda imunização ($p < 0,05$). Destacados por linha tracejada, soros selecionados para purificação de anticorpos, coletados no dia 63, após a terceira imunização.

4.1.8 Purificação de anticorpos

As proteínas recombinantes, transferidas dos géis de SDS-PAGE 1D para membranas de PVDF, foram recortadas das áreas delimitadas na figura 33. Pelo processo de purificação por cromatografia de afinidade (seção 3.2.12), obteve-se 60 µg/mL de anticorpos purificados anti-MIC17B₂₀₋₃₀₃ e anti-MIC17C₂₀₋₃₄₇ que, além de quantificados, foram visualizados em gel de acrilamida corado com prata (figura 33).

Fig. 33 Membranas com proteínas recombinantes usadas na purificação de anticorpos. Membranas de nitrocelulose coradas com DB71 após transferência das proteínas recombinantes separadas por SDS-PAGE 1D. As regiões delimitadas por pontilhado, referentes aos pesos moleculares das proteínas recombinantes (MIC17B₂₀₋₃₀₃, membrana da esquerda; MIC17C₂₀₋₃₄₇, membrana da direita), foram recortadas para incubação com os soros produzidos contra as respectivas proteínas recombinantes. Os anticorpos purificados dos soros são mostrados à direita das membranas, separados por SDS-PAGE 1D em gel de acrilamida (12,5%) corado com prata. M, marcador de peso molecular *Precision Plus Protein All Blue Standards* (BioRad). Poço 1, anticorpo purificado contra MIC17B₂₀₋₃₀₃. Poço 2, anticorpo purificado contra MIC17C₂₀₋₃₄₇. Nas alturas de 25 kDa e 50 kDa são mostradas, respectivamente, as cadeias leve e pesada das imunoglobulinas.

4.1.9 Detecção das proteínas recombinantes e nativas por *Western blot*

O *Western blot* 1D (WB 1D) em que as proteínas recombinantes foram utilizadas como antígenos confirmou a afinidade dos anticorpos produzidos contra elas. MIC17B₂₀₋₃₀₃ foi detectada com peso compatível com o predito de 35.7 kDa, e MIC17C₂₀₋₃₄₇ também, com 40.2 kDa (figura 34.A).

As proteínas nativas MIC17B e MIC17C foram identificadas por WB 1D com os anticorpos purificados anti-MIC17B₂₀₋₃₀₃ e anti-MIC17C₂₀₋₃₄₇ sobre o extrato total de *N. caninum* (NcET) e sobre o extrato secretado produzido a partir do sobrenadante da cultura de *N. caninum* (NcESA_{SN}). O peso molecular predito é de 38 kDa para as proteínas íntegras, 35 kDa sem o peptídeo sinal, e 32 kDa sem o domínio de direcionamento para micronemas. Por comparação com WB 1D sobre os extratos de células Vero (VeET e VeESA_{SN}), a proteína nativa MIC17B foi detectada sobre NcET em três bandas fracas de pesos aproximados de 40 kDa, 35 kDa e 31 kDa; e sobre NcESA_{SN} em duas bandas fortes de aproximadamente 35 kDa e 29 kDa (figura 34.B). Duas bandas com altura de aproximadamente 65 kDa aparecem também nesta detecção com anticorpo anti-MIC17B₂₀₋₃₀₃ sobre NcESA_{SN}, porém

este peso está bastante acima do predito. MIC17C foi detectada sobre NcET em duas bandas fortes com aproximadamente 40 kDa e 35 kDa e outras duas bandas fracas com pesos aproximados de 31 kDa e 27 kDa; sobre NcESA_{SN}, MIC17C aparece em duas bandas fortes de aproximadamente 35 kDa e 27 kDa (figura 34.C).

Fig. 34 Detecção das proteínas recombinantes e nativas por Western blot 1D. Filmes fotográficos com revelação de membranas cujos antígenos foram detectados por anticorpos purificados. **A, detecção das proteínas recombinantes.** 1, MIC17B₂₀₋₃₀₃ detectada por anticorpo anti-MIC17B₂₀₋₃₀₃. 2, MIC17C₂₀₋₃₄₇ detectada por anticorpo anti-MIC17C₂₀₋₃₄₇. **B, detecção da proteína MIC17B nativa,** mostrada em setas na cor rosa, sobre NcET (poço 4: 40 kDa, 35 kDa e 31 kDa) e sobre NcESA_{SN} (poço 2: 35 kDa e 29 kDa). **C, detecção da proteína MIC17C nativa,** mostrada em setas na cor amarela, sobre NcET (poço 4: 40 kDa, 35 kDa, 31 kDa e 27 kDa) e sobre NcESA_{SN} (poço 2: 35 kDa e 27 kDa). Peso molecular predito: 38 kDa para as proteínas íntegras, 35 kDa sem o peptídeo sinal, e 32 kDa sem o domínio de direcionamento para micronemas. **Ac,** anticorpo. **M,** marcador de peso molecular pré-corado *Precision Plus Protein Dual Color* (BioRad). **VeESA_{SN},** extrato secretado de cultura de células Vero. **NcESA_{SN},** extrato secretado de cultura de *N. caninum* sobre células Vero. **VeET,** extrato total de células Vero. **NcET,** extrato total de taquizoítas de *N. caninum*.

Assim como no WB 1D, a proteína MIC17B foi detectada de forma mais intensa apenas no NcESA_{SN} por WB 2D. Em separação por gradiente de pH de 7 a 11 não-linear (NL), a proteína MIC17B nativa foi detectada em dois *spots* com os pontos isoeletricos (pI) 8.1 e 8.3, e mesmo peso molecular (PM) de 31 kDa, o que sugere que a proteína sofra modificação pós-traducional (figura 35.E). MIC17C nativa foi detectada por WB 2D tanto em NcET como em NcESA_{SN}. Em NcET, separado por

gradiente de pH de 3 a 11 (NL), MIC17C aparece em dois *spots* com PM de aproximadamente 30 kDa e pI em torno de 8.0 (figura 35.B). Uma melhor definição do pI de MIC17C foi alcançada nos WB com gradiente de pH de 7 a 11 (NL). Em NcET, MIC17C aparece em dois *spots* (1, pI 8.0 e PM 31 kDa; 2, pI 8.3 e PM 35 kDa), enquanto em NcESA_{SN} aparece em um *spot* de 35 kDa e pI 8.4 (figura 35.D e F). Os PM's e pI's das proteínas destacadas na figura 35 estão reunidos na tabela 16.

Fig. 35 Western blots bidimensionais (WB 2D) para a detecção das proteínas nativas MIC17B e MIC17C. Os três filmes à esquerda (A, C e E) são de WB para a busca de MIC17B nativa com o

anticorpo purificado anti-MIC17B₂₀₋₃₀₃, e os três filmes da direita (B, D e F) para a busca de MIC17C nativa com o anticorpo purificado anti-MIC17B₂₀₋₃₄₇. **A e B**, antígeno extrato total de *N. caninum* (NcET) separado em gradiente de pH 3 a 11 não linear (NL); **C e D**, antígeno NcET separado em gradiente de 7 a 11 NL; **E e F**, antígeno extrato secretado de *N. caninum* (NcESA_{SN}) separado em gradiente de 7 a 11 NL. **MIC17B** nativa foi detectada como dois *spots* em NcESA_{SN}, mostrados em setas rosas no filme E; e não foi detectada nos WB com NcET (A e C), onde os *spots* que aparecem são correspondentes à proteínas de células Vero. **MIC17C** nativa foi detectada em NcET (B e D) e em NcESA_{SN} (F) nos *spots* indicados por setas amarelas; os *spots* não indicados por setas correspondem à proteínas de células Vero. A relação dos pesos moleculares e pontos isoelétricos dos *spots* indicados nesta figura encontra-se na tabela 16. A marcação de peso molecular (kDa) foi feita com *Precision Plus Protein Dual Color* (BioRad). Separação em primeira dimensão em géis a 12,5%; segunda dimensão em *strips* de 7 cm.

Tabela 16. Descrição de peso molecular e ponto isoelétrico das proteínas detectadas por Western blot bidimensional com os anticorpos purificados anti-MIC17B₂₀₋₃₀₃ e anti-MIC17C₂₀₋₃₄₇. A numeração dos *spots* corresponde à localização mostrada na figura 35, especificada entre parênteses na primeira coluna da tabela. PM, Peso molecular em kDa. 2D, bidimensional. pl, ponto isoelétrico.

Western blot 2D (fig. 35)	Spot	pl	PM observado	PM predito forma íntegra/ sem peptídeo sinal/sem domínio de direcionamento para micronemas
MIC17B - NcESA (E)	1	8.1	31	38/35/32
	2	8.3	31	
MIC17C - NcET (D)	1	8.0	31	38/35/32
	2	8.3	35	
MIC17C - NcESA (F)	1	8.4	35	38/35/32

4.1.10 Identificação das proteínas MIC17B e MIC17C nativas por espectrometria de massas

A identificação de MIC17B e MIC17C não foi alcançada pela análise dos *spots* de gel 2D ou bandas de gel 1D por nenhuma das técnicas utilizadas. Foram identificadas algumas proteínas de células Vero excisadas dos géis de *N. caninum*, o que possibilitou desconsiderar o que era contaminação nos géis de extrato total de *N. caninum*.

As proteínas nativas MIC17B e MIC17C foram identificadas no extrato secretado de *N. caninum* (NcESA) por análise em sistema com separação por cromatografia líquida de fase reversa acoplado a espectrômetro de massas (RP-nanoLC-MS/MS). Dois peptídeos únicos foram usados na identificação de MIC17B (tabela 17). Um peptídeo foi usado na identificação de MIC17C, porém não é um peptídeo exclusivo, sendo comum para as três MIC17. Portanto, apesar de o resultado do Mascot ter apontado MIC17C, não há confirmação de que este peptídeo seja correspondente

exclusivamente a esta proteína. As informações sobre a identificação das proteínas e seus respectivos peptídeos são mostradas na tabela 17, e os espectros de fragmentação dos três peptídeos são mostrados na figura 36.

Tabela 17. Peptídeos identificados no extrato secretado de *N. caninum* (ESA) usados na identificação das proteínas MIC17B e MIC17C. Os peptídeos 1 e 2 são exclusivos de MIC17B. O peptídeo 3 é comum a MIC17A, MIC17B e MIC17C. MH+, massa do íon molecular.

	Sequência do peptídeo identificado	Score	Carga	Relação	MH+ (Da)	Identificação	Score /cobertura
				massa/carga monoisotópica m/z (Da)			
1	FFTFVSDKKR	37	3+	425.56815	1274.68988	MIC17B	36.82/5,70%
2	DNSNAFSGPK	25	2+	518.73767	1036.46807		
3	TEVGSFAACQLK	63	2+	655.82361	1310.63944	MIC17C	61.65/3.42%

Figura 36. Espectros de fragmentação (MS/MS) dos três peptídeos usados na identificação das proteínas MIC17 no extrato secretado (ESA) de *N. caninum*. A e B, peptídeos únicos da proteína MIC17B; C, peptídeo encontrado em MIC17A, MIC17B e MIC17C. Nas tabelas são mostrados os valores das massas dos íons (b, c, y e z) junto à sequência de aminoácidos do peptídeo identificado. Os espectros de MS/MS são mostrados em função da relação de massa/carga (m/z) e intensidade dos fragmentos.

4.1.11 Imunolocalização de MIC17B e MIC17C em taquizoítas de *N. caninum*

Na microscopia confocal de taquizoítas em cultura de células Vero, MIC17B e MIC17C nativas apareceram polarizadas para uma das extremidades do taquizoíta, o que sugere localização micronêmica (figura 37). Na imagem para detecção de MIC17C, o taquizoíta inferior apresenta uma constrição na região apical, característica da junção de movimento formada no momento da invasão.

Fig. 37 MIC17B e MIC17C localizadas em taquizoítas de *N. caninum* fixados em cultura de células Vero. Microscopia confocal, **A**, taquizoítas marcados com anticorpo purificado anti-MIC17B₂₀₋₃₀₃ (painel superior) ou anti-MIC17C₂₀₋₃₄₇ (painel inferior) e visualizados com anticorpo conjugado-FITC. **B**, Núcleos de taquizoítas e de células Vero marcados com iodeto de propídeo. **C**, colocalização de A e B. **D**, imagem por contraste de fase. As proteínas MIC17B e MIC17C com polarização no taquizoíta estão indicadas por setas brancas. Objetiva 100X. A escala é representada na régua de 5 µm.

Ainda em cultura de células Vero, a marcação de MIC17B indicou um taquizoíta possivelmente em fase inicial da invasão celular (figura 38). As imagens captadas de um mesmo campo em duas diferentes profundidades (z) mostram, no registro de uma profundidade (camada 1), MIC17B concentrada em uma extremidade do taquizoíta, provavelmente o conóide (figura 38, painel superior). No registro de outra profundidade (camada 2), MIC17B aparece circundando esta extremidade do taquizoíta, o que sugere o movimento das proteínas micronêmicas que vão se transferindo para a porção posterior do taquizoíta, na junção de movimento, durante a penetração da célula hospedeira (figura 38, painel inferior).

Fig. 38 MIC17B em taquizoíta de *N. caninum* em possível fase inicial de invasão celular. Microscopia confocal mostrando imagens geradas a partir de duas profundidades (z) diferentes de um mesmo campo. Figuras na parte superior, imagens da camada 1. Figuras na parte inferior, imagens da camada 2. **A**, MIC17B marcada com anticorpo purificado anti-MIC17B₂₀₋₃₀₃ e visualizada com anticorpo conjugado-FITC. **B**, Núcleos de taquizoítas e de células Vero marcados com iodeto de propídeo. **C**, imagem por contraste de fase. **D**, colocalização de A, B e C. A localização de MIC17B na extremidade do taquizoíta está indicada por setas brancas. No contraste de fase, indicação da região de contato com a célula. Objetiva 100X. A escala é representada na régua de 5 µm.

Em taquizoítas purificados, MIC17B foi nitidamente localizada na região apical do parasita, sugerindo fortemente sua localização nas micronemas (figura 39 A e C, painel superior). A extremidade na qual a proteína se concentra corresponde à mesma extremidade do conóide, por comparação com a imagem por contraste de fase (figura 39 D, painel superior). Alguns taquizoítas não apresentam marcação com MIC17B provavelmente por esta ser uma proteína secretada e a dinâmica de secreção ser rápida. MIC17C apresentou localização na região apical do taquizoíta, mas também distribuída em vários outros pontos, geralmente contornando o núcleo do taquizoíta (figura 39 A e C, painel inferior). Quando localizada na região apical, tem disposição diferente de MIC17B (figura 39 C, painel inferior). Controles negativos, apenas com incubação dos anticorpos conjugados com a cultura de *N. caninum* em Vero ou com os taquizoítas purificados, mostraram não haver reconhecimento destes por células Vero nem por *N. caninum*, tanto para conjugado-FITC como para conjugado-Alexa488 (dados não-mostrados).

Fig. 39 MIC17B e MIC17C nativas em taquizoítas de *N. caninum* purificados. Microscopia confocal, **A**, taquizoítas marcados com anticorpo purificado anti-MIC17B₂₀₋₃₀₃ (parte superior) ou anti-MIC17C₂₀₋₃₄₇ (parte inferior) e visualizados com anticorpo conjugado-Alexa488. **B**, Núcleos de taquizoítas marcados com iodeto de propídeo. **C**, colocalização de A e B. **D**, imagem por contraste de fase. MIC17B localizada na região apical do taquizoíta está indicada por setas brancas. MIC17C localizada na região apical, ao redor do núcleo e em outros pontos do taquizoíta está indicada em setas brancas. Objetiva 100X. A escala é representada na régua de 5 µm.

4.1.12 Ensaios de invasão

Os taquizoítas de *N. caninum* transfectados de forma transiente com o gene LacZ (NcLacZ transientes) foram pré-incubados com os anticorpos purificados ou os soros anti-MIC17B₂₀₋₃₀₃ ou anti-MIC17C₂₀₋₃₄₇, ou com o soro anti-NcET. A inibição da invasão foi de 89% com os soros anti-MIC17B₂₀₋₃₀₃ e anti-MIC17C₂₀₋₃₄₇, e de 83% com o soro anti-NcET. Porém, seus respectivos soros pré-imune também apresentaram alta inibição, de 64%, 74% e 60%, respectivamente (figura 40 A). A justificativa para este efeito com os soros pré-imune foi a lise exacerbada de taquizoítas durante a incubação com as amostras (figura 41). Para as três amostras, tanto os soros pré-imune quanto os anti-soros apresentaram alta absorbância para o sobrenadante removido após a incubação com taquizoítas, mostrando que muitos taquizoítas foram lisados e a β -galactosidase foi liberada no meio. Para eliminar a lise, as amostras foram submetidas à inativação do sistema complemento a 54°C por 20 minutos. A inibição da invasão com as amostras inativadas foi muito menor do que sem inativação, de 11%, 36% e 3% para os soros anti-MIC17B₂₀₋₃₀₃, anti-MIC17C₂₀₋₃₄₇ e anti-NcET, respectivamente (figura 40 B). Este resultado é compatível com a baixa lise de taquizoítas observada para as amostras inativadas (figura 41). Apenas o soro anti-MIC17C₂₀₋₃₄₇ mostrou uma inibição significativa comparada ao controle positivo, porém este efeito não é significativo quando comparado com seu soro pré-imune. Independentemente da inativação, os anticorpos purificados mostraram baixa inibição da invasão (figura 40 A e B). A lise causada pelos anticorpos purificados foi muito baixa, porém, o anticorpo anti-MIC17B₂₀₋₃₀₃ ao invés de inibir, promoveu a invasão, enquanto o anticorpo anti-MIC17C₂₀₋₃₄₇ inibiu por volta de 20% da invasão por *N. caninum* ($p < 0.05$).

Fig. 40 Efeito dos anticorpos e dos soros anti-MIC17B₂₀₋₃₀₃ e anti-MIC17C₂₀₋₃₄₇, e dos soro anti-NcET sobre a invasão celular por taquizoítas de *N. caninum*. **A**, ensaio realizado com soros não-inativados. **B**, ensaio realizado com amostras previamente inativadas para o sistema complemento. **NcLacZ transiente**, taquizoíta de *N. caninum* transfectado de forma estável com gene para β -galactosidase; **C+**, controle positivo (com taquizoítas em meio RPMI); **soro pré-imune**, soros pré-imune das três amostras respectivas; **anti-soro**, soros produzidos contra as proteínas recombinantes MIC17B₂₀₋₃₀₃ ou MIC17C₂₀₋₃₄₇, e contra NcET; **anticorpo purificado**, anticorpos purificados do soro contra as proteínas recombinantes.

Fig. 41 Lise de taquizoítas durante o ensaio de invasão. Os gráficos mostram absorbância a 570 nm para ensaios de invasão e de lise. À esquerda, ensaios com amostras não-inativadas. À direita, ensaios com amostras previamente inativadas para o sistema complemento. **A**, amostras anti-MIC17B₂₀₋₃₀₃; **B**, amostras anti-MIC17C₂₀₋₃₄₇; **C**, amostras anti-NcET. **C+**, controle positivo (com taquizoítas em meio RPMI); **pré-immune**, soros pré-immune das três respectivas amostras; **soro**, soros produzidos contra as proteínas recombinantes MIC17B₂₀₋₃₀₃ ou MIC17C₂₀₋₃₄₇, e contra NcET; **Ac**, anticorpos purificados do soro contra as proteínas recombinantes MIC17B₂₀₋₃₀₃ ou MIC17C₂₀₋₃₄₇.

O ensaio de invasão com o taquizoíta de *N. caninum* transfectado de forma estável com o gene LacZ (NcLacZ estável) garante que todos os taquizoítas produzam β -galactosidase de forma homogênea, evitando possíveis flutuações no resultado. Neste ensaio apenas os anticorpos purificados foram utilizados.

Com a suspeita de que a albumina sérica bovina (BSA), adicionada aos anticorpos recém-purificados para a estabilização destes, estaria proporcionando efeito antioxidante aos taquizoítas e favorecendo a invasão, um novo ensaio foi realizado com um controle positivo adicionado de BSA (C+ BSA) para concentração final igual a das amostras com anticorpos. Este controle positivo foi considerado como 100% para o cálculo de porcentagem de invasão com os anticorpos purificados (figura 42). Desta forma, o anticorpo anti-MIC17B₂₀₋₃₀₃ inibiu 24.2% da invasão, enquanto anti-MIC17C₂₀₋₃₄₇ inibiu 15.5% da invasão pelos taquizoítas. Ambos os valores para inibição são significativos quando comparados ao controle positivo com BSA ($p < 0.05$).

Fig. 42 Efeito dos anticorpos anti-MIC17B₂₀₋₃₀₃ e anti-MIC17C₂₀₋₃₄₇ sobre a invasão por taquizoítas de *N. caninum*. Ensaio com taquizoíta NcLacZ estável e detecção por reação com CPRG. **Eixo Y**, porcentagem de invasão (%) proporcional à absorbância obtida em 570 nm após reação com CPRG. **NcLacZ estável**, taquizoíta de *N. caninum* transfectado de forma estável com gene para β -galactosidase. **C-**, controle negativo, sem taquizoítas; **C+**, controle positivo de taquizoítas em RPMI adicionado de albumina sérica bovina (BSA); **anti-MIC17B**, amostra pré-incubada com anticorpo purificado anti-MIC17B₂₀₋₃₀₃; **anti-MIC17C**, amostra pré-incubada com anticorpo purificado anti-MIC17C₂₀₋₃₄₇. As porcentagens de invasão para as amostras pré-incubadas com anti-MIC17B (75,80%) e com anti-MIC17C (84,50%) são significativamente inferiores ao controle positivo (100%).

4.1.13 Ensaios de adesão (*binding*)

A expressão das proteínas recombinantes MIC17A₅₅₋₃₅₁, MIC17B₅₅₋₃₅₁ e MIC17C₅₅₋₃₅₁ na forma solúvel não foi alcançada. Os fragmentos das três sequências foram amplificados com sucesso por *Nested* PCR e clonados em pGEM. Porém, várias tentativas de clonagem em pET28-SUMO foram realizadas sem sucesso.

Na impossibilidade de expressão das proteínas recombinantes em pET28-SUMO, ensaios de adesão foram desenvolvidos para detectar a adesão das proteínas solúveis presentes no extrato secretado de *N. caninum* (ESA_{SN}), preparado em PBS.

No ensaio por *binding*-ELISA, células Vero foram incubadas com ESA_{SN}, e as proteínas MIC17B e MIC17C (presentes no ESA) foram em seguida detectadas pelos seus respectivos anticorpos purificados. Anticorpos secundários, conjugados à peroxidase, foram usados para a detecção dos anticorpos anti-MIC17B e anti-MIC17C. O controle positivo, com soro anti-NcET (C AcET), e as amostras incubadas com ESA_{SN} (A+Sn+Ac) apresentaram diferença significativa em relação ao controle negativo sem os anticorpos purificados (C -Ac), porém, o controle negativo preparado sem ESA_{SN} e sem anticorpos purificados (C -Sn -Ac) apresentou alta absorbância (figura 43). Os controles negativos sem ESA_{SN} (C -Sn), preparados especificamente para cada amostra, também apresentaram altos valores de absorbância, e também alto desvio padrão. Assim, apesar de em algumas condições haver diferença entre o resultado para as amostras e o resultado do controle sem ESA_{SN} (condições destacadas por setas na figura R17), não foi possível afirmar que a adesão por MIC17B ou MIC17C existe, pois o alto desvio padrão dos resultados torna as diferenças entre amostra e controle não-significativas.

Fig. 43 Ensaio de adesão das proteínas MIC17B e MIC17C, presentes em extrato secretado de *N. caninum* (NcESASN), às células Vero, por *binding*-ELISA. As proteínas aderidas foram detectadas pelos anticorpos purificados anti-MIC17B₂₀₋₃₀₃ e anti-MIC17C₂₀₋₃₄₇, que foram reconhecidos por anticorpos conjugados à peroxidase, cuja reação com TMB permite a leitura da absorbância em 450 nm, mostrada no eixo Y. **Sn**, NcESASN adicionados às células Vero, nas quantidades 0,5 µg, ou 1,0 µg por poço. **Ac**, anticorpo purificado anti-MIC17B₂₀₋₃₀₃ ou anti-MIC17C₂₀₋₃₄₇ adicionado após incubação com NcESASN, nas concentrações 0,2 ou 0,05 µg/mL. Controles: **C - Ac**, controle negativo sem incubação com anticorpo primário (com NcESASN); **C - Sn -Ac**, controle negativo sem NcESASN e sem anticorpo primário; **C AcET**, controle positivo com incubação com soro anti-NcET, como anticorpo primário; **C - Vero**, controle negativo sem células Vero (com NcESASN e com anticorpo purificado nas concentrações específicas para cada amostra); **C - Sn**, controle negativo sem NcESASN (com anticorpo purificado nas concentrações específicas para cada amostra). **A+Sn+Ac**, amostra (células Vero incubadas com NcESASN e com anticorpo anti-MIC17B₂₀₋₃₀₃ ou anti-MIC17C₂₀₋₃₄₇). As setas indicam as condições em que houve diferença entre o controle negativo C -Sn e as amostras (diferenças não significativas).

No ensaio de adesão por *binding*-Western blot, células Vero foram incubadas com NcESASN previamente tratado (t) com os anticorpos anti-MIC17B₂₀₋₃₀₃ ou anti-MIC17C₂₀₋₃₄₇, enquanto os controles foram incubados com NcESASN não-tratado (nt) ou com PBS. O conteúdo solubilizado de células e proteínas aderidas foi separado por SDS-PAGE 1D e transferido para membrana de PVDF para reconhecimento pelos anticorpos anti-MIC17B₂₀₋₃₀₃ ou anti-MIC17C₂₀₋₃₄₇, e também por soro anti-NcET.

Caso as proteínas MIC17B e MIC17C ficassem aderidas às células Vero, no gel em condições desnaturantes, bandas com alturas referentes a essas proteínas deveriam estar presentes no controle não-tratado com anticorpos. Nas amostras

tratadas com anti-MIC17B₂₀₋₃₀₃ ou anti-MIC17C₂₀₋₃₄₇, deveriam ser observadas bandas mais fracas ou ausência de bandas; e no controle com PBS as bandas de MIC17B e MIC17C deveriam estar ausentes. Este resultado não foi observado (figura 44). As bandas referentes à MIC17B e MIC17C, entre 29 e 37 kDa, não foram observadas nem nas amostras tratadas (figura 44 B) tampouco nos controles não-tratados (figura 44 A). As únicas bandas observadas correspondem às proteínas de células Vero reconhecidas pelos anticorpos purificados (figura 44 A) e estão destacadas por setas brancas. Elas apareceram nas fitas detectadas por anti-MIC17B₂₀₋₃₀₃, mais intensas no controle com PBS e um pouco mais fracas no controle não-tratado. No controle não-tratado com detecção por soro anti-NcET, foi observada uma banda de peso um pouco inferior a 50 kDa, que não estava presente no controle com PBS, porém, não é possível saber qual entre as proteínas do ESA foi ali detectada (figura 44 C).

Fig. 44 Ensaio de adesão em células Vero das proteínas MIC17B e MIC17C presentes em extrato secretado de *N. caninum* (NcESA_{SN}) pela técnica de *Western blot*. Na descrição de cada poço encontram-se: a solução em que as células Vero foram incubadas (ESA_{nt}, NcESA_{SN} não-tratado; PBS; ou ESA_t, NcESA_{SN} tratado com anticorpo purificado anti-MIC17B₂₀₋₃₀₃ ou anti-MIC17C₂₀₋₃₄₇), e o anticorpo com o qual foi feita a detecção sobre a membrana de PVDF (Ac M17B, anticorpo purificado anti-MIC17B₂₀₋₃₀₃; Ac M17C, anticorpo purificado anti-MIC17C₂₀₋₃₄₇; Ac NcET, soro contra extrato total de *N. caninum*). **A**, fitas com controle de NcESA_{SN} não-tratado e controle de PBS, e detecção pelos anticorpos purificados; **B**, fitas com NcESA_{SN} tratadas com anti-MIC17B₂₀₋₃₀₃ ou anti-MIC17C₂₀₋₃₄₇, e detecção pelos anticorpos purificados; **C**, fitas com controle de NcESA_{SN} não-tratado e controle de PBS, e detecção por soro anti-NcET. **Setas brancas** indicam bandas referentes a proteínas de células Vero. **M**, marcador de peso molecular *Precision Plus Protein Dual Color* (BioRad).

No ensaio de adesão por *binding*-FACS, NcESA_{SN} foi incubado com células Vero tripsinizadas, e a detecção das proteínas MIC17B e MIC17C possivelmente aderidas a elas foi feita com os anticorpos purificados anti-MIC17B₂₀₋₃₀₃ e anti-MIC17C₂₀₋₃₄₇. Para a leitura em citômetro de fluxo, as células foram ainda incubadas com anticorpo secundário conjugado a FITC. O resultado para MIC17B e MIC17C, expresso em porcentagem de células com fluorescência sobre a população principal (% *Parent*), ficou bastante acima do controle com apenas células Vero (figura 45. 1), do controle com Vero e anticorpo conjugado-FITC (figura 45. 2), e do controle sem anticorpo primário (figura 45. 3). Porém, em relação ao controle sem extrato secretado (figura 45. 4 e 6), as amostras incubadas tanto com o anticorpo anti-MIC17B₂₀₋₃₀₃ como com o anticorpo anti-MIC17C₂₀₋₃₄₇ (figura 45. 5 e 7) apresentaram menor porcentagem de células fluorescentes. Já a amostra incubada com soro contra extrato total (figura 45. 9) apresentou resultado bem acima de seu respectivo controle sem extrato secretado (figura 45. 8).

Fig. 45 Ensaio de adesão a células Vero pelas proteínas MIC17B e MIC17C presentes em extrato secretado de *N. caninum* (NcESA_{SN}), pela técnica de *binding*-FACS. Os resultados são expressos em porcentagem de células com fluorescência sobre a população principal (% *Parent*) e estão dispostos sobre a barra de cada amostra. 1, controle com apenas células Vero; 2, controle com apenas células Vero e anticorpo conjugado-FITC; 3, controle sem anticorpo primário. 4, controle para a amostra MIC17B, sem extrato secretado; 5, amostra MIC17B, com detecção por anticorpo anti-MIC17B₂₀₋₃₀₃; 6, controle para a amostra MIC17C, sem extrato secretado; 7, amostra MIC17C, com detecção por anticorpo anti-MIC17C₂₀₋₃₄₇; 8, controle para a amostra NcET, sem extrato secretado; 9,

amostra NcET, com detecção por soro contra extrato total; **Ve**, células Vero; **ESA**, extrato secretado de *N. caninum*; **Ac**, anticorpo primário; **FITC**, anticorpo secundário conjugado a FITC; **AcM17B**, anticorpo purificado anti-MIC17B₂₀₋₃₀₃; **AcM17C**, anticorpo purificado anti-MIC17C₂₀₋₃₄₇; **AcET**, soro contra extrato total de *N. caninum*.

4.1.14 Construção do plasmídeo pΔApple e seleção de taquizoítas de *N. caninum* nocaute para MIC17A, MIC17B e MIC17C

O plasmídeo pΔApple foi construído com sucesso, como descrito na figura 15. Todas as etapas de conferência por PCR e por digestão da construção deste plasmídeo são mostradas na figura 46.

Fig. 46 Etapas de construção do plasmídeo pΔApple para transfecção de *N. caninum* e nocaute dos genes de MIC17A, MIC17B e MIC17C. Géis de agarose (0,8%). M, marcador de tamanho molecular (pb) *GeneRuler 1kb DNA Ladder* (Fermentas). **A**, PCR dos fragmentos PROM_M17A e 3'UTR_M17C sobre o DNA de *N. caninum*; 1, controle negativo (C-) para PROM_M17A; 2, amplificação de PROM_M17A; 3, C- para 3'UTR_M17C; 4, amplificação de 3'UTR_M17C. **B**, 1, produto da digestão do plasmídeo pGEM após inserção de PROM_M17A, para purificação do fragmento liberado. **C**, 1, produto da digestão do plasmídeo pGEM após inserção de 3'UTR_M17C, para purificação do fragmento liberado. **D**, 1, produto da digestão do plasmídeo pSAGCAT para

remoção de 3'SAG1, e purificação do plasmídeo subtraído. **E, 1**, Digestão para conferência do plasmídeo pSAGCAT após inserção de 3'UTR_M17C, substituindo 3'SAG1. **F, 1**, produto da digestão do plasmídeo pSAGCAT+3'UTR_M17C para remoção de PROM_SAG1, e purificação do plasmídeo subtraído (banda não identificada, suspeita de plasmídeo não digerido). **G**, PCR para conferência da presença de 3'UTR_M17C e da construção correta de 3'UTR_M17C-CAT no plasmídeo pSAGCAT+3'UTR_M17C subtraído de PROM_SAG1; 1, C- para 3'UTR_M17C; 2, amplificação de 3'UTR_M17C; 3, C- para 3'UTR_M17C-CAT; 4, amplificação de 3'UTR_M17C-CAT; 5, controle da amplificação de 3'UTR_M17C-CAT sobre o plasmídeo pSAGCAT original. **H, 1**, Digestão para conferência do plasmídeo pSAGCAT+3'UTR_M17C+PROM_M17A (=pΔApple) após inserção de PROM_M17A, substituindo PROM_SAG1. **I, 1**, Digestão para conferência de 3'UTR_M17C no plasmídeo pΔApple após inserção de PROM_M17A. **J**, PCR para conferência da construção PROM_M17A-CAT no plasmídeo pΔApple; 1, amplificação de PROM_M17A-CAT; 2, C- para PROM_M17A-CAT; 3, controle da amplificação de PROM_M17A-CAT sobre pSAGCAT original. **L**, PCR para conferência da construção CAT-3'UTR_M17C no plasmídeo pΔApple; 1, amplificação de CAT-3'UTR_M17C; 2, C- para CAT-3'UTR_M17C; 3, controle da amplificação de CAT-3'UTR_M17C sobre pSAGCAT original. **M**, Digestão para conferência do plasmídeo pΔApple que, quando digerido em seus dois sítios com EcoRI, libera um fragmento de 1337 pb.

Após a transfecção dos taquizoítas NcLacZ, tanto com o plasmídeo circular (NcΔ1) quanto com o plasmídeo linearizado (NcΔ2), houve uma multiplicação rápida dos taquizoítas nos três primeiros dias. Depois desse período e por todos os meses de seleção com cloranfenicol, a multiplicação ficou excessivamente lenta. Foi possível observar “ninhos” de taquizoítas, porém eles não eram visíveis no sobrenadante da cultura, e para cada reinoculação em novas garrafas foi necessário raspar as células Vero aderidas à garrafa.

Ao final do período de 5 meses de seleção, foi extraído DNA dos taquizoítas transfectados NcΔ1 e NcΔ2, e do taquizoíta NcLacZ como controle. Na PCR para conferência da presença das proteínas Apple no genoma de cada um desses parasitas, foi realizada amplificação da sequência completa de MIC17B, que foi positiva no controle NcLacZ e também nos taquizoítas NcΔ1 e NcΔ2, mostrando que esse gene não foi nocauteado (figura 47. 1 a 4). As bandas referentes à MIC17B, apesar de fracas, devido à baixa quantidade de DNA obtida dos poucos taquizoítas sobreviventes, estão presentes.

Também foram conferidas por PCR as construções do plasmídeo pΔApple, que poderia estar integrado ou não ao genoma. Foram verificadas as construções PROM_M17A-CAT e CAT-3'UTR_M17C nos DNAs extraídos de NcΔ1, NcΔ2 e NcLacZ e, como controle, no plasmídeo pΔApple. O resultado foi negativo para as duas construções nos DNAs dos três taquizoítas, confirmando que o plasmídeo não foi integrado ao genoma e também estava ausente na forma não integrada (figura 47. 5 a 8, e 9 a 12).

Fig. 47 Conferência por PCR do nocaute das proteínas Apple pelo plasmídeo pΔApple. Foi conferida a presença do gene para MIC17B e da construção de pΔApple nos DNAs de NcΔ1, NcΔ2 e NcLacZ. As amplificações no plasmídeo pΔApple foram usadas como controle positivo para as construções PROM_M17A-CAT e CAT-3'UTR_M17C, e no DNA NcLacZ como controle negativo. Para a amplificação do gene MIC17B, NcLacZ foi usado como controle positivo, e o plasmídeo pΔApple como controle negativo. **1 a 4**, amplificação da sequência completa de MIC17B completa sobre os DNAs de NcΔ1 (1), NcΔ2 (2), NcLacZ (3), e sobre pΔApple (4). **5 a 8**, amplificação da construção PROM_M17A-CAT sobre os DNAs de NcΔ1 (1), NcΔ2 (2), NcLacZ (3), e sobre pΔApple (4). **9 a 12**, amplificação da construção CAT-3'UTR_M17C sobre os DNAs de NcΔ1 (1), NcΔ2 (2), NcLacZ (3), e sobre pΔApple (4). As **setas** indicam os resultados positivos para as amplificações. **NcΔ1**, taquizoíta NcLacZ transfectado com pΔApple circular e selecionado com 20mM de cloranfenicol; **NcΔ2**, taquizoíta NcLacZ transfectado com pΔApple linearizado e selecionado com 15mM de cloranfenicol; **NcLacZ**, taquizoíta que possui o gene para β-galactosidase inserido em seu genoma; **(+)**, resultado positivo para amplificação; **(-)**, resultado negativo para amplificação. Gel de agarose (0,8%) . **M**, marcador de tamanho molecular (pb) *GeneRuler 1kb DNA Ladder* (Fermentas).

4.2 Secretoma de *N. caninum*

4.2.1 Viabilidade de taquizoítas

O tratamento por aquecimento de taquizoítas a 85°C por 10 minutos foi eficaz para matar os taquizoítas, resultando em 92,7% de taquizoítas mortos. Utilizando este controle, foi confirmado que na amostra (preparada logo após o processo de purificação) 90,9% dos taquizoítas de *N. caninum* permaneceram vivos, portanto, adequados para a obtenção do extrato secretado (ESA) (figura 48).

Fig. 48 Avaliação por citometria de fluxo da viabilidade de taquizoítas de *N. caninum* recém-purificados. A, amostra composta por taquizoítas recém-purificados marcados com iodeto de propídeo (PI). B, controle composto por taquizoítas mortos por aquecimento (85°C, 10 minutos) marcados com iodeto de propídeo. A(1) e B(1), citogramas de *forward scatter* (FSC, volume celular) por *side scatter* (SSC, complexidade celular). A(2) e B(2), citogramas de contagem (*count*) por emissão de iodeto de propídeo (PI-A). P1, população principal selecionada para análise da viabilidade das células; P2 e P3, porcentagem de células **vivas** (negativo para PI) e **mortas** (positivo para PI), respectivamente, baseado na população de células em P1. Nos taquizoítas recém-purificados (A.2), 90,9% vivos (P2) e 9,1% mortos (P3); nos taquizoítas controle (B.2), 7,3% vivos (P2) e 92,7% mortos (P3).

4.2.2 Identificação de proteínas de *N. caninum* separadas por géis bidimensionais e analisadas por MALDI

De 31 *spots* excisados dos géis de NcET e NcESA, 29 proteínas diferentes foram identificadas. Destas, 22 estavam presentes no gel de NcESA. Na figura 49 são mostrados os géis e o posicionamento dos *spots*; em vermelho estão destacadas as proteínas identificadas como de *N. caninum*, e em azul as proteínas identificadas como de células Vero. No gel de NcET são mostradas algumas proteínas importantes de *N. caninum*. Todas as identificações estão listadas na tabela 18, entre as quais encontram-se a proteína micronêmica MIC1 e as proteínas de grânulos densos GRA2 e GRA7.

Tabela 18 Identificação de proteínas presentes no extrato total (ET) e extrato secretado (ESA) do taquizoíta de *N. caninum*. Colunas: **Spot**, numeração dos spots mostrada nos géis da figura 49; **ET e ESA**, mostra a concentração relativa dos spots nos géis (↑, baixa; ↑↑, média; ↑↑↑, alta; N.O., não-observado); **Identificação**, segundo o ToxoDB; **Descrição**, encontrada na literatura ou no ToxoDB; **PS**, presença (x) de peptídeo sinal; **TM**, predição de domínios transmembrana (quantidade de domínios); **Peso molecular (kDa)** predito e observado nos géis; **pI**, ponto isoelétrico predito.

Spot	ET	ESA	Identificação (ToxoDB)	Descrição (ToxoDB ou literatura)	PS	TM	Peso molecular (kDa)		pI
							Predito	Observado	Predito
1	↑↑↑	N.O.	NCLIV_064380	hypothetical protein		12	81,6	81,9	7,7
2	↑↑↑	N.O.	NCLIV_019110	HSP90-like protein, related	X	1	97,1	81,0	4,6
3s	↑↑↑	↑↑	NCLIV_026710	hypothetical protein (Noc2 family)		0	113,9	61,0	9,3
4	↑	↑	NCLIV_020220	elongation factor 2, putative		0	93,0	60,5	6,3
5s	↑↑↑	↑↑	NCLIV_055490	heat shock protein 70 (HSP70)	X	0	72,9	59,2	4,9
6s	↑↑↑	↑	NCLIV_055490	heat shock protein 70 (HSP70)	X	0	72,9	59,5	4,9
7s	↑↑↑	↑	NCLIV_066960	hypothetical protein (nuclear protein Es2)		0	88,8	51,4	10,7
8	↑↑	↑	NCLIV_009770	patched transmembrane domain-containing protein		11	165,4	50,3	6,9
9	↑	N.O.	NCLIV_058890	tubulin alpha chain		0	50,1	49,5	4,7
10	↑↑	↑	NCLIV_005150	tubulin beta chain		0	50,1	47,5	4,4
11s	↑	↑↑↑	NCLIV_043270	microneme protein MIC1 (MIC1)	X	0	49,1	47,2	4,6
12s	↑	↑↑↑	NCLIV_043270	microneme protein MIC1 (MIC1)	X	0	49,1	47,2	4,6
13s	↑↑	↑↑	NCLIV_037500	enolase related		0	52,0	47,0	6,6
14s	↑↑↑	↑↑	NCLIV_011410	protein disulfide isomerase	X	0	52,9	45,0	4,9
15s	↑	↑↑	NCLIV_011270	hypothetical protein (phosphoglycerate kinase)		0	44,7	42,7	5,6
16	↑↑↑	↑	NCLIV_003440	actin related		0	41,9	41,0	4,8
			NCLIV_025080	hypothetical protein		0	39,9	41,0	9,9
17s	↑	↑↑	NCLIV_050370	fructose-bisphosphate aldolase, class-I		0	33,1	37,5	6,8
18	↑	N.O.	NCLIV_032330	malate dehydrogenase (NAD)		1	35,6	35,2	5,8
19s	↑	↑↑↑	NCLIV_032330	malate dehydrogenase (NAD)		1	35,6	35,6	5,8
20	↑	N.O.	NCLIV_057460	transketolase central region	X	0	44,4	35,1	7,1
21s	↑	↑↑	NCLIV_054700	uridine phosphorylase		0	32,8	31,1	6,5
22	↑	N.O.	NCLIV_046550	elongation factor 1-beta		0	26,8	30,8	4,4
23	↑	N.O.	NCLIV_041850	Hsp20/alpha crystallin domain-containing protein		0	29,7	28,3	5,5
24	↑	N.O.	NCLIV_003650	hypothetical protein (cytochrome c oxidase)	X	0	37,5	26,2	6,5
25s	↑	↑↑	NCLIV_046900	hypothetical protein (triosephosphate isomerase)		0	27,7	26,1	5,9
26s	↑↑	↑↑	NCLIV_032390	hypothetical protein (thioredoxin-like)		0	25,8	25,1	7,5
27	↑	↑↑↑	NCLIV_021640	dense granule protein 7 (GRA7)	X	2	22,5	23,0	4,3
28s	↑	↑↑	NCLIV_045650	28 kDa antigen (GRA2)	X	0	22,4	22,5	9,2
29	↑	N.O.	NCLIV_006720	hypothetical protein	X	0	22,6	22,0	5,2
30	↑↑	↑	NCLIV_062630	thioredoxin-dependent peroxide reductase		0	22,0	21,5	6,4
31s	N.O.	↑	NCLIV_019970	peptidyl-prolyl cis-trans isomerase A		0	18,8	18,2	7,3

4.2.3 Proteínas identificadas no ESA de *N. caninum* por LC-MS

O extrato secretado (ESA), obtido por estímulo do parasita com 1% de etanol, foi separado em 13 bandas por SDS-PAGE 1D e digerido (figura 50). A combinação de diferentes métodos de fragmentação permitiu um grande número de identificações. De um total de 4.941 peptídeos identificados, a maioria foi alcançada por HCD-FT (3.446), seguida por ETD-IT (1.865), CID-IT (1.408), e ETD-FT (475).

Fig. 50 Extrato secretado (ESA), obtido por estímulo secretório com 1% de etanol, separado por SDS-PAGE 1D. Gel de acrilamida (12%) corado com Coomassie Blue G-250. As 13 bandas delimitadas foram excisadas e as proteínas digeridas para análise por espectrometria de massas (LC-MS). Marcador de peso molecular (kDa), *BenchMark Protein Ladder*.

Com os filtros apropriados aplicados aos resultados, um total de 615 proteínas foram identificadas no ESA de taquizoítas de *N. caninum* (figura 51). A lista com as informações: probabilidade de identificação para proteína, número de espectros, número de peptídeos únicos, e abundância relativa no ESA (número de espectros dividido pelo peso molecular) está disponível no Apêndice 1.

Fig. 51 Localizacao predita de proteínas no extrato secretado (ESA) de taquizoítas de *N. caninum*, de um total de 615 proteínas identificadas por LC-MS (em porcentagem). As proteínas

foram classificadas de acordo com a localização no parasita (confirmada ou putativa), e por vias de secreção preditas: clássica, por presença de peptídeo sinal; e não-clássica, por predição no programa *Secretome P*.

Entre as 615 proteínas identificadas no ESA de *N. caninum* (figura 51), 37.4% (n=230) corresponderam a proteínas secretadas ou putativamente secretadas, sendo 3.1% (n=19) de proteínas micronêmicas, 4.2% (n=26) de proteínas de roptrias, 1.5% (n=9) de proteínas de grânulos densos, 7.6% (n=47) de proteínas com localização desconhecida cuja sequência contém peptídeo sinal, e 21% (n=129) de proteínas com localização desconhecida e preditas para secreção não-clássica. Proteínas de superfície, que podem acumular no ESA como resultado de clivagem proteolítica do parasita⁵⁵, representaram 2.0% (n=12) das proteínas. Proteínas originadas em organelas não-secretórias também estavam presentes no ESA, sendo 1.3% de proteínas de apicoplasto (n=8), 1.0% de mitocôndria (n=6), 5.8% de citoplasma (n=36), e 5.0% de núcleo (n=31). Proteínas com localização desconhecida e ausência de sinais para secreção (“desconhecida”) totalizaram 47.5%. As proteínas identificadas com localização nas organelas secretórias e superfície estão listadas na tabela 19; e as proteínas de organelas não secretórias e citoplasmáticas estão na tabela 20.

Tabela 19 Proteínas identificadas por LC-MS no ESA de *N. caninum* com localização relacionada à secreção. Classificação por localização predita ou conhecida para organelas secretórias (micronemas, roptrias e grânulos densos) e para a superfície do parasita. Organização por ordem decrescente de abundância relativa no ESA, calculada por divisão do número de espectros pelo peso molecular da proteína.

ToxoDB ID	Descrição	Abundância relativa
PROTEÍNAS DE MICRONEMAS		
NCLIV_021050	SUB1	33.70
NCLIV_028680	AMA1	33.13
NCLIV_043270	MIC1	23.00
NCLIV_033690	MIC2-like1	5.60
NCLIV_020720	MIC11	5.27
NCLIV_051970	M2AP	4.48
NCLIV_061760	MIC6	3.38
NCLIV_010600	MIC3	3.34
NCLIV_066250	MIC10	3.13
NCLIV_002940	MIC4	2.89
NCLIV_062770	MIC8	2.73
NCLIV_022970	MIC2	1.52
NCLIV_020990	PLP1	0.88
NCLIV_000860	SPATR	0.84
NCLIV_041880	MIC16	0.73
NCLIV_000740	CPL1	0.54
NCLIV_044230	TLN4	0.33
NCLIV_038110	MIC17B	0.13
NCLIV_063730	MIC15	0.02

PROTEÍNAS DE ROPTRIAS		
NCLIV_011700	hypothetical protein (ROP15B)	10.06
NCLIV_031550	predicted rhoptry protein kinase (ROP55)	6.46
NCLIV_001970	ROP7	3.69
NCLIV_011730	ROP26	3.37
NCLIV_018420	ROP9	2.73
NCLIV_012920	ROP40	2.09
NCLIV_007770	rhoptry kinase family protein (BPK1)	1.36
NCLIV_051340	toxofilin	1.25
NCLIV_041930	RAB11	0.96
NCLIV_028170	ROP20	0.87
NCLIV_035860	ROP2L10 (ROP32)	0.45
NCLIV_053840	ROP1	0.36
NCLIV_021100	ROP12	0.35
NCLIV_070010	RON8	0.31
NCLIV_017440	predicted rhoptry protein kinase (ROPK) (ROP19)	0.30
NCLIV_030050	RON4	0.27
NCLIV_011690	ROP15	0.23
NCLIV_055850	ROP13	0.20
NCLIV_044410	ROP35	0.18
NCLIV_048590	RON3	0.17
NCLIV_007800	RON4L1	0.13
NCLIV_017420	predicted rhoptry protein kinase (ROP47)	0.11
NCLIV_064620	RON2	0.08
NCLIV_055360	RON5	0.07
NCLIV_069590	ROP2L8 (ROP24)	0.06
NCLIV_057550	SUB2	0.03
PROTEÍNAS DE GRÂNULOS DENSOS		
NCLIV_066630	GRA9	2.54
NCLIV_063150	serpin peptidase inhibitor, related	2.43
NCLIV_068400	NTPase II	1.81
NCLIV_045650	GRA2	1.59
NCLIV_005560	DG32	0.93
NCLIV_045870	GRA3	0.91
NCLIV_052880	GRA6	0.57
NCLIV_036400	GRA1	0.40
NCLIV_003100	PI-2	0.29
PROTEÍNAS DE SUPERFÍCIE		
NCLIV_033250	SRS2 (SRS29C)	14.26
NCLIV_010720	SRS domain-containing protein	2.66
NCLIV_033230	SAG1 (SRS29B)	1.42
NCLIV_060660	SAG3 (SRS57, P43)	0.87
NCLIV_068920	SRS domain-containing protein	0.76
NCLIV_010730	SRS domain-containing protein	0.46
NCLIV_015870	SRS23	0.16
NCLIV_068872	SRS domain-containing protein	0.16
NCLIV_052740	SRS domain-containing protein	0.14
NCLIV_034381	SRS domain-containing protein	0.12
NCLIV_023620	SRS domain-containing protein	0.12
NCLIV_060700	SRS domain-containing protein	0.05

Tabela 20. Proteínas identificadas por LC-MS no ESA de *N. caninum* com localização não relacionada à secreção. Classificação por localização predita ou conhecida para apicoplasto, mitocôndria, núcleo e citoplasma. Organização por ordem decrescente de abundância relativa no ESA, calculada por divisão do número de espectros pelo peso molecular da proteína. h.p., hypothetical protein.

ToxoDB ID	Descrição (ToxoDB)	Abundância relativa
PROTEÍNAS DE APICOPLASTO		
NCLIV_004190	thioredoxin, putative	3.58
NCLIV_058810	superoxide dismutase	3.00
NCLIV_058830	superoxide dismutase (EC 1.15.1.1), related	1.19
NCLIV_046260	ACN/IRP - aconitate hydratase/iron regulatory protein	0.82
NCLIV_036250	asparaginyl-tRNA synthetase, related	0.73
NCLIV_056480	h.p. (isocitrate dehydrogenase)	0.69
NCLIV_004340	FABZ - beta-hydroxyacyl-ACP dehydratase	0.60
NCLIV_016800	TCP-1/cpn60 chaperonin family protein, putative	0.03
PROTEÍNAS DE MITOCÔNDRIA		
NCLIV_062630	thioredoxin-dependent peroxide reductase, mitochondrial, related	2.77
NCLIV_051840	h.p. (mitochondrial glycoprotein domain-containing protein)	0.84
NCLIV_046260	ACN/IRP - aconitate hydratase/iron regulatory protein	0.82
NCLIV_013180	GM04207p, related	0.17
NCLIV_007730	valyl-tRNA synthetase, mitochondrial, related	0.05
NCLIV_045460	mitochondrial presequence protease (EC 3.4.24.), related	0.05
PROTEÍNAS DO NÚCLEO		
NCLIV_015920	histone H4, related	20.36
NCLIV_043670	high mobility group protein, related	3.27
NCLIV_026140	histone H2A, related	3.00
NCLIV_024230	Zgc:123215, related (HMG box domain-containing protein)	1.91
NCLIV_041100	novel protein (Zgc:66430), related	1.63
NCLIV_033040	Rbm28 protein, related (RNA binding motif-containing protein)	1.56
NCLIV_001070	histone H2B, related (histone H2Ba)	1.08
NCLIV_018560	ras-GTPase-activating protein binding protein, putative	1.02
NCLIV_040540	h.p. (heterogeneous nuclear ribonucleoprotein A3)	1.00
NCLIV_026150	histone H3, related	1.00
NCLIV_035950	h.p. (small nuclear ribonucleoprotein Sm D2)	0.75
NCLIV_038760	mago nashi protein, putative	0.61
NCLIV_038830	Sm protein, related (small nuclear ribonucleoprotein)	0.36
NCLIV_005360	splicing factor 3a protein, putative	0.27
NCLIV_045400	h.p. (splicing factor 3A subunit 2)	0.24
NCLIV_053930	small nuclear ribonucleoprotein U1A, putative	0.19
NCLIV_018710	h.p. (nucleosome assembly protein)	0.16
NCLIV_029040	nucleosome assembly protein, putative	0.14
NCLIV_030120	importin beta-3 subunit, putative	0.13
NCLIV_042270	h.p. (phenylalanyl tRNA synthetase isoform)	0.13
NCLIV_065770	nuclear cap binding protein, putative	0.12
NCLIV_031970	h.p. (pre-mRNA splicing factor PRP8)	0.11
NCLIV_031750	importin, putative	0.11
NCLIV_055800	serine/threonine protein phosphatase 5 (EC 3.1.3.16), related	0.09
NCLIV_020480	h.p. (U1 small nuclear ribonucleoprotein)	0.07
NCLIV_007390	MGC78841 protein, related (importin alpha)	0.03
NCLIV_065560	h.p.	0.03
NCLIV_046750	h.p. (poly(A) polymerase)	0.03
NCLIV_031340	splicing factor 3B subunit 3, putative	0.02
NCLIV_002910	ftsJ-like methyltransferase domain-containing protein, putative	0.01
NCLIV_069920	transcription factor S-II central domain-containing protein, putative	0.01
PROTEÍNAS CITOPLASMÁTICAS		
NCLIV_053860	MGC84926 protein, related	5.56
NCLIV_040860	Tryptophanyl-tRNA synthetase (EC 6.1.1.2), related	2.90

NCLIV_063590	Glutathione reductase, related	2.27
NCLIV_060420	h.p.	2.19
NCLIV_030290	transaldolase, putative	2.02
NCLIV_062230	Glutaminyl-tRNA synthetase (EC 6.1.1.18), related	1.84
NCLIV_062220	Glutaminyl-tRNA synthetase (EC 6.1.1.18), related	1.59
NCLIV_033040	Rbm28 protein, related	1.56
NCLIV_066760	Translationally-controlled tumor protein, related	1.37
NCLIV_049030	h.p.	1.19
NCLIV_066920	h.p.	1.04
NCLIV_036250	Asparaginyl-tRNA synthetase, related	0.73
NCLIV_034160	Glutamine synthetase (EC 6.3.1.2), related	0.63
NCLIV_019960	lysyl-tRNA synthetase, putative	0.63
NCLIV_004340	h.p.	0.60
NCLIV_011900	Arginyl-tRNA synthetase (EC 6.1.1.19), related	0.45
NCLIV_060470	h.p.	0.43
NCLIV_060760	prolyl-tRNA synthetase, putative	0.39
NCLIV_001670	Elongation factor 1-alpha, related	0.37
NCLIV_041650	Methionine--tRNA ligase (EC 6.1.1.10), related	0.25
NCLIV_038170	Hypoxanthine phosphoribosyltransferase (EC 2.4.2.8), related	0.14
NCLIV_061050	h.p.	0.13
NCLIV_042270	h.p.	0.13
NCLIV_030120	importin beta-3 subunit, putative	0.13
NCLIV_031750	importin, putative	0.11
NCLIV_044440	h.p.	0.10
NCLIV_029970	TRNA synthetase Gly, related	0.10
NCLIV_055800	Serine/threonine protein phosphatase 5 (EC 3.1.3.16), related	0.09
NCLIV_061570	D-tyrosyl-tRNA(Tyr) deacylase (EC 3.1.-.-), related	0.08
NCLIV_053190	h.p.	0.05
NCLIV_007730	Valyl-tRNA synthetase, mitochondrial, related	0.05
NCLIV_022400	h.p.	0.05
NCLIV_007390	MGC78841 protein, related	0.03

Um grupo interessante encontrado no ESA, incluído na categoria “desconhecida” na figura 51, é formado por ciclofilinas (1.3%, n=8), em que apenas uma (18 kDa *cyclophilin*) foi previamente identificada e caracterizada⁹⁴ (tabela 21).

Tabela 21 Ciclofilinas identificadas no ESA de *N. caninum*. Organização ordem decrescente de abundância relativa no ESA. Abundância relativa, número de espectros dividido pelo peso molecular.

ToxoDB ID	Descrição	Abundância Relativa
NCLIV_004790	18 kDa cyclophilin	3.63
NCLIV_019970	peptidyl-prolyl cis-trans isomerase A	2.68
NCLIV_036200	cyclophilin, putative	0.43
NCLIV_045550	RNA recognition motif-containing protein, putative	0.26
NCLIV_028870	peptidylprolyl isomerase D (Cyclophilin D), related	0.10
NCLIV_017250	peptidylprolyl isomerase (cyclophilin)-like 2, related	0.08
NCLIV_070210	peptidyl-prolyl cis-trans isomerase, putative	0.03
NCLIV_009630	cyclophilin, putative	0.02

4.2.4 Quantificação relativa de proteínas no taquizoíta *estimulado*

Seguindo-se à cromatografia de troca iônica (SCX), as 15 frações mais intensas (9-24) contendo a mistura dos peptídeos marcados por dimetil-isótopos (figura 52) foram selecionadas para análise por LC-MS.

Figura 52 Cromatograma do fracionamento por cromatografia de troca iônica (SCX) dos peptídeos marcados. A mistura 1:1 entre os peptídeos do extrato total de *N. caninum* marcados com dimetil-isótopos (*controle*, marcação leve; e *estimulado*, marcação intermediária) foi eluída em 50 frações e as mais intensas (9-24) foram submetidas à quantificação por LC-MS. Eixo X, tempo de eluição (minutos). Eixo Y, absorbância em 218 nm.

De um total de 2163 proteínas identificadas, 2011 foram quantificadas. Uma lista completa com as 2011 proteínas, organizadas em ordem decrescente de razão entre as amostras com marcação intermediária/leve (*estimulado/controle*), e contendo informações quanto ao número de peptídeos únicos identificados; número de PSM's (*peptide spectrum matches*); área sob o pico; e razão entre peptídeos com marcação intermediária sobre leve (*estimulado/controle*) está disponível no Apêndice 2.

Considerando-se o limite de duas vezes o valor da razão *estimulado/controle* (*fold-change threshold* = 2), 150 proteínas estavam aumentadas ou *up-regulated* ($\log_2$ da razão maior que +1) e 90 proteínas estavam diminuídas ou *down-regulated* ($\log_2$ da razão menor que -1) no taquizoíta *estimulado*. As tabelas 22 e 23 listam todas as proteínas aumentadas e diminuídas, respectivamente, incluindo informações sobre a razão *estimulado/controle*, localização ou grupo proteico, e identificação no ESA de *N. caninum*. O esperado é que as proteínas secretadas estivessem diminuídas no taquizoíta *estimulado*.

Tabela 22. Proteínas aumentadas no taquizoíta de *N. caninum* submetido ao estímulo secretório por etanol (estimulado). Organização por ordem decrescente de razão estimulado/controle (log₂), portanto com as proteínas mais aumentadas no topo da tabela. Razão maior que 1 configura proteína aumentada. Informações são mostradas sobre a localização ou grupo proteico (preditos ou conhecidos) e presença no extrato secretado de *N. caninum* (ESA, avaliado no presente estudo). h.p., hypothetical protein; rel., related; put, putative; f.transc., fator de transcrição.

ToxoDB ID	Razão (log2)	Descrição (ToxoDB)	Localização ou grupo proteico	Presença no ESA
NCLIV_040870	3.609	h.p., conserved	—	
NCLIV_020380	2.674	h.p., conserved	—	
NCLIV_032250	2.569	h.p., conserved	—	
NCLIV_062830	2.456	h.p., conserved	—	
NCLIV_067860	2.340	h.p., conserved	—	
NCLIV_037770	2.219	erythrocyte membrane protein PFEMP3, rel	—	sim
NCLIV_060430	2.106	Os08g0535200 protein, rel	—	
NCLIV_051150	2.098	h.p., conserved	—	
NCLIV_005640	2.026	h.p., conserved	—	
NCLIV_009220	1.956	h.p., conserved	—	
NCLIV_060180	1.954	cDNA FLJ58794, rel	—	
NCLIV_051890	1.942	h.p., conserved (AP2 transcription factor X-7)	f.transc. AP2	
NCLIV_032340	1.920	poly(A)-specific ribonuclease subunit PAN2, rel	—	
NCLIV_068900	1.918	h.p. (NOT2/NOT3/NOT5)	núcleo	
NCLIV_027450	1.897	h.p., conserved	—	
NCLIV_038470	1.895	h.p., conserved	—	
NCLIV_017110	1.883	h.p., conserved	—	
NCLIV_045520	1.883	h.p., conserved	—	
NCLIV_053220	1.864	PsmB (EC 3.4.25.1), rel	—	sim
NCLIV_065980	1.862	centrosomal protein, put	—	
NCLIV_022380	1.835	NLI interacting factor-like phosphatase, put	—	
NCLIV_070070	1.828	calcyclin binding protein, put	—	sim
NCLIV_026050	1.828	h.p., conserved	—	
NCLIV_017040	1.811	phosphofructokinase, rel	—	sim
NCLIV_067800	1.807	h.p., conserved	—	
NCLIV_037200	1.797	GJ21226, rel	—	sim
NCLIV_010540	1.794	h.p., conserved	—	
NCLIV_067970	1.750	double-strand break repair protein, put	núcleo	
NCLIV_031320	1.709	h.p., conserved	—	
NCLIV_021570	1.701	h.p., conserved	—	
NCLIV_042980	1.670	h.p. (CCR4-Not complex component)	—	sim
NCLIV_067460	1.667	h.p., conserved	—	
NCLIV_039480	1.644	h.p.	—	sim
NCLIV_047820	1.639	h.p., conserved	—	
NCLIV_008490	1.636	RNA recognition motif. family protein, rel	—	
NCLIV_035900	1.633	RNA binding motif-containing protein, put	—	
NCLIV_020970	1.600	h.p., conserved	—	
NCLIV_065290	1.597	h.p., conserved	—	sim
NCLIV_046960	1.588	h.p., conserved	—	
NCLIV_040010	1.583	alpha-1 type II collagen, put	—	
NCLIV_056570	1.576	collagen alpha-1(III) chain (Precursor), rel	—	
NCLIV_002840	1.572	h.p., conserved	—	
NCLIV_015740	1.568	h.p., (transcription initiation factor ia, put)	—	
NCLIV_032740	1.566	h.p., conserved	—	
NCLIV_045200	1.561	h.p., conserved	—	
NCLIV_001290	1.560	h.p., conserved (importin beta, put)	—	
NCLIV_000890	1.557	ER lumen protein retaining receptor 2, rel	—	
NCLIV_038250	1.556	h.p., conserved	—	
NCLIV_046940	1.546	PWWP domain-containing protein, put	—	
NCLIV_017830	1.537	5'-3' exonuclease, put	—	
NCLIV_047470	1.533	h.p., conserved	—	
NCLIV_047980	1.501	h.p., conserved	—	
NCLIV_044390	1.483	h.p., (cold-inducible RNA binding protein, put)	—	
NCLIV_007810	1.470	h.p., conserved (AP2III-2)	f.transc. AP2	
NCLIV_045280	1.466	h.p., conserved	—	
NCLIV_062970	1.460	helicase conserved C-terminal, put	—	
NCLIV_063260	1.459	h.p., conserved	—	
NCLIV_058540	1.453	RNA polymerase II subunit RPB11, put	—	

NCLIV_060560	1.450	h.p., conserved	—	
NCLIV_036010	1.435	h.p. (rbm25 protein, put)	—	sim
NCLIV_050040	1.413	h.p., conserved	—	
NCLIV_031700	1.403	DEHA2G03652p, rel	—	
NCLIV_001920	1.401	ATP-dependent helicase, put	—	
NCLIV_049750	1.393	replication factor C subunit, put	—	
NCLIV_069920	1.392	transcription factor S-II, put	núcleo	sim
NCLIV_055300	1.381	h.p. (WD-40 repeat-containing protein, put)	—	
NCLIV_010090	1.374	h.p., conserved	—	
NCLIV_019560	1.372	h.p., conserved	—	
NCLIV_029660	1.364	proteophosphoglycan 5, rel	—	
NCLIV_046200	1.359	h.p. (glutathione synthetase, put)	—	
NCLIV_014940	1.352	h.p., conserved	—	
NCLIV_036810	1.346	h.p.	núcleo	
NCLIV_036820	1.335	h.p. (RNA polymerase II phosphatase, put)	—	
NCLIV_022030	1.334	h.p., conserved	—	
NCLIV_021200	1.333	h.p., conserved	—	
NCLIV_031430	1.328	protein kinase, put	—	
NCLIV_012100	1.319	vacuolar protein sorting-associated protein, rel	—	
NCLIV_012950	1.317	ubiquitin family domain-containing protein, put	—	
NCLIV_028650	1.316	h.p., conserved	—	
NCLIV_029810	1.315	Tcc1a22.4, rel (NOT2 / NOT3 / NOT5)	núcleo	
NCLIV_025440	1.312	h.p.	—	
NCLIV_063920	1.296	h.p., conserved (AP2XII-4)	f.transc. AP2	
NCLIV_020950	1.294	CMGC kinase, Dyrk family, put	—	
NCLIV_026460	1.291	h.p. (RNA binding)	—	
NCLIV_007880	1.288	GM10776, rel (CMGC/SRPK)	—	
NCLIV_015600	1.286	h.p. (protein phosphatase, put)	—	
NCLIV_004370	1.286	h.p., conserved	—	
NCLIV_022460	1.281	zinc finger (C2H2 type), put	—	
NCLIV_011360	1.279	SET domain-containing protein, put	—	
NCLIV_052600	1.274	YEATS family protein, rel	núcleo	
NCLIV_029290	1.266	h.p. (ankyrin protein)	—	
NCLIV_028800	1.266	h.p., conserved	—	
NCLIV_018190	1.266	h.p., conserved	—	
NCLIV_045660	1.264	h.p., conserved	—	
NCLIV_029260	1.258	h.p.	apicoplast	
NCLIV_017730	1.258	guanylate kinase, rel	—	
NCLIV_001620	1.251	h.p., conserved	—	
NCLIV_020050	1.249	zinc finger (C3HC4 RING finger), rel	—	
NCLIV_066340	1.242	h.p., conserved (AP2XII-8)	f.transc. AP2	
NCLIV_056600	1.234	h.p., conserved	núcleo	
NCLIV_060900	1.233	h.p., conserved	—	
NCLIV_031150	1.228	h.p., conserved	—	
NCLIV_011140	1.225	GL18351, rel	—	
NCLIV_029460	1.212	h.p., conserved	—	
NCLIV_022000	1.208	para-aminobenzoate synthase, put	—	
NCLIV_002350	1.207	h.p.	—	
NCLIV_064180	1.203	h.p., conserved	—	
NCLIV_040850	1.199	nuclear NF-kB activating protein, put	núcleo	
NCLIV_021110	1.189	h.p., conserved	—	sim
NCLIV_031280	1.181	PHD-finger domain-containing protein, put	—	
NCLIV_030010	1.179	DEAD/DEAH box helicase, put	—	
NCLIV_003800	1.177	h.p., conserved	—	
NCLIV_052790	1.172	h.p.	—	
NCLIV_059260	1.171	h.p. (transcriptional co-activator ADA2-A)	—	
NCLIV_001240	1.159	h.p. (protein kinase CMGC/CDK - PK1)	—	
NCLIV_068020	1.149	OTU-like cysteine protease, put	—	
NCLIV_031260	1.147	autophagy protein, put	citoplasma	
NCLIV_048300	1.139	h.p. (PHD-finger domain-containing protein)	—	
NCLIV_010310	1.132	h.p., conserved	—	
NCLIV_061920	1.107	zinc finger (CCCH type) protein, put	—	
NCLIV_059690	1.106	h.p., conserved	—	
NCLIV_049460	1.103	adenosine monophosphate deaminase, put	—	sim
NCLIV_030970	1.100	h.p. (GMP synthase, put)	—	sim
NCLIV_059300	1.093	CD2 antigen cytoplasmic tail-binding prot. 2, put	citoplasma	sim
NCLIV_040510	1.086	h.p. (bromodomain domain-containing protein)	—	

NCLIV_044630	1.085	h.p., conserved	—	
NCLIV_014290	1.078	h.p., conserved	—	sim
NCLIV_018810	1.075	h.p., conserved	núcleo	
NCLIV_058980	1.071	1,4-alpha-glucan branching enzyme, put	—	
NCLIV_044270	1.061	h.p. (CAMK/CDPK - TgCDPK2A)	—	sim
NCLIV_009210	1.061	h.p.	—	sim
NCLIV_010110	1.059	h.p. (ubiquinol-cytochrome C reductase, put)	mitocondria	
NCLIV_048470	1.054	h.p., conserved (AP2X-4)	f.transc. AP2	sim
NCLIV_0260	1.053	h.p., conserved (ARM1)	—	
NCLIV_021260	1.044	h.p.	—	
NCLIV_065460	1.042	CMGC kinase, put	—	sim
NCLIV_004520	1.040	h.p., conserved	—	sim
NCLIV_032930	1.037	h.p., conserved (AP2VIII-2)	f.transc. AP2	
NCLIV_068550	1.037	h.p., conserved	—	
NCLIV_012500	1.036	h.p., conserved	—	
NCLIV_018560	1.036	ras-GTPase-activating binding protein, put	núcleo	sim
NCLIV_052320	1.031	RNA binding protein, put	—	
NCLIV_002470	1.027	h.p., conserved	núcleo	
NCLIV_048740	1.026	h.p., conserved	—	
NCLIV_044340	1.024	h.p., conserved	—	sim
NCLIV_042170	1.022	PHD finger protein BR140/LIN-49, rel	—	
NCLIV_003130	1.018	GTP-binding protein, rel	—	
NCLIV_004480	1.018	myb-like DNA-binding, put	—	
NCLIV_046340	1.010	h.p., conserved	—	
NCLIV_013640	1.009	h.p.	—	

Tabela 23 Proteínas diminuídas no taquizoíta de *N. caninum* submetido ao estímulo secretório por etanol (*estimulado*). Organização por ordem decrescente de razão estimulado/controle (log₂), portanto, com as proteínas mais diminuídas no final da tabela. Razão menor que -1 configura proteína diminuída. Informações são mostradas sobre a localização ou grupo proteico (preditos ou conhecidos) e presença no extrato secretado de *N. caninum* (ESA, avaliado no presente estudo). h.p., hypothetical protein; rel., related; put, putative.

ToxoDB ID	Razão (log2)	Descrição	Localização ou grupo proteico	Presença no ESA
NCLIV_064310	-1.005	GTP-binding nuclear protein Ran, rel	núcleo	
NCLIV_011730	-1.015	roptry kinase family protein ROP26, put	roptrias	sim
NCLIV_003960	-1.019	h.p. (U2 small nuclear ribonucleoprotein, put)	—	
NCLIV_005110	-1.019	ftsJ-like methyltransferase, put	—	
NCLIV_026200	-1.021	h.p.	—	
NCLIV_061600	-1.023	h.p. (phosphoserine aminotransferase, put)	—	sim
NCLIV_066410	-1.024	GG10762, rel (adenosine kinase, put)	—	
NCLIV_031860	-1.027	serine-threonine phosphatase 2C, put	—	sim
NCLIV_004300	-1.027	dynammin-like protein, put	—	sim
NCLIV_002620	-1.028	kinesin-ii motor protein, rel	—	
NCLIV_032320	-1.039	h.p. (protein phosphatase 2C, put)	—	
NCLIV_004380	-1.044	cathepsin L, rel	—	sim
NCLIV_032430	-1.045	h.p., conserved	—	sim
NCLIV_019970	-1.050	peptidyl-prolyl isomerase A, rel	ciclofilinas	sim
NCLIV_012860	-1.050	leucyl-tRNA synthetase 2 (EC 6.1.1.4), rel	citoplasma	
NCLIV_036410	-1.054	cyst matrix protein, put	—	
NCLIV_014760	-1.060	h.p., conserved	—	
NCLIV_068460	-1.060	h.p. (nucleoside-triphosphatase I - NTPasel)	grânulos densos	
NCLIV_011700	-1.072	h.p. (ROP15B)	roptrias	sim
NCLIV_002610	-1.075	isoleucine-tRNA synthetase, put	citoplasma	
NCLIV_015530	-1.075	h.p., conserved	—	
NCLIV_042650	-1.099	GG11844, rel	—	sim
NCLIV_030290	-1.101	transaldolase, put	citoplasma	sim
NCLIV_007730	-1.114	valyl-tRNA synthetase, mitochondrial, rel	citopl / mitoc	sim
NCLIV_068520	-1.114	microneme protein 5 (MIC5)	micronemas	
NCLIV_056520	-1.115	Rab1-prov protein, rel	—	
NCLIV_042740	-1.117	patched family protein, put	—	
NCLIV_000940	-1.134	glucose-6-phosphate dehydrogenase, put	—	sim
NCLIV_040040	-1.137	GrpE protein homolog, rel	mitocôndria	
NCLIV_000840	-1.141	thioredoxin h, rel	—	sim
NCLIV_025130	-1.156	translation initiation factor 1A (AelF-1A), rel	—	

NCLIV_002390	-1.165	nucleoside diphosphate kinase (EC 2.7.4.6), rel	—	sim
NCLIV_028680	-1.177	apical membrane antigen 1 (AMA1)	micronemas	sim
NCLIV_062290	-1.181	h.p., conserved	—	
NCLIV_050630	-1.191	H/ACA ribonucleoprotein, rel	núcleo	
NCLIV_063150	-1.200	serpin peptidase inhibitor, rel	grânulos densos	sim
NCLIV_044850	-1.210	h.p., conserved	—	
NCLIV_063970	-1.211	long chain acyl-CoA synthetase, put	—	sim
NCLIV_006780	-1.216	h.p., conserved (DG32)	grânulos densos	
NCLIV_030420	-1.253	Rcn2-prov protein, rel	—	sim
NCLIV_054830	-1.258	h.p. (GRA4)	grânulos densos	
NCLIV_018500	-1.261	fatty acyl-CoA synthetase 2, rel	—	sim
NCLIV_054150	-1.263	ubiquitin-conjugating enzyme, rel	—	
NCLIV_015200	-1.277	pyruvate kinase (EC 2.7.1.40), rel	—	sim
NCLIV_046540	-1.327	oligoendopeptidase F, put	—	sim
NCLIV_031550	-1.358	h.p. (predicted protein kinase)	roptrias	sim
NCLIV_068340	-1.367	h.p.	—	
NCLIV_061200	-1.376	actin, put (ALP1)	—	
NCLIV_026680	-1.381	protein kinase AKT, rel	—	
NCLIV_021050	-1.398	subtilisin-like protein (SUB1)	micronemas	sim
NCLIV_047660	-1.400	h.p. (adenosylhomocysteinase, put)	—	sim
NCLIV_002480	-1.405	CBR-DPY-23 protein, rel	—	
NCLIV_061170	-1.412	h.p. (heat shock protein, put)	—	sim
NCLIV_018420	-1.440	roptry protein ROP9 (p36)	roptrias	sim
NCLIV_015830	-1.445	h.p.	—	
NCLIV_006060	-1.451	h.p., conserved	—	
NCLIV_049400	-1.524	serine hydroxymethyltransferase, rel	—	sim
NCLIV_060860	-1.527	cytochrome c, rel	—	sim
NCLIV_001080	-1.528	h.p. (vacuolar type Na(+)/H(+) exchanger)	—	
NCLIV_028830	-1.542	glucose-6-phosphate isomerase, rel	—	sim
NCLIV_048240	-1.555	h.p. (aminopeptidase N, put)	—	sim
NCLIV_020990	-1.580	perforin-like protein 1 (PLP1)	micronemas	sim
NCLIV_051340	-1.620	toxofilin	roptrias	sim
NCLIV_022970	-1.637	microneme protein 2 (MIC2)	micronemas	sim
NCLIV_018530	-1.656	h.p., conserved	—	sim
NCLIV_051970	-1.668	MIC2 associated protein M2AP	micronemas	sim
NCLIV_049940	-1.709	translation initiation factor IF-2, rel	—	
NCLIV_042510	-1.712	h.p., conserved	—	sim
NCLIV_068230	-1.762	nucleoside-triphosphatase, put	—	
NCLIV_061760	-1.766	microneme protein 6 (MIC6)	micronemas	
NCLIV_045650	-1.789	dense granule protein GRA2	grânulos densos	sim
NCLIV_020720	-1.798	microneme protein 11 (MIC11)	micronemas	sim
NCLIV_049520	-1.803	h.p., conserved	—	sim
NCLIV_041870	-1.834	h.p. (cathepsin C1)	grânulos densos	
NCLIV_043270	-1.958	microneme protein 1 (MIC1)	micronemas	sim
NCLIV_059510	-2.012	RNase L inhibitor, rel	—	
NCLIV_020480	-2.121	h.p. (U1 small nuclear ribonucleoprotein, put)	núcleo	sim
NCLIV_059380	-2.149	SJCHGC06128 protein, rel	—	
NCLIV_065440	-2.150	h.p. (calmodulin)	—	sim
NCLIV_033690	-2.183	MIC2-like1	micronemas	sim
NCLIV_002940	-2.238	microneme protein 4 (MIC4)	micronemas	sim
NCLIV_036400	-2.310	dense granule protein GRA1	grânulos densos	sim
NCLIV_015380	-2.424	h.p., conserved	—	sim
NCLIV_057330	-2.514	h.p., conserved	—	
NCLIV_000860	-2.517	SPATR	micronemas	
NCLIV_057870	-2.572	h.p., conserved	núcleo	
NCLIV_042480	-2.693	h.p., conserved	—	sim
NCLIV_062770	-2.889	microneme protein 8 (MIC8)	micronemas	sim
NCLIV_066250	-3.053	microneme protein 10 (MIC10)	micronemas	sim
NCLIV_000740	-3.713	chitin-like protein 1 (CPL1)	micronemas	sim

As razões ($\log_2$) das proteínas quantificadas no extrato total (ET) de *N. caninum* (*estimulado/controle*) foram plotadas contra as suas intensidades para visualização da distribuição dessas proteínas no taquizoíta *estimulado*. Nesta distribuição (figura 53.A) as proteínas aumentadas são representadas em azul, enquanto as proteínas diminuídas são representadas em rosa. Em cinza estão as proteínas não diferencialmente expressas. Para as proteínas identificadas em comum no ESA e no ET de *N. caninum* ($n = 497$), foram plotadas suas razões contra suas intensidades, e foi verificada a distribuição destas proteínas no taquizoíta *estimulado*. As proteínas do ESA foram destacadas na cor roxa, e as demais proteínas do ET aparecem em verde (figura 54.B). Foi verificada também a presença de proteínas localizadas na superfície do parasita e nas organelas micronemas, roptrias e grânulos densos entre as proteínas quantificadas, e suas razões foram plotadas contra suas intensidades (figura 53.C).

Fig. 53 Distribuição de proteínas no extrato total (ET) de *N. caninum* após estímulo secretório. Eixo X, Razão ($\log_2$) entre os peptídeos do extrato total de *N. caninum* com marcação intermediária

sobre leve (*estimulado/controle*). Eixo Y, Intensidade, referente a área sob o pico. **A**, Distribuição de proteínas aumentadas e diminuídas no taquizoíta *estimulado* (razão *estimulado/controle*). Em azul, proteínas aumentadas; em rosa, proteínas diminuídas; em cinza, restante das proteínas do ET (não diferencialmente expressas). **B**, Localização de proteínas do ESA sobre a distribuição de proteínas quantificadas no ET. Em roxo, proteínas do ESA; em verde, demais proteínas do ET. **C**, Distribuição de proteínas localizadas na superfície do parasita e nas organelas secretórias identificadas no ET de taquizoíta *estimulado*. MIC, micronemas (amarelo); ROP, roptrias (verde); GRA, grânulos densos (azul); superfície (rosa).

Na tabela 24 estão listadas todas as proteínas de organelas secretórias e de superfície identificadas no ESA e/ou no ET de *N. caninum*. Os valores de abundância relativa no ESA (número de espectros dividido por peso molecular) e da razão *estimulado/controle* (marcação intermediária/marcação leve) das proteínas quantificadas no ET são mostrados para as proteínas identificadas por uma ou por outra abordagem, ou por ambas.

Tabela 24 Proteínas de organelas secretórias (micronemas, roptrias e grânulos densos) e de superfície identificadas no extrato secretado (ESA) e/ou quantificadas no extrato total (ET) de *N. caninum*. As proteínas estão organizadas em ordem decrescente de abundância relativa no ESA (número de espectros dividido por peso molecular). São mostradas as razões entre amostras *estimulado/controle* (marcação intermediária/leve) para as proteínas quantificadas no ET (razão > 2.0, proteínas aumentadas; razão < 0.5, proteínas diminuídas). A ausência de valor (representada por “—”) indica que a proteína não foi identificada por determinada abordagem.

ToxoDB ID	Descrição	Abundância	Razão na Quantificação
		Relativa no ESA	(Interm./Leve)
PROTEÍNAS MICRONÊMICAS			
NCLIV_021050	SUB1	33.70	0.3795
NCLIV_028680	AMA1	33.13	0.4424
NCLIV_043270	MIC1	23.00	0.2573
NCLIV_033690	MIC2-like1	5.60	0.2202
NCLIV_020720	MIC11	5.27	0.2876
NCLIV_051970	M2AP	4.48	0.3147
NCLIV_061760	MIC6	3.38	0.2941
NCLIV_010600	MIC3	3.34	0.5550
NCLIV_066250	MIC10	3.13	0.1205
NCLIV_002940	MIC4	2.89	0.2120
NCLIV_062770	MIC8	2.73	0.1350
NCLIV_022970	MIC2	1.52	0.3215
NCLIV_020990	PLP1	0.88	0.3345
NCLIV_000860	SPATR	0.84	0.1747
NCLIV_041880	MIC16	0.73	—
NCLIV_000740	CPL1	0.54	0.0763
NCLIV_044230	TLN4	0.33	—
NCLIV_038110	MIC17B	0.13	—
NCLIV_063730	MIC15	0.02	0.7029
NCLIV_068520	MIC5	—	0.4619
PROTEÍNAS DE ROPTRIAS			
NCLIV_011700	ROP15B	10.06	0.4757
NCLIV_031550	ROPK (ROP55)	6.46	0.3902
NCLIV_001970	ROP7	3.69	0.8992
NCLIV_011730	ROP26	3.37	0.4948
NCLIV_018420	ROP9	2.73	0.3687
NCLIV_012920	ROP40	2.09	0.5998
NCLIV_007770	(BPK1)	1.36	0.6094
NCLIV_051340	toxofilin	1.25	0.3253

NCLIV_041930	RAB11	0.96	0.7395
NCLIV_028170	ROP20	0.87	0.5091
NCLIV_035860	ROP2L10 (<i>ROP32</i>)	0.45	0.6253
NCLIV_053840	ROP1	0.36	0.6815
NCLIV_021100	ROP12	0.35	—
NCLIV_070010	RON8	0.31	0.9576
NCLIV_017440	ROP19	0.30	0.7750
NCLIV_030050	RON4	0.27	0.9502
NCLIV_011690	ROP15	0.23	0.7583
NCLIV_055850	ROP13	0.20	0.9779
NCLIV_044410	ROP35	0.18	0.8331
NCLIV_048590	RON3	0.17	0.8718
NCLIV_007800	RON4L1	0.13	0.9409
NCLIV_017420	ROP47	0.11	0.9154
NCLIV_064620	RON2	0.08	0.9762
NCLIV_055360	RON5	0.07	0.9004
NCLIV_069590	ROP2L8 (<i>ROP24</i>)	0.06	1.0238
NCLIV_057550	SUB2	0.03	1.1516
NCLIV_065640	ROP50	—	0.5594
NCLIV_009440	NHE2	—	0.8039
NCLIV_057950	ROP14B	—	0.8215
NCLIV_027930	ROP17	—	0.8613
NCLIV_068890	ROP51	—	0.8717
NCLIV_057960	ROP14	—	0.8885
NCLIV_054120	RON1	—	0.8912
NCLIV_060730	ROP5	—	0.9243
NCLIV_020340	RON3L1	—	0.9326
NCLIV_036700	TLN-1	—	0.9423
NCLIV_007450	ROP24	—	1.0616
NCLIV_068850	ROP53	—	1.0708
NCLIV_022270	ROP49	—	1.0785
NCLIV_006850	RON6	—	1.1677
NCLIV_027850	ROP6	—	1.2336
NCLIV_001460	ROP37	—	1.3302
PROTEÍNAS DE GRÁNULOS DENSOS			
NCLIV_066630	GRA9	2.54	0.7452
NCLIV_063150	serpin peptidase inhibitor	2.43	0.4353
NCLIV_068400	NTPase II	1.81	0.5015
NCLIV_045650	GRA2	1.59	0.2894
NCLIV_005560	DG32	0.93	0.6116
NCLIV_045870	GRA3	0.91	0.8770
NCLIV_052880	GRA6	0.57	0.7786
NCLIV_036400	GRA1	0.40	0.2017
NCLIV_003100	PI-2	0.29	—
NCLIV_041870	(<i>cathepsin C1 - CPC1</i>)	—	0.2805
NCLIV_054830	GRA4	—	0.4182
NCLIV_006780	(<i>DG32</i>)	—	0.4304
NCLIV_068460	NTPase I	—	0.4795
NCLIV_008990	GRA8	—	0.7259
NCLIV_016360	GRA14	—	0.9160
NCLIV_014150	GRA5	—	0.9477
NCLIV_021640	GRA7	—	0.9621
PROTEÍNAS DE SUPERFÍCIE			
NCLIV_033250	SRS2 (<i>SRS29C</i>)	14.26	0.886226
NCLIV_010720	SRS domain-containing protein	2.66	0.852343
NCLIV_033230	SAG1 (<i>SRS29B</i>)	1.42	0.81467
NCLIV_060660	SAG3 (<i>SRS57, P43</i>)	0.87	0.953586
NCLIV_068920	SRS domain-containing protein	0.76	0.961652
NCLIV_010730	SRS domain-containing protein	0.46	1.070203
NCLIV_015870	SRS23	0.16	1.032933
NCLIV_068872	SRS domain-containing protein	0.16	1.070327
NCLIV_052740	SRS domain-containing protein	0.14	—
NCLIV_034381	SRS domain-containing protein	0.12	0.874859
NCLIV_023620	SRS domain-containing protein	0.12	1.03756
NCLIV_060700	SRS domain-containing protein	0.05	0.896988
NCLIV_043420	SRS domain-containing protein	—	0.500977
NCLIV_034740	SAG3 protein, related	—	0.684117

NCLIV_067570	SRS domain-containing protein	—	0.733113
NCLIV_034731	SRS domain-containing protein	—	0.742986
NCLIV_010050	SRS domain-containing protein	—	0.766163
NCLIV_020120	hypothetical protein	—	0.905086

4.2.5 Rede de interação entre proteínas de *N. caninum*

As interações físicas entre as 2011 proteínas quantificadas foram preditas com o programa UNISPPPI (Universal Predictor of Protein-Protein Interactions)⁸⁸. Apenas os pares de proteínas com probabilidade de interação maior que 90% foram utilizados em nossas análises. Três redes de interação foram construídas: com pares de proteínas envolvendo proteínas diminuídas (figura 54.A); com pares de proteínas envolvendo proteínas aumentadas (figura 54.B); e uma união destas duas redes, portanto, envolvendo proteínas diminuídas e aumentadas (figura 54.B). Como as redes construídas envolvem um número muito grande de interações (6570 entre as 2011 proteínas quantificadas; 651 interações que envolvem proteínas aumentadas; e 862 interações que envolvem proteínas diminuídas), as listas com os pares interagentes não são disponibilizadas neste documento, mas serão publicadas com o artigo referente a este trabalho.

Fig. 54 Visualização no programa Cytoscape das redes de interação entre proteínas quantificadas no taquizoíta de *N. caninum* após estímulo secretório. **A**, rede de interação envolvendo proteínas diminuídas no taquizoíta *estimulado*; **B**, rede de interação envolvendo proteínas aumentadas no taquizoíta *estimulado*; **C**, união das redes A e B, portanto, envolvendo proteínas diminuídas e aumentadas. **Nodos em rosa**, proteínas diminuídas; **Nodos em azul**, proteínas aumentadas. **Nodos em amarelo**, proteínas não diferencialmente expressas.

4.2.6 Análise proteogenômica: Peptídeos identificados em regiões não-anotadas

Os arquivos de MS gerados por CID das amostras *controle* e *estimulado*, assim como os arquivos gerados por CID e HCD da análise de ESA de *N. caninum*, foram usados na busca por peptídeos presentes em genes não-anotados, empregando a ferramenta *Pepsplince*⁹¹. A figura 55 é um exemplo de como foi feito o reconhecimento destes peptídeos. Uma região de aproximadamente 4000 pb no cromossomo X compreende 6 peptídeos, nomeados 1.X, 2.X, 3.X, 4.X, 5.X e 6.X. Na figura 55 A os peptídeos são representados por retângulos azuis (com setas indicando o sentido reverso), e suas regiões correspondentes no cromossomo X são delimitadas em vermelho. Uma região codificadora anotada (NCLIV_045260) está próxima, porém não sobreposta aos peptídeos. Na figura 55 B, o peptídeo 5.X foi ampliado, confirmando o sentido e a fase de leitura da sequência.

Fig 55. Visualização no programa IGV (Integrative Genomics Viewer) de peptídeos identificados em regiões do genoma de *N. caninum* não anotadas como codificadoras de genes. A, Região de 4673 pb no cromossomo X, onde 6 peptídeos foram identificados. Os peptídeos 1.X, 2.X, 3.X, 4.X, 5.X e 6.X. estão representados por retângulos azuis (com setas indicando o sentido reverso), e suas regiões correspondentes no cromossomo X são mostradas em vermelho. NCLIV_045260 mostra como uma região codificadora anotada é reconhecida nesta visualização. **B,** Ampliação da região do peptídeo 5.X e definição de sua fase de leitura.

Um total de 95 peptídeos foi identificado, distribuídos em todos os cromossomos de *N. caninum* (banco de dados do ToxoDB versão 7.2). Uma lista com todas estas identificações estará disponível em nossa publicação. Seleccionamos para apresentar aqui duas regiões de interesse que podem representar proteínas em potencial. Ambas estão descritas na tabela 25. A região NCLIV_chrlb:1578985-1579756 compreende 5 peptídeos que compartilham o mesmo sentido e mesma fase de leitura, no cromossomo Ib; e a região NCLIV_chrX:398492-402389, no cromossomo X, contém 6 peptídeos, também dispostos no mesmo sentido e mesma fase de leitura. Quando uma busca foi feita com estes peptídeos no banco do NCBI, com a ferramenta BLAST, não foram encontradas proteínas com similaridade significativa.

Tabela 25 Peptídeos identificados em regiões não anotadas como codificadoras. No cromossomo Ib, a região NCLIV_chrlb:1578985-1579756 compreende 5 peptídeos; no cromossomo X, a região NCLIV_chrX:398492-402389 compreende 6 peptídeos. Em ambas as regiões os peptídeos compartilham o mesmo sentido e mesma fase de leitura na sequência genômica. A sequência de aminoácidos de cada peptídeo, sua região codificadora exata no genoma de *N. caninum*, e o número de peptídeos identificados por MS também são mostrados na tabela.

Cromossomo	Descrição do peptídeo	Sequência de aminoácidos	Região no genoma de <i>N. caninum</i>	Sentido	Fase de leitura	Número de peptídeos
Ib	5.Ib	APTAPPFLVDLER	psu NCLIV_chrlb:1578985-1579024	forward	2	5
	6.Ib	EEYTNVLFEAR	psu NCLIV_chrlb:1579024-1579057	forward	2	6
	7.Ib	EMVALWATR	psu NCLIV_chrlb:1579072-1579099	forward	2	1
	8.Ib	NLQEGLLPVSLR	psu NCLIV_chrlb:1579663-1579699	forward	2	4
	9.Ib	FGWPGNAVAENIK	psu NCLIV_chrlb:1579717-1579756	forward	2	4
X	1.X	QEEEAQEEEEANQEEAAK	psu NCLIV_chrX:398582-398636	reverso	3	4
	2.X	QEEEAQEEEEAKQEEAAK	psu NCLIV_chrX:399167-399221	reverso	3	3
	3.X	QEEEAQEEEEANQEEAAK	psu NCLIV_chrX:399221-399275	reverso	3	7
	4.X	QEEEAQEEEEAK	psu NCLIV_chrX:399365-399401	reverso	3	4
	5.X	TQLDNLTAQR	psu NCLIV_chrX:400541-400571	reverso	3	2
	6.X	LQYVDDLYK	psu NCLIV_chrX:402098-402125	reverso	3	1

5. Discussão

5.1 Caracterização das proteínas com domínio Apple

As proteínas de *N. caninum* MIC17A, MIC17B e MIC17C descritas neste trabalho são compostas por um peptídeo sinal para secreção, um domínio característico de direcionamento para micronemas, três domínios Apple (PAN1) e um domínio Apple-like (PAN3). Devido à alta identidade entre suas sequências nucleotídicas e à disposição destes genes no cromossomo, podemos dizer que seus genes são parálogos. MIC17C tem sua sequência codificadora e a região 3'UTR mais semelhantes às de MIC17A (98% de identidade), enquanto sua região 5'UTR é quase idêntica à de MIC17B. Estas semelhanças entre as regiões codificadoras e também entre as regiões não traduzidas podem indicar que esses genes foram gerados por processos de duplicação gênica seguido de divergência⁹⁵. A amplificação de todos os três genes sobre um mesmo cDNA, gerado a partir da polimerização sobre RNA com um *primer* reverso específico para a região 3'UTR de MIC17C, indica que estes genes são transcritos na forma de RNA policistrônico, ou seja, um pré-mRNA é transcrito sobre este *cluster* de genes e, numa segunda etapa, são formados os mRNA maduros para a tradução de cada proteína. Em Apicomplexa, a transcrição por RNA policistrônico foi reportada apenas em *Plasmodium falciparum* para proteínas codificadas a partir de DNA mitocondrial⁹⁶ ou de apicoplasto⁹⁷. Melhor estudado em *Trypanosoma*, o processo para a formação dos mRNAs maduros a partir do RNA policistrônico é chamado *trans-splicing*, no qual uma pequena parte (*mini-exon*) da sequência *spliced leader* (SL) RNA é adicionada ao pré-mRNA, antes da região codificadora de cada proteína, resultando na liberação do mRNA maduro, formado ao final pelo *mini-exon* ligado à região codificadora de determinada proteína⁹⁸. Evolutivamente, a existência deste sistema de transcrição por *trans-splicing* associado à presença de mRNA policistrônicos seria convergente entre kinetoplastídeos (*Trypanosoma*) e dinoflagelados (um subgrupo dos Alveolatas, como os apicomplexas)⁹⁹. E experimentos com o dinoflagelado *Amphidinium carterae* mostraram que genes organizados como repetições em tandem são transcritos na forma de mRNA policistrônico, e em seguida processados por *trans-splicing*¹⁰⁰. Em uma região bem próxima aos genes MIC17 (13,64 kpb), no mesmo cromossomo VIII, existem 3 outros genes potencialmente policistrônicos (NCLIV_038060, NCLIV_038070 e NCLIV_038080). Estes genes possuem alta

identidade entre si e codificam proteínas hipotéticas, com uma leve identidade com o domínio DEK_C (INTERPRO IPR014876), que possui capacidade de interação com DNA¹⁰¹. Foi demonstrado que a proteína contendo este domínio DEK_C, além de outras importantes funções como ser um oncogene, possui uma ação sobre a maquinaria de *splicing*¹⁰². Seria interessante investigar genes como estes, potencialmente envolvidos com o mecanismo de *trans-splicing*, que é muito pouco estudado no Filo Apicomplexa.

Apesar da nomenclatura MIC17A, MIC17B e MIC17C ter sido adotada para proteínas homólogas de *N. caninum* (Nc) e de *T. gondii* (Tg) com base na disposição dos genes na fita de DNA, não existe uma correlação de identidade para as proteínas MIC17A ou MIC17B ou MIC17C entre as duas espécies (tabela 26). Em *T. gondii*, ainda existe uma quarta proteína homóloga a esses genes (TGME49_200270) que não recebeu nomenclatura, provavelmente por ser transcrita no sentido oposto aos outros três genes.

Tabela 26 Relação de identidade e similaridade entre as proteínas Apple de *Neospora caninum* e *Toxoplasma gondii*, baseada nas sequências de aminoácidos. TGME49, código de identificação do ToxoDB para proteínas de *T. gondii* linhagem ME49; NCLIV, código de identificação do ToxoDB para proteínas de *N. caninum* linhagem Liverpool. Em cada célula são mostradas as porcentagens de identidade e similaridade entre duas proteínas (identidade/ similaridade).

<i>N. caninum</i> / <i>T. gondii</i>	TGME49_200270	TGME49_200250 (TgMIC17A)	TGME49_200240 (TgMIC17B)	TGME49_200230 (TgMIC17C)
NCLIV_038120 (NcMIC17A)	58,4/ 83,5%	58,1/ 85,5%	58,1/ 85,7%	58,9/ 85,2%
NCLIV_038110 (NcMIC17B)	57,0/ 86,3%	56,7/ 88,0%	54,4/ 86,6%	55,3/ 86,3%
NCLIV_038100 (NcMIC17C)	59,0/ 83,8%	58,4/ 85,5%	58,7/ 85,7%	59,5/ 85,2%

Também não foi observada para as proteínas de *T. gondii* a mesma relação de alta identidade entre MIC17A e MIC17C existente entre as proteínas de *N. caninum*. Existe uma alta porcentagem de identidade/similaridade entre as quatro proteínas de *T. gondii*, sendo que TGME49_200270 é a que mais diverge das outras três (79 a 83% de identidade). TgMIC17B e TgMIC17C apresentam o maior grau de identidade/similaridade (94,6%/98,6%), que também é alto entre TgMIC17A e TgMIC17B (90,0%/98,6%), enquanto TgMIC17A e TgMIC17C apresentam menor porcentagem de identidade/similaridade entre as três proteínas (88,3%/97,4%).

Estas relações são reforçadas ao se construir uma árvore filogenética não enraizada com todas estas proteínas (figura 56). Estas diferenças nas relações entre as três proteínas Apple de *N. caninum* e de *T. gondii* podem indicar que seus genes sofreram divergência após o processo de especiação, o que pode refletir no desenvolvimento de diferentes funções.

Fig. 56 Reconstrução filogenética das sequências das proteínas Apple MIC17 de *N. caninum* e *T. gondii*. Árvore não enraizada construída com o programa Phylogeny.fr. (<http://www.phylogeny.fr/>) utilizando o modo *one click*. Nc, *Neospora caninum*. Tg, *Toxoplasma gondii*. 1 e 0,96 são os valores de bootstrap.

Os domínios Apple encontrados nas proteínas MIC17 de *N. caninum* seguem o mesmo padrão estrutural dos domínios Apple do fator de coagulação XI e das proteínas da maioria dos organismos Apicomplexas. Ao contrário da predição feita por Sohn e colegas, baseada na interpretação equivocada sobre um alinhamento sequencial (figura 57), em que o segundo domínio Apple (A2) das proteínas MIC17 de *N. caninum* e *T. gondii* possui apenas 5 cisteínas⁴⁷, em nossas predições por modelagem molecular todos os domínios apresentaram seis cisteínas, envolvidas na formação de três pontes dissulfeto intra-domínio.

Fig. 57 Alinhamento sequencial das proteínas MIC17A, MIC17B e MIC17C de *N. caninum* (Nc) e *T. gondii* (Tg). No estudo de Sohn e colegas⁴⁷, os domínios Apple foram identificados como mostra a delimitação por linhas vermelhas, com as cisteínas conservadas destacadas em amarelo. Ao contrário do proposto por eles, em nossas predições a cisteína marcada por uma seta verde corresponde à última cisteína do domínio A2, enquanto a cisteína indicada pela seta vermelha, não reconhecida por eles, corresponde à primeira cisteína do domínio A3. O asterisco verde mostra a posição da cisteína extra que existe apenas no domínio A2 de NcMIC17A e NcMIC17C. (Figura adaptada de Sohn *et al*, 2011⁴⁷).

Apesar da baixa identidade sequencial entre os domínios Apple das NcMIC17 e das proteínas cujas estruturas foram previamente elucidadas, foi verificado por alinhamento que as diferenças residem nas regiões flexíveis desses domínios, possibilitando a predição por modelagem molecular. As estruturas dos domínios foram preditas com sucesso e, após refinamento, apresentaram boa estereoquímica. Todos os domínios seguiram o mesmo padrão de folhas beta ligadas à alfa hélice por duas pontes dissulfeto, entre as cisteínas C2-C5 e C3-C4, e as extremidades N e C terminal dos domínios são ligadas por uma terceira ponte dissulfeto. A diferença mais marcante encontrada foi a presença de uma cisteína extra apenas no domínio A2 das proteínas MIC17A e MIC17C (figura 25). Esta cisteína localiza-se na parte externa do domínio Apple, sugerindo a possibilidade de um papel estrutural e/ou funcional para as proteínas MIC17A e MIC17C, dado o caráter altamente reativo do grupo tiol. O fato de esta cisteína livre não existir em MIC17B pode ser um indício de

que estas proteínas apresentem funções diferentes para o parasita. Futuramente, estas estruturas aqui preditas podem servir também como modelos para o desenho de fármacos contra a neosporose.

Nas análises de superfície eletrostática (figura 26), notou-se uma distribuição de cargas diferente, principalmente dos domínios de MIC17B em relação aos domínios de MIC17A e MIC17C. Também ficou evidente que na região próxima à cisteína livre do domínio A2 de MIC17A e MIC17C há uma concentração de cargas negativas, enquanto na região correspondente em MIC17B a concentração é de cargas positivas. Estas regiões poderiam ter um papel importante na estabilização de uma interação com outra molécula, como ocorre na proteína MIC4 de *T. gondii* (figura 58). No domínio Apple A5 de TgMIC4, que interage com galactose, existem regiões com carga positiva (azul) adjacentes ao bolso de ligação à galactose, que poderiam estabilizar esta ligação, por exemplo, por interação com as cargas negativas do ácido siálico, presente em alguns oligossacarídeos juntamente com a galactose³⁵.

Fig. 58 Superfície eletrostática do domínio Apple A5 da proteína MIC4 de *Toxoplasma gondii* ligado ao dissacarídeo lacto-N-biose. Regiões com carga positiva e negativa são mostradas respectivamente em azul e vermelho. Existe uma concentração de cargas positivas nas regiões adjacentes ao dissacarídeo, que possivelmente atraem a unidade negativamente carregada do ácido siálico presente em alguns oligossacarídeos, estabilizando a interação com galactose. (Figura adaptada de Marchant *et al* 2012³⁵).

A compreensão da forma como são organizados os quatro domínios nas proteínas MIC17 poderia ajudar na investigação de suas funções, porém, por modelagem molecular verificamos apenas que ela não segue a organização da proteína do fator de coagulação XI. O ideal seria a elucidação das estruturas completas por cristalografia de raio-X, já que não dispomos de muitas opções de estruturas previamente elucidadas para a predição por modelagem. A pluralidade de funções adesivas para diferentes domínios Apple^{24,35} e a variabilidade no número de aminoácidos nos intervalos entre os domínios (Tabela 14) sugerem que existam diferentes formas de organização para as proteínas compostas por domínios Apple.

As proteínas nativas MIC17B e MIC17C detectadas por *Western blot* 1D como bandas de 35 kDa nos extratos total e secretado de *N. caninum* (NcET e NcESA) são coerentes com o seu peso predito após clivagem do peptídeo sinal. Já a banda de aproximadamente 40 kDa de MIC17B, apenas em NcET, pode corresponder à proteína não clivada (38.2 kDa) e possivelmente com modificações pós-traducionais, já que existem sítios preditos, por exemplo, para fosforilação (Ser: 3 Thr: 2 Tyr: 5) (www.cbs.dtu.dk/services/NetPhos/). As bandas com menor peso, presentes com maior intensidade no extrato secretado (ESA), indicam possível clivagem dessas proteínas após sua secreção. A diferença nos pesos dessas bandas menores entre MIC17B e MIC17C sugere que elas sejam processadas de forma diferente. Também existe a possibilidade de haver reconhecimento cruzado de MIC17A pelo anticorpo anti-MIC17C. Por *Western blot* 2D foram verificados *spots* da mesma proteína com diferentes pontos isoelétricos (de 8,1 a 8,4). A análise por espectrometria de massas dessas bandas e *spots* poderia contribuir para a elucidação de regiões de clivagens e possíveis modificações pós-traducionais, porém, estas identificações não foram alcançadas, provavelmente devido à baixa concentração dessas proteínas nos extratos. Por RP-nanoLC-MS/MS, mais sensível, foi comprovada a presença de MIC17B em NcESA.

Em um experimento com as proteínas Apple de *T. gondii* foi identificado um domínio de 24 aminoácidos, situado entre o peptídeo sinal e a cisteína C1 do primeiro domínio Apple, descrito como essencial para o direcionamento organelar⁶⁰. Um alinhamento foi feito com as três proteínas MIC17 de *N. caninum*, seus quatro homólogos em *T. gondii*, e uma proteína Apple localizada nos grânulos densos (figura 59). É notável que na região entre o peptídeo sinal e o domínio A1 (delimitada em vermelho), as proteínas NcMIC17 são mais conservadas em relação às proteínas que são direcionadas para as micronemas (62% a 83% de similaridade entre as sequências) e bastante diferentes da proteína direcionada para os grânulos densos (20% a 25% de similaridade).

Fig. 59 Alinhamento de proteínas Apple de *N. caninum* e *T. gondii* (primeiros 81 aminoácidos).

A proteína de *T. gondii* TGME49_0332400 é a única localizada nos grânulos densos, as outras quatro foram localizadas nas micronemas. M, região entre o peptídeo sinal e a cisteína C1 do primeiro domínio Apple, contém a informação de direcionamento organelar⁶⁰. PS, sequência de peptídeo sinal. C1, primeira cisteína do domínio Apple A1. NCLIV_038120 = NcMIC17A; NCLIV_038110 = NcMIC17B; NCLIV_038100 = NcMIC17C; TGME49_200250 = TgMIC17A; TGME49_200240 = TgMIC17B; TGME49_200230 = TgMIC17C. (Figura adaptada de Chen *et al*, 2008⁶⁰)

A informação contida nesta sequência, de que as proteínas NcMIC17 também seriam direcionadas para as micronemas é consistente com a localização de NcMIC17B encontrada por Sohn e colegas⁴⁷ e confirmada no presente trabalho. Porém, para NcMIC17C, que não foi identificada pelo outro grupo, a localização segue um padrão diferente, distribuída por vários pontos do taquizoíta, principalmente ao redor do núcleo. Quando localizada na região apical, mostra um padrão diferente de NcMIC17B. Experimentos de colocalização de MIC17C com proteínas de outras organelas seriam interessantes para confirmar sua localização no parasita, e a sua nomenclatura também deveria ser reavaliada.

O ensaio de invasão com taquizoítas de *N. caninum* transfectados de forma transiente para o gene da β -galactosidase (NcLacZ transiente) não foi bem sucedido devido à alta lise de taquizoítas causada pelos soros e à uma grande variação, mesmo com os soros inativados. O ensaio com NcLacZ estável, descartando o uso dos soros, foi desenvolvido para evitar possíveis flutuações e alcançar resultados mais consistentes. Após a constatação de que o BSA usado na estabilização dos anticorpos purificados promovia a invasão dos taquizoítas, um controle positivo adicionado de BSA foi inserido no ensaio, resultando numa inibição de 24,2% na invasão por bloqueio com o anticorpo anti-MIC17B, e de 15,5% com o anticorpo anti-MIC17C. A única proteína composta exclusivamente por domínios Apple para a qual foi realizado ensaio de invasão foi a proteína P104 de *T. gondii* (TGME49_209920), que contém 10 domínios Apple e apresenta capacidade de ligação à condroitina sulfato, um tipo de glicosaminoglicana¹⁰³. Ao invés de anticorpos, o ensaio de invasão para P104 foi realizado com adição de P104 recombinante (0,6 μ M) sobre cultura de células Vero, que em seguida foi incubada com taquizoítas de *T. gondii*,

alcançando uma inibição de 36.5%. Mesmo para as proteínas MIC4 de *T. gondii* e MIC5 de *Eimeria tenella*, compostas por domínios Apple e conhecidas por formar complexos com outras proteínas micronêmicas^{32,34}, não há resultados de ensaio de invasão. Apenas para TgMIC1, que participa de um complexo com TgMIC4 e TgMIC6, foi realizado ensaio de invasão com um parasita nocaute, mostrando que a ausência do gene de TgMIC1 diminui em 40% a eficiência da invasão³⁵. Ao observarmos os resultados de inibição da invasão para MIC17B e MIC17C devemos considerar a possibilidade de um mecanismo de compensação entre elas, já que realizamos o tratamento com um ou outro anticorpo separadamente, e também no caso de o anticorpo contra MIC17C não inibir ou inibir pouco a proteína MIC17A.

Os ensaios de *binding* visaram verificar a adesão às células hospedeiras, sem isolar um possível receptor. Nos ensaios por *binding-Western blot* e *binding-FACS* a detecção pelos anticorpos anti-MIC17B e anti-MIC17C foi maior nos controles negativos sem ESA, ou seja, usando apenas células Vero. Estes resultados, associados aos de *Western-blot*, indicam que os anticorpos anti-MIC17B e anti-MIC17C possuem afinidade por proteínas de células Vero, já que é detectado um sinal quando o ensaio é feito mesmo na ausência do extrato secretado (ESA). Quando o ESA é adicionado, os receptores de células Vero são ocupados por outras proteínas de *N. caninum*, e o sinal detectado pelos anticorpos anti-MIC17B e anti-MIC17C diminui em relação ao controle. Este problema ficou bastante evidente no ensaio por *binding-FACS* que, apesar do reconhecimento cruzado, foi o ensaio que funcionou de maneira mais adequada, com base no controle com soro anti-extrato total de *N. caninum* (NcET). Alguns spots excisados dos géis de NcET e NcESA, representando proteínas de células Vero que reagiram com os anticorpos purificados, foram identificados por MS como actina, alfa-enolase, ATP sintase e HSP60 de células Vero. Porém, alinhamentos destas sequências com as de MIC17B e MIC17C apresentaram apenas de 7% a 12% de identidade, e não evidenciaram regiões conservadas que pudessem corresponder aos epítomos contra os quais foram gerados os anticorpos.

Para evitar interferência dos anticorpos nos ensaios de adesão, e também para gerar resultados mais irrefutáveis sobre a influência das proteínas MIC17 na invasão, o ideal seria o desenvolvimento de ensaios com o taquizoíta NcLacZ nocauteado para os genes das proteínas MIC17, porém este taquizoíta não foi obtido. Três possíveis causas podem ser responsáveis para este insucesso; uma

delas seria a substituição do promotor da proteína SAG1 pelo promotor de MIC17A. SAG1 é uma proteína expressa em abundância, enquanto MIC17A apresenta níveis baixos de expressão, segundo o ToxoDB, o que pode ter prejudicado a expressão do gene CAT e a resposta esperada à pressão seletiva. Uma segunda causa seria a possibilidade de esta ser uma deleção letal ao taquizoíta, devido ao nocaute simultâneo dos três genes. Uma terceira, e mais provável causa, estaria relacionada ao número de nucleotídeos do plasmídeo (663 pb) que substituiriam uma região de aproximadamente 9000 pb. A deleção específica de cada uma das três proteínas, portanto de uma região menor a ser substituída, não foi testada pois é dificultada devido à alta conservação tanto das regiões traduzidas quanto não traduzidas.

Foram encontradas neste trabalho evidências de que as proteínas MIC17A, MIC17B e MIC17C estão relacionadas à invasão celular, porém, suas funções continuam desconhecidas. Mesmo para a homóloga de MIC17B em *T. gondii*, localizada nas micronemas, não há indícios sobre sua função⁴⁷. As diferenças encontradas entre MIC17B e MIC17C, tanto ao peso das proteínas identificadas por *Western blot*, que sugere formas alternativas de processamento, quanto à localização no taquizoíta e às particularidades no padrão estrutural, evidenciam que estas duas proteínas podem exercer funções diferentes no parasita. Além da possibilidade de divergência entre estes genes após um suposto processo de duplicação, a transcrição por RNA policistrônico não impediria que estas proteínas tivessem funções distintas, uma vez que, em *Trypanosoma*, que também possui este sistema remanescente do *operon* de procariotos, não foi observada ligação funcional entre os genes de uma mesma unidade de transcrição⁹⁸.

5.2 Secretoma de *N. caninum*

A identificação de proteínas no extrato secretado (ESA) de *N. caninum* foi realizada pelas técnicas de gel bidimensional analisado por MALDI (2D-MALDI) e cromatografia líquida acoplada a espectrômetro de massa (LC-MS). Comparando-se os resultados, por ser uma técnica high throughput, a LC-MS apresentou um número muito maior de identificações (615 contra 22). Apesar de por LC-MS não se ter acesso ao pI e ao peso molecular das proteínas, existe a facilidade para o cálculo de abundância relativa das proteínas, por divisão do número de espectros obtidos pelo peso molecular da proteína identificada.

No estudo do secretoma por LC-MS foi utilizado o estímulo secretório por aumento de Ca^{2+} intracelular, deflagrado *in vitro* por etanol. A implementação de duas abordagens diferentes, mas conectadas, empregando técnicas que asseguram alta resolução e sensibilidade, permitiu a identificação de um grande número de proteínas secretadas e a investigação de componentes envolvidos no processo de secreção. Este estudo permitiu, portanto, a pesquisa de novos alvos relacionados ao processo de invasão.

Das 615 proteínas identificadas pela Abordagem Via Extrato Secretado, um número expressivo (n=230; 37.4%) correspondeu a proteínas secretadas ou preditas para serem secretadas, e muitas foram pela primeira vez identificadas. A alta porcentagem de proteínas preditas para secreção por via não-clássica corrobora estudos feitos com *Trypanosoma cruzi* e *Leishmania*, em que rotas alternativas de secreção foram associadas a vesículas^{104,105}. A presença de proteínas cuja secreção não é esperada, como proteínas de apicoplasto, mitocôndria ou núcleo, pode ser devida à liberação destas proteínas de taquizoítas mortos, que representaram aproximadamente 10% dos taquizoítas purificados; ou é possível também que estas proteínas possuam diferentes funções/localizações em diferentes momentos do parasita, e sejam de fato secretadas.

Empregando a Abordagem Quantitativa, a intenção foi contornar a contaminação por proteínas oriundas de taquizoítas mortos e checar se as proteínas secretadas estariam diminuídas no taquizoíta após o estímulo secretório. De fato, as proteínas micronêmicas, cuja secreção em resposta ao aumento de Ca^{2+} foi experimentalmente confirmada⁵⁰, representaram uma grande porcentagem das proteínas diminuídas (16,7%), e a soma de proteínas micronêmicas, de roptrias e de grânulos densos correspondeu a 30,1% do grupo de proteínas diminuídas. Uma

sobreposição entre as proteínas identificadas no ESA com as proteínas quantificadas no ET (figura 53) demonstrou que há uma concentração maior de proteínas do ESA no lado esquerdo do gráfico, portanto, o lado das proteínas diminuídas. Enquanto a maioria das proteínas micronêmicas apresentou razão ($\log_2$) menor que -1,0 (figura 53 C), a maior parte das proteínas de roptrias e de grânulos densos apresentou razão maior que -1,0, porém ainda com valores negativos.

Em uma comparação entre as duas abordagens é notável que proteínas abundantes identificadas no ESA, como MIC3, GRA9 e ROP7, não figuraram entre o grupo de proteínas diminuídas (figura 60 e tabela 24). Isto indica que o limite utilizado para proteínas diferencialmente expressas (*fold-change threshold* = 2.0) pode ter sido muito estrito para apontar proteínas secretadas, uma vez que o mecanismo de liberação de proteínas destas três organelas ainda não é completamente compreendido. Entretanto, a comparação entre as proteínas identificadas no ESA e quantificadas no ET mostrou-se útil para uma busca mais precisa por proteínas mais suscetíveis à secreção.

Vinte proteínas micronêmicas foram identificadas (figura 60); entre elas, três somente no ESA (MIC16, TLN4 e MIC17B) e uma apenas no ET (MIC5). As três proteínas exclusivas do ESA, apesar de não muito abundantes, podem ter seu papel no processo de invasão. MIC16 é uma nova proteína transmembrana que contém domínios trombospondina e é suscetível à clivagem rombóide¹⁰⁶. MIC17B, que

contém domínios adesivos Apple e localização micronêmica, foi detectada por anticorpo monoclonal em fração organelar de *N. caninum*⁴⁷ e seu bloqueio, aqui descrito, foi capaz de inibir 24% da invasão pelo parasita. Toxolisina 4 (TLN4) é uma metaloproteinase putativa, processada pelo sistema secretório em múltiplos fragmentos que permanecem associados em um grande complexo molecular¹⁰⁷. A proteína MIC5 de *T. gondii* ocupa a superfície do parasita durante a invasão, regulando a atividade adesiva de outras proteínas secretadas¹⁰⁸.

Muito abundantes e identificadas por ambas as abordagens, SUB1¹⁰⁹, AMA1¹¹⁰ e MIC1¹¹¹ são proteínas micronêmicas largamente estudadas em parasitas apicomplexas. Outras proteínas micronêmicas de *N. caninum*, como MIC2-like1¹¹², MIC3¹¹³, MIC10¹¹⁴, MIC4¹¹⁵ e MIC2¹¹⁶ também apresentaram abundância elevada no ESA. M2AP, MIC6, e MIC8 são proteínas já bem caracterizadas em outros organismos Apicomplexas^{34,117,118}, em contraste com *N. caninum*, em que foram identificadas apenas em estudos genômicos ou de perfil de expressão gênica^{86,119}. Uma das proteínas mais abundantes no ESA de *N. caninum*, MIC11 foi capaz de estimular células T CD4⁺naive em bovinos⁴⁸. Algumas proteínas menos abundantes, como PLP1, SPATR, CPL1 e MIC15, que pela primeira vez foram identificadas em *N. caninum*, desempenham importante papel em outros organismos Apicomplexas. Em *T. gondii*, PLP1 (proteína relacionada à perforina) auxilia o egresso do parasita de seu vacúolo parasitóforo¹²⁰. SPATR¹²¹, CLP1 e MIC15 (Carruthers 2010 – observações no ToxoDB⁶⁶) são proteínas que possuem domínios relacionados ao processo de invasão celular.

As vias de sinalização para a secreção de roptrias e grânulos densos ainda não estão elucidadas. É presumido que, seguindo-se à adesão aos receptores das células hospedeiras por proteínas micronêmicas, seus domínios citoplasmáticos devem ter um papel na transdução de sinal que resulta na liberação do conteúdo das roptrias⁵³. Em *T. gondii*, foram identificadas várias proteínas de roptrias e grânulos densos em frações de esporocistos/esporozoitos, em contraste com dois estudos proteômicos usando taquizoítas, em que essas proteínas foram escassamente identificadas^{50,55}, provavelmente devido às ferramentas serem menos sensíveis à época do estudo.

Em nosso trabalho, proteínas de roptrias foram numerosamente detectadas (n=42), a maioria presente tanto no ESA quanto no ET (n=25), uma detectada somente no ESA, e outras 16 apenas no ET. Muitas destas proteínas, inclusive

proteínas abundantes como ROP15B, ROP26 e toxofilina, foram detectadas previamente somente em estudos abordando o genoma ou o transcriptoma de *N. caninum*⁸⁶. Os genes de BPK1 e ROP15 tiveram seus níveis de expressão reduzidos em um mutante relativamente avirulento de *N. caninum*, sugerindo a importância destas proteínas para a virulência do parasita¹¹⁹. RON2, RON4, RON5, e RON8 são proteínas presentes no “pescoço” das roptrias (*rhoptry neck*) e fazem parte do complexo da junção de movimento AMA1-RON2/4/5/8 em *T. gondii*¹²². Em um lisado de *N. caninum*, RON2 e RON5 foram coprecipitadas, indicando participação no complexo da junção de movimento, junto com RON4 e RON8¹²³. Algumas proteínas, como as parálogas ROP19 e ROP47, foram preditas como cinases de roptrias, e outras, como TLN1 e SUB2, são serino-proteases^{124,125}.

Apesar de em menor abundância que as proteínas de micronemas e de roptrias, algumas proteínas interessantes de grânulos densos foram encontradas; 8 foram identificadas pelas duas abordagens, 1 somente no ESA, e outras 8 somente no ET. As proteínas de grânulos densos mais abundantes no ESA foram GRA9, serpina (NCLIV_063150), NTPasell, e GRA2. Dois membros da família serpina (inibidores de serino-proteases) foram identificados, PI-2 (*protease inhibitor 2*), que demonstrou atividade contra tripsina em *T. gondii*¹²⁶, e serpin (NCLIV_063150), provavelmente um novo membro da família serpina, ainda não descrito. PI-1, em *T. gondii*, é secretada dos grânulos densos para o interior do vacúolo parasitóforo, inibindo tripsina e quimiotripsina, o que sugere um possível papel protetor para o parasita¹²⁷. A presença destes inibidores de serino-proteases pode ser importante para a capacidade do parasita de sobreviver dentro do hospedeiro¹²⁶.

Doze proteínas de superfície foram identificadas no ESA e, apesar de 11 delas terem sido também quantificadas, nenhuma entrou no grupo de diminuídas nos taquizoítas estimulados. SRS2, uma candidata em potencial à vacina contra neosporose^{128,129}, mostrou-se altamente abundante no ESA de *N. caninum*, assim como SAG1 e SAG3, que são alvos para vacina anti-toxoplasmose^{130,131}. Três proteínas (NCLIV_010720, NCLIV_010730 e NCLIV_068920) compõem um grupo com quatro proteínas de superfície (SRS) parálogas em *N. caninum*. Um outro grupo de proteínas SRS, que inclui NCLIV_068872 e NCLIV_052740, contém 15 parálogos que não possuem homologia em *T. gondii*.

Também foi observada a presença de 8 ciclofilinas no ESA; somente a proteína 18 kDa *cyclophilin* foi previamente identificada⁹⁴. Entre as sete proteínas putativas, apenas NCLIV_019970 encontra-se no grupo de proteínas diminuídas (razão = 0.483), por quantificação no ET, com um potencial, portanto, para ser secretada. Em *N. caninum* e *T. gondii*, 18 kDa *cyclophilin* apresenta propriedades de atração de monócitos, pois mimetiza os ligantes do receptor 5 para CC-quimiocina (CCR5). Desta forma, ela induz migração celular para os sítios de infecção e aumenta a invasão e proliferação do parasita^{132,133}. A presença de várias ciclofilinas no ESA pode estar, portanto, relacionada à regulação de migração celular e proliferação do parasita.

Além de proteínas secretadas, proteínas putativamente envolvidas com o processo de secreção também foram analisadas. Estudos com *Plasmodium falciparum* e *Toxoplasma gondii* indicam que, seguindo-se o contato com a célula hospedeira, os mensageiros inositol trifosfato (IP3) e ADP ribose cíclica (cADPR) são gerados e ligam-se a canais de cálcio putativamente localizados no retículo endoplasmático, liberando o Ca^{2+} ali armazenado^{20,52}. Eventos decorrentes da mobilização de cálcio, como ativação de CDPKs (proteínas cinases dependentes de cálcio) e ativação de outras cinases por cAMP e cGMP, regulam a exocitose de organelas secretórias e a ativação do motor de invasão do parasita^{20,52,53}. Algumas proteínas quantificadas de *N. caninum* foram localizadas dentro de duas vias metabólicas sensíveis ao aumento de cálcio: via do inositol fosfato e via de purinas, descritas em *Library of Apicomplexan Metabolic Pathways*¹³⁴ (LAMP). Entre as 13 proteínas identificadas relacionadas ao metabolismo de inositol fosfato, a enzima fosfatidil-inositídeo 5-fosfatase (NCLIV_000140) apresentou-se diminuída no taquizoíta *estimulado*, o que facilita o acúmulo dos precursores de IP3 (figura 61). Cinco homólogos da enzima fosfatidilinositol-4-fosfato-5-cinase, que favorece a síntese de precursores de IP3, foram identificados. É notável que a síntese de inosina monofosfato (IMP) esteja também favorecida, uma vez que a enzima adenosina cinase foi diminuída, enquanto a enzima AMP deaminase ($\text{AMP} \rightarrow \text{IMP}$) foi aumentada (figura 62). O acúmulo de IMP pode resultar em um aumento de GDP (precursor de cGMP), já que as enzimas GMP sintase e guanilato cinase foram também aumentadas (figura 63).

Metabolismo de inositol fosfato com proteínas identificadas e quantificadas em *N. caninum*

Fig. 61 Metabolismo de inositol fosfato, adaptado de LAMP (*Library of Apicomplexan Metabolic Pathways*). As proteínas quantificadas no taquizoíta estimulado de *N. caninum* estão circulas em cores que designam seu nível de expressão (vermelho, diminuídas; amarelo, não diferencialmente expressas) e são mostradas as suas razões estimulado/controle ($\log_2$).

Metabolismo de purinas: adenosina cinase diminuída e AMP deaminase aumentada

Fig. 62 Metabolismo de purinas, recorte da via mostrando adenosina cinase diminuída e AMP deaminase aumentada. (Adaptado de LAMP - *Library of Apicomplexan Metabolic Pathways*). As proteínas quantificadas no taquizoíta *estimulado* de *N. caninum* estão circuladas em cores que designam seu nível de expressão (azul, aumentadas; vermelho, diminuídas; amarelo, não diferencialmente expressas) e são mostradas as suas razões *estimulado/control* (log₂).

Outras proteínas relacionadas à mobilização por cálcio foram identificadas pela Abordagem Quantitativa, como as homólogas de *T. gondii* que são potencialmente envolvidas com mediação de sinalização de cascata intracelular, regulação de exocitose de organelas secretórias, e controle da mobilidade do parasita⁵² (tabela 27). Nebl e colegas utilizaram o mesmo estímulo por etanol sobre o taquizoíta de *T. gondii* para análise quantitativa do parasita *estimulado*, porém, para identificar sítios de fosforilação cálcio-dependentes⁵². Como um exemplo encontrado nos dois estudos, a proteína ARM1 (*armadillo repeat-containing protein*), encontrada entre as proteínas fosforiladas no estudo com *T. gondii*, mostrou-se aumentada em nossa quantificação. Apesar de ainda não caracterizada em *N. caninum* ou *T. gondii*,

proteínas de outros organismos que contêm domínio ARM estão relacionadas a processos importantes, como sinalização intracelular e regulação do citoesqueleto⁷³.

Tabela 27 Proteínas quantificadas no taquizoítas estimulado de *N. caninum* potencialmente envolvidas com invasão. Os homólogos destas proteínas foram identificados por Nebel and colegas⁵² como potencialmente envolvidos com: mediação de sinalização de cascata intracelular, regulação de exocitose de organelas secretórias, ou controle de mobilidade do parasita. **Razão Interm./Leve** (taquizoítas estimulado/controle). Razão menor que 0.5, proteínas diminuídas; razão maior que 2.0, proteínas aumentadas no taquizoíta estimulado. A única proteína que apresentou-se aumentada foi ARM1, cuja razão maior que 2.0 está destacada com grifo cinza. Todas as proteínas listadas são hipotéticas em *N. caninum*, portanto, são mostrados os códigos das proteínas de *N. caninum* (NCLIV) com a descrição para as proteínas em *T. gondii*, e ao final são mostrados os códigos de seus respectivos homólogos em *T. gondii* (TGME49).

ToxoDB ID	Descrição (<i>T. gondii</i>)	Razão Interm./Leve	Homólogos em <i>T. gondii</i>
NCLIV_0260	ARM1, armadillo repeat-containing protein	2.075	TGME49_061440
NCLIV_017370	PKA-R, cAMP kinase regulatory subunit	1.093	TGME49_042070
NCLIV_034900	CKIIβ, casein kinase subunit II beta subunit	0.879	TGME49_072400
NCLIV_061620	SNAP, soluble NSF-attachment protein	0.848	TGME49_018760
NCLIV_037540	CKIIβ, casein kinase subunit II beta subunit	0.844	TGME49_072400
NCLIV_018710	SET, I2PP2A/SET	0.729	TGME49_044110
NCLIV_009030	PP2C, serine/threonine phosphatase 2C	0.666	TGME49_054770
NCLIV_035310	ANP32, similar to I1PP2A/ANP32	–	TGME49_071810

Os principais componentes do motor de invasão MyoA (NCLIV_049900), MLC1 (NCLIV_029420), GAP40 (NCLIV_065970), GAP45 (NCLIV_048570), GAP50 (NCLIV_061160) e ELC1 (NCLIV_036970) também foram aqui encontrados. Porém, apesar de a mobilidade mostrar-se relacionada à via de transdução sinalizada por cálcio⁵², nenhuma destas proteínas apresentou-se aumentada. Algumas CDPKs (proteínas cinases dependentes de cálcio), implicadas em vias cruciais de transdução de sinal cálcio-dependentes em parasitas Apicomplexas²⁰, foram identificadas em nosso estudo. Em *T. gondii*, CDPK1 é um regulador essencial da exocitose de micronemas¹³⁵; porém nenhum homólogo sintênico é predito em *N. caninum*. A proteína CDPK mais expressa no taquizoíta estimulado de *N. caninum* foi CDPK2A (NCLIV_044270; razão 2.086), seguida por CDPK7 (NCLIV_044860; razão 1.645). Além destas cinases, algumas CMGC (cinases associadas ao ciclo celular), como NCLIV_020950 (razão 2.452), NCLIV_007880 (razão 2.442) e NCLIV_001240 (razão 2.233) mostraram-se também suprarreguladas.

A rede de interação entre as proteínas identificadas e quantificadas revelou proteínas centrais envolvidas com vias induzidas por cálcio e potencialmente com invasão. A exploração de interações proteína-proteína tem sido uma nova estratégia para identificar alvos para medicamentos contra a malária, que devem ter efeito em

nodos de importância funcional, controlando redes de interação cruciais para a sobrevivência do parasita¹³⁶. A proteína aumentada NCLIV_031320 exibiu o maior número de interações (com 466 proteínas). NCLIV_031320 pode ser uma potencial proteína de ligação a RNA, pois contém um sinal para localização nuclear (NLS)^{25,137} e uma região rica em glicina, propensa a promover interações que criam complexos de ribonucleoproteínas (RNP)¹³⁸. Três proteínas aumentadas apresentaram 129 interações: toxofilina (NCLIV_051340); um putativo fator de iniciação de tradução (NCLIV_025130); e uma proteína hipotética (NCLIV_006060). Duas proteínas não-diferencialmente expressas também exibiram alto número de interações: o homólogo da proteína de *T. gondii* small GTPase Rab6 (NCLIV_054540), com 168 interações; e a proteína RACK (NCLIV_059430), um receptor putativo para cinase C, com 148 interações. Um resultado interessante mostra seis proteínas preditas à interação tanto com proteínas diminuídas quanto aumentadas (entre 17 e 38 interações por proteína), logo, forte candidatas a estarem envolvidas com o processo de invasão. Entre elas, apenas a proteína 18 kDa *cyclophilin* (NCLIV_004790) foi previamente caracterizada⁹⁴; as outras cinco são proteínas hipotéticas (NCLIV_006470, NCLIV_002780, NCLIV_019600, NCLIV_016700, NCLIV_022690).

A anotação de um genoma, susceptível a erros e omissões por predição computacional, pode ser melhorada experimentalmente com a identificação de peptídeos em regiões codificadoras não-annotadas, contribuindo para o refinamento dos atuais modelos de genes e para identificação de *splicing* alternativo¹³⁹. Em nossa análise proteogenômica, dos 95 peptídeos não-annotados que identificamos, 5 peptídeos presentes no cromossomo Ib apontam para uma proteína em potencial, que seria composta pelos domínios EGF-1, beta-integrina, e VWFC-1. A região que compreende 6 peptídeos no cromossomo X, quando traduzida, apresentou o domínio ácido metalofosfatase (MPP), seguido por um domínio SbcC, presente em ATPases e envolvido com reparo de DNA, além de um domínio para manutenção estrutural de cromossomos (SMC). Aproximadamente 200 aminoácidos após estes domínios, existe uma região que contém inúmeras sequências ricas em ácido glutâmico, que poderia conferir alta solubilidade à proteína.

A investigação do secretoma de *N. caninum* com as técnicas aqui utilizadas permitiu identificar um grande número de proteínas e suas abundâncias relativas no extrato secretado. A comparação entre a abundância das proteínas identificadas no ESA e sua quantificação no taquizoíta *estimulado* mostrou-se uma ferramenta

poderosa na identificação de proteínas com maior probabilidade de secreção, contornando o problema da presença no ESA de proteínas não secretadas. Os resultados de interação entre proteínas e as informações sobre as proteínas aumentadas em resposta ao estímulo secretório foram importantes para o reconhecimento de proteínas potencialmente envolvidas com a regulação do metabolismo relacionado à secreção. Os resultados aqui obtidos terão grande utilidade para guiar a busca e seleção de alvos a serem investigados para o desenvolvimento de tratamento terapêutico contra *Neospora caninum* e outros organismos apicomplexas.

6. CONCLUSÃO

6.1 Proteínas com domínios Apple de *N. caninum*

1) MIC17A, MIC17B e MIC17C são proteínas compostas por quatro domínios Apple: MIC17A e MIC17C compartilham alta identidade sequencial (97%) entre si, enquanto MIC17B possui apenas 65% de identidade em relação às outras duas;

2) Estruturalmente MIC17A e MIC17C apresentam uma cisteína adicional e um padrão de distribuição de cargas na superfície de seus domínios bastante diferente de MIC17B;

3) A predição estrutural combinando os quatro domínios Apple, baseada no tetrâmero do fator de coagulação XI, resultou em estrutura instável;

4) MIC17B e MIC17C estão presentes no extrato total e no extrato secretado de *N. caninum*, com ponto isoelétrico numa faixa entre 8.0 e 8.4, apresentando pesos moleculares de aproximadamente 29, 31, 35 e 40 kDa para MIC17B, e 27, 31, 35 e 40 kDa para MIC17C;

5) MIC17B localiza-se na região apical do taquizoíta de *N. caninum*, enquanto MIC17C mostra-se distribuída ao longo do parasita, com concentrações ao redor do núcleo;

6) O bloqueio de MIC17B e MIC17C com seus respectivos anticorpos inibe 24,2% e 15,5% respectivamente a invasão pelo parasita;

7) O parasita nocaute para os genes das proteínas Apple não foi obtido e, com os métodos utilizados, a propriedade de adesão das proteínas Apple de *N. caninum* à célula hospedeira não pôde ser comprovada.

8) As diferenças encontradas entre MIC17B e MIC17C quanto às suas estruturas, ao padrão de bandas detectadas por *Western blot*, e à localização no taquizoíta, sugerem possíveis diferentes funções para estas proteínas.

6.2 Secretoma de *N. caninum*

1) De 615 proteínas identificadas no extrato secretado (ESA) de *N. caninum*, 37,4% (n=230) representam proteínas conhecidas ou preditas como secretadas;

2) A técnica de quantificação relativa de proteínas no taquizoíta submetido ao estímulo secretório é inédita para o estudo do secretoma em apicomplexas. De 2011 proteínas quantificadas nos extratos dos parasitas *estimulado* e *controle*, 150 proteínas apresentaram-se aumentadas e 90 diminuídas no taquizoíta *estimulado*;

3) A razão negativa obtida na quantificação *estimulado/controle* presente na maioria das proteínas das organelas secretórias micronemas, roptrias e grânulos densos demonstra que a quantificação pode ser utilizada como auxílio na identificação de proteínas secretadas, porém com um limite para proteínas diminuídas menos rígido do que o adotado neste trabalho;

4) A rede de interação entre as 2011 proteínas quantificadas de *N. caninum* prevê 6570 interações, sendo 651 interações que envolvem proteínas aumentadas; e 862 interações que envolvem proteínas diminuídas. A proteína hipotética NCLIV_031320, aumentada no parasita estimulado, representa a proteína com maior número de interações (466). Entre as proteínas diminuídas, toxofilina (NCLIV_051340) e as proteínas hipotéticas NCLIV_025130 e NCLIV_006060 representam as mais interagentes.

5) Por meio de análise proteogenômica, foram identificados 95 peptídeos em regiões não anotadas do genoma, incluindo duas proteínas em potencial;

6) As proteínas identificadas em abundância no extrato secretado e diminuídas no parasita *estimulado*, assim como as proteínas potencialmente envolvidas com alto número de interações e com a regulação do metabolismo relacionado à secreção, compõem alvos em potencial a serem investigados para o desenvolvimento de tratamento terapêutico contra *Neospora caninum* e outros organismos apicomplexas.

7. REFERÊNCIAS

- ¹ BJERKÅS, I.; MOHN, S. F.; PRESTHUS, J. Unidentified cyst-forming sporozoan causing encephalomyelitis and myositis in dogs. **Z Parasitenkd**, v. 70, n. 2, p. 271-4, 1984.
- ² DUBEY, J. P.; CARPENTER, J. L.; SPEER, C. A.; et al. Newly recognized fatal protozoan disease of dogs. **Journal of the American Veterinary Medical Association**, v. 192, n. 9, p. 1269-1285, May 1 1988.
- ³ DUBEY, J. P.; BARR, B.; BARTA, J.; et al. Redescription of *Neospora caninum* and its differentiation from related coccidia. **International Journal for Parasitology**, v. 32, n. 8, p. 929-946, 2002.
- ⁴ MCCANN, C. M.; VYSE, A. J.; SALMON, R. L.; et al. Lack of serologic evidence of *Neospora caninum* in humans, England. **Emerging Infectious Diseases**, v. 14, n. 6, p. 978-980, 2008.
- ⁵ DUBEY, J. P.; SCHARES, G.; ORTEGA-MORA, L. M. Epidemiology and control of neosporosis and *Neospora caninum*. **Clinical Microbiology Reviews**, v. 20, n. 2, p. 323-367, 2007.
- ⁶ REICHEL, M. P.; ALEJANDRA AYANEGUI-ALCÉRRECA, M.; GONDIM, L. F.; ELLIS, J. T. What is the global economic impact of *Neospora caninum* in cattle - The billion dollar question. **International Journal for Parasitology**, v. 43, n. 2, p. 133-142, 2013.
- ⁷ WILLIAMS, D. J. L.; HARTLEY, C. S.; BJÖRKMAN, C.; et al. Endogenous and exogenous transplacental transmission of *Neospora caninum* - how the route of transmission impacts on epidemiology and control of disease. **Parasitology**, v. 136, n. 14, p. 1895-1900, 2009.
- ⁸ MCALLISTER, M. M.; DUBEY, J.; LINDSAY, D. S.; et al. Dogs are definitive hosts of *Neospora caninum*. **International Journal of Parasitology**, v. 28, n. 9, p. 1473-8, 1998.
- ⁹ LINDSAY, D. S.; DUBEY, J. P.; DUNCAN, R. B. Confirmation that the dog is a definitive host for *Neospora caninum*. **Veterinary Parasitology**, v. 82, n. 4, p. 327-33, 1999.
- ¹⁰ GONDIM, L. F. P.; MCALLISTER, M. M.; PITT, W. C.; et al. Coyotes (*Canis latrans*) are definitive hosts of *Neospora caninum*. **International Journal of Parasitology**, v. 34, n. 2, p. 159-61, 2004.
- ¹¹ KING, J. S.; SLAPETA, J.; JENKINS, D. J.; et al. Australian dingoes are definitive hosts of *Neospora caninum*. **International Journal of Parasitology**, v. 40, n. 8, p. 945-50, 2010.
- ¹² DUBEY, J. P.; SCHARES, G. Neosporosis in animals - the last five years. **Veterinary Parasitology**, v. 180, n. 1-2, p. 90-108, 2011.
- ¹³ DUBEY, J. P.; BUXTON, D.; WOUDA, W. Pathogenesis of bovine neosporosis. **Journal of Comparative Pathology**, v. 134, n. 4, p. 267-289, 2006.

- 14 HEMPHILL, A. Subcellular localization and functional characterization of Nc-p43, a major *Neospora caninum* tachyzoite surface protein. **Infection and Immunity**, v. 64, n. 10, p. 4279-4287, 1996.
- 15 BAUM, J.; GILBERGER, T. W.; FRISCHKNECHT, F.; et al. Host-cell invasion by malaria parasites: insights from *Plasmodium* and *Toxoplasma*. **Trends in Parasitology**, v. 24, n. 12, p. 557-563, 2008.
- 16 CARRUTHERS, V. B.; BOOTHROYD, J. C. Pulling together: an integrated model of *Toxoplasma* cell invasion. **Current Opinion in Microbiology**, v. 10, n. 1, p. 82-89, 2007.
- 17 SANTOS, J. M.; SOLDATI-FAVRE, D. Invasion factors are coupled to key signalling events leading to the establishment of infection in apicomplexan parasites. **Cellular Microbiology**, v. 13, n. 6, p. 787-796, 2011.
- 18 JOINER, K. A.; ROOS, D. S. Secretory traffic in the eukaryotic parasite *Toxoplasma gondii*: less is more. **The Journal of Cell Biology**, v. 157, n. 4, p. 557-63, 2002.
- 19 SOLDATI, D.; MEISSNER, M. *Toxoplasma* as a novel system for motility. **Current Opinion in Cell Biology**, v. 16, n. 1, p. 32-40, 2004.
- 20 BILLKER, O.; LOURIDO, S.; SIBLEY, L. D. Calcium-Dependent Signaling and Kinases in Apicomplexan Parasites. **Cell Host & Microbe**, v. 5, n. 6, p. 612-622, 2009.
- 21 CARRUTHERS, V. B.; TOMLEY, F. M. Microneme proteins in apicomplexans. **Sub-cellular Biochemistry**, v. 47, p. 33-45, 2008.
- 22 TOMLEY, F. M.; SOLDATI, D. S. Mix and match modules: structure and function of microneme proteins in apicomplexan parasites. **Trends in Parasitology**, v. 17, n. 2, p. 81-88, 2001.
- 23 TORDAI, H.; BÁNYAI, L.; PATTHY, L. The PAN module: the N-terminal domains of plasminogen and hepatocyte growth factor are homologous with the apple domains of the prekallikrein family and with a novel domain found in numerous nematode proteins. **FEBS Lettters**, v. 461, n. 1-2, p. 63-67, 1999.
- 24 BROWN, P. J.; GILL, A. C.; NUGENT, P. G.; et al. Domains of invasion organelle proteins from apicomplexan parasites are homologous with the Apple domains of blood coagulation factor XI and plasma pre-kallikrein and are members of the PAN module superfamily. **FEBS Letters**, v. 497, n. 1, p. 31-38, 2001.
- 25 Prosite, Database of protein domains, families and functional sites. Disponível em: < <http://prosite.expasy.org/> >.
- 26 PUNTA, M.; COGGILL, P. C.; EBERHARDT, R. Y.; et al. The Pfam protein families database. **Nucleic Acid Research**, v. 40, p. D290-D301, 2011.
- 27 SMART, Simple Modular Architecture Research Tool. Disponível em: < <http://smart.embl.de/> >.
- 28 ESCHENBACHER, K. H.; KLEIN, H.; SOMMER, I.; et al. Characterization of cDNA clones encoding a major microneme antigen of *sarcocystis-muris* (Apicomplexa) cyst merozoites. **Molecular and Biochemical Parasitology**, v. 62, n. 1, p. 27-36, 1993.

- 29 KLEIN, H.; MEHLHORN, H.; RUGER, W. Characterization of genomic clones encoding two microneme antigens of *Sarcocystis muris* (Apicomplexa). **Parasitology Research**, v. 82, n. 3, p. 230-237, 1996.
- 30 KLEIN, H.; MUELLER, S.; LOESCHNER, B.; et al. Cloning, sequencing and recombinant expression of the open reading frame encoding a novel member of the *Sarcocystis muris* (Apicomplexa) microneme lectin family. **Parasitology Research**, v. 90, n. 1, p. 84-86, 2003.
- 31 BROWN, P. J.; BILLINGTON, K. J.; BUMSTEAD, J. M.; et al. A microneme protein from *Eimeria tenella* with homology to the Apple domains of coagulation factor XI and plasma pre-kallikrein. **Molecular and Biochemical Parasitology**, v. 107, n. 1, p. 91-102, 2000.
- 32 PERIZ, J.; GILL, A. C.; HUNT, L.; et al. The microneme proteins EtMIC4 and EtMIC5 of *Eimeria tenella* form a novel, ultra-high molecular mass protein complex that binds target host cells. **Journal of Biological Chemistry**, v. 282, n. 23, p. 16891-16898, 2007.
- 33 BRECHT, S.; CARRUTHERS, V. B.; FERGUSON, D. J.; et al. The toxoplasma micronemal protein MIC4 is an adhesin composed of six conserved apple domains. **Journal of Biological Chemistry**, v. 276, n. 6, p. 4119-4127, 2001.
- 34 REISS, M.; VIEBIG, N.; BRECHT, S.; et al. Identification and characterization of an escorter for two secretory adhesins in *Toxoplasma gondii*. **Journal of Cell Biology**, v. 152, n. 3, p. 563-578, 2001.
- 35 MARCHANT, J.; COWPER, B.; LIU, Y.; et al. Galactose recognition by the apicomplexan parasite *Toxoplasma gondii*. **Journal of Biological Chemistry**, v. 287, n. 20, p. 16720-16733, 2012.
- 36 BAI, T.; BECKER, M.; GUPTA, A.; et al. Structure of AMA1 from *Plasmodium falciparum* reveals a clustering of polymorphisms that surround a conserved hydrophobic pocket. **Proceedings of the National Academy of Sciences of the United States of America**, v. 102, n. 36, p. 12736-12741, 2005.
- 37 PAPAGRIGORIOU, E.; MCEWAN, P. A.; WALSH, P. N.; et al. Crystal structure of the factor XI zymogen reveals a pathway for transactivation. **Nature Structural & Molecular Biology**, v. 13, n. 6, p. 557-558, 2006.
- 38 BROWN, P. J.; MULVEY, D.; POTTS, J. R.; et al. Solution structure of a PAN module from the apicomplexan parasite *Eimeria tenella*. **Journal of Structural and Functional Genomics**, v. 4, p. 227-234, 2003.
- 39 LEE, E. G.; KIM, J. H.; SHIN, Y. S.; et al. Establishment of a two-dimensional electrophoresis map for *Neospora caninum* tachyzoites by proteomics. **Proteomics**, v. 3, n. 12, p. 2339-2350, 2003.
- 40 LEE, E. G.; KIM, J. H.; SHIN, Y. S.; et al. Two-dimensional gel electrophoresis and immunoblot analysis of *Neospora caninum* tachyzoites. **Journal of Veterinary Science**, v. 5, n. 2, p. 139-45, 2004.
- 41 LEE, E. G.; KIM, J. H.; SHIN, et al. Application of proteomics for comparison of proteome of *Neospora caninum* and *Toxoplasma gondii* tachyzoites. **Journal of**

Chromatography B-Analytical Technologies in the Biomedical and Life Sciences, v. 815, n. 1-2, p. 305-314, 2005.

- 42 SHIN, Y. S.; LEE, E. G.; SHIN, G. W.; et al. Identification of antigenic proteins from *Neospora caninum* recognized by bovine immunoglobulins M, E, A and G using immunoproteomics. **Proteomics**, v. 4, n. 11, p. 3600-3609, 2004.
- 43 SHIN, Y. S.; SHIN, G. W.; KIM, Y. R.; et al. Comparison of proteome and antigenic proteome between two *Neospora caninum* isolates. **Veterinary Parasitology**, v. 134, n. 1-2, p. 41-52, 2005.
- 44 ZHANG, H.; LEE, E. G.; YU, L.; KAWANO, S.; et al. Identification of the cross-reactive and species-specific antigens between *Neospora caninum* and *Toxoplasma gondii* tachyzoites by a proteomics approach. **Parasitology Research**, v. 109, n. 3, p. 899-911, 2011.
- 45 MARUGAN-HERNANDEZ, V. ALVAREZ-GARCÍA, G.; RISCO-CASTILLO, V.; et al. Identification of *Neospora caninum* proteins regulated during the differentiation process from tachyzoite to bradyzoite stage by DIGE. **Proteomics**, v. 10, n. 9, p. 1740-1750, 2010.
- 46 REGIDOR-CERRILLO, J.; ÁLVAREZ-GARCÍA, G.; PASTOR-FERNÁNDEZ, I.; et al. Proteome expression changes among virulent and attenuated *Neospora caninum* isolates. **Journal of Proteomics**, v. 75, n. 8, p. 2306-2318, 2012.
- 47 SOHN, C. S.; CHENG, T. T.; DRUMMOND, M. L.; et al. Identification of novel proteins in *Neospora caninum* using an organelle purification and monoclonal antibody approach. **Plos One**, v. 6, n. 4, 2011.
- 48 ROCCHI, M. S.; BARTLEY, P. M.; INGLIS, N. F.; et al. Selection of *Neospora caninum* antigens stimulating bovine CD4(+ve) T cell responses through immunopotency screening and proteomic approaches. **Veterinary Research**, v. 42, Aug 3 2011.
- 49 MARUGAN-HERNANDEZ, V.; ALVAREZ-GARCÍA, G.; TOMLEY, F.; et al. Identification of novel rhoptry proteins in *Neospora caninum* by LC/MS-MS analysis of subcellular fractions. **Journal of Proteomics**, v. 74, n. 5, p. 629-642, 2011.
- 50 CARRUTHERS, V. B.; SIBLEY, L. D. Mobilization of intracellular calcium stimulates microneme discharge in *Toxoplasma gondii*. **Molecular Microbiology**, v. 31, n. 2, p. 421-428, 1999.
- 51 LOVETT, J. L.; MARCHESINI, N.; MORENO, S. N.; et al. *Toxoplasma gondii* microneme secretion involves intracellular Ca²⁺ release from inositol 1,4,5-triphosphate (IP3)/ryanodine-sensitive stores. **Journal of Biological Chemistry**, v. 277, n. 29, p. 25870-25876, 2002.
- 52 NEBL, T.; PRIETO, J. H.; KAPP, E.; et al. Quantitative in vivo analyses reveal calcium-dependent phosphorylation sites and identifies a novel component of the *Toxoplasma* invasion motor complex. **Plos Pathogens**, v. 7, n. 9, 2011.
- 53 SINGH, S.; CHITNIS, C. E. Signalling mechanisms involved in apical organelle discharge during host cell invasion by apicomplexan parasites. **Microbes and Infection**, v. 14, n. 10, p. 820-824, 2012.

- 54 CARRUTHERS, V. B.; MORENO, S. N. J.; SIBLEY, L. D. Ethanol and acetaldehyde elevate intracellular Ca^{2+} and stimulate microneme discharge in *Toxoplasma gondii*. **Biochemical Journal**, v. 342, p. 379-386, 1999.
- 55 ZHOU, X. W.; KAFSACK, B. F.; COLE, R. N.; et al. The opportunistic pathogen *Toxoplasma gondii* deploys a diverse legion of invasion and survival proteins. **Journal of Biological Chemistry**, v. 280, n. 40, p. 34233-34244, 2005.
- 56 FRESE, C. K.; ALTELAAR, A. F.; HENNRICH, M. L.; et al. improved peptide identification by targeted fragmentation using CID, HCD and ETD on an LTQ-Orbitrap Velos. **Journal of Proteome Research**, v. 10, n. 5, p. 2377-2388, 2011.
- 57 BOERSEMA, P. J.; RAIJMAKERS, R.; LEMEER, S.; et al. Multiplex peptide stable isotope dimethyl labeling for quantitative proteomics. **Nature Protocols**, v. 4, n. 4, p. 484-494, 2009.
- 58 INNES, E. A.; VERMEULEN, A. N. Vaccination as a control strategy against the coccidial parasites *Eimeria*, *Toxoplasma* and *Neospora*. **Parasitology**, v. 133, p. S145-S168, 2006.
- 59 HEMPHILL, A.; DEBACHE, K.; MONNEY, T.; et al. Proteins mediating the *Neospora caninum*-host cell interaction as targets for vaccination. **Front Biosci (Elite Ed)**, v. 5, p. 23-36, 2013.
- 60 CHEN, Z.; HARB, O. S.; ROOS, D. S. In silico identification of specialized secretory-organellar proteins in apicomplexan parasites and in vivo validation in *Toxoplasma gondii*. **Plos One**, v. 3, n. 10, Oct 31 2008.
- 61 DUBEY, J. P.; HATTEL, A. L.; LINDSAY, D. S.; et al. Neonatal *Neospora caninum* infection in dogs: isolation of the causative agent and experimental transmission. **Journal of the American Veterinary Medical Association**, v. 193, n. 10, p. 1259-63, 1988.
- 62 BRADFORD, M. M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. **Analytical Biochemistry**, v. 7, p. 248-254, 1976.
- 63 NEUHOFF, V.; AROLD, N.; TAUBE, D.; et al. Improved staining of proteins in polyacrylamide gels including isoelectric focusing gels with clear background at nanogram sensitivity using Coomassie Brilliant Blue G-250 and R-250. **Electrophoresis**, v. 9, n. 6, p. 255-62, 1988.
- 64 SHEVCHENKO, A.; WILM, M.; VORM, O.; et al. Mass spectrometric sequencing of proteins silver-stained polyacrylamide gels. **Analytical Chemistry**, v. 68, n. 5, p. 850-8, 1996.
- 65 Sanger Institute. Disponível em: < <http://www.sanger.ac.uk/> >
- 66 GAJRIA, B.; BAHL, A.; BRESTELLI, J.; et al. ToxoDB: an integrated *Toxoplasma gondii* database resource. **Nucleic Acids Research**, v. 36, p. D553-D556, 2008.
- 67 ESWAR, N.; WEBB, B.; MARTI-RENO, M. A.; et al. Comparative protein structure modeling using Modeller. **Current Protocols in Bioinformatics**, v. Chapter 5, p. Unit 5.6, 2006.

- 68 RCSB Protein Data Bank - PDB. Disponível em:< <http://www.rcsb.org/pdb/>>
- 69 CORPET, F. Multiple sequence alignment with hierarchical clustering. **Nucleic Acid Res**, v. 16, n. 22, p. 10881-10890, 1988.
- 70 CHEN, V. B.; ARENDALL, W. B.; HEADD, J. J.; et al. MolProbity: all-atom structure validation for macromolecular crystallography. **Acta Crystallographica**, v. D66, p. 12-21, 2010.
- 71 SCHRÖDINGER, L. **The PyMOL Molecular Graphics System**.
- 72 LESSEL, U.; SCHOMBURG, D. Similarities between protein 3-D structures. **Protein Engineering**, v. 7, n. 10, p. 1175-87, 1994.
- 73 WINN, M. D.; BALLARD, C. C.; COWTAN, K. D.; et al. Overview of the CCP4 suite and current developments. **Acta Crystallographica, section D, Biological Crystallography**, v. 67, n. Pt 4, p. 235-42, 2011.
- 74 BAKER, N. A.; SEPT, D.; JOSEPH, S.; et al. Electrostatics of nanosystems: application to microtubules and the ribosome. **Proceedings of the National Academy of Sciences of the United States of America**, v. 98, n. 18, p. 10037-41, 2001.
- 75 DOLINSKY, T. J.; NIELSEN, J. E.; MCCAMMON, J. A.; et al. PDB2PQR: an automated pipeline for the setup of Poisson-Boltzmann electrostatics calculations. **Nucleic Acids Research**, v. 32, p. W665-W667, 2004.
- 76 BRESOLIN, I. T.; BORSOI-RIBEIRO, M.; TAMASHIRO, W. M.; et al. Evaluation of immobilized metal-ion affinity chromatography (IMAC) as a technique for IgG(1) monoclonal antibodies purification: the effect of chelating ligand and support. **Applied Biochemistry and Biotechnology**, v. 160, n. 7, p. 2148-65, 2010.
- 77 HEMPHILL, A.; FUCHS, N.; SONDA, S.; et al. Identification and partial characterization of a 36 kDa surface protein on *Neospora caninum* tachyzoites. **Parasitology**, v. 115 (Pt 4), p. 371-80, 1997.
- 78 HOWE, D. K.; MERCIER, C.; MESSINA, M.; et al. Expression of *Toxoplasma gondii* genes in the closely-related apicomplexan parasite *Neospora caninum*. **Molecular and Biochemical Parasitology**, v. 86, n. 1, p. 29-36, 1997.
- 79 SEEGER, F.; BOOTHROYD, J. C. *Escherichia coli* beta-galactosidase as an *in vitro* and *in vivo* reporter enzyme and stable transfection marker in the intracellular protozoan parasite *Toxoplasma gondii*. **Gene**, v. 169, n. 1, p. 39-45, 1996.
- 80 PEREIRA, L. L. **Avaliação funcional da proteína anônima relacionada à trombospondina 2 de *Neospora caninum* (NcMIC2-like1)**. 2013. (PhD). Faculdade de Ciências Farmacêuticas de Ribeirão Preto., Universidade de São Paulo
- 81 VAN POPPEL, N. F.; WELAGEN, J.; DUISTERS, R. F.; et al. Tight control of transcription in *Toxoplasma gondii* using an alternative tet repressor. **International Journal of Parasitology**, v. 36, n. 4, p. 443-52, 2006.

- 82 GARCIA-RÉGUET, N.; LEBRUN, M.; FOURMAUX, M. N.; et al. The microneme protein MIC3 of *Toxoplasma gondii* is a secretory adhesin that binds to both the surface of the host cells and the surface of the parasite. **Cellular Microbiology**, v. 2, n. 4, p. 353-64, 2000.
- 83 KELLER, A.; NESVIZHSHKII, A. I.; KOLKER, E.; et al. Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. **Analytical Chemistry**, v. 74, n. 20, p. 5383-5392, 2002.
- 84 NESVIZHSHKII, A. I.; KELLER, A.; KOLKER, E.; et al. A statistical model for identifying proteins by tandem mass spectrometry. **Analytical Chemistry**, v. 75, n. 17, p. 4646-4658, 2003.
- 85 BENDTSEN, J. D.; JENSEN, L. J.; BLOM, N.; et al. Feature-based prediction of non-classical and leaderless protein secretion. **Protein Engineering Design & Selection**, v. 17, n. 4, p. 349-356, 2004.
- 86 REID, A. J.; VERMONT, S. J.; COTTON, J. A.; et al. Comparative genomics of the apicomplexan parasites *Toxoplasma gondii* and *Neospora caninum*: coccidia differing in host range and transmission strategy. **Plos Pathogens**, v. 8, n. 3, 2012.
- 87 GAUCI, S.; HELBIG, A. O.; SLIJPER, M.; et al. Lys-N and trypsin cover complementary parts of the phosphoproteome in a refined SCX-based approach. **Analytical Chemistry**, v. 81, n. 11, p. 4493-501, 2009.
- 88 VALENTE, G. T.; ACENCIO, M. L.; MARTINS, C.; et al. The development of a universal *in silico* predictor of protein-protein interactions. **Plos One**, v. 8, n. 5, p. e65587, 2013.
- 89 RAO, H. B.; ZHU, F.; YANG, G. B.; et al. Update of PROFEAT: a web server for computing structural and physicochemical features of proteins and peptides from amino acid sequence. **Nucleic Acids Research**, v. 39, p. W385-W390, 2011.
- 90 SMOOT, M. E.; ONO, K.; RUSCHEINSKI, J.; et al. Cytoscape 2.8: new features for data integration and network visualization. **Bioinformatics**, v. 27, n. 3, p. 431-432, 2011.
- 91 ROOS, F. F.; JACOB, R.; GROSSMANN, J.; et al. PepSplice: cache-efficient search algorithms for comprehensive identification of tandem mass spectra. **Bioinformatics**, v. 23, n. 22, p. 3016-3023, 2007.
- 92 THORVALDSDOTTIR, H.; ROBINSON, J. T.; MESIROV, J. P. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. **Briefings in Bioinformatics**, v. 14, n. 2, p. 178-192, 2013.
- 93 UniProt. Disponível em: < <http://www.uniprot.org/?tab=align>>
- 94 TUO, W. B.; FETTERER, R.; JENKINS, M.; et al. Identification and characterization of *Neospora caninum* cyclophilin that elicits gamma interferon production. **Infection and Immunity**, v. 73, n. 8, p. 5093-5100, 2005.
- 95 ALBERTS, B.; BRAY, D.; LEWIS, J.; et al. **Biologia Molecular da Célula**. 5 Ed. Porto Alegre: Artmed, 2010.

- 96 REHKOPF, D. H.; GILLESPIE, D. E.; HARRELL, M. I.; et al. Transcriptional mapping and RNA processing of the *Plasmodium falciparum* mitochondrial mRNAs. **Molecular and Biochemical Parasitology**, v. 105, n. 1, p. 91-103, 2000.
- 97 WILSON, R. J.; DENNY, P. W.; PREISER, P. R.; et al. Complete gene map of the plastid-like DNA of the malaria parasite *Plasmodium falciparum*. **Journal of Molecular Biology**, v. 261, n. 2, p. 155-72, 1996.
- 98 PREUßER, C.; JAÉ, N.; BINDEREIF, A. mRNA splicing in trypanosomes. **International Journal of Medical Microbiology**, v. 302, n. 4-5, p. 221-4, 2012.
- 99 LUKES, J.; LEANDER, B. S.; KEELING, P. J. Cascades of convergent evolution: the corresponding evolutionary histories of euglenozoans and dinoflagellates. **Proceedings of the National Academy of Sciences of the United States of America**, v. 106 Suppl 1, p. 9963-70, 2009.
- 100 BACHVAROFF, T. R.; PLACE, A. R. From stop to start: tandem gene arrangement, copy number and trans-splicing sites in the dinoflagellate *Amphidinium carterae*. **PLoS One**, v. 3, n. 8, p. e2929, 2008.
- 101 DEVANY, M.; KOTHARU, N. P.; MATSUO, H. Solution NMR structure of the C-terminal domain of the human protein DEK. **Protein Science**, v. 13, n. 8, p. 2252-9, 2004.
- 102 SOARES, L. M.; ZANIER, K.; MACKERETH, C.; et al. Intron removal requires proofreading of U2AF/3' splice site recognition by DEK. **Science**, v. 312, n. 5782, p. 1961-5, 2006.
- 103 GONG, H.; GONG, H.; KOBAYASHI, K.; et al. A novel PAN/apple domain-containing protein from *Toxoplasma gondii*: characterization and receptor identification. **PLoS One**, v. 7, n. 1, p. e30169, 2012.
- 104 BAYER-SANTOS, E.; AGUILAR-BONAVIDES, C.; RODRIGUES, S. P.; et al. Proteomic analysis of *Trypanosoma cruzi* secretome: characterization of two populations of extracellular vesicles and soluble proteins. **Journal of Proteome Research**, v. 12, n. 2, p. 883-897, 2013.
- 105 CORRALES, R. M.; SERENO, D.; MATHIEU-DAUDE, F. Deciphering the Leishmania exoproteome: what we know and what we can learn. **Fems Immunology and Medical Microbiology**, v. 58, n. 1, p. 27-38, 2010.
- 106 SHEINER, L.; SANTOS, J. M.; KLAGES, N.; PARUSSINI, F.; JEMMELY, N.; FRIEDRICH, N.; WARD, G. E.; SOLDATI-FAVRE, D. *Toxoplasma gondii* transmembrane microneme proteins and their modular design. **Molecular Microbiology**, v. 77, n. 4, p. 912-929, 2010.
- 107 LALIBERTE, J.; CARRUTHERS, V. B. *Toxoplasma gondii* toxolysin 4 is an extensively processed putative metalloproteinase secreted from micronemes. **Molecular and Biochemical Parasitology**, v. 177, n. 1, p. 49-56, 2011.
- 108 BRYDGES, S. D.; ZHOU, X. W.; HUYNH, M. H.; et al. Targeted deletion of MIC5 enhances trimming proteolysis of *Toxoplasma* invasion proteins. **Eukaryotic Cell**, v. 5, n. 12, p. 2174-2183, 2006.

- 109 LAGAL, V.; BINDER, E. M.; HUYNH, M. H.; et al. *Toxoplasma gondii* protease TgSUB1 is required for cell surface processing of micronemal adhesive complexes and efficient adhesion of tachyzoites. **Cellular Microbiology**, v. 12, n. 12, p. 1792-1808, 2010.
- 110 ZHANG, H.; COMPAORE, M. K.; LEE, E. G.; et al. Apical membrane antigen 1 is a cross-reactive antigen between *Neospora caninum* and *Toxoplasma gondii*, and the anti-NcAMA1 antibody inhibits host cell invasion by both parasites*. **Molecular and Biochemical Parasitology**, v. 151, n. 2, p. 205-212, 2007.
- 111 KELLER, N.; NAGULESWARAN, A.; CANNAS, A.; et al. Identification of a *Neospora caninum* microneme protein (NcMIC1) which interacts with sulfated host cell surface glycosaminoglycans. **Infection and Immunity**, v. 70, n. 6, p. 3187-3198, 2002.
- 112 PEREIRA, L. M.; CANDIDO-SILVA, J. A.; DE VRIES, E.; et al. A new thrombospondin-related anonymous protein homologue in *Neospora caninum* (NcMIC2-like1). **Parasitology**, v. 138, n. 3, p. 287-297, 2011.
- 113 NAGULESWARAN, A.; CANNAS, A.; KELLER, N.; et al. *Neospora caninum* microneme protein NcMIC3: Secretion, subcellular localization, and functional involvement in host cell interaction. **Infection and Immunity**, v. 69, n. 10, p. 6483-6494, 2001.
- 114 ELLIS, J.; MILLER, C.; QUINN, H.; et al. Evaluation of recombinant proteins of *Neospora caninum* as vaccine candidates (in a mouse model). **Vaccine**, v. 26, n. 47, p. 5989-5996, 2008.
- 115 KELLER, N.; RIESEN, M.; NAGULESWARAN, A.; et al. Identification and characterization of a *Neospora caninum* microneme-associated protein (NcMIC4) that exhibits unique lactose-binding properties. **Infection and Immunity**, v. 72, n. 8, p. 4791-4800, 2004.
- 116 LOVETT, J. L.; HOWE, D. K.; SIBLEY, L. D. Molecular characterization of a thrombospondin-related anonymous protein homologue in *Neospora caninum*. **Molecular and Biochemical Parasitology**, v. 107, n. 1, p. 33-43, 2000.
- 117 JEWETT, T. J.; SIBLEY, L. D. The toxoplasma proteins MIC2 and M2AP form a hexameric complex necessary for intracellular survival. **Journal of Biological Chemistry**, v. 279, n. 10, p. 9362-9369, 2004.
- 118 KESSLER, H.; HERM-GÖTZ, A.; HEGGE, S.; et al. Microneme protein 8 - a new essential invasion factor in *Toxoplasma gondii*. **Journal of Cell Science**, v. 121, n. 7, p. 947-956, 2008.
- 119 LI, R. W.; TUO, W. *Neospora caninum*: Comparative gene expression profiling of *Neospora caninum* wild type and a temperature sensitive clone. **Experimental Parasitology**, v. 129, n. 4, p. 346-354, 2011.
- 120 KAFSACK, B. F. C.; PENA, J. D.; COPPENS, I.; et al. Rapid membrane disruption by a perforin-like protein facilitates parasite exit from host cells. **Science**, v. 323, n. 5913, p. 530-533, 2009.
- 121 KAWASE, O.; PENA, J. D.; COPPENS, I.; et al. Characterization of a novel thrombospondin-related protein in *Toxoplasma gondii*. **Parasitology International**, v. 59, n. 2, p. 211-216, 2010.

- 122 TYLER, J. S.; BOOTHROYD, J. C. The C-terminus of *Toxoplasma* RON2 provides the crucial link between AMA1 and the host-associated invasion complex. **Plos Pathogens**, v. 7, n. 2, 2011.
- 123 STRAUB, K. W.; CHENG, S. J.; SOHN, C. S.; et al. Novel components of the Apicomplexan moving junction reveal conserved and coccidia-restricted elements. **Cellular Microbiology**, v. 11, n. 4, p. 590-603, 2009.
- 124 HAJAGOS, B. E.; TURETZKY, J. M.; PENG, E. D.; et al. Molecular dissection of novel trafficking and processing of the *Toxoplasma gondii* rhoptry metalloprotease toxolysin-1. **Traffic**, v. 13, n. 2, p. 292-304, 2012.
- 125 MILLER, S. A.; THATHY, V.; AJIOKA, J. W.; et al. TgSUB2 is a *Toxoplasma gondii* rhoptry organelle processing proteinase. **Molecular Microbiology**, v. 49, n. 4, p. 883-894, 2003.
- 126 MORRIS, M. T.; CARRUTHERS, V. B. Identification and partial characterization of a second Kazal inhibitor in *Toxoplasma gondii*. **Molecular and Biochemical Parasitology**, v. 128, n. 1, p. 119-122, 2003.
- 127 PSZENNY, V.; LEDESMA, B. E.; MATRAJT, M.; et al. Subcellular localization and post-secretory targeting of TgPI, a serine proteinase inhibitor from *Toxoplasma gondii*. **Molecular and Biochemical Parasitology**, v. 121, n. 2, p. 283-286, 2002.
- 128 HALDORSON, G. J.; MATHISON, B. A.; WENBERG, K.; et al. Immunization with native surface protein NcSRS2 induces a Th2 immune response and reduces congenital *Neospora caninum* transmission in mice. **International Journal for Parasitology**, v. 35, n. 13, p. 1407-1415, 2005.
- 129 BASZLER, T. V.; SHKAP, V.; MWANGI, W.; et al. Bovine immune response to inoculation with *Neospora caninum* surface antigen SRS2 lipopeptides mimics immune response to infection with live parasites. **Clinical and Vaccine Immunology**, v. 15, n. 4, p. 659-667, 2008.
- 130 WU, X.-N.; LIN, J.; LIN, X.; et al. Multicomponent DNA vaccine-encoding *Toxoplasma gondii* GRA1 and SAG1 primes: anti-*Toxoplasma* immune response in mice. **Parasitology Research**, v. 111, n. 5, p. 2001-2009, 2012.
- 131 LEE, Y. H.; Shin, D. W.; Lee, J. H.; et al. Vaccination against murine toxoplasmosis using recombinant *Toxoplasma gondii* SAG3 antigen alone or in combination with Quil A. **Yonsei Med J.**, v. 48, n. 3, p. 396-404, 2007.
- 132 MINEO, T. W. P.; OLIVEIRA, C. J.; SILVA, D. A.; et al. *Neospora caninum* excreted/secreted antigens trigger CC-chemokine receptor 5-dependent cell migration. **International Journal for Parasitology**, v. 40, n. 7, p. 797-805, 2010.
- 133 IBRAHIM, H. M.; XUAN, X.; NISHIKAWA, Y. *Toxoplasma gondii* Cyclophilin 18 Regulates the Proliferation and Migration of Murine Macrophages and Spleen Cells. **Clinical and Vaccine Immunology**, v. 17, n. 9, p. 1322-1329, 2010.
- 134 SHANMUGASUNDRAM, A.; GONZALEZ-GALARZA, F. F.; WASTLING, J. M.; et al. Library of Apicomplexan Metabolic Pathways: a manually curated database for metabolic pathways of apicomplexan parasites. **Nucleic Acids Research**, v. 41, n. D1, p. D706-D713, 2013.

- 135 LOURIDO, S.; SHUMAN, J.; ZHANG, C.; et al. Calcium-dependent protein kinase 1 is an essential regulator of exocytosis in *Toxoplasma*. **Nature**, v. 465, n. 7296, p. 359-U118, 2010.
- 136 PIERROT, C.; FRÉVILLE, A.; OLIVIER, C.; et al. Inhibition of Protein-protein Interactions in *Plasmodium falciparum*: Future Drug Targets. **Current Pharmaceutical Design**, v. 18, n. 24, p. 3522-3530, 2012.
- 137 CASSOLA, A.; FRASCH, A. C. An RNA recognition motif mediates the nucleocytoplasmic transport of a Trypanosome RNA-binding protein. **Journal of Biological Chemistry**, v. 284, n. 50, p. 35015-35028, 2009.
- 138 ROGELJ, B.; GONDIN, K. S.; SHAW, C. E.; et al. The Functions of Glycine-Rich Regions in TDP-43, FUS And Related Rna-Binding Proteins. In: LORKOVIĆ, Z. J. (Ed.). **RNA Binding Proteins**: Landes Bioscience, 2012. cap. *Chapter Category: RNA*.
- 139 KIM, W.; KIM, W.; SILBY, M. W.; et al. Proteomic detection of non-annotated protein-coding genes in *Pseudomonas fluorescens* Pf0-1. **Plos One**, v. 4, n. 12, 2009.

Neospora caninum: estudo do secretoma e caracterização molecular de três proteínas com domínios Apple

Apêndice 1

Apêndice 1 Relação das 615 proteínas identificadas por LC-MS no ESA de taquizoítas de *N. caninum*. Organização em ordem decrescente de abundância relativa no ESA (número de espectros dividido por peso molecular da proteína). Informação nas colunas: descrição de acordo com ToxoDB ou literatura; peso molecular predito (PM); probabilidade de identificação da proteína; número de espectros; abundância relativa; e números de peptídeos únicos. h.p., hypothetical protein; rel, related; put, putative.

	ToxoDB ID (NCLIV)	Descrição (ToxoDB ou literatura)	PM	proba- bilidade	número de espectros	abundância relativa	número de peptídeos únicos
1	021050	SUB1 - subtilisin protease	93	100%	3134	33.70	29
2	028680	AMA1 - apical membrane antigen 1	63	100%	2087	33.13	53
3	002390	nucleoside diphosphate kinase, rel	18	100%	429	23.83	10
4	043270	MIC1 - microneme protein 1	49	100%	1127	23.00	34
5	015920	histone H4, rel	11	100%	224	20.36	10
6	050370	h.p. (<i>fructose-1,6-bisphosphate aldolase</i>)	33	100%	607	18.39	22
7	033250	SRS2 (SRS29C)	42	100%	599	14.26	18
8	011830	h.p.	22	100%	311	14.14	20
9	032240	catalase	59	100%	825	13.98	53
10	042000	ubiquitin CEP52 fusion protein, put.	15	100%	199	13.27	9
11	033950	HSP70 - heat shock protein 70	73	100%	836	11.45	51
12	011410	protein disulfide isomerase	53	100%	604	11.40	38
13	011700	h.p. (<i>ROP15B</i>)	35	100%	352	10.06	22
14	049520	h.p., conserved	39	100%	365	9.36	25
15	060860	cytochrome c, rel	12	100%	112	9.33	13
16	024820	14-3-3 protein homolog	31	100%	285	9.19	23
17	006170	phosphoglycerate mutase, rel	28	100%	193	6.89	19
18	011120	malate dehydrogenase, rel	34	100%	221	6.50	19
19	031550	predicted rhoptry protein kinase	48	100%	310	6.46	20
20	019770	h.p.	59	100%	354	6.00	27
21	033690	MIC2-like1 - microneme protein	80	100%	448	5.60	28
22	053860	MGC84926 protein	57	100%	317	5.56	21
23	020720	MIC11 - microneme protein 11	22	100%	116	5.27	11
24	051000	ethylene-inducible protein hever, rel	33	100%	160	4.85	16
25	029690	h.p., conserved	19	100%	91	4.79	5
26	001350	EGF-like domain-containing protein, put.	124	100%	561	4.52	58
27	051970	M2AP - MIC2-associated protein	25	100%	112	4.48	6
28	065440	h.p. (<i>calmodulin</i>)	17	100%	72	4.24	9
29	055490	heat shock protein 70 (precursor), rel	73	100%	283	3.88	32
30	013260	h.p., conserved	20	100%	76	3.80	8
31	007260	p97 protein, put	98	100%	366	3.73	57
32	037500	enolase 2	52	100%	194	3.73	18
33	001970	ROP7 - rhoptry protein 7	67	100%	247	3.69	22
34	004790	CyP - 18 kDa cyclophilin	19	100%	69	3.63	11
35	004190	thioredoxin, put	48	100%	172	3.58	18
36	032110	h.p., conserved	49	100%	170	3.47	22
37	061760	MIC6 - microneme protein 6	34	100%	115	3.38	10
38	011730	ROP26 - rhoptry kinase family protein 26	60	100%	202	3.37	24
39	010600	MIC3 - microneme protein 3	38	100%	127	3.34	19
40	043670	high mobility group protein, rel	11	100%	36	3.27	10
41	066250	MIC10 - microneme protein 10	23	100%	72	3.13	17
42	058810	superoxide dismutase	23	100%	69	3.00	5
43	026140	histone H2A, rel	13	100%	39	3.00	3
44	040860	tryptophanyl-tRNA synthetase, rel	68	100%	197	2.90	28
45	002940	MIC4 - microneme protein 4	64	100%	185	2.89	18
46	011270	h.p. (<i>phosphoglycerate kinase</i>)	45	100%	130	2.89	25
47	054720	h.p. (<i>glycogen phosphorylase</i>)	107	100%	297	2.78	48

48	062630	thioredoxin reductase,rel.	22	100%	61	2.77	10
49	000430	h.p., conserved	43	100%	119	2.77	17
50	040880	HSP90, rel	82	100%	224	2.73	30
51	062770	MIC8 - microneme protein 8	74	100%	202	2.73	19
52	018420	ROP9 (p36 protein)	37	100%	101	2.73	16
53	019970	peptidyl-prolyl cis-trans isomerase A	19	100%	51	2.68	8
54	030420	Rcn2-prov protein, rel	39	100%	104	2.67	14
55	010720	SRS domain-containing protein	41	100%	109	2.66	18
56	024110	h.p., conserved	23	100%	60	2.61	7
57	066630	GRA9 - dense granule protein 9	35	100%	89	2.54	9
58	015380	h.p., conserved	38	100%	96	2.53	14
59	019110	HSP90-like protein, rel	97	100%	245	2.53	42
60	052190	h.p., conserved	48	100%	117	2.44	18
61	063150	serpin peptidase inhibitor, rel	46	100%	112	2.43	26
62	042660	probable cytosol aminopeptidase, rel.	60	100%	146	2.43	26
63	028540	h.p., conserved	30	100%	72	2.40	12
64	032270	h.p., conserved	35	100%	83	2.37	17
65	043760	h.p., conserved	17	100%	39	2.29	9
66	063590	glutathione reductase, rel	52	100%	118	2.27	16
67	070280	h.p. (CAMK/CDPK)	54	100%	119	2.20	21
68	060420	h.p. (threonyl-tRNA synthetase)	88	100%	193	2.19	27
69	015460	h.p. (PKC-cousin of thioredoxin)	26	100%	57	2.19	8
70	002770	MORN repeat-containing protein, put.	42	100%	91	2.17	10
71	063740	h.p., conserved	50	100%	108	2.16	17
72	010740	kelch motif domain protein, put.	44	100%	94	2.14	15
73	012920	ROP40 - rhoptry kinase family protein 40	43	100%	90	2.09	16
74	065090	h.p., conserved	75	100%	156	2.08	18
75	064950	h.p.	55	100%	113	2.05	18
76	059600	KH domain-containing protein, put	101	100%	207	2.05	31
77	004380	cathepsin L, rel	46	100%	94	2.04	11
78	014330	prolyl oligopeptidase (precursor), rel	83	100%	169	2.04	23
79	030290	transaldolase, put	41	100%	83	2.02	20
80	004140	h.p.	18	100%	35	1.94	8
81	046900	h.p. (triosephosphate isomerase)	28	100%	54	1.93	14
82	024230	Zgc:123215, rel	11	100%	21	1.91	7
83	062230	glutaminyt-tRNA synthetase, rel	57	100%	105	1.84	12
84	068400	NTPase	69	100%	125	1.81	17
85	041100	novel protein, rel	32	100%	52	1.63	7
86	044380	ubiquitin carrier protein, rel	18	100%	29	1.61	8
87	006860	h.p. (aspartyl aminopeptidase)	54	100%	86	1.59	21
88	045650	GRA2 - dense granule protein 2	22	100%	35	1.59	9
89	046800	AGC kinase, put	39	100%	62	1.59	14
90	062220	glutaminyt-tRNA synthetase, rel	46	100%	73	1.59	15
91	005010	h.p., conserved	62	100%	98	1.58	19
92	033040	Rbm28 protein, rel (RNA binding)	18	100%	28	1.56	5
93	022970	MIC2 - microneme protein 2	83	100%	126	1.52	7
94	012830	MORN repeat-containing protein, put	48	100%	72	1.50	14
95	067180	glucose-6-P-1-dehydrogenase,rel	82	100%	120	1.46	21
96	004270	protein kinase, put	124	100%	180	1.45	27
97	033230	SAG1 - surface antigen 1 (SRS29B)	33	100%	47	1.42	8
98	042400	macrophage migration inhibitory factor	12	100%	17	1.42	5
99	066760	translationally-controlled tumor protein	19	100%	26	1.37	8
100	050390	pyrroline-5-carboxylate reductase, rel	30	100%	41	1.37	8
101	007770	rhoptry kinase family protein (BPK1)	42	100%	57	1.36	14
102	011980	calmodulin-like protein (CAMK/CDPK)	57	100%	77	1.35	16
103	008680	h.p., conserved	73	100%	96	1.32	17
104	036830	h.p., conserved	98	100%	128	1.31	30
105	020220	elongation factor 2, put	93	100%	118	1.27	13
106	051340	toxofilin, put	32	100%	40	1.25	9
107	037400	la domain-containing protein, put	61	100%	75	1.23	24
108	050620	lysine decarboxylase protein, put	44	100%	54	1.23	14

109	039820	hexokinase, put	52	100%	63	1.21	15
110	019480	proteasome protein, put.	30	100%	36	1.20	6
111	049030	h.p. (<i>eIF4E</i>)	26	100%	31	1.19	5
112	058830	superoxide dismutase, rel	32	100%	38	1.19	7
113	046540	oligoendopeptidase F, put	71	100%	83	1.17	17
114	017370	CAMP-dependent kinase , put	43	100%	50	1.16	16
115	052350	h.p., conserved	28	100%	32	1.14	9
116	056590	phosphatidylinositol4phosphate 5kinase	51	100%	57	1.12	15
117	055230	serine/threonine protein phosphatase	54	100%	60	1.11	15
118	046170	heat shock protein 70, ER lumen, rel	104	100%	114	1.10	27
119	001070	histone H2B, rel (<i>histone H2Ba</i>)	13	100%	14	1.08	7
120	066920	h.p. (<i>tyrosyl-tRNA synthetase, put</i>)	48	100%	50	1.04	16
121	018560	ras-GTPase-activating binding protein	88	100%	90	1.02	16
122	013910	Pcd4-prov protein, rel	58	100%	58	1.00	20
123	040540	h.p. (<i>nuclear ribonucleoprotein A3</i>)	45	100%	45	1.00	10
124	026150	histone H3, rel	15	100%	15	1.00	4
125	041930	RAB11 - Rab-GTPase	23	100%	22	0.96	8
126	024740	myosin, rel (<i>myosin D</i>)	85	100%	81	0.95	15
127	047860	h.p. (<i>polyadenylate-binding protein</i>)	82	100%	78	0.95	12
128	032950	triphosphate nucleotidohydrolase, put	20	100%	19	0.95	5
129	005560	DG32 - dense-granule antigen 32	28	100%	26	0.93	6
130	008410	gamma-aminobutyric acid receptor, rel	14	100%	13	0.93	5
131	065470	h.p. (<i>thioredoxin</i>)	51	100%	47	0.92	11
132	061170	h.p. (<i>heat shock protein</i>)	93	100%	85	0.91	24
133	045870	GRA3 - dense granule protein 3	23	100%	21	0.91	4
134	011210	transketolase, rel	108	100%	98	0.91	21
135	049900	h.p. (<i>myosin A</i>)	93	100%	83	0.89	22
136	020990	PLP1 - perforin-like protein 1	115	100%	101	0.88	23
137	028170	ROP20 - rhopty kinase protein 20	63	100%	55	0.87	14
138	060660	SAG3 - surface antigen 3 (SRS57, P43)	39	100%	34	0.87	10
139	054520	MORN1	41	100%	35	0.85	16
140	000860	SPATR, rel	58	100%	49	0.84	13
141	051840	h.p. (<i>mitochondrial glycoprotein</i>)	32	100%	27	0.84	8
142	003580	h.p., conserved	32	100%	27	0.84	8
143	046260	ACN/IRP	99	100%	81	0.82	22
144	032330	LDH1-lactate dehydrogenase	36	100%	29	0.81	10
145	031860	serine-threonine phosphatase 2C, put	36	100%	29	0.81	10
146	000280	RanBP1 domain containing protein, rel	25	100%	20	0.80	8
147	028830	glucose-6-phosphate isomerase, rel	126	100%	100	0.79	14
148	018530	h.p., conserved	101	100%	80	0.79	16
149	045430	DNA-binding protein HU, put	24	100%	19	0.79	7
150	013780	h.p. (<i>proteasome subunit alpha type 2</i>)	26	100%	20	0.77	6
151	068920	SRS domain-containing protein	41	100%	31	0.76	5
152	035950	h.p. (<i>small nuclear ribonucleoprotein</i>)	12	100%	9	0.75	5
153	000840	thioredoxin h, rel	12	100%	9	0.75	6
154	066190	caltractin, put (<i>centrin2</i>)	19	100%	14	0.74	5
155	036250	asparaginyl-tRNA synthetase, rel	60	100%	44	0.73	14
156	041880	MIC16 - microneme protein 16	41	100%	30	0.73	9
157	064580	h.p. (<i>elongation factor 1-gamma</i>)	44	100%	32	0.73	9
158	022550	GK15875, rel	51	100%	37	0.73	8
159	024380	h.p., conserved	61	100%	44	0.72	10
160	050590	GL11864, rel	53	100%	38	0.72	13
161	026070	armadillo/beta-catenin- protein, put	31	100%	22	0.71	6
162	048240	h.p. (<i>aminopeptidase N</i>)	106	100%	74	0.70	22
163	056480	h.p. (<i>isocitrate dehydrogenase</i>)	59	100%	41	0.69	11
164	036130	CBR-RSP-4 protein, rel	22	100%	15	0.68	4
165	005900	translation initiation factor, rel	89	100%	59	0.66	19
166	009640	choline kinase, put	62	100%	41	0.66	14
167	003570	h.p., conserved	52	100%	34	0.65	15
168	060810	h.p., conserved	102	100%	66	0.65	22
169	047660	h.p. (<i>adenosylhomocysteinase</i>)	52	100%	33	0.63	10

170	034160	glutamine synthetase (EC 6.3.1.2), rel	51	100%	32	0.63	9
171	019960	lysyl-tRNA synthetase, put	67	100%	42	0.63	14
172	025770	ankyrin repeat-containing protein, put	24	100%	15	0.63	6
173	054540	RAB5C, rel	21	100%	13	0.62	8
174	038680	h.p.	26	100%	16	0.62	4
175	038760	mago nashi protein, put	18	100%	11	0.61	7
176	004340	FABZ	25	100%	15	0.60	6
177	055170	DnaJ domain containing protein, rel	48	100%	28	0.58	12
178	024840	h.p. (<i>CMGC kinase, CK2 family</i>)	52	100%	30	0.58	10
179	052880	GRA6 - granule antigen protein 6	21	100%	12	0.57	1
180	046100	glycogen debranching enzyme, put	205	100%	117	0.57	36
181	044270	h.p. (<i>CAMK/CDPK</i>)	79	100%	45	0.57	13
182	003050	myosin heavy chain, put	125	100%	71	0.57	21
183	000940	glucose-6-phosphate dehydrogenase	55	100%	31	0.56	6
184	025240	Gbp1p protein, put	32	100%	18	0.56	9
185	027230	dihydropteroate synthase, rel	73	100%	40	0.55	9
186	000740	CPL1 - Chitin-like protein 1	74	100%	40	0.54	9
187	009030	metallophosphoesterase, rel	37	100%	20	0.54	7
188	007330	(<i>Hsc70/Hsp90-organizing protein</i>)	63	100%	34	0.54	15
189	041940	glyceraldehyde 3-P dehydrogenase, rel	36	100%	19	0.53	4
190	005440	mog1 protein, put	19	100%	10	0.53	5
191	067490	protein phosphatase 2C, put	37	100%	19	0.51	7
192	029340	EGF-like domain-containing protein, put	155	100%	78	0.50	18
193	048570	GAP45 - gliding-associated protein	28	100%	14	0.50	7
194	056140	h.p., conserved	296	100%	140	0.47	46
195	012130	eukaryotic translation initiation factor 3	60	100%	28	0.47	9
196	012220	h.p., conserved	28	100%	13	0.46	5
197	010730	SRS domain-containing protein	41	100%	19	0.46	5
198	015010	h.p., conserved	70	100%	32	0.46	10
199	041740	h.p., conserved	103	100%	47	0.46	8
200	065590	NAD-specific glutamate dehydrogenase	127	100%	57	0.45	22
201	011900	arginyl-tRNA synthetase, rel	67	100%	30	0.45	13
202	035860	ROP2L10 (<i>ROP32</i>)	65	100%	29	0.45	11
203	042820	clathrin heavy chain, put	195	100%	86	0.44	29
204	035310	inhibitor-1 of protein phosphatase 2A	32	100%	14	0.44	7
205	043300	nucleolar phosphoprotein nucleolin, put	74	100%	32	0.43	8
206	036200	cyclophilin, put	21	100%	9	0.43	5
207	017510	RPS29 - 40S ribosomal protein S29	7	98%	3	0.43	2
208	060470	h.p. (<i>cysteinyI-tRNA synthetase</i>)	75	100%	32	0.43	7
209	018620	Os08g0464000 protein, rel	19	100%	8	0.42	5
210	034130	h.p. (<i>proteasome PCI protein</i>)	67	100%	27	0.40	9
211	054750	PP1 - protein phosphatase type 1	35	100%	14	0.40	6
212	010200	RPL5 - ribosomal protein L5	35	100%	14	0.40	7
213	036400	GRA1 - dense granule protein 1	20	98%	8	0.40	2
214	010120	h.p., conserved	10	98%	4	0.40	2
215	016120	proteasome subunit alpha 4, subunit, put	28	100%	11	0.39	6
216	060760	prolyl-tRNA synthetase, put	74	100%	29	0.39	5
217	016150	serine-pyruvate transaminase, rel	44	100%	17	0.39	9
218	012510	ADF - actin depolymerizing factor	13	100%	5	0.38	3
219	069960	RNA binding motif- protein, put	73	100%	28	0.38	14
220	032430	h.p., conserved	55	100%	21	0.38	14
221	001670	elongation factor 1-alpha, rel	49	100%	18	0.37	7
222	017470	proteasome subunit alpha type, rel	33	100%	12	0.36	7
223	039480	h.p. (<i>ubiquitin-like protein SMT3</i>)	11	100%	4	0.36	3
224	004300	dynammin-like protein, put	97	100%	35	0.36	11
225	053840	ROP1 - rhoptry protein 1	39	100%	14	0.36	2
226	046890	h.p., conserved	39	100%	14	0.36	5
227	023090	h.p. (<i>protein phosphatase 2C</i>)	42	100%	15	0.36	6
228	038830	Sm prot rel	14	98%	5	0.36	2
229	022140	GA11385, rel (<i>RNA-binding protein, put</i>)	59	100%	21	0.36	7
230	001260	RNA binding protein, put	17	100%	6	0.35	3

231	002590	phthalate dioxygenase reductase, rel	74	100%	26	0.35	9
232	007580	h.p. (<i>alpha-glucosidase II</i>)	180	100%	63	0.35	23
233	064760	An16c0190	46	100%	16	0.35	7
234	041790	h.p., conserved	98	100%	34	0.35	15
235	042650	GG11844, rel	49	100%	17	0.35	7
236	034190	eukaryotic translation initiat. factor 3, put	75	100%	26	0.35	7
237	021100	ROP12 - rhoptry protein 12	26	100%	9	0.35	4
238	061600	h.p. (<i>phosphoserine aminotransferase</i>)	51	100%	17	0.33	8
239	020310	fructose-1,6-bisphosphatase, put	42	100%	14	0.33	4
240	028230	proteasome subunit alpha type-7, rel	27	100%	9	0.33	5
241	015440	h.p. (<i>staphylococcal nuclease protein</i>)	103	100%	34	0.33	11
242	044230	TLN4 - toxolysin 4	237	100%	78	0.33	14
243	020070	h.p., conserved	52	100%	17	0.33	6
244	029420	MLC1 - myosin light chain 1	25	100%	8	0.32	4
245	003160	nicotinate phosphoribosyltransferase, rel	66	100%	21	0.32	7
246	066840	h.p. (<i>heat shock protein 70</i>)	79	100%	25	0.32	12
247	015410	RPL23A - ribosomal protein L23A	19	100%	6	0.32	3
248	051590	GTP binding protein, put	45	100%	14	0.31	6
249	057700	RPL7 - ribosomal protein L7	29	100%	9	0.31	5
250	070010	RON8 - rhoptry neck protein 8	328	100%	101	0.31	34
251	005500	h.p., conserved	127	100%	38	0.30	15
252	017440	predicted ROPK (<i>ROP19</i>)	61	100%	18	0.30	9
253	000710	h.p. (<i>SCP-like extracellular protein</i>)	24	100%	7	0.29	4
254	003100	PI2 - serine proteinase inhibitor PI2	28	100%	8	0.29	4
255	063690	h.p., conserved	28	100%	8	0.29	4
256	020140	h.p.	21	98%	6	0.29	2
257	064700	RPL18 - ribosomal protein L18	21	100%	6	0.29	4
258	050070	fructose-1,6-bisphosphatase, rel	14	98%	4	0.29	2
259	046970	h.p., conserved	117	100%	33	0.28	13
260	021510	h.p., conserved	25	100%	7	0.28	5
261	012890	h.p.	782	100%	218	0.28	55
262	000130	h.p. (<i>NADP-glutamate dehydrogenase</i>)	54	100%	15	0.28	8
263	030340	cytochrome c, put	18	95%	5	0.28	1
264	015550	26S proteasome non-ATPase, put	47	100%	13	0.28	6
265	009510	DEHA2E04510p, rel	11	100%	3	0.27	3
266	005360	splicing factor 3a protein, put	59	100%	16	0.27	10
267	038330	adenylosuccinate synthetase, put	52	100%	14	0.27	9
268	030050	RON4 - rhoptry neck protein 4	97	100%	26	0.27	7
269	036970	calmodulin, put (<i>ELC1</i>)	15	100%	4	0.27	4
270	001300	calmodulin, put	15	100%	4	0.27	3
271	061830	60S acidic ribosomal protein P0	34	100%	9	0.26	4
272	006090	h.p., conserved	19	100%	5	0.26	3
273	063750	calmodulin, rel (<i>centrin 1</i>)	19	100%	5	0.26	4
274	057680	h.p. (<i>small nuclear ribonucleoprotein</i>)	19	98%	5	0.26	2
275	045550	RNA recognition protein, put	66	100%	17	0.26	12
276	042500	ubiquitin-activating enzyme E1, rel	121	100%	31	0.26	12
277	003890	h.p., conserved	43	100%	11	0.26	5
278	041650	methionine-tRNA ligase, rel	107	100%	27	0.25	11
279	064370	lactoylglutathione lyase, rel	36	100%	9	0.25	5
280	065170	snRNP protein Lsm5, put	12	98%	3	0.25	2
281	005890	h.p., conserved	12	100%	3	0.25	2
282	023550	hydroxyacylglutathione hydrolase, put	29	98%	7	0.24	2
283	045400	h.p. (<i>splicing factor 3A subunit 2</i>)	25	100%	6	0.24	4
284	066080	h.p., conserved	25	100%	6	0.24	3
285	056110	small heat shock protein 21, put	21	100%	5	0.24	3
286	016840	AFL091Wp, rel (<i>kinase CK1/CK1</i>)	38	100%	9	0.24	4
287	018300	HIT family protein, rel	17	100%	4	0.24	4
288	055610	h.p., conserved	17	100%	4	0.24	3
289	034460	h.p. (<i>cell division protein 48</i>)	65	100%	15	0.23	5
290	062570	An13c0020, rel (<i>proteasome PCI</i>)	52	100%	12	0.23	5
291	009670	cold-shock protein, DNA-binding, rel	13	98%	3	0.23	2

292	011690	ROP15 - rhoptry protein 15	31	98%	7	0.23	2
293	039940	h.p. (<i>protein kinase family CMGC/GSK</i>)	45	100%	10	0.22	5
294	004920	h.p.	27	100%	6	0.22	4
295	000610	profilin family protein, put	18	100%	4	0.22	3
296	010240	h.p., conserved	97	100%	21	0.22	10
297	053870	h.p. (<i>small nuclear ribonucleoprotein</i>)	14	98%	3	0.21	2
298	049600	RPS15A - ribosomal protein RPS15A	19	100%	4	0.21	3
299	035540	h.p. (<i>folylpolyglutamate synthase</i>)	62	100%	13	0.21	3
300	039960	h.p. (<i>heat shock protein 40</i>)	34	100%	7	0.21	4
301	006040	coproporphyrinogen oxidase, put	39	100%	8	0.21	3
302	038400	methionine aminopeptidase, rel	44	100%	9	0.20	4
303	055850	ROP13 - rhoptry protein 13	45	100%	9	0.20	4
304	011840	h.p. (<i>RNA-binding protein</i>)	31	100%	6	0.19	3
305	025480	kelch motif containing protein, put	88	100%	17	0.19	9
306	053930	small nuclear ribonucleoprotein U1A, put	42	100%	8	0.19	5
307	034470	h.p.	21	100%	4	0.19	3
308	043950	h.p. (<i>methionine aminopeptidase</i>)	53	100%	10	0.19	7
309	038290	h.p., conserved	48	100%	9	0.19	3
310	031510	RPS25 - ribosomal protein S25	16	100%	3	0.19	3
311	018120	h.p., conserved	75	100%	14	0.19	9
312	006180	duplicated carbonic anhydrase, put	59	100%	11	0.19	4
313	053590	TBL1- transducin-beta-like 1	86	100%	16	0.19	8
314	046550	elongation factor 1-beta, rel	27	100%	5	0.19	3
315	044410	ROP35	65	100%	12	0.18	7
316	026180	h.p. (<i>26S proteasome non-ATPase</i>)	55	100%	10	0.18	3
317	067060	h.p. (<i>U6 snRNA associated LSm4</i>)	11	98%	2	0.18	2
318	070100	GF20490, rel (<i>exoribonuclease, put</i>)	139	100%	25	0.18	12
319	018830	h.p., conserved	78	100%	14	0.18	4
320	045260	h.p. (<i>protein phosphatase</i>)	67	100%	12	0.18	5
321	051450	centromere/microtubule bind. protein, put	56	100%	10	0.18	5
322	065910	h.p., conserved	34	100%	6	0.18	4
323	038410	predict. hydrolases / acyltransferases, rel	34	100%	6	0.18	4
324	036500	TXNL4A/DIM1	17	100%	3	0.18	2
325	008850	h.p., conserved	222	100%	39	0.18	14
326	037250	transport protein particle Bet3, put	23	100%	4	0.17	3
327	037590	h.p., conserved	23	100%	4	0.17	2
328	042480	h.p., conserved	23	100%	4	0.17	3
329	023050	eukaryotic translation initiat. factor 3, put	121	100%	21	0.17	10
330	013180	GM04207p, rel	81	100%	14	0.17	7
331	036010	h.p.	87	100%	15	0.17	11
332	049270	h.p., conserved	77	100%	13	0.17	5
333	048590	h.p., conserved (<i>RON3</i>)	220	100%	37	0.17	15
334	019040	peptidase M16, rel	125	100%	21	0.17	13
335	064880	h.p., conserved	132	100%	22	0.17	4
336	033100	prolidase, put	66	100%	11	0.17	9
337	002850	h.p.	24	100%	4	0.17	4
338	049100	RPS19 - ribosomal protein RPS19	18	98%	3	0.17	2
339	015870	SRS23	31	100%	5	0.16	4
340	046050	h.p. (<i>phosphofructokinase</i>)	131	100%	21	0.16	5
341	054140	adenylyl cyclase associated protein, put	25	98%	4	0.16	2
342	048880	proteasome subunit beta type-7, rel	25	100%	4	0.16	4
343	053640	peroxidoxin 2, put	25	98%	4	0.16	2
344	068872	SRS domain-containing protein	38	100%	6	0.16	4
345	018710	h.p. (<i>nucleosome assembly protein</i>)	51	100%	8	0.16	4
346	026350	h.p., conserved	32	100%	5	0.16	5
347	059620	h.p., conserved	45	100%	7	0.16	5
348	065450	RPL35A - ribosomal protein RPL35A	13	98%	2	0.15	2
349	036640	h.p., conserved	98	100%	15	0.15	10
350	013460	h.p. h.p., conserved	46	100%	7	0.15	5
351	039400	h.p. (<i>acetyl-coenzyme A synthetase</i>)	79	100%	12	0.15	6
352	008170	h.p.	40	100%	6	0.15	5

353	056830	60S ribosomal protein L7a, put	20	98%	3	0.15	2
354	044340	h.p., conserved	217	100%	32	0.15	13
355	010320	dihydrolipoamide transacylase, put	69	100%	10	0.14	5
356	058000	alanine dehydrogenase, put	42	100%	6	0.14	4
357	052740	SRS domain-containing protein	35	100%	5	0.14	3
358	070070	calcyclin binding protein, put	28	100%	4	0.14	3
359	038170	hypoxanthine phosphoribosyltransferase, rel	28	98%	4	0.14	2
360	024140	h.p.	21	100%	3	0.14	3
361	038710	Lin-28 homolog B, rel	14	98%	2	0.14	2
362	013360	plectin, put	260	100%	37	0.14	21
363	028140	h.p., conserved	57	100%	8	0.14	5
364	015200	pyruvate kinase (EC 2.7.1.40), rel	58	100%	8	0.14	5
365	014430	h.p., conserved	80	100%	11	0.14	5
366	048580	h.p. (<i>RNA binding protein</i>)	44	100%	6	0.14	3
367	017500	RPS5 - ribosomal protein RPS5	22	100%	3	0.14	2
368	029040	nucleosome assembly protein, put	96	100%	13	0.14	5
369	045800	RPL3 - 60S ribosomal protein L3, rel	45	100%	6	0.13	4
370	063760	h.p. (<i>ubiquitin carboxyl- isozyme L5</i>)	45	100%	6	0.13	3
371	055710	RPL23 - ribosomal protein RPL23	15	98%	2	0.13	2
372	033850	h.p. (<i>chromatin remodelling SNF2L</i>)	181	100%	24	0.13	10
373	007800	RON4L1	212	100%	28	0.13	11
374	061050	h.p. (<i>alanyl-tRNA synthetase</i>)	114	100%	15	0.13	5
375	038110	MIC17B - microneme protein 17B	38	100%	5	0.13	4
376	048460	thioredoxin, put	69	100%	9	0.13	7
377	050210	KH domain-containing protein, put	69	100%	9	0.13	5
378	060160	Tgd057, rel	23	100%	3	0.13	3
379	022080	h.p. (<i>C2 domain-containing protein</i>)	23	100%	3	0.13	3
380	022540	h.p., conserved	62	100%	8	0.13	4
381	043400	proteasome subunit p58, rel	62	100%	8	0.13	6
382	037150	h.p. (<i>26S proteasome reg subunit</i>)	39	100%	5	0.13	3
383	042270	h.p. (<i>phenylalanyl tRNA synthetase</i>)	328	100%	41	0.13	13
384	030120	importin beta-3 subunit, put	128	100%	16	0.13	10
385	051810	phosphoglycerate dehydrogenase, put	40	100%	5	0.13	3
386	027080	h.p., conserved	32	100%	4	0.13	3
387	021110	h.p., conserved	195	100%	24	0.12	6
388	039490	h.p. (<i>lysine decarboxylase protein</i>)	41	100%	5	0.12	4
389	034381	SRS domain-containing protein	41	98%	5	0.12	2
390	055260	AGC kinase TgPKG1, put	115	100%	14	0.12	7
391	065800	h.p. (<i>apolipoprotein A-I binding protein</i>)	25	96%	3	0.12	1
392	043110	interferon gamma-inducible protein30, put	42	100%	5	0.12	3
393	065770	nuclear cap binding protein, put	135	100%	16	0.12	6
394	058360	h.p. (<i>DnaJ domain-containing protein</i>)	34	98%	4	0.12	2
395	023620	SRS domain-containing protein	34	98%	4	0.12	2
396	013320	RPS15 - ribosomal protein RPS15	17	98%	2	0.12	2
397	007110	RPS10 - ribosomal protein RPS10	17	98%	2	0.12	2
398	017420	roptry protein kinase (ROPK)	61	100%	7	0.11	2
399	018510	hydrolase Cof, rel	35	98%	4	0.11	2
400	001450	h.p. (<i>dynein heavy chain</i>)	546	100%	62	0.11	36
401	031970	h.p. (<i>pre-mRNA splicing factor PRP8</i>)	294	100%	33	0.11	16
402	043880	h.p. (<i>vacuolar ATP synthase subunit E</i>)	27	100%	3	0.11	3
403	038700	h.p.	74	100%	8	0.11	6
404	036230	ubiquitin fusion-degradation protein, rel	37	98%	4	0.11	2
405	001650	h.p., conserved	28	91%	3	0.11	1
406	031750	importin, put	131	100%	14	0.11	7
407	023100	h.p., conserved (<i>SP18</i>)	47	100%	5	0.11	5
408	042700	h.p., conserved	19	98%	2	0.11	2
409	044440	h.p. (<i>eukaryotic peptide release factor</i>)	48	100%	5	0.10	3
410	056700	26S proteasome regulatory subunit, rel	116	100%	12	0.10	6
411	028870	peptidylprolyl isomerase D (Cyclophilin D), rel	58	100%	6	0.10	5
412	047810	h.p. (<i>adenylate kinase</i>)	29	98%	3	0.10	2

413	049110	h.p.	39	100%	4	0.10	2
414	000020	M protein, put	88	96%	9	0.10	1
415	055070	ubiquitin carboxyl-terminal hydrolase, rel	69	98%	7	0.10	2
416	007510	h.p., conserved	89	100%	9	0.10	4
417	004520	h.p., conserved	210	100%	21	0.10	11
418	029970	TRNA synthetase Gly, rel	80	100%	8	0.10	7
419	034900	casein kinase II beta chain, put	40	100%	4	0.10	4
420	019950	adenosine deaminase (EC 3.5.4.4), rel	40	100%	4	0.10	4
421	066600	DEHA2F06798p,rel	50	100%	5	0.10	3
422	070060	RNA binding protein, put	40	100%	4	0.10	3
423	047500	trichohyalin, put	163	100%	16	0.10	8
424	039080	adaptin N terminal region protein, rel	123	100%	12	0.10	6
425	009780	beta-lactamase (precursor), rel	125	100%	12	0.10	4
426	061340	h.p. (<i>elongation factor 1</i>)	21	98%	2	0.10	2
427	059830	h.p., conserved	21	98%	2	0.10	2
428	051650	M16 family peptidase, put	148	100%	14	0.09	10
429	055370	h.p., conserved	43	98%	4	0.09	2
430	037770	erythrocyte membrane prot. PFEMP3,rel	54	98%	5	0.09	2
431	062880	h.p., conserved	87	100%	8	0.09	6
432	041780	RPL19 - ribosomal protein RPL19	22	100%	2	0.09	2
433	054700	uridine phosphorylase, put	33	100%	3	0.09	3
434	026390	centrin, put	22	98%	2	0.09	2
435	055800	serine/threon. protein phosphatase 5, rel	67	98%	6	0.09	2
436	060040	h.p. (<i>AKMT</i>)	79	100%	7	0.09	4
437	024010	h.p. (<i>proteasome PCI</i>)	46	98%	4	0.09	2
438	066900	serine/threon. protein phosphatase 5,put	117	100%	10	0.09	5
439	032620	Cs1 protein, rel	47	100%	4	0.09	4
440	064620	RON2 - rhopty neck protein 2	165	100%	14	0.08	6
441	033860	h.p., conserved	36	100%	3	0.08	3
442	061570	D-tyrosyl-tRNA(Tyr) deacylase, rel	24	100%	2	0.08	2
443	032220	RPL30 - ribosomal protein RPL30	12	97%	1	0.08	1
444	011550	novel protein (Zgc:77155), rel	24	98%	2	0.08	2
445	065210	HSP60 - heat shock protein 60	61	100%	5	0.08	5
446	017780	trichohyalin, put	199	100%	16	0.08	7
447	046030	h.p.	25	100%	2	0.08	2
448	005980	14kDa phosphohistidine phosphataseput	25	98%	2	0.08	2
449	032850	h.p.	25	98%	2	0.08	2
450	022920	renin, rel (<i>eukaryotic aspartyl protease</i>)	64	100%	5	0.08	2
451	056340	ATPdependent RNA helicase, rel	64	98%	5	0.08	2
452	017250	peptidylprolyl isomerase cyclophilin 2,rel	65	100%	5	0.08	3
453	033320	h.p. (<i>nuclear movement protein</i>)	39	100%	3	0.08	3
454	048470	AP2 domain transcription factor X-4	93	100%	7	0.08	4
455	067980	proteasome subunit alpha type, rel	27	100%	2	0.07	2
456	054910	h.p., conserved	41	98%	3	0.07	2
457	048040	h.p., conserved	55	100%	4	0.07	3
458	049400	serine hydroxymethyltransferase, rel	55	98%	4	0.07	2
459	060600	6-phosphogluconate dehydrogenase, rel	55	98%	4	0.07	2
460	068090	thrombospondin type 1 protein, put	83	100%	6	0.07	5
461	055360	RON5 - rhopty neck protein 5	181	100%	13	0.07	8
462	003440	actin	42	100%	3	0.07	3
463	024470	h.p.	29	100%	2	0.07	2
464	039000	probable 26S protease, rel	29	98%	2	0.07	2
465	058760	h.p., conserved	44	100%	3	0.07	2
466	038320	h.p. (<i>PAN domain-containing protein</i>)	119	100%	8	0.07	3
467	020480	h.p. (<i>U1 small nuclear ribonucleoprotein</i>)	30	98%	2	0.07	1
468	068970	succinate dehydrogenase flavoprot., rel	45	98%	3	0.07	2
469	053220	PsmB, rel (<i>proteasome subunit beta</i>)	31	94%	2	0.06	1
470	049710	h.p. (<i>coatomer protein complex beta</i>)	141	100%	9	0.06	5
471	029990	vacuolar ATP synthase, put	63	98%	4	0.06	2
472	069590	ROP2L8 (<i>ROP24</i>)	63	98%	4	0.06	2
473	024500	6-phosphogluconate dehydrogenase,put	34	96%	2	0.06	1

474	042510	h.p., conserved	34	98%	2	0.06	2
475	038000	CUG-BP- and ETR-3-like factor 3, rel	69	100%	4	0.06	2
476	057060	dynactin p25, put	70	100%	4	0.06	4
477	047390	h.p., conserved	53	98%	3	0.06	2
478	037200	GJ21226, rel (<i>MIF4G protein</i>)	269	100%	15	0.06	6
479	020180	h.p. (<i>nucleolar protein NOP5</i>)	54	98%	3	0.06	2
480	062950	50S ribosomal protein L21e, rel	18	97%	1	0.06	1
481	038360	TSP1 protein TSP12 (Precursor), rel	145	100%	8	0.06	4
482	053190	h.p. (<i>phenylalanyl-tRNA synthetase</i>)	74	100%	4	0.05	3
483	000570	h.p., conserved	93	100%	5	0.05	4
484	054200	Zgc:92083, rel	94	100%	5	0.05	4
485	043320	RNA recognition motif (RRM), rel	57	100%	3	0.05	3
486	005510	h.p. (<i>UBA/TS-N containing protein</i>)	38	100%	2	0.05	2
487	050080	fructose-1,6-bisphosphatase class 1, rel	19	98%	1	0.05	1
488	060700	SRS domain-containing protein	38	98%	2	0.05	2
489	063570	peptidase, S9A/B/C family, rel	97	100%	5	0.05	4
490	038580	RNA recognition motif-containing protein, put	78	100%	4	0.05	3
491	007730	valyl-tRNA synthetase, mitochondrial, rel	119	98%	6	0.05	2
492	056180	h.p., conserved	40	98%	2	0.05	2
493	043140	aspartate carbamoyltransferase, put	40	98%	2	0.05	2
494	045460	mitochondrial presequence protease, rel	143	100%	7	0.05	4
495	033270	h.p. (<i>triosephosphate isomerase</i>)	41	100%	2	0.05	2
496	022400	h.p. (<i>aspartyl-tRNA synthetase</i>)	62	98%	3	0.05	2
497	058530	tetratricopeptide TPR_2 protein, rel	62	98%	3	0.05	2
498	043930	kelch prot. fused HSP90-like ATPase, rel	207	100%	10	0.05	5
499	054250	acyl-CoA synthetase, rel	85	100%	4	0.05	3
500	016110	h.p. (<i>mitochondrial branched- E1</i>)	43	98%	2	0.05	2
501	043390	h.p. (<i>poly(A) binding protein</i>)	220	100%	10	0.05	5
502	001440	NOL1/NOP2/sun methyltransferase, rel	110	100%	5	0.05	4
503	045000	h.p.	45	100%	2	0.04	2
504	049460	adenosine monophosphate deaminase, put	136	100%	6	0.04	5
505	054800	h.p., conserved	68	98%	3	0.04	2
506	049050	RNA helicase-for pre-mRNA splicing, rel	251	100%	11	0.04	8
507	050000	probable deoxyhypusine synthase, rel	47	98%	2	0.04	2
508	042790	2OG-Fe(II) oxygenase protein, put	98	100%	4	0.04	2
509	010390	novel protein (Zgc:77804), rel	49	98%	2	0.04	2
510	049500	h.p. (<i>DNA topoisomerase I</i>)	100	100%	4	0.04	3
511	051870	DNA repair enzyme, put	50	98%	2	0.04	2
512	057470	h.p., conserved	251	100%	10	0.04	6
513	014750	h.p.	176	100%	7	0.04	4
514	051420	KLLA0B04587p, rel (<i>EPN3 protein</i>)	76	100%	3	0.04	3
515	061040	h.p.	51	98%	2	0.04	2
516	021040	SNARE domain-containing protein, put	54	98%	2	0.04	2
517	005400	Brx1, rel (<i>ribosome biogenesis protein</i>)	55	98%	2	0.04	2
518	041690	ubiquitin - hydrolase, rel	252	100%	9	0.04	4
519	007540	h.p., conserved	313	100%	11	0.04	7
520	030970	h.p. (<i>GMP synthase</i>)	59	100%	2	0.03	2
521	007390	MGC78841 protein, rel (<i>importin alpha</i>)	60	98%	2	0.03	2
522	041850	Hsp20/alpha crystallin, put	30	97%	1	0.03	1
523	063970	long chain acyl-CoA synthetase, put	91	98%	3	0.03	2
524	041380	h.p., conserved	122	96%	4	0.03	1
525	019550	histidyl-tRNA synthetase, put	122	98%	4	0.03	2
526	056660	h.p., conserved	184	100%	6	0.03	3
527	059300	CD2 antigen-binding protein 2, put	62	98%	2	0.03	2
528	037460	h.p. (<i>citrate synthase</i>)	62	98%	2	0.03	2
529	001380	ubiquitin carboxyl-terminal hydrolase, rel	63	97%	2	0.03	1
530	050680	h.p. (<i>DEAD/DEAH box helicase</i>)	63	98%	2	0.03	2
531	065560	h.p.	63	98%	2	0.03	2
532	016800	TCP-1/cpn60 chaperonin protein, put	63	98%	2	0.03	2
533	041480	h.p., conserved	64	100%	2	0.03	2

534	014290	h.p., conserved	291	100%	9	0.03	6
535	016900	h.p.	98	100%	3	0.03	2
536	038850	fumarase, put	66	98%	2	0.03	2
537	010520	h.p., conserved	101	100%	3	0.03	2
538	061460	family T1, threonine peptidase, rel	34	95%	1	0.03	1
539	061660	CAM kinase, put (<i>CDPK6</i>)	69	97%	2	0.03	1
540	019240	h.p., conserved	70	98%	2	0.03	2
541	028180	DNA-directed RNA polymerase sub.beta	141	100%	4	0.03	3
542	057550	SUB2 - subtilisin-like protease SUB2	141	100%	4	0.03	3
543	010960	phosphoglucomutase, put	72	100%	2	0.03	2
544	027530	lectin-domain protein, put	74	100%	2	0.03	2
545	065520	h.p.	111	98%	3	0.03	2
546	070050	h.p. (<i>formin binding protein</i>)	75	98%	2	0.03	2
547	025000	h.p. (<i>trypsin</i>)	113	100%	3	0.03	3
548	046750	h.p. (<i>poly(A) polymerase</i>)	77	100%	2	0.03	2
549	033280	h.p. (<i>ATP-dependent RNA helicase</i>)	234	100%	6	0.03	3
550	042980	h.p. (<i>CCR4 - Not1 containing protein</i>)	274	100%	7	0.03	5
551	014450	phosphoglucomutase 2, rel	79	98%	2	0.03	2
552	034090	kinesin heavy chain, put	199	100%	5	0.03	4
553	060220	h.p., conserved	320	100%	8	0.03	7
554	005140	ubiquitin carboxyl-term hydrolase, put	120	98%	3	0.03	2
555	070210	peptidyl-prolyl cis-trans isomerase, put	80	98%	2	0.03	2
556	063730	MIC15 - microneme protein 15	285	100%	7	0.02	7
557	037180	h.p. (<i>crooked neck-like protein 1</i>)	82	100%	2	0.02	2
558	060330	molybdopterin cofactor sulfurase, put	82	98%	2	0.02	2
559	065290	h.p., conserved	253	100%	6	0.02	4
560	051510	GM14051,rel(<i>translation initiat. factor2</i>)	128	100%	3	0.02	3
561	009630	cyclophilin, put	86	98%	2	0.02	2
562	018500	fatty acyl-CoA synthetase 2, rel	90	98%	2	0.02	2
563	031340	splicing factor 3B subunit 3, put	136	98%	3	0.02	2
564	016240	5'-nucleotidase domain protein, rel	96	100%	2	0.02	2
565	059450	h.p. (<i>ATP-dependent RNA helicase</i>)	49	97%	1	0.02	1
566	049920	h.p., conserved	101	98%	2	0.02	2
567	017040	phosphofructokinase, rel	155	100%	3	0.02	3
568	041660	h.p. (<i>cullin family protein</i>)	104	100%	2	0.02	2
569	022070	zinc finger (C3HC4 type) / FHA, put	104	98%	2	0.02	2
570	036570	YALI0D21604p, rel	105	100%	2	0.02	2
571	043960	h.p., conserved	269	92%	5	0.02	1
572	033430	h.p., conserved	54	96%	1	0.02	1
573	034530	TCP-1/cpn60 family chaperonin, put	58	95%	1	0.02	1
574	062780	dipeptidyl peptidase IV, put	195	98%	3	0.02	2
575	036080	h.p., conserved	131	100%	2	0.02	2
576	051350	h.p.	203	100%	3	0.01	3
577	054950	AP2XI-3 - AP2	136	98%	2	0.01	2
578	002910	ftsJ-like methyltransferase, put	137	100%	2	0.01	2
579	005200	h.p.	212	98%	3	0.01	2
580	015430	h.p. (<i>thioredoxin</i>)	212	98%	3	0.01	2
581	051140	TLD domain-containing protein, put	222	100%	3	0.01	2
582	035920	adenylate/guanylate cyclase, rel	224	100%	3	0.01	3
583	020650	splicing factor 3B subunit 1, put	152	100%	2	0.01	2
584	032840	protein kinase CMGC/MAPK/ERK7	76	98%	1	0.01	1
585	004350	calcium homeostasis endopl retic protein, rel	81	91%	1	0.01	1
586	000060	cyclic nucleotide phosphodiesterase, put	245	98%	3	0.01	2
587	050050	M16 family peptidase, put	172	98%	2	0.01	2
588	055820	h.p. (<i>DNA topoisomerase II</i>)	184	100%	2	0.01	2
589	050870	SNF2 family, put	276	100%	3	0.01	3
590	004580	h.p.	374	97%	4	0.01	1
591	009210	h.p.	381	100%	4	0.01	3
592	017750	h.p., conserved	485	100%	5	0.01	1
593	001570	eukaryotic translation initiat.factor 3, rel	206	98%	2	0.01	2

594	023830	AP2Vlla-9 - AP2	112	90%	1	0.01	1
595	027010	h.p., conserved	126	96%	1	0.01	1
596	001370	DEAD/DEAH box helicase, put	255	100%	2	0.01	2
597	039670	200 kDa antigen p200, put	134	100%	1	0.01	1
598	069920	transcription factor S-II, put	146	95%	1	0.01	1
599	035850	h.p.	323	98%	2	0.01	2
600	056040	h.p., conserved	168	93%	1	0.01	1
601	017880	h.p., conserved	362	98%	2	0.01	2
602	068070	h.p., conserved	192	100%	1	0.01	1
603	021920	h.p. (<i>protein kinase Other/Aur</i>)	216	91%	1	0.00	1
604	062460	h.p., conserved	217	97%	1	0.00	1
605	065460	CMGC kinase, put	219	99%	1	0.00	1
606	047840	h.p., conserved	918	100%	4	0.00	2
607	014560	h.p., conserved	294	100%	1	0.00	1
608	001690	dehydrogenase/reductase SDR, rel	1071	100%	3	0.00	1
609	053360	h.p., conserved	368	95%	1	0.00	1
610	013830	dynein heavy chain family protein, rel	409	99%	1	0.00	1
611	014200	h.p., conserved	563	99%	1	0.00	1
612	043020	h.p.	578	100%	1	0.00	1
613	019860	h.p., conserved	767	100%	1	0.00	1
614	037840	h.p.	920	98%	1	0.00	1
615	017650	h.p.	959	98%	1	0.00	1

***Neospora caninum*: estudo do secretoma e caracterização molecular de três proteínas com domínios Apple**

Apêndice 2

Apêndice 2 Proteínas identificadas e quantificadas no extrato total de *N. caninum* submetido ao estímulo secretório por etanol (taquizoíta estimulado). Organização decrescente de acordo com a razão estimulado/controle. Informações são apresentadas quanto ao número de peptídeos únicos identificados; número de PSM's (peptide spectrum matches, número de espectros de fragmentação experimental que correspondem ao espectro teórico, gerado a partir do banco de dados); área sob pico, que corresponde à intensidade; razão entre pepídeos com marcação intermediária sobre leve (*estimulado/controle*); e o log na base 2 ($\log_2$) da razão *estimulado/controle*. As razões em negrito, maior que 1 e menor que 1, são referentes às proteínas aumentadas ediminuídas, respectivamente, no taquizoíta *estimulado*. rel, related; put, putative.

	ToxoDB ID	Description	Número peptídeos únicos	número PSM's	área sob o pico (intensidade)	Razão <i>estimul/controle</i>	Razão ($\log_2$)
1	NCLIV_040870	hypothetical protein, conserved	1	1	3.665E7	12.201	3.609
2	NCLIV_020380	hypothetical protein, conserved	2	2	3.424E7	6.383	2.674
3	NCLIV_032250	hypothetical protein, conserved	1	4	7.821E7	5.933	2.569
4	NCLIV_062830	hypothetical protein, conserved	1	4	5.771E7	5.486	2.456
5	NCLIV_067860	hypothetical protein, conserved	2	8	2.103E7	5.061	2.340
6	NCLIV_037770	ERYTHROCYTE MEMBRANE PROTEIN PFEMP3, related	1	2	3.751E6	4.655	2.219
7	NCLIV_060430	Os08g0535200 protein, related	1	1	5.183E7	4.304	2.106
8	NCLIV_051150	hypothetical protein, conserved	1	2	3.019E8	4.282	2.098
9	NCLIV_005640	hypothetical protein, conserved	1	2	5.452E7	4.074	2.026
10	NCLIV_009220	hypothetical protein, conserved	1	2	2.545E8	3.879	1.956
11	NCLIV_060180	cDNA FLJ58794, highly similar to Pogo transposable element with KRAB domain, related	1	1	2.295E7	3.875	1.954
12	NCLIV_051890	hypothetical protein, conserved	5	18	2.386E8	3.841	1.942
13	NCLIV_032340	PAB-dependent poly(A)-specific ribonuclease subunit PAN2, rel	1	3	3.048E7	3.783	1.920
14	NCLIV_068900	hypothetical protein	1	4	1.602E7	3.779	1.918
15	NCLIV_027450	hypothetical protein, conserved	1	12	2.107E8	3.723	1.897
16	NCLIV_038470	hypothetical protein, conserved	3	8	4.256E7	3.718	1.895
17	NCLIV_017110	hypothetical protein, conserved	1	5	6.112E7	3.689	1.883
18	NCLIV_045520	hypothetical protein, conserved	1	1	2.852E6	3.688	1.883
19	NCLIV_053220	PsmB (EC 3.4.25.1), related	1	7	4.217E7	3.641	1.864
20	NCLIV_065980	centrosomal protein, putative	1	1	5.631E7	3.634	1.862
21	NCLIV_022380	NLI interacting factor-like phosphatase containing protein, put	3	9	3.939E7	3.567	1.835
22	NCLIV_070070	calcyclin binding protein, putative	1	2	2.231E7	3.550	1.828
23	NCLIV_026050	hypothetical protein, conserved	1	5	2.419E7	3.549	1.828
24	NCLIV_017040	Phosphofructokinase, related	3	12	1.522E8	3.509	1.811
25	NCLIV_067800	hypothetical protein, conserved	4	8	3.971E7	3.499	1.807
26	NCLIV_037200	GJ21226, related	6	17	6.471E7	3.474	1.797
27	NCLIV_010540	hypothetical protein, conserved	1	2	1.483E7	3.468	1.794

28	NCLIV_067970	double-strand break repair protein, putative	1	1	1.253E6	3.364	1.750
29	NCLIV_031320	hypothetical protein, conserved	8	30	2.152E8	3.268	1.709
30	NCLIV_021570	hypothetical protein, conserved	1	5	8.352E7	3.251	1.701
31	NCLIV_042980	hypothetical protein	3	7	5.205E7	3.183	1.670
32	NCLIV_067460	hypothetical protein, conserved	1	2	5.960E7	3.175	1.667
33	NCLIV_039480	hypothetical protein	1	1	1.152E8	3.126	1.644
34	NCLIV_047820	hypothetical protein, conserved	7	19	1.697E8	3.115	1.639
35	NCLIV_008490	RNA recognition motif. family protein, related	1	2	3.530E7	3.107	1.636
36	NCLIV_035900	RNA binding motif-containing protein, putative	3	6	3.185E7	3.101	1.633
37	NCLIV_020970	hypothetical protein, conserved	2	9	6.326E7	3.031	1.600
38	NCLIV_065290	hypothetical protein, conserved	2	8	5.563E7	3.025	1.597
39	NCLIV_046960	hypothetical protein, conserved	4	7	5.789E7	3.006	1.588
40	NCLIV_040010	alpha-1 type II collagen, putative	9	17	8.613E7	2.996	1.583
41	NCLIV_056570	Collagen alpha-1(III) chain (Precursor), related	6	11	4.055E7	2.981	1.576
42	NCLIV_002840	hypothetical protein, conserved	1	4	1.131E8	2.974	1.572
43	NCLIV_015740	hypothetical protein, conserved	2	5	1.672E7	2.965	1.568
44	NCLIV_032740	hypothetical protein, conserved	2	7	2.394E7	2.961	1.566
45	NCLIV_045200	hypothetical protein, conserved	1	2	2.956E7	2.951	1.561
46	NCLIV_001290	hypothetical protein, conserved	3	12	4.613E8	2.949	1.560
47	NCLIV_000890	ER lumen protein retaining receptor 2, related	1	4	1.480E7	2.942	1.557
48	NCLIV_038250	hypothetical protein, conserved	1	1	5.118E7	2.941	1.556
49	NCLIV_046940	PWWP domain-containing protein, putative	1	3	3.025E7	2.920	1.546
50	NCLIV_017830	5'-3' exonuclease, putative	1	1	1.958E7	2.902	1.537
51	NCLIV_047470	hypothetical protein, conserved	6	11	1.248E8	2.894	1.533
52	NCLIV_047980	hypothetical protein, conserved	1	5	7.362E5	2.830	1.501
53	NCLIV_044390	hypothetical protein, conserved	2	14	3.851E8	2.795	1.483
54	NCLIV_007810	hypothetical protein, conserved	5	9	3.086E8	2.770	1.470
55	NCLIV_045280	hypothetical protein, conserved	1	2	9.084E6	2.763	1.466
56	NCLIV_062970	helicase conserved C-terminal domain-containing protein, put	1	2	7.497E6	2.751	1.460
57	NCLIV_063260	hypothetical protein, conserved	1	3	6.694E7	2.750	1.459
58	NCLIV_058540	RNA polymerase II subunit RPB11, putative	1	6	7.279E7	2.738	1.453
59	NCLIV_060560	hypothetical protein, conserved	3	6	7.500E7	2.733	1.450
60	NCLIV_036010	hypothetical protein	1	4	1.319E8	2.705	1.435
61	NCLIV_050040	hypothetical protein, conserved	1	5	1.885E8	2.663	1.413
62	NCLIV_031700	DEHA2G03652p, related	4	10	1.104E8	2.644	1.403
63	NCLIV_001920	ATP-dependent helicase, putative	1	1	2.884E7	2.641	1.401
64	NCLIV_049750	replication factor c subunit, putative	1	4	2.175E7	2.626	1.393
65	NCLIV_069920	transcription factor S-II central domain- containing protein, put	2	6	3.393E7	2.624	1.392
66	NCLIV_055300	hypothetical protein	14	49	5.092E8	2.604	1.381
67	NCLIV_010090	hypothetical protein, conserved	2	5	4.499E7	2.592	1.374
68	NCLIV_019560	hypothetical protein, conserved	4	8	2.630E7	2.589	1.372
69	NCLIV_029660	Proteophosphoglycan 5, related	1	10	3.401E9	2.574	1.364
70	NCLIV_046200	hypothetical protein	1	3	1.067E7	2.565	1.359

71	NCLIV_014940	hypothetical protein, conserved	3	6	5.622E7	2.553	1.352
72	NCLIV_036810	hypothetical protein	5	9	9.375E7	2.542	1.346
73	NCLIV_036820	hypothetical protein	3	5	4.234E7	2.523	1.335
74	NCLIV_022030	hypothetical protein, conserved	3	8	1.107E8	2.522	1.334
75	NCLIV_021200	hypothetical protein, conserved	1	2	1.790E8	2.519	1.333
76	NCLIV_031430	protein kinase, putative	2	3	1.483E8	2.511	1.328
77	NCLIV_012100	Vacuolar protein sorting-associated protein, related	1	2	2.495E7	2.495	1.319
78	NCLIV_012950	ubiquitin family domain-containing protein, putative	4	13	1.376E8	2.492	1.317
79	NCLIV_028650	hypothetical protein, conserved	2	6	1.520E8	2.489	1.316
80	NCLIV_029810	Tcc1a22.4, related	3	7	2.479E7	2.487	1.315
81	NCLIV_025440	hypothetical protein	1	7	2.844E7	2.483	1.312
82	NCLIV_063920	hypothetical protein, conserved	3	10	6.544E7	2.455	1.296
83	NCLIV_020950	CMGC kinase, Dyrk family, putative	2	5	5.852E7	2.452	1.294
84	NCLIV_026460	hypothetical protein	2	5	2.933E7	2.447	1.291
85	NCLIV_007880	GM10776, related	5	12	8.999E7	2.442	1.288
86	NCLIV_015600	hypothetical protein	1	2	2.874E7	2.439	1.286
87	NCLIV_004370	hypothetical protein, conserved	1	1	3.317E7	2.438	1.286
88	NCLIV_022460	zinc finger (C2H2 type) domain-containing protein, putative	1	1	1.823E6	2.431	1.281
89	NCLIV_011360	SET domain-containing protein, putative	1	1	6.616E6	2.427	1.279
90	NCLIV_052600	YEATS family protein, related	1	2	1.133E7	2.419	1.274
91	NCLIV_029290	hypothetical protein	1	1	6.404E6	2.406	1.266
92	NCLIV_028800	hypothetical protein, conserved	1	1	3.699E6	2.405	1.266
93	NCLIV_018190	hypothetical protein, conserved	2	6	1.082E7	2.405	1.266
94	NCLIV_045660	hypothetical protein, conserved	1	9	5.507E8	2.401	1.264
95	NCLIV_029260	hypothetical protein	1	12	1.470E7	2.392	1.258
96	NCLIV_017730	Guanylate kinase, related	1	1	4.135E7	2.391	1.258
97	NCLIV_001620	hypothetical protein, conserved	5	11	1.718E8	2.380	1.251
98	NCLIV_020050	Zinc finger (C3HC4 RING finger) protein, related	1	3	1.620E7	2.376	1.249
99	NCLIV_066340	hypothetical protein, conserved	1	4	5.528E8	2.365	1.242
100	NCLIV_056600	hypothetical protein, conserved	1	4	3.981E7	2.352	1.234
101	NCLIV_060900	hypothetical protein, conserved	10	36	2.713E8	2.351	1.233
102	NCLIV_031150	hypothetical protein, conserved	1	2	9.495E6	2.343	1.228
103	NCLIV_011140	GL18351, related	2	6	8.533E7	2.338	1.225
104	NCLIV_029460	hypothetical protein, conserved	1	2	3.102E7	2.317	1.212
105	NCLIV_022000	para-aminobenzoate synthase, putative	1	2	1.106E8	2.311	1.208
106	NCLIV_002350	hypothetical protein	1	2	5.938E6	2.309	1.207
107	NCLIV_064180	hypothetical protein, conserved	1	2	1.483E7	2.302	1.203
108	NCLIV_040850	nuclear NF-kB activating protein, putative	1	2	3.676E7	2.296	1.199
109	NCLIV_021110	hypothetical protein, conserved	2	13	9.016E7	2.280	1.189
110	NCLIV_031280	PHD-finger domain-containing protein, putative	1	5	3.271E7	2.267	1.181
111	NCLIV_030010	DEAD/DEAH box helicase, putative	1	1	8.218E6	2.265	1.179
112	NCLIV_003800	hypothetical protein, conserved	3	7	6.440E7	2.261	1.177
113	NCLIV_052790	hypothetical protein	4	7	3.979E7	2.253	1.172

114	NCLIV_059260	hypothetical protein	1	4	5.380E7	2.252	1.171
115	NCLIV_001240	hypothetical protein	2	4	1.422E7	2.233	1.159
116	NCLIV_068020	OTU-like cysteine protease domain-containing protein, putative	1	1	4.920E6	2.217	1.149
117	NCLIV_031260	autophagy protein, putative	1	1	1.144E7	2.215	1.147
118	NCLIV_048300	hypothetical protein	3	12	5.613E7	2.202	1.139
119	NCLIV_010310	hypothetical protein, conserved	1	2	5.357E6	2.191	1.132
120	NCLIV_061920	zinc finger (CCCH type) protein, putative	1	2	5.088E7	2.154	1.107
121	NCLIV_059690	hypothetical protein, conserved	1	2	2.347E7	2.153	1.106
122	NCLIV_049460	adenosine monophosphate deaminase, putative	3	7	4.044E7	2.148	1.103
123	NCLIV_030970	hypothetical protein	4	13	1.832E8	2.143	1.100
124	NCLIV_059300	CD2 antigen cytoplasmic tail-binding protein 2, putative	1	2	2.824E7	2.133	1.093
125	NCLIV_040510	hypothetical protein	3	3	4.411E7	2.123	1.086
126	NCLIV_044630	hypothetical protein, conserved	1	3	2.917E7	2.121	1.085
127	NCLIV_014290	hypothetical protein, conserved	1	2	6.714E7	2.111	1.078
128	NCLIV_018810	hypothetical protein, conserved	1	2	4.381E5	2.106	1.075
129	NCLIV_058980	1,4-alpha-glucan branching enzyme, putative	1	2	4.000E6	2.102	1.071
130	NCLIV_044270	hypothetical protein	10	41	8.148E8	2.086	1.061
131	NCLIV_009210	hypothetical protein	6	12	7.845E7	2.086	1.061
132	NCLIV_010110	hypothetical protein	1	8	1.400E8	2.083	1.059
133	NCLIV_048470	hypothetical protein, conserved	2	7	1.374E8	2.077	1.054
134	NCLIV_026000	hypothetical protein, conserved	1	4	8.243E8	2.075	1.053
135	NCLIV_021260	hypothetical protein	2	7	8.619E7	2.061	1.044
136	NCLIV_065460	CMGC kinase, putative	2	12	1.183E8	2.059	1.042
137	NCLIV_004520	hypothetical protein, conserved	4	14	1.211E8	2.056	1.040
138	NCLIV_032930	hypothetical protein, conserved	8	23	1.033E8	2.052	1.037
139	NCLIV_068550	hypothetical protein, conserved	1	1	2.891E7	2.052	1.037
140	NCLIV_012500	hypothetical protein, conserved	2	7	3.225E7	2.051	1.036
141	NCLIV_018560	ras-GTPase-activating protein binding protein, putative	13	47	3.805E8	2.050	1.036
142	NCLIV_052320	RNA binding protein, putative	1	4	3.197E6	2.043	1.031
143	NCLIV_002470	hypothetical protein, conserved	6	12	3.341E8	2.038	1.027
144	NCLIV_048740	hypothetical protein, conserved	3	7	1.559E8	2.036	1.026
145	NCLIV_044340	hypothetical protein, conserved	9	38	2.820E8	2.034	1.024
146	NCLIV_042170	PHD finger protein BR140/LIN-49, related	4	14	5.082E7	2.031	1.022
147	NCLIV_003130	GTP-binding protein, related	1	2	1.396E7	2.025	1.018
148	NCLIV_004480	myb-like DNA-binding domain-containing protein, putative	4	10	1.472E8	2.025	1.018
149	NCLIV_046340	hypothetical protein, conserved	6	17	1.633E8	2.014	1.010
150	NCLIV_013640	hypothetical protein	1	2	7.431E7	2.012	1.009
151	NCLIV_030400	SNF2 family helicase, putative	1	1	6.609E6	1.998	0.999
152	NCLIV_055550	platelet binding protein GspB, putative	11	20	2.124E8	1.995	0.996
153	NCLIV_032350	CW-type zinc finger domain-containing protein, putative	3	5	4.995E7	1.990	0.993
154	NCLIV_056060	hypothetical protein	1	10	7.802E7	1.986	0.990
155	NCLIV_051990	hypothetical protein, conserved	2	8	2.780E8	1.984	0.988
156	NCLIV_062780	dipeptidyl peptidase IV domain-containing protein, putative	7	17	1.328E8	1.980	0.986

157	NCLIV_013160	hypothetical protein, conserved	3	3	2.205E7	1.979	0.985
158	NCLIV_044930	Sec7 domain containing protein, related	1	11	1.041E8	1.978	0.984
159	NCLIV_015670	hypothetical protein, conserved	2	4	2.056E7	1.975	0.982
160	NCLIV_035750	vacuolar protein sorting-associated protein, putative	2	2	3.334E7	1.969	0.977
161	NCLIV_048820	50S ribosomal protein L4, putative	1	4	5.323E7	1.966	0.975
162	NCLIV_019390	SNF2 family N-terminal domain-containing protein, putative	1	3	7.316E7	1.953	0.966
163	NCLIV_044830	RNA pseudouridine synthase domain containing protein, putative	3	9	5.873E7	1.951	0.964
164	NCLIV_046990	hypothetical protein, conserved	1	2	2.452E7	1.949	0.963
165	NCLIV_036550	hypothetical protein, conserved	1	2	6.018E7	1.949	0.962
166	NCLIV_016490	hypothetical protein, conserved	3	5	1.285E8	1.944	0.959
167	NCLIV_010960	phosphoglucomutase, putative	5	14	7.519E7	1.942	0.957
168	NCLIV_049430	hypothetical protein	7	24	1.540E8	1.940	0.956
169	NCLIV_049870	GH22120, related	2	9	1.459E8	1.933	0.951
170	NCLIV_035470	hypothetical protein, conserved	1	10	8.496E7	1.927	0.946
171	NCLIV_051260	hypothetical protein, conserved	7	13	1.176E8	1.926	0.945
172	NCLIV_061420	hypothetical protein, conserved	7	24	5.761E7	1.924	0.944
173	NCLIV_043260	Salivary proline-rich protein, related	2	7	4.635E7	1.911	0.934
174	NCLIV_053590	WD-40 repeat protein, related	2	8	8.319E7	1.910	0.934
175	NCLIV_017610	iron only hydrogenase large subunit, C-terminal domain-containing protein, putative	2	4	1.810E7	1.909	0.933
176	NCLIV_045060	hypothetical protein, conserved	1	3	1.670E7	1.906	0.930
177	NCLIV_030230	hypothetical protein, conserved	4	8	2.868E7	1.903	0.928
178	NCLIV_008160	hypothetical protein, conserved	1	3	5.758E7	1.903	0.928
179	NCLIV_070180	hypothetical protein, conserved	1	2	6.324E7	1.899	0.925
180	NCLIV_023250	hypothetical protein	1	2	1.945E8	1.893	0.921
181	NCLIV_016060	protein kinase, putative	1	1	4.935E6	1.893	0.920
182	NCLIV_038370	hypothetical protein, conserved	1	1	8.854E7	1.891	0.919
183	NCLIV_054920	hypothetical protein, conserved	6	12	7.527E7	1.888	0.917
184	NCLIV_003520	mRNA decapping enzyme, putative	2	10	4.180E7	1.880	0.911
185	NCLIV_023510	hypothetical protein, conserved	1	1	5.732E7	1.873	0.905
186	NCLIV_054720	hypothetical protein	12	64	1.134E9	1.870	0.903
187	NCLIV_020530	hypothetical protein	2	5	3.223E7	1.869	0.902
188	NCLIV_021630	Large protein with a UBR1-like ring finger related treble clef, rel	1	3	9.863E6	1.867	0.901
189	NCLIV_063010	hypothetical protein, conserved	1	2	2.014E7	1.861	0.896
190	NCLIV_049030	hypothetical protein	1	1	1.210E8	1.861	0.896
191	NCLIV_027310	ATP-dependent nucleolar RNA helicase, putative	3	8	4.651E7	1.857	0.893
192	NCLIV_022960	GM13279, related	1	3	7.089E7	1.855	0.891
193	NCLIV_006460	beta-tubulin cofactor d, putative	1	2	1.563E8	1.852	0.889
194	NCLIV_010120	hypothetical protein, conserved	1	1	7.681E7	1.844	0.883
195	NCLIV_025150	TPR Domain containing protein, putative	1	3	2.666E7	1.840	0.880
196	NCLIV_018600	DEAD/DEAH box helicase, putative	1	3	2.705E7	1.840	0.880
197	NCLIV_059210	hypothetical protein	3	7	4.994E7	1.840	0.880
198	NCLIV_052480	Protein arginine methyltransferase 5, related	1	7	6.415E7	1.835	0.876

199	NCLIV_001650	hypothetical protein, conserved	2	6	7.654E7	1.827	0.869
200	NCLIV_004200	hypothetical protein	2	7	8.192E7	1.826	0.869
201	NCLIV_063810	hypothetical protein, conserved	4	7	1.362E8	1.825	0.868
202	NCLIV_056470	hypothetical protein, conserved	2	4	3.692E7	1.824	0.867
203	NCLIV_017780	Trichohyalin, putative	1	6	3.691E7	1.818	0.862
204	NCLIV_029180	WD-40 repeat-containing protein, putative	4	5	3.893E7	1.814	0.860
205	NCLIV_037370	hypothetical protein, conserved	3	7	1.244E7	1.814	0.859
206	NCLIV_024920	hypothetical protein	3	10	1.047E8	1.814	0.859
207	NCLIV_052910	Myb1 protein, related	4	12	1.542E8	1.808	0.854
208	NCLIV_028300	signal recognition particle, putative	2	7	7.385E7	1.807	0.854
209	NCLIV_044660	Zgc:158610 protein, related	1	3	5.314E7	1.806	0.853
210	NCLIV_030470	hypothetical protein, conserved	19	67	7.524E8	1.804	0.852
211	NCLIV_005180	hypothetical protein	2	4	3.593E7	1.804	0.851
212	NCLIV_048390	hypothetical protein, conserved	1	2	9.009E6	1.800	0.848
213	NCLIV_038450	hypothetical protein, conserved	2	4	4.863E7	1.799	0.847
214	NCLIV_040610	Virulent strain associated lipoprotein, related	15	77	8.507E8	1.797	0.845
215	NCLIV_054950	hypothetical protein	4	13	2.398E8	1.796	0.845
216	NCLIV_009790	Rab 11b, putative	2	7	8.604E7	1.793	0.842
217	NCLIV_064820	hypothetical protein, conserved	1	3	3.477E7	1.792	0.841
218	NCLIV_051140	TLD domain-containing protein, putative	2	8	4.109E7	1.785	0.836
219	NCLIV_040950	hypothetical protein, conserved	1	1	1.488E6	1.784	0.835
220	NCLIV_032610	hypothetical protein	1	2	4.080E7	1.779	0.831
221	NCLIV_000260	hypothetical protein, conserved	4	18	2.041E8	1.777	0.829
222	NCLIV_005550	hypothetical protein	7	37	1.767E9	1.768	0.822
223	NCLIV_034270	coatamer gamma 2-subunit protein, putative	8	31	3.951E8	1.765	0.820
224	NCLIV_000130	hypothetical protein	4	18	4.701E7	1.764	0.819
225	NCLIV_041000	Ribosomal RNA small subunit methyltransferase F (RRNA(Cytosine-C)), related	7	9	1.009E8	1.762	0.818
226	NCLIV_030630	p25-alpha family protein, related	1	7	1.264E8	1.762	0.817
227	NCLIV_008570	hypothetical protein, conserved	2	5	8.295E7	1.760	0.816
228	NCLIV_012320	hypothetical protein, conserved	1	3	1.480E8	1.759	0.815
229	NCLIV_027010	hypothetical protein, conserved	1	2	2.103E7	1.754	0.811
230	NCLIV_060980	hypothetical protein, conserved	1	3	4.966E6	1.754	0.810
231	NCLIV_023890	hypothetical protein	1	4	2.501E7	1.754	0.810
232	NCLIV_019020	hypothetical protein, conserved	5	16	8.219E7	1.752	0.809
233	NCLIV_012360	leucine rich repeat protein, putative	1	1	1.882E7	1.752	0.809
234	NCLIV_068240	hypothetical protein, conserved	2	3	2.952E7	1.748	0.806
235	NCLIV_033550	hypothetical protein	2	10	4.925E7	1.748	0.806
236	NCLIV_017680	hypothetical protein	1	10	4.671E7	1.746	0.804
237	NCLIV_020130	hypothetical protein	2	4	1.059E7	1.745	0.803
238	NCLIV_025650	hypothetical protein, conserved	2	9	1.181E8	1.742	0.801
239	NCLIV_019550	histidyl-tRNA synthetase, putative	7	20	2.248E8	1.741	0.800
240	NCLIV_047620	hypothetical protein, conserved	1	2	4.311E6	1.741	0.800

241	NCLIV_005630	hypothetical protein, conserved	2	6	1.453E7	1.739	0.798
242	NCLIV_070210	peptidyl-prolyl cis-trans isomerase, putative	1	1	1.972E7	1.737	0.797
243	NCLIV_006890	DEAD-box helicase family protein, related	1	1	1.107E7	1.734	0.794
244	NCLIV_045100	FK506-binding protein 1, putative	3	17	1.251E8	1.733	0.793
245	NCLIV_010150	RNA recognition motif-containing protein, putative	3	9	1.320E8	1.732	0.792
246	NCLIV_068580	hypothetical protein	3	6	3.049E7	1.732	0.792
247	NCLIV_033850	hypothetical protein	10	32	2.821E8	1.728	0.789
248	NCLIV_007230	hypothetical protein	2	7	1.015E8	1.725	0.786
249	NCLIV_018150	CBR-FKB-6 protein, related	1	1	2.334E7	1.724	0.786
250	NCLIV_021290	Proteophosphoglycan 5, related	7	21	2.271E8	1.724	0.786
251	NCLIV_056660	hypothetical protein, conserved	4	13	1.283E8	1.723	0.785
252	NCLIV_028160	hypothetical protein	3	7	8.546E7	1.721	0.783
253	NCLIV_021490	GTPase domain containing protein, putative	2	4	6.215E7	1.720	0.782
254	NCLIV_005860	IMP-specific 5'-nucleotidase, related	1	4	6.298E7	1.719	0.781
255	NCLIV_051070	membrane fatty acid desaturase, putative	2	5	1.139E8	1.717	0.779
256	NCLIV_046190	Multidomain chromatinic protein 3x PHD-bromo-3xPHD-SET associated cysteine cluster at the C-terminus, related	1	3	6.561E7	1.714	0.777
257	NCLIV_043030	hypothetical protein	3	8	4.286E8	1.712	0.776
258	NCLIV_064110	RSC6/BAF60A ortholog with a SWIB domain, related	2	9	9.184E7	1.711	0.775
259	NCLIV_032970	hypothetical protein, conserved	4	12	2.074E8	1.710	0.774
260	NCLIV_055160	hypothetical protein, conserved	26	126	1.017E9	1.710	0.774
261	NCLIV_023040	hypothetical protein	1	7	6.863E8	1.709	0.773
262	NCLIV_039290	hypothetical protein	1	3	4.961E7	1.708	0.772
263	NCLIV_020310	fructose-1,6-bisphosphatase, putative	3	7	1.813E8	1.705	0.769
264	NCLIV_039760	hypothetical protein, conserved	2	4	2.027E7	1.703	0.768
265	NCLIV_063460	hypothetical protein, conserved	1	1	1.270E8	1.703	0.768
266	NCLIV_024610	hypothetical protein, conserved	6	21	3.576E8	1.701	0.767
267	NCLIV_045860	hypothetical protein	3	12	1.060E8	1.701	0.766
268	NCLIV_046360	hypothetical protein, conserved	1	5	4.472E8	1.699	0.765
269	NCLIV_038290	hypothetical protein, conserved	1	3	2.443E7	1.698	0.764
270	NCLIV_025110	hypothetical protein, conserved	1	3	2.310E7	1.694	0.760
271	NCLIV_049220	hypothetical protein, conserved	1	1	1.294E7	1.693	0.760
272	NCLIV_053120	3-dehydroquinate synthase (EC 4.2.3.4), related	4	13	1.013E8	1.691	0.758
273	NCLIV_002120	hypothetical protein	26	73	1.621E8	1.688	0.756
274	NCLIV_063570	Peptidase, S9A/B/C family, catalytic domain protein, related	1	3	2.605E7	1.687	0.754
275	NCLIV_055860	hypothetical protein	2	10	1.009E8	1.681	0.750
276	NCLIV_019800	hypothetical protein	1	2	2.330E7	1.679	0.748
277	NCLIV_065560	hypothetical protein	2	4	9.175E6	1.673	0.743
278	NCLIV_018930	Gl15163, related	1	14	2.701E8	1.669	0.739
279	NCLIV_030690	hypothetical protein, conserved	2	5	5.908E7	1.669	0.739
280	NCLIV_056360	Eukaryotic initiation factor, related	3	8	1.939E8	1.669	0.739
281	NCLIV_001600	Possible RNA-binding protein, related	7	20	6.541E7	1.668	0.738
282	NCLIV_062320	hypothetical protein, conserved	1	2	4.607E6	1.667	0.737

283	NCLIV_001840	sec1 family domain-containing protein, putative	3	4	4.334E6	1.667	0.737
284	NCLIV_000060	3',5'-cyclic nucleotide phosphodiesterase containing protein, put	4	19	8.655E7	1.666	0.737
285	NCLIV_063490	hypothetical protein, conserved	2	4	4.015E7	1.665	0.736
286	NCLIV_070240	Thiosulfate sulfurtransferase, related	1	4	4.590E7	1.659	0.730
287	NCLIV_043390	hypothetical protein	3	11	7.451E7	1.655	0.727
288	NCLIV_054690	hypothetical protein, conserved	7	15	1.501E8	1.653	0.725
289	NCLIV_057580	hypothetical protein, conserved	3	6	3.491E7	1.652	0.724
290	NCLIV_053740	hypothetical protein, conserved	1	2	1.007E7	1.652	0.724
291	NCLIV_021300	hypothetical protein, conserved	2	6	1.225E8	1.646	0.719
292	NCLIV_044860	hypothetical protein	5	16	1.001E8	1.645	0.718
293	NCLIV_031210	ATPase, AAA family domain-containing protein, putative	1	2	2.752E7	1.643	0.716
294	NCLIV_013690	GF22565, related	1	2	6.403E7	1.642	0.715
295	NCLIV_026520	hypothetical protein, conserved	3	11	5.965E8	1.640	0.714
296	NCLIV_070060	RNA binding protein, putative	8	30	9.553E8	1.639	0.713
297	NCLIV_014470	hypothetical protein, conserved	1	8	7.823E7	1.639	0.713
298	NCLIV_015000	Os05g0488800 protein, related	3	5	8.109E7	1.638	0.712
299	NCLIV_061290	DEAD/DEAH box RNA helicase, putative	2	13	3.463E8	1.636	0.710
300	NCLIV_068440	hypothetical protein, conserved	6	9	9.269E7	1.635	0.709
301	NCLIV_037840	hypothetical protein	8	17	1.428E8	1.632	0.707
302	NCLIV_006620	Trehalose-6-phosphate synthase of likely plant origin, related	5	7	5.393E7	1.631	0.706
303	NCLIV_003550	hypothetical protein	1	3	3.302E7	1.629	0.704
304	NCLIV_032840	Mitogen-activated protein kinase 2, related	2	7	7.788E7	1.626	0.701
305	NCLIV_003620	KH domain-containing protein, putative	6	17	4.340E8	1.625	0.701
306	NCLIV_016520	hypothetical protein	1	9	4.472E7	1.624	0.700
307	NCLIV_013650	hypothetical protein	2	4	2.982E7	1.624	0.699
308	NCLIV_030870	hypothetical protein, conserved	2	7	5.535E7	1.622	0.698
309	NCLIV_069600	hypothetical protein, conserved	2	16	1.669E8	1.622	0.697
310	NCLIV_066680	hypothetical protein	1	2	1.851E7	1.621	0.697
311	NCLIV_017270	hypothetical protein, conserved	2	9	3.901E7	1.620	0.696
312	NCLIV_012680	hypothetical protein, conserved	8	18	5.476E7	1.620	0.696
313	NCLIV_065910	hypothetical protein, conserved	2	9	4.738E8	1.619	0.695
314	NCLIV_046690	VASA RNA helicase, related	11	53	5.708E8	1.613	0.690
315	NCLIV_031410	hypothetical protein	1	1	8.903E6	1.611	0.688
316	NCLIV_012930	hypothetical protein	1	2	1.022E7	1.609	0.686
317	NCLIV_024640	Rhomboid family protein, related	1	5	4.240E7	1.606	0.684
318	NCLIV_041140	endonuclease V, putative	2	8	4.652E7	1.606	0.683
319	NCLIV_009170	Proteasome (Prosome, macropain) subunit, beta type, 1, related	1	12	2.346E8	1.604	0.682
320	NCLIV_030680	hypothetical protein	1	2	5.664E7	1.600	0.678
321	NCLIV_065920	hypothetical protein, conserved	1	2	1.043E7	1.598	0.676
322	NCLIV_044610	hypothetical protein, conserved	4	16	1.797E8	1.597	0.676
323	NCLIV_004350	Calcium homeostasis endoplasmic reticulum protein, related	1	1	2.215E7	1.595	0.673
324	NCLIV_025100	hypothetical protein	1	1	8.565E6	1.593	0.671
325	NCLIV_054940	hypothetical protein, conserved	3	6	4.860E7	1.588	0.667

326	NCLIV_042660	Probable cytosol aminopeptidase, related	8	53	8.409E8	1.585	0.664
327	NCLIV_040640	Forkhead-associated domain-containing protein, related	5	12	1.045E8	1.584	0.664
328	NCLIV_051350	hypothetical protein	2	9	9.700E7	1.583	0.663
329	NCLIV_030670	AGAP010490-PA, related	2	8	3.539E7	1.582	0.662
330	NCLIV_035160	hypothetical protein, conserved	1	4	4.522E7	1.581	0.661
331	NCLIV_056580	hypothetical protein	1	1	1.596E7	1.580	0.660
332	NCLIV_059540	hypothetical protein, conserved	1	3	1.586E7	1.579	0.659
333	NCLIV_024050	hypothetical protein, conserved	1	3	7.371E7	1.578	0.658
334	NCLIV_040470	hypothetical protein, conserved	1	1	2.388E6	1.577	0.657
335	NCLIV_048670	cDNA FLJ55447, highly similar to ATP-citrate synthase, related	1	4	3.119E8	1.575	0.655
336	NCLIV_040770	Small GTP-binding protein domain containing protein, related	1	1	5.942E6	1.573	0.653
337	NCLIV_022990	hypothetical protein, conserved	1	5	1.136E8	1.569	0.650
338	NCLIV_062810	hypothetical protein, conserved	12	42	2.625E8	1.568	0.649
339	NCLIV_015710	hypothetical protein, conserved	1	3	2.803E7	1.567	0.648
340	NCLIV_004100	nucleolar phosphoprotein p130, putative	1	3	2.358E8	1.566	0.647
341	NCLIV_011320	Vesicle-associated membrane protein-associated protein A, rel	5	29	1.062E9	1.565	0.646
342	NCLIV_017020	hypothetical protein	3	20	4.269E8	1.561	0.643
343	NCLIV_033870	Hypothetical low complexity protein, related	1	2	1.031E8	1.561	0.642
344	NCLIV_020510	Sporulation protein RMD1, related	3	15	2.158E7	1.558	0.640
345	NCLIV_041820	hypothetical protein	1	2	2.041E7	1.557	0.639
346	NCLIV_000030	hypothetical protein, conserved	1	2	7.484E7	1.554	0.636
347	NCLIV_048930	hypothetical protein	2	4	6.741E7	1.552	0.635
348	NCLIV_024360	hypothetical protein	9	27	2.133E8	1.544	0.627
349	NCLIV_028320	actin-like family protein ARP4a, putative	2	2	1.161E8	1.543	0.626
350	NCLIV_054310	hypothetical protein	3	6	3.263E8	1.542	0.624
351	NCLIV_041690	Ubiquitin carboxyl-terminal hydrolase (EC 3.1.2.15), related	9	29	1.858E8	1.541	0.624
352	NCLIV_013070	hypothetical protein, conserved	1	2	1.302E7	1.539	0.622
353	NCLIV_053670	AT hook motif-containing protein, putative	1	8	1.114E8	1.539	0.622
354	NCLIV_007680	hypothetical protein, conserved	1	4	4.959E7	1.538	0.621
355	NCLIV_023480	zinc finger (C3HC4 RING finger) protein, putative	1	3	3.805E6	1.538	0.621
356	NCLIV_061570	D-tyrosyl-tRNA(Tyr) deacylase (EC 3.1.-.-), related	1	2	5.114E7	1.536	0.619
357	NCLIV_063840	hypothetical protein, conserved	1	3	4.261E7	1.535	0.618
358	NCLIV_001570	Eukaryotic translation initiation factor 3 subunit G-2, related	3	21	6.399E8	1.534	0.618
359	NCLIV_002510	ATP-dependent RNA helicase DDX family protein, putative	1	3	5.618E6	1.533	0.617
360	NCLIV_035850	hypothetical protein	5	20	2.443E8	1.533	0.616
361	NCLIV_017570	amine oxidase, flavin-containing domain- containing protein, put	1	5	1.288E9	1.532	0.616
362	NCLIV_057470	hypothetical protein, conserved	7	22	1.849E8	1.531	0.615
363	NCLIV_006510	TCP-1/cpn60 family chaperonin, putative	2	8	2.945E7	1.530	0.613
364	NCLIV_003680	40S ribosomal protein S28, related	3	12	1.641E8	1.529	0.613
365	NCLIV_019480	proteasome A-type and B-type domain-containing protein, put	2	8	5.688E8	1.528	0.612
366	NCLIV_012090	ATP-dependent protease ATP-binding subunit, putative	1	3	1.696E7	1.525	0.609
367	NCLIV_045300	Chloroquine resistance marker protein, related	10	25	1.160E8	1.523	0.607
368	NCLIV_004090	hypothetical protein	1	3	4.487E8	1.522	0.606

369	NCLIV_070040	RNA polymerase II accessory factor Cdc73p, related	1	2	5.358E7	1.516	0.600
370	NCLIV_013840	hypothetical protein	11	39	2.645E8	1.513	0.598
371	NCLIV_018520	DNA replication licensing factor, putative	2	11	1.255E8	1.512	0.597
372	NCLIV_062980	phosphatidylinositol-4-phosphate 5-kinase, putative	1	2	3.546E7	1.511	0.596
373	NCLIV_046380	syntaxin, putative	1	8	1.159E8	1.508	0.593
374	NCLIV_014190	Kynurenine 3-monooxygenase and related flavoprotein monooxygenases (ISS), related	1	4	2.611E7	1.505	0.590
375	NCLIV_027020	hypothetical protein, conserved	2	3	2.893E7	1.505	0.590
376	NCLIV_004130	hypothetical protein	1	2	5.321E6	1.502	0.587
377	NCLIV_061460	Family T1, proteasome beta subunit, threonine peptidase, rel	1	4	4.084E8	1.502	0.587
378	NCLIV_043930	Kelch repeat proteins that is fused to a HSP90-like ATPase, rel	19	70	5.733E8	1.501	0.586
379	NCLIV_066400	Os02g0819100 protein, related	1	4	1.288E7	1.499	0.584
380	NCLIV_006980	hypothetical protein	2	5	5.676E7	1.498	0.583
381	NCLIV_028630	hypothetical protein, conserved	1	1	1.420E7	1.498	0.583
382	NCLIV_033000	hypothetical protein, conserved	3	8	3.891E7	1.497	0.582
383	NCLIV_019100	hypothetical protein, conserved	3	4	9.175E7	1.497	0.582
384	NCLIV_039400	hypothetical protein	12	93	1.920E9	1.495	0.580
385	NCLIV_020250	hypothetical protein	6	27	4.057E8	1.495	0.580
386	NCLIV_024420	hypothetical protein	7	35	1.239E9	1.495	0.580
387	NCLIV_007850	hypothetical protein	2	4	2.612E7	1.494	0.579
388	NCLIV_057640	mRNA splicing factor 3bA, related	4	11	2.528E8	1.494	0.579
389	NCLIV_046550	Elongation factor 1-beta, related	2	20	6.365E8	1.491	0.577
390	NCLIV_023980	hypothetical protein, conserved	4	6	1.046E8	1.491	0.576
391	NCLIV_056440	Signal recognition particle GTPase, related	1	4	5.014E7	1.490	0.575
392	NCLIV_001520	Eukaryotic translation initiation factor 3 subunit C, related	13	67	1.135E9	1.484	0.569
393	NCLIV_069290	hypothetical protein, conserved	2	8	4.976E7	1.483	0.568
394	NCLIV_064260	WD domain-containing protein, putative	1	5	5.494E8	1.482	0.568
395	NCLIV_059620	hypothetical protein, conserved	6	19	6.317E8	1.481	0.567
396	NCLIV_011110	DEAH-box RNA/DNA helicase, putative	8	25	2.686E8	1.480	0.565
397	NCLIV_013470	Spermine/spermidine synthase family protein, related	3	4	1.067E7	1.476	0.561
398	NCLIV_048880	Proteasome subunit beta type-7, related	1	5	8.167E7	1.475	0.561
399	NCLIV_068130	hypothetical protein, conserved	2	7	8.308E7	1.474	0.560
400	NCLIV_066360	hypothetical protein, conserved	13	37	3.722E8	1.473	0.559
401	NCLIV_053360	hypothetical protein, conserved	7	15	1.100E8	1.471	0.557
402	NCLIV_028940	Transformer-2-like protein, related	3	9	2.623E8	1.469	0.555
403	NCLIV_027290	Ribosomal protein S21-maize (ISS), related	2	11	1.782E8	1.468	0.553
404	NCLIV_064790	hypothetical protein	1	8	2.673E8	1.466	0.552
405	NCLIV_011820	AGAP011504-PA, related	4	39	1.941E9	1.465	0.551
406	NCLIV_067550	hypothetical protein	2	8	5.639E7	1.464	0.550
407	NCLIV_027070	hypothetical protein, conserved	2	6	8.324E6	1.462	0.548
408	NCLIV_037870	hypothetical protein, conserved	2	4	4.852E7	1.458	0.544
409	NCLIV_027230	Dihydropteroate synthase, related	7	23	3.532E8	1.456	0.542
410	NCLIV_017190	hemolysin, putative	1	4	5.277E7	1.456	0.542

411	NCLIV_019240	hypothetical protein, conserved	2	7	5.223E7	1.453	0.539
412	NCLIV_000430	hypothetical protein, conserved	11	38	1.125E9	1.453	0.539
413	NCLIV_015910	hypothetical protein	2	5	5.782E7	1.450	0.536
414	NCLIV_012910	Novel protein (Zgc:63649), related	6	18	1.603E8	1.450	0.536
415	NCLIV_057340	hypothetical protein	1	1	2.275E7	1.449	0.535
416	NCLIV_007120	hypothetical protein	1	5	5.522E7	1.447	0.533
417	NCLIV_020200	integral membrane protein, putative	2	3	1.543E7	1.441	0.527
418	NCLIV_013170	WD domain-containing protein, putative	9	18	1.178E8	1.441	0.527
419	NCLIV_030550	hypothetical protein, conserved	2	12	7.215E7	1.440	0.526
420	NCLIV_041900	Phosphoenolpyruvate carboxykinase (ATP), related	2	15	3.371E8	1.438	0.525
421	NCLIV_061070	hypothetical protein	3	5	1.976E8	1.437	0.523
422	NCLIV_066900	serine/threonine protein phosphatase 5, putative	12	56	4.434E8	1.436	0.522
423	NCLIV_016120	proteasome subunit alpha type 4, subunit, putative	2	8	3.160E7	1.432	0.518
424	NCLIV_030430	TPR domain-containing protein, putative	2	6	1.578E8	1.432	0.518
425	NCLIV_045290	cDNA FLJ30053 fis, clone ADRGL1000144, highly similar to StAR-related lipid transfer protein 7, related	1	4	1.427E8	1.428	0.514
426	NCLIV_004250	nuclear RNA binding protein, putative	2	5	1.945E8	1.426	0.512
427	NCLIV_033770	corA-like Mg2 transporter domain-containing protein, putative	1	3	4.028E7	1.425	0.511
428	NCLIV_011740	hypothetical protein, conserved	2	6	7.542E7	1.424	0.510
429	NCLIV_015440	hypothetical protein	8	41	4.448E8	1.424	0.509
430	NCLIV_031730	translational activator, putative	11	42	2.466E8	1.421	0.507
431	NCLIV_008850	hypothetical protein, conserved	5	19	1.139E8	1.420	0.506
432	NCLIV_042330	hypothetical protein, conserved	1	2	1.102E7	1.419	0.505
433	NCLIV_021650	hypothetical protein	1	2	4.661E7	1.418	0.504
434	NCLIV_051420	KLLA0B04587p, related	1	1	2.782E6	1.417	0.503
435	NCLIV_052370	hypothetical protein	1	8	1.346E8	1.417	0.503
436	NCLIV_062310	human SRp75 can complement a splicing- deficient extract, rel	5	17	2.418E8	1.416	0.502
437	NCLIV_038620	protein phosphatases pp1 regulatory subunit, putative	1	1	3.407E6	1.415	0.501
438	NCLIV_014990	hypothetical protein	2	6	6.912E7	1.414	0.500
439	NCLIV_058700	FF domain-containing protein, putative	1	2	5.641E6	1.413	0.499
440	NCLIV_004490	hypothetical protein	3	12	5.030E7	1.413	0.498
441	NCLIV_020810	hypothetical protein, conserved	1	7	1.741E9	1.410	0.496
442	NCLIV_065710	hypothetical protein, conserved	1	1	6.085E6	1.408	0.494
443	NCLIV_034130	hypothetical protein	6	34	9.487E8	1.407	0.492
444	NCLIV_045190	Proteophosphoglycan ppg1, related	6	29	4.856E8	1.406	0.492
445	NCLIV_034800	hypothetical protein, conserved	5	25	1.468E9	1.404	0.489
446	NCLIV_002430	hypothetical protein, conserved	2	5	7.216E7	1.403	0.489
447	NCLIV_043320	RNA recognition motif (RRM)-containing protein, related	3	6	1.107E8	1.401	0.486
448	NCLIV_020400	dual-specificity phosphatase laforin, putative	1	2	6.120E7	1.401	0.486
449	NCLIV_041500	glutamic acid-rich protein, putative	1	5	3.192E7	1.398	0.483
450	NCLIV_017470	Proteasome subunit alpha type (EC 3.4.25.1), related	4	16	6.893E8	1.397	0.482
451	NCLIV_060260	Trichohyalin, putative	1	2	2.765E7	1.395	0.480
452	NCLIV_059600	KH domain-containing protein, putative	22	138	1.320E9	1.394	0.479

453	NCLIV_011250	TAF5-like RNA polymerase II, p300/CBP-associated factor (PCAF)-associated factor, 65kDa, related	2	3	1.613E8	1.394	0.479
454	NCLIV_043640	Transmembrane protein 184A, related	2	5	3.255E7	1.394	0.479
455	NCLIV_016860	hypothetical protein, conserved	1	3	2.487E7	1.390	0.475
456	NCLIV_065770	nuclear cap binding protein, putative	3	11	1.207E8	1.389	0.474
457	NCLIV_051500	hypothetical protein	1	2	4.519E7	1.389	0.474
458	NCLIV_020160	hypothetical protein	2	15	8.745E7	1.386	0.471
459	NCLIV_063290	hypothetical protein	1	5	6.932E7	1.384	0.469
460	NCLIV_048550	26S proteasome non-ATPase regulatory subunit 4, putative	4	15	1.320E8	1.383	0.468
461	NCLIV_019780	KH domain containing protein, putative	9	48	1.083E9	1.379	0.464
462	NCLIV_010870	Probable RNA-binding protein 19, related	1	3	1.125E7	1.379	0.464
463	NCLIV_050000	Probable deoxyhypusine synthase (EC 2.5.1.46), related	2	9	6.458E7	1.379	0.463
464	NCLIV_063250	hypothetical protein, conserved	2	10	7.883E7	1.378	0.463
465	NCLIV_061340	hypothetical protein	1	5	1.127E8	1.377	0.462
466	NCLIV_041710	hypothetical protein	3	10	4.983E7	1.377	0.462
467	NCLIV_006420	hypothetical protein, conserved	1	5	3.239E7	1.375	0.459
468	NCLIV_008120	hypothetical protein, conserved	29	83	2.139E8	1.372	0.457
469	NCLIV_028490	hypothetical protein, conserved	1	4	1.255E8	1.370	0.454
470	NCLIV_024140	hypothetical protein	1	1	5.358E7	1.368	0.452
471	NCLIV_001890	hypothetical protein	2	6	1.718E8	1.367	0.451
472	NCLIV_000050	hypothetical protein, conserved	1	5	6.878E6	1.365	0.449
473	NCLIV_013660	regulator of chromosome condensation, putative	7	17	2.148E8	1.365	0.449
474	NCLIV_016890	hypothetical protein	1	1	1.060E6	1.363	0.447
475	NCLIV_028610	hypothetical protein	3	12	1.084E8	1.362	0.446
476	NCLIV_008660	adenylate and guanylate cyclase catalytic domain-containing protein, putative	7	16	1.152E8	1.361	0.445
477	NCLIV_019460	hypothetical protein	1	4	1.333E8	1.360	0.444
478	NCLIV_068960	hypothetical protein	1	4	4.679E7	1.360	0.444
479	NCLIV_007500	hypothetical protein, conserved	1	1	4.518E7	1.358	0.441
480	NCLIV_048170	Articulon family protein, related	2	9	3.362E8	1.357	0.441
481	NCLIV_047110	hypothetical protein, conserved	8	32	2.237E8	1.357	0.440
482	NCLIV_065020	actin-like protein 3b, putative	1	5	1.047E8	1.356	0.439
483	NCLIV_005050	coatamer delta subunit, putative	4	13	1.620E8	1.356	0.439
484	NCLIV_005900	TRANSLATION INITIATION FACTOR 3 SUBUNIT 9-like protein, rel	6	27	6.618E8	1.356	0.439
485	NCLIV_043180	hypothetical protein	7	24	6.017E8	1.355	0.438
486	NCLIV_053810	Os03g0795800 protein, related	3	14	1.068E8	1.354	0.437
487	NCLIV_005950	trafficking protein particle complex protein, putative	1	1	6.041E7	1.353	0.436
488	NCLIV_041660	hypothetical protein	2	3	3.269E7	1.353	0.436
489	NCLIV_051240	protein phosphatase 2C, putative	3	8	6.238E7	1.351	0.434
490	NCLIV_046520	hypothetical protein	1	5	1.502E7	1.350	0.433
491	NCLIV_042530	chloride channel protein, putative	1	2	2.022E7	1.350	0.433
492	NCLIV_068000	hypothetical protein, conserved	4	14	3.925E7	1.350	0.433
493	NCLIV_005200	hypothetical protein	2	8	2.452E8	1.350	0.432

494	NCLIV_012130	Eukaryotic translation initiation factor 3 subunit 7-like protein, related	3	22	5.104E8	1.347	0.430
495	NCLIV_004580	hypothetical protein	1	5	3.272E7	1.346	0.428
496	NCLIV_023760	hypothetical protein, conserved	1	2	6.274E7	1.344	0.427
497	NCLIV_013360	plectin, putative	11	27	2.708E8	1.343	0.425
498	NCLIV_011950	hypothetical protein, conserved	1	1	5.167E7	1.342	0.424
499	NCLIV_028360	hypothetical protein, conserved	3	6	4.376E8	1.342	0.424
500	NCLIV_038700	hypothetical protein	2	7	2.652E8	1.341	0.423
501	NCLIV_033950	Heat shock protein 70, related	17	215	4.453E9	1.340	0.422
502	NCLIV_034330	mpv17 / PMP22 family domain-containing protein, putative	1	5	2.762E7	1.338	0.420
503	NCLIV_051530	hypothetical protein	2	8	2.221E8	1.337	0.419
504	NCLIV_066390	CMGC kinase, putative	1	7	4.998E7	1.335	0.417
505	NCLIV_014750	hypothetical protein	4	7	6.556E7	1.335	0.416
506	NCLIV_015970	Ribose 5-phosphate isomerase A (EC 5.3.1.6), related	1	2	4.081E7	1.334	0.416
507	NCLIV_032830	hypothetical protein	3	8	1.893E8	1.334	0.416
508	NCLIV_013790	hypothetical protein	10	42	2.595E8	1.334	0.415
509	NCLIV_058430	hypothetical protein, conserved	5	10	2.795E7	1.334	0.415
510	NCLIV_017010	Lysophosphatidic acid acyltransferase, related	2	6	2.182E7	1.333	0.414
511	NCLIV_025000	hypothetical protein	3	14	2.574E8	1.330	0.412
512	NCLIV_001460	ROP37 (incomplete catalytic triad), putative	1	2	1.196E8	1.330	0.412
513	NCLIV_068380	hypothetical protein	3	22	2.266E8	1.329	0.410
514	NCLIV_043020	hypothetical protein	6	34	1.189E8	1.328	0.410
515	NCLIV_063180	ARID/BRIGHT DNA-binding domain-containing protein, putative	6	19	1.385E8	1.327	0.408
516	NCLIV_026880	hypothetical protein, conserved	1	1	3.096E7	1.325	0.406
517	NCLIV_060830	Similar to Q03338 S cerevisiae YDR473c splicing factor, related	4	10	4.656E7	1.325	0.406
518	NCLIV_006150	tRNA binding domain containing protein, related	3	14	1.502E8	1.325	0.406
519	NCLIV_038510	hypothetical protein	2	5	1.370E7	1.323	0.403
520	NCLIV_059190	hypothetical protein	1	1	7.986E7	1.323	0.403
521	NCLIV_014660	hypothetical protein	1	4	9.138E7	1.322	0.403
522	NCLIV_069360	hypothetical protein	2	17	1.044E8	1.322	0.403
523	NCLIV_039670	200 kDa antigen p200, putative	4	11	1.430E8	1.322	0.403
524	NCLIV_032180	GK24228, related	3	9	5.766E7	1.321	0.402
525	NCLIV_012470	FAD-dependent monooxygenase, putative	4	10	8.323E7	1.320	0.400
526	NCLIV_049920	hypothetical protein, conserved	2	6	1.626E8	1.319	0.399
527	NCLIV_042790	2OG-Fe(II) oxygenase family protein, putative	2	4	3.860E7	1.318	0.398
528	NCLIV_040680	hypothetical protein, conserved	4	9	3.348E7	1.318	0.398
529	NCLIV_064880	hypothetical protein, conserved	12	77	1.273E9	1.317	0.398
530	NCLIV_030490	30S ribosomal protein S12, related	2	14	1.771E9	1.316	0.396
531	NCLIV_042370	hypothetical protein	2	6	3.077E8	1.315	0.395
532	NCLIV_067980	Proteasome subunit alpha type (EC 3.4.25.1), related	2	10	3.750E8	1.315	0.395
533	NCLIV_016420	pescadillo family protein, putative	1	4	7.321E6	1.315	0.395
534	NCLIV_018400	TCP-1/cpn60 family chaperonin, putative	5	17	3.904E8	1.315	0.395
535	NCLIV_060040	hypothetical protein	2	4	4.882E7	1.312	0.392

536	NCLIV_027820	5-aminolevulinic acid synthase, related	1	2	1.514E7	1.311	0.390
537	NCLIV_018650	hypothetical protein	1	5	1.277E7	1.311	0.390
538	NCLIV_059920	eukaryotic initiation factor-2B gamma subunit, putative	4	13	2.746E8	1.310	0.390
539	NCLIV_040800	hypothetical protein, conserved	2	2	1.163E7	1.310	0.390
540	NCLIV_033620	hypothetical protein, conserved	2	8	6.646E7	1.309	0.388
541	NCLIV_045120	elongation factor Tu GTP binding domain- containing protein, put	2	3	1.174E7	1.307	0.387
542	NCLIV_029610	prefoldin subunit 3, putative	1	3	4.434E7	1.307	0.386
543	NCLIV_046720	hypothetical protein, conserved	2	7	1.607E8	1.305	0.384
544	NCLIV_001120	XPA-binding protein, putative	2	11	5.191E8	1.304	0.383
545	NCLIV_045500	hypothetical protein, conserved	2	7	2.757E7	1.303	0.382
546	NCLIV_029920	Receptor expression-enhancing protein 3, related	1	1	2.812E8	1.303	0.382
547	NCLIV_066060	MAK16 protein, putative	1	2	6.789E7	1.303	0.382
548	NCLIV_045240	eukaryotic translation initiation factor 3 subunit 5, putative	4	28	3.642E8	1.303	0.382
549	NCLIV_051550	hypothetical protein, conserved	1	2	4.278E6	1.302	0.380
550	NCLIV_021940	hypothetical protein, conserved	2	8	8.003E7	1.302	0.380
551	NCLIV_007590	actin-like family protein ARP4a, putative	2	9	6.344E7	1.300	0.379
552	NCLIV_025960	nucleolar GTP-binding protein, putative	1	3	2.405E7	1.300	0.379
553	NCLIV_060910	DNA replication licensing factor, putative	3	4	5.166E7	1.300	0.379
554	NCLIV_061040	hypothetical protein	5	61	1.792E9	1.299	0.377
555	NCLIV_023660	Phosphatidate cytidyltransferase (EC 2.7.7.41), related	1	2	5.966E7	1.299	0.377
556	NCLIV_009450	60S ribosomal protein L17, related	3	46	3.108E9	1.295	0.373
557	NCLIV_009630	cyclophilin, putative	1	1	6.267E7	1.295	0.373
558	NCLIV_013280	hypothetical protein, conserved	1	6	2.024E8	1.295	0.373
559	NCLIV_059730	hypothetical protein, conserved	4	14	8.406E7	1.295	0.373
560	NCLIV_002330	hypothetical protein, conserved	2	3	6.286E7	1.295	0.373
561	NCLIV_058260	hypothetical protein	1	2	1.698E7	1.292	0.370
562	NCLIV_061320	hypothetical protein, conserved	5	19	5.850E7	1.291	0.368
563	NCLIV_052520	hypothetical protein, conserved	2	2	3.094E7	1.290	0.367
564	NCLIV_021440	RNA binding protein, putative	1	5	8.875E7	1.288	0.365
565	NCLIV_053750	Smc ABC ATPase, related	1	4	9.546E7	1.286	0.363
566	NCLIV_001800	leucine rich repeat protein, putative	2	8	8.666E7	1.286	0.363
567	NCLIV_059040	hypothetical protein	2	5	4.141E7	1.285	0.362
568	NCLIV_011130	30S ribosomal protein S16, related	1	3	4.543E7	1.283	0.359
569	NCLIV_002670	hypothetical protein	3	8	6.303E7	1.282	0.359
570	NCLIV_011790	Cysteine desulfurase (EC 2.8.1.7), related	1	9	3.080E8	1.281	0.357
571	NCLIV_009850	SNF2 family N-terminal domain-containing protein, related	4	15	2.449E8	1.278	0.354
572	NCLIV_025400	4-hydroxy-3-methylbut-2-en-1-y	2	4	2.847E7	1.277	0.353
573	NCLIV_008080	hypothetical protein	6	25	3.121E8	1.276	0.352
574	NCLIV_053720	adaptin, putative	2	3	4.920E7	1.275	0.351
575	NCLIV_008100	importin-alpha re-exporter, putative	2	5	9.635E7	1.275	0.350
576	NCLIV_023100	hypothetical protein, conserved	2	7	2.446E8	1.274	0.350
577	NCLIV_060930	Pyruvate dehydrogenase kinase, isoenzyme 2, related	1	4	2.112E7	1.273	0.349
578	NCLIV_005750	hypothetical protein, conserved	2	3	2.487E7	1.272	0.347

579	NCLIV_033030	hypothetical protein, conserved	11	40	4.178E8	1.271	0.346
580	NCLIV_045410	ATP-dependent Clp protease proteolytic subunit, related	1	6	2.856E7	1.271	0.346
581	NCLIV_042430	hypothetical protein, conserved	1	1	2.419E7	1.269	0.344
582	NCLIV_064580	hypothetical protein	7	36	1.457E9	1.268	0.343
583	NCLIV_030660	TCP-1/cpn60 family chaperonin, putative	7	28	7.781E8	1.268	0.342
584	NCLIV_030890	high molecular mass nuclear antigen, putative	6	47	1.201E9	1.268	0.342
585	NCLIV_047130	hypothetical protein	1	2	1.896E7	1.267	0.341
586	NCLIV_032880	hypothetical protein	1	1	2.908E7	1.266	0.340
587	NCLIV_043100	zinc finger MYND domain-containing protein, putative	1	1	1.093E7	1.264	0.338
588	NCLIV_001820	hypothetical protein, conserved	1	2	1.307E8	1.264	0.338
589	NCLIV_031510	hypothetical protein	3	26	1.144E9	1.263	0.337
590	NCLIV_015280	mRNA processing protein, putative	2	6	1.117E8	1.261	0.334
591	NCLIV_001930	hypothetical protein, conserved	2	9	1.488E8	1.261	0.334
592	NCLIV_046060	hypothetical protein, conserved	1	3	6.772E7	1.260	0.333
593	NCLIV_031940	hypothetical protein, conserved	1	1	1.056E7	1.259	0.333
594	NCLIV_014540	hypothetical protein	2	5	1.442E7	1.259	0.332
595	NCLIV_017650	hypothetical protein	7	23	9.724E7	1.259	0.332
596	NCLIV_012770	Tyrosine kinase-like (TKL) protein, putative	2	11	6.888E8	1.259	0.332
597	NCLIV_008310	hypothetical protein, conserved	2	11	1.688E8	1.258	0.331
598	NCLIV_002530	UPF0027 protein PH1602, related	4	9	6.387E7	1.258	0.331
599	NCLIV_066620	hypothetical protein, conserved	1	1	3.287E7	1.258	0.331
600	NCLIV_036590	hypothetical protein, conserved	1	1	8.330E5	1.257	0.330
601	NCLIV_008150	ribosomal protein L15, putative	1	2	3.857E7	1.257	0.330
602	NCLIV_065180	GA28568, related	5	8	7.057E7	1.256	0.329
603	NCLIV_028990	hypothetical protein	1	4	3.342E7	1.256	0.328
604	NCLIV_032360	WD-40 repeat protein, putative	3	8	2.069E7	1.254	0.327
605	NCLIV_060880	hypothetical protein, conserved	2	5	3.455E6	1.252	0.324
606	NCLIV_024860	Proteasome/cyclosome repeat family protein, related	1	11	3.404E7	1.252	0.324
607	NCLIV_047290	BIS(5'-nucleosyl)-tetrphosphatase, putative	1	4	2.130E8	1.251	0.324
608	NCLIV_051760	hypothetical protein, conserved	3	9	4.994E7	1.251	0.323
609	NCLIV_060390	SNARE domain-containing protein, putative	2	3	6.084E7	1.251	0.323
610	NCLIV_053910	hypothetical protein, conserved	1	2	3.556E7	1.247	0.318
611	NCLIV_045320	hypothetical protein, conserved	2	7	6.917E8	1.246	0.318
612	NCLIV_068950	hypothetical protein	1	2	5.715E7	1.244	0.315
613	NCLIV_008920	hypothetical protein	1	2	4.185E7	1.243	0.313
614	NCLIV_022280	vacuolar protein sorting 29, putative	3	5	1.515E8	1.243	0.313
615	NCLIV_045980	hypothetical protein, conserved	1	6	1.267E8	1.242	0.313
616	NCLIV_056300	hypothetical protein, conserved	3	14	1.679E8	1.240	0.311
617	NCLIV_066370	hypothetical protein	2	10	2.707E8	1.240	0.311
618	NCLIV_057190	hypothetical protein, conserved	4	16	1.288E9	1.240	0.310
619	NCLIV_002660	hypothetical protein	2	8	9.124E7	1.240	0.310
620	NCLIV_014050	hypothetical protein, conserved	3	24	5.523E8	1.238	0.309
621	NCLIV_049150	hypothetical protein, conserved	9	39	2.330E8	1.237	0.307

622	NCLIV_035940	hypothetical protein	4	12	5.169E8	1.237	0.307
623	NCLIV_057050	hypothetical protein, conserved	1	2	2.191E7	1.236	0.305
624	NCLIV_027850	rhostry protein 6, putative	4	23	2.050E8	1.234	0.303
625	NCLIV_004140	hypothetical protein	2	32	8.023E8	1.233	0.302
626	NCLIV_047230	hypothetical protein, conserved	3	10	2.690E7	1.233	0.302
627	NCLIV_065750	Alpha 2 subunit of 20S proteasome (ISS), related	7	15	3.026E8	1.232	0.302
628	NCLIV_034190	eukaryotic translation initiation factor 3 sub 6 interact. protein, put	8	45	8.574E8	1.231	0.300
629	NCLIV_013290	hypothetical protein, conserved	1	8	8.618E7	1.230	0.299
630	NCLIV_010850	MGC81714 protein, related	2	6	1.241E7	1.230	0.298
631	NCLIV_002190	hypothetical protein, conserved	1	7	1.245E8	1.229	0.298
632	NCLIV_038170	Hypoxanthine phosphoribosyltransferase (EC 2.4.2.8), related	3	5	2.511E8	1.229	0.298
633	NCLIV_067380	hypothetical protein, conserved	3	5	1.554E8	1.228	0.296
634	NCLIV_042720	WD-40 repeat protein, putative	1	8	1.535E8	1.228	0.296
635	NCLIV_030170	hypothetical protein	2	6	1.186E8	1.227	0.295
636	NCLIV_008790	hypothetical protein, conserved	2	3	9.388E7	1.226	0.294
637	NCLIV_062630	Thioredoxin-dependent peroxide reductase, mitochondrial, rel	6	81	9.892E9	1.226	0.294
638	NCLIV_016040	hypothetical protein, conserved	8	31	1.822E8	1.226	0.294
639	NCLIV_046670	AGAP008757-PA, related	1	17	2.235E7	1.225	0.293
640	NCLIV_024850	PrsS1, related	4	13	1.470E8	1.224	0.291
641	NCLIV_006140	Ubiquitin, related	1	1	3.876E6	1.223	0.290
642	NCLIV_031010	Phosphatidylinositol-4-phosphate 5-kinase family protein, related	2	7	1.331E8	1.223	0.290
643	NCLIV_055260	AGC kinase TgPKG1, putative	2	10	1.753E8	1.222	0.290
644	NCLIV_052280	hypothetical protein, conserved	3	12	5.793E8	1.222	0.289
645	NCLIV_023050	eukaryotic translation initiation factor 3 subunit 10, putative	14	64	8.674E8	1.221	0.289
646	NCLIV_006470	hypothetical protein, conserved	1	2	7.735E7	1.220	0.287
647	NCLIV_028540	hypothetical protein, conserved	4	35	7.877E8	1.220	0.287
648	NCLIV_026710	hypothetical protein, conserved	5	15	6.755E7	1.220	0.287
649	NCLIV_024480	hypothetical protein, conserved	2	3	8.458E7	1.219	0.286
650	NCLIV_020360	40S ribosomal protein S12, related	4	22	5.465E8	1.216	0.282
651	NCLIV_037500	Enolase (EC 4.2.1.11), related	8	38	1.271E9	1.216	0.282
652	NCLIV_037780	zinc finger (CCCH type) protein, putative	2	7	4.976E8	1.215	0.281
653	NCLIV_005770	hypothetical protein	4	10	1.759E8	1.214	0.279
654	NCLIV_001090	CorA-like Mg2 transporter domain-containing protein, related	3	9	1.774E8	1.212	0.278
655	NCLIV_053580	50S ribosomal protein L4P, related	12	103	3.404E9	1.212	0.277
656	NCLIV_042190	hypothetical protein, conserved	3	12	8.804E7	1.211	0.277
657	NCLIV_018050	hypothetical protein, conserved	5	13	1.582E8	1.211	0.276
658	NCLIV_036230	Ubiquitin fusion-degradation protein (ISS), related	2	5	8.047E7	1.211	0.276
659	NCLIV_046770	transmembrane amino acid transporter, putative	5	18	2.906E8	1.208	0.273
660	NCLIV_042860	hypothetical protein, conserved	1	3	1.329E7	1.207	0.272
661	NCLIV_012660	hypothetical protein, conserved	7	30	6.043E8	1.205	0.270
662	NCLIV_011930	ribosome biogenesis protein BOP1, putative	1	1	1.639E8	1.205	0.269
663	NCLIV_008030	hypothetical protein, conserved	4	53	3.291E9	1.204	0.268
664	NCLIV_053640	peroxidoxin 2, putative	3	12	4.713E8	1.204	0.268

665	NCLIV_041780	LSU ribosomal protein L19E, related	3	22	2.433E8	1.203	0.266
666	NCLIV_070100	GF20490, related	3	6	8.010E7	1.202	0.266
667	NCLIV_062570	Contig An13c0020, complete genome, related	3	11	4.031E8	1.201	0.265
668	NCLIV_059960	hypothetical protein, conserved	4	15	1.000E8	1.201	0.265
669	NCLIV_048180	hypothetical protein, conserved	2	4	1.972E7	1.201	0.264
670	NCLIV_045440	proteasome PCI domain-containing protein, putative	5	16	3.477E8	1.201	0.264
671	NCLIV_033350	UDP-glucose 4-epimerase Gal10, related	1	1	1.104E8	1.200	0.263
672	NCLIV_045550	RNA recognition motif-containing protein, putative	1	6	5.441E7	1.199	0.261
673	NCLIV_050870	SNF2 family N-terminal domain containing protein, putative	3	7	2.683E8	1.197	0.260
674	NCLIV_039680	hypothetical protein, conserved	2	14	1.026E8	1.197	0.259
675	NCLIV_053820	DEAD box RNA helicase, related	4	8	1.581E8	1.197	0.259
676	NCLIV_060380	hypothetical protein	1	11	2.228E8	1.196	0.258
677	NCLIV_039520	hypothetical protein, conserved	2	9	8.817E7	1.196	0.258
678	NCLIV_034530	TCP-1/cpn60 family chaperonin, putative	5	18	2.487E8	1.196	0.258
679	NCLIV_006740	armadillo/beta-catenin-like repeat-containing protein, putative	1	9	4.348E8	1.196	0.258
680	NCLIV_040840	UBX domain-containing protein, putative	3	13	1.783E8	1.196	0.258
681	NCLIV_019220	hypothetical protein, conserved	2	7	2.720E7	1.195	0.257
682	NCLIV_039910	hypothetical protein, conserved	1	4	1.641E8	1.195	0.257
683	NCLIV_056020	hypothetical protein	3	20	1.724E8	1.195	0.257
684	NCLIV_063510	GH23102, related	3	5	1.149E8	1.195	0.257
685	NCLIV_067260	Diaminopimelate decarboxylase protein, related	3	11	8.687E7	1.194	0.256
686	NCLIV_040900	hypothetical protein	3	23	2.910E8	1.193	0.255
687	NCLIV_004280	hypothetical protein	2	17	1.403E8	1.190	0.251
688	NCLIV_039080	Adaptin N terminal region family protein, related	12	45	4.358E8	1.189	0.250
689	NCLIV_027330	Eukaryotic Na /H exchanger, related	1	3	6.096E7	1.189	0.250
690	NCLIV_065380	Ras family domain-containing protein, putative	3	10	1.191E8	1.188	0.249
691	NCLIV_034160	Glutamine synthetase (EC 6.3.1.2), related	3	8	1.587E8	1.186	0.246
692	NCLIV_069350	hypothetical protein	1	8	8.129E7	1.186	0.246
693	NCLIV_021970	hypothetical protein, conserved	1	6	9.836E7	1.184	0.244
694	NCLIV_053190	hypothetical protein	2	7	1.846E8	1.184	0.244
695	NCLIV_066290	hypothetical protein	1	3	6.466E6	1.184	0.243
696	NCLIV_038330	adenylosuccinate synthetase, putative	1	2	8.549E7	1.183	0.242
697	NCLIV_028870	Peptidylprolyl isomerase D (Cyclophilin D), related	7	29	4.253E8	1.182	0.241
698	NCLIV_067050	hypothetical protein, conserved	2	18	2.736E8	1.182	0.241
699	NCLIV_000300	hypothetical protein, conserved	2	14	2.577E8	1.182	0.241
700	NCLIV_065240	Fructose-1 6-biphosphatase, related	7	56	1.323E9	1.181	0.241
701	NCLIV_057670	myosin G, putative	6	14	2.072E8	1.181	0.240
702	NCLIV_054520	hypothetical protein	5	42	1.603E9	1.181	0.240
703	NCLIV_049270	hypothetical protein, conserved	3	13	5.380E7	1.180	0.238
704	NCLIV_028710	Os07g0659500 protein, related	1	5	3.764E7	1.179	0.238
705	NCLIV_057170	hypothetical protein	2	5	2.959E7	1.179	0.237
706	NCLIV_014690	dihydrolipoamide acyltransferase, putative	1	6	1.921E8	1.178	0.237
707	NCLIV_019540	signal recognition particle protein, putative	1	4	1.948E8	1.178	0.237

708	NCLIV_007630	ATP-dependent RNA helicase, putative	1	5	2.646E7	1.176	0.233
709	NCLIV_050160	cDNA, FLJ92180, highly similar to brain protein BRP44, related	2	5	2.011E7	1.175	0.233
710	NCLIV_038000	CUG-BP- and ETR-3-like factor 3, related	4	29	4.517E8	1.174	0.232
711	NCLIV_044820	hypothetical protein, conserved	2	9	1.591E8	1.171	0.228
712	NCLIV_061820	Testis cDNA:QtsA14439 similar human pleiotropic regulator 1,rel	2	6	1.549E8	1.171	0.228
713	NCLIV_063420	hypothetical protein, conserved	1	4	1.693E8	1.169	0.226
714	NCLIV_052390	hypothetical protein	7	68	1.089E9	1.169	0.225
715	NCLIV_006850	Hornerin, related	5	24	1.099E8	1.168	0.224
716	NCLIV_010630	cDNA FLJ53078 similar toSplicing factor, arginine/serine-rich1,rel	3	8	1.155E8	1.166	0.222
717	NCLIV_018700	SnoRNA binding domain, related	2	14	5.411E8	1.165	0.221
718	NCLIV_020830	endonuclease/exonuclease/phosphatase containing protein, put	1	3	1.435E7	1.165	0.220
719	NCLIV_034070	hypothetical protein, conserved	1	2	4.123E8	1.165	0.220
720	NCLIV_021740	ATP-dependent RNA helicase, putative	2	5	6.144E7	1.164	0.219
721	NCLIV_052220	hypothetical protein	1	3	2.501E8	1.164	0.219
722	NCLIV_041680	Casein kinase one (CK1) TgCK1b, putative	5	13	1.914E8	1.164	0.219
723	NCLIV_010320	dihydrolipoamide branched chain transacylase, E2 subunit, put	8	58	2.449E9	1.163	0.217
724	NCLIV_066030	mitochondria-associated granulocyte macrophage CSF signaling molecule, putative	1	3	8.455E7	1.162	0.216
725	NCLIV_056030	hypothetical protein, conserved	1	1	4.783E7	1.161	0.216
726	NCLIV_009540	hypothetical protein, conserved	1	3	8.768E6	1.160	0.215
727	NCLIV_062080	hypothetical protein, conserved	3	9	8.599E7	1.160	0.214
728	NCLIV_032390	hypothetical protein, conserved	10	169	9.212E9	1.158	0.212
729	NCLIV_032170	hypothetical protein	1	1	4.662E7	1.157	0.210
730	NCLIV_008680	hypothetical protein, conserved	3	9	2.128E8	1.157	0.210
731	NCLIV_049420	Serine/threonine protein phosphatase (EC 3.1.3.16), related	1	6	3.710E7	1.156	0.209
732	NCLIV_055660	hypothetical protein, conserved	5	16	1.300E8	1.156	0.209
733	NCLIV_043620	hypothetical protein	1	3	6.610E7	1.156	0.209
734	NCLIV_003990	hypothetical protein, conserved	1	8	1.759E8	1.153	0.206
735	NCLIV_037920	Pyrrolidone-carboxylate peptidase (EC 3.4.19.3), related	1	5	2.945E7	1.153	0.206
736	NCLIV_007860	hypothetical protein	10	69	1.944E9	1.153	0.205
737	NCLIV_055330	Fibrillarin superfamily, related	9	59	1.327E9	1.152	0.205
738	NCLIV_040360	hypothetical protein, conserved	2	6	5.850E7	1.152	0.204
739	NCLIV_057550	subtilisin-like protease TgSUB2, putative	6	20	2.197E8	1.152	0.204
740	NCLIV_036510	hypothetical protein	2	14	2.195E7	1.152	0.204
741	NCLIV_003160	Nicotinate phosphoribosyltransferase (EC 2.4.2.11), related	7	26	3.523E8	1.151	0.203
742	NCLIV_055370	hypothetical protein, conserved	5	23	8.304E7	1.151	0.203
743	NCLIV_056720	nucleoporin FG repeat-containing protein, putative	5	18	1.073E8	1.151	0.203
744	NCLIV_033510	hypothetical protein	2	5	6.152E7	1.150	0.201
745	NCLIV_039750	hypothetical protein	2	5	3.694E8	1.148	0.199
746	NCLIV_056770	hypothetical protein, conserved	3	3	3.760E7	1.146	0.196
747	NCLIV_006790	hypothetical protein, conserved	1	3	1.284E7	1.145	0.196
748	NCLIV_006570	serine/threonine protein phosphatase, putative	3	6	7.257E7	1.145	0.196
749	NCLIV_055760	hypothetical protein, conserved	2	13	5.672E7	1.145	0.195

750	NCLIV_057680	hypothetical protein	1	6	4.086E7	1.145	0.195
751	NCLIV_054540	RAB5C, member RAS oncogene family, related	3	44	5.333E8	1.143	0.193
752	NCLIV_046100	glycogen debranching enzyme, putative	4	26	2.033E8	1.143	0.192
753	NCLIV_031340	Splicing factor 3B subunit 3, putative	10	43	4.993E8	1.142	0.192
754	NCLIV_057650	T2D23.8 protein, related	3	17	3.526E8	1.141	0.191
755	NCLIV_045810	Niemann-Pick type C1 disease protein/ patched like cholesterol transporter of the SecD family, 12 transmembrane domain, rel	1	3	2.033E7	1.141	0.190
756	NCLIV_027420	hypothetical protein	1	1	4.869E6	1.140	0.190
757	NCLIV_028230	Proteasome subunit alpha type-7, related	1	13	6.740E8	1.140	0.189
758	NCLIV_016850	AT3G15980 protein, related	9	31	7.612E8	1.140	0.189
759	NCLIV_059980	hypothetical protein, conserved	2	14	1.727E8	1.139	0.187
760	NCLIV_015160	hypothetical protein	4	34	4.503E9	1.138	0.187
761	NCLIV_032240	Catalase (EC 1.11.1.6), related	4	24	3.827E8	1.137	0.186
762	NCLIV_046410	hypothetical protein, conserved	1	3	5.106E7	1.137	0.185
763	NCLIV_053020	hypothetical protein	2	7	1.851E7	1.137	0.185
764	NCLIV_066840	hypothetical protein	19	156	3.052E9	1.137	0.185
765	NCLIV_037180	hypothetical protein	3	16	1.415E8	1.137	0.185
766	NCLIV_032490	brix domain-containing protein, putative	2	10	1.241E7	1.136	0.184
767	NCLIV_041100	Novel protein (Zgc:66430), related	3	21	1.603E8	1.136	0.184
768	NCLIV_024870	hypothetical protein	4	32	1.987E8	1.134	0.182
769	NCLIV_052900	hypothetical protein, conserved	3	12	4.948E8	1.134	0.181
770	NCLIV_024250	hypothetical protein	5	54	2.082E9	1.133	0.181
771	NCLIV_061670	Predicted P-loop ATPase fused to an acetyltransferase (ISS), rel	7	32	2.045E8	1.131	0.178
772	NCLIV_041180	60S ribosomal protein L10, related	4	33	1.351E9	1.131	0.178
773	NCLIV_019030	viral A-type inclusion protein repeat family, putative	1	8	6.613E7	1.131	0.178
774	NCLIV_016900	hypothetical protein	1	7	2.249E8	1.130	0.177
775	NCLIV_003560	mitochondrial alternative NADH dehydrogenase 1, putative	3	6	1.064E8	1.130	0.177
776	NCLIV_042270	hypothetical protein	5	16	8.402E7	1.130	0.176
777	NCLIV_036840	acyl-CoA carboxyltransferase beta chain, putative	1	1	7.173E7	1.130	0.176
778	NCLIV_043400	Proteasome subunit p58, related	4	19	1.773E8	1.128	0.173
779	NCLIV_056980	hypothetical protein, conserved	1	2	2.091E7	1.128	0.173
780	NCLIV_011150	hypothetical protein, conserved	1	1	1.385E6	1.127	0.172
781	NCLIV_053660	CorA-like Mg2 transporter domain-containing protein, related	3	16	1.994E8	1.126	0.171
782	NCLIV_067140	Myosin, related	18	63	6.078E8	1.126	0.171
783	NCLIV_003340	hypothetical protein, conserved	1	7	3.566E8	1.126	0.171
784	NCLIV_002130	ubiquinone/menaquinone biosynthesis methyltransferase, putative	1	1	6.023E6	1.126	0.171
785	NCLIV_005280	hypothetical protein, conserved	9	44	4.357E8	1.125	0.170
786	NCLIV_030140	U2 small nuclear ribonucleoprotein, putative	2	6	8.535E6	1.125	0.170
787	NCLIV_019950	Adenosine deaminase (EC 3.5.4.4), related	1	1	7.539E6	1.125	0.169
788	NCLIV_058500	EF hand domain-containing protein, putative	3	7	2.216E7	1.123	0.168
789	NCLIV_028890	hypothetical protein	3	5	9.374E7	1.123	0.167
790	NCLIV_034850	hypothetical protein, conserved	1	1	5.271E7	1.123	0.167
791	NCLIV_041810	nuclear cap-binding protein, putative	2	7	6.583E7	1.123	0.167

792	NCLIV_003430	hypothetical protein, conserved	1	1	1.338E7	1.123	0.167
793	NCLIV_055020	hypothetical protein, conserved	6	17	1.572E8	1.123	0.167
794	NCLIV_046970	hypothetical protein, conserved	7	37	1.089E8	1.121	0.165
795	NCLIV_020180	hypothetical protein	9	47	8.946E8	1.121	0.165
796	NCLIV_018170	Bromodomain containing protein, related	3	5	7.706E7	1.120	0.164
797	NCLIV_053330	hypothetical protein, conserved	8	43	2.363E9	1.120	0.163
798	NCLIV_003950	hypothetical protein, conserved	2	10	8.251E7	1.118	0.161
799	NCLIV_018800	hypothetical protein	14	83	1.161E9	1.118	0.161
800	NCLIV_048640	RNA-binding protein, putative	1	1	5.888E7	1.118	0.161
801	NCLIV_020650	splicing factor 3B subunit 1, putative	6	22	1.855E8	1.118	0.160
802	NCLIV_014040	hypothetical protein, conserved	12	59	9.383E8	1.116	0.159
803	NCLIV_028280	hypothetical protein, conserved	7	26	1.719E8	1.112	0.154
804	NCLIV_000010	heat shock protein 90, putative	6	30	6.618E8	1.112	0.153
805	NCLIV_025160	hypothetical protein	4	41	1.031E9	1.111	0.152
806	NCLIV_056960	hypothetical protein, conserved	1	8	1.702E7	1.109	0.149
807	NCLIV_026040	structure specific recognition protein I, putative	3	9	1.828E7	1.108	0.148
808	NCLIV_018950	coatamer epsilon subunit, putative	3	13	1.329E8	1.108	0.148
809	NCLIV_045990	NAD-dependent deacetylase, putative	1	3	3.738E7	1.107	0.147
810	NCLIV_047070	hypothetical protein	6	16	2.249E7	1.107	0.146
811	NCLIV_003010	hypothetical protein	1	4	4.657E7	1.106	0.146
812	NCLIV_036720	CBR-CSN-5 protein, related	3	13	3.112E8	1.106	0.146
813	NCLIV_018890	L24, related	2	14	1.553E8	1.105	0.144
814	NCLIV_007040	hypothetical protein, conserved	1	2	7.331E8	1.104	0.143
815	NCLIV_054440	hypothetical protein	2	21	1.222E9	1.104	0.142
816	NCLIV_045670	hypothetical protein	2	15	1.075E9	1.102	0.141
817	NCLIV_003600	ABC transporter, putative	4	9	8.743E7	1.101	0.139
818	NCLIV_042810	hypothetical protein	2	5	4.031E8	1.100	0.138
819	NCLIV_025450	elongation factor Tu, putative	8	40	9.481E8	1.099	0.136
820	NCLIV_010610	hypothetical protein	1	5	9.733E7	1.099	0.136
821	NCLIV_012830	MORN repeat-containing protein, putative	5	21	2.797E8	1.099	0.136
822	NCLIV_049710	hypothetical protein	11	46	7.222E8	1.099	0.136
823	NCLIV_057540	hypothetical protein, conserved	1	7	1.238E8	1.098	0.135
824	NCLIV_020430	hypothetical protein	5	14	1.231E8	1.096	0.132
825	NCLIV_050910	hypothetical protein, conserved	8	30	3.534E8	1.096	0.132
826	NCLIV_021550	hypothetical protein	4	11	1.865E8	1.096	0.132
827	NCLIV_063890	hypothetical protein, conserved	2	6	7.038E7	1.095	0.131
828	NCLIV_018510	Hydrolase Cof, related	4	14	3.916E7	1.094	0.129
829	NCLIV_038870	hypothetical protein, conserved	3	5	5.065E7	1.093	0.129
830	NCLIV_017370	CAMP-dependent protein kinase regulatory subunit, putative	14	95	2.588E9	1.093	0.128
831	NCLIV_066580	ATP-dependent RNA helicase, putative	2	4	3.966E7	1.092	0.126
832	NCLIV_041770	hypothetical protein, conserved	2	4	3.597E7	1.091	0.126
833	NCLIV_019740	3',5'-cyclic phosphodiesterase, putative	1	2	1.803E8	1.091	0.125
834	NCLIV_008170	hypothetical protein	3	14	1.585E8	1.091	0.125

835	NCLIV_015570	alpha/beta hydrolase fold domain containing protein, putative	5	17	2.947E8	1.090	0.125
836	NCLIV_026850	hypothetical protein, conserved	1	1	4.039E6	1.090	0.124
837	NCLIV_050680	hypothetical protein	13	63	9.426E8	1.088	0.121
838	NCLIV_056910	hypothetical protein, conserved	2	11	4.813E8	1.088	0.121
839	NCLIV_016330	phosphomannomutase 2, putative	1	6	2.730E7	1.087	0.121
840	NCLIV_012510	hypothetical protein	4	27	1.599E9	1.087	0.120
841	NCLIV_050500	hypothetical protein, conserved	1	1	1.012E7	1.087	0.120
842	NCLIV_028180	DNA-directed RNA polymerase subunit beta	6	13	3.454E8	1.086	0.119
843	NCLIV_030650	26S protease regulatory subunit 6b, putative	1	15	8.330E8	1.086	0.119
844	NCLIV_025380	hypothetical protein, conserved	1	5	4.461E7	1.086	0.119
845	NCLIV_069330	hypothetical protein, conserved	2	4	8.950E7	1.086	0.119
846	NCLIV_009810	hypothetical protein	4	20	1.073E8	1.085	0.118
847	NCLIV_019000	adenosine transporter, putative	4	19	2.622E8	1.085	0.117
848	NCLIV_046450	sec23/Sec24 helical domain-containing protein, putative	5	15	1.330E8	1.084	0.117
849	NCLIV_026540	hypothetical protein, conserved	3	53	1.152E9	1.084	0.117
850	NCLIV_037570	hypothetical protein, conserved	4	12	1.152E8	1.084	0.116
851	NCLIV_049100	Ribosomal protein S19e, related	2	20	1.067E9	1.084	0.116
852	NCLIV_001360	hypothetical protein, conserved	3	17	6.078E7	1.083	0.116
853	NCLIV_041040	hypothetical protein, conserved	2	4	1.872E7	1.083	0.115
854	NCLIV_003220	vacuolar ATP synthase subunit h, putative	5	20	1.961E8	1.083	0.115
855	NCLIV_070050	hypothetical protein	2	3	7.960E6	1.083	0.115
856	NCLIV_036280	30S ribosomal protein S15P/S13e, related	2	10	2.882E8	1.082	0.114
857	NCLIV_041730	hypothetical protein	1	3	8.375E7	1.082	0.114
858	NCLIV_066280	DEHA2F02992p, related	1	4	8.759E7	1.082	0.114
859	NCLIV_050130	ULK kinase, putative	1	2	3.743E7	1.082	0.114
860	NCLIV_012460	Lysyl-tRNA synthetase, related	1	2	1.865E7	1.081	0.112
861	NCLIV_057030	hypothetical protein, conserved	2	5	2.601E8	1.080	0.111
862	NCLIV_021080	hypothetical protein	5	56	1.629E9	1.080	0.110
863	NCLIV_049500	hypothetical protein	7	15	2.051E8	1.079	0.109
864	NCLIV_022270	hypothetical protein, conserved	10	33	5.122E8	1.079	0.109
865	NCLIV_015060	hypothetical protein	4	16	1.005E8	1.078	0.109
866	NCLIV_003580	hypothetical protein, conserved	2	10	3.259E8	1.077	0.108
867	NCLIV_043790	hypothetical protein	5	23	6.032E7	1.077	0.107
868	NCLIV_033340	sulfate transporter, putative	3	8	8.784E7	1.076	0.106
869	NCLIV_050320	hypothetical protein, conserved	1	3	1.977E7	1.076	0.106
870	NCLIV_058990	myosin heavy chain, putative	1	3	8.597E7	1.076	0.105
871	NCLIV_011980	Calmodulin-like domain protein kinase isoenzyme gamma, rel	11	58	2.076E9	1.075	0.105
872	NCLIV_063220	hypothetical protein, conserved	3	11	9.928E7	1.074	0.103
873	NCLIV_026250	NADPH:adrenodoxin oxidoreductase FprB, related	2	3	4.919E7	1.073	0.102
874	NCLIV_061030	hypothetical protein	2	5	2.515E8	1.072	0.100
875	NCLIV_064360	50S ribosomal protein L24P, related	3	17	5.379E8	1.072	0.100
876	NCLIV_024090	glutamyl-tRNA synthetase, putative	7	28	2.110E8	1.071	0.099
877	NCLIV_068850	hypothetical protein	11	133	1.352E9	1.071	0.099

878	NCLIV_068872	SRS domain-containing protein	3	10	1.890E8	1.070	0.098
879	NCLIV_068190	hypothetical protein	2	15	2.704E7	1.070	0.098
880	NCLIV_010730	SRS domain-containing protein	8	195	1.966E10	1.070	0.098
881	NCLIV_026390	centrin, putative	4	11	2.365E8	1.069	0.097
882	NCLIV_024880	30S ribosomal protein S9P, related	4	35	6.861E8	1.069	0.096
883	NCLIV_045740	lytB domain-containing protein, putative	1	3	4.640E7	1.068	0.095
884	NCLIV_069320	hypothetical protein, conserved	1	1	3.206E7	1.067	0.093
885	NCLIV_042050	hypothetical protein	2	21	7.041E8	1.067	0.093
886	NCLIV_026210	hypothetical protein	1	3	1.392E8	1.066	0.093
887	NCLIV_045710	hypothetical protein	1	5	2.570E7	1.066	0.092
888	NCLIV_066140	RNA polymerase, putative	1	2	4.444E6	1.066	0.092
889	NCLIV_040440	hypothetical protein, conserved	8	32	4.261E8	1.065	0.091
890	NCLIV_016310	hypothetical protein, conserved	1	6	2.436E8	1.065	0.091
891	NCLIV_047870	cAMP phosphodiesterase, putative	6	34	4.062E8	1.065	0.091
892	NCLIV_067280	hypothetical protein, conserved	1	4	2.084E7	1.065	0.091
893	NCLIV_047500	Trichohyalin, putative	9	26	1.193E8	1.064	0.089
894	NCLIV_034460	hypothetical protein	13	68	5.093E8	1.064	0.089
895	NCLIV_062910	hypothetical protein, conserved	4	8	2.725E8	1.064	0.089
896	NCLIV_001370	DEAD/DEAH box helicase, putative	8	35	2.223E8	1.063	0.088
897	NCLIV_015460	hypothetical protein	1	9	1.317E8	1.063	0.088
898	NCLIV_014360	hypothetical protein	6	25	1.731E8	1.063	0.088
899	NCLIV_031030	hypothetical protein, conserved	2	5	1.083E8	1.063	0.088
900	NCLIV_065390	Bifunctional dihydrofolate reductase-thymidylate synthase, rel	5	10	1.914E8	1.062	0.087
901	NCLIV_010520	hypothetical protein, conserved	5	18	2.228E8	1.062	0.087
902	NCLIV_063650	hypothetical protein, conserved	2	6	2.508E7	1.062	0.087
903	NCLIV_026270	26S proteasome regulatory subunit S4 like AAA ATPase, related	5	16	4.083E8	1.062	0.086
904	NCLIV_007450	hypothetical protein	1	40	2.194E9	1.062	0.086
905	NCLIV_021150	ABC1 domain-containing protein, putative	1	2	2.349E7	1.061	0.086
906	NCLIV_044200	hypothetical protein	9	80	2.065E9	1.061	0.085
907	NCLIV_047040	hypothetical protein, conserved	3	20	4.825E8	1.061	0.085
908	NCLIV_024780	citrate synthase, putative	2	8	3.432E8	1.061	0.085
909	NCLIV_002380	hypothetical protein, conserved	5	34	7.427E8	1.060	0.084
910	NCLIV_049830	hypothetical protein, conserved	8	18	4.675E8	1.060	0.084
911	NCLIV_010230	hypothetical protein	4	15	5.485E8	1.060	0.084
912	NCLIV_054800	hypothetical protein, conserved	7	74	5.521E8	1.059	0.083
913	NCLIV_059140	C2 domain-containing protein, putative	1	4	2.221E7	1.059	0.083
914	NCLIV_051450	centromere/microtubule binding protein, putative	7	31	6.622E8	1.058	0.082
915	NCLIV_034090	kinesin heavy chain, putative	13	71	4.736E8	1.058	0.081
916	NCLIV_024210	ATP-binding cassette, putative	2	4	5.305E7	1.057	0.081
917	NCLIV_044840	hypothetical protein, conserved	2	6	2.496E8	1.057	0.080
918	NCLIV_045050	tubulin-tyrosine ligase family protein, putative	1	15	5.217E8	1.057	0.080
919	NCLIV_013860	F2J6.5 protein, related	1	1	1.500E7	1.057	0.080
920	NCLIV_067180	Glucose-6-phosphate 1-dehydrogenase (EC 1.1.1.49), related	8	26	4.194E8	1.056	0.079

921	NCLIV_006260	hypothetical protein, conserved	1	7	9.961E6	1.056	0.078
922	NCLIV_060050	hypothetical protein, conserved	1	3	1.148E7	1.055	0.078
923	NCLIV_000610	profilin family protein, putative	4	33	3.661E8	1.055	0.077
924	NCLIV_068510	hypothetical protein	1	2	1.803E7	1.055	0.077
925	NCLIV_051510	GM14051, related	3	13	3.972E8	1.054	0.075
926	NCLIV_009300	GK17769, related	4	17	1.405E8	1.053	0.075
927	NCLIV_069130	hypothetical protein, conserved	7	33	8.400E8	1.053	0.074
928	NCLIV_068300	GD23551, related	2	5	5.624E7	1.053	0.074
929	NCLIV_029830	Integral membrane protein, related	2	7	9.420E7	1.052	0.074
930	NCLIV_045590	hypothetical protein, conserved	1	5	2.118E7	1.052	0.073
931	NCLIV_027100	orotidine-monophosphate-decarboxylase, putative	1	1	3.552E6	1.052	0.073
932	NCLIV_030950	hypothetical protein	3	8	4.757E8	1.052	0.073
933	NCLIV_040120	hypothetical protein	1	4	3.745E7	1.052	0.073
934	NCLIV_065790	glutaredoxin domain containing protein, putative	2	5	2.514E8	1.052	0.073
935	NCLIV_038780	60S ribosomal protein L32, related	4	31	1.026E9	1.051	0.072
936	NCLIV_037720	Os07g0694800 protein, related	2	11	5.248E7	1.050	0.071
937	NCLIV_061990	hypothetical protein, conserved	7	31	3.348E8	1.049	0.069
938	NCLIV_022690	hypothetical protein, conserved	3	11	4.967E8	1.049	0.069
939	NCLIV_016750	Cytochrome c oxidase copper chaperone, related	1	1	1.582E8	1.049	0.069
940	NCLIV_025580	eukaryotic translation initiation factor 3 subunit 11, putative	1	4	2.885E7	1.048	0.068
941	NCLIV_048090	sterol-regulatory element binding protein site 2 protease, putative	1	2	4.749E7	1.048	0.067
942	NCLIV_038320	hypothetical protein	4	26	1.741E8	1.047	0.066
943	NCLIV_064600	cysteine desulfurase, putative	16	129	2.110E9	1.046	0.066
944	NCLIV_042500	Ubiquitin-activating enzyme E1, related	6	20	3.081E8	1.046	0.065
945	NCLIV_045010	hypothetical protein	2	27	2.486E9	1.046	0.065
946	NCLIV_047630	40S ribosomal protein S18, putative	4	29	1.084E9	1.045	0.063
947	NCLIV_063480	hypothetical protein, conserved	1	4	5.296E7	1.044	0.062
948	NCLIV_010010	HET-R, related	9	36	1.160E9	1.044	0.062
949	NCLIV_062250	hypothetical protein, conserved	2	6	5.398E7	1.042	0.060
950	NCLIV_026180	hypothetical protein	2	13	3.379E8	1.042	0.060
951	NCLIV_010200	hypothetical protein	4	33	7.603E8	1.042	0.059
952	NCLIV_013120	activator 1 36 kDa, putative	3	10	1.333E8	1.042	0.059
953	NCLIV_027150	hypothetical protein	5	15	2.975E8	1.042	0.059
954	NCLIV_060820	V-type ATP synthase beta chain, related	11	62	5.145E9	1.041	0.059
955	NCLIV_054990	hypothetical protein, conserved	1	5	1.906E7	1.040	0.057
956	NCLIV_053950	hypothetical protein	2	38	2.240E9	1.040	0.057
957	NCLIV_057700	hypothetical protein	4	37	2.111E9	1.038	0.054
958	NCLIV_028240	Ras family domain-containing protein, putative	1	39	4.879E8	1.038	0.053
959	NCLIV_023620	SRS domain-containing protein	3	12	2.917E8	1.038	0.053
960	NCLIV_031440	60S ribosomal protein L38, related	2	15	2.481E8	1.037	0.052
961	NCLIV_019490	lysine decarboxylase, putative	2	3	7.606E6	1.037	0.052
962	NCLIV_019890	hypothetical protein	1	7	1.069E8	1.036	0.052
963	NCLIV_037910	hypothetical protein	1	2	1.255E8	1.036	0.051

964	NCLIV_009150	alpha-galactosidase A, putative	3	15	1.999E8	1.036	0.050
965	NCLIV_014240	hypothetical protein, conserved	4	6	1.027E8	1.035	0.050
966	NCLIV_064440	hypothetical protein	5	21	6.311E8	1.035	0.049
967	NCLIV_065590	NAD-specific glutamate dehydrogenase, putative	21	109	1.382E9	1.033	0.047
968	NCLIV_046830	ATP synthase, putative	3	20	1.640E8	1.033	0.047
969	NCLIV_015870	SRS domain-containing protein,SRS23, putative	4	17	7.517E8	1.033	0.047
970	NCLIV_025570	protein kinase, putative	2	3	4.457E7	1.032	0.046
971	NCLIV_038550	Strain CBS138 chromosome D complete sequence, related	3	8	2.208E8	1.032	0.045
972	NCLIV_060610	hypothetical protein	1	4	9.148E6	1.032	0.045
973	NCLIV_016540	hypothetical protein, conserved	9	41	9.843E8	1.031	0.044
974	NCLIV_006070	30S ribosomal protein S10P, related	4	28	1.305E9	1.030	0.043
975	NCLIV_003750	TLD domain-containing protein, putative	2	6	1.778E8	1.030	0.042
976	NCLIV_035700	endonuclease/exonuclease/phosphatase containing protein, putative	2	9	9.115E7	1.030	0.042
977	NCLIV_008650	hypothetical protein	1	8	1.402E8	1.029	0.042
978	NCLIV_058890	tubulin alpha chain	15	235	2.139E10	1.029	0.041
979	NCLIV_046800	AGC kinase, putative	6	34	2.848E9	1.029	0.041
980	NCLIV_057460	Transketolase central region, related	4	27	2.430E9	1.028	0.040
981	NCLIV_070080	hypothetical protein, conserved	3	7	3.254E8	1.028	0.040
982	NCLIV_024830	hypothetical protein, conserved	4	16	8.513E8	1.028	0.039
983	NCLIV_051560	Glucose transporter 1A, related	3	26	3.493E8	1.027	0.039
984	NCLIV_029040	nucleosome assembly protein, putative	2	9	7.688E7	1.027	0.038
985	NCLIV_032370	rna 3' terminal phosphate cyclase, putative	1	3	1.400E8	1.026	0.038
986	NCLIV_023390	hypothetical protein, conserved	7	19	8.714E7	1.026	0.037
987	NCLIV_019810	3-ketodihydrosphingosine reductase TSC10, related	1	6	1.537E8	1.026	0.037
988	NCLIV_028500	GJ17676, related	5	21	6.926E7	1.026	0.036
989	NCLIV_002770	MORN repeat-containing protein, putative	7	46	4.667E8	1.025	0.035
990	NCLIV_069590	hypothetical protein	9	107	2.581E9	1.024	0.034
991	NCLIV_019340	DNA-directed rna polymerase I largest subunit, putative	11	34	1.024E8	1.024	0.034
992	NCLIV_062950	50S ribosomal protein L21e, related	5	46	1.598E9	1.023	0.033
993	NCLIV_046400	hypothetical protein, conserved	2	11	9.759E7	1.023	0.032
994	NCLIV_033810	hypothetical protein	1	5	2.340E8	1.021	0.031
995	NCLIV_056640	WD-40 repeat-containing protein, putative	1	3	3.803E6	1.021	0.030
996	NCLIV_012120	hypothetical protein	4	35	8.615E8	1.021	0.029
997	NCLIV_046390	hypothetical protein	2	18	1.359E9	1.020	0.029
998	NCLIV_056130	hypothetical protein, conserved	5	23	4.274E8	1.019	0.028
999	NCLIV_045970	pyruvate dehydrogenase kinase, putative	1	5	2.774E8	1.019	0.027
1000	NCLIV_001550	ribonucleoside-diphosphate reductase, large subunit, putative	1	2	3.322E7	1.019	0.027
1001	NCLIV_051840	hypothetical protein	3	30	2.479E9	1.018	0.026
1002	NCLIV_063190	hypothetical protein, conserved	2	12	5.431E8	1.018	0.026
1003	NCLIV_061830	60S acidic ribosomal protein P0	10	86	3.257E9	1.018	0.025
1004	NCLIV_065690	hypothetical protein	4	18	2.773E8	1.017	0.025
1005	NCLIV_008320	Protein F23B2.12, partially confirmed by transcript evidence, related	1	3	3.456E8	1.017	0.024
1006	NCLIV_025190	LOC549444 protein, related	5	29	1.086E9	1.016	0.024

1007	NCLIV_070280	hypothetical protein	6	50	1.511E9	1.016	0.023
1008	NCLIV_006540	DNA-directed RNA polymerase (EC 2.7.7.6), related	4	8	8.834E7	1.016	0.023
1009	NCLIV_012400	Articulon family protein, related	7	55	2.634E9	1.016	0.022
1010	NCLIV_054500	hypothetical protein	1	6	1.441E8	1.015	0.022
1011	NCLIV_004760	transmembrane domain-containing protein, putative	1	8	9.210E7	1.015	0.022
1012	NCLIV_045960	hypothetical protein, conserved	3	5	7.878E7	1.015	0.021
1013	NCLIV_038410	Predicted hydrolases or acyltransferases, related	1	3	1.402E8	1.015	0.021
1014	NCLIV_043880	hypothetical protein	1	8	7.338E7	1.015	0.021
1015	NCLIV_016450	hypothetical protein	4	42	3.456E9	1.013	0.019
1016	NCLIV_022950	RNA-binding protein, putative	6	35	3.533E8	1.013	0.018
1017	NCLIV_034010	Wee kinase, putative	1	2	1.614E7	1.013	0.018
1018	NCLIV_004920	hypothetical protein	4	48	1.047E9	1.012	0.017
1019	NCLIV_045230	vacuolar (H ⁺)-ATPase G subunit domain-containing protein, putative	2	12	4.606E7	1.011	0.016
1020	NCLIV_052500	hypothetical protein	6	55	1.566E9	1.011	0.016
1021	NCLIV_033910	hypothetical protein	1	5	9.349E7	1.011	0.015
1022	NCLIV_046160	hypothetical protein, conserved	2	6	6.514E7	1.010	0.014
1023	NCLIV_011400	ATP-dependent helicase, putative	6	28	2.370E8	1.010	0.014
1024	NCLIV_032770	hypothetical protein	7	48	1.571E8	1.009	0.014
1025	NCLIV_058830	Superoxide dismutase (EC 1.15.1.1), related	4	30	1.370E9	1.009	0.013
1026	NCLIV_033430	hypothetical protein, conserved	1	2	7.134E7	1.009	0.013
1027	NCLIV_062790	P-type ATPase, putative	7	23	2.086E8	1.008	0.012
1028	NCLIV_044810	tRNA (adenine-N(1)-)-methyltransferase catalytic subunit, putative	2	5	3.947E7	1.008	0.011
1029	NCLIV_051360	Similarity to HYPOTHETICAL INTEGRAL MEMBRANE PROTEIN YIPA_yeast, related	1	1	1.086E7	1.008	0.011
1030	NCLIV_031190	hypothetical protein	1	4	9.119E7	1.006	0.009
1031	NCLIV_032940	adenosine transporter, putative	1	6	2.965E7	1.006	0.009
1032	NCLIV_000960	hypothetical protein, conserved	2	6	1.665E8	1.006	0.008
1033	NCLIV_041320	CCAAT-box DNA binding protein subunit B, related	1	2	3.855E7	1.005	0.007
1034	NCLIV_037460	hypothetical protein	6	22	5.597E8	1.004	0.006
1035	NCLIV_052410	transporter, major facilitator family domain containing protein, putative	3	36	2.401E8	1.004	0.006
1036	NCLIV_051040	hypothetical protein, conserved	3	18	5.986E8	1.004	0.006
1037	NCLIV_001540	PH domain-containing protein, putative	3	18	2.297E8	1.004	0.006
1038	NCLIV_005540	hypothetical protein	1	3	7.796E7	1.004	0.005
1039	NCLIV_054240	Probable lysosomal cobalamin transporter, related	1	3	3.399E7	1.003	0.005
1040	NCLIV_038560	hypothetical protein, conserved	1	5	9.852E7	1.003	0.004
1041	NCLIV_010220	hypothetical protein, conserved	1	4	5.530E7	1.003	0.004
1042	NCLIV_064780	actin, putative	3	13	2.871E8	1.003	0.004
1043	NCLIV_056950	hypothetical protein	2	10	1.225E8	1.001	0.002
1044	NCLIV_007900	PV1H14125_P, related	5	14	4.872E8	1.001	0.001
1045	NCLIV_035360	thioredoxin domain-containing protein, putative	1	1	1.115E7	1.001	0.001
1046	NCLIV_042070	hypothetical protein	2	14	1.115E9	1.000	0.000
1047	NCLIV_025670	ATP synthase subunit beta (EC 3.6.3.14), related	17	216	6.833E9	0.999	-0.001
1048	NCLIV_006480	hypothetical protein	9	30	5.637E8	0.999	-0.001

1049	NCLIV_065140	endonuclease/exonuclease/phosphatase containing protein, putative	6	26	4.625E8	0.999	-0.001
1050	NCLIV_065210	KLLA0F09449p, related	14	118	3.085E9	0.999	-0.002
1051	NCLIV_034770	hypothetical protein	1	1	2.456E7	0.997	-0.005
1052	NCLIV_004060	hypothetical protein, conserved	3	16	5.259E8	0.996	-0.005
1053	NCLIV_045720	mRNA decapping protein, putative	1	4	5.791E6	0.996	-0.005
1054	NCLIV_020140	hypothetical protein	1	8	3.811E8	0.996	-0.006
1055	NCLIV_022400	hypothetical protein	9	39	3.325E8	0.996	-0.006
1056	NCLIV_031670	hypothetical protein, conserved	9	46	1.078E9	0.994	-0.009
1057	NCLIV_066830	hypothetical protein, conserved	1	1	3.092E7	0.994	-0.009
1058	NCLIV_033480	hypothetical protein, conserved	2	12	1.856E8	0.994	-0.009
1059	NCLIV_024630	porin, putative	12	98	2.966E9	0.994	-0.009
1060	NCLIV_029640	SCO1/SenC domain-containing protein, putative	1	1	6.282E7	0.994	-0.009
1061	NCLIV_025720	ethanolaminephosphotransferase, putative	1	4	1.408E8	0.994	-0.009
1062	NCLIV_006640	hypothetical protein	23	89	4.372E8	0.993	-0.010
1063	NCLIV_064700	Ribosomal protein L18, related	4	36	5.880E8	0.993	-0.010
1064	NCLIV_015430	hypothetical protein	8	64	4.713E8	0.993	-0.010
1065	NCLIV_015550	26S proteasome non-ATPase regulatory subunit, putative	3	15	1.603E8	0.993	-0.010
1066	NCLIV_013670	hypothetical protein, conserved	1	1	7.715E7	0.993	-0.010
1067	NCLIV_057970	hypothetical protein, conserved	2	4	7.081E6	0.993	-0.010
1068	NCLIV_026570	hypothetical protein	1	9	7.604E7	0.992	-0.011
1069	NCLIV_016700	hypothetical protein, conserved	1	1	3.824E7	0.992	-0.011
1070	NCLIV_060750	DNA replication licensing factor, putative	2	6	3.190E7	0.992	-0.012
1071	NCLIV_049160	hypothetical protein, conserved	2	6	1.014E8	0.991	-0.012
1072	NCLIV_031770	membrane skeletal protein IMC1, putative	8	124	9.196E9	0.991	-0.012
1073	NCLIV_044060	hypothetical protein	5	16	7.881E8	0.991	-0.013
1074	NCLIV_005940	phosphoglycerate mutase family protein, putative	3	14	2.411E8	0.991	-0.013
1075	NCLIV_036750	ADP-ribosylation factor, putative	1	3	6.961E7	0.990	-0.015
1076	NCLIV_053870	hypothetical protein	1	12	3.286E8	0.989	-0.016
1077	NCLIV_040600	hypothetical protein	4	26	4.371E8	0.989	-0.016
1078	NCLIV_031500	hypothetical protein	2	16	2.386E8	0.989	-0.016
1079	NCLIV_017340	hypothetical protein	2	11	1.199E8	0.989	-0.017
1080	NCLIV_000240	WW domain containing protein, putative	3	12	3.425E7	0.988	-0.017
1081	NCLIV_022560	Thermosome subunit, related	1	19	3.948E8	0.988	-0.017
1082	NCLIV_049900	hypothetical protein	33	477	1.071E10	0.988	-0.018
1083	NCLIV_065450	hypothetical protein	2	25	1.629E9	0.987	-0.018
1084	NCLIV_010390	Novel protein (Zgc:77804), related	2	9	2.580E7	0.987	-0.019
1085	NCLIV_006900	hypothetical protein, conserved	1	4	2.756E6	0.987	-0.019
1086	NCLIV_002520	hypothetical protein	8	74	3.462E9	0.986	-0.020
1087	NCLIV_020390	hypothetical protein	2	5	1.088E8	0.986	-0.021
1088	NCLIV_049050	RNA helicase-related protein required for pre- mRNA splicing, related	12	49	4.141E8	0.985	-0.022
1089	NCLIV_004750	peptidase family M48 domain-containing protein, putative	2	32	8.069E8	0.985	-0.022
1090	NCLIV_056700	26S proteasome regulatory subunit rpn1, related	10	32	5.126E8	0.985	-0.022
1091	NCLIV_011270	hypothetical protein	9	51	2.123E9	0.984	-0.023

1092	NCLIV_026350	hypothetical protein, conserved	2	4	3.617E7	0.984	-0.024
1093	NCLIV_005830	hypothetical protein	3	12	5.865E7	0.984	-0.024
1094	NCLIV_025910	Histone H2A, related	1	20	2.585E8	0.982	-0.026
1095	NCLIV_014780	hypothetical protein, conserved	2	5	2.162E7	0.982	-0.026
1096	NCLIV_060470	hypothetical protein	2	3	2.868E7	0.982	-0.026
1097	NCLIV_031970	hypothetical protein	19	85	5.690E8	0.981	-0.027
1098	NCLIV_010020	Ubiquinol-cytochrome c reductase iron-sulfur subunit, related	8	74	4.655E9	0.980	-0.029
1099	NCLIV_067300	hypothetical protein, conserved	3	8	8.626E7	0.980	-0.029
1100	NCLIV_058440	Os02g0824100 protein, related	10	45	2.164E9	0.980	-0.029
1101	NCLIV_039410	hypothetical protein	1	2	2.560E7	0.980	-0.029
1102	NCLIV_042390	hypothetical protein, conserved	4	40	6.123E8	0.980	-0.030
1103	NCLIV_026010	Phosphatidylserine synthase 2 (EC 2.7.8.-), related	1	4	7.262E7	0.979	-0.031
1104	NCLIV_022430	hypothetical protein	3	14	2.293E8	0.979	-0.031
1105	NCLIV_001450	hypothetical protein	20	71	2.721E8	0.978	-0.032
1106	NCLIV_018220	Mitochondrial protein-like, related	2	9	2.340E7	0.978	-0.032
1107	NCLIV_055850	hypothetical protein, conserved	8	49	1.092E9	0.978	-0.032
1108	NCLIV_042670	hypothetical protein, conserved	5	8	3.430E7	0.977	-0.033
1109	NCLIV_019590	Selenophosphate synthase (EC 2.7.9.3), related	11	36	2.871E8	0.977	-0.034
1110	NCLIV_064620	rhoptry neck protein 2, putative	33	229	5.804E9	0.976	-0.035
1111	NCLIV_063370	hypothetical protein, conserved	3	14	3.789E8	0.976	-0.035
1112	NCLIV_001180	CBR-GPDH-1 protein, related	3	13	3.719E8	0.976	-0.035
1113	NCLIV_042710	Peptidase S1, chymotrypsin:PDZ/DHR/GLGF domain, related	2	6	5.434E7	0.975	-0.037
1114	NCLIV_044100	hypothetical protein, conserved	1	3	3.320E7	0.975	-0.037
1115	NCLIV_056680	hypothetical protein	2	27	1.914E9	0.974	-0.037
1116	NCLIV_047080	hypothetical protein, conserved	3	26	1.111E9	0.974	-0.037
1117	NCLIV_016430	hypothetical protein, conserved	1	5	3.871E7	0.974	-0.037
1118	NCLIV_015480	emp24/gp25L/p24 family domain, transmembrane protein, putative	3	20	6.404E8	0.974	-0.038
1119	NCLIV_048050	hypothetical protein, conserved	2	5	4.315E8	0.974	-0.038
1120	NCLIV_066920	hypothetical protein	3	18	4.148E8	0.974	-0.038
1121	NCLIV_031920	ATP synthase gamma chain, putative	5	40	9.638E8	0.973	-0.039
1122	NCLIV_022310	Cell division protease FtsH homolog (EC 3.4.24.-), related	2	12	2.407E7	0.973	-0.040
1123	NCLIV_055720	hypothetical protein	10	32	4.070E8	0.972	-0.041
1124	NCLIV_037950	hypothetical protein	1	6	5.641E7	0.971	-0.042
1125	NCLIV_056120	hypothetical protein, conserved	1	2	7.854E7	0.971	-0.042
1126	NCLIV_002020	hypothetical protein	4	18	2.020E8	0.971	-0.042
1127	NCLIV_032270	hypothetical protein, conserved	9	73	2.892E9	0.971	-0.043
1128	NCLIV_006490	myosin light chain 2, putative	4	20	4.057E8	0.970	-0.044
1129	NCLIV_035590	hypothetical protein, conserved	2	18	1.507E8	0.970	-0.044
1130	NCLIV_019940	hypothetical protein, conserved	1	5	1.898E7	0.970	-0.044
1131	NCLIV_044950	hypothetical protein, conserved	1	3	3.638E8	0.970	-0.045
1132	NCLIV_060190	acetyl-CoA transporter, putative	1	6	1.229E8	0.969	-0.045
1133	NCLIV_019450	hypothetical protein	4	22	3.230E8	0.968	-0.047
1134	NCLIV_051820	hypothetical protein	3	81	1.989E9	0.967	-0.048

1135	NCLIV_037350	Abc transporter, related	3	11	1.656E8	0.966	-0.049
1136	NCLIV_008860	hypothetical protein	3	10	8.843E7	0.966	-0.050
1137	NCLIV_024600	hypothetical protein	2	7	2.645E8	0.966	-0.050
1138	NCLIV_066080	hypothetical protein, conserved	3	12	1.988E8	0.966	-0.050
1139	NCLIV_025480	kelch motif domain-containing protein, putative	4	23	1.604E8	0.966	-0.051
1140	NCLIV_063980	hypothetical protein, conserved	2	20	1.074E9	0.965	-0.051
1141	NCLIV_034950	poly(A) RNA export protein, putative	2	4	4.027E7	0.965	-0.051
1142	NCLIV_013320	hypothetical protein	1	21	1.369E8	0.964	-0.052
1143	NCLIV_013130	hypothetical protein, conserved	2	4	1.922E7	0.964	-0.053
1144	NCLIV_041650	Methionine--tRNA ligase (EC 6.1.1.10), related	8	23	3.927E8	0.964	-0.053
1145	NCLIV_058200	hypothetical protein, conserved	1	8	5.629E7	0.964	-0.053
1146	NCLIV_010770	glycosylphosphatidylinositol anchor attachment 1 protein, put	2	8	2.915E7	0.964	-0.054
1147	NCLIV_025220	hypothetical protein, conserved	3	24	1.402E9	0.964	-0.054
1148	NCLIV_030340	cytochrome c, putative	2	7	1.493E8	0.963	-0.055
1149	NCLIV_016110	hypothetical protein	5	41	7.953E8	0.962	-0.055
1150	NCLIV_021640	dense granule protein 7, putative	2	21	6.615E8	0.962	-0.056
1151	NCLIV_053860	MGC84926 protein, related	4	19	1.918E9	0.962	-0.056
1152	NCLIV_068920	SRS domain-containing protein	7	298	1.757E10	0.962	-0.056
1153	NCLIV_038810	kinesin motor domain-containing protein, putative	3	5	2.113E7	0.961	-0.057
1154	NCLIV_055170	DnaJ domain containing protein, related	7	27	1.507E9	0.960	-0.058
1155	NCLIV_020840	hypothetical protein	20	280	9.557E9	0.960	-0.058
1156	NCLIV_053370	U2 small nuclear ribonucleoprotein auxiliary factor U2AF, put	1	2	1.359E8	0.959	-0.061
1157	NCLIV_062720	hypothetical protein	7	75	2.857E9	0.958	-0.061
1158	NCLIV_052230	Succinate dehydrogenase iron-sulfur subunit, related	6	49	1.740E9	0.958	-0.062
1159	NCLIV_009390	Cleft lip and palate transmembrane protein 1, putative	6	26	4.765E8	0.958	-0.062
1160	NCLIV_003230	DNA/RNA non-specific endonuclease containing protein, putative	1	2	7.296E6	0.958	-0.062
1161	NCLIV_060990	hypothetical protein	1	2	7.615E5	0.958	-0.062
1162	NCLIV_070010	hypothetical protein, conserved	52	312	2.887E9	0.958	-0.062
1163	NCLIV_025920	hypothetical protein	13	77	2.341E9	0.958	-0.063
1164	NCLIV_047560	hypothetical protein	7	19	1.680E8	0.957	-0.063
1165	NCLIV_006800	hypothetical protein, conserved	5	10	1.424E8	0.957	-0.064
1166	NCLIV_010240	hypothetical protein, conserved	3	7	1.034E8	0.956	-0.065
1167	NCLIV_007990	hypothetical protein, conserved	1	4	6.366E7	0.956	-0.065
1168	NCLIV_067010	Mitochondrial phosphate carrier protein, related	10	57	1.643E9	0.956	-0.065
1169	NCLIV_008730	hypothetical protein	2	11	1.213E8	0.955	-0.066
1170	NCLIV_018290	Ribosomal protein S26E, related	1	17	2.222E8	0.955	-0.066
1171	NCLIV_055690	hypothetical protein	8	29	8.376E8	0.955	-0.067
1172	NCLIV_051110	hypothetical protein, conserved	2	17	5.186E8	0.955	-0.067
1173	NCLIV_023790	hypothetical protein, conserved	4	19	1.602E9	0.955	-0.067
1174	NCLIV_061440	hypothetical protein	3	19	1.180E9	0.954	-0.067
1175	NCLIV_002780	YML024Wp-like protein, related	3	16	1.718E9	0.954	-0.067
1176	NCLIV_008760	hypothetical protein, conserved	1	2	5.552E7	0.954	-0.068
1177	NCLIV_032070	hypothetical protein, conserved	18	56	9.578E8	0.954	-0.068

1178	NCLIV_005360	splicing factor 3a protein, putative	4	12	5.842E7	0.954	-0.068
1179	NCLIV_047520	hypothetical protein, conserved	6	23	4.647E8	0.954	-0.068
1180	NCLIV_015210	ATP-dependent helicase, putative, putative	7	71	1.405E9	0.954	-0.068
1181	NCLIV_060660	SRS domain-containing protein, SRS57, putative	6	72	1.010E9	0.954	-0.069
1182	NCLIV_028140	hypothetical protein, conserved	1	4	4.470E8	0.952	-0.071
1183	NCLIV_050440	signal recognition particle domain-containing protein, putative	5	16	1.107E8	0.952	-0.071
1184	NCLIV_050050	M16 family peptidase, putative	2	9	1.323E8	0.952	-0.071
1185	NCLIV_024740	Myosin, heavy polypeptide 1, skeletal muscle, adult, related	11	66	9.618E8	0.951	-0.072
1186	NCLIV_065270	hypothetical protein	9	58	1.441E9	0.951	-0.072
1187	NCLIV_019330	NADPH:quinone reductase and related Zn-dependent oxidoreductases, related	1	3	5.505E7	0.951	-0.072
1188	NCLIV_036540	vesicle trafficking protein, putative	3	12	4.364E8	0.951	-0.073
1189	NCLIV_052380	hypothetical protein	2	23	1.136E9	0.951	-0.073
1190	NCLIV_050470	hypothetical protein	9	73	3.462E9	0.950	-0.074
1191	NCLIV_030050	hypothetical protein, conserved	14	118	2.316E9	0.950	-0.074
1192	NCLIV_010480	hypothetical protein, conserved	1	1	2.587E7	0.950	-0.074
1193	NCLIV_036610	hypothetical protein	9	50	1.195E9	0.950	-0.074
1194	NCLIV_018020	hypothetical protein	14	41	4.582E8	0.950	-0.074
1195	NCLIV_034910	hypothetical protein, conserved	1	1	4.907E7	0.949	-0.075
1196	NCLIV_039100	hypothetical protein	6	359	1.170E10	0.949	-0.075
1197	NCLIV_050030	hypothetical protein	3	12	1.931E8	0.949	-0.076
1198	NCLIV_042690	vacuolar proton-translocating ATPase subunit, putative	7	24	2.209E8	0.949	-0.076
1199	NCLIV_033410	hypothetical protein	2	7	1.244E7	0.949	-0.076
1200	NCLIV_055070	Ubiquitin carboxyl-terminal hydrolase (EC 3.1.2.15), related	1	4	4.985E8	0.948	-0.076
1201	NCLIV_007470	hypothetical protein, conserved	6	15	3.215E7	0.948	-0.077
1202	NCLIV_032680	diacylglycerol acyltransferase, putative	3	7	5.828E7	0.948	-0.077
1203	NCLIV_014150	hypothetical protein	3	36	2.034E8	0.948	-0.077
1204	NCLIV_056430	hypothetical protein, conserved	6	59	2.327E9	0.947	-0.079
1205	NCLIV_062880	hypothetical protein, conserved	8	43	9.173E8	0.946	-0.079
1206	NCLIV_051650	M16 family peptidase, putative	1	3	4.800E6	0.946	-0.079
1207	NCLIV_052000	DNA replication licensing factor, putative	3	6	9.012E7	0.946	-0.080
1208	NCLIV_057710	ATP synthase epsilon chain, putative	1	15	2.051E8	0.946	-0.080
1209	NCLIV_026820	hypothetical protein, conserved	2	14	1.929E8	0.946	-0.080
1210	NCLIV_032900	hypothetical protein	1	2	2.365E7	0.945	-0.081
1211	NCLIV_044690	hypothetical protein, conserved	3	17	9.892E7	0.945	-0.081
1212	NCLIV_054570	60S ribosomal protein L128B 27a, related	1	17	1.168E8	0.945	-0.082
1213	NCLIV_033680	Solute carrier family 25 (Mitochondrial carrier, dicarboxylate transporter), member 10, related	6	52	2.637E8	0.944	-0.083
1214	NCLIV_001135	hypothetical protein	7	25	3.315E8	0.944	-0.083
1215	NCLIV_046010	hypothetical protein, conserved	4	16	4.991E7	0.944	-0.083
1216	NCLIV_006860	hypothetical protein	3	16	2.638E8	0.944	-0.084
1217	NCLIV_037150	hypothetical protein	5	17	1.677E8	0.942	-0.086
1218	NCLIV_036700	M16 family peptidase, putative	22	131	2.276E9	0.942	-0.086

1219	NCLIV_060600	6-phosphogluconate dehydrogenase, decarboxylating, related	6	27	4.415E8	0.942	-0.086
1220	NCLIV_014670	YGR048Wp-like protein, related	1	1	1.419E7	0.942	-0.086
1221	NCLIV_065540	hypothetical protein, conserved	1	6	2.477E7	0.942	-0.086
1222	NCLIV_028750	hypothetical protein	12	45	4.575E8	0.942	-0.086
1223	NCLIV_035910	hypothetical protein	5	30	2.203E8	0.942	-0.086
1224	NCLIV_067220	hypothetical protein, conserved	2	12	4.143E8	0.941	-0.087
1225	NCLIV_007800	Tg65, related	10	41	5.308E8	0.941	-0.088
1226	NCLIV_007640	hypothetical protein, conserved	1	2	5.456E7	0.941	-0.088
1227	NCLIV_010980	3-methyl-2-oxobutanoate dehydrogenase (lipoamide) kinase, put	2	5	1.978E7	0.941	-0.088
1228	NCLIV_048520	hypothetical protein, conserved	4	11	1.160E8	0.940	-0.090
1229	NCLIV_022240	ATPase, related	4	19	5.891E8	0.939	-0.091
1230	NCLIV_066600	DEHA2F06798p, related	2	8	1.646E8	0.939	-0.091
1231	NCLIV_025990	hypothetical protein, conserved	2	5	7.828E7	0.939	-0.091
1232	NCLIV_020300	hypothetical protein, conserved	1	3	4.527E7	0.939	-0.091
1233	NCLIV_061940	hypothetical protein	7	45	8.744E8	0.938	-0.092
1234	NCLIV_064330	hypothetical protein, conserved	1	2	1.782E8	0.938	-0.092
1235	NCLIV_025320	cyclophilin, putative	1	2	5.817E7	0.938	-0.092
1236	NCLIV_036630	14-3-3 protein, putative	5	12	1.862E8	0.938	-0.093
1237	NCLIV_046050	hypothetical protein	28	210	3.513E9	0.938	-0.093
1238	NCLIV_000340	hypothetical protein	5	19	3.177E8	0.938	-0.093
1239	NCLIV_023610	vacuolar ATP synthase subunit d, putative	1	2	1.429E8	0.937	-0.093
1240	NCLIV_041850	Hsp20/alpha crystallin domain-containing protein, putative	8	48	9.646E8	0.937	-0.094
1241	NCLIV_062230	Glutaminyl-tRNA synthetase (EC 6.1.1.18), related	5	13	3.238E8	0.937	-0.094
1242	NCLIV_032130	hypothetical protein, conserved	1	4	1.609E8	0.936	-0.095
1243	NCLIV_049080	hypothetical protein	1	22	2.971E9	0.936	-0.095
1244	NCLIV_001660	hypothetical protein, conserved	4	31	4.924E8	0.936	-0.095
1245	NCLIV_018550	YGR231Cp-like protein, related	4	16	2.914E8	0.936	-0.095
1246	NCLIV_054110	YHL017Wp-like protein, related	2	10	2.574E8	0.936	-0.096
1247	NCLIV_032050	DnaJ domain-containing protein, putative	8	45	1.179E9	0.935	-0.096
1248	NCLIV_066270	ribosomal protein L7, putative	1	3	2.884E7	0.935	-0.097
1249	NCLIV_004240	hypothetical protein, conserved	3	7	7.746E7	0.935	-0.098
1250	NCLIV_055820	hypothetical protein	3	21	8.288E7	0.934	-0.098
1251	NCLIV_015310	longevity-assurance (LAG1) domain-containing protein, putative	1	6	3.106E8	0.934	-0.098
1252	NCLIV_013150	hypothetical protein, conserved	7	36	4.851E8	0.933	-0.099
1253	NCLIV_066020	hypothetical protein	17	236	6.894E9	0.933	-0.100
1254	NCLIV_035220	hypothetical protein	2	5	5.402E7	0.933	-0.100
1255	NCLIV_020340	hypothetical protein	12	32	2.395E8	0.933	-0.101
1256	NCLIV_011340	Os02g0511500 protein, related	1	2	1.171E7	0.932	-0.101
1257	NCLIV_024220	glycerol-3-phosphate dehydrogenase, putative	13	45	9.343E8	0.932	-0.101
1258	NCLIV_052510	hypothetical protein	2	12	9.756E7	0.932	-0.102
1259	NCLIV_049790	hypothetical protein, conserved	10	44	1.344E9	0.932	-0.102
1260	NCLIV_048870	hypothetical protein, conserved	2	3	1.312E7	0.931	-0.103
1261	NCLIV_065830	MGC79800 protein, related	4	16	6.364E7	0.931	-0.103

1262	NCLIV_010860	eukaryotic translation initiation factor 6, putative	1	2	3.161E8	0.931	-0.104
1263	NCLIV_046260	Iron regulatory protein-like protein, related	16	96	1.736E9	0.931	-0.104
1264	NCLIV_007110	hypothetical protein	4	19	1.118E9	0.930	-0.105
1265	NCLIV_028050	hypothetical protein, conserved	7	30	1.290E9	0.930	-0.105
1266	NCLIV_032000	hypothetical protein, conserved	1	2	8.569E7	0.930	-0.105
1267	NCLIV_058840	hypothetical protein, conserved	4	41	3.202E9	0.929	-0.107
1268	NCLIV_025010	hypothetical protein	5	28	2.776E8	0.929	-0.107
1269	NCLIV_003510	hypothetical protein, conserved	2	4	1.073E8	0.929	-0.107
1270	NCLIV_041240	NADH dehydrogenase (EC 1.6.99.3), related	13	54	7.485E8	0.928	-0.108
1271	NCLIV_026400	ribosomal protein L23, putative	1	2	2.357E7	0.927	-0.109
1272	NCLIV_054190	hypothetical protein	3	23	4.339E8	0.926	-0.111
1273	NCLIV_005190	transcription elongation factor FACT 140 kDa, putative	7	42	3.335E8	0.926	-0.111
1274	NCLIV_058670	hypothetical protein, conserved	1	3	1.798E8	0.925	-0.112
1275	NCLIV_048020	hypothetical protein	1	7	7.134E8	0.925	-0.112
1276	NCLIV_045800	60S ribosomal protein L3, related	4	42	1.900E9	0.925	-0.112
1277	NCLIV_032660	hypothetical protein	7	83	2.626E9	0.925	-0.112
1278	NCLIV_060400	hypothetical protein	1	3	1.731E8	0.925	-0.113
1279	NCLIV_060730	hypothetical protein	23	180	4.004E9	0.924	-0.114
1280	NCLIV_003050	myosin heavy chain, putative	22	104	1.377E9	0.923	-0.115
1281	NCLIV_013530	hypothetical protein, conserved	1	2	3.185E7	0.923	-0.115
1282	NCLIV_013180	GM04207p, related	9	34	1.641E8	0.923	-0.116
1283	NCLIV_034830	Protoporphyrinogen oxidase, related	1	4	1.685E8	0.923	-0.116
1284	NCLIV_031780	hypothetical protein	19	275	1.226E10	0.923	-0.116
1285	NCLIV_056590	phosphatidylinositol-4-phosphate 5-kinase, putative	5	27	4.807E8	0.923	-0.116
1286	NCLIV_025140	ctr copper transporter domain-containing protein, putative	3	8	1.138E8	0.923	-0.116
1287	NCLIV_001440	NOL1/NOP2/sun family methyltransferase, related	1	2	3.801E7	0.922	-0.117
1288	NCLIV_055980	hypothetical protein, conserved	2	6	4.947E7	0.922	-0.118
1289	NCLIV_062280	hypothetical protein, conserved	5	45	3.111E9	0.921	-0.119
1290	NCLIV_035820	hypothetical protein	2	6	5.521E7	0.921	-0.119
1291	NCLIV_053880	cDNA FLJ54097, highly similar to Succinyl-CoA ligase (ADP-forming) beta-chain, mitochondrial (EC 6.2.1.5), related	10	74	1.219E9	0.920	-0.120
1292	NCLIV_039420	thioredoxin, putative	2	4	2.841E8	0.919	-0.121
1293	NCLIV_031750	importin, putative	2	7	1.188E8	0.919	-0.121
1294	NCLIV_061160	acid phosphatase, putative	10	176	1.189E10	0.919	-0.123
1295	NCLIV_065580	Zinc finger DHHC domain containing 12, related	1	5	5.691E7	0.918	-0.123
1296	NCLIV_002680	AGAP001651-PA, related	1	2	1.712E7	0.918	-0.124
1297	NCLIV_048350	ER lumen protein retaining receptor, related	1	2	3.084E8	0.918	-0.124
1298	NCLIV_030900	hypothetical protein	13	55	9.377E8	0.917	-0.126
1299	NCLIV_001250	guanylate binding protein, putative	8	32	2.270E8	0.916	-0.126
1300	NCLIV_035140	hypothetical protein	2	5	4.532E7	0.916	-0.126
1301	NCLIV_016360	hypothetical protein, conserved	4	14	3.334E8	0.916	-0.127
1302	NCLIV_006160	Multi-pass transmembrane protein, related	6	14	1.122E8	0.916	-0.127
1303	NCLIV_053460	hypothetical protein, conserved	2	2	6.756E7	0.915	-0.127

1304	NCLIV_017420	Zgc:55863, related	1	37	1.354E9	0.915	-0.128
1305	NCLIV_008960	hypothetical protein	4	14	2.533E8	0.915	-0.128
1306	NCLIV_005710	hypothetical protein, conserved	1	1	1.864E8	0.915	-0.129
1307	NCLIV_052920	hypothetical protein, conserved	5	13	3.393E8	0.914	-0.129
1308	NCLIV_030930	peroxiredoxin 3, putative	5	31	8.432E8	0.914	-0.129
1309	NCLIV_062220	Glutaminyl-tRNA synthetase (EC 6.1.1.18), related	2	8	3.689E8	0.914	-0.130
1310	NCLIV_032110	hypothetical protein, conserved	12	58	1.575E9	0.913	-0.131
1311	NCLIV_001040	der1-like family domain-containing protein, conserved, putative	1	2	4.244E8	0.912	-0.132
1312	NCLIV_045480	hypothetical protein	2	4	3.017E7	0.912	-0.133
1313	NCLIV_004730	hypothetical protein	5	25	1.092E9	0.912	-0.133
1314	NCLIV_015880	hypothetical protein	4	38	2.579E9	0.912	-0.133
1315	NCLIV_067350	P-type Ca(2)-ATPase, putative	15	87	1.574E9	0.912	-0.133
1316	NCLIV_032290	SSU ribosomal protein S3P, related	8	88	2.808E9	0.911	-0.134
1317	NCLIV_027530	lectin-domain protein, putative	9	39	6.256E8	0.911	-0.135
1318	NCLIV_005790	hypothetical protein, conserved	2	5	1.367E9	0.910	-0.136
1319	NCLIV_055710	60S ribosomal protein L23, putative	4	27	6.420E8	0.910	-0.136
1320	NCLIV_039500	hypothetical protein	6	33	5.290E8	0.909	-0.137
1321	NCLIV_020920	hypothetical protein, conserved	3	5	7.242E7	0.909	-0.137
1322	NCLIV_040670	hypothetical protein, conserved	1	2	1.820E8	0.909	-0.137
1323	NCLIV_061230	acetyltransferase domain-containing protein, putative	2	6	3.906E7	0.909	-0.138
1324	NCLIV_048600	hypothetical protein	2	7	1.861E8	0.909	-0.138
1325	NCLIV_018130	hypothetical protein, conserved	1	3	3.336E7	0.909	-0.138
1326	NCLIV_003310	hypothetical protein	4	32	1.370E8	0.908	-0.139
1327	NCLIV_026830	Translation elongation factor G, related	4	19	2.273E8	0.908	-0.139
1328	NCLIV_002560	T-cell activation protein phosphatase 2C, putative	1	3	1.685E8	0.907	-0.140
1329	NCLIV_046460	hypothetical protein, conserved	2	17	1.513E8	0.907	-0.140
1330	NCLIV_027600	hypothetical protein, conserved	9	57	3.401E8	0.907	-0.141
1331	NCLIV_058310	vacuolar ATP synthase subunit c, putative	1	10	1.273E8	0.907	-0.142
1332	NCLIV_056110	small heat shock protein 21, putative	3	20	4.572E8	0.906	-0.142
1333	NCLIV_047390	hypothetical protein, conserved	4	21	1.415E9	0.906	-0.142
1334	NCLIV_042410	sortilin, putative	7	54	1.330E9	0.906	-0.143
1335	NCLIV_020120	hypothetical protein, conserved	7	25	3.618E8	0.905	-0.144
1336	NCLIV_034280	ABC transporter, putative	2	6	1.029E8	0.905	-0.144
1337	NCLIV_065030	hypothetical protein, conserved	5	12	1.626E8	0.905	-0.144
1338	NCLIV_005740	hypothetical protein, conserved	2	8	7.706E7	0.905	-0.144
1339	NCLIV_029590	hypothetical protein	2	7	1.270E8	0.904	-0.145
1340	NCLIV_029010	glycerophosphoryl diester phosphodiesterase protein, putative	2	4	2.627E8	0.904	-0.146
1341	NCLIV_000390	hypothetical protein, conserved	9	62	1.116E9	0.904	-0.146
1342	NCLIV_058000	alanine dehydrogenase, putative	1	2	1.850E8	0.903	-0.146
1343	NCLIV_047860	hypothetical protein	17	87	2.869E9	0.903	-0.147
1344	NCLIV_045220	hypothetical protein	3	21	1.680E8	0.902	-0.149
1345	NCLIV_064950	hypothetical protein	14	84	1.490E9	0.902	-0.149
1346	NCLIV_000400	hypothetical protein, conserved	1	5	2.947E8	0.902	-0.149

1347	NCLIV_029170	hypothetical protein, conserved	1	1	7.608E7	0.901	-0.150
1348	NCLIV_020440	U5 snRNP-associated 102 kDa protein, putative	4	18	2.955E8	0.901	-0.150
1349	NCLIV_056380	hypothetical protein, conserved	3	6	1.664E7	0.900	-0.151
1350	NCLIV_055360	hypothetical protein, conserved	28	274	3.174E9	0.900	-0.151
1351	NCLIV_030070	hypothetical protein, conserved	3	9	4.973E7	0.899	-0.153
1352	NCLIV_038460	DEAD box polypeptide 27, related	1	3	4.957E7	0.899	-0.153
1353	NCLIV_001970	hypothetical protein	20	666	3.650E10	0.899	-0.153
1354	NCLIV_056560	DEAD/DEAH box helicase, putative	1	2	6.652E7	0.899	-0.154
1355	NCLIV_048420	hypothetical protein	2	4	3.312E8	0.899	-0.154
1356	NCLIV_014000	hypothetical protein, conserved	3	16	5.765E7	0.898	-0.155
1357	NCLIV_014770	hypothetical protein	2	5	2.012E8	0.898	-0.155
1358	NCLIV_030940	hypothetical protein	10	59	7.907E8	0.898	-0.155
1359	NCLIV_043330	hypothetical protein	7	69	3.454E9	0.898	-0.156
1360	NCLIV_060700	SRS domain-containing protein	7	37	6.585E8	0.897	-0.157
1361	NCLIV_030090	sugar permease, putative	2	4	1.018E8	0.897	-0.157
1362	NCLIV_016390	ERYTHROCYTE MEMBRANE PROTEIN PFEMP3, related	1	4	2.588E8	0.896	-0.158
1363	NCLIV_020220	elongation factor 2, putative	16	130	1.925E9	0.896	-0.158
1364	NCLIV_001670	Elongation factor 1-alpha, related	11	182	1.132E10	0.896	-0.158
1365	NCLIV_062460	hypothetical protein, conserved	23	93	8.736E8	0.896	-0.158
1366	NCLIV_019750	hypothetical protein, conserved	7	31	1.689E8	0.896	-0.159
1367	NCLIV_012230	Ribose-phosphate pyrophosphokinase (EC 2.7.6.1), related	6	21	7.816E8	0.895	-0.160
1368	NCLIV_035100	hypothetical protein, conserved	1	5	8.829E7	0.894	-0.161
1369	NCLIV_014060	lysophospholipase, putative	28	300	4.241E9	0.894	-0.162
1370	NCLIV_042680	hypothetical protein, conserved	5	39	5.286E8	0.893	-0.162
1371	NCLIV_005620	articulin 4, putative	13	79	1.791E9	0.893	-0.163
1372	NCLIV_051590	GTP binding protein, putative	1	9	1.071E8	0.893	-0.163
1373	NCLIV_054120	sushi domain protein / SCR repeat- containing protein, putative	11	59	5.507E8	0.891	-0.166
1374	NCLIV_035190	hypothetical protein, conserved	1	20	5.724E9	0.891	-0.166
1375	NCLIV_019830	hypothetical protein	5	23	1.085E9	0.891	-0.167
1376	NCLIV_019270	hypothetical protein, conserved	1	5	6.116E7	0.891	-0.167
1377	NCLIV_063910	hypothetical protein	2	4	3.969E7	0.890	-0.169
1378	NCLIV_058420	hypothetical protein, conserved	2	17	2.468E8	0.889	-0.170
1379	NCLIV_068640	protein phosphatase 2C, putative	3	8	9.443E7	0.889	-0.170
1380	NCLIV_047160	hypothetical protein	5	20	5.600E8	0.889	-0.170
1381	NCLIV_016380	Ribosomal protein L22, related	2	17	1.379E9	0.889	-0.170
1382	NCLIV_046590	Diacylglycerol acyl transferase, related	2	4	1.982E7	0.889	-0.170
1383	NCLIV_038360	TSP1 domain-containing protein TSP12 (Precursor), related	14	83	1.405E9	0.889	-0.170
1384	NCLIV_057960	hypothetical protein	8	30	6.984E8	0.889	-0.171
1385	NCLIV_015920	Histone H4, related	4	135	1.262E10	0.889	-0.171
1386	NCLIV_017500	hypothetical protein	8	57	2.851E9	0.888	-0.171
1387	NCLIV_003470	thrombospondin type 1 domain-containing protein, putative	3	17	1.753E8	0.888	-0.171
1388	NCLIV_037520	hypothetical protein, conserved	13	71	7.211E8	0.887	-0.172
1389	NCLIV_061220	cation-transporting ATPase, putative	2	4	2.930E6	0.887	-0.174

1390	NCLIV_043940	acetyltransferase domain-containing protein, putative	4	19	1.730E8	0.886	-0.174
1391	NCLIV_038440	hypothetical protein, conserved	2	12	1.420E8	0.886	-0.174
1392	NCLIV_033250	SRS domain-containing protein	16	449	2.622E10	0.886	-0.174
1393	NCLIV_059650	EF hand domain-containing protein, putative	10	43	2.289E8	0.886	-0.174
1394	NCLIV_015990	hypothetical protein	1	7	1.622E8	0.886	-0.175
1395	NCLIV_026340	hypothetical protein	2	6	4.930E8	0.885	-0.176
1396	NCLIV_042570	ABC transporter, putative	3	10	8.153E7	0.885	-0.176
1397	NCLIV_043950	hypothetical protein	2	3	2.423E8	0.885	-0.176
1398	NCLIV_054020	hypothetical protein, conserved	1	2	4.222E6	0.885	-0.176
1399	NCLIV_020770	hypothetical protein	6	27	6.322E8	0.885	-0.176
1400	NCLIV_010130	Helicase, DEAD/DEAH family protein, related	1	3	4.221E7	0.885	-0.177
1401	NCLIV_006280	hypothetical protein, conserved	2	7	6.791E7	0.885	-0.177
1402	NCLIV_000510	translocation protein sec62, putative	7	35	1.586E9	0.884	-0.178
1403	NCLIV_037400	la domain-containing protein, putative	6	36	7.427E8	0.884	-0.178
1404	NCLIV_011800	hypothetical protein, conserved	1	3	4.694E7	0.884	-0.179
1405	NCLIV_061350	glutathione reductase, putative	2	10	1.140E8	0.884	-0.179
1406	NCLIV_057260	hypothetical protein, conserved	1	11	7.081E7	0.883	-0.179
1407	NCLIV_017840	hypothetical protein, conserved	2	7	7.193E8	0.882	-0.180
1408	NCLIV_062960	MORN repeat protein, putative	4	27	1.782E8	0.882	-0.182
1409	NCLIV_047440	DNA-directed RNA polymerase II largest subunit, putative	2	7	4.931E7	0.882	-0.182
1410	NCLIV_031250	mucin, putative	6	25	3.970E8	0.882	-0.182
1411	NCLIV_026600	46 kDa FK506-binding nuclear protein, putative	12	74	8.323E8	0.881	-0.183
1412	NCLIV_001000	hypothetical protein, conserved	1	5	2.105E8	0.881	-0.183
1413	NCLIV_011960	hypothetical protein, conserved	4	24	5.331E8	0.880	-0.185
1414	NCLIV_038590	hypothetical protein, conserved	3	14	1.870E8	0.880	-0.185
1415	NCLIV_060120	hypothetical protein, conserved	1	1	3.821E7	0.879	-0.186
1416	NCLIV_032910	hypothetical protein	3	22	8.488E8	0.879	-0.186
1417	NCLIV_034900	casein kinase II beta chain, putative	4	12	1.917E8	0.879	-0.187
1418	NCLIV_031040	Peptidyl-prolyl cis-trans isomerase (EC 5.2.1.8), related	6	41	4.558E8	0.879	-0.187
1419	NCLIV_070170	hypothetical protein	1	11	1.749E8	0.878	-0.187
1420	NCLIV_042020	hypothetical protein	1	8	6.990E7	0.878	-0.188
1421	NCLIV_004860	hypothetical protein	4	23	3.489E8	0.877	-0.189
1422	NCLIV_045870	Putative dense granule protein 3	1	22	2.369E9	0.877	-0.189
1423	NCLIV_053890	hypothetical protein, conserved	4	32	2.419E8	0.877	-0.190
1424	NCLIV_040970	malate:quinone oxidoreductase, putative	15	98	2.247E9	0.877	-0.190
1425	NCLIV_057900	WD domain, G-beta repeat-containing protein, putative	5	13	7.948E7	0.876	-0.191
1426	NCLIV_065500	hypothetical protein, conserved	5	21	6.382E8	0.876	-0.191
1427	NCLIV_003890	hypothetical protein, conserved	3	15	4.149E8	0.876	-0.191
1428	NCLIV_057210	hypothetical protein, conserved	2	4	4.661E7	0.875	-0.192
1429	NCLIV_032920	hypothetical protein	7	29	7.099E8	0.875	-0.192
1430	NCLIV_044000	hypothetical protein	8	73	1.952E9	0.875	-0.193
1431	NCLIV_034381	SRS domain-containing protein	2	19	3.863E8	0.875	-0.193
1432	NCLIV_036570	YALI0D21604p, related	10	43	6.453E8	0.874	-0.194

1433	NCLIV_005700	hypothetical protein, conserved	5	16	2.551E8	0.874	-0.194
1434	NCLIV_051010	signal peptide peptidase domain-containing protein, putative	4	17	5.757E8	0.874	-0.194
1435	NCLIV_024030	hypothetical protein, conserved	7	48	2.146E8	0.874	-0.194
1436	NCLIV_020550	glyoxalase, putative	3	9	1.191E8	0.874	-0.194
1437	NCLIV_048370	phospholipid-transporting ATPase, P-type, putative	1	5	8.311E7	0.874	-0.195
1438	NCLIV_067490	protein phosphatase 2C, putative	7	29	6.112E8	0.873	-0.195
1439	NCLIV_055770	hypothetical protein, conserved	1	7	1.230E9	0.873	-0.196
1440	NCLIV_046530	reticulon domain-containing protein, putative	1	9	2.555E8	0.873	-0.196
1441	NCLIV_047010	hypothetical protein, conserved	2	8	7.067E7	0.872	-0.197
1442	NCLIV_011840	hypothetical protein	1	3	7.459E8	0.872	-0.197
1443	NCLIV_048590	hypothetical protein, conserved	34	186	2.251E9	0.872	-0.198
1444	NCLIV_068890	hypothetical protein	2	42	1.497E9	0.872	-0.198
1445	NCLIV_029090	hypothetical protein, conserved	6	24	2.005E8	0.872	-0.198
1446	NCLIV_038570	hypothetical protein, conserved	6	24	4.850E8	0.872	-0.198
1447	NCLIV_003190	mitochondrial carrier domain-containing protein, putative	3	18	1.378E8	0.871	-0.199
1448	NCLIV_048460	thioredoxin, putative	11	52	7.067E8	0.871	-0.199
1449	NCLIV_028110	DnaJ protein, putative	4	20	5.648E8	0.870	-0.201
1450	NCLIV_043760	hypothetical protein, conserved	6	91	4.109E9	0.870	-0.201
1451	NCLIV_020040	hypothetical protein, conserved	3	8	7.169E7	0.869	-0.202
1452	NCLIV_032780	small heat shock protein 20, putative	9	123	6.842E9	0.869	-0.202
1453	NCLIV_056340	ATP-dependent RNA helicase DDX55 (EC 3.6.1.-), related	3	11	3.312E8	0.869	-0.202
1454	NCLIV_022370	Ribophorin I, related	2	7	1.834E8	0.869	-0.203
1455	NCLIV_068970	Succinate dehydrogenase flavoprotein subunit, related	6	48	9.024E8	0.869	-0.203
1456	NCLIV_067160	hypothetical protein	1	5	1.104E8	0.868	-0.204
1457	NCLIV_007260	p97 protein, putative	11	55	7.630E8	0.868	-0.205
1458	NCLIV_051800	hypothetical protein, conserved	4	14	1.310E8	0.866	-0.208
1459	NCLIV_016480	hypothetical protein	1	1	1.612E7	0.865	-0.208
1460	NCLIV_025030	hypothetical protein, conserved	1	3	6.117E7	0.865	-0.209
1461	NCLIV_021280	hypothetical protein, conserved	1	4	4.434E8	0.865	-0.209
1462	NCLIV_057820	hypothetical protein	13	37	8.549E8	0.865	-0.209
1463	NCLIV_000540	50S ribosomal protein L27, chloroplastic, related	1	1	9.129E7	0.864	-0.211
1464	NCLIV_004710	hypothetical protein	2	14	4.298E8	0.864	-0.211
1465	NCLIV_021720	hypothetical protein, conserved	5	13	3.341E8	0.863	-0.212
1466	NCLIV_015180	ATP synthase, related	2	11	1.708E8	0.863	-0.213
1467	NCLIV_049530	hypothetical protein, conserved	1	1	7.037E7	0.862	-0.214
1468	NCLIV_026920	hypothetical protein	2	4	2.373E7	0.861	-0.215
1469	NCLIV_027930	Rhoptry kinase family protein ROP17, putative	5	18	5.022E8	0.861	-0.215
1470	NCLIV_041790	hypothetical protein, conserved	3	9	4.638E7	0.861	-0.216
1471	NCLIV_054790	hypothetical protein, conserved	1	4	8.900E6	0.861	-0.216
1472	NCLIV_038710	Lin-28 homolog B, related	4	16	3.751E8	0.860	-0.217
1473	NCLIV_026500	ALVEOLIN1, related	2	20	1.590E9	0.860	-0.218
1474	NCLIV_026430	DnaJ domain containing protein, related	10	43	3.930E8	0.860	-0.218
1475	NCLIV_023240	Heavy metal translocating P-type ATPase, related	2	6	1.943E7	0.859	-0.219

1476	NCLIV_065010	hypothetical protein	9	69	9.929E8	0.859	-0.220
1477	NCLIV_062940	pyruvate dehydrogenase, putative	6	24	4.312E8	0.857	-0.222
1478	NCLIV_039900	hypothetical protein, conserved	3	24	1.852E9	0.857	-0.222
1479	NCLIV_052030	hypothetical protein, conserved	3	6	1.132E8	0.855	-0.226
1480	NCLIV_060140	inner membrane complex protein IMC3, putative	11	129	8.751E9	0.855	-0.226
1481	NCLIV_068630	hypothetical protein, conserved	4	25	1.683E8	0.854	-0.227
1482	NCLIV_024940	hypothetical protein, conserved	1	2	6.316E7	0.854	-0.227
1483	NCLIV_001070	Histone H2B, related	5	98	7.488E9	0.854	-0.228
1484	NCLIV_061740	Hydrolase, alpha/beta fold family, related	2	7	8.547E7	0.854	-0.228
1485	NCLIV_011330	major facilitator superfamily domain-containing protein, putative	1	2	2.954E6	0.854	-0.228
1486	NCLIV_054250	Acyl-CoA synthetase, related	8	43	4.358E8	0.854	-0.228
1487	NCLIV_063240	DnaJ homologue, related	1	6	6.697E7	0.853	-0.229
1488	NCLIV_055230	serine/threonine protein phosphatase, putative	1	6	9.585E7	0.853	-0.229
1489	NCLIV_038850	fumarase, putative	3	18	3.632E8	0.853	-0.229
1490	NCLIV_012890	hypothetical protein	7	46	2.602E8	0.853	-0.229
1491	NCLIV_019820	hypothetical protein	4	12	1.983E8	0.853	-0.230
1492	NCLIV_010720	SRS domain-containing protein	3	222	1.757E10	0.852	-0.230
1493	NCLIV_064640	Probable signal peptidase complex subunit 3 (EC 3.4.-.-), related	2	9	8.050E8	0.852	-0.231
1494	NCLIV_041830	Os07g0543600 protein, related	3	14	4.883E8	0.852	-0.231
1495	NCLIV_027520	hypothetical protein, conserved	6	25	1.335E8	0.852	-0.232
1496	NCLIV_019670	calmodulin, putative	3	7	1.556E8	0.852	-0.232
1497	NCLIV_048030	hypothetical protein	2	19	1.893E9	0.851	-0.233
1498	NCLIV_069400	hypothetical protein, conserved	1	2	2.114E7	0.851	-0.233
1499	NCLIV_040550	hypothetical protein, conserved	2	17	1.247E8	0.850	-0.235
1500	NCLIV_047200	hypothetical protein	1	4	9.553E7	0.850	-0.235
1501	NCLIV_051750	hypothetical protein, conserved	2	10	1.203E8	0.849	-0.237
1502	NCLIV_060500	hypothetical protein, conserved	10	45	6.110E8	0.848	-0.238
1503	NCLIV_061620	hypothetical protein	2	6	2.184E7	0.848	-0.238
1504	NCLIV_024840	hypothetical protein	10	38	1.058E9	0.847	-0.239
1505	NCLIV_055490	Heat shock protein 70 (Precursor), related	26	297	1.201E10	0.847	-0.239
1506	NCLIV_053290	ORF73, related	13	60	1.115E9	0.847	-0.239
1507	NCLIV_043070	hypothetical protein	1	5	1.304E8	0.847	-0.239
1508	NCLIV_038830	Sm protein, related	1	2	1.192E8	0.847	-0.239
1509	NCLIV_019930	hypothetical protein, conserved	6	25	2.972E8	0.847	-0.240
1510	NCLIV_059000	hypothetical protein, conserved	6	56	2.446E9	0.846	-0.241
1511	NCLIV_026150	Histone H3, related	5	88	1.394E10	0.846	-0.241
1512	NCLIV_005610	Membrane protein, related	4	18	8.208E8	0.846	-0.241
1513	NCLIV_031890	glutamine/glutamic acid rich protein, putative	5	27	2.951E8	0.846	-0.241
1514	NCLIV_036760	hypothetical protein	1	8	1.218E8	0.845	-0.243
1515	NCLIV_039090	hypothetical protein, conserved	3	23	1.835E9	0.845	-0.243
1516	NCLIV_068210	hypothetical protein, conserved	4	8	1.359E8	0.845	-0.243
1517	NCLIV_056730	hypothetical protein, conserved	4	6	1.299E8	0.844	-0.244
1518	NCLIV_030620	hypothetical protein, conserved	6	36	1.024E9	0.844	-0.245

1519	NCLIV_037540	YOR039Wp-like protein, related	2	9	1.080E8	0.844	-0.245
1520	NCLIV_015260	hypothetical protein, conserved	2	7	1.980E8	0.844	-0.246
1521	NCLIV_046810	transmembrane domain-containing protein, putative	2	7	9.384E7	0.843	-0.246
1522	NCLIV_060060	hypothetical protein, conserved	6	22	1.155E8	0.843	-0.246
1523	NCLIV_029850	transcription factor btf3	1	5	9.932E6	0.842	-0.248
1524	NCLIV_015290	hypothetical protein	4	15	1.669E8	0.842	-0.248
1525	NCLIV_011040	N-ethylmaleimide-sensitive factor, putative	5	22	2.593E8	0.841	-0.251
1526	NCLIV_001050	hypothetical protein	2	4	5.784E6	0.839	-0.253
1527	NCLIV_025240	Gbp1p protein, putative	10	114	2.575E9	0.839	-0.253
1528	NCLIV_023560	hypothetical protein, conserved	2	5	6.002E7	0.838	-0.255
1529	NCLIV_026910	SAC3/GANP domain-containing protein, putative	1	2	1.638E8	0.838	-0.255
1530	NCLIV_041260	hypothetical protein, conserved	1	1	1.945E7	0.837	-0.256
1531	NCLIV_030860	hypothetical protein, conserved	9	63	4.302E9	0.836	-0.258
1532	NCLIV_014950	trans-2,3-enoyl-CoA reductase, putative	2	10	1.010E8	0.836	-0.258
1533	NCLIV_031610	Glucosamine--fructose-6-phosphate aminotransferase, related	5	20	3.535E8	0.835	-0.260
1534	NCLIV_026290	hypothetical protein, conserved	1	4	3.955E7	0.834	-0.261
1535	NCLIV_064840	hypothetical protein, conserved	6	22	2.766E9	0.834	-0.262
1536	NCLIV_023090	hypothetical protein	10	54	2.833E9	0.834	-0.263
1537	NCLIV_053040	hypothetical protein, conserved	1	6	7.487E8	0.833	-0.263
1538	NCLIV_004400	hypothetical protein	5	30	1.156E9	0.833	-0.263
1539	NCLIV_025280	hypothetical protein, conserved	2	10	8.514E7	0.833	-0.263
1540	NCLIV_044410	Rhoptry kinase family protein ROP35, putative	6	17	1.772E8	0.833	-0.264
1541	NCLIV_029800	hypothetical protein, conserved	1	2	5.494E7	0.832	-0.266
1542	NCLIV_006180	duplicated carbonic anhydrase, putative	8	120	9.263E9	0.832	-0.266
1543	NCLIV_005120	DsPTP1 protein, related	2	10	1.179E8	0.831	-0.266
1544	NCLIV_029720	eukaryotic translation initiation factor, putative	1	2	1.596E8	0.831	-0.267
1545	NCLIV_049110	hypothetical protein	3	11	6.374E8	0.830	-0.269
1546	NCLIV_016840	AFL091Wp, related	1	6	2.027E8	0.830	-0.269
1547	NCLIV_032080	3-hydroxybutyryl-coA dehydrogenase, putative	4	9	1.358E8	0.829	-0.271
1548	NCLIV_038680	hypothetical protein	2	15	2.480E8	0.827	-0.275
1549	NCLIV_026080	hypothetical protein, conserved	2	10	8.604E7	0.826	-0.275
1550	NCLIV_036290	hypothetical protein	1	1	3.670E7	0.826	-0.276
1551	NCLIV_036830	hypothetical protein, conserved	15	61	6.041E8	0.826	-0.277
1552	NCLIV_035920	Adenylate/guanylate cyclase with GAF sensor / FHA domain, rel	5	23	2.822E8	0.825	-0.277
1553	NCLIV_063330	Ubiquinol-cytochrome c reductase cytochrome c1 subunit, rel	2	10	2.271E9	0.825	-0.277
1554	NCLIV_061540	mismatch repair protein, putative	2	6	1.967E8	0.824	-0.279
1555	NCLIV_046250	MGC83131 protein, related	3	15	1.460E8	0.824	-0.280
1556	NCLIV_032220	50S ribosomal protein L30e, related	4	25	1.037E9	0.823	-0.280
1557	NCLIV_051210	aminomethyltransferase, mitochondrial, putative	2	4	6.757E7	0.822	-0.283
1558	NCLIV_057950	Lipase maturation factor 2, related	10	35	1.725E9	0.822	-0.284
1559	NCLIV_046570	PST-A protein, related	1	1	5.416E7	0.821	-0.284
1560	NCLIV_035540	hypothetical protein	5	14	4.672E8	0.821	-0.285
1561	NCLIV_029980	glycerol-3-phosphate acyltransferase, putative	4	16	2.356E8	0.821	-0.285

1562	NCLIV_038750	DnaJ domain-containing protein, putative	4	29	9.381E8	0.821	-0.285
1563	NCLIV_056480	hypothetical protein	13	86	2.138E9	0.819	-0.287
1564	NCLIV_016370	Galactosyltransferase, related	2	4	2.038E7	0.819	-0.288
1565	NCLIV_057920	hypothetical protein	1	4	3.244E6	0.819	-0.289
1566	NCLIV_066100	microtubule-binding protein, putative	12	47	1.068E9	0.818	-0.290
1567	NCLIV_065280	Proliferating cell nuclear antigen, related	3	13	3.290E8	0.817	-0.291
1568	NCLIV_054210	GPI mannosyltransferase 1, putative	1	2	1.094E8	0.817	-0.292
1569	NCLIV_043300	nucleolar phosphoprotein nucleolin, putative	10	48	4.513E8	0.815	-0.295
1570	NCLIV_047600	hypothetical protein	1	4	9.456E7	0.815	-0.295
1571	NCLIV_033230	SRS domain-containing protein,SRS29B, putative	11	661	1.617E10	0.815	-0.296
1572	NCLIV_045170	hypothetical protein	3	9	1.307E8	0.814	-0.296
1573	NCLIV_064420	ubiquitin, putative	2	7	6.176E8	0.814	-0.296
1574	NCLIV_022260	hypothetical protein	1	2	5.853E7	0.814	-0.297
1575	NCLIV_007510	hypothetical protein, conserved	6	18	1.460E8	0.813	-0.299
1576	NCLIV_022220	hypothetical protein	8	97	1.513E9	0.812	-0.300
1577	NCLIV_035300	lipoate-protein ligase a, putative	3	5	7.717E7	0.812	-0.301
1578	NCLIV_041120	hypothetical protein, conserved	5	96	1.225E10	0.810	-0.303
1579	NCLIV_013870	pinin/SDK/memA/ domain-containing protein, putative	1	3	5.031E7	0.810	-0.304
1580	NCLIV_029570	hypothetical protein, conserved	4	24	4.429E8	0.810	-0.305
1581	NCLIV_070190	Dihydrolipoyl dehydrogenase (EC 1.8.1.4), related	4	11	4.028E8	0.809	-0.306
1582	NCLIV_022100	DnaJ domain containing protein, related	1	3	6.157E7	0.809	-0.306
1583	NCLIV_036250	Asparaginyl-tRNA synthetase, related	8	58	1.541E9	0.808	-0.308
1584	NCLIV_054140	adenyl cyclase associated protein, putative	8	100	3.533E9	0.807	-0.309
1585	NCLIV_029970	TRNA synthetase Gly, related	10	48	5.561E8	0.806	-0.310
1586	NCLIV_029000	hypothetical protein, conserved	1	3	4.909E7	0.806	-0.311
1587	NCLIV_027270	cell division protein, putative	1	6	2.931E8	0.806	-0.312
1588	NCLIV_020070	hypothetical protein, conserved	3	10	9.980E7	0.805	-0.312
1589	NCLIV_063860	thioredoxin, putative	5	38	5.789E8	0.805	-0.312
1590	NCLIV_027090	hypothetical protein, conserved	1	3	1.557E7	0.805	-0.313
1591	NCLIV_037060	hypothetical protein	1	2	6.528E6	0.805	-0.314
1592	NCLIV_037190	glyceraldehyde-3-phosphate dehydrogenase, putative	10	43	2.299E9	0.804	-0.315
1593	NCLIV_032160	UBA/TS-N domain-containing protein, putative	1	3	7.862E6	0.804	-0.315
1594	NCLIV_009440	hypothetical protein	2	15	5.178E8	0.804	-0.315
1595	NCLIV_041940	Glyceraldehyde 3-phosphate dehydrogenase (EC 1.2.1.12), rel	6	67	3.630E9	0.803	-0.316
1596	NCLIV_038540	RAB22A, member RAS oncogene family, related	6	28	2.854E8	0.803	-0.316
1597	NCLIV_024410	10 kDa chaperonin, related	1	5	9.461E7	0.803	-0.316
1598	NCLIV_025730	hypothetical protein, conserved	1	15	7.196E8	0.802	-0.317
1599	NCLIV_013940	Heat shock protein 17.4, related	4	12	8.566E7	0.802	-0.318
1600	NCLIV_039820	cDNA FLJ57348, highly similar to Homo sapiens hexokinase domain containing 1 (HKDC1), mRNA, related	7	42	7.838E8	0.801	-0.320
1601	NCLIV_056670	hypothetical protein, conserved	12	134	4.375E9	0.799	-0.323
1602	NCLIV_022390	hypothetical protein	5	11	1.484E8	0.799	-0.323
1603	NCLIV_030820	hypothetical protein, conserved	6	70	2.491E9	0.798	-0.325

1604	NCLIV_048960	hypothetical protein	1	1	5.925E7	0.797	-0.327
1605	NCLIV_069630	hypothetical protein, conserved	5	22	5.724E8	0.797	-0.328
1606	NCLIV_026140	Histone H2A, related	2	76	9.745E9	0.797	-0.328
1607	NCLIV_030450	hypothetical protein, conserved	1	2	7.270E7	0.797	-0.328
1608	NCLIV_000710	hypothetical protein, conserved	6	44	1.453E9	0.796	-0.329
1609	NCLIV_053680	hypothetical protein	3	6	1.125E8	0.796	-0.330
1610	NCLIV_038210	hypothetical protein, conserved	1	6	1.583E8	0.795	-0.330
1611	NCLIV_059590	Serine/threonine phosphatase 6 regulatory ankyrin repeat, rel	3	24	1.030E8	0.795	-0.331
1612	NCLIV_066710	Cation efflux system permease, related	1	8	5.120E7	0.795	-0.331
1613	NCLIV_042820	cDNA FLJ58099, highly similar to Homo sapiens clathrin, heavy polypeptide-like 1 (CLTCL1), transcript variant 1, mRNA, related	34	238	3.587E9	0.795	-0.331
1614	NCLIV_037760	Rhomboid-6, isoform A (EC 3.4.21.-), related	9	44	8.957E8	0.795	-0.331
1615	NCLIV_029420	myosin light chain TgMLC1, putative	7	104	2.157E9	0.794	-0.333
1616	NCLIV_024290	hypothetical protein, conserved	2	9	1.098E8	0.794	-0.333
1617	NCLIV_005400	Brx1, related	2	6	3.383E7	0.793	-0.334
1618	NCLIV_011090	cation-transporting ATPase, putative	3	15	1.007E8	0.793	-0.335
1619	NCLIV_012040	dihydroorotate dehydrogenase, putative	1	10	1.622E8	0.793	-0.335
1620	NCLIV_027080	hypothetical protein, conserved	1	6	1.841E8	0.792	-0.336
1621	NCLIV_012610	hypothetical protein, conserved	2	7	7.977E7	0.792	-0.336
1622	NCLIV_062730	hypothetical protein	1	5	9.216E7	0.792	-0.336
1623	NCLIV_039240	integral membrane protein, putative	3	12	3.816E8	0.792	-0.337
1624	NCLIV_009780	Beta-lactamase domain protein (Precursor), related	7	16	3.080E8	0.792	-0.337
1625	NCLIV_011500	hypothetical protein, conserved	3	12	1.258E8	0.792	-0.337
1626	NCLIV_022920	Renin, related	1	8	2.724E7	0.791	-0.339
1627	NCLIV_049570	hypothetical protein	5	27	4.066E8	0.790	-0.339
1628	NCLIV_044870	hypothetical protein, conserved	1	1	4.417E6	0.790	-0.340
1629	NCLIV_007010	cathepsin C2 (TgCPC2), putative	2	11	1.210E8	0.789	-0.342
1630	NCLIV_016970	hypothetical protein, conserved	6	43	7.317E8	0.789	-0.342
1631	NCLIV_018260	U4/U6 small nuclear ribonucleoprotein, putative	4	8	1.207E8	0.788	-0.344
1632	NCLIV_014020	Peroxisomal protein 2E-1 (EC 1.11.1.15), related	4	10	5.191E8	0.787	-0.345
1633	NCLIV_056400	short-chain dehydrogenase/reductase family protein, putative	8	38	1.112E9	0.787	-0.345
1634	NCLIV_043870	hypothetical protein, conserved	3	7	1.255E8	0.786	-0.348
1635	NCLIV_034930	hypothetical protein, conserved	3	10	1.331E8	0.786	-0.348
1636	NCLIV_016800	TCP-1/cpn60 chaperonin family protein, putative	7	30	1.771E9	0.785	-0.349
1637	NCLIV_017510	40S ribosomal protein S29, putative	1	6	7.524E8	0.784	-0.350
1638	NCLIV_042450	opine dehydrogenase, putative	6	17	8.026E8	0.784	-0.350
1639	NCLIV_036960	hypothetical protein, conserved	3	11	1.222E8	0.784	-0.352
1640	NCLIV_039000	Probable 26S protease regulatory subunit 6B, related	3	8	4.463E8	0.782	-0.355
1641	NCLIV_049600	30S ribosomal protein S8P, related	6	46	1.695E9	0.781	-0.357
1642	NCLIV_007220	hypothetical protein	4	24	1.040E9	0.780	-0.358
1643	NCLIV_004160	Histone H2B, related	4	61	4.714E9	0.780	-0.358
1644	NCLIV_061560	hypothetical protein, conserved	2	17	2.457E8	0.780	-0.359
1645	NCLIV_045260	hypothetical protein	4	24	3.683E8	0.779	-0.360

1646	NCLIV_052350	hypothetical protein, conserved	5	43	7.788E8	0.779	-0.361
1647	NCLIV_052880	granule antigen protein GRA6, putative	2	26	5.686E8	0.779	-0.361
1648	NCLIV_066990	hypothetical protein	1	1	5.209E7	0.779	-0.361
1649	NCLIV_026420	autophagy protein APG9, putative	1	1	3.099E7	0.778	-0.361
1650	NCLIV_046760	DnaJ domain-containing protein, related	2	12	1.369E8	0.778	-0.361
1651	NCLIV_016150	Serine--pyruvate transaminase (EC 2.6.1.51), related	5	15	4.340E8	0.778	-0.362
1652	NCLIV_012700	TCP-1/cpn60 family chaperonin, putative	5	22	6.848E8	0.778	-0.362
1653	NCLIV_061050	hypothetical protein	3	20	6.149E8	0.777	-0.364
1654	NCLIV_011550	Novel protein (Zgc:77155), related	2	4	3.096E7	0.777	-0.364
1655	NCLIV_051690	hypothetical protein	1	4	5.178E6	0.777	-0.365
1656	NCLIV_051300	Nifu-pending protein, related	1	5	4.886E7	0.776	-0.365
1657	NCLIV_051960	hypothetical protein, conserved	2	14	1.500E8	0.775	-0.367
1658	NCLIV_064290	mitochondrial carrier family protein, putative	2	8	3.467E7	0.775	-0.367
1659	NCLIV_017440	hypothetical protein	5	17	8.928E8	0.775	-0.368
1660	NCLIV_011580	helicase, putative	1	1	2.734E7	0.775	-0.368
1661	NCLIV_064810	hypothetical protein	3	17	8.258E8	0.774	-0.370
1662	NCLIV_054550	Cytochrome c oxidase subunit 2, related	2	9	9.557E8	0.774	-0.370
1663	NCLIV_008410	Gamma-aminobutyric acid receptor-associated protein-like 1, rel	1	4	4.143E7	0.773	-0.372
1664	NCLIV_055730	hypothetical protein	9	40	9.379E8	0.772	-0.373
1665	NCLIV_039280	Isocitrate dehydrogenase-like protein, related	2	54	3.507E9	0.771	-0.374
1666	NCLIV_033790	replication factor c, putative	1	5	1.024E8	0.771	-0.375
1667	NCLIV_038140	hypothetical protein	1	6	4.864E8	0.771	-0.375
1668	NCLIV_047680	hypothetical protein, conserved	2	5	5.033E6	0.771	-0.375
1669	NCLIV_005420	Phosphoglycerate kinase (EC 2.7.2.3), related	1	7	6.771E7	0.771	-0.376
1670	NCLIV_004810	hypothetical protein, conserved	4	14	1.044E8	0.771	-0.376
1671	NCLIV_000530	3-oxoacyl-(Acyl-carrier-protein) synthase II, related	4	7	9.971E7	0.770	-0.377
1672	NCLIV_040650	hypothetical protein, conserved	22	52	6.206E8	0.770	-0.377
1673	NCLIV_065970	hypothetical protein, conserved	2	7	5.313E8	0.770	-0.377
1674	NCLIV_011380	TB2/DP1, HVA22 domain-containing protein, putative	2	11	1.767E9	0.770	-0.378
1675	NCLIV_026070	armadillo/beta-catenin-like repeat-containing protein, putative	2	18	1.892E8	0.769	-0.378
1676	NCLIV_004870	DNA gyrase subunit A, putative	2	3	1.129E9	0.768	-0.380
1677	NCLIV_030120	importin beta-3 subunit, putative	4	8	5.409E7	0.767	-0.382
1678	NCLIV_061690	hypothetical protein	3	8	4.853E7	0.767	-0.382
1679	NCLIV_032510	FAD binding domain containing protein, related	1	3	3.590E6	0.767	-0.383
1680	NCLIV_051490	hypothetical protein, conserved	2	11	1.880E8	0.767	-0.384
1681	NCLIV_050590	GL11864, related	8	42	1.006E9	0.766	-0.384
1682	NCLIV_010050	SRS domain-containing protein	7	27	1.586E8	0.766	-0.384
1683	NCLIV_058810	superoxide dismutase	2	19	4.822E8	0.766	-0.385
1684	NCLIV_014710	DEAD/DEAH box helicase domain-containing protein, putative	1	2	1.422E7	0.765	-0.386
1685	NCLIV_069460	hypothetical protein	4	18	4.260E8	0.765	-0.386
1686	NCLIV_014100	hypothetical protein, conserved	1	16	2.249E8	0.764	-0.388
1687	NCLIV_031480	ribosomal protein S11, putative	1	1	2.885E6	0.763	-0.390
1688	NCLIV_001150	Elongation factor Ts, related	1	3	1.714E7	0.763	-0.390

1689	NCLIV_033270	hypothetical protein	7	25	4.474E8	0.763	-0.391
1690	NCLIV_027160	hypothetical protein, conserved	6	31	7.429E8	0.763	-0.391
1691	NCLIV_024320	hypothetical protein, conserved	3	20	5.221E8	0.762	-0.393
1692	NCLIV_031900	3-oxoacyl-(acyl-carrier-protein) synthase III family protein, put	1	1	8.964E7	0.761	-0.394
1693	NCLIV_031460	hypothetical protein, conserved	3	10	3.040E8	0.761	-0.394
1694	NCLIV_022470	diacylglycerol kinase, putative	3	10	4.560E7	0.760	-0.396
1695	NCLIV_011690	ROP15 protein, related	1	7	2.372E8	0.758	-0.399
1696	NCLIV_003350	ATP-dependent DNA helicase, putative	1	5	7.499E7	0.755	-0.405
1697	NCLIV_061210	hypothetical protein, conserved	1	1	7.587E7	0.754	-0.407
1698	NCLIV_007670	hypothetical protein	1	2	1.156E7	0.754	-0.407
1699	NCLIV_009670	Cold-shock protein, DNA-binding, related	3	26	9.951E8	0.753	-0.409
1700	NCLIV_054910	hypothetical protein, conserved	2	3	8.510E7	0.753	-0.409
1701	NCLIV_027480	hypothetical protein	5	41	6.884E8	0.753	-0.409
1702	NCLIV_062890	hypothetical protein	1	12	1.138E9	0.753	-0.410
1703	NCLIV_002590	Phthalate dioxygenase reductase subunit, related	4	32	1.625E8	0.752	-0.410
1704	NCLIV_011410	protein disulfide isomerase	19	270	7.178E9	0.752	-0.411
1705	NCLIV_003570	hypothetical protein	2	7	4.120E7	0.751	-0.413
1706	NCLIV_036050	hypothetical protein, conserved	1	3	9.314E8	0.751	-0.414
1707	NCLIV_054510	heat shock protein 90, putative	2	8	7.678E7	0.750	-0.415
1708	NCLIV_014160	Tudor domain-containing protein, related	2	6	5.285E7	0.749	-0.416
1709	NCLIV_064520	hypothetical protein, conserved	6	18	3.279E8	0.749	-0.418
1710	NCLIV_016080	protein kinase (incomplete catalytic triad), putative	1	1	2.757E7	0.747	-0.421
1711	NCLIV_023540	aldo/keto reductase family oxidoreductase, putative	3	10	9.819E7	0.746	-0.422
1712	NCLIV_064680	ATP-dependent metalloprotease involved in cell division, related	1	3	1.398E7	0.746	-0.423
1713	NCLIV_056830	60S ribosomal protein L7a, putative	1	3	1.274E8	0.745	-0.424
1714	NCLIV_066630	GRA9 protein, putative	11	67	2.533E9	0.745	-0.424
1715	NCLIV_045450	14-3-3-like protein, related	1	3	8.408E6	0.744	-0.427
1716	NCLIV_034731	SRS domain-containing protein	2	19	7.644E8	0.743	-0.429
1717	NCLIV_026750	cyclin H, putative	1	12	8.147E7	0.743	-0.429
1718	NCLIV_059680	CBR-IMB-2 protein, related	5	22	4.344E8	0.742	-0.430
1719	NCLIV_068670	hypothetical protein, conserved	1	7	1.550E8	0.742	-0.431
1720	NCLIV_039030	hypothetical protein	2	18	4.864E8	0.741	-0.433
1721	NCLIV_024060	hypothetical protein	2	5	2.088E8	0.740	-0.434
1722	NCLIV_066770	seryl-tRNA synthetase, putative	4	24	2.453E8	0.740	-0.434
1723	NCLIV_015520	hypothetical protein, conserved	1	14	1.009E8	0.740	-0.434
1724	NCLIV_002280	hypothetical protein	2	6	9.371E7	0.740	-0.434
1725	NCLIV_017460	NUDIX hydrolase domain-containing protein, putative	1	2	2.453E7	0.740	-0.435
1726	NCLIV_007390	MGC78841 protein, related	3	15	5.090E7	0.740	-0.435
1727	NCLIV_041930	GE25707, related	6	45	9.342E8	0.740	-0.435
1728	NCLIV_042590	2-oxoglutarate dehydrogenase E1 component, related	6	58	2.857E9	0.739	-0.437
1729	NCLIV_045580	hypothetical protein, conserved	9	65	1.068E9	0.738	-0.437
1730	NCLIV_014450	Phosphoglucomutase 2, related	6	27	8.358E8	0.737	-0.441
1731	NCLIV_008340	AGAP004885-PA, related	2	17	8.391E8	0.736	-0.442

1732	NCLIV_046040	hypothetical protein	6	64	5.460E9	0.735	-0.444
1733	NCLIV_032560	hypothetical protein	3	8	3.301E7	0.735	-0.445
1734	NCLIV_063340	hypothetical protein	3	8	3.643E8	0.735	-0.445
1735	NCLIV_062520	3-ketoacyl-(Acyl-carrier-protein) reductase , related	7	23	9.074E8	0.734	-0.446
1736	NCLIV_029730	Ras family domain-containing protein, putative	4	38	8.754E8	0.733	-0.448
1737	NCLIV_067570	SRS domain-containing protein	1	2	7.393E6	0.733	-0.448
1738	NCLIV_025430	lipase domain-containing protein, putative	2	5	5.527E7	0.733	-0.448
1739	NCLIV_007560	ELMO/CED-12 family domain-containing protein, putative	1	1	7.929E6	0.733	-0.449
1740	NCLIV_015790	fatty acyl-CoA desaturase, putative	3	10	1.537E8	0.733	-0.449
1741	NCLIV_060640	Gpi16 subunit, GPI transamidase domain- containing protein, put	2	11	7.251E7	0.732	-0.450
1742	NCLIV_016770	hypothetical protein, conserved	2	4	5.593E7	0.732	-0.450
1743	NCLIV_060630	hypothetical protein, conserved	2	3	3.558E7	0.732	-0.451
1744	NCLIV_003650	hypothetical protein	5	42	2.768E9	0.731	-0.453
1745	NCLIV_063740	hypothetical protein, conserved	5	21	2.964E8	0.730	-0.453
1746	NCLIV_018710	hypothetical protein	3	10	2.961E8	0.729	-0.456
1747	NCLIV_043200	ADP-ribosylation factor-like protein 1, putative	1	3	2.542E7	0.728	-0.458
1748	NCLIV_032330	Malate dehydrogenase (NAD) (EC 1.1.1.37) (Precursor), related	11	63	3.478E9	0.727	-0.459
1749	NCLIV_065470	hypothetical protein	5	52	4.344E8	0.727	-0.461
1750	NCLIV_013260	hypothetical protein, conserved	5	90	1.271E9	0.726	-0.461
1751	NCLIV_008990	hypothetical protein	1	7	2.261E8	0.726	-0.462
1752	NCLIV_046790	Calmodulin, related	1	5	7.668E6	0.726	-0.462
1753	NCLIV_035050	heat shock protein hslv, putative	1	5	1.821E8	0.726	-0.462
1754	NCLIV_067500	ftsJ-like methyltransferase domain-containing protein, putative	2	5	4.168E7	0.726	-0.463
1755	NCLIV_058380	hypothetical protein, conserved	1	1	1.004E6	0.725	-0.463
1756	NCLIV_056140	hypothetical protein, conserved	13	49	7.487E8	0.725	-0.464
1757	NCLIV_015410	hypothetical protein	5	30	5.421E8	0.725	-0.464
1758	NCLIV_023780	hypothetical protein, conserved	1	7	1.165E8	0.722	-0.469
1759	NCLIV_047050	hypothetical protein, conserved	1	2	3.891E7	0.721	-0.471
1760	NCLIV_035250	GK18150, related	6	31	8.646E8	0.720	-0.473
1761	NCLIV_051620	4'-phosphopantetheinyl transferase, putative	2	6	1.252E8	0.719	-0.475
1762	NCLIV_019600	hypothetical protein	1	2	7.544E6	0.719	-0.475
1763	NCLIV_048040	hypothetical protein, conserved	5	28	6.047E8	0.715	-0.483
1764	NCLIV_045490	hypothetical protein, conserved	2	5	6.287E8	0.714	-0.485
1765	NCLIV_004340	hypothetical protein	1	2	1.980E8	0.714	-0.485
1766	NCLIV_013730	hypothetical protein, conserved	2	6	2.741E7	0.714	-0.486
1767	NCLIV_060320	mitochondrial carrier domain-containing protein, putative	1	17	1.446E9	0.713	-0.488
1768	NCLIV_035950	hypothetical protein	3	7	1.854E8	0.711	-0.491
1769	NCLIV_036970	calmodulin, putative	2	9	1.688E8	0.708	-0.498
1770	NCLIV_034610	hypothetical protein, conserved	1	3	2.099E7	0.708	-0.499
1771	NCLIV_058550	hypothetical protein, conserved	1	3	1.110E9	0.706	-0.501
1772	NCLIV_044350	hypothetical protein, conserved	2	18	6.680E8	0.706	-0.502
1773	NCLIV_065410	ctr copper transporter domain-containing protein, putative	1	3	1.845E7	0.706	-0.502
1774	NCLIV_057720	hypothetical protein	1	3	2.952E8	0.704	-0.506

1775	NCLIV_039610	hypothetical protein, conserved	1	3	3.658E7	0.704	-0.507
1776	NCLIV_048570	hypothetical protein, conserved	9	61	1.653E9	0.703	-0.508
1777	NCLIV_063730	thrombospondin type 1 domain-containing protein, putative	7	23	5.796E7	0.703	-0.509
1778	NCLIV_036300	hypothetical protein, conserved	2	8	1.080E8	0.700	-0.514
1779	NCLIV_027730	hypothetical protein	1	5	9.493E6	0.700	-0.515
1780	NCLIV_014330	Prolyl oligopeptidase (EC 3.4.21.26) (Precursor), related	5	21	3.968E8	0.700	-0.515
1781	NCLIV_034470	hypothetical protein	4	45	3.560E8	0.699	-0.517
1782	NCLIV_046950	UDP-glucose 4-epimerase (EC 5.1.3.2), related	1	1	2.357E7	0.698	-0.518
1783	NCLIV_060760	prolyl-tRNA synthetase, putative	9	40	6.958E8	0.698	-0.518
1784	NCLIV_001300	calmodulin, putative	8	56	3.528E9	0.698	-0.519
1785	NCLIV_005040	hypothetical protein, conserved	2	7	6.176E8	0.696	-0.524
1786	NCLIV_027540	peptidyl-prolyl cis-trans isomerase NIMA- interacting 1, putative	2	5	5.677E7	0.695	-0.524
1787	NCLIV_063610	hypothetical protein, conserved	1	8	9.932E7	0.692	-0.530
1788	NCLIV_038400	Methionine aminopeptidase (EC 3.4.11.18), related	4	26	4.023E8	0.690	-0.535
1789	NCLIV_007000	ADP-ribosylation factor 4, related	3	45	4.143E8	0.690	-0.536
1790	NCLIV_058150	hypothetical protein, conserved	1	1	5.955E6	0.689	-0.538
1791	NCLIV_060810	hypothetical protein, conserved	4	9	1.058E8	0.688	-0.539
1792	NCLIV_029990	vacuolar ATP synthase catalytic subunit A, putative	11	36	9.913E8	0.688	-0.540
1793	NCLIV_031200	hypothetical protein, conserved	1	2	3.593E6	0.686	-0.543
1794	NCLIV_034740	SRS domain-containing protein,SAG3 protein, related	3	22	7.774E8	0.684	-0.548
1795	NCLIV_050210	KH domain-containing protein, putative	7	44	6.593E8	0.683	-0.550
1796	NCLIV_053840	hypothetical protein	7	93	2.792E9	0.681	-0.553
1797	NCLIV_045460	Mitochondrial presequence protease, rel	7	26	2.750E8	0.678	-0.560
1798	NCLIV_016880	S-adenosylmethionine synthetase (EC 2.5.1.6), related	2	6	2.714E8	0.678	-0.561
1799	NCLIV_056250	hypothetical protein	5	22	1.656E8	0.677	-0.563
1800	NCLIV_000280	RanBP1 domain containing protein, related	2	7	1.547E8	0.676	-0.565
1801	NCLIV_027780	hypothetical protein, conserved	6	19	3.281E8	0.675	-0.566
1802	NCLIV_025600	calmodulin, putative	1	2	7.435E6	0.674	-0.569
1803	NCLIV_051440	WD-40 repeat protein, putative	3	6	9.561E7	0.671	-0.576
1804	NCLIV_040830	hypothetical protein, conserved	1	1	1.046E8	0.670	-0.577
1805	NCLIV_033280	hypothetical protein	4	13	7.790E7	0.670	-0.579
1806	NCLIV_021410	hypothetical protein, conserved	4	24	7.966E7	0.668	-0.582
1807	NCLIV_061650	hypothetical protein, conserved	1	2	4.464E8	0.668	-0.582
1808	NCLIV_052610	hypothetical protein, conserved	1	1	4.342E7	0.667	-0.584
1809	NCLIV_009030	Metallophosphoesterase, related	3	12	3.602E8	0.666	-0.586
1810	NCLIV_057740	hypothetical protein, conserved	1	1	2.531E5	0.665	-0.588
1811	NCLIV_019960	lysyl-tRNA synthetase, putative	2	8	3.391E8	0.665	-0.589
1812	NCLIV_046170	Heat Shock Protein 70, ER lumen, related	11	61	1.451E9	0.664	-0.592
1813	NCLIV_043650	WD domain, G-beta repeat-containing protein, putative	1	4	1.283E7	0.660	-0.598
1814	NCLIV_054200	Zgc:92083, related	9	36	2.217E8	0.659	-0.601
1815	NCLIV_059430	hypothetical protein	9	59	3.565E9	0.659	-0.602
1816	NCLIV_057890	Protein-L-isoaspartate O-methyltransferase, related	2	4	9.962E7	0.658	-0.604
1817	NCLIV_054700	uridine phosphorylase, putative	5	55	1.768E9	0.658	-0.604

1818	NCLIV_033370	hypothetical protein, conserved	1	4	5.792E7	0.658	-0.605
1819	NCLIV_052150	hypothetical protein, conserved	2	5	2.332E7	0.657	-0.606
1820	NCLIV_019110	HSP90-like protein, related	28	299	8.962E9	0.652	-0.617
1821	NCLIV_029860	hypothetical protein	2	9	3.348E8	0.652	-0.617
1822	NCLIV_040880	HSP90, related	18	95	2.658E9	0.651	-0.619
1823	NCLIV_022540	hypothetical protein, conserved	5	31	1.633E8	0.649	-0.624
1824	NCLIV_028090	hypothetical protein, conserved	4	35	2.122E9	0.644	-0.635
1825	NCLIV_011900	Arginyl-tRNA synthetase (EC 6.1.1.19), related	3	15	3.508E8	0.641	-0.642
1826	NCLIV_013380	hypothetical protein	1	11	8.527E8	0.640	-0.645
1827	NCLIV_064760	Contig An16c0190, complete genome (Precursor), related	1	3	8.383E7	0.639	-0.645
1828	NCLIV_015010	hypothetical protein, conserved	3	29	7.180E8	0.639	-0.646
1829	NCLIV_063500	hypothetical protein, conserved	1	2	7.950E7	0.639	-0.646
1830	NCLIV_012190	hypothetical protein	1	12	1.479E8	0.633	-0.659
1831	NCLIV_052240	saccharopine dehydrogenase, putative	4	21	7.755E8	0.633	-0.660
1832	NCLIV_060220	hypothetical protein, conserved	22	81	4.747E8	0.632	-0.662
1833	NCLIV_058760	hypothetical protein, conserved	3	11	2.694E8	0.630	-0.666
1834	NCLIV_024470	hypothetical protein	1	2	3.669E6	0.630	-0.667
1835	NCLIV_047810	hypothetical protein	7	35	7.994E8	0.629	-0.669
1836	NCLIV_063590	Glutathione reductase, related	1	8	4.909E8	0.629	-0.670
1837	NCLIV_005010	hypothetical protein, conserved	3	13	4.075E8	0.629	-0.670
1838	NCLIV_066190	caltractin, putative	2	3	1.829E8	0.628	-0.672
1839	NCLIV_035860	Rhoptry kinase family protein ROP32, putative	2	6	5.914E7	0.625	-0.677
1840	NCLIV_018120	hypothetical protein, conserved	12	49	2.538E9	0.625	-0.678
1841	NCLIV_002830	Calcium/proton antiporter, related	1	2	4.652E7	0.625	-0.679
1842	NCLIV_011750	hypothetical protein	2	9	1.669E8	0.624	-0.679
1843	NCLIV_066310	DEAD-box ATP-dependent RNA helicase 34 (EC 3.6.1.-), related	9	66	1.540E9	0.624	-0.681
1844	NCLIV_004190	thioredoxin, putative	6	32	8.339E8	0.621	-0.688
1845	NCLIV_027660	hypothetical protein, conserved	1	5	3.898E7	0.619	-0.691
1846	NCLIV_066960	hypothetical protein, conserved	1	1	9.541E7	0.619	-0.691
1847	NCLIV_045430	DNA-binding protein HU, putative	3	8	6.458E7	0.618	-0.694
1848	NCLIV_060800	Ubiquitin, related	2	57	2.397E9	0.618	-0.694
1849	NCLIV_005150	hypothetical protein	6	347	1.373E10	0.618	-0.694
1850	NCLIV_037110	hypothetical protein, conserved	3	18	7.712E7	0.617	-0.697
1851	NCLIV_043110	interferon gamma-inducible protein 30, putative	6	34	1.840E9	0.617	-0.697
1852	NCLIV_008890	tim10/DDP zinc finger domain-containing protein, putative	1	1	8.045E6	0.616	-0.699
1853	NCLIV_004790	18 kDa cyclophilin, putative	4	30	1.312E9	0.615	-0.700
1854	NCLIV_067450	hypothetical protein	1	2	3.661E8	0.612	-0.708
1855	NCLIV_005560	dense-granule antigen DG32, putative	4	20	5.058E8	0.612	-0.709
1856	NCLIV_044600	hypothetical protein, conserved	2	9	3.444E8	0.611	-0.711
1857	NCLIV_060420	hypothetical protein	19	94	1.882E9	0.609	-0.714
1858	NCLIV_054750	hypothetical protein	6	35	1.566E9	0.609	-0.714
1859	NCLIV_007770	Rhoptry kinase family protein, truncated, putative	9	37	4.738E8	0.609	-0.715
1860	NCLIV_032850	hypothetical protein	1	3	4.727E7	0.607	-0.720

1861	NCLIV_050300	GH18750, related	3	22	2.809E8	0.607	-0.720
1862	NCLIV_063600	calmodulin, putative	1	2	1.043E6	0.605	-0.724
1863	NCLIV_059340	hypothetical protein, conserved	4	19	1.518E8	0.603	-0.730
1864	NCLIV_007580	hypothetical protein	6	20	4.718E8	0.603	-0.730
1865	NCLIV_010370	Ubiquitin carrier protein (EC 6.3.2.-), related	1	2	2.695E8	0.602	-0.731
1866	NCLIV_018830	hypothetical protein, conserved	1	6	9.191E7	0.602	-0.732
1867	NCLIV_012920	Rhoptry kinase family protein ROP40, putative	5	19	3.662E8	0.600	-0.737
1868	NCLIV_006940	hypothetical protein, conserved	2	6	9.986E7	0.598	-0.742
1869	NCLIV_040860	Tryptophanyl-tRNA synthetase (EC 6.1.1.2), related	10	61	1.375E9	0.597	-0.744
1870	NCLIV_060330	molybdopterin cofactor sulfurase, putative	3	7	6.402E7	0.596	-0.747
1871	NCLIV_056420	hypothetical protein, conserved	1	10	1.128E8	0.595	-0.748
1872	NCLIV_019770	hypothetical protein	13	89	2.218E9	0.594	-0.752
1873	NCLIV_046580	hypothetical protein, conserved	3	6	5.246E8	0.593	-0.753
1874	NCLIV_050620	lysine decarboxylase domain-containing protein, putative	6	27	5.414E8	0.585	-0.772
1875	NCLIV_052070	hypothetical protein	3	8	4.515E8	0.585	-0.773
1876	NCLIV_024820	14-3-3 protein homolog	9	107	3.759E9	0.584	-0.775
1877	NCLIV_053940	60S acidic ribosomal protein P2, related	2	14	1.295E9	0.584	-0.776
1878	NCLIV_039960	hypothetical protein	2	6	1.195E8	0.583	-0.778
1879	NCLIV_065510	hypothetical protein, conserved	1	1	2.383E7	0.582	-0.781
1880	NCLIV_047400	hypothetical protein	1	5	2.214E8	0.577	-0.793
1881	NCLIV_036130	CBR-RSP-4 protein, related	1	7	6.094E7	0.576	-0.796
1882	NCLIV_003440	Actin, related	16	369	1.478E10	0.576	-0.797
1883	NCLIV_065090	hypothetical protein, conserved	7	62	1.534E9	0.576	-0.797
1884	NCLIV_033780	YALI0B05610p, related	2	5	2.889E8	0.573	-0.802
1885	NCLIV_006720	hypothetical protein, conserved	6	28	1.692E9	0.573	-0.804
1886	NCLIV_069500	Thermostable carboxypeptidase 1 (EC 3.4.17.-), related	1	2	6.295E7	0.571	-0.808
1887	NCLIV_044440	hypothetical protein	2	11	6.900E7	0.570	-0.810
1888	NCLIV_042700	hypothetical protein, conserved	1	8	1.704E8	0.567	-0.819
1889	NCLIV_009700	Elongation factor 1-alpha, related	2	8	2.186E8	0.563	-0.828
1890	NCLIV_065640	Rhoptry kinase family protein, truncated, putative	5	18	2.150E8	0.559	-0.838
1891	NCLIV_010600	microneme protein MIC3, putative	13	102	6.023E8	0.555	-0.849
1892	NCLIV_038760	mago nashi protein, putative	1	5	1.529E8	0.554	-0.852
1893	NCLIV_005890	hypothetical protein, conserved	2	4	7.305E8	0.554	-0.852
1894	NCLIV_046130	Beta-tubulin, related	1	2	5.186E8	0.553	-0.853
1895	NCLIV_052190	hypothetical protein, conserved	4	20	5.077E8	0.552	-0.859
1896	NCLIV_029080	hypothetical protein	1	9	1.737E8	0.550	-0.862
1897	NCLIV_033140	hypothetical protein	1	2	4.330E8	0.546	-0.872
1898	NCLIV_023130	hypothetical protein, conserved	2	6	5.483E7	0.546	-0.873
1899	NCLIV_051460	hypothetical protein, conserved	3	12	5.276E8	0.544	-0.879
1900	NCLIV_010740	kelch motif domain-containing protein, putative	10	48	2.594E9	0.543	-0.880
1901	NCLIV_048430	hypothetical protein	1	9	2.917E8	0.542	-0.882
1902	NCLIV_011120	Malate dehydrogenase (EC 1.1.1.37), related	2	9	3.752E8	0.542	-0.883
1903	NCLIV_041740	hypothetical protein, conserved	7	31	6.328E8	0.540	-0.889

1904	NCLIV_042640	hypothetical protein, conserved	1	1	2.561E7	0.536	-0.901
1905	NCLIV_066820	proteasome activator subunit, putative	1	1	2.495E7	0.535	-0.904
1906	NCLIV_000140	Inositol-1,4,5-trisphosphate 5-phosphatase 2, related	1	3	4.286E7	0.530	-0.917
1907	NCLIV_004270	protein kinase, putative	1	5	8.282E7	0.529	-0.918
1908	NCLIV_040260	hypothetical protein, conserved	1	3	1.249E7	0.527	-0.924
1909	NCLIV_024110	hypothetical protein, conserved	1	3	1.120E8	0.526	-0.926
1910	NCLIV_036640	hypothetical protein, conserved	2	6	1.086E8	0.525	-0.930
1911	NCLIV_040540	hypothetical protein	7	22	3.720E8	0.523	-0.935
1912	NCLIV_009640	choline kinase, putative	4	8	2.961E8	0.522	-0.937
1913	NCLIV_066870	hypothetical protein	1	1	9.505E7	0.521	-0.940
1914	NCLIV_041270	hypothetical protein	5	19	8.189E8	0.519	-0.947
1915	NCLIV_028170	Rhoptry kinase family protein ROP20, putative	4	15	2.876E8	0.509	-0.974
1916	NCLIV_042400	Macrophage migration inhibitory factor, related	2	9	1.054E8	0.509	-0.976
1917	NCLIV_059450	hypothetical protein	8	29	2.256E9	0.505	-0.985
1918	NCLIV_052780	penicillin amidase domain-containing protein, putative	1	3	3.405E7	0.502	-0.993
1919	NCLIV_068400	NTPase	2	72	2.200E9	0.502	-0.996
1920	NCLIV_043420	SRS domain-containing protein	1	2	1.943E7	0.501	-0.997
1921	NCLIV_050370	hypothetical protein	8	44	2.561E9	0.500	-0.999
1922	NCLIV_064310	GTP-binding nuclear protein Ran, related	6	16	6.744E8	0.498	-1.005
1923	NCLIV_011730	Rhoptry kinase family protein ROP26, putative	5	21	3.825E8	0.495	-1.015
1924	NCLIV_003960	hypothetical protein	1	2	1.109E7	0.494	-1.019
1925	NCLIV_005110	ftsJ-like methyltransferase domain-containing protein, putative	1	2	7.081E8	0.494	-1.019
1926	NCLIV_026200	hypothetical protein	5	11	6.248E7	0.493	-1.021
1927	NCLIV_061600	hypothetical protein	5	11	2.311E8	0.492	-1.023
1928	NCLIV_066410	GG10762, related	4	13	2.200E8	0.492	-1.024
1929	NCLIV_031860	serine-threonine phosphatase 2C, putative	3	14	3.807E8	0.491	-1.027
1930	NCLIV_004300	dynammin-like protein, putative	6	17	2.646E8	0.491	-1.027
1931	NCLIV_002620	Kinesin-ii motor protein, related	1	1	2.050E7	0.491	-1.028
1932	NCLIV_032320	hypothetical protein	2	3	2.048E7	0.487	-1.039
1933	NCLIV_004380	Cathepsin L, related	4	13	6.014E8	0.485	-1.044
1934	NCLIV_032430	hypothetical protein, conserved	1	3	1.629E8	0.485	-1.045
1935	NCLIV_019970	Peptidyl-prolyl cis-trans isomerase A, related	5	33	4.208E8	0.483	-1.050
1936	NCLIV_012860	Leucyl-tRNA synthetase 2 (EC 6.1.1.4), related	1	7	7.696E6	0.483	-1.050
1937	NCLIV_036410	cyst matrix protein, putative	5	28	9.665E8	0.482	-1.054
1938	NCLIV_014760	hypothetical protein, conserved	3	14	3.708E8	0.480	-1.060
1939	NCLIV_068460	hypothetical protein	2	83	2.982E9	0.480	-1.060
1940	NCLIV_011700	hypothetical protein	7	37	1.141E9	0.476	-1.072
1941	NCLIV_002610	isoleucine-tRNA synthetase, putative	4	14	3.112E8	0.475	-1.075
1942	NCLIV_015530	hypothetical protein, conserved	3	11	4.240E8	0.475	-1.075
1943	NCLIV_042650	GG11844, related	1	5	2.218E7	0.467	-1.099
1944	NCLIV_030290	transaldolase, putative	2	6	2.937E7	0.466	-1.101
1945	NCLIV_007730	Valyl-tRNA synthetase, mitochondrial, related	10	34	3.049E8	0.462	-1.114
1946	NCLIV_068520	microneme TgMIC5 protein, putative	2	11	7.410E8	0.462	-1.114

1947	NCLIV_056520	Rab1-prov protein, related	1	25	8.932E8	0.462	-1.115
1948	NCLIV_042740	patched family domain-containing protein, conserved, putative	1	1	6.404E6	0.461	-1.117
1949	NCLIV_000940	Glucose-6-phosphate dehydrogenase, putative	7	36	9.096E8	0.456	-1.134
1950	NCLIV_040040	GrpE protein homolog, related	3	3	3.659E7	0.455	-1.137
1951	NCLIV_000840	Thioredoxin h, related	2	12	7.602E8	0.454	-1.141
1952	NCLIV_025130	Translation initiation factor 1A (AelF-1A), related	1	1	1.119E8	0.449	-1.156
1953	NCLIV_002390	Nucleoside diphosphate kinase (EC 2.7.4.6), related	7	51	1.582E9	0.446	-1.165
1954	NCLIV_028680	apical membrane antigen 1, putative	9	55	1.096E9	0.442	-1.177
1955	NCLIV_062290	hypothetical protein, conserved	1	2	6.882E7	0.441	-1.181
1956	NCLIV_050630	H/ACA ribonucleoprotein complex subunit 2-like protein, related	1	11	7.981E8	0.438	-1.191
1957	NCLIV_063150	Serpin peptidase inhibitor, clade B, member 1, like 3, related	8	44	4.262E9	0.435	-1.200
1958	NCLIV_044850	hypothetical protein, conserved	2	5	1.521E8	0.432	-1.210
1959	NCLIV_063970	long chain acyl-CoA synthetase, putative	2	5	7.770E7	0.432	-1.211
1960	NCLIV_006780	hypothetical protein, conserved	6	25	3.505E9	0.430	-1.216
1961	NCLIV_030420	Rcn2-prov protein, related	6	55	2.207E9	0.420	-1.253
1962	NCLIV_054830	hypothetical protein	1	3	8.334E7	0.418	-1.258
1963	NCLIV_018500	Fatty acyl-CoA synthetase 2, related	7	41	2.401E9	0.417	-1.261
1964	NCLIV_054150	Ubiquitin-conjugating enzyme, related	1	2	1.809E8	0.417	-1.263
1965	NCLIV_015200	Pyruvate kinase (EC 2.7.1.40), related	7	19	1.025E9	0.413	-1.277
1966	NCLIV_046540	ligoendopeptidase F, putative	2	2	1.539E8	0.399	-1.327
1967	NCLIV_031550	hypothetical protein	8	50	5.821E8	0.390	-1.358
1968	NCLIV_068340	hypothetical protein	1	6	1.246E8	0.388	-1.367
1969	NCLIV_061200	actin, putative	2	10	1.256E8	0.385	-1.376
1970	NCLIV_026680	Protein kinase AKT, related	1	5	3.953E8	0.384	-1.381
1971	NCLIV_021050	hypothetical protein	12	124	8.332E9	0.379	-1.398
1972	NCLIV_047660	hypothetical protein	8	39	1.413E9	0.379	-1.400
1973	NCLIV_002480	CBR-DPY-23 protein, related	1	2	1.350E8	0.378	-1.405
1974	NCLIV_061170	hypothetical protein	9	29	6.333E8	0.376	-1.412
1975	NCLIV_018420	p36 protein, putative	5	27	8.685E8	0.369	-1.440
1976	NCLIV_015830	hypothetical protein	1	8	2.505E8	0.367	-1.445
1977	NCLIV_006060	hypothetical protein, conserved	2	10	7.150E7	0.366	-1.451
1978	NCLIV_049400	Serine hydroxymethyltransferase (EC 2.1.2.1), related	1	5	2.733E8	0.348	-1.524
1979	NCLIV_060860	Cytochrome c, related	1	3	3.183E8	0.347	-1.527
1980	NCLIV_001080	hypothetical protein	1	2	9.382E5	0.347	-1.528
1981	NCLIV_028830	Glucose-6-phosphate isomerase (EC 5.3.1.9), related	3	5	2.938E8	0.343	-1.542
1982	NCLIV_048240	hypothetical protein	6	16	1.420E8	0.340	-1.555
1983	NCLIV_020990	hypothetical protein	6	20	6.609E8	0.335	-1.580
1984	NCLIV_051340	toxofilin, putative	6	35	7.764E8	0.325	-1.620
1985	NCLIV_022970	microneme protein 2, putative	6	42	6.724E8	0.321	-1.637
1986	NCLIV_018530	hypothetical protein, conserved	5	11	1.620E8	0.317	-1.656
1987	NCLIV_051970	MIC2-associated protein M2AP, putative	3	31	4.212E9	0.315	-1.668
1988	NCLIV_049940	Translation initiation factor IF-2 beta subunit, related	2	6	2.578E7	0.306	-1.709
1989	NCLIV_042510	hypothetical protein, conserved	2	6	2.214E7	0.305	-1.712

1990	NCLIV_068230	nucleoside-triphosphatase, putative	5	10	2.076E8	0.295	-1.762
1991	NCLIV_061760	microneme protein MIC6, putative	4	21	4.411E8	0.294	-1.766
1992	NCLIV_045650	28 kDa antigen, putative	8	75	8.402E9	0.289	-1.789
1993	NCLIV_020720	microneme protein MIC11, putative	2	16	6.190E8	0.288	-1.798
1994	NCLIV_049520	hypothetical protein, conserved	8	48	1.634E9	0.287	-1.803
1995	NCLIV_041870	hypothetical protein	1	1	8.809E6	0.281	-1.834
1996	NCLIV_043270	microneme protein MIC1, putative	12	123	1.324E10	0.257	-1.958
1997	NCLIV_059510	RNase L inhibitor, related	1	2	6.140E7	0.248	-2.012
1998	NCLIV_020480	hypothetical protein	1	4	1.236E9	0.230	-2.121
1999	NCLIV_059380	SJCHGC06128 protein, related	1	5	3.588E7	0.226	-2.149
2000	NCLIV_065440	hypothetical protein	3	11	5.215E8	0.225	-2.150
2001	NCLIV_033690	hypothetical protein	6	42	1.270E9	0.220	-2.183
2002	NCLIV_002940	microneme protein MIC4, putative	11	78	4.240E9	0.212	-2.238
2003	NCLIV_036400	dense granule protein 1 / major antigenp24, putative	4	17	2.205E9	0.202	-2.310
2004	NCLIV_015380	hypothetical protein, conserved	6	29	7.569E8	0.186	-2.424
2005	NCLIV_057330	hypothetical protein, conserved	1	1	3.404E7	0.175	-2.514
2006	NCLIV_000860	SPATR, related	3	6	2.257E7	0.175	-2.517
2007	NCLIV_057870	hypothetical protein, conserved	1	2	2.026E7	0.168	-2.572
2008	NCLIV_042480	hypothetical protein, conserved	1	2	5.389E6	0.155	-2.693
2009	NCLIV_062770	microneme protein 8, putative	3	6	2.392E7	0.135	-2.889
2010	NCLIV_066250	microneme protein 10, putative	1	2	9.347E7	0.120	-3.053
2011	NCLIV_000740	Class I chitinase, related	2	2	1.238E8	0.076	-3.713